Rational Design of Self-Assembling Crystal Scaffolds and DNA-Peptide Hybrid Materials

by

Tara MacCulloch

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved March 2021 by the
Graduate Supervisory Committee:

Nicholas Stephanopoulos, Chair
Ian Gould
Chad Borges

ARIZONA STATE UNIVERSITY

May 2021

ABSTRACT

Since the conception of DNA nanotechnology, the field has evolved towards the development of complex, dynamic 3D structures. The predictability of Watson-Crick base pairing makes DNA an unparalleled building block, and enables exceptional programmability in nanostructure shape and size. The work presented in this dissertation focuses on expanding two facets of the field: (1) introducing functionality through the incorporation of peptides to create DNA-peptide hybrid materials, and (2) the development of self-assembling DNA crystal lattices for scaffolding biomolecules.

DNA nanostructures have long been proposed as drug delivery vehicles; however, they are not biocompatible because of their low stability in low salt environments and entrapment within the endosome. To address these issues, a functionalized peptide coating was designed to act as a counterion to a six-helix bundle, while simultaneously displaying numerous copies of an endosomal escape peptide to enable cytosolic delivery. This functionalized peptide coating creates a DNA-peptide hybrid material, but does not allow specific positioning or orientation of the peptides. The ability to control those aspects required the synthesis of DNA-peptide or DNA-peptide-DNA conjugates that can be incorporated into the nanostructure. The approach was utilized to produce a synbody where three peptides that bind transferrin with micromolar affinity, which were presented for multivalent binding to optimize affinity. Additionally, two DNA handle was attached to an enzymatically cleavable peptide to link two unique nanostructures. The second DNA handle was also used to constrain the peptide in a cyclic fashion to mimic the cell-adhesive conformations of RGD and PHSRN in fibronectin.

The original goal of DNA nanotechnology was to use a crystalline lattice made of DNA to host proteins for their structural determination using X-ray crystallography. The work presented here takes significant steps towards achieving this goal, including elucidating design rules to control cavity size within the scaffold for accommodating guest molecules of unique sizes, approaches to improve the atomic detail of the scaffold, and strategies to modulate the symmetry of each unique lattice. Finally, this work surveys methodologies towards the incorporation of several guest molecules, with promising preliminary results that constitute a significant advancement towards the ultimate goal of the field.

ACKNOWLEDGMENTS

To my advisor, Dr. Nicholas Stephanopoulos, I am grateful that you took me into your lab and allowed me to a part of getting it up and running. It has been quite the adventure working with you over the last few years while I figure out how to become a better scientist and while you figure out how to become a mentor. Thank you for making meetings slightly more enjoyable by spending the extra few minutes to joke around instead of taking everything seriously, even if it meant having to hear terrible puns every day. I would also like to thank the rest of my supervisory committee, Dr. Gould, and Dr. Borges, for taking the time out of your busy schedules to help me through the big milestones of my graduate degree, reaching the point of writing this dissertation would not have been possible without you.

To the two graduate students who started the Stephanopoulos lab with me, Alex Buchberger and Raghu Pradeep, I am thankful that I was not going through this journey alone but as Taraleghu. Alex, even though we butted heads a lot, you were also my closest ally in lab, always willing to help me fix an instrument when it was on the fritz and for the tacos every Tuesday. Raghu, your energy and ability to put 100% of effort into anything that you tried was very admirable. Thanks for putting in the effort to try and be a part of my world while also sharing your world and culture with open arms, I look forward to coming to visit you in India in the future so I have a chance to wear all of my Kurtas.

To other members of the Stephanopoulos lab who were present to give advice and lend a hand, Drs. Minghui Liu and Yang Xu, thank you guys for always taking the time

to help me out whenever your experience was greater than mine. Chad Simmons, thank you for being a mentor, even when that was not your official job. Working on the crystal projects and spending hours arguing over what specific shade of PyMOL color we wanted to use for a figure, taught me to put as much effort and pride into my work as possible. I also appreciate all the time you took to go and grab a Hazy IPA with me when I needed a break from lab and all the dog pictures you sent.

To the younger graduate students in my lab; Julio, Alamin, Skylær, and Ally; I want to thank you guys for coming in and bringing some joy and love of science to help brighten up all our days as we got old, jaded, and bitter with the passing years. The group chat between the four of us dedicated to PhD life memes helped to laugh off the mishaps that occur during grad school. To Tyler Rockwood, even though you started out as an undergraduate and worked your way through your masters with me, I never viewed you as a someone I had to mentor, but more of someone I could collaborate with. To all of the collaborators I was able to work with throughout my time at ASU, thank you for taking the time to work with me on things that I could not have done by myself. Particularly, Fallon Fumasi, I always enjoyed the time we spent catching up while exchanging materials.

Thank you to the three older graduate students (Claire Crowther, Marely Tejada-Ferrari, and Nick Hollaran) that I was fortunate to have taken me in as a roommate during my first year. My labmates and I would not have made it through my first year or my oral exams without the advice and guidance from you guys on how the program worked or what steps to take when. Marely, I owe you for getting me out of my room and out to different events, you helped me to make all of my friends that I could turn to when I

needed to get out of lab and have some fun. The memories of Strong Beer, Tour de Fat, Pride, and showing you around Colorado are some of my favorite from the last six years, not to mention you made the best 30-year-old flower/bubble girl any wedding has ever seen. Claire, even though we are so different, I could not have asked for a better roommate for four years. From binge watching survivor and the Great British Bake-Off to having late night tea after a rough day, you were the biggest support system during my time at ASU, even after you graduated and left you were still willing to help me with all the applications and my dissertation. To the other piece of Friendship Village, Jesse Coe, I had the best time jamming to old 90's ska with you in our backyard and look forward to being your neighbor again in our old age. Thanks to my other friends in the department including Rafa Alcala-Torano and Zach Dobson, who were always up to go grab a beer at happy hour, and Amanda Young-Gonzales for the countless JVN videos to lift my mood.

To my family, particularly my husband Tim, thank you for dealing with my weird schedule, I know that trying to plan anything with me was impossible. To my parents, thank you for supporting my education that allowed me to get to graduate school, in order to make this dissertation a possibility. And to my entire In-law family, thank you for being understanding when I only showed up to half of the family occasions because I was too busy carrying out experiments in lab.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION TO DNA NANOTECHNOLOGY AND DNA-PEPTDE HYBRID MATERIALS

1.1 Nanotechnology Background and Basics

The origin of nanotechnology, a field that concerns itself with building materials on the nanometer scale, is attributed to a speech given by Richard Feynman in 1959. His talk titled "There's Plenty of Room at the Bottom" discussed producing materials on a small scale.[1] Following his speech, he posed a challenge to write the entire *Encyclopedia Britannica* on the head of the pin. This feat was accomplished in 1985, by Tom Newman when he was able to write a paragraph of *A Tale of Two Cities* by Charles Dickens at 1/25,000th of its original size using electron lithography. [2] Since then, nanotechnology has grown into a vast field that utilizes the principle of bottom-up assembly or building larger materials from small building blocks. Bottom-up assembly often exploits the same intermolecular forces that Nature has developed throughout time to create complex systems such as the extracellular matrix (to take just one example), a complex, three-dimensional (3D) network of proteins (e.g.) collagen, glycoproteins, enzymes, and polysaccharides.[3] These forces allow the small building blocks to spontaneously self-assemble into larger systems through a careful balance of repulsive and attractive forces.[4]

To understand the principles of self-assembly, each of these forces must first be examined individually. Some important driving forces include the hydrophobic effect, which is the propensity of non-polar materials to aggregate in aqueous solutions, hydrogen bonding (the dipole-dipole interaction between nitrogen, oxygen, fluorine, and hydrogen); π-π stacking; which is the attraction between two aromatic rings, and wan der

Waals forces, comprised of momentary dipoles of a molecule. When considering repulsive forces, two major considerations include: the steric effect which stops two molecules from occupying the same space; and solvation/hydration, which is the phenomenon of the solvent interacting with the molecule to form a stable complex. [4-8] While each of these forces are easy to comprehend individually, the ability to predict how each of their enthalpic and entropic properties will dictate spontaneous self-assembly in a predictable manner remains challenging.

DNA has emerged as a well understood, and attractive natural building block to address these challenges because of its ability to self-assemble using specific hydrogen bonding networks. Due to its predictable nature, the sub-field of DNA Nanotechnology has been highly successful in the bottom-up assembly of larger structures made entirely of nucleic acids. Additionally, the invention of solid phase synthesis by Merrifield (which was later extended to produce oligos), and the dramatic advancements with the technique has made the production of synthetic oligonucleotides facile and relatively inexpensive.[9, 10]

1.2 Introduction and Basics of DNA Nanotechnology

1.2.1 Inspiration from Art

DNA nanotechnology was born in 1980 when Nadrian Seeman, a protein crystallographer, was sitting in a pub looking at the woodcarving "Depth" by M.C. Escher. He envisioned the fish in the artwork (Figure 1.1A) as an impression of a branched DNA junction, with a repetitive pattern that inspired him to make a similar DNA junction-based 3D lattice (Figure 1.1B). He posited that by using this lattice to

scaffold proteins, the process of obtaining protein crystals for structure determination with X-ray crystallography, would no longer be unpredictable and arbitrary.

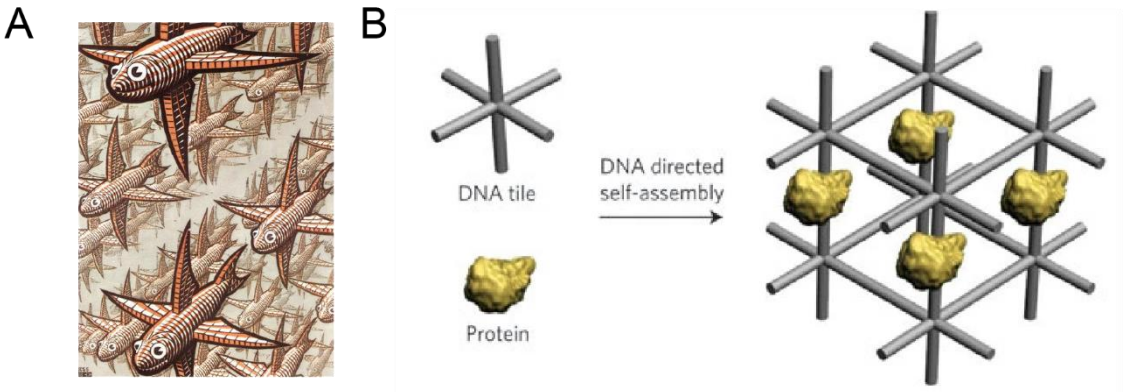


Figure 1.1 Conceptualization of DNA Nanotechnology. A) A single fish that makes up the repetitive pattern in "Depth".[11] B) Depiction of the DNA lattice hosting proteins that Seeman envisioned.[12]

While Seeman's goal of using DNA structures to solve the structure of a protein has yet to be accomplished, his vision served as the foundation for an entirely new branch of DNA-based nanotechnology. Forty years after its conception, DNA nanotechnology is as rich and flourishing area, with thousands of publications showing the diverse range of structures and methodologies from over 500 laboratories that are actively working in the field.[13]

1.2.2 Parameters and Predictability

The predictability of DNA comes from its base pairing, where guanine (G) pairs with cytosine (C), and adenine (A) pairs with thymine (T) to create a right-handed, double stranded helix. This phenomenon was first reported by Watson and Crick who revealed that two strands of DNA assemble together to form a helix based on its complementary base pairing.[14, 15]

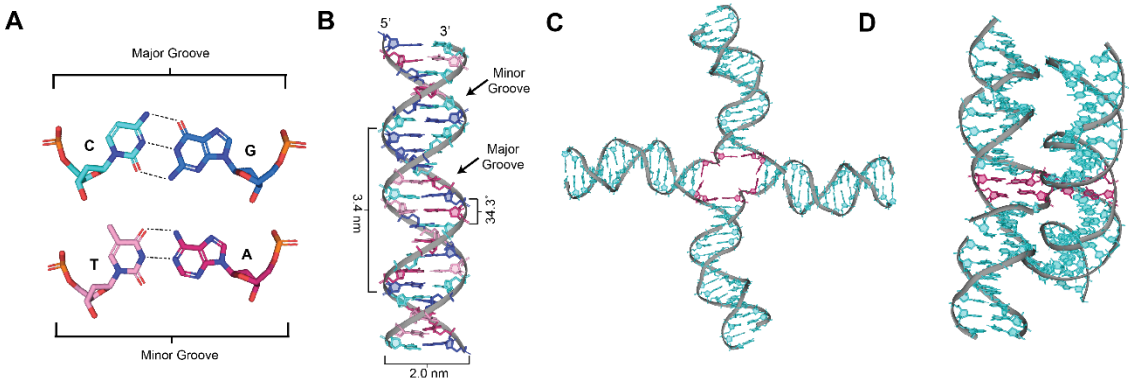


Figure 1.2. Structure of the Watson-Crick base pairing and B-form DNA duplex. A) The base pairs are shown with their H-bonds (dashed lines) with the side corresponding to the major and minor grooves labeled. Nitrogens are shown in blue and oxygens in orange. B) B-form DNA with key parameters are labeled, and the bases retain the same coloring as panel (A) (T is light pink, A is dark pink, C is light blue, and G is dark blue). (C) Open Holliday junction structure with the 8 junction bases highlighted in pink and (D) a corresponding stacked junction that allows DNA nanostructures to access the third dimension.

Additional studies demonstrated that the single stranded polymers are comprised of a sugar phosphate backbone connecting each adjacent nucleobase. Each strand has an inherent directionality, where the 5' end corresponds to the phosphate group and the 3' end corresponds to the hydroxyl group of each ribose moiety. When a duplex forms, the two strands anneal in an antiparallel fashion, with the 5' end forming a perfect hydrogen pair with the 3' end of the complementary strand creating a right-handed helix.[16] G/C base pairs (bp) create three hydrogen bonds, and thus form the strongest bonding partners, while an A/T bp only creates two (Figure 1.2A). Most DNA duplexes follow the rules of B-form DNA; 10.5 bp per turn, corresponding to a helical pitch of 3.4 nm, and a diameter of ~2.0 nm (Figure 1.2B) that is comprised of two strands that twist around to create a helical structure with a major and minor groove.[17] One full helical turn (360°) results from 10.5 consecutive bases (~34.3° per base). Because of each of these

well-defined characteristics, it makes it possible to design structures based on simple geometric principles.

The overall flexibility of double stranded DNA of increasing lengths becomes a concern for the design of higher order structures with greater complexity. To provide additional rigidity, crossover regions (when a strand bridges two adjacent duplexes) may be introduced to mitigate this issue to create larger, rigid structures.[18] Another DNA structure that occurs in biological systems is the Holliday junction (HJ) which is a phenomenon that occurs during genetic recombination.[19] The junction is a four armed branch structure (Figure 1.2C) with eight bases that interact at each crossing point, and can exist in either an open or stacked (Figure 1.2D) formation, which leads to complex structures by adding a multi-dimensional aspect[19-21] These inherent parameters and structural aspects of DNA account for its ability to be used as an effective building block in nanotechnology.

1.2.3 DNA Nanostructures Throughout the Years

While several structural aspects are inherent to DNA, they were not necessarily completely amenable to the production of more complex, stable nanostructures, and required some engineering before they could be applied. One of the first major milestones in DNA nanotechnology was the invention of the *immobile* Holliday junction in 1983.[22] The traditional junction contains symmetry within the central 8 bases, allowing the bases participating in the HJ to slide along the duplex, a phenomenon that occurs in branch migration during replication, Seeman was able to render the junctions immobile by introducing asymmetry to the 8 junction bases (Figure 1.3A). With four nucleotides (A,C,G,T) and 8 positions in a junction, there are 36 possible immobile

junction sequences. The technique of sticky-end cohesion was first developed in 1973 by Cohen to insert a DNA segment into a plasmid, and Seeman hypothesized that the same concept could be applied to assemble larger structures and arrays by designing the branched DNA junctions to contain complementary overhangs (Figure 1.3B). [23] Several examples of multi-branched junctions with 3-6 different arms have been used to create both 2D and 3D structures.[24-30]

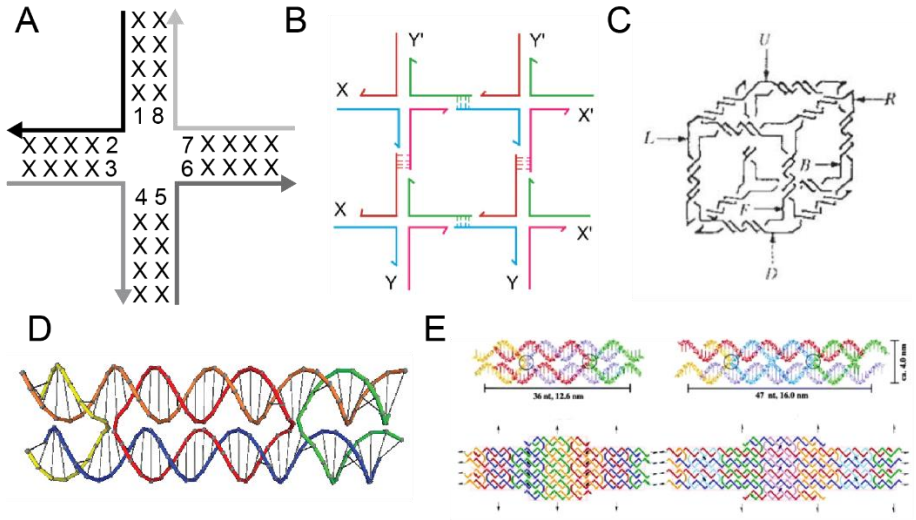


Figure 1.3. Early DNA nanotechnology structures. (A) Immobile Holliday junction where the 8 central bases (all numbered) lack symmetry to prevent migration (B) Four branched DNA structures assembling via sticky end cohesion, where four single stranded regions are complementary, shown by lines[31] (C) Cartoon figure of the first 3D polyhedral structure in the form of a wireframe cube.[32] (D) Representative figure of a DX tile, where two crossover events happen between adjacent duplexes[33] (E) First 2D DNA "crystal" comprised of two DX tiles that can assemble to form a predictable and repetitive pattern[34]

The first 3D polyhedral structure created was a cube that was comprised of six catenated strands, each of which outlined a face of the cube.[32] It was constructed by introducing each strand individually and subsequently ligating them, where the 5' phosphate reacts with the 3' hydroxyl to create a continuous strand with a T4 ligase. This process formed a circular DNA strand, which was iteratively purified and ligated after

each strand was added (Figure 1.3C). Two years after the polyhedral cube, the staple structure of a double crossover (DX) tile was introduced. The DX tile comes in five distinct variations, two of which consist of adjacent antiparallel duplexes (DA) and three of parallel duplexes (DP). The two duplexes are tethered together by two different "crossovers", each of which occurs at intervals related to a half turn (~5 bases) of DNA..[33] One example of this was the production of the first 2D lattices that were periodic using sticky end cohesion to assemble several DX tiles.[34]

A novel approach to producing larger nanostructures was created by Rothemund in 2006 to produce large structures using a large viral DNA scaffold and small staple strands (Figure 1.4A). This methodology, termed "DNA origami", allowed for much larger 2D nanostructures (>100 nm) than previous constructs, which required a multitude of smaller tiles that assemble using sticky ended cohesion. DNA origami uses a large (~7kbp) single stranded scaffold (that comes from the M13 bacteriophage) that is routed to create the overall shape as a single stranded entity, which is then "stapled" together using short oligonucleotides to form duplexes.[35] Three years later, the Shih lab expanded origami to the 3rd dimension by stacking the helices in a honeycomb pattern (Figure 1.4B).[36]

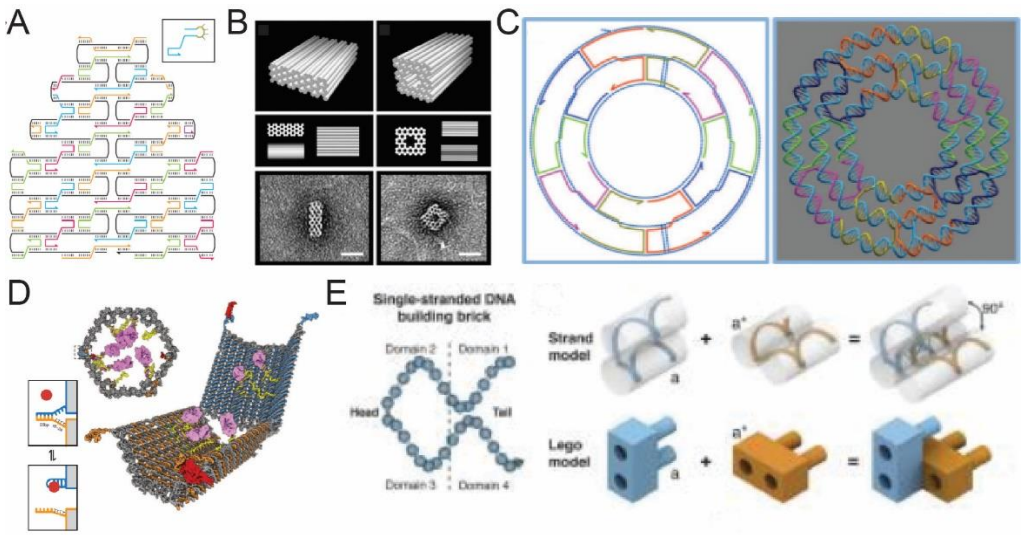


Figure 1.4. DNA nanotechnology milestones in later years. (A) Schematic of DNA origami methodology[35] (B) First 3D DNA origami structures with honeycomb packing[37] (C) Introduction of curvature to DNA nanostructures[38] (D) DNA robot to deliver antibody fragments to cells[39] (E) Using short DNA bricks to assemble large structures.[40]

Twenty nine years after Seeman's proposal of using DNA to create a 3D crystal to scaffold proteins for their structural determination, his lab published the first rationally designed 3D crystal. [41] The design employed a "tensegrity triangle" as the basis for the lattice, and will be discussed in detail in chapter 4. The next major contribution to the field came in 2011, when the Yan lab developed several design rules to allow for the development of curved structures by strategically placing crossovers and nicks (Figure 1.4C).[42] The following year, the Church lab demonstrated the first example of a DNA "robot" that was designed to deliver molecular payloads to cells (Figure 1.4D). The "robot" consisted of a DNA origami-based barrel that was held in a closed configuration by a DNA aptamer (short DNA sequence that selectively binds a target) based locks that would open when introduced to their target antigens. To increase the specificity of when the robot would open, two different aptamers were used in concert with one another to act as an "AND" gate, only opening when both had bound their target.[39] The same year, an

alternative approach from origami for producing large DNA structures was developed which utilized short oligos of 32 nt each to create "bricks" that could assemble into the large structure (25x25x27 nm) without having to undergo the time-consuming process of redesigning the scaffold routing that origami requires. Each of the bricks contains a unique DNA sequence that is comprised of four 8 base sticky end domains that can anneal to a neighboring brick at a 90° rotation (Figure 1.4E).[40]

The examples that were covered in this section only give a very brief overview of the expansive amount of progress that has been achieved in the field. The chosen studies highlight major accomplishments towards designing specific geometric parameters (dimension, curvature, size, etc.), the diverse shape and sizes of nanostructures that the DNA has constructed, and many others have been chronicled in the broad array of published DNA nanotechnology reviews.[12, 17, 27, 31, 43, 44] Since the development of the vast array of fundamental concepts, and structures of increasing complexity, DNA nanotechnology has started to focus on introducing these structures for a variety of disciplines such as biosensing, nanoelectronics, biophysics, drug delivery, and synthetic biology.[25, 44-49] Several examples of the successful applications are functionalization for targeted cell delivery, scaffolding of molecules to increase enzyme efficiency through proximity, and templation of liposomes.[24, 50-53]

1.2.4 Shortcomings and Limitations

While it is evident that DNA is an ideal material for building complex structures, it lacks any inherent function on its own, and therefore, encounters several limitations for utilization in biological applications, in part because of the anionic nature of its backbone due to the phosphate group.[12] To mitigate charge-charge repulsion that can occur

between two adjacent backbones, DNA nanostructures are traditionally constructed in buffer solutions that contain much higher cationic salt concentrations (≥12.5 mM Mg^{2+}) than the sub-millimolar divalent concentrations that are typically found in biological environments. The anionic nature of DNA nanostructures can also cause issues with their uptake by cells because they largely prefer positively charged species.[54] Additionally, DNA is susceptible to degradation by nucleases, enzymes that digest DNA, which are ubiquitous in both plant and animal cells. These factors create obstacles that are currently being addressed to usher in the next era of using DNA nanotechnology for use in a broad array of *in vivo* applications.

1.3 Functionalization of DNA Nanostructures with Peptides

1.3.1 Basics of Peptides

DNA nanostructures are limited to the physicochemical properties that are inherent to the nucleic acid building blocks. This limitation can be overcome by incorporating different classes of molecules to create hybrid materials. In recent years, one such material that has been incorporated is another commonly found biopolymer which utilizes amino acids (AAs) as its repeating unit, and unlike the nucleotides that make up DNA, the side chains contain a variety of properties. There are 20 naturally occurring amino acids, each of which contains a backbone with an amine and a carboxylic acid separated by a carbon from which a distinctive side chain that varies in size, composition, and charge. These properties are summarized in Figure 1.5 along with the accepted three and one letter abbreviations for each individual AA.

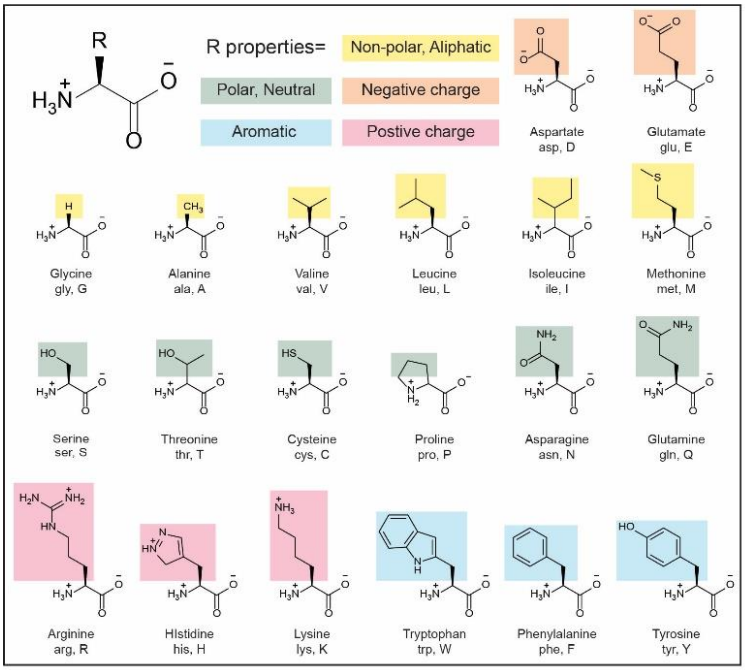


Figure 1.5. Amino acid properties and abbreviations. Each of the amino acids is sorted by their physicochemical properties: yellow is non-polar, green is polar but neutral, orange is negatively charged side chains, pink is positively charged side chains, and blue is aromatic.

While the diverse side chains found in amino acids allow for peptides to be programmed to contain specific functions, they also complicate their ability to self-assemble. For this, only a handful of types of nanostructures constructed entirely of peptides have been realized.[55-58] Similar to the methodology used to produce short oligonucleotides, peptides can be synthesized using solid phase peptide synthesis (SPPS). The most common approach builds sequentially off a solid resin that is functionalized with an amine group to facilitate the first coupling and adds amino acids from the C- to the N-terminus by using a base-labile Fmoc- (fluorenylmethoxycarbonyl) protecting group on the amine group of the amino acid to ensure that only one residue is added at a time. Any reactive side chains are protected by acid-labile groups which are removed

simultaneously as the peptide is cleaved from the resin. After the addition of each amino acid, the Fmoc group is removed to produce a free amine that can undergo a condensation reaction with the next amino acid's free carboxylic acid group. This cycle is repeated until the full peptide has been completed (Figure 1.6)

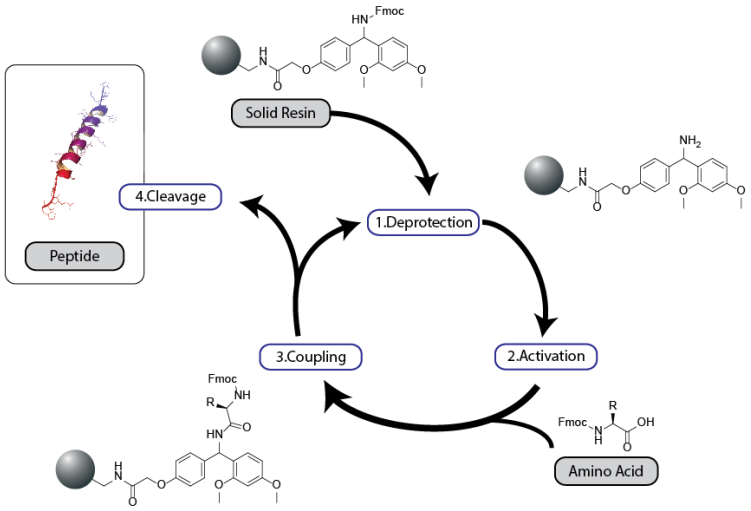


Figure 1.6. Schematic for a generalized Fmoc based solid phase peptide synthesis cycle.

1.3.2 Peptides in Biomaterials

Peptides can range anywhere from 2 to ~50 amino acids in length, and considering that each residue contains one of the 20 AAs it becomes apparent that there are nearly an infinite number of combinations. Therefore, when incorporating peptides into an experimental design, inspiration is most often taken from naturally occurring systems.[59] One example of this is the well-studied arginine-glycine-serine (RGD) peptide segment that is found in the extracellular matrix (ECM). This three-residue motif has been shown to be responsible for cell binding and is therefore often incorporated into various biomaterials.[60-63] In its natural environment, the RGD motif is found on a loop

of the ECM protein fibronectin that binds specifically to integrins, which facilitates cell adhesion.[64-66] Since RGD is typically found in a constrained form innately, several studies have looked at the effect on binding affinity of a linear (flexible) versus a cyclized (constrained) RGD peptide. The results have demonstrated that the cyclized version not only binds at a 200x higher affinity which can be attributed to the lack of entropic penalty for its binding but that it is also more resistant to degradation.[67-71]

Additional peptide motifs and sequence that can be taken from Nature include known targeting peptides which can direct the peptide bearing the molecule to a specific cell receptor. This concept has been exploited in drug delivery applications to reduce off target effects, specifically in targeting a cancer cell that overexpresses a particular receptor. One example of this is the computationally designed peptide P51 which was found to target the HER2 receptor which has been shown to be overexpressed in 25% of invasive breast cancers.[72] One study found that the P51 peptide bound to the HER2 receptor with high specificity at an 18.6 nmol/L affinity.[73] While this is only a singular example of a targeting peptide, the concept could be extended to other cell receptors to minimize off target effects and increase the cell delivery efficiency.

Cell penetrating peptides (CPPs) and endosome escape peptides (EEPs) are also routinely incorporated into drug delivery systems to help them traverse the membrane, and subsequently be released from the endosome to allow for cytosolic delivery. CPPs were discovered based on the natural Tat protein of HIV and the *Drosophila melanogaster* Antennapedia homeodomain protein structures. Both proteins were observed to overcome a cell membrane barrier, and the proteins were each truncated incrementally to find the shortest peptide sequences that could still carry out the function

efficiently outside the context of the full protein. When examining the composition of the Tat peptide, it becomes apparent that it is due to the highly cationic nature of the six arginine and two lysine residues. Further studies indicated that the success of this CPP could be attributed solely to its high arginine content as alternative cationic residues did not cross the cell membrane as efficiently.[74-77] These arginine rich peptides have also been shown to avoid endosomal entrapment.[78] Unfortunately, highly cationic molecules have also been shown to be toxic to cells at higher concentrations[79, 80] which effectively disallows their use for *in vitro and in vivo* applications. An alternative 13- residue EEP has been reported and shown to increase cytosolic delivery by 5-fold when incorporated into a system. This peptide, aurein 1.2, is relatively neutral and was not found to induce cytotoxic effects to the cells.[81]

Another peptide motif of interest to the scientific community are the short binding regions of antibodies, which are large proteins used by the immune system to neutralize foreign materials. IgG antibodies contain a complementarity determining region (CDR) that presents three short peptide loops that contain sequences that bind to a specific antigen. Each of the three loops is only 4-8 amino acids long but they work in concert to achieve high specificity. The ability to scaffold these binding loops onto novel biomaterials has become an area of interest since antibodies can be expensive, and difficult to produce and modify.[82-84]

1.4 A Look at DNA-peptide Materials

The ability to synthesize both DNA and peptides make them ideal candidates for the creation of novel hybrid materials. There are two main approaches to making a DNA- peptide hybrid: (1) using electrostatic interactions that facilitate the assembly of

the negatively charged DNA and positively charged peptides to come together, and (2) creation of peptide-oligonucleotide conjugates (POC) as a novel building block. The latter approach is more common because there are many forces at play when all the psychochemical properties of both the nucleotide and amino acids are taken into consideration, making the final outcome difficult to predict. One way to remedy this issue is to find a way to control which of these forces will predominate. Initial studies carried out by the Lim lab took a POC and showed that they could alter what shape was formed by altering the assembly conditions to control whether the DNA structure of the self-assembling peptide portion would assemble first.[85] While this field is still relatively new, it has been successfully applied to a variety of applications including creating protein mimetics[86, 87], antisense and other gene therapeutics[88-90], templated protein and peptide synthesis[91, 92], as imaging agents[93], display and control over cell signals[62, 63, 94, 95], and the formation of nanowires[96] amongst many others. The additional applications of improving DNA stability *in vivo*[97-100] and using POCs as a method to orient peptides (often in a multivalent fashion) [83, 101-105] will all be discussed in later chapters.

1.4.1 DNA-Peptide Hybrids from Electrostatic Forces

One way to produce a DNA-peptide hybrid material is to exploit the anionic nature of the phosphate backbone and the canonical cationic amino acid residues (arginine, lysine, histidine). By simply mixing the two materials they will spontaneously self-assemble according to their electrostatic interactions. While this approach is facile and requires no additional reactions or purification steps to create the hybrid material, it lacks the ability to control the presentation of the peptides and the DNA.

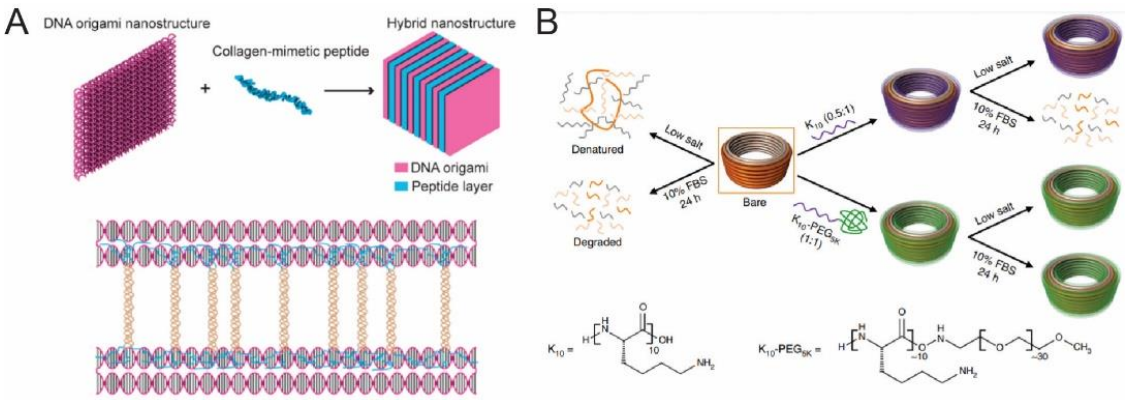


Figure 1.7. Examples of DNA-peptide hybrids based on electrostatic interactions. A) A fiber composed of DNA origami rectangles and collagen mimetic peptides with cationic termini.[106] B) Cationic peptide coating for stabilization of a DNA nanostructure in a low salt environment.[99]

There are several examples for this methodology, all of which use a positively charged amine- or guanidinium- based AA side chain to associate the peptide to the anionic DNA structure. One study used a rectangular origami structure and a collagen mimetic peptide (a triple helical structure) with several arginine residues on the termini that could facilitate the co-assembly. The formation of the hybrid structure was controlled by the dimensions of the DNA origami structure and the periodicity of the repeating unit was determined to be a combination of the height of the DNA and peptide helix length (Figure 1.7A). [106] Additional studies examined the use of both polyarginine and polylysine peptides to coat 3D DNA structures. In these instances, the peptides behaved as an intramolecular counterion to mitigate the charge-charge repulsion of the backbone to overcome the need for high concentrations of divalent salts. This peptide "coating" was found to stabilize the nanostructure in low salt solutions for over 24 hours; however, it was susceptible to enzymatic degradation until the oligolysine was modified with a polyethylene glycol (PEG) tail (Figure 1.7B). These oligolysine coatings were

also studied through molecular simulations to determine the packing of the hybrid material and found that the number of amino acids is the polyamine determines whether or not the molecule acts simply as a counterion or if it also allows for the additional stabilizing aspect of bridge adjacent DNA duplexes.[99, 100, 107, 108]

The electrostatic interactions that occur between peptides and DNA duplexes causing their co-assembly are routinely found in nature in the form of DNA- binding peptides/proteins. Traditionally, this class of molecules contain transcription factors such as homeodomains, AT-hooks, and zinc fingers.[109] Unlike the previous examples that utilized poly-cationic peptides to bind randomly, these motifs have preferential nucleotide binding sequences, allowing for their incorporation into DNA nanostructures in a specific fashion. Instances where exploitation of these properties could be desirable will be discussed in a future chapter.

1.4.2 Covalent DNA-Peptide Hybrid Materials

The second approach to creating DNA-peptide hybrids requires the two biopolymers to be covalently linked to create POCs, an approach that offers several advantages over those which are simply mixed and allowed to self-assemble. One approach is to control the number of copies of both the peptide and the oligonucleotide that are incorporated into the structure, with a second possibility to design the orientation of both components with respect to one another. The synthetic route for POCs can be grouped into two distinct categories: on- resin synthesis for the entire peptide-oligonucleotide hybrid, and a fragmented approach where each component is synthesized individually, and then subsequently linked.[97]

Complete on-resin synthesis is used infrequently because peptide and DNA solid phase synthesis approaches are not compatible with one another, so additional modifications are necessary. This approach *does* have the advantage, however, that it avoids the use of linkers that allow for further control over the orientation of the two materials. Additionally, all functional groups that are present are traditionally protected during the entire synthesis which prohibits any potential side reactions from occurring. The biggest obstacle towards making this approach feasible is that DNA can undergo a depurination reaction when exposed to an acid, such as the trifluoroacetic acid commonly used to cleave peptides from the resin and remove the amino acid side chain protecting groups. Despite this incompatibility, there have been several publications that have demonstrated that POCs can be successfully synthesized using this method. These approaches typically use a branched linker off the resin that contain an amine that can be utilized for the peptide, and a hydroxyl that can be used for the DNA (Figure 1.8A). [110-112] An alternative approach for complete on resin synthesis uses a form of conjugation for the attachment of the two materials together before being cleaved from the resin, such as a copper click[85, 113, 114] or condensation reaction.[88]

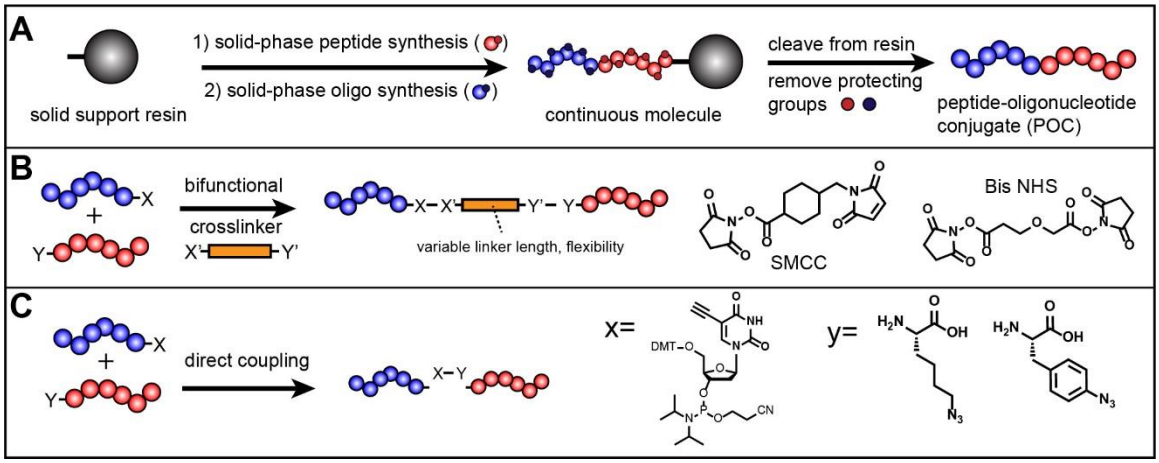


Figure 1.8. Common approaches to POC synthesis. (A) Schematic for a complete on resin approach where the peptide and oligonucleotide is synthesized, consecutively. (B) Schematic showing the use of a bifunctional linker with the examples of SMCC which connects an amine and a thiol and a bis-NHS which connects two amines. (C) Schematic demonstrating the fragmented approach that incorporates functional groups using noncanonical amino acids and phosphoramidites with an example showing an alkyne modified thymine and an azide modified lysine and phenylalanine for copper click chemistry.

The more common approach uses a fragmented synthesis, where the DNA and the peptide are synthesized separately, and then conjugated to produce the POC by utilizing either a bifunctional linker or functionalization with a bio-orthogonal group on the phosphoramidites or amino acids. Several bifunctional linkers (Figure 1.8B) that are employed are a bis-NHS (N-hydroxysuccinimyl) which reacts with amine groups[115], SMCC which uses an NHS group on one side and a maleimide on the other to react with a thiol[83, 91, 116, 117], and a bifunctional alkyne that could be utilized for click chemistry.[118] Other reported methods included the synthesis of POCs using common coupling techniques such as DCC[87, 119] or NHS-EDC couplings.[98, 120] However, employing a bifunctional linker does limit what functional groups can be present in the amino acid

side chains, for example, any time a maleimide group is present, there can only be a single cysteine present to ensure site specificity of the linkage.

One way to overcome this shortcoming is by the addition of a bio-orthogonal functional group to the peptide or oligonucleotide (Figure 1.8C). A common example of this approach uses click chemistry, a class of high-yielding reactions that are fast, simple, and regiospecific.[121] Specific examples of this are the copper assisted azide alkyne cycloaddition or "copper click" and the strain promoted azide alkyne cycloaddition (SPaaC) or "copper-free click", which will be discussed in length in chapter 3. While the copper click struggles to work on two larger biomolecules such as a peptide and oligonucleotide, it has been used several times.[114] The copper free click however, works well and traditionally uses the noncanonical amino acids azido phenylalanine or azido lysine as the source of azide to react with the cyclooctyne (CO) that can be introduced to the DNA using an amine modification and NHS-CO.[63, 86, 122] Additional functionalization examples have used an oxime and thiazolidine formation[123], addition of an aldehyde onto a uracil base[124], enzymatic ligation,[120, 125] and introduction of a thiol moiety onto the DNA.[126-129]

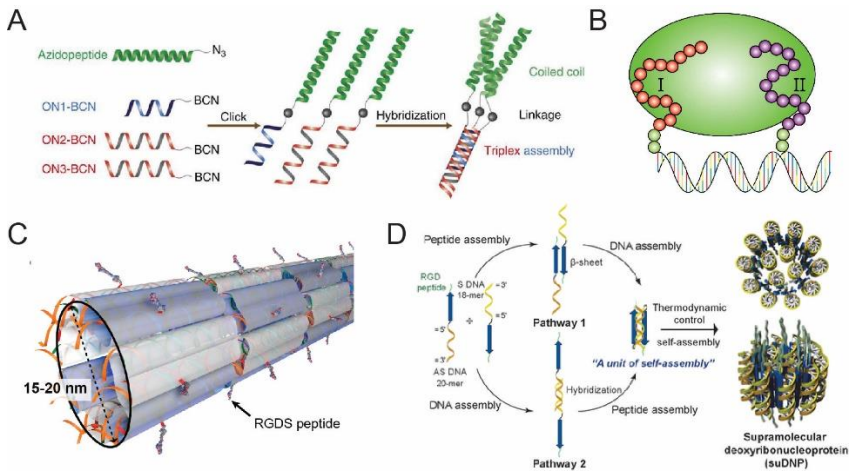


Figure 1.9. Examples of POCs in the literature. (A) Protein mimetic composed of a singular POC to form a DNA triplex-coiled-coil nanostructure.[86] (B) A POC based synbody where the DNA duplex is used to determine the proper distance for two binding peptides to obtain optimal binding affinity.[83] (C) DNA nanotube that presents the cell adhesion peptide RGD.[62] (D) A POC that was studied under different assembly conditions to control whether the DNA structure or self-assembling peptide could assemble first.[85]

Several examples of POCs have illustrated current applications for which they can be utilized.[62, 95, 97, 105, 111, 122, 130] The first example uses three POCs that assembled two orthogonal assembly principles, one of which forms a DNA triplex, and an alternate method using a coiled-coil containing three α-helices. This design was described as a mimetic to a protein with higher order structure. Upon further investigation, it was revealed that this hybrid material was more thermodynamically stable than each of the components individually (Figure 1.9A).[86] A second example used a simple DNA duplex to scaffold two different binding peptides at discrete distances to determine the optimal spacing for the highest binding affinity. This multivalent scaffolding of the peptides is an early example of a synthetic antibody or "synbody" (Figure 1.9B). The two binding peptides were originally found to bind the Gal80 protein through a screening of 4000 different 12- residue peptides. Upon discovery of two peptides that weakly bound the

protein, the two peptides were conjugated with the DNA at distances ranging from 4.3 to 9.2 nanometers. Each of these POC synbodies were then tested for binding affinity using surface plasmon resonance (SPR).[83]

A third example created a POC that contained RGD. The DNA-RGD conjugate was incorporated into a DNA nanotube so that the binding peptide could be presented in a multivalent fashion throughout the structure (Figure 1.9C). It was then demonstrated that these binding peptides were sufficient to adhere neural stem cells to surfaces coated with the DNA tube to enhance their differentiation into neurons.[62] A final example used a POC to demonstrate how the self-assembly forces could be controlled. One POC that was synthesized contained a short oligonucleotide for the assembly of a nanostructure attached to a β-sheet forming peptide that would self-assemble to create another protein mimetic in the form of a deoxyribonucleoprotein. Some of the peptides also contained the RGD epitope to aid in cellular uptake by facilitating cell adhesion. A second POC, which also contained the β-sheet forming peptide, used a 20-nt antisense DNA strand for green fluorescent protein (GFP). The authors then determined which conditions would facilitate the initial assembly of the DNA or peptide portion would be preferable. Interestingly, they observed that regardless of which self-assembling motif was the leading source, the same nanostructure was obtained.[85]

1.5 Overview of Dissertation

This dissertation will describe the work that has been carried out throughout my doctoral studies relating to the field of DNA nanotechnology. This chapter has provided an overview for the start and current state of the field, to date. Additionally, it gives a look at a sub-field that incorporates peptides to create hybrid DNA-peptide materials.

The presentation provides a foundation necessary to understand the importance and scope of the work that will be discussed in the following chapters. Chapter 2 will discuss the work that has been carried out using cationic peptides to electrostatically coat DNA nanostructures. These coatings were designed to not only increase the DNA structure's stability in biological environments by behaving as a counterion, but also to introduce functional peptides that allow for endosomal escape. Chapter 3 will also focus on DNA-peptide hybrid materials that are covalently linked, and will describe several pathways used to produce POCs, and their intended applications which include the multivalent scaffolding of binding peptides to create synbodies. It will also look at a continuation of the conjugation to produce a DNA-peptide-DNA (DPD) tri- and penta-block polymers to produce peptides at discrete locations.

Chapter 4 will pivot from DNA-peptide hybrid materials to introduce a detailed look at the foundational principles of self-assembling 3D DNA crystals. Specifically, it will introduce the basics of DNA based crystallography, and discuss what design parameters of these lattices need to be evaluated to fully understand how they assemble. The work that has been carried out in order to understand several of these parameters will then be discussed in chapters 5-7. In chapter 5, the third published DNA crystal motif (the 4x6) will be detailed. This study was the first step towards the elucidation of the effect that the number of interjunction bases plays on the overall packing of the DNA lattice. In addition to elucidating the structural aspects, this chapter will also discuss additional design characteristics that were explored with the motif, namely the ability to design and control the handedness of the lattice, and the ability to design an asymmetric lattice with unique sequences with virtues that will be discussed later. Both of these

parameters are important steps towards expanding the application possibilities for the crystals. Chapter 6 will describe a detailed study to determine the effect that a unique Holliday junction sequence had upon the packing of the crystal. The work is the first to describe the structure, preferred angle, and preference to selectively bind ions on all possible 36 immobile junction sequence combinations. To date, all self-assembling DNA crystals utilized a singular junction termed "J1" as the ideal junction, but the work proved that this notion was incorrect, indicating that the choice of nucleotides at each junction can dramatically impact the ability to successfully design and construct all 3D DNA nanostructures.

The comprehensive investigation discussed in chapter 6 was then applied to produce a novel rhombohedral crystal lattice that will be introduced in chapter 7. The detailed work informed a penultimate study that reported that by transforming the sequence of the original "holy grail" J1 junction resulted in a crystal lattice that diffracted to the highest reported resolution (2.7Å) in DNA crystals, thus far. This study also explored the capability of altering cavity size by adding additional an additional interhelical turn, as well as determining the effect of sticky end length and sequence on crystal quality and resolution. The design parameters of the 3D DNA crystals that were uncovered in the previous chapter will then be applied toward reaching the ultimate goal of DNA nanotechnology by utilizing the scaffolds as a way to encapsulate biological guests for high resolution structural determination, and thus achieving the foundational goal of structural DNA nanotechnology. Chapter 8 will summarize an entire collection of projects that have systematically approached the realization of Seeman's proposal, including the incorporation and structural determination of several well studied minor

groove binders that serve as a proof of concept that the proposed application is feasible, and the incorporation of various DNA-binding proteins and peptides show significant progress towards the advancement of the 40-year-old goal.

To conclude, Chapter 9 will coalesce around all of the work that was carried out and discussed in previous chapters, while offering a perspective for broader impacts of each individual study. Both arms of DNA nanotechnology that are discussed in my comprehensive thesis for DNA-peptide hybrid materials and DNA crystalline scaffolds are both in their infancy; however, they describe a comprehensive map towards achieving many long sought-after goals of DNA nanotechnology.

1.6 References

1. Feynman, R. P., There's plenty of room at the bottom [data storage]. *Journal of microelectromechanical systems* **1992,** *1* (1), 60-66.

2. Kaiser, D., Richard Feynman: A Life in Science. John Gribbin , Mary Gribbin. *Isis* **2001,** *92* (1), 207-208.

3. Yurchenco, P. D.; Birk, D. E.; Mecham, R. P., *Extracellular matrix assembly and structure*. Academic Press: San Diego, California, 1994.

4. Bishop, K. J. M.; Wilmer, C. E.; Soh, S.; Grzybowski, B. A., Nanoscale Forces and Their Uses in Self-Assembly. *Small (Weinheim an der Bergstrasse, Germany)* **2009,** *5* (14), 1600-1630.

5. Yadav, S.; Sharma, A. K.; Kumar, P., Nanoscale Self-Assembly for Therapeutic Delivery. *Frontiers in bioengineering and biotechnology* **2020,** *8*, 127-127.

6. Ninham, B. W., *Molecular forces and self assembly in colloid, nano sciences and biology*. Cambridge University Press: Cambridge, U.K, 2010.

7. Lee, Y. S., *Self-Assembly and Nanotechnology Systems: Design, Characterization, and Applications*. John Wiley & Sons, Incorporated: New York, 2011.

8. Bensaude-Vincent, B., Self-Assembly, Self-Organization: Nanotechnology and Vitalism. *Nanoethics* **2009,** *3* (1), 31-42.

9. Merrifield, R. B., Solid Phase Synthesis (Nobel Lecture). *Angewandte Chemie (International Edition in English)* **1985,** *24(10),* 799-810.

10. Ito, H.; Ike, Y.; Ikuta, S.; Itakura, K., Solid phase synthesis of polynucleotides. VI. Farther studies on polystyrene copolymers for the solid support. *Nucleic acids research* **1982,** *10* (5), 1755-1769.

11. Eastaugh, B.; Sternal-Johnson, C., Recontruction of M.C.Escher Depth. Exploriment Design, 2011.

12. Pinheiro, A. V.; Han, D. R.; Shih, W. M.; Yan, H., Challenges and opportunities for structural DNA nanotechnology. *Nature Nanotechnology* **2011,** *6* (12), 763-772.

13. Seeman, N. C., DNA Nanotechnology at 40. *Nano letters* **2020,** *20* (3), 1477-1478.

14. Pray, L., Discovery of DNA Structure and Function: Watson and Crick. *Nature Education* **2008,** *1:1(100)*.

15. Watson, J. D.; Crick, F. H., Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. J.D. Watson and F.H.C. Crick. Published in Nature, number 4356 April 25, 1953. *Nature (London)* **1974,** *248* (5451), 765-765.

16. Qiu, H.; Dewan, J. C.; Seeman, N. C., A DNA decamer with a sticky end: the crystal structure of d-CGACGATCGT. *Journal of molecular biology* **1997,** *267* (4), 881-898.

17. Seeman, N. C., Biochemistry and structural DNA nanotechnology: An evolving symbiotic relationship. *Biochemistry* **2003,** *42* (24), 7259-7269.

18. Kowalczykowski, S. C.; Dixon, D. A.; Eggleston, A. K.; Lauder, S. D.; Rehrauer, W. M., Biochemistry of homologous recombination in Escherichia coli. *Microbiological reviews* **1994,** *58* (3), 401-465.

19. Holliday, R., A mechanism for gene conversion in fungi (Reprinted). *Genetics Research* **2007,** *89* (5-6), 285-307.

20. Clegg, R. M.; Murchie, A. I. H.; Zechel, A.; Carlberg, C.; Diekmann, S.; Lilley, D. M. J., FLUORESCENCE RESONANCE ENERGY-TRANSFER ANALYSIS OF THE STRUCTURE OF THE 4-WAY DNA JUNCTION. *Biochemistry* **1992,** *31* (20), 4846-4856.

21. Duckett, D. R.; Murchie, A. I. H.; Diekmann, S.; Vonkitzing, E.; Kemper, B.; Lilley, D. M. J., THE STRUCTURE OF THE HOLLIDAY JUNCTION, AND ITS RESOLUTION. *Cell* **1988,** *55* (1), 79-89.

22. Kallenbach, N. R.; Ma, R. I.; Seeman, N. C., AN IMMOBILE NUCLEIC-ACID JUNCTION CONSTRUCTED FROM OLIGONUCLEOTIDES. *Nature* **1983,** *305* (5937), 829-831.

23. Seeman, N. C., NUCLEIC-ACID JUNCTIONS AND LATTICES. *Journal of Theoretical Biology* **1982,** *99* (2), 237-247.

24. Zadegan, R. M.; Norton, M. L., Structural DNA Nanotechnology: From Design to Applications. *International Journal of Molecular Sciences* **2012,** *13* (6), 7149-7162.

25. Zhang, F.; Nangreave, J.; Liu, Y.; Yan, H., Structural DNA Nanotechnology: State of the Art and Future Perspective. *Journal of the American Chemical Society* **2014,** *136* (32), 11198-11211.

26. Ke, Y.; Castro, C.; Choi, J. H., Structural DNA Nanotechnology: Artificial Nanostructures for Biomedical Research. *Annual review of biomedical engineering* **2018,** *20* (1), 375-401.

27. Zhang, C. A.; He, Y.; Su, M.; Ko, S. H.; Ye, T.; Leng, Y. J.; Sun, X. P.; Ribbe, A. E.; Jiang, W.; Mao, C. D., DNA self-assembly: from 2D to 3D. *Faraday Discussions* **2009,** *143*, 221-233.

28. Zhang, F.; Jiang, S. X.; Wu, S. Y.; Li, Y. L.; Mao, C. D.; Liu, Y.; Yan, H., Complex wireframe DNA origami nanostructures with multi-arm junction vertices. *Nature Nanotechnology* **2015,** *10* (9), 779-+.

29. He, Y.; Tian, Y.; Ribbe, A. E.; Mao, C. D., Highly connected two-dimensional crystals of DNA six-point-stars. *Journal of the American Chemical Society* **2006,** *128* (50), 15978-15979.

30. Ma Ri Fau - Kallenbach, N. R.; Kallenbach Nr Fau - Sheardy, R. D.; Sheardy Rd Fau - Petrillo, M. L.; Petrillo Ml Fau - Seeman, N. C.; Seeman, N. C., Three-arm nucleic acid junctions are flexible. (0305-1048 (Print)).

31. Seeman, N. C., DNA nanotechnology: Novel DNA constructions. *Annual Review of Biophysics and Biomolecular Structure* **1998,** *27*, 225-248.

32. Chen, J. H.; Seeman, N. C., SYNTHESIS FROM DNA OF A MOLECULE WITH THE CONNECTIVITY OF A CUBE. *Nature* **1991,** *350* (6319), 631-633.

33. Fu, T. J.; Seeman, N. C., DNA DOUBLE-CROSSOVER MOLECULES. *Biochemistry* **1993,** *32* (13), 3211-3220.

34. Winfree, E.; Liu, F. R.; Wenzler, L. A.; Seeman, N. C., Design and self-assembly of two-dimensional DNA crystals. *Nature* **1998,** *394* (6693), 539-544.

35. Rothemund, P. W. K., Folding DNA to create nanoscale shapes and patterns. *Nature* **2006,** *440* (7082), 297-302.

36. Douglas, S. M.; Dietz, H.; Liedl, T.; Hogberg, B.; Graf, F.; Shih, W. M., Self-assembly of DNA into nanoscale three-dimensional shapes (vol 459, pg 414, 2009). *Nature* **2009,** *459* (7250), 1154-1154.

37. Douglas, S. M.; Dietz, H.; Liedl, T.; Hogberg, B.; Graf, F.; Shih, W. M., Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **2009,** *459* (7245), 414-418.

38. Han, D. R.; Pal, S.; Nangreave, J.; Deng, Z. T.; Liu, Y.; Yan, H., DNA Origami with Complex Curvatures in Three-Dimensional Space. *Science* **2011,** *332* (6027), 342-346.

39. Douglas, S. M.; Bachelet, I.; Church, G. M., A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science* **2012,** *335* (6070), 831-834.

40. Ke, Y. G.; Ong, L. L.; Shih, W. M.; Yin, P., Three-Dimensional Structures Self-Assembled from DNA Bricks. *Science* **2012,** *338* (6111), 1177-1183.

41. Zheng, J. P.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R. J.; Constantinou, P. E.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **2009,** *461* (7260), 74-77.

42. Dongran, H.; Suchetan, P.; Jeanette, N.; Zhengtao, D.; Yan, L.; Hao, Y., DNA Origami with Complex Curvatures in Three-Dimensional Space. *Science (American Association for the Advancement of Science)* **2011,** *332* (6027), 342-346.

43. Seeman, N. C., DNA engineering and its application to nanotechnology. *Trends in Biotechnology* **1999,** *17* (11), 437-443.

44. Seeman, N. C.; Sleiman, H. F., DNA nanotechnology. *Nature Reviews Materials* **2018,** *3* (1).

45. Zahid, M.; Kim, B.; Hussain, R.; Amin, R.; Park, S. H., DNA nanotechnology: a future perspective. *Nanoscale research letters* **2013,** *8* (1), 1-13.

46. Aldaye, F. A.; Palmer, A. L.; Sleiman, H. F., Assembling materials with DNA as the guide. *Science* **2008,** *321* (5897), 1795-1799.

47. Aldaye, F. A.; Sleiman, H. F., Supramolecular DNA nanotechnology. *Pure and Applied Chemistry* **2009,** *81* (12), 2157-2181.

48. Zhang, Q.; Jiang, Q.; Li, N.; Dai, L. R.; Liu, Q.; Song, L. L.; Wang, J. Y.; Li, Y. Q.; Tian, J.; Ding, B. Q.; Du, Y., DNA Origami as an In Vivo Drug Delivery Vehicle for Cancer Therapy. *Acs Nano* **2014,** *8* (7), 6633-6643.

49. Madhanagopal, B. R.; Zhang, S. Q.; Demirel, E.; Wady, H.; Chandrasekaran, A. R., DNA Nanocarriers: Programmed to Deliver. *Trends in Biochemical Sciences* **2018,** *43* (12), 997-1013.

50. Hu, Y.; Niemeyer, C. M., From DNA Nanotechnology to Material Systems Engineering. *Advanced Materials* **2019,** *31* (26).

51. Liu, M.; Fu, J.; Hejesen, C.; Yang, Y.; Woodbury, N. W.; Gothelf, K.; Liu, Y.; Yan, H., A DNA tweezer-actuated enzyme nanoreactor. *Nature communications* **2013,** *4* (1), 2127-2127.

52. Liu, M.; Fu, J.; Qi, X.; Wootten, S.; Woodbury, N. W.; Liu, Y.; Yan, H., A Three-Enzyme Pathway with an Optimised Geometric Arrangement to Facilitate Substrate Transfer. *Chembiochem : a European journal of chemical biology* **2016,** *17* (12), 1097-1101.

53. Yang, Y.; Wang, J.; Shigematsu, H.; Xu, W.; Shih, W. M.; Rothman, J. E.; Lin, C., Self-assembly of size-controlled liposomes on DNA nanotemplates. *Nature chemistry* **2016,** *8* (5), 476-483.

54. Fröhlich, E., The role of surface charge in cellular uptake and cytotoxicity of medical nanoparticles. *International journal of nanomedicine* **2012,** *7*, 5577-5591.

55. Gradišar, H.; Božič, S.; Doles, T.; Vengust, D.; Hafner-Bratkovič, I.; Mertelj, A.; Webb, B.; Šali, A.; Klavžar, S.; Jerala, R., Design of a single-chain polypeptide tetrahedron assembled from coiled-coil segments. *Nature chemical biology* **2013,** *9* (6), 362-366.

56. Habibi, N.; Kamaly, N.; Memic, A.; Shafiee, H., Self-assembled peptide-based nanostructures: Smart nanomaterials toward targeted drug delivery. *Nano today* **2016,** *11* (1), 41-60.

57. Jennifer, E. P.; Christos, C.; Todd, O. Y., Nanohedra: Using Symmetry to Design Self Assembling Protein Cages, Layers, Crystals, and Filaments. *Proceedings of the National Academy of Sciences - PNAS* **2001,** *98* (5), 2217-2221.

58. Doles, T.; Božič, S.; Gradišar, H.; Jerala, R., Functional self-assembling polypeptide bionanomaterials. *Biochemical Society transactions* **2012,** *40* (4), 629-634.

59. Vassilev, T. L.; Kazatchkine, M. D.; Van Huyen, J. P. D.; Mekrache, M.; Bonnin, E.; Mani, J. C.; Lecroubier, C.; Korinth, D.; Baruch, D.; Schriever, F.; Kaveri, S. V., Inhibition of cell adhesion by antibodies to Arg-Gly-Asp (RGD) in normal immunoglobulin for therapeutic use (intravenous immunoglobulin, IVIg). *Blood* **1999,** *93* (11), 3624-3631.

60. Van Agthoven, J. F.; Xiong, J.-P.; Alonso, J. L.; Rui, X.; Adair, B. D.; Goodman, S. L.; Arnaout, M. A., Structural basis for pure antagonism of integrin αVβ3

by a high-affinity form of fibronectin. *Nature structural & molecular biology* **2014,** *21* (4), 383-388.

61. Assa-Munt, N.; Jia, X.; Laakkonen, P.; Ruoslahti, E., Solution structures and integrin binding activities of an RGD peptide with two isomers. *Biochemistry* **2001,** *40* (8), 2373-2378.

62. Stephanopoulos, N.; Freeman, R.; North, H. A.; Sur, S.; Jeong, S. J.; Tantakitti, F.; Kessler, J. A.; Stupp, S. I., Bioactive DNA-Peptide Nanotubes Enhance the Differentiation of Neural Stem Cells Into Neurons. *Nano Letters* **2015,** *15* (1), 603-609.

63. Freeman, R.; Stephanopoulos, N.; Alvarez, Z.; Lewis, J. A.; Sur, S.; Serrano, C. M.; Boekhoven, J.; Lee, S. S.; Stupp, S. I., Instructing cells with programmable peptide DNA hybrids. *Nature Communications* **2017,** *8*.

64. Liskamp, R.; Rijkers, D.; Bakker, S., *Bioactive Macrocyclic Peptides and Peptide Mimics*. Wiley-VCH: 2008.

65. Finke, A.; Busskamp, H.; Manea, M.; Marx, A., Designer Extracellular Matrix Based on DNA-Peptide Networks Generated by Polymerase Chain Reaction. *Angewandte Chemie-International Edition* **2016,** *55* (34), 10136-10140.

66. Ruoslahti, E., Fibronectin and its receptors. *Annual Review of Biochemistry* **1988,** *57*, 375-413.

67. Bogdanowich-Knipp, S. J.; Jois, D. S. S.; Siahaan, T. J., The effect of conformation on the solution stability of linear vs. cyclic RGD peptides. *Journal of Peptide Research* **1999,** *53* (5), 523-529.

68. Kapp, T. G.; Rechenmacher, F.; Neubauer, S.; Maltsev, O. V.; Cavalcanti-Adam, E. A.; Zarka, R.; Reuning, U.; Notni, J.; Wester, H.-J.; Mas-Moruno, C.; Spatz, J.; Geiger, B.; Kessler, H., A comprehensive evaluation of the activity and selectivity profile of ligands for RGD-binding integrins. *Scientific reports* **2017,** *7* (1), 39805-39805.

69. Roxin, A.; Zheng, G., Flexible or fixed: a comparative review of

linear and cyclic cancer-targeting peptides. *Future Med. Chem.* **2012,** *4(12)*, 1601-1618.

70. Patel, P. R.; Kiser, R. C.; Lu, Y. Y.; Fong, E.; Ho, W. C.; Tirrell, D. A.; Grubbs, R. H., Synthesis and Cell Adhesive Properties of Linear and Cyclic RGD Functionalized Polynorbornene Thin Films. *Biomacromolecules* **2012,** *13* (8), 2546-2553.

71. Koivunen, E.; Wang, B.; Ruoslahti, E., Phage Libraries Displaying Cyclic Peptides with Different Ring Sizes: Ligand Specificities of the RGD-Directed Integrins. *Nature biotechnology* **1995,** *13* (3), 265-270.

72. Subik, K.; Lee, J.-F.; Baxter, L.; Strzepek, T.; Costello, D.; Crowley, P.; Xing, L.; Hung, M.-C.; Bonfiglio, T.; Hicks, D. G.; Tang, P., The expression patterns of ER, PR, HER2, CK5/6, EGFR, KI-67 and AR by immunohistochemical analysis in breast cancer cell lines. *Breast cancer : basic and clinical research* **2010,** *4* (1), 35-41.

73. Geng, L.; Wang, Z.; Jia, X.; Han, Q.; Xiang, Z.; Li, D.; Yang, X.; Zhang, D.; Bu, X.; Wang, W.; Hu, Z.; Fang, Q., HER2 targeting peptides screening and applications in tumor imaging and drug delivery. *Theranostics* **2016,** *6* (8), 1261-1273.

74. Tian, Y.; Zhou, M. X.; Shi, H. G.; Gao, S. J.; Xie, G. C.; Zhu, M.; Wu, M.; Chen, J.; Niu, Z. W., Integration of Cell-Penetrating Peptides with Rod-like Bionanoparticles: Virus-Inspired Gene-Silencing Technology. *Nano Letters* **2018,** *18* (9), 5453-5460.

75. Lebleu, B.; Moulton, H. M.; Abes, R.; Ivanova, G. D.; Abes, S.; Stein, D. A.; Iversen, P. L.; Arzumanov, A. A.; Gait, M. J., Cell penetrating peptide conjugates of steric block oligonucleotides. *Advanced Drug Delivery Reviews* **2008,** *60* (4-5), 517-529.

76. Vives, E.; Schmidt, J.; Pelegrin, A., Cell-penetrating and cell-targeting peptides in drug delivery. *Biochimica Et Biophysica Acta-Reviews on Cancer* **2008,** *1786* (2), 126-138.

77. Wright, L.; Rothbard, J.; Wender, P., Guanidinium rich peptide transporters and drug delivery. *Current Protein Pept Sci* **2003,** *4(2)*, 105-124.

78. El-Sayed, A.; Futaki, S.; Harashima, H., Delivery of Macromolecules Using Arginine-Rich Cell-Penetrating Peptides: Ways to Overcome Endosomal Entrapment. *Aaps Journal* **2009,** *11* (1), 13-22.

79. Neundorf, I.; Rennert, R.; Hoyer, J.; Schramm, F.; Löbner, K.; Kitanovic, I.; Wölfl, S., Fusion of a short HA2-derived peptide sequence to cell-penetrating peptides improves cytosolic uptake, but enhances cytotoxic activity. *Pharmaceuticals (Basel, Switzerland)* **2009,** *2* (2), 49-65.

80. Sugita, T.; Yoshikawa, T.; Mukai, Y.; Yamanada, N.; Imai, S.; Nagano, K.; Yoshida, Y.; Shibata, H.; Yoshioka, Y.; Nakagawa, S.; Kamada, H.; Tsunoda, S. i.; Tsutsumi, Y., Comparative study on transduction and toxicity of protein transduction domains. *British journal of pharmacology* **2008,** *153* (6), 1143-1152.

81. Li, M.; Tao, Y.; Shu, Y. L.; LaRochelle, J. R.; Steinauer, A.; Thompson, D.; Schepartz, A.; Chen, Z. Y.; Liu, D. R., Discovery and Characterization of a Peptide That Enhances Endosomal Escape of Delivered Proteins in Vitro and in Vivo. *Journal of the American Chemical Society* **2015,** *137* (44), 14084-14093.

82. Strohl, W. R., Current progress in innovative engineered antibodies. *Protein & cell* **2018,** *9* (1), 86-120.

83. Williams, B. A. R.; Diehnelt, C. W.; Belcher, P.; Greving, M.; Woodbury, N. W.; Johnston, S. A.; Chaput, J. C., Creating Protein Affinity Reagents by Combining Peptide Ligands on Synthetic DNA Scaffolds. *Journal of the American Chemical Society* **2009,** *131* (47), 17233-17241.

84. Gupta, N.; Lainson, J. C.; Belcher, P. E.; Shen, L. H.; Mason, H. S.; Johnston, S. A.; Diehnelt, C. W., Cross-Reactive Synbody Affinity Ligands for Capturing Diverse Noroviruses. *Analytical Chemistry* **2017,** *89* (13), 7174-7181.

85. Kye, M.; Lim, Y. B., Reciprocal Self-Assembly of Peptide-DNA Conjugates into a Programmable Sub-10-nm Supramolecular Deoxyribonucleoprotein. *Angewandte Chemie-International Edition* **2016,** *55* (39), 12003-12007.

86. Lou, C. G.; Martos-Maldonado, M. C.; Madsen, C. S.; Thomsen, R. P.; Midtgaard, S. R.; Christensen, N. J.; Kjems, J.; Thulstrup, P. W.; Wengel, J.; Jensen, K. J., Peptide-oligonucleotide conjugates as nanoscale building blocks for assembly of an artificial three-helix protein mimic. *Nature Communications* **2016,** *7*.

87. Patutina, O. A.; Bichenkova, E. V.; Miroshnichenko, S. K.; Mironova, N. L.; Trivoluzzi, L. T.; Burusco, K. K.; Bryce, R. A.; Vlassov, V. V.; Zenkova, M. A., miRNases: Novel peptide-oligonucleotide bioconjugates that silence miR-21 in lymphosarcoma cells. *Biomaterials* **2017,** *122*, 163-178.

88. Zubin, E. M.; Romanova, E. A.; Volkov, E. M.; Tashlitsky, V. N.; Korshunova, G. A.; Shabarova, Z. A.; Oretskaya, T. S., Oligonucleotide-peptide conjugates as potential antisense agents. *Febs Letters* **1999,** *456* (1), 59-62.

89. Arar, K.; Aubertin, A. M.; Roche, A. C.; Monsigny, M.; Mayer, R., SYNTHESIS AND ANTIVIRAL ACTIVITY OF PEPTIDE-OLIGONUCLEOTIDE CONJUGATES PREPARED BY USING N-ALPHA-(BROMOACETYL)PEPTIDES. *Bioconjugate Chemistry* **1995,** *6* (5), 573-577.

90. Astakhova, K.; Ray, R.; Taskova, M.; Uhd, J.; Carstens, A.; Morris, K., "Clicking" Gene Therapeutics: A Successful Union of Chemistry and Biomedicine for New Solutions. *Molecular pharmaceutics* **2018,** *15* (8), 2892-2899.

91. Spruijt, E.; Tusk, S. E.; Bayley, H., DNA scaffolds support stable and uniform peptide nanopores. *Nature Nanotechnology* **2018,** *13* (8), 739-+.

92. Bruick, R. K.; Dawson, P. E.; Kent, S. B.; Usman, N.; Joyce, G. F., Template-directed ligation of peptides to oligonucleotides. *Chemistry & Biology* **1996,** *3* (1), 49-56.

93. Dovgan, I.; Koniev, O.; Kolodych, S.; Wagner, A., Antibody–Oligonucleotide Conjugates as Therapeutic, Imaging, and Detection Agents. *Bioconjugate chemistry* **2019,** *30* (10), 2483-2501.

94. Stephanopoulos, N.; Ortony, J. H.; Stupp, S. I., Self-assembly for the synthesis of functional biomaterials. *Acta Materialia* **2013,** *61* (3), 912-930.

95. Stephanopoulos, N., Peptide–Oligonucleotide Hybrid Molecules for Bioactive Nanomaterials. *Bioconjugate chemistry* **2019,** *30* (7), 1915-1922.

96. Nazari, Z. E.; Herrero, J. G.; Fojan, P.; Gurevich, L., Formation of Conductive DNA-Based Nanowires via Conjugation of dsDNA with Cationic Peptide. *Nanomaterials* **2017,** *7* (6).

97. Tung, C.-H.; Stein, S., Preparation and Applications of Peptide−Oligonucleotide Conjugates. *Bioconjugate chemistry* **2000,** *11* (5), 605-618.

98. Charbgoo, F.; Alibolandi, M.; Taghdisi, S. M.; Abnous, K.; Soltani, F.; Ramezani, M., MUC1 aptamer-targeted DNA micelles for dual tumor therapy using doxorubicin and KLA peptide. *Nanomedicine-Nanotechnology Biology and Medicine* **2018,** *14* (3), 685-697.

99. Ponnuswamy, N.; Bastings, M. M. C.; Nathwani, B.; Ryu, J. H.; Chou, L. Y. T.; Vinther, M.; Li, W. A.; Anastassacos, F. M.; Mooney, D. J.; Shih, W. M., Oligolysine-based coating protects DNA nanostructures from low-salt denaturation and nuclease degradation. *Nature Communications* **2017,** *8*.

100. Wang, S.-T.; Gray, M. A.; Xuan, S.; Lin, Y.; Byrnes, J.; Nguyen, A. I.; Todorova, N.; Stevens, M. M.; Bertozzi, C. R.; Zuckermann, R. N.; Gang, O., DNA origami protection and molecular interfacing through engineered sequence-defined peptoids. *Proceedings of the National Academy of Sciences - PNAS* **2020,** *117* (12), 6339-6348.

101. Flory, J. D.; Shinde, S.; Lin, S.; Liu, Y.; Yan, H.; Ghirlanda, G.; Fromme, P., PNA-Peptide Assembly in a 3D DNA Nanocage at Room Temperature. *Journal of the American Chemical Society* **2013,** *135* (18), 6985-6993.

102. Janssen, B. M. G.; Lempens, E. H. M.; Olijve, L. L. C.; Voets, I. K.; Dongen, v. J. L. J.; Greef, d. T. F. A.; Merkx, M., Reversible blocking of antibodies using bivalent peptide-DNA conjugates allows protease-activatable targeting. *Chemical science (Cambridge)* **2013,** *4* (4), 1442-1450.

103. Marczynke, M.; Gröger, K.; Seitz, O., Selective Binders of the Tandem Src Homology 2 Domains in Syk and Zap70 Protein Kinases by DNA-Programmed Spatial Screening. *Bioconjugate chemistry* **2017,** *28* (9), 2384-2392.

104. Eberhard, H.; Diezmann, F.; Seitz, O., DNA as a Molecular Ruler: Interrogation of a Tandem SH2 Domain with Self-Assembled, Bivalent DNA–Peptide Complexes. *Angewandte Chemie (International ed.)* **2011,** *50* (18), 4146-4150.

105. MacCulloch, T.; Buchberger, A.; Stephanopoulos, N., Emerging applications of peptide-oligonucleotide conjugates: bioactive scaffolds, self-assembling systems, and hybrid nanomaterials. *Organic & Biomolecular Chemistry* **2019,** *17* (7), 1668-1682.

106. Jiang, T.; Meyer, T. A.; Modlin, C.; Zuo, X.; Conticello, V. P.; Ke, Y., Structurally Ordered Nanowire Formation from Co-Assembly of DNA Origami and Collagen-Mimetic Peptides. *Journal of the American Chemical Society* **2017,** *139* (40), 14025-14028.

107. Anastassacos, F. M.; Zhao, Z.; Zeng, Y.; Shih, W. M., Glutaraldehyde Cross-Linking of Oligolysines Coating DNA Origami Greatly Reduces Susceptibility to Nuclease Degradation. *Journal of the American Chemical Society* **2020,** *142* (7), 3311-3315.

108. Roodhuizen, J. A. L.; Hendrikx, P.; Hilbers, P. A. J.; de Greef, T. F. A.; Markvoort, A. J., Counterion-Dependent Mechanisms of DNA Origami Nanostructure Stabilization Revealed by Atomistic Molecular Simulation. *Acs Nano* **2019,** *13* (9), 10798-10809.

109. Pazos, E.; Mosquera, J.; Vazquez, E.; Mascarenas, J., DNA Recognition by Synthetic Constructs. *ChemBioChem* **2011,** *12*, 1958-1973.

110. Bergmann, F.; Bannwarth, W., Solid phase synthesis of directly linked peptide-oligodeoxynucleotide hybrids using standard synthesis protocols. *Tetrahedron letters* **1995,** *36* (11), 1839-1842.

111. Eritja, R.; Pons, A.; Escarceller, M.; Giralt, E.; Albericio, F., SYNTHESIS OF DEFINED PEPTIDE-OLIGONUCLEOTIDE HYBRIDS CONTAINING A NUCLEAR TRANSPORT SIGNAL SEQUENCE. *Tetrahedron* **1991,** *47* (24), 4113-4120.

112. Lu, K.; Duan, Q. P.; Ma, L.; Zhao, D. X., Chemical Strategies for the Synthesis of Peptide-Oligonucleotide Conjugates. *Bioconjugate Chemistry* **2010,** *21* (2), 187-202.

113. Kye, M.; Lim, Y. B., Synthesis and purification of self-assembling peptide-oligonucleotide conjugates by solid-phase peptide fragment condensation. *Journal of Peptide Science* **2018,** *24* (8-9).

114. Gogoi, K.; Mane, M. V.; Kunte, S. S.; Kumar, V. A., A versatile method for the preparation of conjugates of peptides with DNA/PNA/analog by employing chemo-selective click reaction in water. *Nucleic Acids Research* **2007,** *35* (21).

115. Kong, D. H.; Lei, Y.; Yeung, W.; Hili, R., Enzymatic Synthesis of Sequence-Defined Synthetic Nucleic Acid Polymers with Diverse Functional Groups. *Angewandte Chemie-International Edition* **2016,** *55* (42), 13164-13168.

116. Williams, B. A. R.; Lund, K.; Liu, Y.; Yan, H.; Chaput, J. C., Self-assembled peptide nanoarrays: An approach to studying protein-protein interactions. *Angewandte Chemie-International Edition* **2007,** *46* (17), 3051-3054.

117. Harrison, J. G.; Balasubramanian, S., Synthesis and hybridization analysis of a small library of peptide-oligonucleotide conjugates. *Nucleic Acids Research* **1998,** *26* (13), 3136-3145.

118. Merg, A. D.; Thaner, R. V.; Mokashi-Punekar, S.; Nguyen, S. T.; Rosi, N. L., Triblock peptide-oligonucleotide chimeras (POCs): Programmable biomolecules for the assembly of morphologically tunable and responsive hybrid materials. *Chemical communications (Cambridge, England)* **2017,** *53* (90), 12221-12224.

119. Abraham, J. N.; Gour, N.; Bolisetty, S.; Mezzenga, R.; Nardin, C., Controlled aggregation of peptide-DNA hybrids into amyloid-like fibrils. *European Polymer Journal* **2015,** *65*, 268-275.

120. Carter, J.; LaBean, T., Coupling Strategies for the Synthesis of Peptide-Oligonucleotide Conjugates for Patterned Synthetic Biomineralization. *Journal of Nucleic Acids* **2011**.

121. Hein, C. D.; Liu, X.-M.; Wang, D., Click Chemistry, A Powerful Tool for Pharmaceutical Sciences. *Pharmaceutical research* **2008,** *25* (10), 2216-2230.

122. Buchberger, A.; Simmons, C. R.; Fahmi, N. E.; Freeman, R.; Stephanopoulos, N., Hierarchical Assembly of Nucleic Acid/Coiled-Coil Peptide Nanostructures. *Journal of the American Chemical Society* **2020,** *142* (3), 1406-1416.

123. Forget, D.; Boturyn, D.; Defrancq, E.; Lhomme, J.; Dumy, P., Highly efficient synthesis of peptide-oligonucleotide conjugates: Chemoselective oxime and thiazolidine formation. *Chemistry-a European Journal* **2001,** *7* (18), 3976-3984.

124. Zatsepin, T. S.; Stetsenko, D. A.; Arzumanov, A. A.; Romanova, E. A.; Gait, M. J.; Oretskaya, T. S., Synthesis of peptide-oligonucleotide conjugates with single and multiple peptides attached to 2 '-aldehydes through thiazolidine, oxime, and hydrazine linkages. *Bioconjugate Chemistry* **2002,** *13* (4), 822-830.

125. Serva, S.; Lagunavicius, A., Direct Conjugation of Peptides and 5-Hydroxymethylcytosine in DNA. *Bioconjugate Chemistry* **2015,** *26* (6), 1008-1012.

126. Rogers, F. A.; Manoharan, M.; Rabinovitch, P.; Ward, D. C.; Glazer, P. M., Peptide conjugates for chromosomal gene targeting by triplex-forming oligonucleotides. *Nucleic Acids Research* **2004,** *32* (22), 6595-6604.

127. Stanojevic, D.; Verdine, G., Deconstruction of GCN4/GCRE into a monomeric peptide-DNA complex. *Nat Struct Mol Biol* **1995,** *2*, 450-457.

128. Chotera, A.; Sadihov, H.; Cohen-Luria, R.; Monnard, P. A.; Ashkenasy, G., Functional Assemblies Emerging in Complex Mixtures of Peptides and Nucleic Acid-Peptide Chimeras. *Chemistry-a European Journal* **2018,** *24* (40), 10128-10135.

129. Pazos, E.; Portela, C.; Penas, C.; Vázquez, M. E.; Mascareñas, J. L., Peptide-DNA conjugates as tailored bivalent binders of the oncoprotein c-Jun. *Organic & biomolecular chemistry* **2015,** *13* (19), 5385-5390.

130. Taskova, M.; Mantsiou, A.; Astakhova, K., Synthetic Nucleic Acid Analogues in Gene Therapy: An Update for Peptide-Oligonucleotide Conjugates. *Chembiochem* **2017,** *18* (17), 1671-1682.

CHAPTER 2

STABILIZATION AND FUNCTIONALIZATION OF DNA NANOSTRUCTURES THROUGH ELECTROSTATIC PEPTIDE COATINGS

2.1 Introduction

2.1.1 Introduction to Drug Delivery Vehicles

DNA nanostructures have long been proposed as ways to deliver cargo to cells because they can be customized with defined shapes and sizes, and several recent studies demonstrated the that these properties can have a large impact on the intake efficiency of the drug delivery vehicles.[1-3] They have also been designed in conformations that contain a cavity that can be loaded with the desired cargo at specific locations with various modifications. However, utilizing DNA based drug delivery vehicles has several downfalls that must be addressed before they can routinely be employed. The first issue is to mitigate charge-charge repulsion between two adjacent phosphate backbones in a compact structure. These systems are assembled in divalent salt conditions (~12.5mM) that far exceed the divalent concentration of typical biological media (~0.6 mM). Additionally, they are susceptible to degradation by nucleases found in cell serum or *in vivo*. While these stability concerns can be problematic when using DNA nanostructures for delivery, there has recently been significant progress towards addressing these issues.[4] Remarkably, only a handful of studies probing DNA nanostructure-cell interactions have been reported. These studies will be necessary before DNA nanostructures can be successfully utilized as delivery vehicles to fully understand their cell internalization pathways and the factors that influence its efficiency.

Nanoparticles can be made from a variety of materials including metals, lipids, and carbon, and are an alternative drug delivery vehicle that has been widely used for various biomedical applications.[5-8] Unlike DNA nanostructures, nanoparticles are not susceptible to degradation; however, their shape and size can be more difficult to control.[9-11] While nanoparticles have found good success for cellular uptake and ability to deliver cargo in a laboratory setting, the results have not translated well to the clinic which may be attributed to a lack of ability to control off-target effects.[12-15] In order to understand why this is the case, experiments that look at nanoparticle-cell interactions must be carried out, which poses many challenges due to the lack of ability to control the size, shape, and modifications in a specific fashion.[5, 16, 17] This impediment indicates that while nanoparticles have long been considered the ideal delivery vehicle, alternative approaches should be explored.

The majority of delivery vehicles that are internalized by cells enter *via* endocytosis. This process occurs when a molecule encounters the cell membrane, where it becomes internalized by the membrane, and then subsequently enveloped into a vesicle that encapsulates the molecule allowing it to enter the intracellular fluid. Upon entering the cell, the vesicle is considered an early endosome with a mildly acidic environment which then travels through the cytosol towards its specific intracellular target or is trafficked back to the cell surface to release its cargo extracellularly. When being further internalized, its environment becomes increasingly acidic to help release any bound signals, before ultimately freeing its cargo via endosomal escape. Alternatively, any cargo that does not escape the endosomal membrane is degraded when it ultimately becomes a lysosome.[18-21] A variety of additives that help to facilitate endosomal escape

have been identified, however the vast majority of the studies identifying these endosomal escape agents are carried out using serum-free media which does not accurately represent conditions *in vivo*.[22-26]

2.1.2 Precedents in DNA Nanostructure Cell Delivery and Stabilization

DNA nanostructures are inherently negatively charged which can lead to issues with cellular uptake as they prefer positively charged species.[27, 28] A variety of modifications have been explored to increase the uptake efficiency of single stranded DNA into cells including addition of lipid anchors,[29-31] incorporating polymers such as peptides,[32, 33] and neutralizing the anionic backbone,[34] however, little work has been done towards increasing the uptake efficiency of DNA nanostructures themselves. Previous work that aimed to probe the interaction of these DNA nanostructures with cells found that the addition of cholesterol modifications on a six-helix bundle increase the cellular association by 10-fold. It was shown that the internalization was due to increased membrane binding since the cholesterol modifications could embed into the hydrophobic lipids of the membrane. Additionally, the authors that with three cholesterol modifications, ~62% of the total nanostructures had been internalized after 4 hours; despite this, after monitoring these bundles for 24 hours they remained colocalized with an endolysosomal dye indicating that they were unable to escape the endsome.[35]

An additional study carried out by the Shih laboratory took a different approach to facilitate DNA nanostructure uptake by neutralizing the DNA backbone using polyamine polymers. [36] This approach not only helps to increase the uptake efficiency, but also helps to stabilize the DNA structures in the low salt environments of biological media by screening the charge repulsion typically observed with adjacent DNA helices. Spermine,

spermidine, oligoarginine, branched polyethylamine, and oligolysine polyamines were all explored. Both spermine and spermidine required very high concentrations to be able to stabilize the origami structures in a short period of time, while both the branched polyethylamine and oligoarginine deformed the structures, leaving oligolysine as the ideal polyamine with an optimized length of 10 residues (K10). To further stabilize the DNA nanostructures from degradation, the K10 was modified with a 5 kD polyethylene glycol (PEG) tail. Subsequently, the coated origami structures were monitored for stability using TEM and FRET experiments, and were found to increase resistance to degradation by ~1000 fold more than the uncoated structures,[36] Additional research carried out by the same lab demonstrated that they could further stabilize the structures by crosslinking the oligolysine coatings with glutaraldehyde.[37]

In concert with the polyamine work carried out by the Shih lab, an atomistic molecular simulation study looked at the mechanism of DNA structure stabilization by spermine and oligolysine. The simulations emphasized that oligolysine was a superior polyamine because K10 could not only shield the charge of the helix, but also bridge two adjacent helices by adhering to the two backbones, whereas spermine was too short and could only attach to a singular helix. Additionally, the work discovered that K10 specifically adheres to the backbone and is unable to bind in the major or minor grooves, which some polyamines are known to do,[38-40] indicating that a K10 molecule will simply coat the structure.[41]

2.2 Results and Discussion

2.2.1 Synthesis and Characterization of Electrostatic Coating Molecules

Inspired by the work carried out by the Shih and Howorka labs, this work aimed to introduce a peptide-based coating that would not only act as a counterion for the phosphate backbone, but also as a way to introduce functional peptides that would be able to facilitate endosomal escape. Most peptides that enable a particle to cross a membrane carry a high positive charge which can be toxic to cells at high concentrations.[42, 43] Alternatively, there is a class of peptides, antimicrobial peptides (AMP), that are known to cross membranes with negligible toxic effects. Screening of these AMP's demonstrated that several were also capable of increasing the efficiency of cytosolic delivery by assisting in endosomal escape. The best AMP was determined to be aurein 1.2, a 13-residue peptide (GLFDIIKKIAESF), which interestingly, is neutral, unlike most known endosome escape peptides.[44] This peptide had previously been attached to DNA nanostructures within the Yan laboratory; however, the researchers were never observed to escape the endosome (work carried out by Xiaodong Qi, unpublished). In these studies, the aurein 1.2 was only incorporated 1-3 times. For a highly multivalent approach, a coating molecule with two copies of the aurein 1.2 peptide was designed. Besides the two endosome escape (EE) peptides (one of the N- and one on the C- terminus, both with a GSG linker), the coating molecule (EE-K10) contained a 10 residue polylysine that would electrostatically adhere to the DNA nanostructure. A plain K10 was also explored to act as a negative control for endosome escape while still containing the counterion stabilizing factor for the coating molecule.

Both the plain K10 and EE-K10 were synthesized using traditional Fmoc based microwave-assisted solid phase peptide synthesis.[45] Briefly, each amino acid was added

to the rink amide resin in a cycle that consisted of N-terminal Fmoc deprotection of the previous residue using a 20% piperidine solution followed by the introduction of the desired amino acid that was activated and coupled to the free N-terminus using diisopropylcarbodiimide, diisopropylethylamine, and oxyma. After addition of the final N-terminal amino acid, the peptide was cleaved from the resin using a 95% solution of trifluoroacetic acid (TFA) with 2.5% triisopropylsilane and 2.5% water which simultaneously removed any side chain protecting groups. The crude peptide was precipitated with cold diethyl ether, pelleted, and resuspended with a 0.1% TFA/water solution. The peptide was then purified using reverse phase high pressure liquid chromatography (RP-HPLC) using a C-18 column and a mobile phase gradient of 0 to 100% acetonitrile. Following purification, the peptides were characterized using matrix assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) (Figure 2.1).

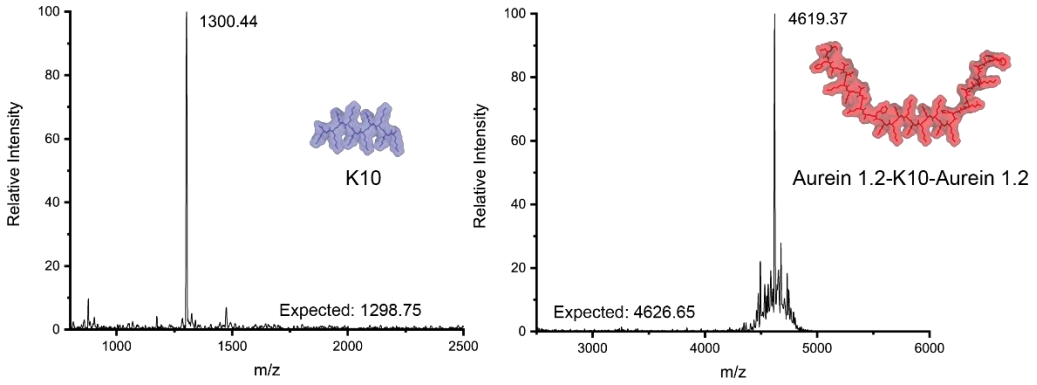


Figure 2.1. Coating Molecule Characterization. Spectra obtained using MALDI-TOF-MS are shown for both the plain K10 coating and the EE-K10 coating molecules alongside 3D models of both.

2.2.2 Coating and Characterization of Functionalized DNA Bundle

For the DNA based model delivery vehicle, a simple six helix bundle (6HB) was chosen as a starting structure, demonstrated that it could effectively be taken up through endocytosis, but would not be released to the cytosol.[35, 46] The structure consists of six different strands that are 50 nt long, each of which participates in two adjacent two-turn duplexes that are connected with a 4-nt poly-T linker region to form a 6x7 nm structure (Figure 2.2).

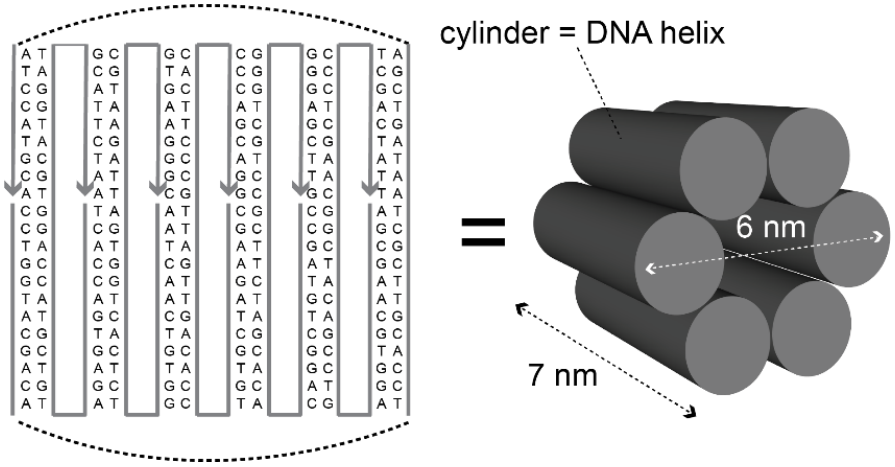


Figure 2.2. Six-Helix Bundle Design. A 2D topology showing the routing for the six strands that makeup the bundle alongside a cartoon depiction showing the dimensions.

The simplicity of this structure allows for any purification steps to be skipped post annealing. Additionally, no aggregation was observed even when the structure was prepared at high concentrations (5-10 μM). The structure was formed by mixing the six individual strands at an equal stoichiometric ratio in a buffer solution (Tris-acetic acid-EDTA pH 8.0) and then annealed using a temperature gradient of 95 to 4°C over 2 hours.

With both the DNA nanostructure and the coating molecules in hand, the next step was to determine the ratio of peptide to DNA in order to fully coat and neutralize the

anionic charge of the phosphate backbone. A simple way to look at this is to consider the nitrogen to phosphate (N:P) ratio where the nitrogen comes from the amine side of the lysine group and the phosphate comes from each nucleotide in the structure. The nanostructure and plain K10 were mixed to screen different N to P ratios (0-2.5:1) with the DNA structure concentration being held constant.

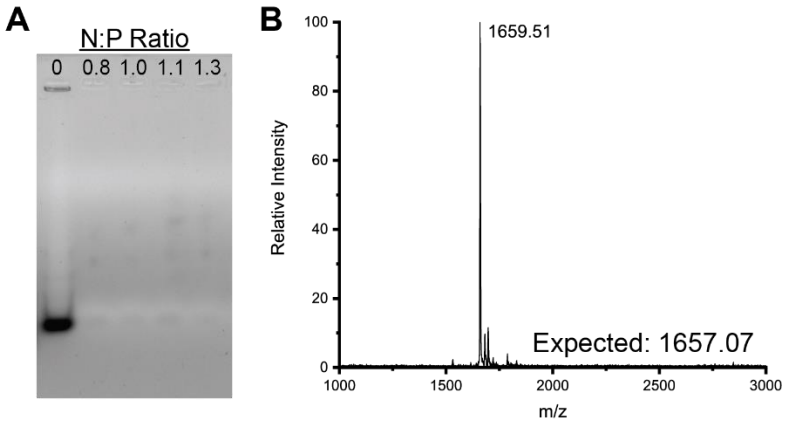


Figure 2.3. K10 coating prevents staining. Agarose gel analysis revealed that near a 1 N:P ratio, the coating molecules prevent intercalation of DNA stains (A). In order to visualize the coated molecules using gel electrophoresis, a K10 labeled with a fluorescein was synthesized (B).

Each of the different ratios was then run on an agarose gel to determine if it had been neutralized, or not. Unfortunately, the K10 coating blocked the DNA stain in the gel (Sybr-Gold) from being able to intercalate into the structure for analysis. To resolve this issue, a new K10 was synthesized with a fluorescein molecule added onto the N-terminus (Figure 2.3A&B). This allowed for the DNA and coated structures to be imaged simultaneously as SYBR-gold and fluorescein have similar excitation and emission wavelengths (495→537 nm and 492→517 nm, respectively). With the new fluorescein labeled K10, the ratio analysis was rerun on a 1.5% agarose gel (Figure 2.4) and consistent with previously described work, at a 1:1 N:P ratio the bundle showed an

upward gel mobility shift indicating that the structure had become larger and less negatively charged. Higher ratios showed signs of aggregation as evidenced by the band smearing, and by the inability to enter the gel, signaling that a 1:1 ratio was optimal. Each of the coating molecules themselves contain 10 lysine residues, while one nanostructure has 300 nucleotides (6 strands x 50 bases), so at a 1:1 ratio there should be ~30 copies of the coating molecule per bundle.

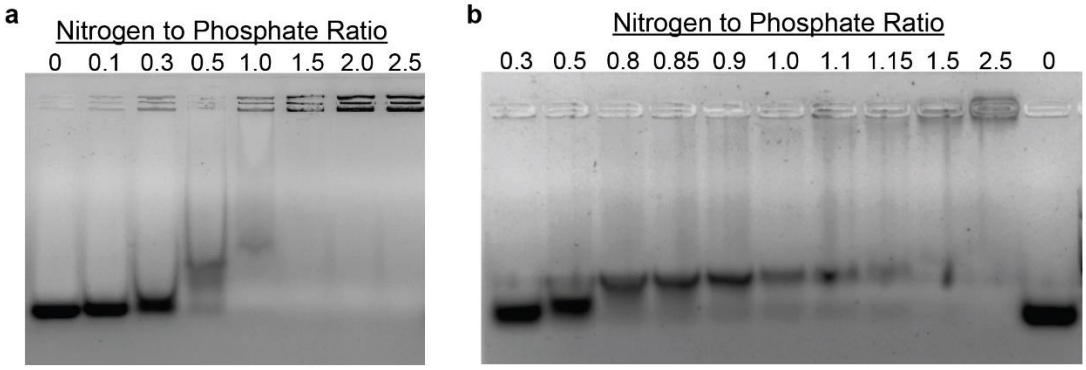


Figure 2.4. Optimization of N:P ratio. A screen of a wide (A) and narrow (B) range of N:P ratios between coating molecule and DNA bundle were characterized on a 1.5% agarose gel using both SYBR Gold and fluorescein labeled K10 for visualization.

With an optimal ratio of N:P determined, a new batch of the 6HB was annealed with an Alexa Fluor 488 dye incorporated on one of the strands so the bundle could be imaged during cell studies. Two additional peptides were also synthesized: (1) a plain K10 with a C-terminal cysteine modified using a maleimide-C2-pHrodo Red fluorophore, and (2) a "scrambled" version of the EE-K10 where the K10 was unaltered, but the sequence for the two aurein 1.2 peptides were rearranged to act as a negative control for endosome escape (Figure 2.5). The pHrodo dye allows for the tracking of endocytosis by monitoring emission intensity, which strengthens as the pH drops, thus indicating a switch from early to late endosome.

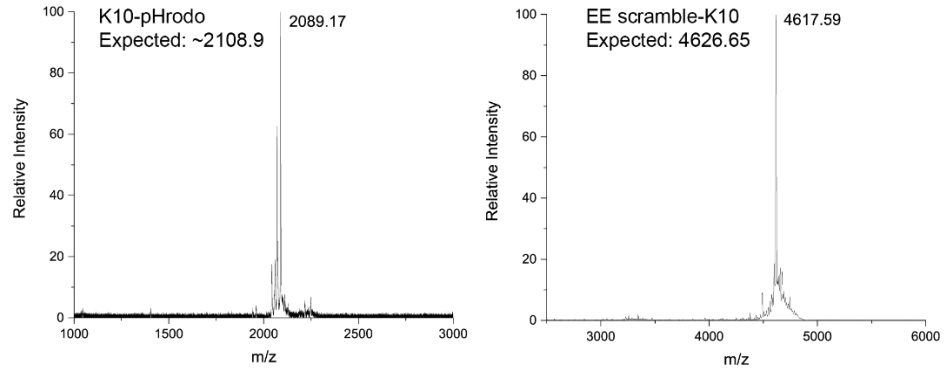


Figure 2.5. Characterization of endocytosis tracking and scrambled peptides. MALDI-MS spectra for the pHrodo labeled K10 and aurein scramble coating molecules. The observed masses correspond well with the expected mass.

Four distinct samples were prepared: (1) uncoated 6HB; or 6HB coated with (2) plain K10; (3) EE-K10; and (4) EE scramble-K10 (Figure 2.6). All three of the coated samples (2-4) incorporated the pHrodo-K10 into the coating mixture at a 20% ratio of the total coating molecules to keep the fluorophore intensity consistent between all samples. Because each bundle has ~30 copies of the molecule, it was estimated that there was ~6 pHrodo dyes and ~48 copies of the aurein or aurein scramble per bundle (24 EE-K10 molecules, each with 2 copies of the bioactive peptide).

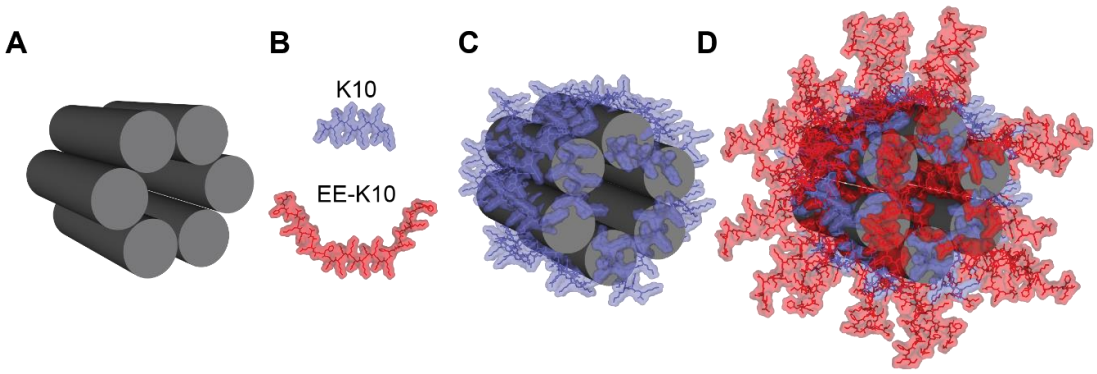


Figure 2.6. Coated DNA nanostructure. Cartoon depictions of the samples of interest prepared for cell analysis: the bare 6HB (A), the coating molecules (B), plain K10 coated 6HB (C), and EE-K10 coated 6HB (D).

2.2.3 DNA Bundle and Cell Interactions

To probe DNA nanostructure delivery, three different hepatic cell were chosen, even though there was no previously reported precedent for comprehensive DNA nanostructure-liver cell interaction studies. However, the majority of materials that are injected into the body are known to be sequestered in the liver, and due to the fact that the most materials that have been previously utilized in the past have displayed interaction with hepatocytes, it was reasonable to hypothesize that DNA nanostructures would also undergo cellular uptake. [47-49] The three cell lines that were chosen were HepG2, Huh7, and Alexander cells as they are readily available and well-understood. Prior to any of the uptake experiments, the potential toxicity of the 6HB, K10-6HB, and EE-K10-6HB was investigated. These experiments were performed by incubating the cell lines with the different structures and comparing the results to cells that were cultured without any addition of the delivery vehicles after a 24-hour incubation period. The cells were then analyzed using an alamarBlue viability assay using cells that had been incubated with

30% methanol (which is toxic) as a positive control. The analysis showed no significant decrease in cell viability upon incubation with any of the three samples of interest (Figure 2.7).

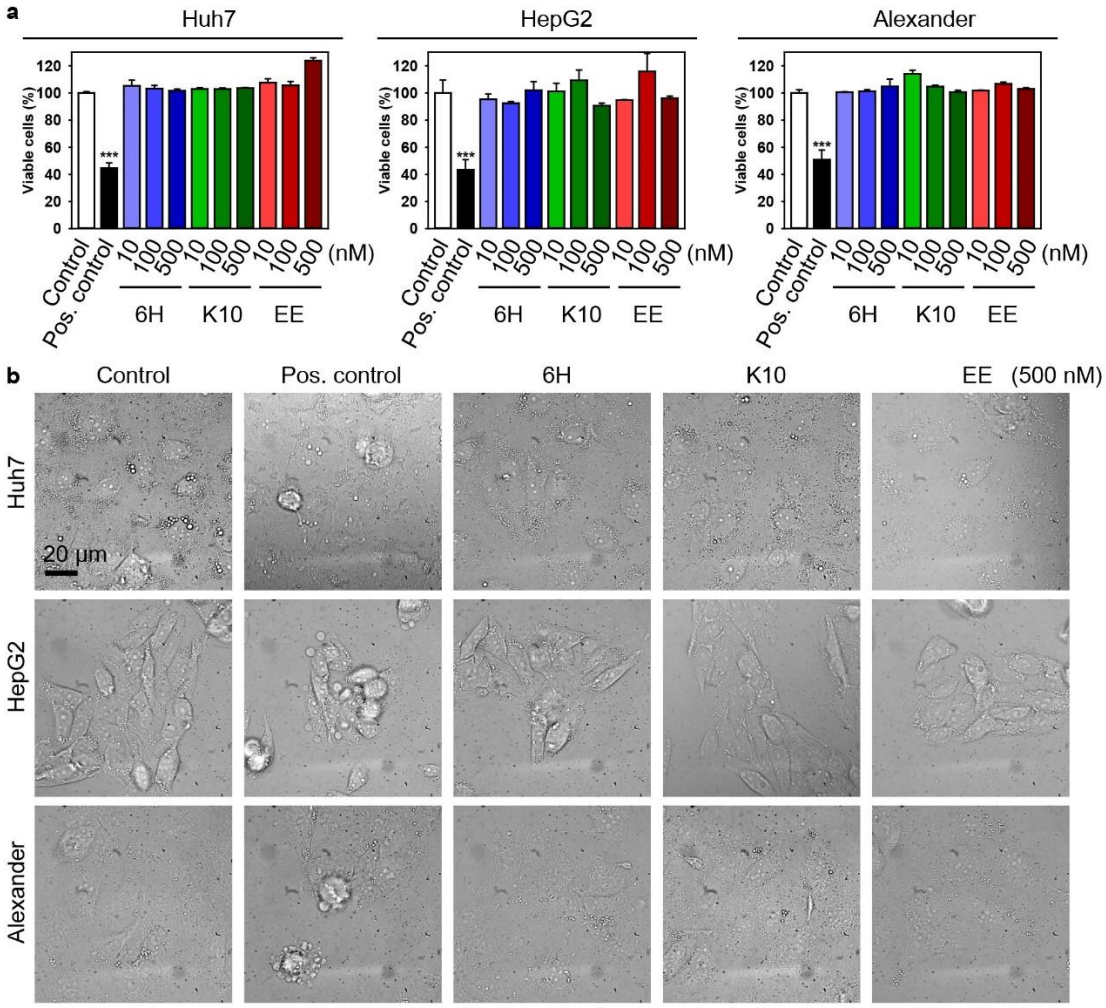


Figure 2.7. Determining the toxic effects of coated nanostructures. (A) Bar graphs showing the percentage of viable cells after incubation with the bare and coated structures at different concentrations in each of the three cell lines. (B) The values were obtained using an alamarBlue assay. Representative images of the fixed cells used to perform the viability assay.

After confirming that the coated structures were biocompatible, their uptake kinetics and efficiencies were examined across the three cells lines using high-resolution spinning disc confocal microscopy run by Dr. Oleg Lunov. The first uptake study that

was carried out simply looked at the concentration of samples that had been internalized across the three different cell lines after the bare and coated structures (10-500 nM screens) were incubated for 24 hours. The number of internalized structures was monitored using the fluorescence of the alexafluor-488 (green) label on the DNA bundle itself by using corrected total cell fluorescence[50] with a sample size of 30 different cells for each variation with representative images displayed in (Figure 2.8A). While there are no observable differences between the uptake efficiency of the samples (6HB, K10-6HB, EE-K10-6HB) there was a significant difference between the three cell lines with Alexander cells having the highest efficiency, followed by Huh7 and HepG2. With each of the cell lines being phenotypically similar, the difference can be attributed to cell size and morphology, with the Alexander cells being the largest, and having the highest uptake efficiency (Figure 2.8C). Each of the three lines displayed a distinct morphology with Alexander being a hexagonal epithelial-like, Huh7 being cuboidal epithelial-like, and HepG2 being elongated. When the internalization was observed at various time points (1, 6, and 24 hours) it was shown that the uptake occurred within the first hour of incubation, and the amount taken up was directly correlated to the cell size which is consistent with previous work that has monitored the uptake efficiency of various materials.[51-54]

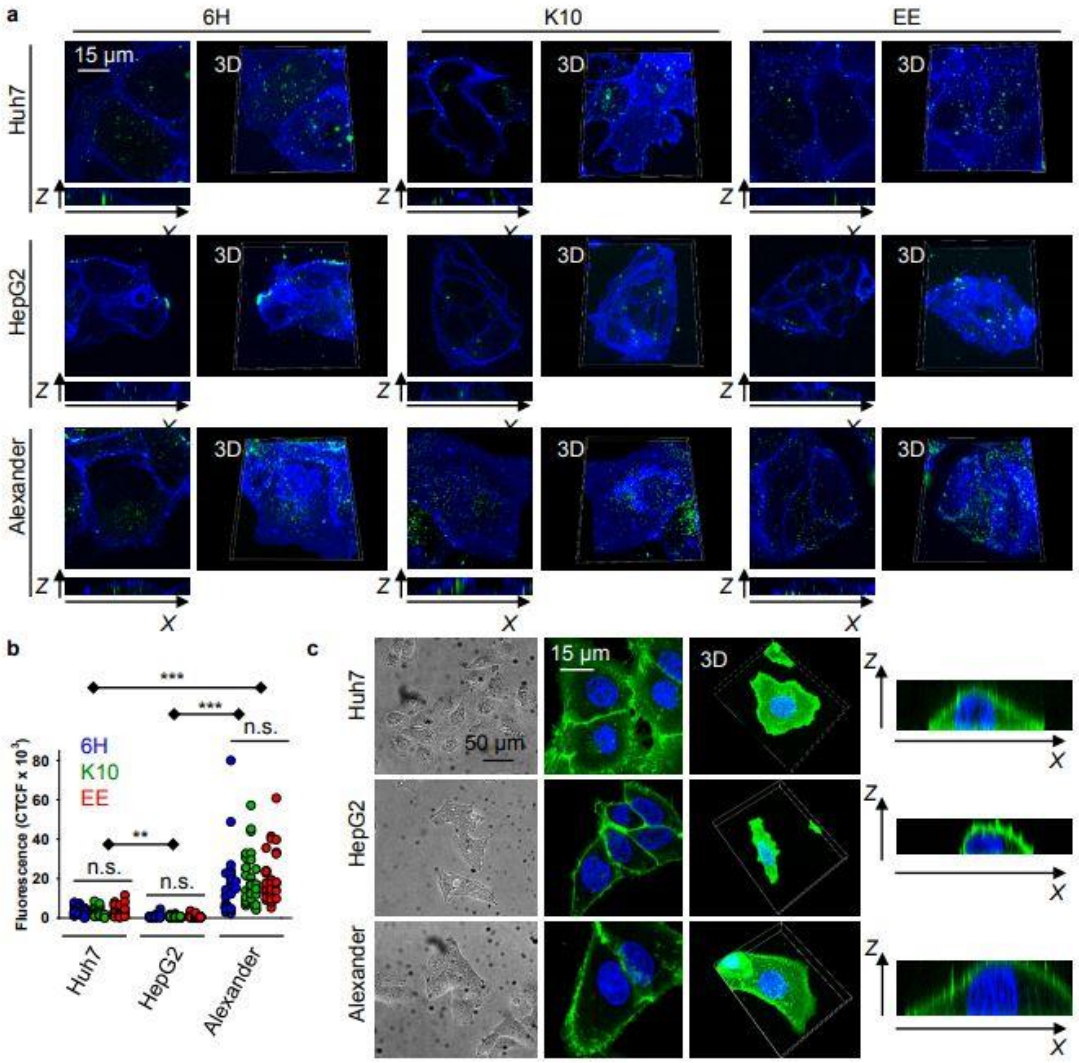


Figure 2.8. Uptake of bare and coated 6HB. Representative images used for calculating the total internalization for each of the three samples across the three different cell lines with the DNA bundles being shown in green and the cell membrane stained with Cellbrite Blue (A) . Quantification of the amount of samples internalized with the 6HB shown in blue, K10-6HB in green, and EE-K10-6HB in red for all three cell lines with statistical shown by (**) $P < 0.01$ and (***) $P < 0.001$ (B). Representative images showing the different cell morphologies with the membranes stained with CellMask Green and nuclei stained with Hoechst 33342 (Blue) along with 3D renderings (C).

2.2.4 Effectiveness of Cytosolic Delivery

With previous experiments that monitored cell delivery using a 6HB that resulted in the nanostructure becoming trapped in the endosome,[35] a common phenomenon with

DNA nanostructures,[55-57] the main focus of this work was to determine if the EE modified K10 coating would enable endosomal escape. For this, the three hepatic cells lines were incubated in serum-free media and media with 10% fetal bovine serum (FBS) with bare 6HB, plain K10-6HB, and EE-K10-6HB for 6 hours, at which point the cells were stained using LysoTracker Blue to visualize the endosomes.

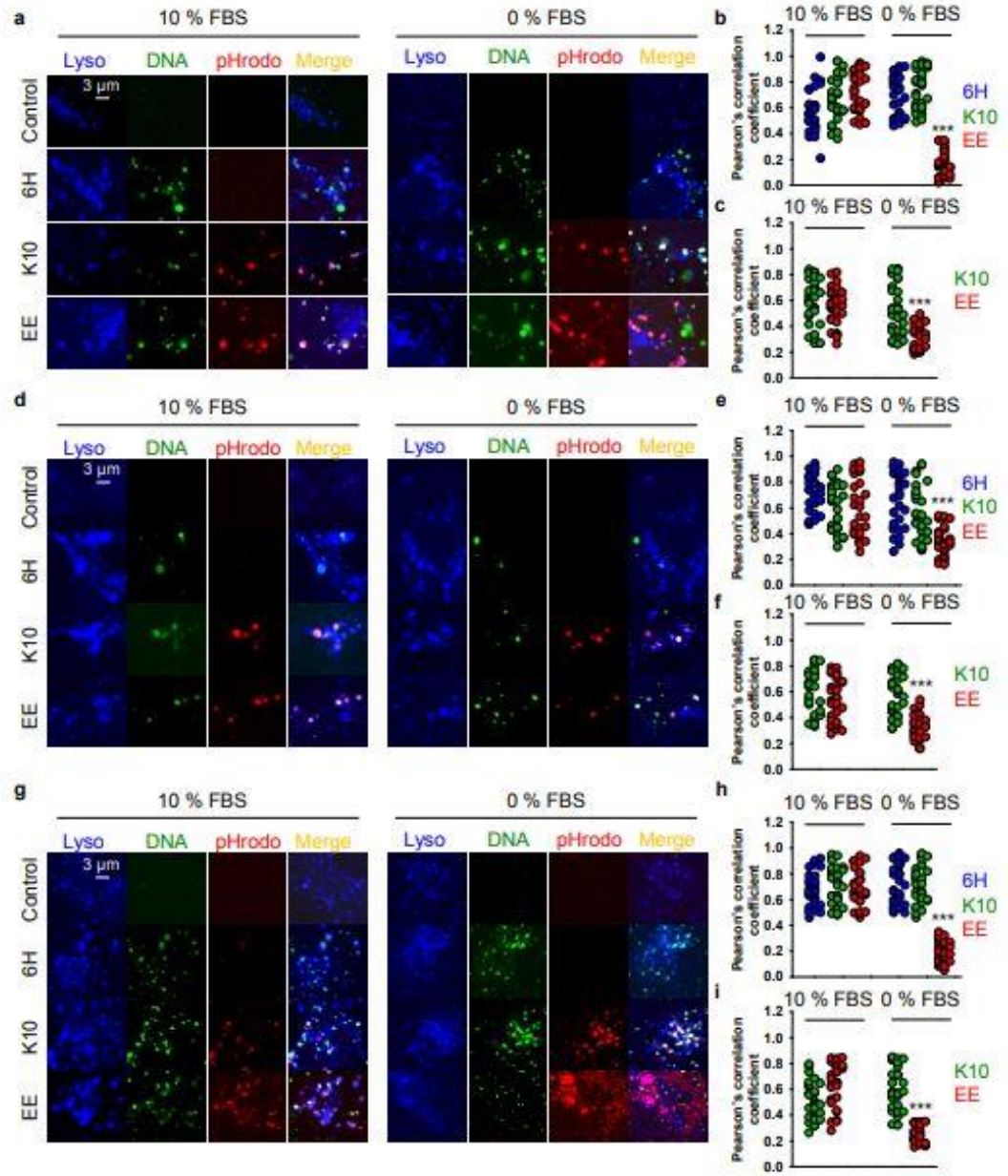


Figure 2.9. Determination of intracellular location of DNA bundles. Each cell line was incubated with each of the sample types with and without serum for 6 hrs. before fluorescence confocal microscopy was used to determine endosomal vesical overlap. Endosomes are stained in blue, 6HB are labeled green, and K10 coatings are red. Fluorescence overlap was analyzed using Pearson's correlations with (***) $P < 0.001$ significance. (A-C) Huh7 cells (D-F) HepG2 and (G-I) Alexander.

Each condition was then analyzed to determine the percentage of overlap between the green (alexafluro-488) labeled DNA bundle, red (pHrodo) labeled K10 coating that was incorporated at 20% ratio of coated structures, and the blue endosomes. The stained cells were then imaged using confocal fluorescence microscopy to determine whether or not the bundles colocalized with the endosomal compartments. A merged image of all the three different fluorophores for the 0% FBS samples revealed that both the bare 6HB and K10-6HB remained in the endosome as indicated by the overlap of the blue, green, and red colors across all three cells (Figure 2.9 A, D,G). On the other hand, there was significant endosomal escape observed with the EE-K10-6HB. Interestingly, the green and red signals also did not overlap, suggesting that upon release into the cytosol the coating molecules are separated from the DNA nanostructures. However, in the experiments that were carried out with the addition of FBS, no endosomal escape was observed. The amount of overlap between fluorophore pairs was analyzed using a Pearson's correlation, and the values were calculated with all samples showing values > 0.5, with the exception of the EE-K10-6HB in the serum free medium. This value was consistantly below 0.5 which indicated that the majority of the structures had in fact made it to the cytosol (Figure 2.9B-I).

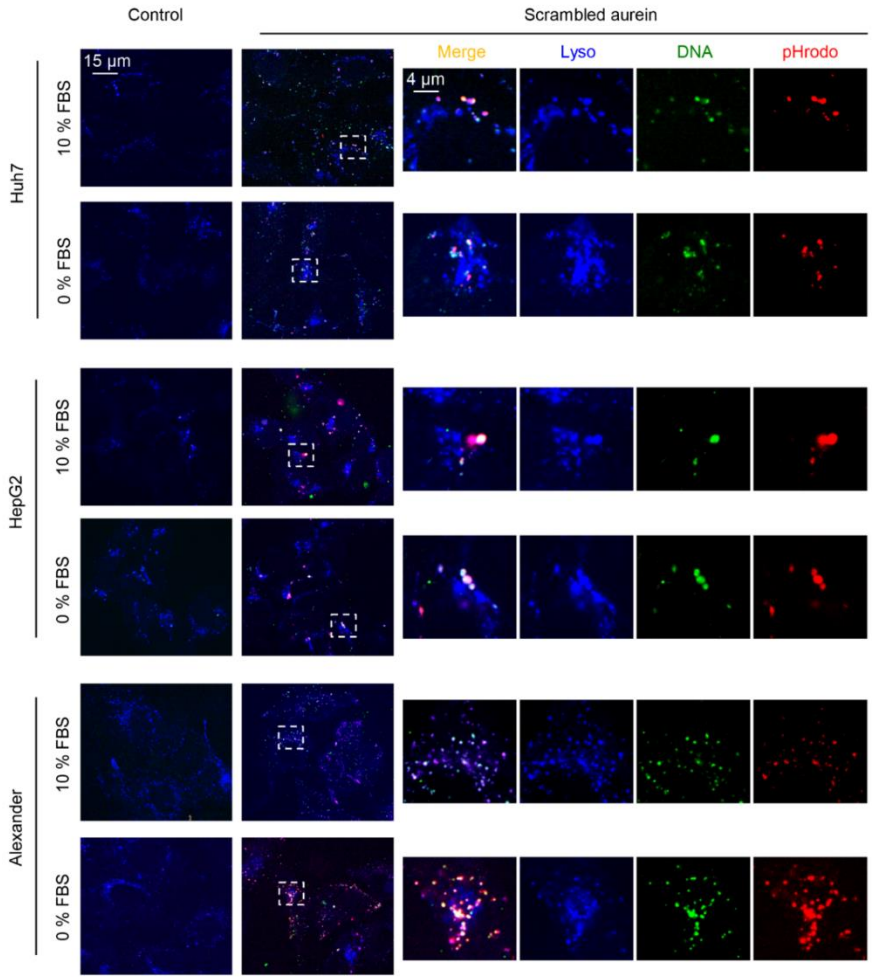
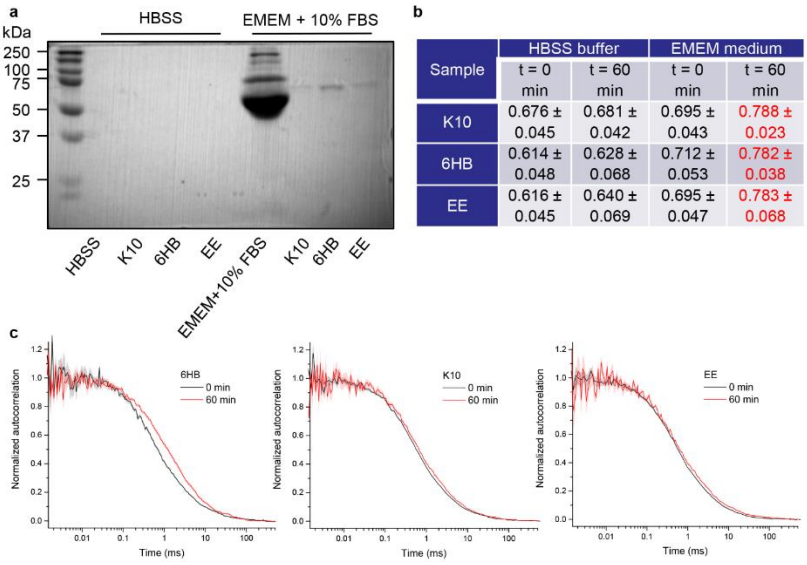


Figure 2.10. Figure 2.9. Determination of sequence specificity of EE-K10 coating. Each cell line was incubated with EE scramble coated sample with and without serum for 6 hrs. before fluorescence confocal microscopy was used to determine endosomal vesical overlap. Endosomes are stained in blue, 6HB are labeled green, and K10 coatings are red. Fluorescence overlap was analyzed using Pearson's correlations with (***) $P < 0.001$ significance. (A-C) Huh7 cells (D-F) HepG2 and (G-I) Alexander.

To verify the specificity of the EE-K10 coating, the scrambled variation was also monitored using fluorescence microscopy (Figure 2.10). Most cell-penetrating and endosomal escape peptides use a mechanism based on charge; however, because aurein 1.2 is neutral, it is hypothesized that the membrane disruption action is sequence specific.

Neither the serum free or serum containing cell studies with the EE scramble-K10 coating bundles showed an enhanced endosomal escape compared to the plain K10 coated or bare 6HB samples. These findings support the previously reports that describe the aurein 1.2 mechanism is highly sequence specific.[44] To further examine why the addition of the serum prevented endosomal escape, the bare and coated samples were incubated in the media and then analyzed using SDS-PAGE. After staining with Coomassie blue, a protein stain, it became apparent that protein corona had formed around the 6HB, K10-6HB, and EE-K10-6HB by the appearance of bands (Figure 2.11 A). This phenomenon had also previously been observed with other DNA nanostructures and nanoparticle delivery systems.[35, 58, 59] To further corroborate the protein corona formation, each of the samples was analyzed using fluorescence correlation spectroscopy (FCS) which allows for a determination of the particles size based on diffusion time.[60] The bare and coated structures that had been incubated in the serum-containing medium displayed a significantly longer diffusion time suggesting that they were larger than their corresponding structures that had been incubated in serum free media (Figure 2.11 B,C). It is likely that the protein corona is thicker than the EE-K10 coating molecule (since the aurein 1.2 portion is only ~3 nm when the backbone is fully stretched out) thereby blocking the peptide once the corona has formed.



Sample	HBSS buffer		EMEM medium	
	t = 0 min	t = 60 min	t = 0 min	t = 60 min
K10	0.676 ± 0.045	0.681 ± 0.042	0.695 ± 0.043	0.788 ± 0.023
6HB	0.614 ± 0.048	0.628 ± 0.068	0.712 ± 0.053	0.782 ± 0.038
EE	0.616 ± 0.045	0.640 ± 0.069	0.695 ± 0.047	0.783 ± 0.068

Figure 2.11. Analysis of protein corona formation. (A) SDS-page showing the formation bands indicating that proteins from the serum had formed a corona around the bare and coated 6HB. (B) Diffusion times for the FCS carried out on the samples had been incubated with serum and (C) their corresponding spectra.

2.3 Conclusion and Future Directions

In this chapter, a functionalized oligolysine coating for DNA nanostructures was explored to determine if they would enhance endosomal escape. The functionalized coating incorporated two copies of aurein 1.2 per molecule, allowing for an unprecedented high-density display of the peptide on the DNA structure itself. The coating molecule was found to optimally coat the 6HB bundle at a 1:1 nitrogen to phosphate ratio and was found to successfully facilitate cytosolic delivery in serum free conditions, but not in the presence of FBS due to the formation of a protein corona. Additionally, this work looked in depth at DNA nanostructure-hepatic cell interactions which had previously not been explored, demonstrating that the uptake efficiency directly correlates to cell size and morphology.

Further, it was shown that these K10 coatings may be a facile way to functionalize, while simultaneously stabilizing the DNA nanostructures for drug delivery purposes. It is likely that the approach could be extended to also incorporate targeting peptides to further increase cellular uptake and minimize off target effects. Before switching to liver cells, previous work focused on incorporating an anti-HER2 (human epidermal growth receptor 2) targeting peptide.[61] HER2 is often upregulated in aggressive forms of breast cancer and is often used as a model for target drug delivery.[62, 63] This peptide (P51) was synthesized to be on the N-terminus of the polylysine, and was intended to be incorporated at an optimized ratio with the EE-K10 coating to increase the specificity of cell delivery while still allowing for endosomal escape.

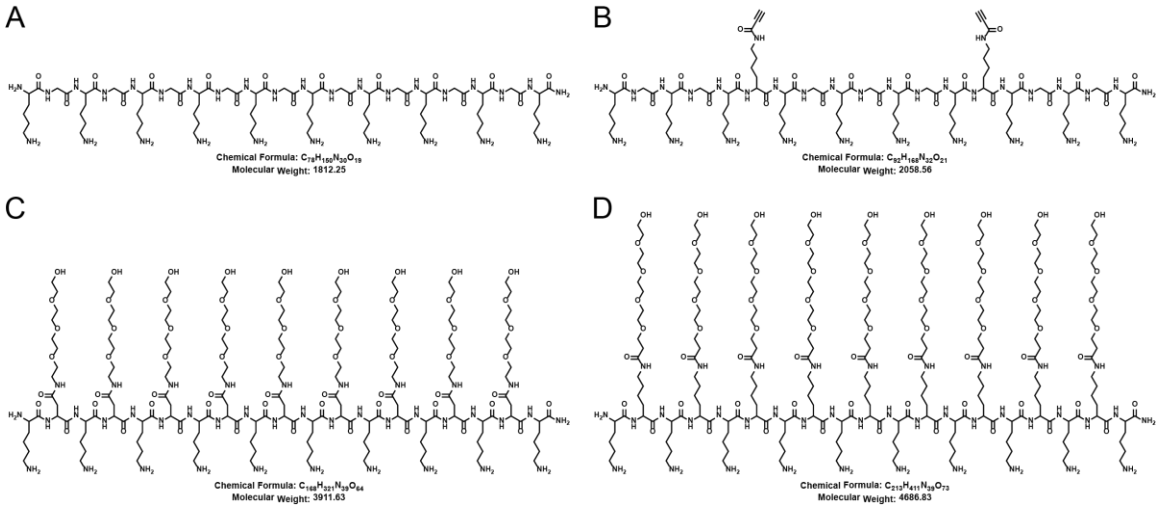


Figure 2.12. Alternative coating molecules. Additional oligolysine based molecules based on alternating lys-gly patterns (A). Alkyne modified K10 for attachment of targeting peptides using copper click (B). K10-PEG approach using a PEGylated amino acid (C) and K10-PEG after adding PEG onto selective lysines.

Additional methods looked at incorporating additional coating molecules for degradation resistance, crosslinking, and introduction of peptides after coating the DNA structures if they themselves are positively charged. These approaches were based off an alternating $(Lys\text{-}Gly)_{10}$ design which was later shown to cause aggregation because it could easily bridge two small 6HBs; however, it could possibly be more successful using larger origami structures (Figure 2.12A). To incorporate peptides after the fact, two of the glycine residues were switched to lysine with methyltrityl (MTT) groups which can be selectively deprotected before cleaving the peptide from resin. This reaction yields an extra reactive amine group which could then be coupled with propiolic acid to incorporate two alkyne groups (Figure 2.12B). Multiple variations with the additional cysteine at the N- and C- terminus could subsequently be used for crosslinking the coatings by using an oxidation agent. Furthermore, variations that had alternating lysine and non-canonical amino acids with PEG side chains were also explored. Two different approaches were taken, the first of which an Fmoc-lysine-OH amino acid was converted to an Fmoc- protected PEG_4 amino acid prior to synthesis, and the second again utilized lys(mtt) to incorporate a carboxylic acid functionalized-PEG_4 at selected sites (Figure 2.11C, D). While these PEGylated coating molecules were never incorporated into the *in vitro* experiments, a study with similar peptoid molecules was later published which verified that they successfully impede degradation.[64] Furthermore, work is being carried to incorporate non-fouling polymers to stop the formation of a protein corona.

2.4 References

1. Bastings, M. M. C.; Anastassacos, F. M.; Ponnuswamy, N.; Leifer, F. G.; Cuneo, G.; Lin, C. X.; Ingber, D. E.; Ryu, J. H.; Shih, W. M., Modulation of the Cellular Uptake of DNA Origami through Control over Mass and Shape. *Nano Letters* **2018,** *18* (6), 3557-3564.

2. Raniolo, S.; Croce, S.; Thomsen, R. P.; Okholm, A. H.; Unida, V.; Iacovelli, F.; Manetto, A.; Kjems, J.; Desideri, A.; Biocca, S., Cellular uptake of covalent and non-covalent DNA nanostructures with different sizes and geometries. *Nanoscale* **2019,** *11* (22), 10808-10818.

3. Mishra, S.; Feng, Y.; Endo, M.; Sugiyama, H., Advances in DNA Origami–Cell Interfaces. *ChemBioChem* **2019,** *21* (1-2), 33-44.

4. Stephanopoulos, N., Strategies for Stabilizing DNA Nanostructures to Biological Conditions. *ChemBioChem* **2019,** *20* (17), 2191-2197.

5. Shi, J. J.; Kantoff, P. W.; Wooster, R.; Farokhzad, O. C., Cancer nanomedicine: progress, challenges and opportunities. *Nature Reviews Cancer* **2017,** *17* (1), 20-37.

6. Ahrens, E. T.; Bulte, J. W. M., Tracking immune cells in vivo using magnetic resonance imaging. *Nature Reviews Immunology* **2013,** *13* (10), 755-763.

7. Lammers, T.; Ferrari, M., The success of nanomedicine. *Nano Today* **2020,** *31*.

8. Pelaz, B.; Alexiou, C. H.; Alvarez -Puebla, R. A.; Alves, F.; Andrews, A. M.; Ashraf, S.; Balogh, L. P.; Ballerini, L.; Bestetti, A.; Brendel, C.; Bosi, S.; Carril, M.; Chan, W. C. W.; Chen, C. Y.; Chen, X. D.; Chen, X. Y.; Cheng, Z.; Cui, D. X.; Du, J. Z.; Dullin, C.; Escudero, A.; Feliu, N.; Gao, M. Y.; George, M.; Gogotsi, Y.; Grunweller, A.; Gu, Z. W.; Halas, N. J.; Hampp, N.; Hartmann, R. K.; Hersam, M. C.; Hunziker, P.; Jian, J.; Jiang, X. Y.; Jungebluth, P.; Kadhiresan, P.; Kataoka, K.; Khademhosseini, A.; Kopecek, J.; Kotov, N. A.; Krug, H. F.; Lee, D. S.; Lehr, C. M.; Leong, K. W.; Liang, X. J.; Lim, M. L.; Liz-Marzan, L. M.; Ma, X. M.; Macchiarini, P.; Meng, H.; Mohwald, H.; Mulvaney, P.; Nel, A. E.; Nie, S. M.; Nordlander, P.; Okano, T.; Oliveira, J.; Park, T. H.; Penner, R. M.; Prato, M.; Puntes, V.; Rotello, V. M.; Samarakoon, A.; Schaak, R. E.; Shen, Y. Q.; Sjoqvist, S.; Skirtach, A. G.; Soliman, M. G.; Stevens, M. M.; Sung, H. W.; Tang, B. Z.; Tietze, R.; Udugama, B. N.; VanEpps, J. S.; Weil, T.; Weiss, P. S.; Willner, I.; Wu, Y. Z.; Yang, L. L.; Yue, Z.; Zhang, Q.; Zhang, X. E.; Zhao, Y. L.; Zhou, X.; Parak, W. J., Diverse Applications of Nanomedicine. *Acs Nano* **2017,** *11* (3), 2313-2381.

9. Abedini, A.; Bakar, A. A. A.; Larki, F.; Menon, P. S.; Islam, M. S.; Shaari, S., Recent Advances in Shape-Controlled Synthesis of Noble Metal Nanoparticles by Radiolysis Route. *Nanoscale Research Letters* **2016,** *11*, 1-13.

10. da Silva, A. G. M.; Rodrigues, T. S.; Slater, T. J. A.; Lewis, E. A.; Alves, R. S.; Fajardo, H. V.; Balzer, R.; da Silva, A. H. M.; de Freitas, I. C.; Oliveira, D. C.; Assaf, J. M.; Probst, L. F. D.; Haigh, S. J.; Camargo, P. H. C., Controlling Size, Morphology, and Surface Composition of AgAu Nanodendrites in 15 s for Improved Environmental Catalysis under Low Metal Loadings. *Acs Applied Materials & Interfaces* **2015,** *7* (46), 25624-25632.

11. Groeneveld, E.; de Mello Donegá, C., The Challenge of Colloidal Nanoparticle Synthesis. *Nanoparticles: Workhorses of Nanoscience* **2014**, 145-189.

12. Cheng, Y. H.; He, C. L.; Riviere, J. E.; Monteiro-Riviere, N. A.; Lin, Z. M., Meta-Analysis of Nanoparticle Delivery to Tumors Using a Physiologically Based Pharmacokinetic Modeling and Simulation Approach. *Acs Nano* **2020,** *14* (3), 3075-3095.

13. Wilhelm, S.; Tavares, A. J.; Dai, Q.; Ohta, S.; Audet, J.; Dvorak, H. F.; Chan, W. C. W., Analysis of nanoparticle delivery to tumours. *Nature Reviews Materials* **2016,** *1* (5).

14. Venditto, V. J.; Szoka, F. C., Cancer nanomedicines: So many papers and so few drugs! *Advanced Drug Delivery Reviews* **2013,** *65* (1), 80-88.

15. Park, K., The beginning of the end of the nanomedicine hype. *Journal of Controlled Release* **2019,** *305*, 221-222.

16. Gause, K. T.; Wheatley, A. K.; Cui, J. W.; Yan, Y.; Kent, S. J.; Caruso, F., Immunological Principles Guiding the Rational Design of Particles for Vaccine Delivery. *Acs Nano* **2017,** *11* (1), 54-68.

17. Anchordoquy, T. J.; Barenholz, Y.; Boraschi, D.; Chorny, M.; Decuzzi, P.; Dobrovolskaia, M. A.; Farhangrazi, Z. S.; Farrell, D.; Gabizon, A.; Ghandehari, H.; Godin, B.; La-Beck, N. M.; Ljubimova, J.; Moghimi, S. M.; Pagliaro, L.; Park, J.-H.; Peer, D.; Ruoslahti, E.; Serkova, N. J.; Simberg, D., Mechanisms and Barriers in Cancer Nanomedicine: Addressing Challenges, Looking for Solutions. *ACS nano* **2017,** *11* (1), 12-18.

18. Cooper, G. M., *The cell : a molecular approach*. 4th ed. ed.; ASM Press: Washington, D.C, 2007.

19. Poteryaev, D.; Datta, S.; Ackema, K.; Zerial, M.; Spang, A., Identification of the Switch in Early-to-Late Endosome Transition. *Cell (Cambridge)* **2010,** *141* (3), 497-508.

20. Endocytosis. *Nature reviews. Molecular cell biology* **2007,** *8* (6), 427.

21. Huotari, J.; Helenius, A., Endosome maturation. *Embo Journal* **2011,** *30* (17), 3481-3500.

22. Pei, D. H.; Buyanova, M., Overcoming Endosomal Entrapment in Drug Delivery. *Bioconjugate Chemistry* **2019,** *30* (2), 273-283.

23. Yang, X.; Fan, B.; Gao, W.; Li, L. P.; Li, T. T.; Sun, J. H.; Peng, X. Y.; Li, X. Y.; Wang, Z. J.; Wang, B. Q.; Zhang, R. P.; Xie, J., Enhanced endosomal escape by photothermal activation for improved small interfering RNA delivery and antitumor effect. *International Journal of Nanomedicine* **2018,** *13*, 4333-4344.

24. Ahmed, S.; Fujitab, S.; Matsumura, K., Enhanced protein internalization and efficient endosomal escape using polyampholyte-modified liposomes and freeze concentration. *Nanoscale* **2016,** *8* (35), 15888-15901.

25. Gao, Y. J.; Li, Y. F.; Li, Y. S.; Yuan, L.; Zhou, Y. X.; Li, J. W.; Zhao, L.; Zhang, C.; Li, X. R.; Liu, Y., PSMA-mediated endosome escape-accelerating polymeric micelles for targeted therapy of prostate cancer and the real time tracing of their intracellular trafficking. *Nanoscale* **2015,** *7* (2), 597-612.

26. Huang, G. J.; Chen, Q.; Wu, W. T.; Wang, J. W.; Chu, P. K.; Bai, H. Z.; Tang, G. P., Reconstructed chitosan with alkylamine for enhanced gene delivery by promoting endosomal escape. *Carbohydrate Polymers* **2020,** *227*.

27. Fröhlich, E., The role of surface charge in cellular uptake and cytotoxicity of medical nanoparticles. *International journal of nanomedicine* **2012,** *7*, 5577-5591.

28. Cristina, M.; Manuela, M.; Gheorghita, M.; Vasile, O., Understanding the Factors Influencing Chitosan-Based Nanoparticles-Protein Corona Interaction and Drug Delivery Applications. *Molecules (Basel, Switzerland)* **2020,** *25* (20), 4758.

29. Raouane, M.; Desmaële, D.; Urbinati, G.; Massaad-Massade, L.; Couvreur, P., Lipid Conjugated Oligonucleotides: A Useful Strategy for Delivery. *Bioconjugate chemistry* **2012,** *23* (6), 1091-1104.

30. Gissot, A.; Camplo, M.; Grinstaff, M. W.; Barthélémy, P., Nucleoside, nucleotide and oligonucleotide based amphiphiles: A successful marriage of nucleic acids with lipids. *Organic & biomolecular chemistry* **2008,** *6* (8), 1324-1333.

31. Shmushkovich, T.; Monopoli, K. R.; Homsy, D.; Leyfer, D.; Betancur-Boissel, M.; Khvorova, A.; Wolfson, A. D., Functional features defining the efficacy of cholesterol-conjugated, self-deliverable, chemically modified siRNAs. *Nucleic acids research* **2018,** *46* (20), 10905-10916.

32. Jeong, J. H.; Kim, S. W.; Park, T. G., Novel Intracellular Delivery System of Antisense Oligonucleotide by Self-Assembled Hybrid Micelles Composed of DNA/PEG Conjugate and Cationic Fusogenic Peptide. *Bioconjugate chemistry* **2003,** *14* (2), 473-479.

33. Kundu, A.; Nandi, S.; Nandi, A. K., Nucleic acid based polymer and nanoparticle conjugates: Synthesis, properties and applications. *Progress in materials science* **2017,** *88*, 136-185.

34. Nielsen, P. E., Peptide nucleic acid: a versatile tool in genetic diagnostics and molecular biology. *Current Opinion in Biotechnology* **2001,** *12* (1), 16-20.

35. Whitehouse, W. L.; Noble, J. E.; Ryadnov, M. G.; Howorka, S., Cholesterol Anchors Enable Efficient Binding and Intracellular Uptake of DNA Nanostructures. *Bioconjugate Chemistry* **2019,** *30* (7), 1836-1844.

36. Ponnuswamy, N.; Bastings, M. M. C.; Nathwani, B.; Ryu, J. H.; Chou, L. Y. T.; Vinther, M.; Li, W. A.; Anastassacos, F. M.; Mooney, D. J.; Shih, W. M., Oligolysine-based coating protects DNA nanostructures from low-salt denaturation and nuclease degradation. *Nature Communications* **2017,** *8*.

37. Anastassacos, F. M.; Zhao, Z.; Zeng, Y.; Shih, W. M., Glutaraldehyde Cross-Linking of Oligolysines Coating DNA Origami Greatly Reduces Susceptibility to Nuclease Degradation. *Journal of the American Chemical Society* **2020,** *142* (7), 3311-3315.

38. Dervan, P. B., A Personal Perspective on Chemical Biology: Before the Beginning. *Israel journal of chemistry* **2019,** *59* (1-2), 71-83.

39. Alniss, H. Y., Thermodynamics of DNA Minor Groove Binders. *Journal of medicinal chemistry* **2019,** *62* (2), 385-402.

40. Pazos, E.; Mosquera, J.; Vazquez, E.; Mascarenas, J., DNA Recognition by Synthetic Constructs. *ChemBioChem* **2011,** *12*, 1958-1973.

41. Roodhuizen, J. A. L.; Hendrikx, P.; Hilbers, P. A. J.; de Greef, T. F. A.; Markvoort, A. J., Counterion-Dependent Mechanisms of DNA Origami Nanostructure Stabilization Revealed by Atomistic Molecular Simulation. *Acs Nano* **2019,** *13* (9), 10798-10809.

42. Neundorf, I.; Rennert, R.; Hoyer, J.; Schramm, F.; Löbner, K.; Kitanovic, I.; Wölfl, S., Fusion of a short HA2-derived peptide sequence to cell-penetrating peptides improves cytosolic uptake, but enhances cytotoxic activity. *Pharmaceuticals (Basel, Switzerland)* **2009,** *2* (2), 49-65.

43. Sugita, T.; Yoshikawa, T.; Mukai, Y.; Yamanada, N.; Imai, S.; Nagano, K.; Yoshida, Y.; Shibata, H.; Yoshioka, Y.; Nakagawa, S.; Kamada, H.; Tsunoda, S. i.; Tsutsumi, Y., Comparative study on transduction and toxicity of protein transduction domains. *British journal of pharmacology* **2008,** *153* (6), 1143-1152.

44. Li, M.; Tao, Y.; Shu, Y. L.; LaRochelle, J. R.; Steinauer, A.; Thompson, D.; Schepartz, A.; Chen, Z. Y.; Liu, D. R., Discovery and Characterization of a Peptide That

Enhances Endosomal Escape of Delivered Proteins in Vitro and in Vivo. *Journal of the American Chemical Society* **2015,** *137* (44), 14084-14093.

45. Amblard, M.; Fehrentz, J.-A.; Martinez, J.; Subra, G., Methods and protocols of modern solid phase peptide synthesis. *Molecular biotechnology* **2006,** *33* (3), 239-254.

46. Burns, J. R.; Seifert, A.; Fertig, N.; Howorka, S., A biomimetic DNA-based channel for the ligand-controlled transport of charged molecular cargo across a biological membrane. *Nature Nanotechnology* **2016,** *11* (2), 152-156.

47. Zhang, Y. N.; Poon, W.; Tavares, A. J.; McGilvray, I. D.; Chan, W. C. W., Nanoparticle-liver interactions: Cellular uptake and hepatobiliary elimination. *Journal of Controlled Release* **2016,** *240*, 332-348.

48. Frtús, A.; Smolková, B.; Uzhytchak, M.; Lunova, M.; Jirsa, M.; Kubinová, Š.; Dejneka, A.; Lunov, O., Analyzing the mechanisms of iron oxide nanoparticles interactions with cells: A road from failure to success in clinical applications. *Journal of controlled release* **2020,** *328*, 59-77.

49. Tsoi, K. M.; MacParland, S. A.; Ma, X. Z.; Spetzler, V. N.; Echeverri, J.; Ouyang, B.; Fadel, S. M.; Sykes, E. A.; Goldaracena, N.; Kaths, J. M.; Conneely, J. B.; Alman, B. A.; Selzner, M.; Ostrowski, M. A.; Adeyi, O. A.; Zilman, A.; McGilvray, I. D.; Chan, W. C. W., Mechanism of hard-nanomaterial clearance by the liver. *Nature Materials* **2016,** *15* (11), 1212-1221.

50. Lunova, M.; Prokhorov, A.; Jirsa, M.; Hof, M.; Olzynska, A.; Jurkiewicz, P.; Kubinova, S.; Lunov, O.; Dejneka, A., Nanoparticle core stability and surface functionalization drive the mTOR signaling pathway in hepatocellular cell lines. *Scientific Reports* **2017,** *7*.

51. Lunova, M.; Zablotskii, V.; Dempsey, N. M.; Devillers, T.; Jirsa, M.; Sykova, E.; Kubinova, S.; Lunov, O.; Dejneka, A., Modulation of collective cell behaviour by geometrical constraints. *Integrative Biology* **2016,** *8* (11), 1099-1110.

52. Bade, N. D.; Kamien, R. D.; Assoian, R. K.; Stebe, K. J., Curvature and Rho activation differentially control the alignment of cells and stress fibers. *Science Advances* **2017,** *3* (9).

53. Amodeo, A. A.; Skotheim, J. M., Cell-size control. *Cold Spring Harbor perspectives in biology* **2016,** *8* (4), a019083-a019083.

54. Khetan, J.; Shahinuzzaman, M.; Barua, S.; Barua, D., Quantitative Analysis of the Correlation between Cell Size and Cellular Uptake of Particles. *Biophysical Journal* **2019,** *116* (2), 347-359.

55. Ni, R.; Feng, R.; Chau, Y., Synthetic Approaches for Nucleic Acid Delivery: Choosing the Right Carriers. *Life (Basel, Switzerland)* **2019,** *9* (3), 59.

56. Auvinen, H.; Zhang, H.; Nonappa; Kopilow, A.; Niemelä, E.; Nummelin, S.; Correia, A.; Santos, H.; Linko, V.; Kostiainen , M., Protein Coating of DNA Nanostructures for Enhanced Stability and Immunocompatibility. *advanced healthcare materials* **2017,** *6*.

57. Juliano, R. L., Intracellular Trafficking and Endosomal Release of Oligonucleotides: What We Know and What We Don't. *Nucleic acid therapeutics* **2018,** *28* (3), 166-177.

58. Del Pino, P.; Pelaz, B.; Zhang, Q.; Maffre, P.; Nienhaus, G. U.; Parak, W. J., Protein corona formation around nanoparticles - from the past to the future. *Materials Horizons* **2014,** *1* (3), 301-313.

59. Ke, P. C.; Lin, S.; Parak, W. J.; Davis, T. P.; Caruso, F., A Decade of the Protein Corona. *Acs Nano* **2017,** *11* (12), 11773-11776.

60. Shang, L.; Nienhaus, G. U., In Situ Characterization of Protein Adsorption onto Nanoparticles by Fluorescence Correlation Spectroscopy. *Accounts of Chemical Research* **2017,** *50* (2), 387-395.

61. Geng, L.; Wang, Z.; Jia, X.; Han, Q.; Xiang, Z.; Li, D.; Yang, X.; Zhang, D.; Bu, X.; Wang, W.; Hu, Z.; Fang, Q., HER2 targeting peptides screening and applications in tumor imaging and drug delivery. *Theranostics* **2016,** *6* (8), 1261-1273.

62. Neve, R.; Chin, K.; Fridlyand, J.; Yeh, J.; Gazdar, A.; Gray, J., A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **2006,** *10(6)*, 515-27.

63. Subik, K.; Lee, J.-F.; Baxter, L.; Strzepek, T.; Costello, D.; Crowley, P.; Xing, L.; Hung, M.-C.; Bonfiglio, T.; Hicks, D. G.; Tang, P., The expression patterns of ER, PR, HER2, CK5/6, EGFR, KI-67 and AR by immunohistochemical analysis in breast cancer cell lines. *Breast cancer : basic and clinical research* **2010,** *4* (1), 35-41.

64. Wang, S.-T.; Gray, M. A.; Xuan, S.; Lin, Y.; Byrnes, J.; Nguyen, A. I.; Todorova, N.; Stevens, M. M.; Bertozzi, C. R.; Zuckermann, R. N.; Gang, O., DNA origami protection and molecular interfacing through engineered sequence-defined peptoids. *Proceedings of the National Academy of Sciences - PNAS* **2020,** *117* (12), 6339-6348.

CHAPTER 3

DNA-PEPTIDE CONJUGATES TO CREATE SYNBODIES, DYNAMIC NANOSTRUCTURES, AND PROTEIN MIMETICS

3.1 Introduction to click chemistry

In order to control the orientation and location of the peptides with respect to a DNA nanostructure, they must be covalently linked to a unique DNA sequence, or "handle", that is integrated directly into the overall structure. There are three main approaches to producing peptide-oligonucleotide conjugates (POCs): total on-resin synthesis, utilizing a bifunctional linker, or incorporating non-canonical amino acids and phosphoramidites with functional groups amenable to conjugation.[1-5] The latter two routes often employ various "click" reactions which were first identified by Sharpless in 2001.[6] This class of reactions are characterized as being high yield under mild conditions (aqueous solution, room temperature, etc.) leading to a desired bond with specificity while producing little to no byproducts. All click reactions are thermodynamically driven, and consist of a mixture of nucleophilic substitutions, cycloadditions, non-aldol carbonyl chemistry, and Michael additions.[7, 8] One common click reaction used to link oligonucleotides and peptides exploits the thiol group of cysteine which reacts with a maleimide moiety (Figure 3.1A). There are countless commercially available heterobifunctional linkers with an NHS (n-hydroxysuccinimidyl) ester to react with an amine, some form of linker arm, and a maleimide group.[9] However, this reaction can only be utilized with specificity if only one reactive cysteine occurs naturally, or a single cysteine can be site-specifically introduced.

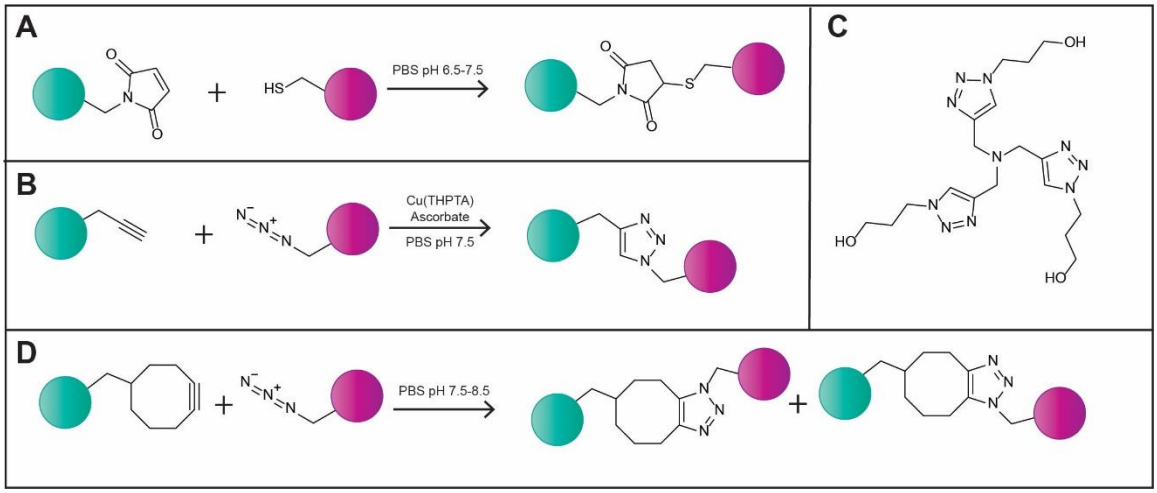


Figure 3.1. Common click reactions for bioconjugation. (A) A reaction between a maleimide and a thiol and (B) the Cu(I) catalyzed click along with a (C) common ligand, THPTA. (D) SPaaC starting material and the two isomers that are produced.

The copper free and copper (Cu(I)) catalyzed clicks both occur between an azide and an alkyne, which undergo a [3+2] Huisgen cycloaddition. The Cu(I) click uses a linear alkyne, which does not react spontaneously with the azide, thus requiring a Cu(I) catalyst (Figure 3.1B). The copper must be in its 1st oxidation state, and is often complexed with some form of triazole ligand such as THPTA (tris-hydroxypropyl-triazolylmethylamine) at a 1:5 ratio to minimize any cytotoxic effects the reaction itself can have (Figure 3.1C).[10] Ascorbate, a common reducing agent, is traditionally added in 100-fold excess to the CuTHPTA solution to ensure it is in the proper oxidation state. One advantage to using the Cu(I) click is that it is regiospecific, resulting in only the 1,4-isomer, whereas the copper free click produces both the 1,4- and 1,5- isomers.[11-13] Unlike the Cu(I) click, the copper free click uses some form of cyclooctyne (Figure 3.1D), which allows for a spontaneous strain promoted azide-alkyne cycloaddition (SPaaC), but can be more challenging to synthesize.[14]

3.2 Peptide Oligonucleotide Conjugates (POCs) for Tribodies

3.2.1 POC Tribody Motivation and Background

The ability to bind a protein of interest (POI) with high affinity and specificity is highly desirable because it can be used for a variety of applications including cell imaging, ELISAs, Western blots, and protein immobilization.[15] One way that biological systems achieve this protein binding is in the form of an IgG (immunoglobin G) antibody which contains three short complementarity determining regions (CDRs) peptide loops in its antigen- binding site.[16, 17] Unfortunately, antibodies have a complex structure with multiple disulfide bond and glycosylation patterns, making them difficult to produce or mutate, so as a result they are extremely costly to produce. Recently, there has been an effort to make synthetic antibodies or "synbodies" which use some form of scaffold to position multiple binding peptides at discrete distances.[18-20]

One example of these synbodies was comprised of peptide oligonucleotide conjugates (POCs), where a simple DNA duplex was utilized to position two Gal80 binding peptides at varying distances, with spacing being determined by total number of base pairs to obtain a synbody with ~1000-fold higher binding affinity than the individual peptides alone (5.6 nM for the synbody, vs. ~3.5 μM for the individual peptides).[21] This improved affinity comes from a synergistic multivalent effect between the two peptides, much like the effect seen with the three CDR loops in an antibody. The ability to control the shape and size of a DNA nanostructure make them an attractive scaffold to position multiple binding peptides simultaneously at precise distances to create antibody mimetics.

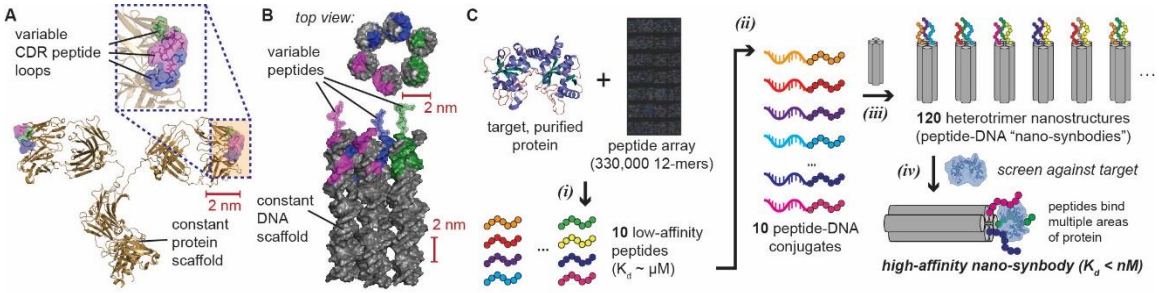


Figure 3.2. Tribody overview. (A) Crystal structure of an IgG antibody with the three CDR peptide loops highlighted alongside (B) the starting nanostructure of a six-helix bundle displaying three binding peptides. (C) Schematic showing a representative target protein and peptide array that is used to determining several mediocre binding peptides for the POI that will be turned into peptide-DNA conjugates, three of which will be screen in concert with one another for higher binding affinities.

This section will explore utilizing POCs, to create a "tribody", to orient three unique binding peptides on a DNA nanostructure—once again, a six helix bundle (6HB), to mimic the CDR loops of a traditional IgG (Figure 3.2A, B). The possible binding peptides were identified using a peptide array and the POI, each individually displaying low micromolar affinity. After synthesizing POCs with the identified peptides, each with their own unique DNA handle, different combinations of the peptides were incorporated into the 6HB to determine if any of the tribodies displayed nanomolar binding affinities (Figure 3.2C).

3.2.2 Choosing a Target and Identifying Peptides

Transferrin was chosen as a model starting target because it is stable, well-studied, and commercially available. Transferrin is a glycoprotein that binds and transports iron, and is found to be present at increased levels in a variety of diseases, such as anemia and prostate cancer.[22-24] Potential peptide targets were identified by using a fluorescently-labeled transferrin to visualize which spots on a peptide array bound the protein. The array that was employed, screened 330,000 different peptides each

containing 12 residues (Figure 3.3A). Additionally, *E. coli* lysate, which contains over 800 unique proteins, was also introduced to the same peptide array to ensure that the peptides were binding specifically to transferrin and not non-specifically to any off-target proteins in the lysate (Figure 3.3B). Following the identification of the potential peptide binders, their binding affinity was determined using surface plasmon resonance (SPR) with the POI immobilized on the surface while introducing each binding peptide individually (Figure 3.3C).

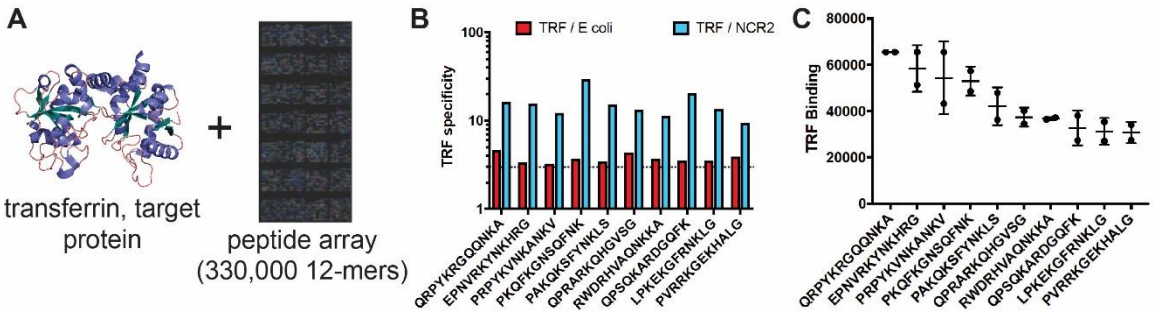


Figure 3.3. Transferrin as a target. (A) Structure of the transferrin protein with the random loops shown in pink, α-helices shown in purple, and β-sheet shown in teal and an image of the peptide array chip containing 330,000 different sequences. (B) Bar graph showing the specificity of the chosen peptide sequences by looking at the binding factors to transferrin (blue) versus an *E. coli* lysate in red and (C) the binding units for the same 10 peptides as determined by SPR.

From the peptide array screening, the 10 peptides that displayed the best binding affinities were selected. Each of these peptides was then synthesized with a GSG linker, a commonly employed sequence because it is small, flexible, and non-reactive, and an azidolysine (AzK) that could be utilized in a copper free click reaction. Following synthesis, the peptides were purified using RP-HPLC and characterized using MALDI-MS to ensure the synthesis was successful and the peptide was pure (Figure 3.4A, B). Additionally, each of the synthetic peptides was tested again using SPR to ensure that they bound to transferrin as efficiently with the GSG-AzK modification (Figure 3.3C, D).

The SPR data showed that every peptide had binding affinities between 1-9 µM, with the exception of TRF104 which was between 10-19 µM.

A

Peptide	Expected	Observed
101	1793.93	1792.69
102	1709.99	1705.83
103	1718.91	1715.1
104	1706.94	1706.66
105	1616.87	1613.9
106	1805.01	1804.43
108	1682.96	1681.55
109	1643.89	1640.02

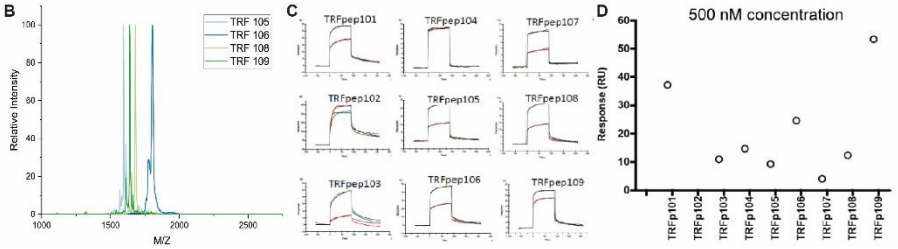
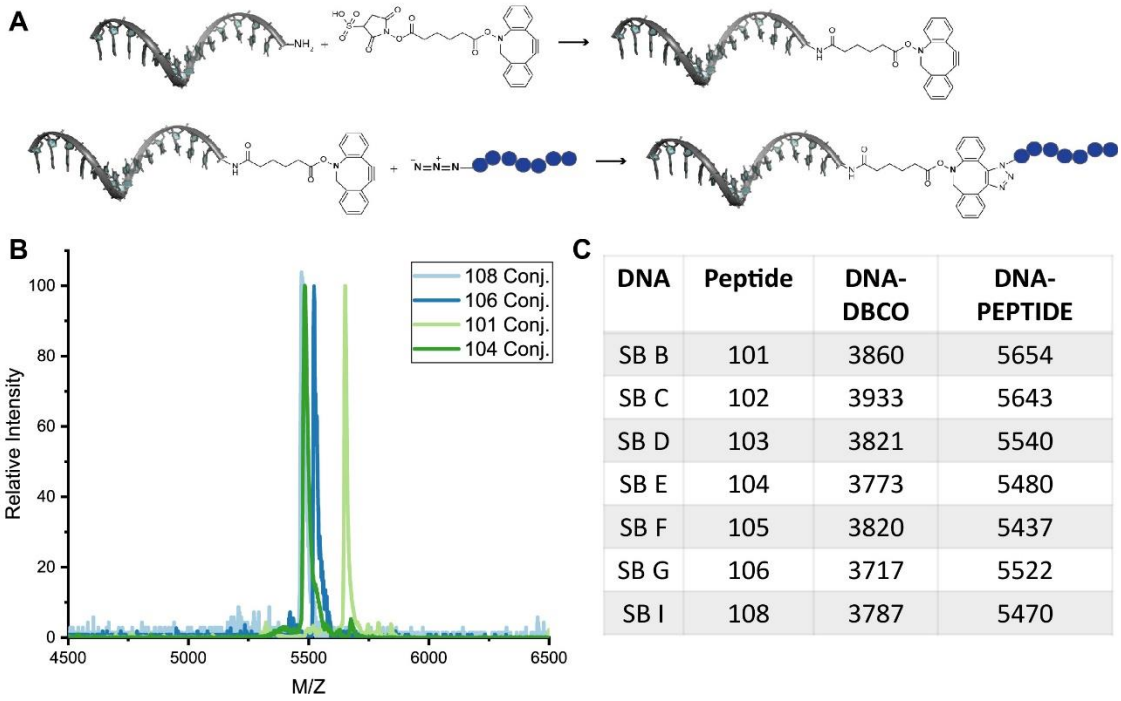


Figure 3.4. Synthesis and characterization of transferrin peptides. (A) Expected vs. observed masses for the synthesized peptides and (B) a few representative MALDI-TOF-MS spectra. (C) Binding curves of the peptides obtained using SPR used to determine binding affinities along with a (D) graph showing the response of binding during a 500 nM screen of the synthetic peptides containing an azidolysine.

3.2.3 Synthesis of POCs

Once each of the transferrin binding peptides had been prepared, the next step was to conjugate each of them to its own unique DNA handle. For this, each DNA strand had to be modified with a moiety that would undergo a SPaaC with the non-canonical azidolysine in each peptide (Figure 3.5A). The DNA handles contained an amine modification that could react with a dibenzylcyclooctyne (DBCO) NHS ester. This reaction was carried out in phosphate buffered saline (PBS) pH 7.5 buffer with the DBCO added at a 100-fold excess at room temperature. The DBCO-modified DNA was separated from the unmodified DNA using RP-HPLC using a gradient of 10-80% methanol (with 50 mM triethylammonium acetate as the aqueous buffer) over 45 minutes. The wavelengths of 260 nm (DNA absorbance) and 309 nm (DBCO absorbance) were observed in tandem and used to determine that the second peak did indeed correspond to the DBCO-DNA. The peak fractions were collected, and buffer exchanged back into 10mM PBS using molecular weight cut off spin filtration and

concentrating it to a final volume of 200 µL, at which point the corresponding peptide was added at a 5x molar excess and allowed to react overnight at room temperature.



DNA	Peptide	DNA-DBCO	DNA-PEPTIDE
SB B	101	3860	5654
SB C	102	3933	5643
SB D	103	3821	5540
SB E	104	3773	5480
SB F	105	3820	5437
SB G	106	3717	5522
SB I	108	3787	5470

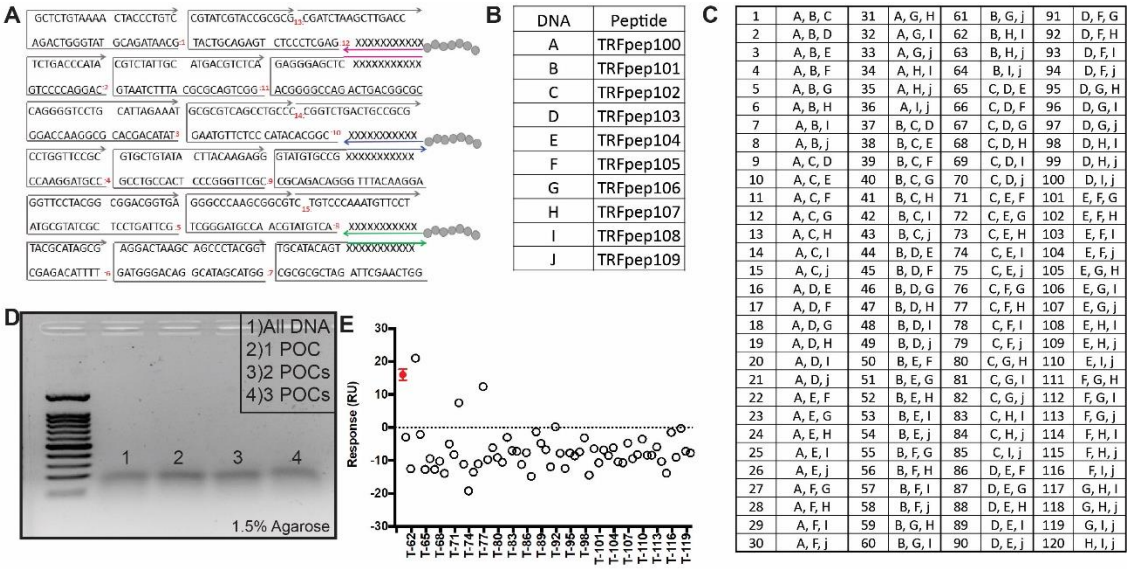
Figure 3.5. Schematic of click and characterization of POCs. (A)Full scheme showing the modification of an amine modified DNA strand with DBCO and its reaction with an azide containing peptide to create a POC. (B) MALDI-TOF mass spectrum for a select few conjugates showing the good correlation with the (C) expected masses.

While the yield of the DNA-peptide conjugates was near 100%, each sample was purified a second time using the same RP-HPLC method to ensure that any residual DNA or DNA-DBCO was removed. Each of the POCs was then buffer exchanged back into water and characterized using MALDI-MS (Figure 3.5B, C). The concentration of the purified POCs was determined using UV-Vis spectroscopy, by determining the absorbance at 260 nm and calculating extinction coefficients based on the contribution of

the DNA sequence itself and any partial absorbance that may be contributed by the aromatic amino acids present in some of the sequences.

3.2.4 Assembly Characterization of POC Synbody Variants

Each of the POCs was designed to have a unique DNA handle, to serve as a marker specific to each peptide that was being displayed in each tribody. As a starting point, a six-helix bundle (6HB) was designed using a total of 21 oligonucleotides. This was a modified version of a 6HB originally reported by the Reif laboratory,[25] which had a diameter of ~5.5 nm, and could display the three peptides in a symmetric fashion (with respect to rotational symmetry), simplifying the process of making all combinations of three peptides. The structure was comprised of 21 total strands; 15 of those strands (shown in gray) with sequences that were unperturbed from the original design, to act as the scaffold that would display the three different peptides, while three of the remaining strands contained the complementary sequence to the desired POC handle to act as an anchoring point, with the remaining three reserved for the POCs (Figure 3.6A, B).



DNA	Peptide
A	TRFpep100
B	TRFpep101
C	TRFpep102
D	TRFpep103
E	TRFpep104
F	TRFpep105
G	TRFpep106
H	TRFpep107
I	TRFpep108
J	TRFpep109

1	A, B, C	31	A, G, H	61	B, G, j	91	D, F, G
2	A, B, D	32	A, G, I	62	B, H, I	92	D, F, H
3	A, B, E	33	A, G, j	63	B, H, j	93	D, F, I
4	A, B, F	34	A, H, I	64	B, I, j	94	D, F, j
5	A, B, G	35	A, H, j	65	C, D, E	95	D, G, H
6	A, B, H	36	A, I, j	66	C, D, F	96	D, G, I
7	A, B, I	37	B, C, D	67	C, D, G	97	D, G, j
8	A, B, j	38	B, C, E	68	C, D, H	98	D, H, I
9	A, C, D	39	B, C, F	69	C, D, I	99	D, H, j
10	A, C, E	40	B, C, G	70	C, D, j	100	D, I, j
11	A, C, F	41	B, C, H	71	C, E, F	101	E, F, G
12	A, C, G	42	B, C, I	72	C, E, G	102	E, F, H
13	A, C, H	43	B, C, j	73	C, E, H	103	E, F, I
14	A, C, I	44	B, D, E	74	C, E, I	104	E, F, j
15	A, C, j	45	B, D, F	75	C, E, j	105	E, G, H
16	A, D, E	46	B, D, G	76	C, F, G	106	E, G, I
17	A, D, F	47	B, D, H	77	C, F, H	107	E, G, j
18	A, D, G	48	B, D, I	78	C, F, I	108	E, H, I
19	A, D, H	49	B, D, j	79	C, F, j	109	E, H, j
20	A, D, I	50	B, E, F	80	C, G, H	110	E, I, j
21	A, D, j	51	B, E, G	81	C, G, I	111	F, G, H
22	A, E, F	52	B, E, H	82	C, G, j	112	F, G, I
23	A, E, G	53	B, E, I	83	C, H, I	113	F, G, j
24	A, E, H	54	B, E, j	84	C, H, j	114	F, H, I
25	A, E, I	55	B, F, G	85	C, I, j	115	F, H, j
26	A, E, j	56	B, F, H	86	D, E, F	116	F, I, j
27	A, F, G	57	B, F, I	87	D, E, G	117	G, H, I
28	A, F, H	58	B, F, j	88	D, E, H	118	G, H, j
29	A, F, I	59	B, G, H	89	D, E, I	119	G, I, j
30	A, F, j	60	B, G, I	90	D, E, j	120	H, I, j

Figure 3.6. Tribody design and characterization. (A) Body of the six-helix bundle consists of the same 12 strands (gray) along with 3 POCs and their three complements to anchor them into the nanostructure. (B) A list of the DNA strand and its corresponding peptide to make a unique POC (C) and the 120 different combinations when choosing 3 POCs. (D) Agarose gel showing the successful assemble of the bundle make out of all DNA, 1,2, and 3 POCs and the (E) SPR data from some of the tribodies alongside a positive control of peptide104 shown in red.

By displaying three different peptides simultaneously, with a total of 10 different options, there are 120 different possible combinations (Figure 3.6C). To produce the 120 different tribodies, an 800 nM master mix containing the 15 strands comprising the DNA scaffold was prepared in annealing buffer (1x TAE Mg). The master mix was then split into 120 tubes and the 6 other strands (3 complementary anchors and 3 POCs) were added to each corresponding tube, and water was added to each tribody to a final concentration of 100 nM in 100 µL. To form the tribodies from the 120 mixtures were annealed from 95 to 12°C over 10 hours. The assembly of the tribodies were visualized using a 1.5% agarose gel and analyzed (25 nM samples) using SPR with an immobilized transferrin for binding (Figure 3.6D, E). Unfortunately, very few tribodies exhibited

transferrin binding at 25 nM and the select few that did, only remained bound for a short time, indicating it was not highly specific.

3.2.5 POC Future Direction

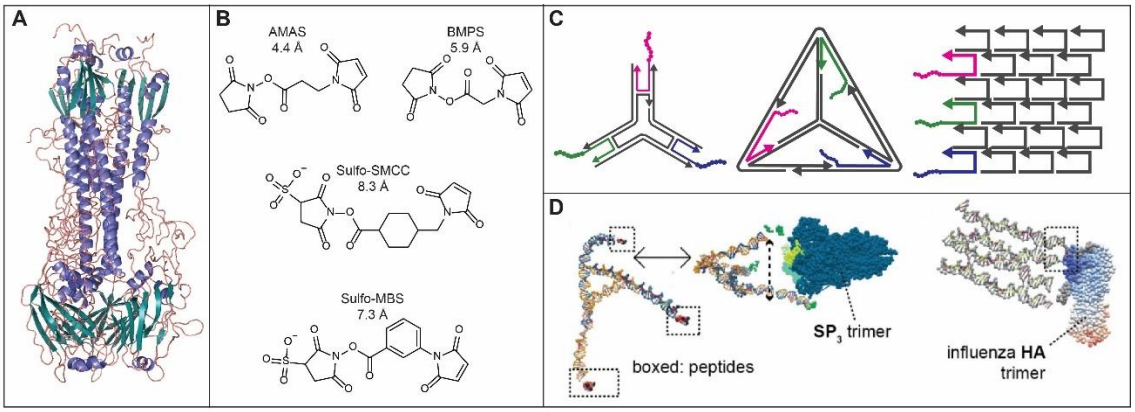


Figure 3.7. Future directions of tribodies. (A) Structure of Influenza A with the random loops shown in pink, α-helices shown in purple, and β-sheet shown in teal. (B) Four different linkers that can be used to link amine modified DNA to cysteine containing peptides with different linker lengths. (C) Representative 2D topologies demonstrating that the same POCs can be positioned at different distances by changing the nanostructure (D) and preliminary coarse grained molecular dynamic simulations on influenza A to determine optimal linker length and DNA nanostructure.

The lack of binding at 25 nM could be attributed to several factors: improper linker length, interpeptide distance, or non-cooperative binding between the three peptides. Additionally, at this point several new POIs were identified as HER2 (human epithelial growth factor receptor 2) which is upregulated in breast cancer tissue and hemagglutinin (HA) from influenza A (Figure 3.7A). Peptide array experiments were performed against both targets, identifying 13 anti-HER2 peptides and 5 for HA. In order to explore linker length as a variable, each of the 18 identified peptides were synthesized with the addition of a GSG-C linker, allowing for a thiol-maleimide based click reaction (Figure 3.1A). A wide variety of NHS-ester maleimide bifunctional linkers were

commercially available with different linker arm lengths and degrees of freedom, so each POC could be altered for a specific binding distance readily (Figure 3.7B). Simultaneously, the DNA handles were designed with the flexibility to be incorporated into a variety of nanostructures to allow for variable distances between the peptides to allow them to be tested without synthesizing the POCs themselves (Figure 3.7). In addition, a collaborative effort sought to optimize DNA nanostructure and linker length *in situ* (Figure 3.7D).

3.3 Synthesis and Application of DNA-Peptide-DNA conjugates

3.3.1 DPD Motivation and Background

To date, there have been few reported DNA-peptide conjugates containing a singular DNA handle; however, the addition of a second DNA handle on a peptide, making a DNA-peptide-DNA triblock (DPD), could offer some distinct advantages. One such advantage would be to use the two DNA handles to position and constrain the peptide into a cyclic conformation, much like the antigen binding peptide loops seen in proteins. Another application would allow for the peptide to act as a functional linker between different DNA structures, to lock a DNA nanostructure lid in a dynamic fashion, or within a DNA based hydrogel. One example of a biologically functional peptide 'linker" is GPQGIWGQ, a peptide sequence naturally found in type I collagen, that can be cleaved enzymatically. In the presence of matrix metalloproteinase (MMP) enzymes, the peptide sequence cleaves between the glycine and isoleucine.[26-29] While the sequence is compatible to most MMPs, its highest activity occurs when in the presence MMP-8.[30]

In order to conjugate two different DNA handles at a specific terminus of the peptide, two orthogonal reactions are needed. There are very few reported DPD

molecules in the literature and approaches are challenging synthetically and/or expensive. One such example used a linear solid phase synthesis to generate a DPD; however, this approach had significant draw backs because it limits amino acids that can be present in the peptide due to inherent differences between phosphoramidites and amino acid chemistries.[31] Alternatively, several of these triblock molecules have been synthesized using peptide nucleic acid (PNA) instead of DNA, which can be costly.[3] Recently, the Hayashi lab has also reported using both a copper free and Cu(I) click simultaneously, similar to the approach that was employed in this work.[32]

3.3.2 Proximity Enhanced Click Reactions

For the two orthogonal conjugation reactions, an MMP-cleavable peptide (MMP_{pep}) was synthesized with an azidolysine on the N-terminus for a copper free click and a propargyl alanine (PrA) on the C-terminus for the Cu(I) click reaction (Figure 3.8A, B). The two DNA handles were both functionalized using NHS ester-amine chemistry with one bearing a DBCO and the other a PEG-azide (Figure 3.8C, D). Following purification using RP-HPLC, the DNA-DBCO was mixed with the peptide at a 1:5 molar ratio to create a DNA-peptide conjugate (Figure 3.8 E, F). Following the protocol reported by the Hayashi lab, a Cu(I) click was carried out in solution to attach the second DNA handle bearing the azide but very low yields of the full DPD were achieved, and the molecule was difficult to purify.

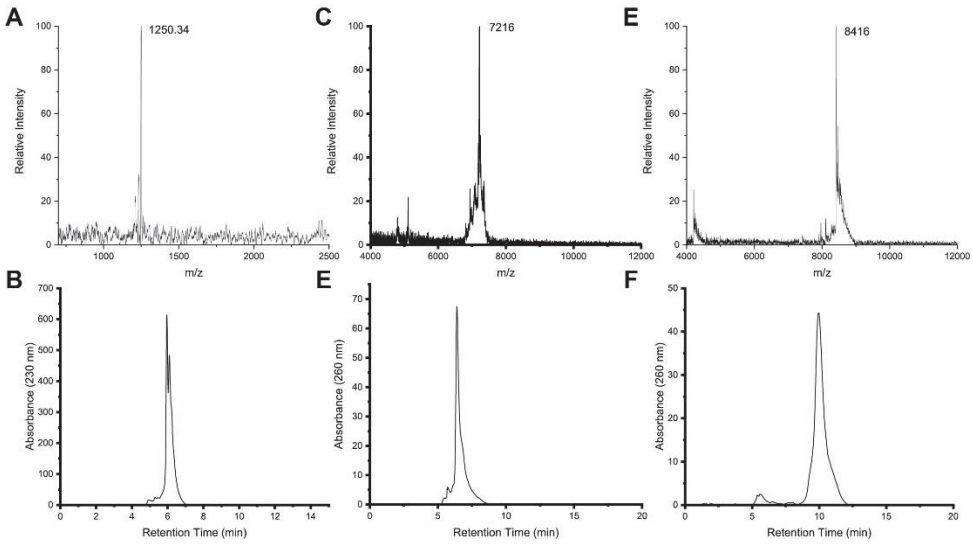


Figure 3.8. Characterization of MMP cleavable peptide and DNA handles. MALDI-TOF MS spectra and analytical RP-HPLC chromatograms confirming the successful synthesis and purify of (A, B) the MMP cleavable peptide with an expected mass of 1217.36 (C,D) the DNA handle with the azide with an expected mass of 7191 (E,F) and the DNA-DBCO-MMP peptide conjugate with an expected mass of 8407.

One way to increase the yield of the second click reaction is to exploit a proximity- aided effect. For this, the two DNA handles were redesigned to be partially complementary. The complementary region was optimized to be 9 nucleotides long so it could form a stable duplex at room temperature, while also being easy to denature upon adding a fully complementary strand (Figure 3.9B). The DNA-MMP_{pep} conjugate and DNA-az were both added to a buffer solution (20 mM PBS pH 7.5) at a concentration of 5 µM and annealed from 95-20°C over 30 minutes to form the duplex.

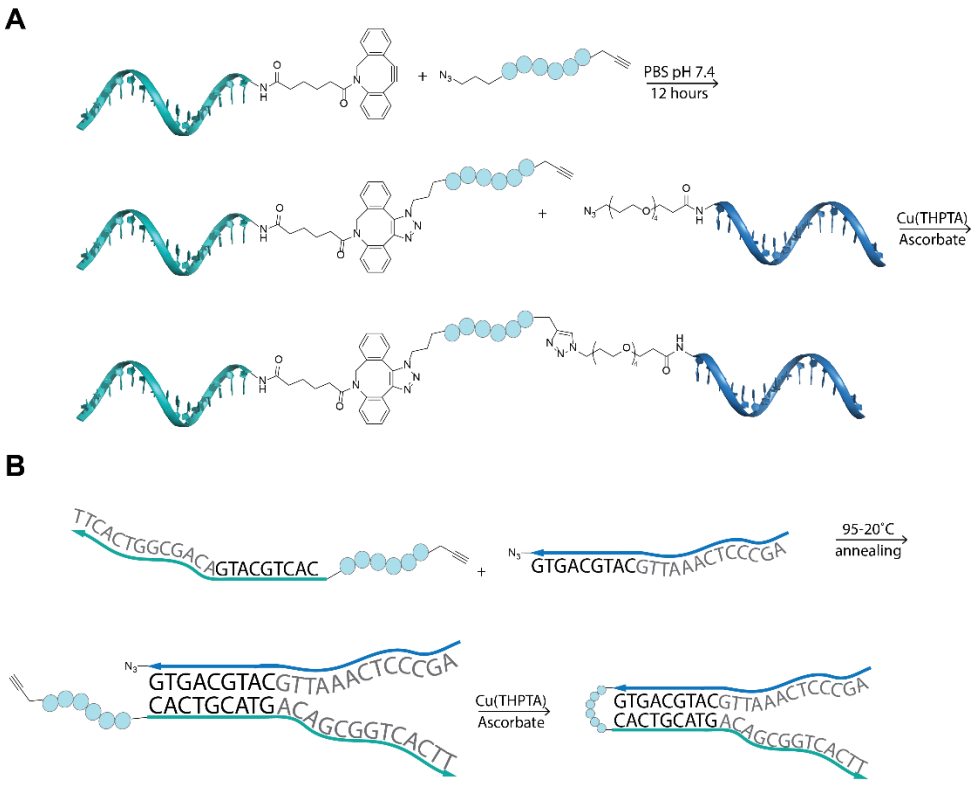


Figure 3.9. Schematic of proximity enhanced synthesis of DPD. Full schematic of the addition of the first handle using SPaaC followed by the addition of the second handle using CuAAC (A) and a schematic to highlight that the two DNA handles were partially complementary, forming a duplex, to increase yield of the CuAAC (B).

With the free alkyne on the peptide now proximal to the azide on the second DNA handle, a Cu(I) click reaction was carried out. Briefly, this was done by adding a pre-incubated solution of Cu:THPTA (1:5) (0.4 µM) with the duplex at 4 µM before adding a freshly prepared reducing solution containing sodium ascorbate to total 40 mM. The ascorbate reduces Cu(II) to Cu(I) which can catalyze the azide-alkyne cycloaddition. After 4 hours, the Cu(I) catalyst was chelated out of solution by adding an excess (250 mM) of ethylenediaminetetraacetic acid (EDTA) and carrying out a buffer exchange using molecular weight cut off spin filtration. The formation of the DPD was characterized using gel electrophoresis on a urea based denaturing polyacrylamide gel

(Figure 3.10). The DNA-az and DNA- MMP_{pep} conjugate are in lanes 1 and 2, with a control that lacked Cu(I), and lacked the proximity aided effect shown in lanes 3 and 4, and the unpurified DPD in lane 5 showing a high yield conversion from POC to DPD.

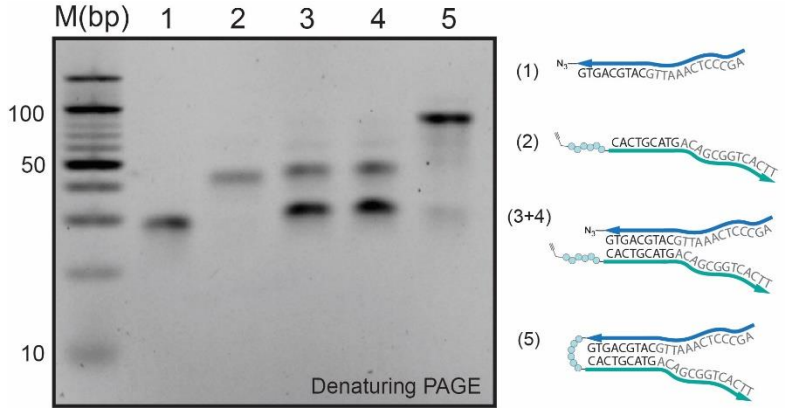


Figure 3.10. DNA-peptide-DNA characterization. 6% denaturing PAGE with a corresponding key on the side to highlight the high yield of DPD with the proximity aided effect.

3.3.3 DPDs for Dynamic Nanostructures

With the overall goal to create dynamic nanostructures that could be linked by an enzymatically cleavable peptide (e.g., a DNA "robot" that opens its lid in the presence of the enzyme), the next step was to incorporate each of the DNA handles into separate nanostructures. As model DNA nanostructures, two simple double crossover (DX) tiles were designed to contain unique sequences, with each one the DPD handles. The optimal working temperature of the matrix metalloproteinase (MMP) is 37°C, therefore the two tiles were checked for stability after being incubated at 37°C for 24 hours (Figure 3.11A). Since no denaturation was observed during incubation, the DPD was mixed into a solution containing the eight other strands that that comprise each tile at an equal stoichiometric ratio (1 μM), and then annealed. The formation of the two DX tiles linked by the peptide in the DPD was charactered using native PAGE (Figure 3.11B).

According to the gel, it was apparent that a larger structure (lane 3) formed due to the clear band shift compared to the individual tiles (lanes 1 and 2). The MMP cleavable peptide linked tiles were then incubated with MMP-8 at 37°C for 48 hours, resulting in two individual tiles each containing half of the peptide linker (lane 4).

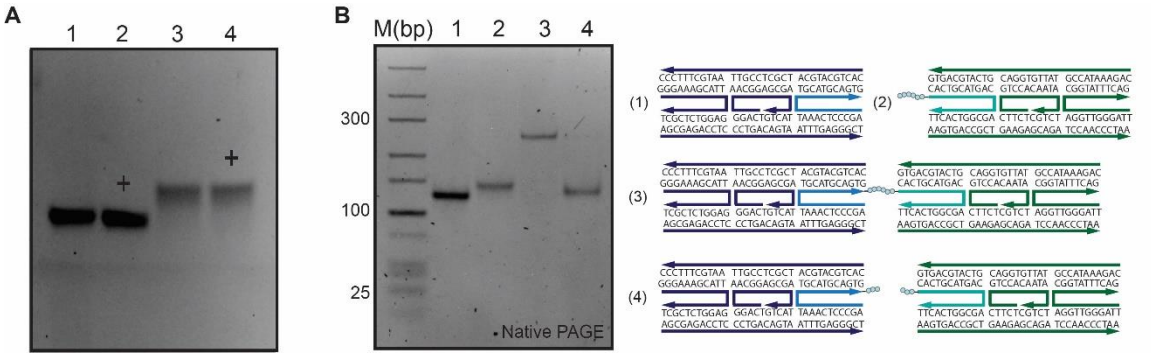


Figure 3.11. Characterization of MMP_{pep} linked DX tiles and the enzymatic cleavage. (A) Thermal incubation of the two DX tiles at 37°C to ensure that they were stable at the enzymes working temperature with the + indicating the dx tile after incubation. (B) 5% native PAGE showing the formation of the tile with the azide, tile with the DNA-peptide conjugate, two tiles linked by the peptide portion of the DPD, and the two tiles with half the peptide after cleavage.

To demonstrate that the enzymatic cleavage was a sequence specific phenomenon, a "scrambled" version (same amino acid composition in a different order) of the peptide (AzK-GQGIPQGWGG-PrA, MMP_{peps}) was synthesized, and the DNA-peptide conjugate was synthesized using a copper free click reaction as previously described (Figure 3.12A). The DNA- MMP_{peps} conjugate was then made into a DPD and incorporated into the same DX tiles. The original and scrambled DX- MMP_{pep}-DX structures were then both incubated with the MMP-8 but cleavage was only observed in the original peptide sample (lane 4 vs. lane 6, Figure 3.12B)

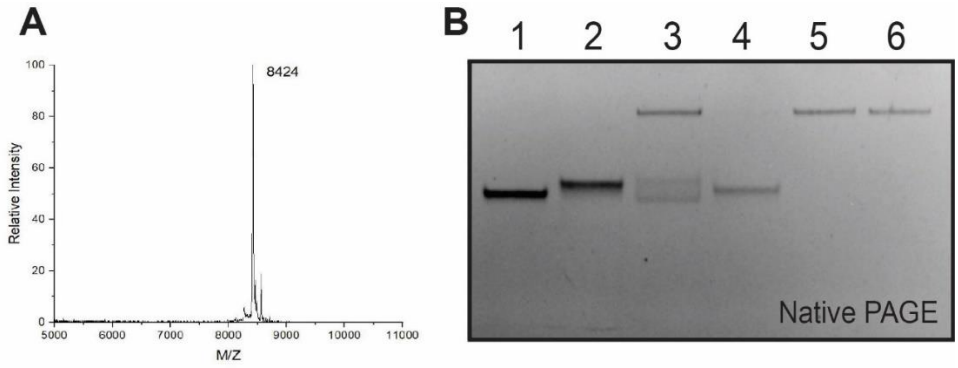


Figure 3.12. Scrambled MMP DPD. (A) Characterization of the DNA-scrambled MMP cleavable peptide using MALDI-MS with an expected mass of 8407 and (B) a native PAGE showing that the enzymatic cleavage is specific to the peptide sequence with the lanes containing;1)DX-azide, 2)DX-peptide, 3)DX-MMP_{pep}-DX (unpurified), 4) lane 3 after MMP-8 incubation, 5-DX- MMP_{pep}s-DX, and 6) MMP_{pep}s after MMP-8 incubation.

3.4 Synthesis and Application of DNA-Peptide Co-block Polymers

3.4.1 Fibronectin and Cyclic Peptides

Another application of a DPD is to use the two DNA handles to induce a conformational restriction on the peptide to mimic an antigen binding loop that naturally occurs in proteins. It is well known that cyclizing a peptide often leads to increased binding affinity and stability when compared to a linear peptide.[33-35] One example of this increased binding affinity that has been widely studied is the cell-adhesion peptide RGD, a short peptide loop found in fibronectin,[36-38] where the cyclization of the sequence lead to a 10-fold increase in binding affinity.[39-41]

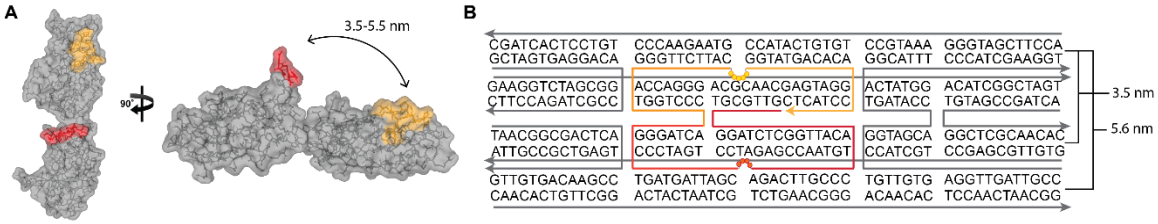


Figure 3.13. RGD and PHSRN loops in fibronectin. (A) Crystal structure of fibronectin with the RGD loop shown in red and PHSRN loop in yellow alongside (B) a DNA nanostructure that could be used to position the two loops at discrete distances.

For cell adhesion, the RGD loop works synergistically with another peptide loop (PHSRN) that is located approximately 3.2-5.5 nm apart in fibronectin depending on its conformation (Figure 3.13A).[42-44] With the structural predictability of DNA nanostructures, it is possible to design a nanostructure that displays these two peptide loops using DPDs at defined distances to mimic the native fibronectin positioning (Figure 3.13B) in future work.[45] To synthesize two different DPDs, an RGDS peptide and a PHSRN peptide were both synthesized with the addition of an PrA-G- on the N-terminus and a -G-AzK on the C-terminus for the copper free and Cu(I) clicks to add the two DNA handles (Figure 3.14).

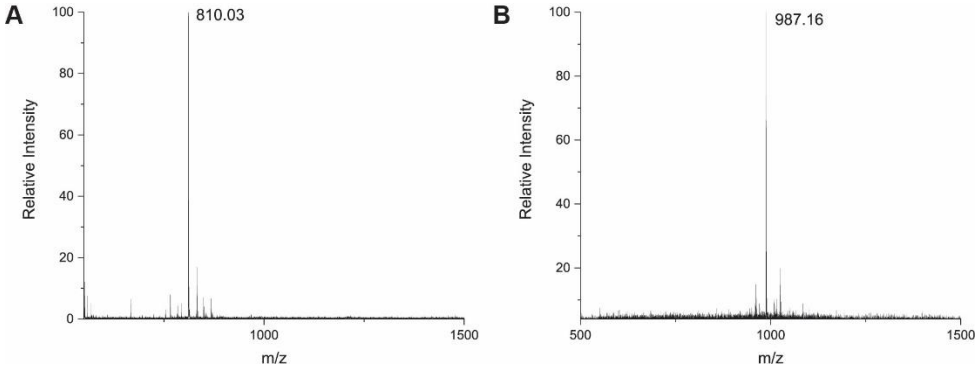


Figure 3.14. Characterization of RGDS and PHSRN peptides. (A) MALDI-MS spectra for the RGDS peptide and (B) PHSRN peptides used to create the two DPDS with expected masses of 809.84 and 987.07, respectively.

3.4.2 Synthetic Approach

It is possible to further mimic a protein by making a co-block polymer with alternating DNA-peptide sections that could intramolecularly fold using the DNA position to display the peptide loops at designed locations. This can be done by first synthesizing two DPDs and adding a 5' phosphorylation to one handle and using a ligase

to seal the nick. Instead of adding complementarity to the DNA handles themselves, a single complementary strand was used to position each of the 4 handles, positioning each of the reactive groups in close enough proximity to one another, allowing for only one annealing step (Figure 3.15).

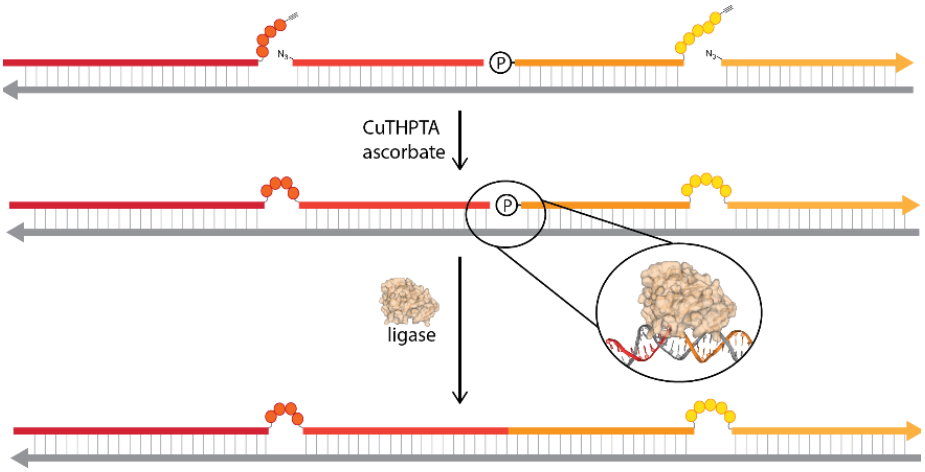


Figure 3.15. DNA peptide coblock polymer synthesis. The four DNA handles, two functionalized with a peptide (RGDS in red, PHSRN in yellow) and two functionalized with an azide and their singular complementary strand (gray) used to add proximity for the two Cu(I) clicks and the ligation. The P denotes 5'-phosphorylation

Each of the four individual parts were first synthesized separately to produce a DNA-RGDS conjugate, an az-DNA handle, a PO_4-DNA-PHSRN conjugate, and a second az-DNA handle. They were all characterized using MALDI-MS to ensure each DNA segment was properly synthesized before carrying out the Cu(I) click and ligation steps (Figure 3.16A-D). A stock solution (8 µM) of the 4 DNA handles and their complementary strand was prepared and annealed from 95-12°C over 2.5 hours. The full duplex then underwent a Cu(I) click by adding CuTHPTA at 0.1 mM, ascorbate at 10 mM, and aminoguanidine (1 mM) to help protect the arginine from the byproducts of ascorbate oxidation.[46] The Cu(I) click was run for 4 hours at room temperature before

chelating the catalyst with EDTA (250 mM) and buffer exchanging to water.

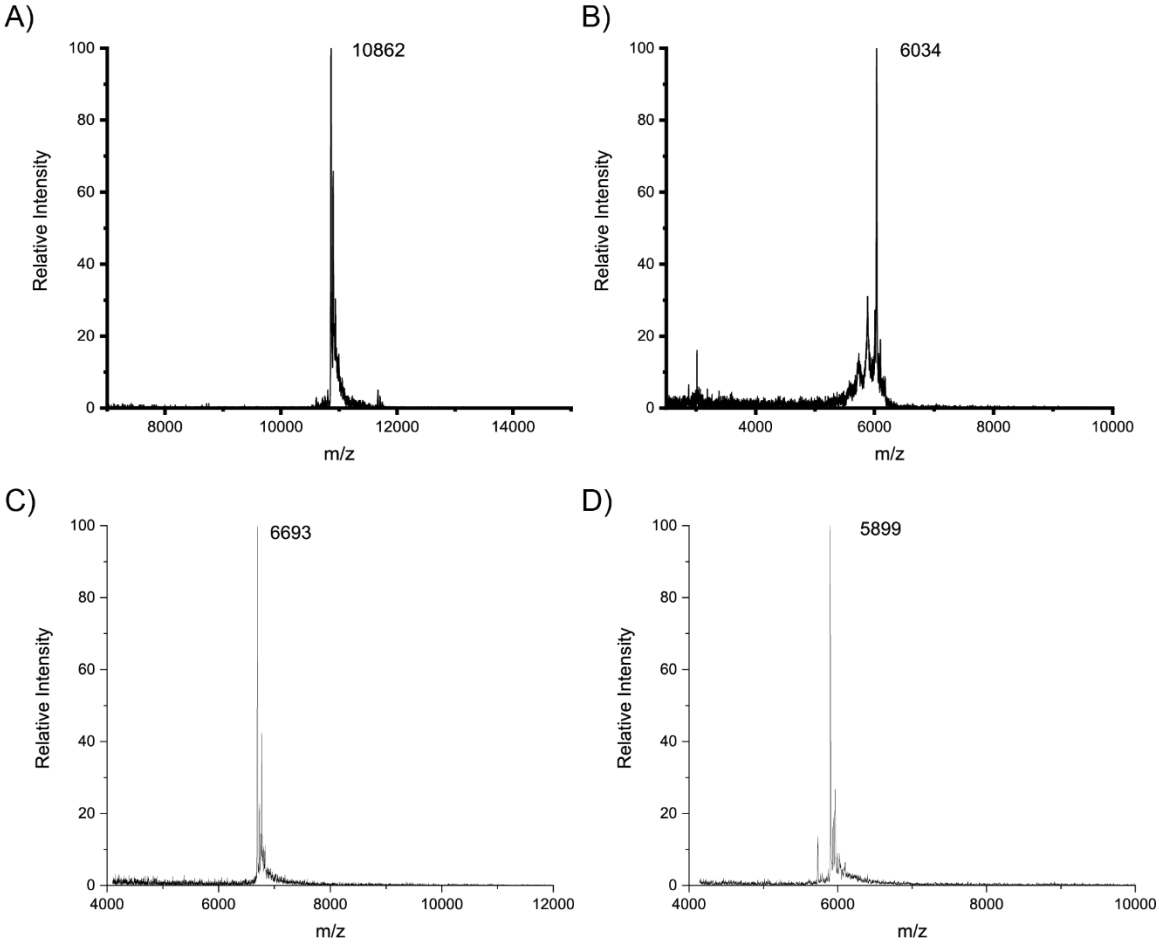


Figure 3.16. Segments used to make a DNA peptide copolymer. MALDI-MS spectra for the DNA-RGDS with an expected mass of 10855 (A) and its corresponding DNA-az with an expected mass of 6032 (B), and the DNA-PHSRN with an expected mass of 6760 (C) and its corresponding DNA-az with an expected mass of 5898 (D).

3.4.3 Characterization of the Block Copolymer

Following the Cu(I) click and buffer exchange back into water, the duplex which was composed of two DPDs and their complementary strand, was added to ligation buffer (5 mM Tris-HCl, 1 mM $MgCl_2$, 0.1 mM ATP, 1 mM DTT) along with 200 units of T4 ligase. The ligation was carried out at 16°C overnight, after which it was deactivated by incubating at 65°C for 10 minutes. The formation of the full DPDPD pentablock

copolymer was characterized using denaturing PAGE (Figure 3.17). Lanes 1-4 contain each of the individual DNA handles alone (DNA-RGDS, DNA-az, DNA-PHSRN, DNA-az), while lane 5 contains the full complementary strand. Lane 6 contains the duplex with no added Cu(I) or ligase, lane 7 shows the duplex after the click only, and lane 8 shows the full DPD (red box) after both the Cu(I) click and the ligation.

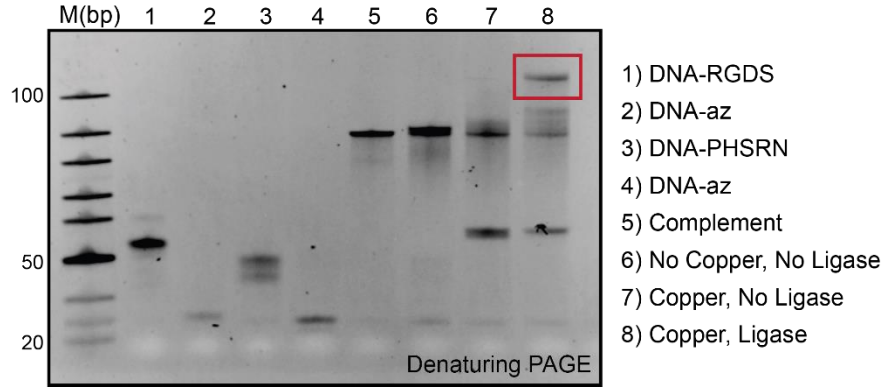


Figure 3.17. DNA-peptide copolymer characterization. A 6% denaturing page with the individual DNA handles in lanes 1-4, the complementary strand in lane 5, and the full duplex after each step, annealing (lane 6), CuAAC (lane 7), and ligation (lane 8).

Unexpectedly, the full product proved difficult to characterize due to a lack of absorbance at 260 nm with the intermittent peptide segments, a phenomenon whose origin is still mysterious. Additionally, each of the DNA handles were too short to stain individually, making characterization difficult. Without the ability to quantify each product individually, each reaction needed to be carried out sequentially without purification using a single complementary strand to ensure the proper stoichiometry. The final product can be isolated through gel extraction after UV shadowing to visualize each of the induvial bands but will require further modification, such as the addition of a fluorophore onto a handle in order to be able to determine its concentration prior to incorporating it into a larger DNA nanostructure.

3.5 Conclusion

The data reported in this chapter shows several potential applications of DNA-peptide conjugates with both one and two DNA handles. A single DNA handle results in a linear peptide dangling from the DNA nanostructure while two handles allow for the conformational control of the peptide to mimic a binding loop. The singular handle conjugates were employed to create synbodies for transferrin, by using a multivalent effect by positioning three unique binding peptides off the same 6HB to try and obtain higher binding affinity than each of the individual peptides alone. However, upon screening the 120 different tribodies, each containing three transferrin binding peptides displayed off the 6HB that were screened did not bind to transferrin at the concentration tested (25 nM). This result could be due to several reasons: (1) the peptides could all bind at the same location on the protein, disallowing for a multivalent effect; (2) the linker length between the DNA and peptide could be too long or short, allowing for too many or few little degrees of freedom for the peptides to bind efficiently; or (3) the DNA nanostructure could be positioning the three peptides at the wrong spacing and geometry.

DNA-peptide-DNA conjugates have proven difficult to synthesize because two orthogonal conjugation chemistries are required. Both a Cu(I) click, and a copper free click were utilized to attach two unique handles at distinct positions. This work showed that carrying out the second conjugation reaction with a proximity-aided effect, drastically increases the yield of the second handle addition. When using an enzymatically cleavable peptide as a linker between the two DNA handles, two different DX tiles were tethered together using the DPD, and subsequently separated by incubation with the cleaving enzyme (MMP-8) to produce two individual tiles, each bearing half the

peptide. These DPDs with cleavable linkers could potentially act as a way to close a DNA lid, and specifically open when introducing the specific enzymes. Furthermore, it was shown these DPD can be ligated together to form DNA-peptide block copolymers. This work utilized both an RGDS loop and PHSRN loop that can in future studies, be incorporated into a DNA nanostructure to position these two loops at a distance similar to their position within the native fibronectin protein.

3.6 References

1. MacCulloch, T.; Buchberger, A.; Stephanopoulos, N., Emerging applications of peptide-oligonucleotide conjugates: bioactive scaffolds, self-assembling systems, and hybrid nanomaterials. *Organic & Biomolecular Chemistry* **2019,** *17* (7), 1668-1682.

2. Venkatesan, N.; Kim, B. H., Peptide conjugates of oligonucleotides: Synthesis and applications. *Chemical Reviews* **2006,** *106* (9), 3712-3761.

3. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O., SEQUENCE-SELECTIVE RECOGNITION OF DNA BY STRAND DISPLACEMENT WITH A THYMINE-SUBSTITUTED POLYAMIDE. *Science* **1991,** *254* (5037), 1497-1500.

4. Lonnberg, H., Solid-Phase Synthesis of Oligonucleotide Conjugates Useful for Delivery and Targeting of Potential Nucleic Acid Therapeutics. *Bioconjugate Chemistry* **2009,** *20* (6), 1065-1094.

5. Lu, K.; Duan, Q.-P.; Ma, L.; Zhao, D.-X., Chemical Strategies for the Synthesis of Peptide−Oligonucleotide Conjugates. *Bioconjugate chemistry* **2010,** *21* (2), 187-202.

6. Kolb, H. C.; Finn, M. G.; Sharpless, K. B., Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angewandte Chemie International Edition* **2001,** *40* (11), 2004-2021.

7. Service, R. F., Chemistry: Click chemistry clicks along. *Science (American Association for the Advancement of Science)* **2008,** *320* (5878), 868-869.

8. Hein, C. D.; Liu, X.-M.; Wang, D., Click Chemistry, A Powerful Tool for Pharmaceutical Sciences. *Pharmaceutical research* **2008,** *25* (10), 2216-2230.

9. Northrop, B. H.; Frayne, S. H.; Choudhary, U., Thiol-maleimide "click" chemistry: Evaluating the influence of solvent, initiator, and thiol on the reaction mechanism, kinetics, and selectivity. *Polymer chemistry* **2015,** *6* (18), 3415-3430.

10. Besanceney-Webler, C.; Jiang, H.; Zheng, T.; Feng, L.; Soriano del Amo, D.; Wang, W.; Klivansky, L.; Marlow, F.; Liu, Y.; Wu, P., Increasing the Efficacy of Bioorthogonal Click Reactions for Bioconjugation: A Comparative Study. *Angewandte Chemie International Edition* **2011,** *50* (35), 8051-8056.

11. Liang, L.; Astruc, D., The copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) "click" reaction and its applications. An overview. *Coordination chemistry reviews* **2011,** *255* (23), 2933-2945.

12. Presolski, S. I., Ligand-Accelerated Copper(I)-Catalyzed Azide-Alkyne Cycloaddition: Kinetic Studies and Reaction Mechanism. ProQuest Dissertations Publishing: 2011.

13. Presolski, S.; Hong, V. P.; Finn, M. G., Copper-Catalyzed Azide–Alkyne Click

Chemistry for Bioconjugation. *Current Protocols Chemical Biology* **2011,** *3*, 153-162.

14. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R., A Strain-Promoted [3 + 2] Azide−Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *Journal of the American Chemical Society* **2004,** *126* (46), 15046-15047.

15. Sapra, P.; Shor, B., Monoclonal antibody-based therapies in cancer: Advances and challenges. *Pharmacology & therapeutics (Oxford)* **2013,** *138* (3), 452-469.

16. Chiu, M. L.; Goulet, D. R.; Teplyakov, A.; Gilliland, G. L., Antibody Structure and Function: The Basis for Engineering Therapeutics. *Antibodies (Basel)* **2019,** *8* (4), 55.

17. Strohl, W. R., Current progress in innovative engineered antibodies. *Protein & cell* **2018,** *9* (1), 86-120.

18. Miersch, S.; Sidhu, S. S., Synthetic antibodies: Concepts, potential and practical considerations. *Methods (San Diego, Calif.)* **2012,** *57* (4), 486-498.

19. Deyev, S. M.; Lebedenko, E. N., Modern Technologies for Creating Synthetic Antibodies for Clinical Application. *Actanaturae* **2009,** *1* (1), 32-50.

20. Owens, B., Faster, deeper, smaller-the rise of antibody-like scaffolds (vol 35, pg 602, 2017). *Nature Biotechnology* **2017,** *35* (11), 1004-1004.

21. Williams, B. A. R.; Diehnelt, C. W.; Belcher, P.; Greving, M.; Woodbury, N. W.; Johnston, S. A.; Chaput, J. C., Creating Protein Affinity Reagents by Combining Peptide Ligands on Synthetic DNA Scaffolds. *Journal of the American Chemical Society* **2009,** *131* (47), 17233-17241.

22. Miller, J. L., Iron deficiency anemia: A common and curable disease. *Cold Spring Harbor perspectives in biology* **2013,** *5* (7), a011866-a011866.

23. Ulbrich, K.; Hekmatara, T.; Herbert, E.; Kreuter, J., Transferrin- and transferrin-receptor-antibody-modified nanoparticles enable drug delivery across the blood–brain barrier (BBB). *European journal of pharmaceutics and biopharmaceutics* **2009,** *71* (2), 251-256.

24. Scognamiglio, I.; Di Martino, M. T.; Campani, V.; Virgilio, A.; Galeone, A.; Gullà, A.; Gallo Cantafio, M. E.; Misso, G.; Tagliaferri, P.; Tassone, P.; Caraglia, M.; De Rosa, G., Transferrin-Conjugated SNALPs Encapsulating 2′-O-Methylated miR-34a for the Treatment of Multiple Myeloma. *BioMed research international* **2014,** *2014*, 217365-7.

25. Peng, Y.; Rizal, F. H.; Sudheer, S.; Harry, M. T. C.; Sung Ha, P.; Thomas, H. L.; John, H. R., Programming DNA Tube Circumferences. *Science (American Association for the Advancement of Science)* **2008,** *321* (5890), 824-826.

26. Lutolf, M. P.; Hubbell, J. A., Synthetic biomaterials as instructive extracellular microenvironments for morphogenesis in tissue engineering. *Nature Biotechnology* **2005,** *23* (1), 47-55.

27. Chau, Y.; Luo, Y.; Cheung, A. C. Y.; Nagai, Y.; Zhang, S. G.; Kobler, J. B.; Zeitels, S. M.; Langer, R., Incorporation of a matrix metalloproteinase-sensitive substrate into self-assembling peptides - A model for biofunctional scaffolds. *Biomaterials* **2008,** *29* (11), 1713-1719.

28. Chau, Y.; Tan, F. E.; Langer, R., Synthesis and characterization of dextran-peptide-methotrexate conjugates for tumor targeting via mediation by matrix metalloproteinase II and matrix metalloproteinase IX. *Bioconjugate Chemistry* **2004,** *15* (4), 931-941.

29. Aimes, R. T.; Quigley, J. P., MATRIX METALLOPROTEINASE-2 IS AN INTERSTITIAL COLLAGENASE - INHIBITOR-FREE ENZYME CATALYZES THE CLEAVAGE OF COLLAGEN FIBRILS AND SOLUBLE NATIVE TYPE-I COLLAGEN GENERATING THE SPECIFIC 3/4-LENGTH AND 1/4-LENGTH FRAGMENTS. *Journal of Biological Chemistry* **1995,** *270* (11), 5872-5876.

30. Patterson, J.; Hubbell, J. A., Enhanced proteolytic degradation of molecularly engineered PEG hydrogels in response to MMP-1 and MMP-2. *Biomaterials* **2010,** *31* (30), 7836-7845.

31. Bergmann, F.; Bannwarth, W., Solid phase synthesis of directly linked peptide-oligodeoxynucleotide hybrids using standard synthesis protocols. *Tetrahedron letters* **1995,** *36* (11), 1839-1842.

32. Hayashi, G.; Yanase, M.; Nakatsuka, Y.; Okamoto, A., Simultaneous and Traceless Ligation of Peptide Fragments on DNA Scaffold. *Biomacromolecules* **2019,** *20* (3), 1246-1253.

33. Horton, D. A.; Bourne, G. T.; Smythe, M. L., Exploring privileged structures: The combinatorial synthesis of cyclic peptides. *Molecular diversity* **2000,** *5* (4), 289-304.

34. Joo, S. H., Cyclic peptides as therapeutic agents and biochemical tools. *Biomolecules & therapeutics* **2012,** *20* (1), 19-26.

35. Lee, A. C.-L.; Harris, J. L.; Khanna, K. K.; Hong, J.-H., A comprehensive review on current advances in peptide drug development and design. *International journal of molecular sciences* **2019,** *20* (10), 2383.

36. Liskamp, R.; Rijkers, D.; Bakker, S., *Bioactive Macrocyclic Peptides and Peptide Mimics*. Wiley-VCH: 2008.

37. Finke, A.; Busskamp, H.; Manea, M.; Marx, A., Designer Extracellular Matrix Based on DNA-Peptide Networks Generated by Polymerase Chain Reaction. *Angewandte Chemie-International Edition* **2016,** *55* (34), 10136-10140.

38. Campbell, I. D.; Humphries, M. J., Integrin structure, activation, and interactions. *Cold Spring Harbor perspectives in biology* **2011,** *3* (3), 1-14.

39. Assa-Munt, N.; Jia, X.; Laakkonen, P.; Ruoslahti, E., Solution structures and integrin binding activities of an RGD peptide with two isomers. *Biochemistry* **2001,** *40* (8), 2373-2378.

40. Bogdanowich-Knipp, S. J.; Jois, D. S. S.; Siahaan, T. J., The effect of conformation on the solution stability of linear vs. cyclic RGD peptides. *Journal of Peptide Research* **1999,** *53* (5), 523-529.

41. Bogdanowich-Knipp, S. J.; Chakrabarti, S.; Williams, T. D.; Dillmall, R. K.; Siahaan, T. J., Solution stability of linear vs. cyclic RGD peptides. *Journal of Peptide Research* **1999,** *53* (5), 530-541.

42. Krammer, A.; Craig, D.; Thomas, W. E.; Schulten, K.; Vogel, V., A structural model for force regulated integrin binding to fibronectin's RGD-synergy site. *Matrix biology* **2002,** *21* (2), 139-147.

43. Leahy, D. J.; Aukhil, I.; Erickson, H. P., 2.0 Å Crystal Structure of a Four-Domain Segment of Human Fibronectin Encompassing the RGD Loop and Synergy Region. *Cell (Cambridge)* **1996,** *84* (1), 155-164.

44. Aota, S.-i.; Nomizu, M.; Yamada, K. M., The Short Amino Acid Sequence Pro-His-Ser-Arg-Asn in Human

Fibronectin Enhances Cell-adhesive Function. *The Journal of Biological Chemistry* **1994,** *269*, 24756-24761.

45. Liu, Y.; Ke, Y.; Yan, H.; Rinker, S.; Chhabra, R., Self-assembled DNA nanostructures for distance-dependent multivalent ligand-protein binding. *Nature nanotechnology* **2008,** *3* (7), 418-422.

46. Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. â. G., Analysis and Optimization of Copper-Catalyzed Azideâ□“Alkyne Cycloaddition for Bioconjugation. *Angewandte Chemie International Edition* **2009,** *48* (52), 9879-9883.

CHAPTER 4

OVERVIEW AND BACKGROUND OF SELF-ASSEMBLING DNA CRYSTAL: DESIGNS AND METHODOLOGY

4.1 Historical perspective and fundamentals of structural DNA nanotechnology

The structural determination of proteins is a vital step in fully understanding how they function. Protein structures are traditionally solved using X-ray diffraction; however, this requires the laborious task of determining crystallization conditions for each individual protein. This process is arbitrary and does not always result in crystals amenable to diffraction. Ned Seeman, a protein crystallographer who was frustrated with the capriciousness of this process, was (in his retelling of the story) sitting in a bar when he was inspired by the wood carving, "Depth" by M.C. Escher. To him, the fish resembled a DNA junction with six branches, allowing him to conceive of the possibility that a rationally designed three-dimensional (3D) lattice could be constructed from nucleic acids.[1] This insight stemmed from the ability to utilize the predictable Watson-Crick base pairing that occurs between different DNA strands and the multi-dimensional construction that can be facilitated by Holliday junctions.[2, 3]

Since Seeman's original proposal, several key advancements have been made, with the first occurring in the following year with the discovery of the *immobile* Holliday junction. This construct took the four arm branched structure that is traditionally found in a replication fork during genetic recombination, and imparted an asymmetric (unique base pairing) sequence in the 8 participating junction bases to prevent branch migration (Figure 1.3C).[4] The junction consists of four arms, which are known to stack over each other in an X conformation[5] when in the presence of divalent salts[2], and provides the

ability for the self-assembled crystal scaffold to grow in the third dimension.[6-9] Seeman thus proposed that one could use these DNA based crystals for the site-specific placement at addressable sites for the scaffolding of biomolecules of unknown structures that proved difficult to crystallize on their own, effectively "crystallizing" them for the atomic resolution solution of the guest's structure.

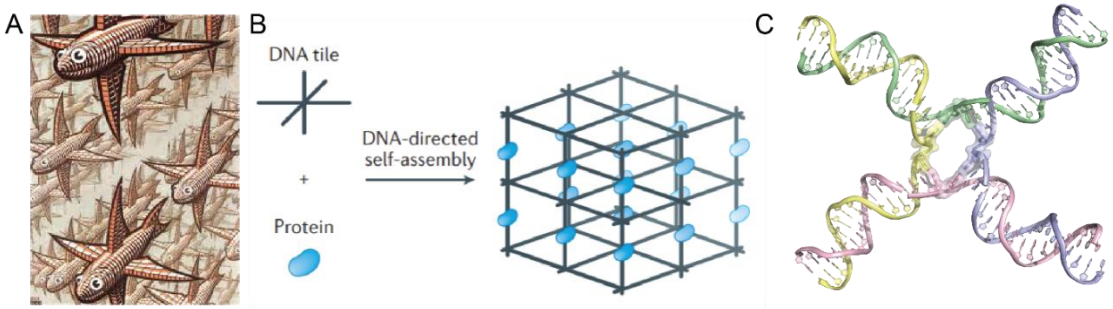


Figure 4.1. Conceptualization of DNA nanotechnology. (A) M.C. Escher's 1955 wood carving "Depth" that inspired Seeman[10] followed by a (B) generalized depiction of his vision to use a six-branch DNA tile to scaffold proteins.[11] (C) Four-arm Holliday junction with the central eight bases that require asymmetry to prevent migration are highlighted in the center of the motif.

Another important aspect towards the design of self-assembling DNA crystals is their ability to properly form contiguous arrays using "sticky end" cohesion. Sticky ends were first employed in molecular biology, where restriction enzymes were used to cut double stranded DNA resulting in two complementary single stranded bases to create an overhang.[12] In DNA nanotechnology, the same notion is employed in order to connect two or more small structures. One example used a self-complementary sticky end to assemble a decamer (10bp duplex) that contained a 2bp sticky end on the 5' end to create an ordered system whose structure was determined using X-ray crystallography.[8] The crystal structure confirmed that these overhangs could help to assemble larger DNA structures without disrupting the known parameters of standard B-form DNA. In 1998,

Winfree and colleagues produced 2D crystalline materials made from the periodic assembly of a DNA structure. They employed a simple double crossover tile as the building block for the periodic array that could form intermolecular interactions with each other using complementary sticky ends at explicitly prescribed sites. In order to determine if the tiles assembled in the intended fashion, a stem-loop or hairpin (a self-complementary region on the same strand) was included, and its location was identified using atomic force microscopy (AFM).[13]

The final significant achievement towards the design of the first 3D lattice was carried out by Chengde Mao and co-workers, which looked at the concept of "tensegrity", the combination of tension and integrity, to construct a rigid DNA structure.[14] Tensegrity uses rigid struts that push out and flexible tendons that pull in, resulting in a stable architecture. In his design, three DNA duplexes were fused together using three four arm junctions to create a triangle. The four arm junctions act as the flexible tendons while the duplexes that connect them act as the rigid struts, resulting in an equilateral triangle with 60° angles. While the individual triangle was difficult to be fully resolved using AFM due to the short 13-nm length of each edge, the addition of sticky ends at the end of each duplex allowed the triangles to create a regular array (200x200 nm) that confirmed the successful assembly of the designed structure.

Since this proposal, the predictable and programmable nature of the Watson-Crick base pair[3] combined with a modified version of the Holliday junction [15] has been exploited to produce various structures.[16, 17] To date, all of the well-established parameters and programmable features of DNA have allowed for the design and

construction of a multitude of nanostructures[11, 19-26] and have provided the foundation towards the design of 3D DNA crystal lattices.

4.2 The Tensegrity Triangle

While Seeman had originally proposed the use of a self-assembling 3D DNA lattice to crystallize biomolecules in 1980,[27] the first successful example of this proposal took 29 years to fully realize.[28]

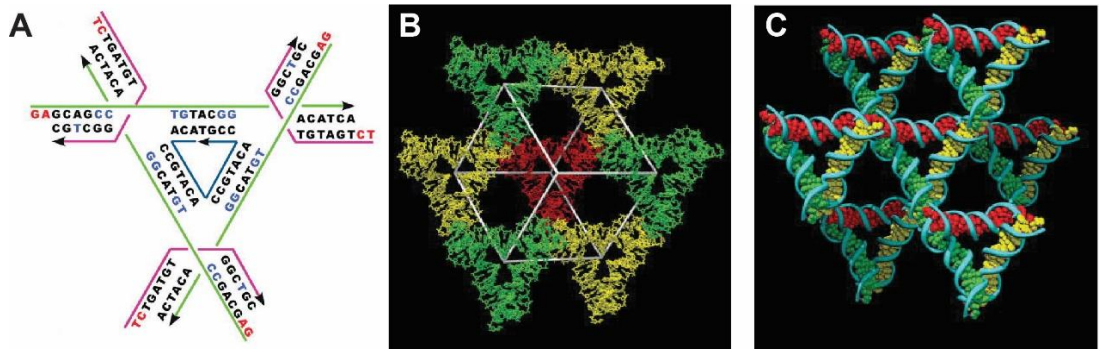


Figure 4.2. Overview of the Tensegrity Triangle system. (A)The basic unit is made of three duplexes, each of which contains component sequences from three oligonucleotides; a linear strand (green), a crossover strand connecting one half of each 4-arm junction (pink), and a "central" strand containing 3 repeats of 7 bases which connect each arm of the motif (blue). The 2bp sticky ends at the end of the 5' of the pink and green strands are highlighted in red. (B) A single triangular unit shown in red, connects to 6 other units at the sticky ends to create a lattice with cavities. (C) The crystal structure showing how individual triangles come together to create a lattice. All images are adapted from Seeman's work.[28]

The "tensegrity triangle" crystal design consists of three duplexes that weave over and under each other in three-dimensional space, with each edge lying on a different axis. Each edge contained identical sequence, and were designed to be 21 bps (2 full helical turns) to allow for a full 720° rotation between complementary 2bp- sticky ends (shown in red). Each of the duplexes contains three different strands: (1) a linear strand (green) that extends across the whole duplex and yields one component sticky end on each edge, (2) a crossover strand (pink) completing one half of each corner, and leaving the sticky

ended complement, and (3) the "central" strand (blue) containing 3 repeats of 7 bases that completes the internal portion of the tensegrity motif (Figure 4.2A). The resulting triangular unit contains a total of 6 arms which connect to the complementary sticky ends of 6 other adjacent triangles (Figure 6.2B), yielding a crystal with rhombohedral symmetry (*R*3). The crystals were formed using hanging drop vapor diffusion (methodology discussed below) in a buffer consisting of cacodylic acid and tris (pH 8.5), $(NH_4)_2SO_4$, and $Mg(CH_3COO)_2$ and $MgCl_2$; the crystals diffracted to 4 Å, and were solved using iodine derivatives. In addition, 3- and 4-turn/edge tensegrity motifs were also designed, each of which resulted in crystals retaining the rhombohedral symmetry; however, the resolutions (6.1-11.0Å) were much lower than the original design, and no structures were reported.

Since the original study, the tensegrity triangle has been modified in a variety of ways including changing the length and sequence of the sticky ends[29], ligating the sticky ends to try and stabilize them in low salts conditions[30], and addition of terminal phosphates in attempts to improve resolution[29, 31]. Other researchers looked at the effect of impurities in the crystallization buffers and in the component DNA strands.[32] Additionally, fluorophores have been added to the units in several ways including diffusion of an external strand with a complementary overhang[33], addition of a triplex strand[34], and introducing two distinct molecules within the asymmetric unit[35]. However, do date no additional density had been observed in any of the structures that were attributable to any of the chemical adducts.

4.3 Methodology for Obtaining DNA Crystals

Obtaining crystals requires two distinct events: nucleation and precipitation. This process occurs when the molecule(s) that make up the crystal begin to assemble in an orderly manner (nucleation) and then reach a threshold where they are no longer soluble (precipitation). The most common technique used for crystallization is vapor diffusion, where the molecule of interest is placed in an aqueous solution (mother liquor) that contains a lower salt concentration than a second much larger mixture that is located separate from it within the same sealed container. The water within the solution containing the molecule of interest will slowly vaporize and relocate to the high salt buffer until the two mixtures equilibrate to equivalent salt concentrations. This equilibration event occurs effectively as a time-dependent gradient based on the starting salt concentration, which can be tuned to enable a gradual increase in the material in the drop to allow for successful crystallization (ordered precipitation). Vapor diffusion can either be carried out by hanging drop, where the solution containing the sample is inverted on a coverslip over a vacuum sealed reservoir, or sitting drop, where the mixture is located up on a pedestal. Since salt concentration is a vital part of producing crystals, the molecule of interest is traditionally screened using a sparse matrix screen which samples a large variety of buffers containing various concentrations of additional salts or additives. Some components of a buffer that are typically screened (fine matrix screen) are pH, ionic strength, addition of external polymers or surfactants, and percent of organic solvents. Another vital aspect for successful crystallization is the purity of the sample, which cannot be understated because even very minor impurities can prevent the initial nucleation event for crystal formation.

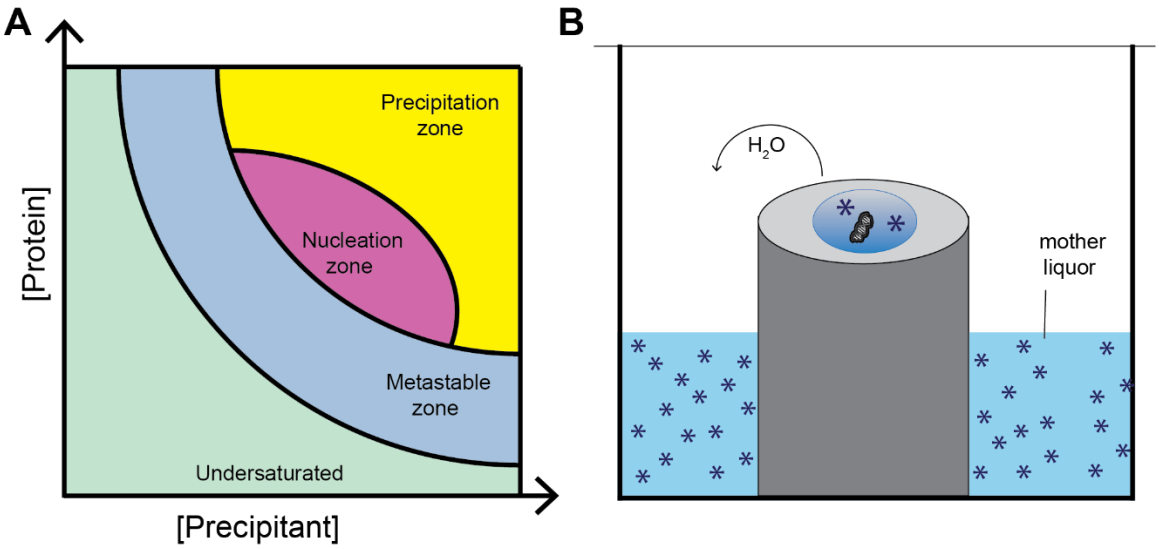


Figure 4.3. Methodology to growing crystals. (A) Phase diagram for crystal growth mediated by precipitant concentration where the undersaturated zone (green) has too low of a concentration for either the protein or precipitant or both, the yellow precipitation zone is supersaturated leading to the protein precipitating from the solution, and the pink nucleation zone where the crystal begins after which the growth occurs in the metastable zone. (B) Schematic showing sitting drop vapor diffusion where a mixture containing the molecule that needs to be crystallized and the crystallization buffer with precipitates are placed onto the pedestal. A reservoir solution with higher concentration of ions allows for water in the mixture to slowly diffuse to the reservoir until the two ion concentrations are in equilibrium, facilitating crystallization.

Obtaining DNA crystals required a few modifications to the conventional techniques used most commonly for proteins. Namely, the DNA crystals also need to go through a temperature gradient (facilitating the annealing process required for DNA) where the mixture is heated to denature the single strands before slowly cooling with an explicitly defined temperature gradient (60 to 23°C). Additionally, the DNA strand sequences are designed (using the NuPack software[36]) so they do not have any secondary structure due to intrastrand interactions 60°C. The component strands can then assemble using their designed sequence complementarity to form the intended lattice as the temperature is gradually decreased. Additionally, the buffer screens were adapted to include additives that are known to stabilize DNA and neutralize the negative charge

found on the phosphate backbone (e.g., divalent salts such as magnesium or polyamines like spermine and spermidine). Two common techniques to introduce a ligand of interest is through soaking, where the ligand is added directly to fully formed crystals and allowed to diffuse into the cavities to its proper binding site, or through co-crystallization where it is introduced to the stock mixture prior to crystallization.

4.4 Basics of Crystallography

A crystal, by definition, is a material that comes together via noncovalent forces to form an ordered three-dimensional array of molecules. It is comprised of copies of a unit cell, which is the smallest group of molecules that repeats to make up the full lattice. A unit cell can also contain an asymmetric unit (ASU) component related through symmetry to comprise the molecular contents of the unit cell (e.g. proteins, DNA, or small molecules).[37]

A basic workflow in crystallography starts with crystal formation, followed by crystal harvesting and cryo-protection, X-ray diffraction, translation of the diffraction data from reciprocal space to real space, and finally model building and refinement. Following crystal growth, the crystals must be retrieved from the crystallization buffer in a process known as harvesting. During this process, a "cryoprotectant" is added to the crystal to ensure that ice formation does not harm the crystalline lattice. Some commonly used cryoprotectants include: sugars (i.e. glucose, trehalose, etc.), paraffin oil, surfactants (i.e. polyethylene glycol (PEG) of varying chain lengths), and often sugar alcohols such as glycerol. The crystal is then removed from the drop using a loop tuned in diameter to the size of the crystals, and instantly submerged ("plunged") into liquid nitrogen to "cryo-cool" it. The crystals must remain in liquid nitrogen during the X-ray experiments to

protect them from the high intensity laser beam during the subsequent diffraction step that occurs at a synchrotron. The X-ray beam is reflected by the atoms within the structure, creating a distinct pattern of Bragg peaks, or "reflections", that is referred to as a diffraction pattern (Figure 4.4B). This pattern is a direct result of the conditions matching Bragg's law: $2d \sin\theta = n\lambda$, where d is the interplanar distance, n is a positive integer, and λ is the wavelength of the wave (Figure 6.4d).[37] Following data collection, the resulting diffraction pattern can be translated from the reciprocal space to real space using a Fourier transform (Figure 6.4A) using several different crystallographic software programs to obtain the three-dimensional electron density maps (Figure 6.4C). For the solution of *de novo* structures, heavy atoms such as selenomethionine (most commonly used in proteins), bromine, iodine, mercury, etc. are incorporated at specified locations in the sample. Locating these positions within the crystal allows for initial determination of preliminary electron density maps, and subsequently aids in the solution and model building of the resulting structure. Once all atomic positions have been roughly fit within the electron density, the model undergoes iterative rounds of "refinement" to ensure that each atom is in the proper position, and all of the obtained density is interpreted properly (Figure 4.4C).

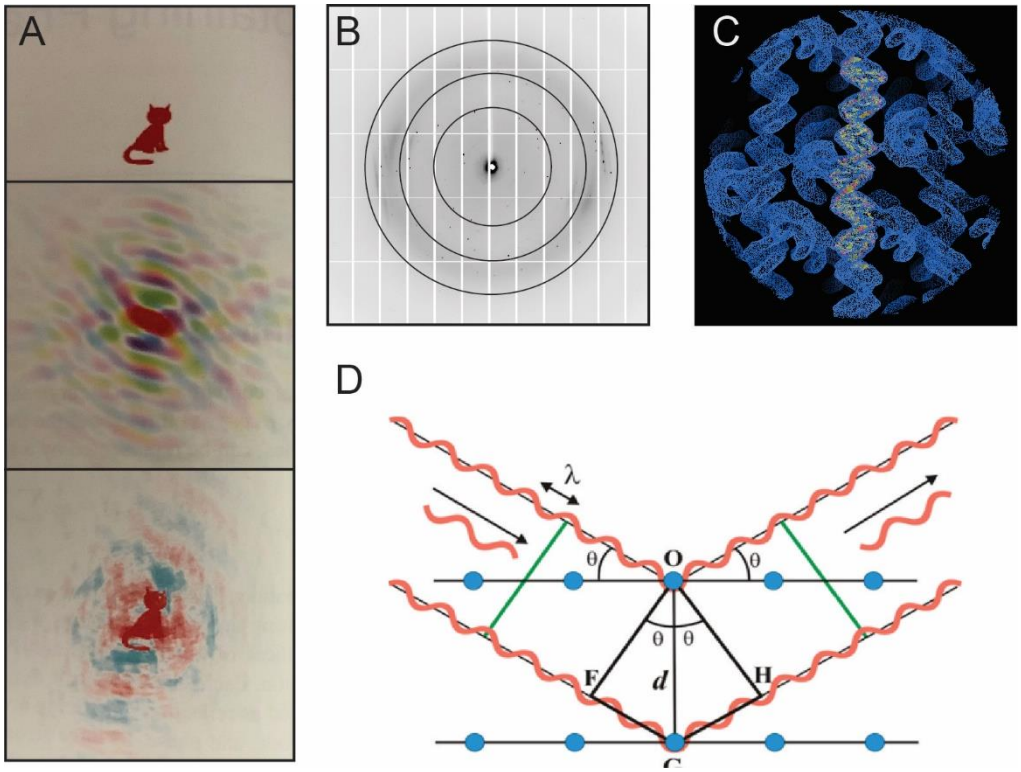


Figure 4.4. Basics of crystallography concepts. (A) Schematic demonstrating the transition from reciprocal to real space to reveal a cat. [37] (B) Example of a collected diffraction pattern with resolution shells that is obtained during electron diffraction. (C)Example of the DNA model being built into the electron density (D) Schematic of Bragg's Law with the variables defined.

Crystals are often classified by their space group, which defines the symmetry of the unit cell. Some examples for these classifications are *P*-primitive, *R*-rhombohedral and, *H*-hexagonal. Following the lattice classification is a number that describes the internal symmetry of the unit cell. One example is rotational symmetry of the screw-axis which is defined as the number of times (n) it contains an identical unit while being rotated up to 360° or $\frac{360}{n}$. A subscript number following the internal symmetry defines the handedness of that symmetry with 1 corresponding to left-handedness and 2 corresponding to right-handedness. For example, $P3_2$ is a primitive (trigonal) packing with a right-handed threefold screw axis.

4.5 The 4x5 Central Weaving Strand System

In 2016, the Yan lab set out to expand on the tensegrity triangle, designing a "tensegrity square" by adding an extra repeat to the central strand so it contained 4 repeats of 7 bases; however, this "4x7" design always yielded poorly ordered crystals that did not diffract well (Figure 4.1A).[38] The central strand was then modified to only contain 5 bp per repeat (the "4x5" design), which readily crystallized and diffracted to 3.1 Å with $P3_221$ symmetry. Since this was a novel structure, heavy atoms were necessary to obtain initial electron density maps so two bromines were incorporated at the C5 position of specified thymines (bromo-dU) into each turn of the duplex.

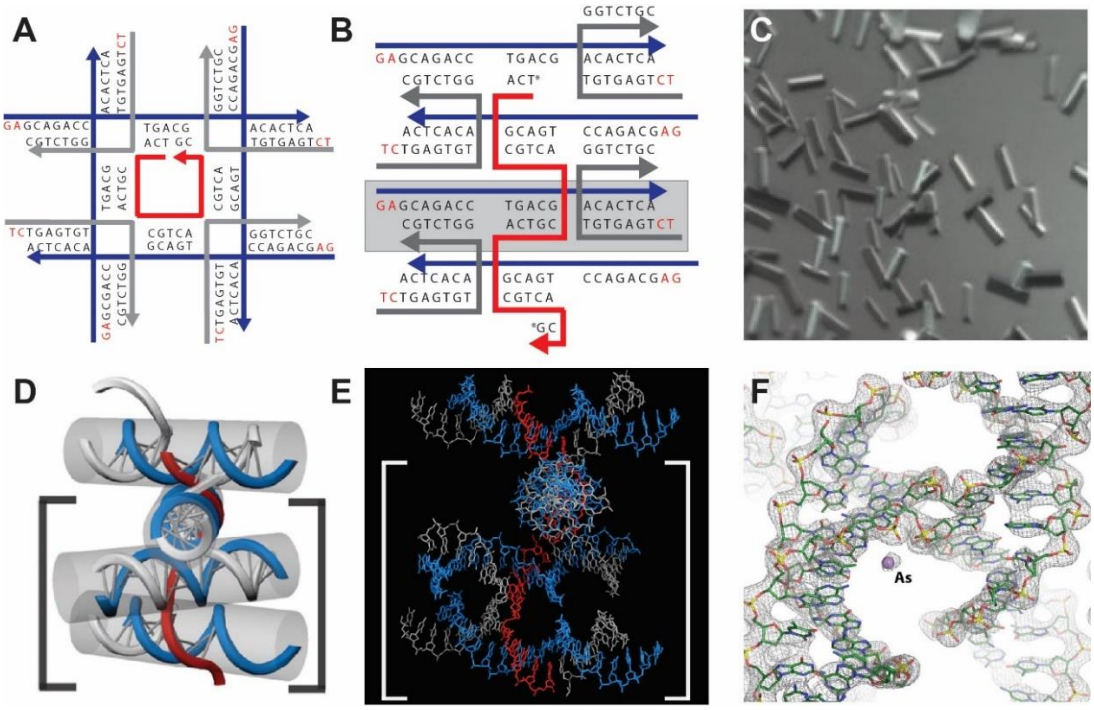


Figure 4.5. Design and structure of the tensegrity square. (A) A 2D topology with the expected tensegrity square design for the 4x5 system alongside (B) the actual topology observed with four duplexes stacked on top of one another, held together by the 4x5 central weaving strand (red). (C) Representative bright field image of the crystals obtained and (D) a cartoon depiction of the central building block with the unit cell bracketed. (E) The actual model built from the electron density highlighting the central

block being four parallel duplexes stacked and rotated from one another and (F) the model built into the density with the arsenic ion that comes from the cacodylic acid buffer.

Unlike the intended designed square structure, the crystal revealed a drastically different lattice system. The structure was defined by a "block" unit containing four duplexes stacked on top of one another, rotated 120° from each adjacent duplex, with the central strand weaving between each of them to effectively "tether" each constituent duplex together (Figure 4.5B, D). Like the tensegrity triangle, each of the blocks is comprised of three distinct component strands, the central weaving strand (S1), the linear strand that makes up one half of each duplex (S2), and a third that forms crossover at each four arm Holliday junction (S3). Since there are four duplexes per block, these strands are at a 1:4:4 stoichiometric ratio. Each block was designed to assemble via sticky end cohesion of complementary 2 bp overhangs that tail each duplex to connect to 8 other units. The unit cell with dimensions of a=b=68 Å, c=59 Å, α=β=90°, and γ=120°, was comprised of three of the four duplexes in the block. While these values were determined through the crystallographic data, they also correspond to the DNA parameters where a and b = 6.8 nm resulting precisely from two helical turns in each duplex and c = 5.9 nm corresponding to the three stacked helices (each duplex has a diameter of 2 nm).

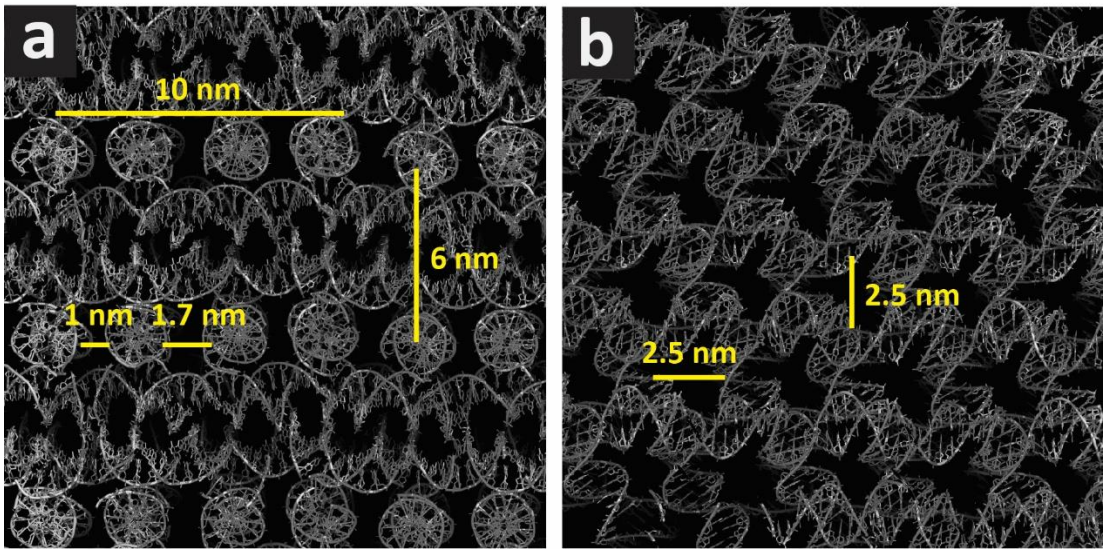


Figure 4.6. Full lattice of the 4x5 system highlighting the aperiodic and densely packed solvent channels. The two views are a 90° rotation of the crystal lattice.

When examining the full symmetry related lattice, it was apparent that the densely packed, aperiodic cavities would eventually only be amenable to very small guests. It was proposed that these aperiodic cavities were the result of torsional strain in the 20 bp central weaving strand since it is one base short of two full helical rotations (21 bp). 20 bases only achieve a 686° rotation which would cause it to over-twist in order complete a full 720° rotation required for each 3' and 5' connection (shown by asterisks in Figure 4.5B). Unaccounted for density in the crystal revealed two additional peaks that were not directly linked to the DNA, but were located at opposing corners near the junction crossover (Figure 6.5F). After examining the buffer components for the crystals used for the structures, it was determined that the only heavy atom that could account for this extra density was the arsenic in the cacodylic acid since no other buffer components could account for the density. While the full structure of cacodylic acid, $(CH_3)_2AsO_2H$,

did not fit within the density, just the arsenic ion was still in good agreement with these positions.

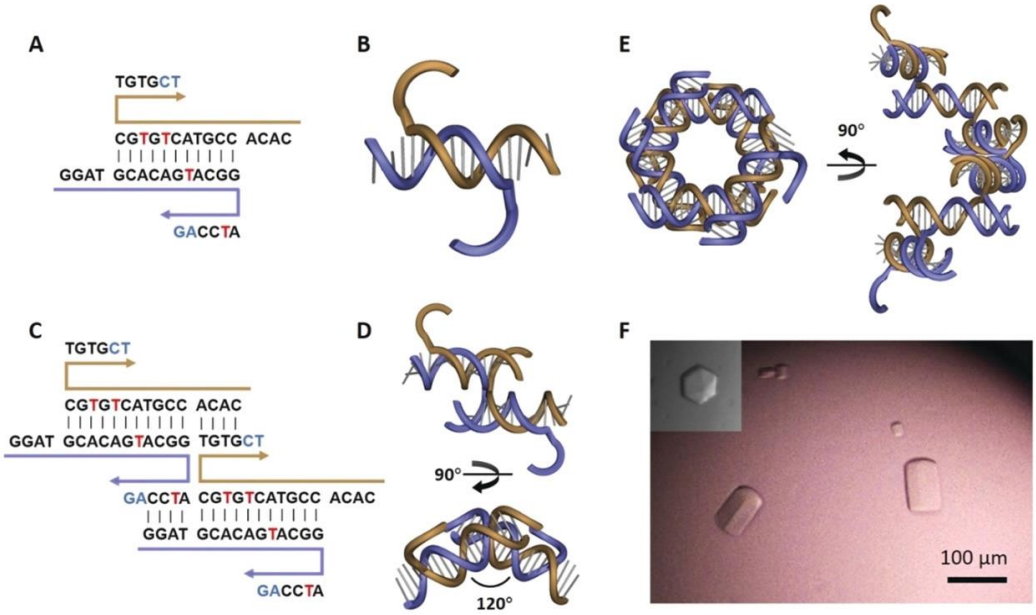


Figure 4.7. Rationally designed six-fold symmetry DNA crystal. (A) Schematic of the duplex motif used in the six-fold system comprised of two strands, each 21 bases in length (brown and purple), with 2bp sticky ends shown in blue. The two DNA strands form an 11 bp duplex leaving four unpaired ssDNA regions. (B)3D cartoon model of the duplex and (C) schematic demonstrating how two of the duplex motifs come together with their single stranded regions to create a Holliday junction. (D) Two orientations rotated 90° from one another, of a 3D model showing the interactions of two duplexes to highlight the 120° angle they create. (E) A 3D model showing the interactions of six duplex motifs shown at two orientations. (F) Optical images of self-assembled hexagonal prism DNA crystals.

Since 2016, the Yan lab has built upon the design principles learned from the 4x5 system which they were able to apply towards the design of a novel crystal motif containing six-fold crystal symmetry.[39] Unlike the other crystal designs, the structure is comprised of only two unique strands, which make up an 11- bp duplex that creates a layered hexagonal lattice with *P*6 symmetry. In addition to the 11- bp duplex, the two strands have four single stranded regions, two are 4 bp and two that are 6 bp. The single

stranded regions are designed to pair with the four bp complement within the six bp regions leaving a 2-bp sticky end (Figure 4.7A, C). Each block in this design creates a duplex that subsequently layers with another block unit creating a 120° angle between each adjacent layer, which results in 6 consecutive layers that complete a full 360° rotation, resulting in the designed 6-fold symmetry (Figure 4.7D, E).

4.6 Proposed Applications

DNA crystals were originally proposed to scaffold biomolecules in precise arrays, allowing for their structural determination through X-ray crystallography. Biomolecules such as proteins can be difficult to crystallize on their own, so it was proposed that scaffolds that could crystallize "automatically", while simultaneously immobilizing the target at its assigned location, and could allow for the guest's structure to be determined.[40] In order for this application to be feasible, several major challenges must first be addressed, the most difficult of which includes scaffolding the biomolecule with deliberate orientation and location. This in turn, requires it to either naturally bind to a specific DNA sequence (e.g. A transcription factor), or a form to be specifically conjugated to a DNA strand (with methods such as those described in Chapter 3). Additionally, the cavity size of the lattice needs to be large enough to host the target. The biomolecule could be introduced through two different techniques, either through diffusion (soaking) after the crystal lattice has been formed, or through co-crystallization. Both options have shortcomings: diffusion experiments requires that the protein be small enough to diffuse into the solvent channels of the crystal to reach its designated location, while co-crystallization requires the protein to be thermally stable at minimum of 60°, so that it remains properly folded during the entire thermal annealing process required for

crystallization of the DNA scaffold. For both techniques, the buffer components (salt concentration, pH, volatility) in which the crystal is grown also need to be compatible with maintenance of the protein integrity.

In addition to the originally proposed application, 3D DNA crystals have the potential to be used as molecular sieves as initially demonstrated by Paukestelis[41], or as a scaffold for catalysis.[42, 43] They may also be employed for nanoelectronics, plasmonics, or photonics by templating nanoparticles.[44-46] There have been several examples of the crystals being used as a scaffold for small molecules, such as fluorophores and polyamines.[33-35, 47,30, 48-50] With this subfield of DNA nanotechnology still in its infancy, it is expected that additional applications for these lattices will continue to be developed.

4.7 Building a DNA Crystal Toolbox

While the tensegrity triangle, 4x5, and six-fold systems have all elucidated several design rules for self-assembling DNA crystals, there is still a vast amount of research that needs to be done to fully understand aspects like; the size of cavity, sequence asymmetry, overall packing, and the symmetry of lattice required to fully realize future applications. It is proposed that there are several ways to change the overall packing and symmetry, several of which will be explored in the following chapters. While one obvious way to change the lattice is to design an entirely new system, as demonstrated with the six-fold symmetry design[39], it is also possible to utilize the modularity of the existing motifs by altering multiple parameters (Figure 4.8).

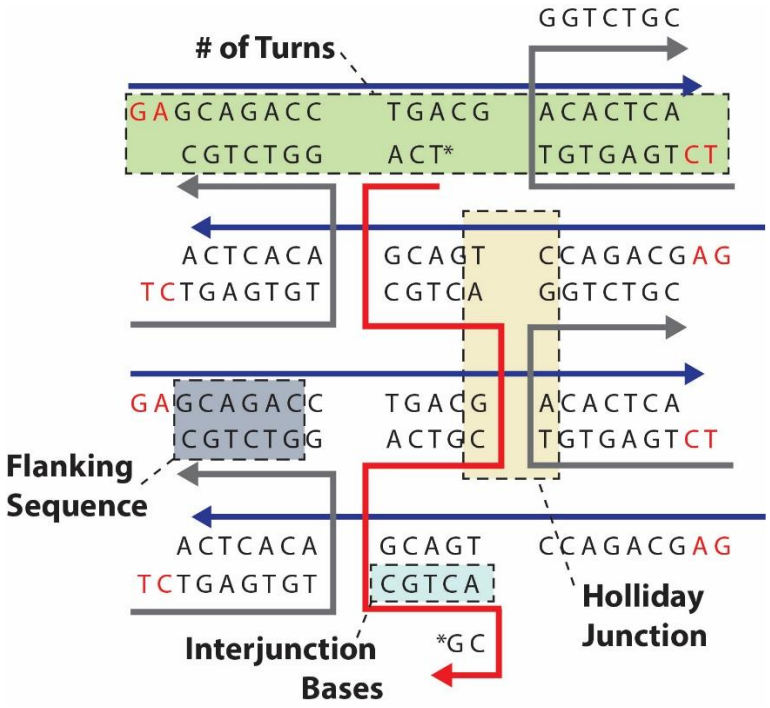


Figure 4.8. Variables to alter for crystal design. A number of parameters that can be explored to fully understand crystal design rules are highlighted with a colored box with number of helical turns (green), HJ sequence (orange), flanking sequence (gray), and number of interjunction bases (blue).

Modifications to the original 4x5 system could include changing the number of bases between the Holliday junctions (interjunction bases) or changing the number of helical turns of each constituent helix which could lead to larger cavities. It is possible, however, that additional turns may lead to issues with obtaining high resolution structures due to the higher solvent content and reduced rigidity of the crystal lattice. Additional parameters that could have an impact on resolution and packing is the sequence of the immobile Holliday junction (to date only a single sequence had been used), as well as the bases that flank it. It is expected that these variations could potentially lead to higher resolution structures, which would enable a more comprehensive understanding of the molecular details of the lattice, and that changes in the packing and cavity size could

prove vital to the discovery of effective constructs for hosting a broad variety of guest molecules.

4.8 References

1. Seeman, N. C., NUCLEIC-ACID JUNCTIONS AND LATTICES. *Journal of Theoretical Biology* **1982,** *99* (2), 237-247.

2. Duckett, D. R.; Murchie, A. I. H.; Diekmann, S.; Vonkitzing, E.; Kemper, B.; Lilley, D. M. J., THE STRUCTURE OF THE HOLLIDAY JUNCTION, AND ITS RESOLUTION. *Cell* **1988,** *55* (1), 79-89.

3. Watson, J. D.; Crick, F. H., Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. J.D. Watson and F.H.C. Crick. Published in Nature, number 4356 April 25, 1953. *Nature (London)* **1974,** *248* (5451), 765-765.

4. Kallenbach, N. R.; Ma, R. I.; Seeman, N. C., AN IMMOBILE NUCLEIC-ACID JUNCTION CONSTRUCTED FROM OLIGONUCLEOTIDES. *Nature* **1983,** *305* (5937), 829-831.

5. Eis, P. S.; Millar, D. P., CONFORMATIONAL DISTRIBUTIONS OF A 4-WAY DNA JUNCTION REVEALED BY TIME-RESOLVED FLUORESCENCE RESONANCE ENERGY-TRANSFER. *Biochemistry* **1993,** *32* (50), 13852-13860.

6. Zhang, S. W.; Seeman, N. C., SYMMETRICAL HOLLIDAY JUNCTION CROSSOVER ISOMERS. *Journal of Molecular Biology* **1994,** *238* (5), 658-668.

7. Fu, T.-J.; Tse-Dinh, Y.-C.; Seeman, N. C., Holliday Junction Crossover Topology. *Journal of molecular biology* **1994,** *236* (1), 91-105.

8. Miick, S. M.; Fee, R. S.; Millar, D. P.; Chazin, W. J., Crossover isomer bias is the primary sequence-dependent property of immobilized Holliday junctions. *Proceedings of the National Academy of Sciences of the United States of America* **1997,** *94* (17), 9080-9084.

9. Ho, P. S.; Eichman, B. F., The crystal structures of DNA Holliday junctions. *Current Opinion in Structural Biology* **2001,** *11* (3), 302-308.

10. Nadrian, C. S., Art as a Stimulus for Structural DNA Nanotechnology. *Leonardo (Oxford)* **2014,** *47* (2), 142-149.

11. Seeman, N. C.; Sleiman, H. F., DNA nanotechnology. *Nature Reviews Materials* **2018,** *3* (1).

12. Stanley, N. C.; Annie, C. Y. C.; Herbert, W. B.; Robert, B. H., Construction of Biologically Functional Bacterial Plasmids In Vitro. *Proceedings of the National Academy of Sciences - PNAS* **1973,** *70* (11), 3240-3244.

13. Winfree, E.; Liu, F. R.; Wenzler, L. A.; Seeman, N. C., Design and self-assembly of two-dimensional DNA crystals. *Nature* **1998,** *394* (6693), 539-544.

14. Liu, D.; Wang, M.; Deng, Z.; Walulu, R.; Mao, C., Tensegrity: Construction of Rigid DNA Triangles with Flexible Four-Arm DNA Junctions. *Journal of the American Chemical Society* **2004,** *126* (8), 2324-2325.

15. Holliday, R., A mechanism for gene conversion in fungi (Reprinted). *Genetics Research* **2007,** *89* (5-6), 285-307.

16. Ho, P. S., Structure of the Holliday junction: applications beyond recombination. *Biochemical Society Transactions* **2017,** *45*, 1149-1158.

17. Shrestha, P.; Emura, T.; Koirala, D.; Cui, Y. X.; Hidaka, K.; Maximuck, W. J.; Endo, M.; Sugiyama, H.; Mao, H. B., Mechanical properties of DNA origami nanoassemblies are determined by Holliday junction mechanophores. *Nucleic Acids Research* **2016,** *44* (14), 6574-6582.

18. Gelbin, A.; Schneider, B.; Clowney, L.; Hsieh, S.-H.; Olson, W. K.; Berman, H. M., Geometric Parameters in Nucleic Acids: Sugar and Phosphate Constituents. *Journal of the American Chemical Society* **1996,** *118* (3), 519-529.

19. Seeman, N. C., DNA Nanotechnology at 40. *Nano letters* **2020,** *20* (3), 1477-1478.

20. Hu, Y.; Niemeyer, C. M., From DNA Nanotechnology to Material Systems Engineering. *Advanced Materials* **2019,** *31* (26).

21. Veneziano, R.; Ratanalert, S.; Zhang, K. M.; Zhang, F.; Yan, H.; Chiu, W.; Bathe, M., DNA NANOTECHNOLOGY Designer nanoscale DNA assemblies programmed from the top down. *Science* **2016,** *352* (6293).

22. Seeman, N. C., DNA nanotechnology: Novel DNA constructions. *Annual Review of Biophysics and Biomolecular Structure* **1998,** *27*, 225-248.

23. Seeman, N. C., DNA in a material world. *Nature* **2003,** *421* (6921), 427-431.

24. Deng, Z. X.; Lee, S. H.; Mao, C. D., DNA as nanoscale building blocks. *Journal of Nanoscience and Nanotechnology* **2005,** *5* (12), 1954-1963.

25. Pinheiro, A. V.; Han, D. R.; Shih, W. M.; Yan, H., Challenges and opportunities for structural DNA nanotechnology. *Nature Nanotechnology* **2011,** *6* (12), 763-772.

26. Zhang, C. A.; He, Y.; Su, M.; Ko, S. H.; Ye, T.; Leng, Y. J.; Sun, X. P.; Ribbe, A. E.; Jiang, W.; Mao, C. D., DNA self-assembly: from 2D to 3D. *Faraday Discussions* **2009,** *143*, 221-233.

27. Weiss, P. S., A conversation with Prof. Ned Seeman: Founder of DNA nanotechnology. *ACS nano* **2008,** *2* (6), 1089-1096.

28. Zheng, J. P.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R. J.; Constantinou, P. E.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **2009,** *461* (7260), 74-77.

29. Ohayon, Y. P.; Hernandez, C.; Chandrasekaran, A. R.; Wang, X. Y.; Abdallah, H. O.; Jong, M. A.; Mohsen, M. G.; Sha, R. J.; Birktoft, J. J.; Lukeman, P. S.; Chaikin, P. M.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., Designing Higher Resolution Self-Assembled 3D DNA Crystals via Strand Terminus Modifications. *Acs Nano* **2019,** *13* (7), 7957-7965.

30. Zhao, J. M.; Chandrasekaran, A. R.; Li, Q.; Li, X.; Sha, R. J.; Seeman, N. C.; Mao, C. D., Post-Assembly Stabilization of Rationally Designed DNA Crystals. *Angewandte Chemie-International Edition* **2015,** *54* (34), 9936-9939.

31. Sha, R. J.; Birktoft, J. J.; Nguyen, N.; Chandrasekaran, A. R.; Zheng, J. P.; Zhao, X. S.; Mao, C. D.; Seeman, N. C., Self-Assembled DNA Crystals: The Impact on Resolution of 5 '-Phosphates and the DNA Source. *Nano Letters* **2013,** *13* (2), 793-797.

32. Stahl, E.; Praetorius, F.; Mann, C. C. D.; Hopfner, K. P.; Dietz, H., Impact of Heterogeneity and Lattice Bond Strength on DNA Triangle Crystal Growth. *Acs Nano* **2016,** *10* (10), 9156-9164.

33. Hao, Y. D.; Kristiansen, M.; Sha, R. J.; Birktoft, J. J.; Hernandez, C.; Mao, C. D.; Seeman, N. C., A device that operates within a self-assembled 3D DNA crystal. *Nature Chemistry* **2017,** *9* (8), 824-827.

34. Rusling, D. A.; Chandrasekaran, A. R.; Ohayon, Y. P.; Brown, T.; Fox, K. R.; Sha, R. J.; Mao, C. D.; Seeman, N. C., Functionalizing Designer DNACrystals with a Triple-Helical Veneer. *Angewandte Chemie-International Edition* **2014,** *53* (15), 3979-3982.

35. Wang, T.; Sha, R. J.; Birktoft, J.; Zheng, J. P.; Mao, C. D.; Seeman, N. C., A DNA Crystal Designed to Contain Two Molecules per Asymmetric Unit. *Journal of the American Chemical Society* **2010,** *132* (44), 15471-15473.

36. Zadeh, J. N.; Steenberg, C. D.; Bois, J. S.; Wolfe, B. R.; Pierce, M. B.; Khan, A. R.; Dirks, R. M.; Pierce, N. A., NUPACK: analysis and design of nucleic acid systems. *J Comput Chem* **2011,** *37*, 170-173.

37. Rhodes, G., *Crystallography made crystal clear a guide for users of macromolecular models*. 3rd ed. ed.; Elsevier/Academic Press: Amsterdam ;, 2006.

38. Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* **2016,** *138* (38), 12690-12690.

39. Zhang, F.; Simmons, C. R.; Gates, J.; Liu, Y.; Yan, H., Self-Assembly of a 3D DNA Crystal Structure with Rationally Designed Six-Fold Symmetry. *Angewandte Chemie-International Edition* **2018,** *57* (38), 12504-12507.

40. Liu, D.; Wang, M. S.; Deng, Z. X.; Walulu, R.; Mao, C. D., Tensegrity: Construction of rigid DNA triangles with flexible four-arm DNA junctions. *Journal of the American Chemical Society* **2004,** *126* (8), 2324-2325.

41. Paukstelis, P. J., Three-dimensional DNA crystals as molecular sieves. *Journal of the American Chemical Society* **2006,** *128* (21), 6794-6795.

42. Alexis, T. B., The Impact of Nanoscience on Heterogeneous Catalysis. *Science (American Association for the Advancement of Science)* **2003,** *299* (5613), 1688-1691.

43. Geng, C.; Paukstelis, P. J., DNA Crystals as Vehicles for Biocatalysis. *Journal of the American Chemical Society* **2014,** *136* (22), 7817-7820.

44. Stebe, K. J.; Lewandowski, E.; Ghosh, M., Materials science. Oriented assembly of metamaterials. *Science (American Association for the Advancement of Science)* **2009,** *325* (5937), 159-160.

45. Robert, J. M.; Byeongdu, L.; Matthew, R. J.; Nadine, H.; George, C. S.; Chad, A. M., Nanoparticle Superlattice Engineering with DNA. *Science (American Association for the Advancement of Science)* **2011,** *334* (6053), 204-208.

46. Jones, M. R.; Osberg, K. D.; Macfarlane, R. J.; Langille, M. R.; Mirkin, C. A., Templated Techniques for the Synthesis and Assembly of Plasmonic Nanostructures. *Chemical reviews* **2011,** *111* (6), 3736-3827.

47. Zhao, J. M.; Zhao, Y.; Li, Z.; Wang, Y.; Sha, R. J.; Seeman, N. C.; Mao, C. D., Modulating Self-Assembly of DNA Crystals with Rationally Designed Agents. *Angewandte Chemie-International Edition* **2018,** *57* (50), 16529-16532.

48. Melinger, J. S.; Sha, R.; Mao, C.; Seeman, N. C.; Ancona, M. G., Fluorescence and Energy Transfer in Dye-Labeled DNA Crystals. *The journal of physical chemistry. B* **2016,** *120* (48), 12287-12292.

49. Gottesfeld, J. M.; Melander, C.; Suto, R. K.; Raviol, H.; Luger, K.; Dervan, P. B., Sequence-specific Recognition of DNA in the Nucleosome by Pyrrole-Imidazole Polyamides. *Journal of molecular biology* **2001,** *309* (3), 615-629.

50. Kang, J. S.; Meier, J. L.; Dervan, P. B., Design of Sequence-Specific DNA Binding Molecules for DNA Methyltransferase Inhibition. *Journal of the American Chemical Society* **2014,** *136* (9), 3687-3694.

CHAPTER 5

SELF-ASSEMBLING DNA CRYSTALS WITH TUNABLE CAVITY SIZE AND CHIRALITY

5.1 Introduction

Rational design of all self-assembled DNA crystal systems was inspired by the "tensegrity triangle", which provided the conceptualization for the discovery of the "4 x N" crystal systems. Specifically, the crystals were comprised of three component oligonucleotides that self-assemble into continuous linear layers of inter-connected duplexes tethered together by one of the constituent strands (see details in Chapter 4.5). The 4x5 system was defined by a central "weaving" oligonucleotide that contained four identical repeats of five nucleotides which mediated the assembly of the 3D layered helical arrays (See Chapter 4.5 for full design details). The crystal lattice of the 4x5 structure yielded an aperiodic array of small cavities with volumes untenable for hosting guest molecules of any appreciable size, and homogeneously organized to achieve the eventual goal of scaffolding proteins, peptides, etc. to solve *de novo* crystal structures of the guest species. It was hypothesized that the aperiodic arrays (Figure 4.6) observed were the effect of strain originating from the strand being comprised of 20 nucleotides, one nucleotide short of two full turns (21 bp/turn), and thus rendering it incapable of completing two full 360° rotations, as seen in the tensegrity triangle. The tensegrity triangle employs a 21 nt central strand (three repeats of seven nucleotides) to centrally connect the three duplexes that make up its structure without adding any strain. A single base deficit results in 686° rotation, ~34° short of the desired 720° for the 5' and 3' of

that central weaving strand, causing torsional strain as it overtwists to compensate for the deficit causing the aperiodicity of the structure.

Although the 4x5 system yielded an unintended result, the structure enabled important insights into the principles governing the rational design of novel self-assembling 3D DNA crystals. However, many fundamental concepts and applications are still not well established, and further studies are necessary for the design of structures with higher complexity. It was hypothesized that the addition of an extra base between each junction would lead to a tethered array of four repeats of a six-base sequence (the "4x6" system) that would relax the torsionally strained lattice observed in the 4x5 system due to the additional ~34° twist between junctions.

Naturally occurring DNA exclusively forms a right-handed helix (D-DNA)[1], however, its enantiomeric form (L-DNA) can be synthesized using standard solid phase DNA synthesis techniques using left-handed phosphoramidites, and displays identical structural properties (helical periodicity, diameter, etc.) as its counterpart[2]. With the capability of producing left-handed DNA readily, this study also sought to demonstrate that chirality is also a controllable design parameter of self-assembling DNA crystals where identical sequences could be explored in parallel with right-handed DNA to yield enantiomeric lattices (Figure 5.1B, C).

Due to the fact that L-DNA does not naturally appear in any biological context, it is completely resistant to nuclease digestion. Because of this property, it provides the ability to assemble mirror image crystals that are significantly more stable and resistant to degradation within biological environments compared to D-DNA crystals.[3-7] Since the fundamental goal of using these crystal systems is for eventually scaffolding various

biomolecules, the inherent resistance of L-DNA to nuclease degradation provides an attractive attribute as a protective measure for protecting guest molecules positioned within the lattice. Furthermore, rendering the DNA crystals stable to biological environments could allow for additional applications, such as using them to densely organize drug molecules, and act as a delivery vehicle *in vivo*. To test this concept, both enantiomers of the 4x6 crystals were produced and characterized to compare resistance to degradation in buffers containing nucleases.

5.2 Design and Characterization of the 3D Lattice

The 4x6 motif is comprised of three component strands: S1, the "central weaving strand", which contains four repeats of six bases (so 24 nt in total length); S2, which is a continuous strand that comprises one half of the 2-turn duplex (21 nt), and S3, which contains 15 bases, which serve as crossover strands complementary to S2 sequences in adjacent layers. These three strands are assembled at a 1:4:4 (S1:S2:S3) stoichiometry to yield a building block containing four duplexes tethered together by the central weaving strand (S1) forming four-arm "Holliday" junctions[8-11] with each S3 strand. Each of these duplexes is flanked by complementary 2bp sticky ends, which allow these building blocks to assemble to form the overall lattice. (Figure 5.2). With the four identical duplexes that make up the building block, the 4x6 is considered a symmetric design; differentiating each of the duplexes by giving each its own unique symmetry would render the block asymmetric, a concept that will explored in section 5.4.

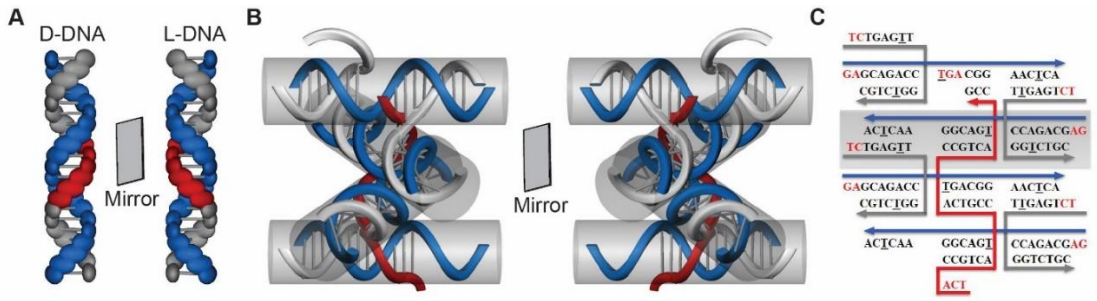


Figure 5.1. Topology of the building block design and cartoon depiction of the enantiomeric forms of DNA in the context of the crystal building blocks. of the difference between the right- and left-handed DNA as a duplex (A) Mirror image duplex representation showing the handedness of D- and L-DNA, right- and left-handed, respectively. (B) The structural features leading to the assembly of the crystal are identical, however the chirality of each layer, mediated by the central strand, controls the handedness of the rotation of each layer with respect to one another. (C) Schematic of the assembly of the 3 component strands with S1 shown in red, S2 shown in blue, and S3 in gray. The modified seleno-dU bases are underlined, and the sticky end sequences are highlighted in red. The crystallographic asymmetric unit (ASU) is comprised of one duplex outlined in a gray box.

Crystals were obtained for each enantiomer of the 4x6 motif and were found to contain $P3_2$ symmetry rather than the $P3_221$ symmetry of the 4x5 system, which indicated that packing was profoundly influenced by the additional base in each repeat. This unique symmetry meant that derivative crystals were required to obtain initial maps to solve each of the structures. Both the left-handed and right-handed structures contained a heavy atom on specified thymines at the C5 position, replacing the methyl group (Figure 5.1C) to facilitate crystallographic phase determination. The L-DNA crystals incorporated commercially obtained bromo-dU while the D-DNA crystals used a seleno-dU phosphoramidite that was synthesized in the lab. Both versions were successfully crystallized (Figure 5.2 A&E) and the resulting structures were determined to ~ 3.1Å resolution, and contained the designed right- and left-handed symmetries, $P3_2$ and $P3_1$, respectively. The full collection and data statistics are shown in Table A.1 in Appendix

A. Each respective model was built into the resulting electron density maps from the derivative crystals (Fig. 5.3 C&F), and the refined models were deposited in the Protein Data Bank (PDB) under accession codes 5VY6 and 5VY7.

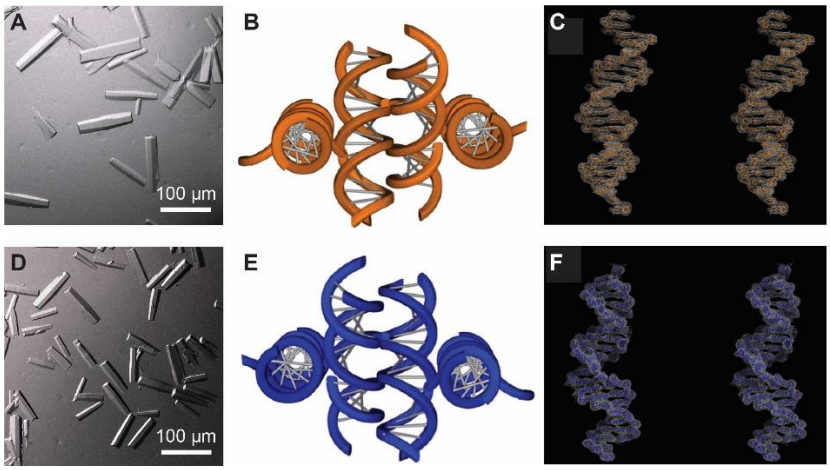


Figure 5.2. Right- and left-handed DNA crystal morphology, handedness, and models in density. (A) Representative light images of the crystals obtained for the D-DNA with a scale bar of 100 µm scale bar. (B) Cartoon depiction of the right- handed central building block and (C) a stereo view of the D-DNA model built into its electron density contoured at σ=1.8 (D) Representative light images of the crystals obtained for the L-DNA with a scale bar of 100 µm scale bar. (B) Cartoon depiction of the left- handed central building block and (C) a stereo view of the L-DNA model built into its electron density contoured at σ=1.8

All duplexes within each unit cell are related by trigonal (*P*3) symmetry, and the full lattice can be visualized using the 3D molecular visualization software PyMOL[12]. It was apparent that the additional base in each repeat of the central weaving strand led to the assembly of the designed four-duplex block and the crystal did in fact contain the expected periodic array of cavities not previously observed in the 4x5 motif (Figure 5.3). As dictated by the design, each layer was rotated 120° from one another, and based on measurements performed using PyMOL, the cavities were approximately 3.4 x 2 x 5 nm in size (Figure 5.3A & B), corresponding to one helical turn of B-form DNA, the width

of each duplex in the layer, and the cross-section of the cavity, respectively, when viewed along the three-fold axis (Figure 5.3B). Although the 4x5 and 4x6 systems contained nearly identical cell dimensions, the lattice packing was starkly different, which accounted for the difference in space groups ($P3_221$ vs. $P3_2$). The single base deficit in the 4x5 repeats clearly lead to a structurally strained lattice that contained aperiodic arrays of cavities with spacing between duplexes at intervals of 1.0 and 1.7 nm and pores of only 2.5 nm along the three-fold axis (one half of those in the 4x6). Volumes of the cavities were calculated by classifying the 4x5 cavities as triangular prisms and the 4x6 cavities as hexagonal prisms. The parameter for height was considered to be 6 nm in both systems (Figure 5.3 A, C) which corresponds to the height of three stacked layers that make up the unit cell. Final calculations showed that the 4x5 system only contained a cavity volume of ~23 nm^3 while the 4x6 increased the volume more than 25-fold, to ~615 nm^3. Additionally the crystal content values for the two systems were compared and it was found that the 4x6 solvent content for these crystals was approximately 70% (compared to ~55% in the 4x5 lattice)[13]. This result is noteworthy because it was originally thought increasing solvent content would lead to reduced diffraction quality, but the resulting resolution was equivalent (3.1 Å).

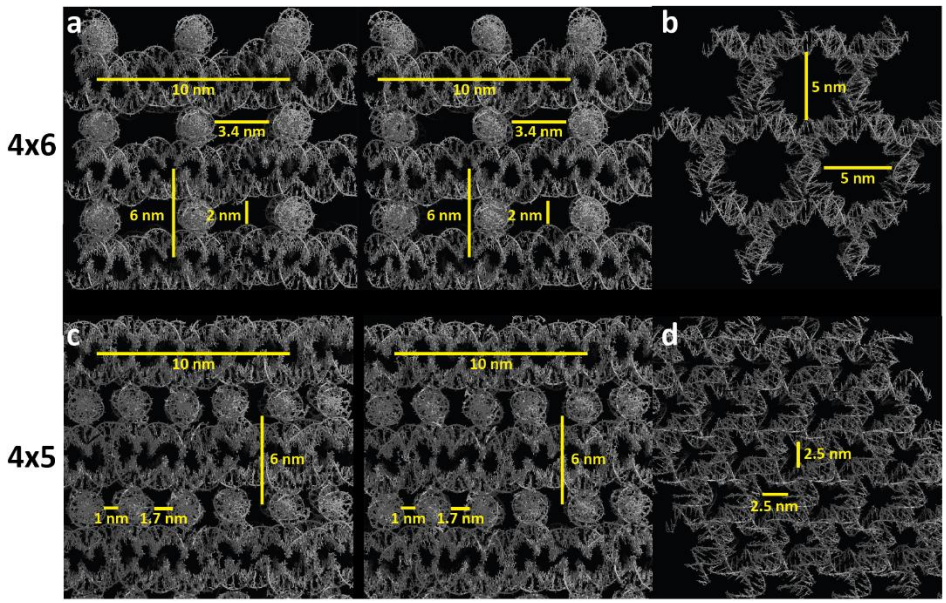


Figure 5.3. Comparison of the 4x6 and 4x5 scaffolds to crystal cavity periodicity and sizes. (A) Crystal packing of the 4x6 lattice with $P3_2$ symmetry. Yellow bars are indicated for measurements of cavity sizes. Distances were measured in PyMOL to provide values that approximate the sizes of the cavities from edge to edge of neighboring helices. The 6 and 10 nm scale bars are used as references for distance. Regular periodicity of each cavity is evident. (B) 90-degree rotation with respect to (A) with a view down the three-fold symmetry axis with pores of ~5 nm in size. (C) Crystal packing of the 4x5 lattice with $P3_221$ symmetry. Yellow bars are indicated for measurements of cavity sizes. The 6 and 10 nm scale bars are used for reference. The crystal structure revealed an aperiodic array of densely packed cavities. The resulting symmetries of each lattice yield vastly different packing. (D) 90-degree rotation with respect to (C) along the three-fold symmetry axis showing ~2.5 nm size pores.

5.3 Nuclease Resistance of the L-DNA Crystal

To demonstrate the resistance to nuclease degradation of the L- vs. D-DNA systems, crystals were obtained by broad matrix screening across 48 buffers to identify identical conditions that produced crystals with adequate size and morphology to allow for a direct comparison for crystals of each corresponding handedness. Additionally, since the buffers often contained components (i.e., high salt or volatile solvents) that do not naturally exist in native environments of the DNA nuclease, the activity of the nuclease had to be tested in each of the representative crystallization buffers to ensure

that the components did not inhibit or reduce its ability to degrade. For this, a simple right-handed 21 bp duplex containing the sequence contained in the crystal was utilized and monitored in the presence of DNase 1.

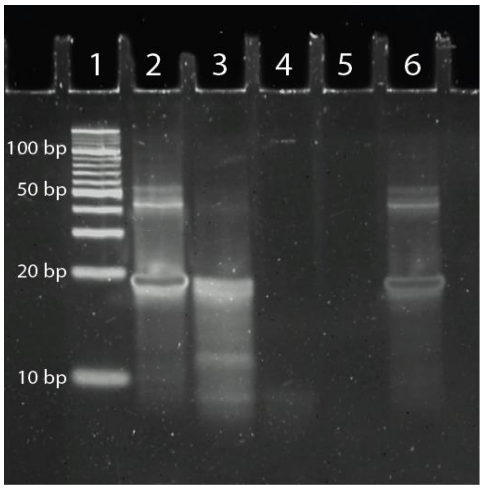


Figure 5.4 15% native PAGE testing the activity of DNase I in crystallization buffer. Lane 1 is a double stranded DNA ladder, lane 2 is a 21 bp D-DNA duplex with no added DNase, lanes 3-5 5-, 10-, and 15-minute incubation of the duplex with DNase; and lane 6 was incubated with heat inactivated DNase for 15 minutes.

The duplex was annealed by mixing the two oligonucleotides at 100 µM each, in the crystallization buffer containing 50 mM HEPES (pH 7.5) with 80 mM $MgCl_2$ and 2.5 mM spermine and annealed by heating at 95°C and cooling to 4°C over 30 minutes. The stock was split into 5 x 10 µL aliquots, and then subsequently tested for efficacy at varying conditions. The first aliquot only contained the duplex and was incubated at 37°C (the optimal temperature for the nuclease) to determine the ability of the duplex to remain intact at the increased temperature without denaturing (Lane 2). While no degradation of the duplex was observed, there did seem to be some aggregation of the duplex itself (upper band between 40 and 50bp), likely due to the strands themselves not being purified beforehand. The second, third, and fourth aliquots all received 1 µL (2 units) of

DNase 1 and were incubated at 37°C for 5, 10 and 15 minutes, respectively (Lanes 3-5). Lane 3, the 5-minute incubation, showed the initial signs of degradation with the aggregate band disappearing and the start of smearing indicating the presence of many duplexes of varying lengths. Lane 4 (10-minute incubation) showed near complete degradation with only a bit of smearing remaining, followed by the complete digestion at 15 minutes, indicated by no visible band remaining. The final aliquot contained 2 units of heat inactivated DNase 1 to serve as a negative control to confirm that the disappearance of the bands in lane 5 was solely attributed to nuclease degradation and not another factor such as aggregation of the nuclease and the duplex which could have prevented it from migrating into the gel.

Upon confirmation of the activity of DNase 1, crystals of both enantiomeric forms were prepared as described above. The thermal stability of the crystals was then probed to ensure that they were stable at the working temperature of the nuclease. This experiment was performed by monitoring the stability of the crystals by incrementally increasing the incubation temperature, beginning at room temperature, and increasing the temperature at 2°C intervals every 12 hours. The crystals were monitored under the light microscope at each temperature point to observe their ability to maintain their morphology compared to T_0 until they became damaged and eventually dissolved. The crystals were found to be stable at the optimal nuclease temperature of the nuclease (37°) and remained undisturbed until incubation at 44°C at which point began to show signs of cracking and started to dissolve. A time course study was then carried out to demonstrate how the left-handed crystals could prevent degradation of any biomolecule it hosted by taking bright images of

the D- and L- crystals at various timepoints until complete degradation of the right-handed crystals was observed (Figure 5.5).

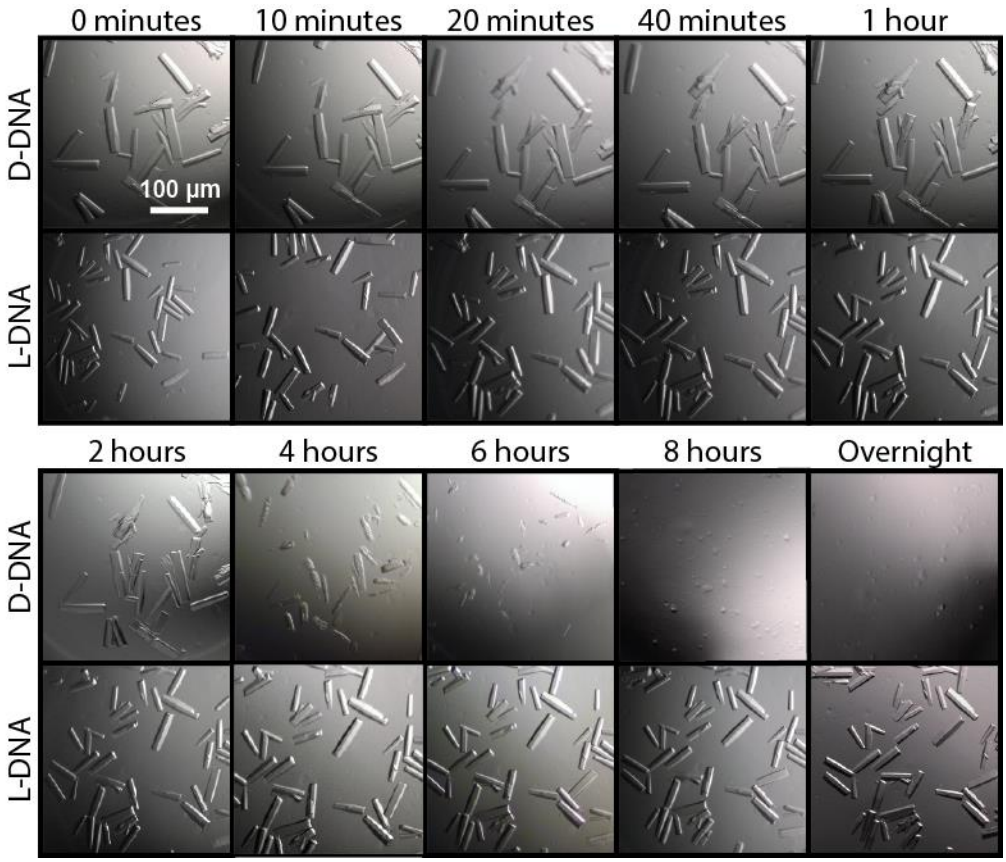


Figure 5.5. Time course of incubation of D-DNA crystals (top panels) and L-DNA crystals (bottom panels) with DNase I in a buffer containing 50 mM HEPES (pH 7.5) with 80 mM $MgCl_2$ and 2.5 mM spermine over 24 hours to demonstrate crystal stability at 37°C. Light images of representative time points of 0, 4, 6, and 24 hours are shown. Degradation of the D-DNA crystals became apparent at 4 hours, with nearly complete dissolution at 6 hours, whereas the L-DNA crystals suffered no visible damage throughout the experiment. Scale bar: 100 µm.

The initial time point (T_0) shows the integrity and quality of the crystals before incubation with any DNase. After imaging the starting point, 1 µL (2 Units) of DNase 1 was added to each drop before it was resealed and incubated at 37°C. The tray was removed from incubation and quickly imaged at each remaining time point (10, 20, 40 minutes, and 1, 2 ,6 ,8 hours and overnight) before continuing the incubation. The

morphology of the crystals in both the D- and L- samples remained intact until the 4-hour timepoint, when the D-DNA (top row) crystals showed initial signs of degradation. These crystals showed even more degradation over the next few timepoints and ended in complete disintegration after overnight incubation. By contrast, the L-DNA crystals remained completely intact over the entire time course with no observable damage. In addition, to demonstrate that the left-handed DNA was indeed impervious to DNase digestion, the crystals were continually monitored, and no appreciable damage after two weeks was observed. While these studies were all carried out in the buffer that provided the best hits for observing crystal morphology, the initial screen did provide several other options which were also explored, and which provided consistent stability regardless of the crystallization buffer.

In addition to monitoring crystal morphology, crystals (where intact) were also removed from the drop at each of the major time points and analyzed with a urea-based denaturing polyacrylamide gel to show that the DNA was being degraded by the DNase I and was not just simply redissolving. Prior to running each sample (Figure 5.6), the crystals or crystal containing solutions were subjected to heating to dissolve any remaining crystals. Each of the three component strands (S1, S2, S3) for both the D- and L- crystals are clearly visible at T_0 (Lanes 1 and 2) in accordance with their respective lengths. The 4- hour timepoint for the D-DNA sample only shows slight degradation, as the band is slightly lighter; however, after 6 hours (lane 5), the bands were barely visible, demonstrating that the nuclease digestion is nearly complete, corresponding to the images shown in Figure 5.5. By 24 hours, no bands were visible for the right-handed DNA (lane 7) while the left-handed DNA bands (Lane 8) show no signs of degradation over the

complete time course, thus demonstrating that the disappearance of the D-DNA crystals was a direct result of nuclease degradation, and not dissolution of the crystals.

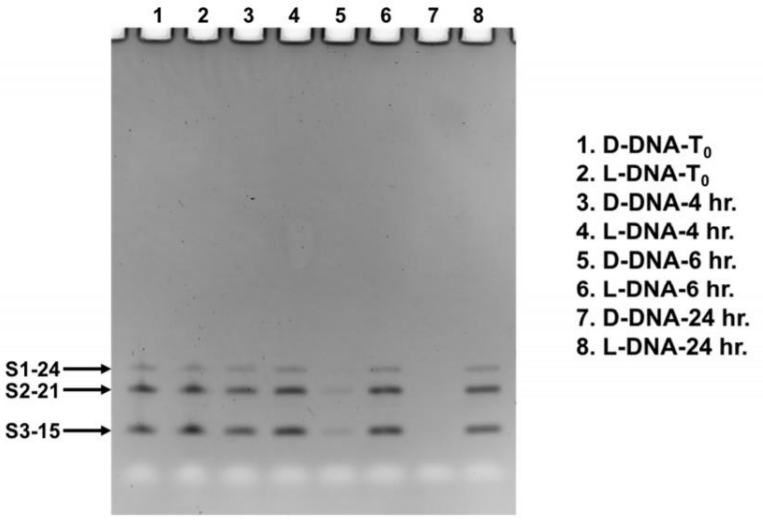


Figure 5.6. 12% Denaturing PAGE to probe the degradation of the three component strands that make up the crystal at several time points for both the D- and L-DNA systems

The results from these studies indicate that a designed L-DNA crystalline lattice is completely resistant to the nucleases that are ubiquitous in nature. Additionally, this resistance to nuclease could offer protection to any biomolecules hosted by the lattice. While only resistance to nuclease degradation was investigated, it is reasonable to expect that this resistance could extend to proteases such as thrombin, since the enzyme would be excluded from entering the lattice due to its dimensions (4.8 x 3.6 x 5.0 nm) which are larger than the cavity sizes. Attempts to model thrombin into the lattice makes it apparent that it would not be able to successfully enter the scaffold as it overlaps with the DNA duplexes in the lattice (Figure 5.7A). It is likely that the smaller pore sizes would protect any hosted proteins, such as an engrailed homeodomain (Figure 5.7B) from degradation due to size exclusion of the protease (Figure 5.7C).

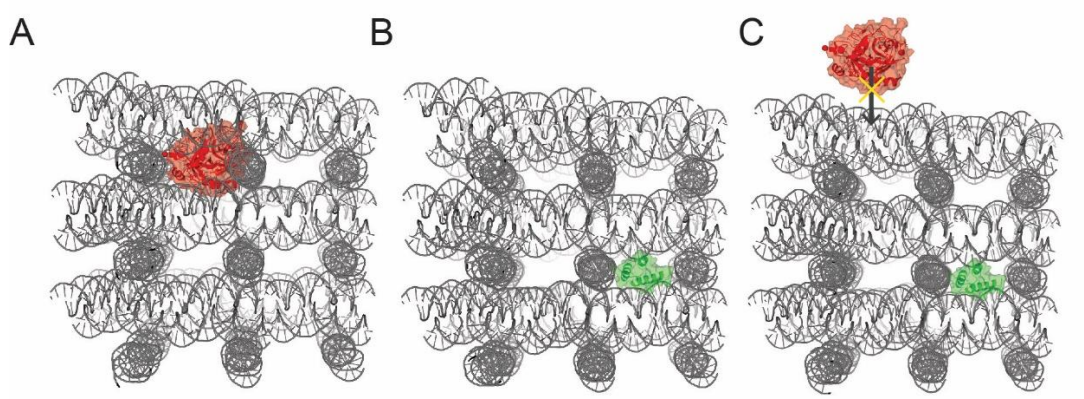


Figure 5.7. Size exclusion of protease. (A) A model protease, thrombin (PDB code: 3U69), is shown docked into the 4x6 cavity to highlight that it is not size compatible. (B) A representative small protein (2HDD) that could be hosted in the lattice and (C) its protection from degradation due to the size exclusion of thrombin.

5.4 Transition to an Asymmetric System

Both central weaving strand designs, the 4x5 and 4x6 systems, are comprised of four duplexes containing identical sequences. While this symmetry has certain advantages, namely the simplicity of having only three component strands and thus easier control over stoichiometry, it also has some notable shortcomings. In a symmetric system, it is impossible to address a specified duplex in the design because each of the four identical duplexes that make up the central building block are averaged together in the solved structure. Due to the statistical averaging of the 5' to 3' nicks throughout the lattice, there are no observable breaks in density at those positions, with the same concept being applicable to any guest molecules in the lattice. To remedy these issues, the 4x5 and 4x6 systems were designed to contain 4 asymmetric duplexes containing unique sequences. The motif was designed using a central scaffolding strand containing four segments of five or six bases with entirely unique sequences, while at the same time maintaining the bases that participate in each constituent Holliday junction to avoid any

possible perturbation to the system. Additionally, the flanking regions and sticky ends of each duplex all contained unique sequences. An example of the sequence modification that were made to render the 4x6 motif asymmetric are highlighted in Figure 5.8.

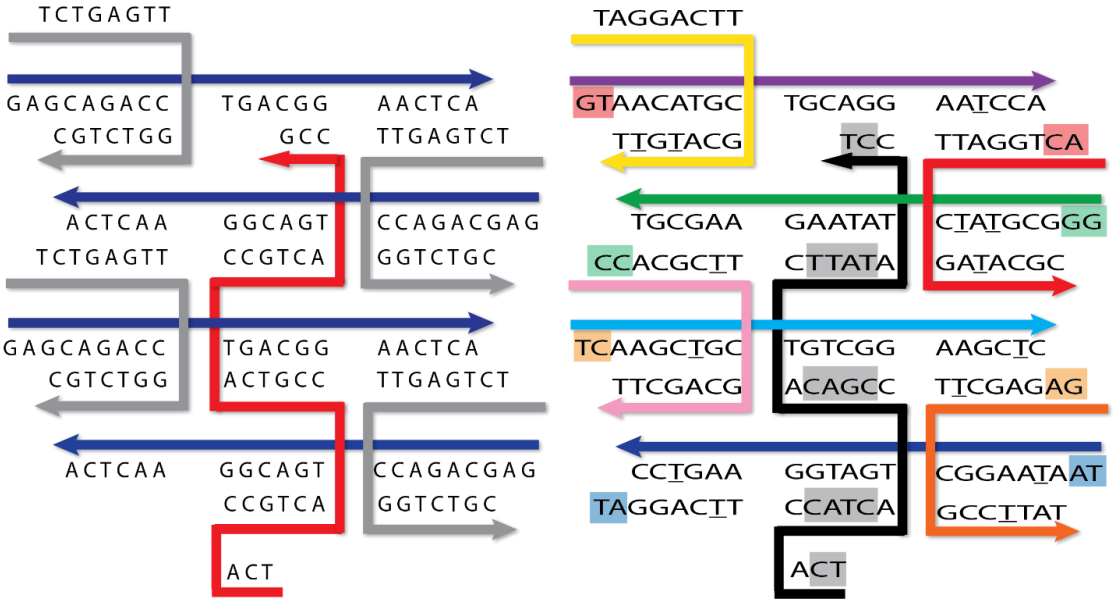


Figure 5.8. Comparison of the 4x6 symmetric and asymmetric designs. The left panel shows the symmetric design containing only three unique strands (red, blue, gray) and the same GA/CT sticky end on each duplex. The right panel displays the sequences in the asymmetric version with each unique sticky end pair boxed in a unique color. Additionally, the gray boxes highlight the four central nucleotides in the 4x6 central strand that are unique to their designed duplex, the remaining 2 bases in the central strand are part of the junction sequence and were not modified from the symmetric version.

Instead of being comprised of three component strands, the asymmetric design contains nine unique strands: one S1 shown in black, four different S2 strands shown in cold colors (purple, green, light blue, and dark blue), and four different S3 strands shown in warm colors (yellow, red, pink, and orange). Each of the nine strands are incorporated at an equal (1:1) stoichiometric ratio. However, increasing the number of component strands from three to nine introduces many additional places for error in the stoichiometry or possible impurities from individual oligonucleotide stocks, either of

which could prove fatal for crystallization. This fact was made apparent when neither the 4x5 or 4x6 asymmetric systems crystallized as readily as their symmetric versions, requiring several iterations of buffer optimization and additional purification of each component strand before obtaining crystals with adequate morphology and size (Figure 5.9).

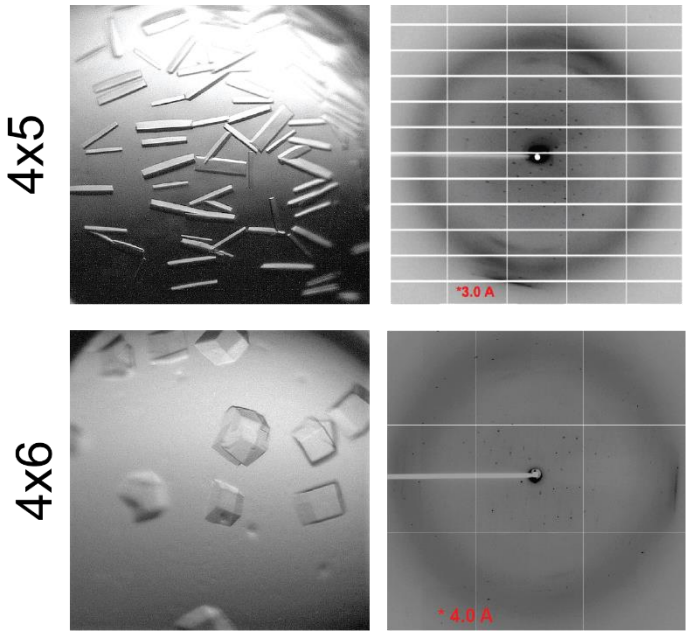


Figure 5.9. Representative bright field image and diffraction patterns for the 4x5 and 4x6 asymmetric systems.

After the diffracting crystals were indexed, it was apparent that the motif was indeed asymmetric; however, this meant that derivative data sets were required to solve the structure. To accomplish this, each duplex was designed to contain either three or four heavy atom incorporations so that each individual duplex could be differentiable in the resulting electron density. In the 4x5 system, a derivative dataset at a resolution of 3.1 Å has been collected; however, building the four duplex model has proven to be difficult and requires more rounds of refinement, or potentially a new set of better phases.

Interestingly, the 4x5 asymmetric system was shown to exhibit $P3_2$ symmetry like the original symmetric 4x6 system, indicating that the asymmetric central strand causes less torsional strain than the symmetric version. The 4x6 asymmetric motif, also exhibited $P3_2$ symmetry, but thus far has only diffracted to a resolution of 4 Å making initial model building even more difficult. The best electron density maps that have been collected to date are displayed in figure 5.10A. Despite the resolution being suboptimal, an initial model was built into the electron density based on the observable junctions and helical nature of the maps (Figure 5.10 B, C). Based on the fit shown here, it is clear that better maps will be essential to obtaining the atomic detail that is needed for the structure. To work with the initial data obtained with the asymmetric system, a molecular model based on the symmetric structure was then recolored to match the intended design (Figure 5.8C) to highlight the extra complexity of the asymmetric system.

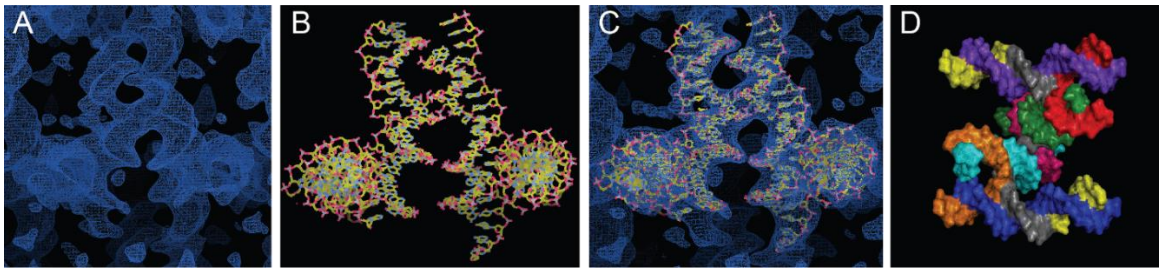


Figure 5.10. Initial electron density maps obtained from the 4x6 derivative system that diffracted to 4.0 Å resolution (A). The initial model that was built into the electron density (B) along with an overlay of the density and the model to indicate that while unfinished the design does produce the intended packing (C). The initial model was recolored to correspond with the 2D topology shown in figure 5.8 (D) to show the complexity of the 9-component strand-based design.

Despite neither of the asymmetric models being completed, this project has significant promise. Once completed, these designs will expand the possible applications of 3D DNA crystals by allowing multiple guests to be scaffolded at discrete locations

within the crystal. Additionally, it was fortuitously discovered that by giving each segment a unique sequence, the asymmetric 4x5 system produced periodic cavities, thus making the resulting scaffold entirely unique from its symmetric counterpart. The asymmetric system will be the first example of a self-assembling DNA crystal with all unique component strands, adding another design parameter to the "toolbox".

5.5 Conclusion

This structural determination of the 4x6 motif was only the third reported self-assembling 3D DNA crystal design, and provided a potentially new route towards the organization of biomolecules for structural solution of a guest. The motif was found to have larger and periodic cavities that are better suited for the discrete scaffolding of guests when compared to the 4x5 motif. Additionally, the cavities were shown to have accessible solvent channels that could allow for diffusion of larger guests, such as proteins, uniformly throughout the scaffold. In this work, it was also shown that lattices of a desired handedness can be readily designed by using L-DNA, which in the future could serve as a nuclease- or protease-protective framework for the encapsulated target molecule. Eventually, the increased stability achieved by left-handed crystals could allow them to serve as scaffolds for delivery of functional or therapeutic proteins *in vivo.* A novel 4x6 asymmetric system was also designed and crystallized, and initial diffraction and model building efforts show significant promise towards an assembly which could be used for a wider variety of applications that require discrete positioning of multiple entities.

5.6 References

1. Watson, J. D.; Crick, F. H., Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. J.D. Watson and F.H.C. Crick. Published in Nature, number 4356 April 25, 1953. *Nature (London)* **1974,** *248* (5451), 765-765.

2. Mandal, P. K.; Collie, G. W.; Kauffmann, B.; Huc, I., Racemic DNA Crystallography. *Angewandte Chemie-International Edition* **2014,** *53* (52), 14424-14427.

3. Hauser, N. C.; Martinez, R.; Jacob, A.; Rupp, S.; Hoheisel, J. D.; Matysiak, S., Utilising the left-helical conformation of L-DNA for analysing different marker types on a single universal microarray platform. *Nucleic Acids Research* **2006,** *34* (18), 5101-5111.

4. Urata, H.; Shinohara, K.; Ogura, E.; Ueda, Y.; Akagi, M., MIRROR-IMAGE DNA. *Journal of the American Chemical Society* **1991,** *113* (21), 8174-8175.

5. Urata, H.; Ogura, E.; Shinohara, K.; Ueda, Y.; Akagi, M., SYNTHESIS AND PROPERTIES OF MIRROR-IMAGE DNA. *Nucleic Acids Research* **1992,** *20* (13), 3325-3332.

6. Ashley, G. W., MODELING, SYNTHESIS, AND HYBRIDIZATION PROPERTIES OF (L)-RIBONUCLEIC ACID. *Journal of the American Chemical Society* **1992,** *114* (25), 9731-9736.

7. Damha, m. J.; giannaris, p. A.; marfey, p., antisense l/d-oligodeoxynucleotide chimeras - nuclease stability, base-pairing properties, and activity at directing ribonuclease-h. *Biochemistry* **1994,** *33* (25), 7877-7885.

8. Duckett, d. R.; murchie, a. I. H.; diekmann, s.; vonkitzing, e.; kemper, b.; lilley, d. M. J., the structure of the holliday junction, and its resolution. *Cell* **1988,** *55* (1), 79-89.

9. Seeman, N. C., DNA structural engineering using immobile junctions. *Current Opinion in Structural Biology* **1991,** *1* (4), 653-661.

10. Liu, Y. L.; West, S. C., Timeline - Happy Hollidays: 40th anniversary of the Holliday junction. *Nature Reviews Molecular Cell Biology* **2004,** *5* (11), 937-U21.

11. Holliday, R., A mechanism for gene conversion in fungi (Reprinted). *Genetics Research* **2007,** *89* (5-6), 285-307.

12. Delano, W. L. *The PYMOL Molecular Graphics System*, DeLano Scientific: San Carlos, CA, 2002.

13. Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction

and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* **2016,** *138* (38), 12690-12690.

CHAPTER 6

A COMPREHESIVE CRYSTALLOGRAPHIC STUDY OF ALL 36 IMMOBILE HOLLIDAY JUNCTION SEQUENCES

6.1 Background of Holliday Junctions

All self-assembling DNA crystals utilize "Holliday" junctions as the central building block for 3D assemblies. The Holliday junction (HJ) was first described in 1963 by Robin Holliday as the mechanism for gene conversion in fungi, whereby two homologous chromosomes undergo strand exchange to create a four-arm branched structure.[1] The naturally occurring HJ undergoes branch migration due to symmetry within the nucleotides at the branch point of each arm. HJs were rendered immobile (unable to undergo migration) in 1983 when Seeman introduced asymmetry into eight junction bases (Figure 6.1A). Since then, the structure has been studied in detail using a variety of techniques including FRET (fluorescence resonance energy transfer),[2-5] gel electrophoresis,[6, 7] and molecular dynamic simulations.[8, 9] HJs have three distinct conformations when in solution, including an open structure (Figure 6.1A), which occurs when no divalent salts are present, and two stacked isomer conformations (Figure 6.1B).[10-12]

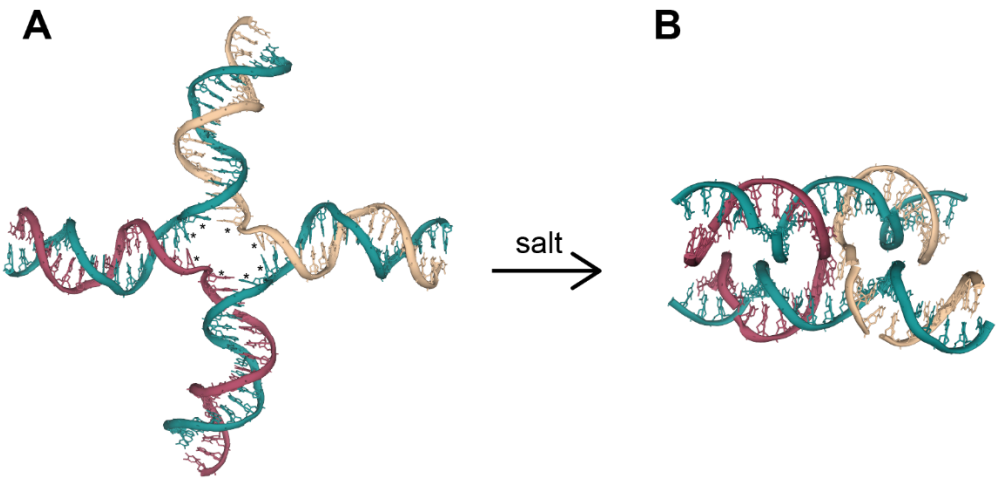


Figure 6.1. Open and Stacked HJ Conformations. (A) The open HJ conformation that occurs when in low salt solutions due to charge repulsion with the eight bases that need to be asymmetric (labeled with an asterisk) and (B) the stacked conformation that occurs when salts are present with two linear strands (teal) and two crossover strands (red and tan).

In addition to these solution-based studies, several experiments analyzed the structure of HJs using X-ray crystallography.[13, 14] The findings showed that a stacked junction adopts an interduplex angle of approximately 60°, but the local structure of each structure was highly dependent on the sequence of the junction, as well as the flanking sequences.[14-19] Shing Ho defined several structural parameters of HJs in 2004, including the interduplex angle (IDA), J_{roll}, J_{twist}, and J_{slide}.[20] The distance from the center of the junction to the end of the duplex is considered the IDA (Figure 6.2). A similar parameter is J_{twist} which projects the distance onto a resolving plane, and is independent of arm length.

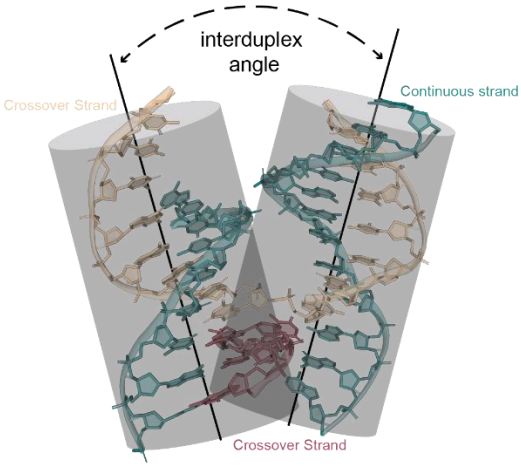


Figure 6.2. Interduplex angle (IDA). A common geometric parameter that is studied in different Holliday junctions, the IDA defines the distance between the center of the two arms with the planes intersecting at the center of the junction.

The J1 immobile junction sequence was the first to be described and characterized and has historically been used as the most favored building block for self-assembled 3D crystal systems throughout DNA nanotechnology for decades. To date, every single published structure (tensegrity triangle, 4x5, 4x6, and six-fold) all utilize the J1 without exception, despite the fact that there are a total of 36 immobile junction sequence combinations available.[21-24] It is not particularly clear why the lack of diversity in the junction sequences used in these systems has never been explored, aside from the original reports nearly 40 years ago describing J1 as “ideal”. Nevertheless, it was hypothesized that the junction sequence can in fact have a major effect on the overall lattice of these self-assembling crystals, thus a comprehensive survey for each individual junction would be exceptionally useful.

6.2 Experimental Design, Crystallization, and Initial Results

Two model systems, the 4x5[22] and the 4x6,[23] were chosen to explore the effect of all 36 immobile junction sequences. The DNA crystals self-assemble into continuous arrays (Figure 6.3A), containing three component strands: the central weaving strand (S1) that contains the 4xN motif, the linear strand (S2) that spans each duplex, and the strand that completes each HJ by crossing over between two adjacent duplexes (S3). These three component strands make up the central building block (Figure 6.3B) which contains two base "sticky ends" that tail each constituent duplex (and that facilitate the growth of each layer in the lattice), as well as a Holliday junction that mediates the assembly of the crystal in 3D space (Figure 6.3C). Within these designed lattices, the HJ is designed to adopt a single isomer,[25, 26] with S1 and S3 always participating as the crossover strands. A uniform numbering system was assigned to both the open and closed junction conformers, to define the sequence positions on each strand that comprise the junction (Fig. 1D), and to subsequently define all possible immobile junction sequences (J1-J36) for both the 4x5 and 4x6 systems (Figure 6.3E). The original sequences of the 4x5 and 4x6 systems were left unaltered with the exception of the eight bases that participate in the junction, which were replaced systematically to contain all 36 junction sequences (Figure 6.4).

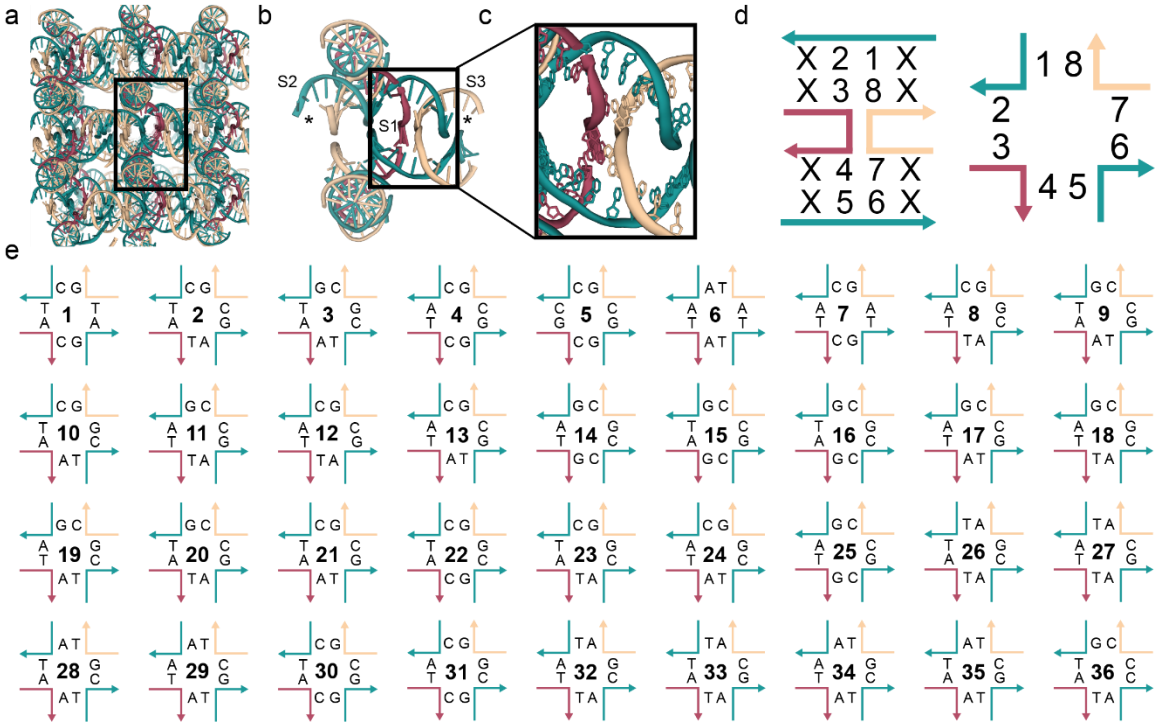


Figure 6.3. Experimental Design of the Junction Study and HJ Sequences. (A) The 4x6 lattice with the central building block boxed along with a blown-up version of that block (B) with the three component strands labeled and sticky ends marked with an asterisk. One of the HJs in the central building block is boxed off and shown in the zoom in, (C) with the S2 (teal) strands remaining linear and the S1(red) and S3 (tan) strands creating the crossover. (D) Defining each of the eight base positions within the 4xN systems in both a stacked and open conformation with consistent strand coloring and (E) all 36 immobile junction sequences defined as 1-36 in the open conformation.

As an initial metric for the influence of the junction, the remaining 35 possible junction sequences were screened across all 48 crystallization buffers in both the 4x5 and 4x6 system to screen for candidates capable of crystallizing, and to rule out those that proved “fatal”. Exhaustive screening attempts were made for those junctions that proved challenging in order to make an ultimate determination of the fate that each sequence rendered, and these results will be summarized when describing the results from each of the 4x5 and 4x6 systems, respectively, at the end of each section below.

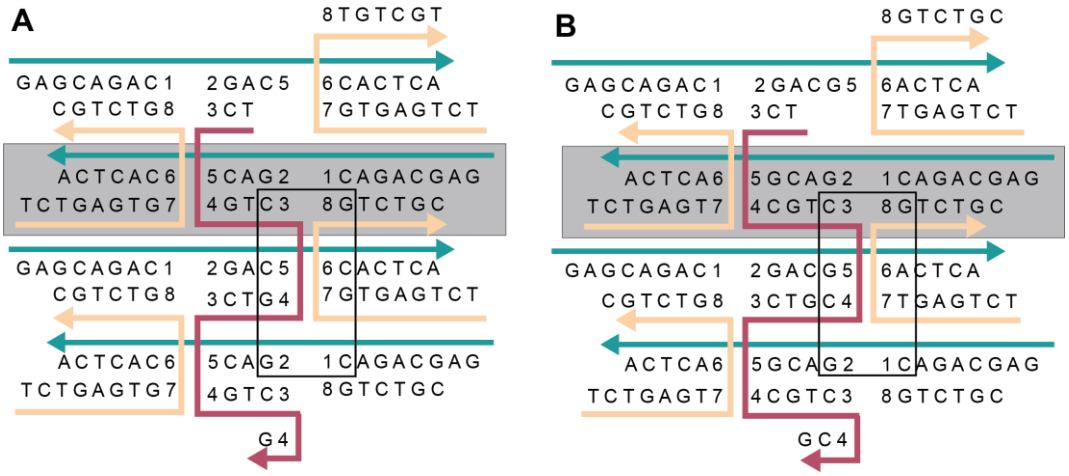


Figure 6.4. 2D topologies of 4x5 and 4x6 with 8 junction positions. (A) The sequences of the original J1 4x5 system and the (B) original J1 4x6 systems were left unaltered with exception of the eight junction base positions. In both systems, the three component strands are at a 1:4:4 ratio (S1-red:S2-teal:S3-tan) to create four stacked duplexes with 2bp-sticky ends. One of the duplexes is boxed in gray and a full HJ is outlined in black.

6.2.1 Junctions in the 4x5 System

The 4x5 system was modified to contain all 36 different immobile junction sequences, each of which was screened across the 48 different crystallization buffers. While 27 of the junction variants crystallized readily, 9 failed to produce quality crystals despite several screens and optimization attempts (Figure 6.5). The "fatal" junctions included J2, 4, 11, 12, 13, 17, 18, 27 and 30, and though both J2 and J30 readily crystallized, they diffracted poorly, and were ultimately classified as unsuccessful.

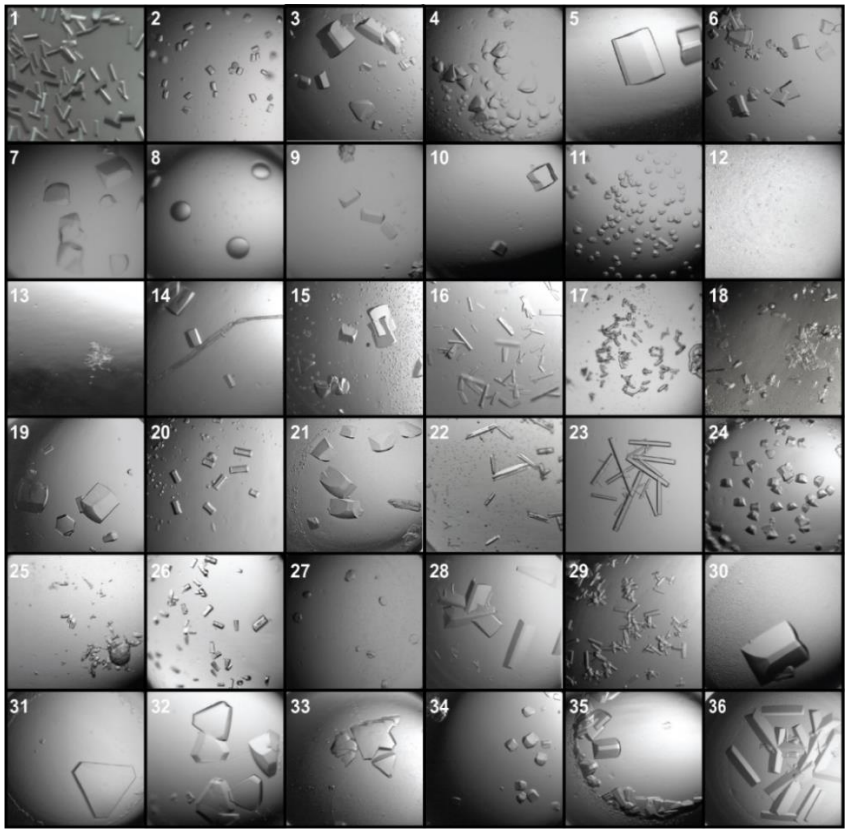


Figure 6.5. All 36 junctions in the 4x5 system crystal images. Representative bright-field images from the crystallization screen of all 36 immobile Holliday junctions in the 4x5 system.

The remaining 27 4x5 junction variants that did crystallize, were all analyzed using X-ray diffraction. As the data were processed it became obvious that the junctions were affecting crystal symmetry. The previously described J1 4x5 crystals diffracted to 3.1Å with $P3_221$ symmetry. However, in 18 of the 36 junctions, the crystals exhibited $P3_2$ symmetry (average cell dimensions a = b = 68.85 Å c = 60.09 Å), which was consistent with the lattice packing in the original J1 4x6 system, and the 9 remaining all retained the original $P3_221$ symmetry (average cell dimensions a = b = 68.17 Å c = 60.60 Å), consistent with J1 4x5 (Table 6.1). The results were surprising because of the fact that in the majority of junctions, the 4x5 system had $P3_2$ symmetry rather than the undesired $P3_221$ symmetry that was originally observed with J1. Further, it was also

reasonable to hypothesize that the junction in fact, could be a determinant that could relieve the torsional strain previously observed, and yield a similar homogeneous lattice to that observed in the 4x6 crystals, potentially as a result of minor angular differences induced by junction sequence. Additionally, 14 of the junctions resulted in higher resolution structures with the best (J19) diffracting to 2.75Å.

Table 6.1. Symmetry and resolution of the 36 immobile junctions in the 4x5 system

Junction	Space Group	Resolution
1	$P3_221$	3.10
2	—	—
3	$P3_2$	3.00
4	—	—
5	$P3_221$	3.15
6	$P3_2$	2.90
7	$P3_221$	3.05
8	$P3_2$	3.10
9	$P3_2$	3.05
10	$P3_221$	3.05
11	—	—
12	—	—
13	—	—
14	$P3_221$	2.85
15	$P3_2$	3.00
16	$P3_2$	3.05
17	—	—
18	—	—

Junction	Space Group	Resolution
19	$P3_221$	2.75
20	$P3_221$	3.10
21	$P3_2$	3.10
22	$P3_2$	3.10
23	$P3_221$	3.05
24	$P3_2$	3.10
25	$P3_2$	3.10
26	$P3_221$	3.10
27	—	—
28	$P3_2$	3.05
29	$P3_2$	3.10
30	—	—
31	$P3_2$	3.15
32	$P3_2$	3.05
33	$P3_2$	3.10
34	$P3_2$	3.00
35	$P3_2$	3.05
36	$P3_2$	3.15

Although the cell parameters between the two symmetries were virtually indistinguishable, as evidenced by superimposing a representative junction structure (J1 vs. J9) from each symmetry (Figure 6.6), the differences between the resulting lattices were drastic (to be discussed later). Although the local geometries of both structures show no apparent differences, it is reasonable to expect that only a slight perturbation in angle can have a significant effect on the overall lattice.

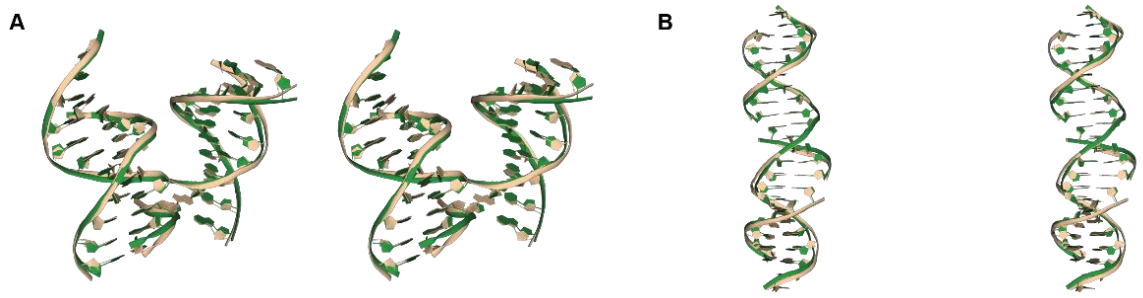


Figure 6.6. Overlay of 4x5 symmetries. Representative junctions for both the $P3_221$ shown in green (J1) and $P3_2$ shown in tan (J9) symmetries were aligned to show the minor discrepancies are shown in stereo with both the (A) junction and (B) duplex models.

6.2.2 Junctions in the 4x6 System

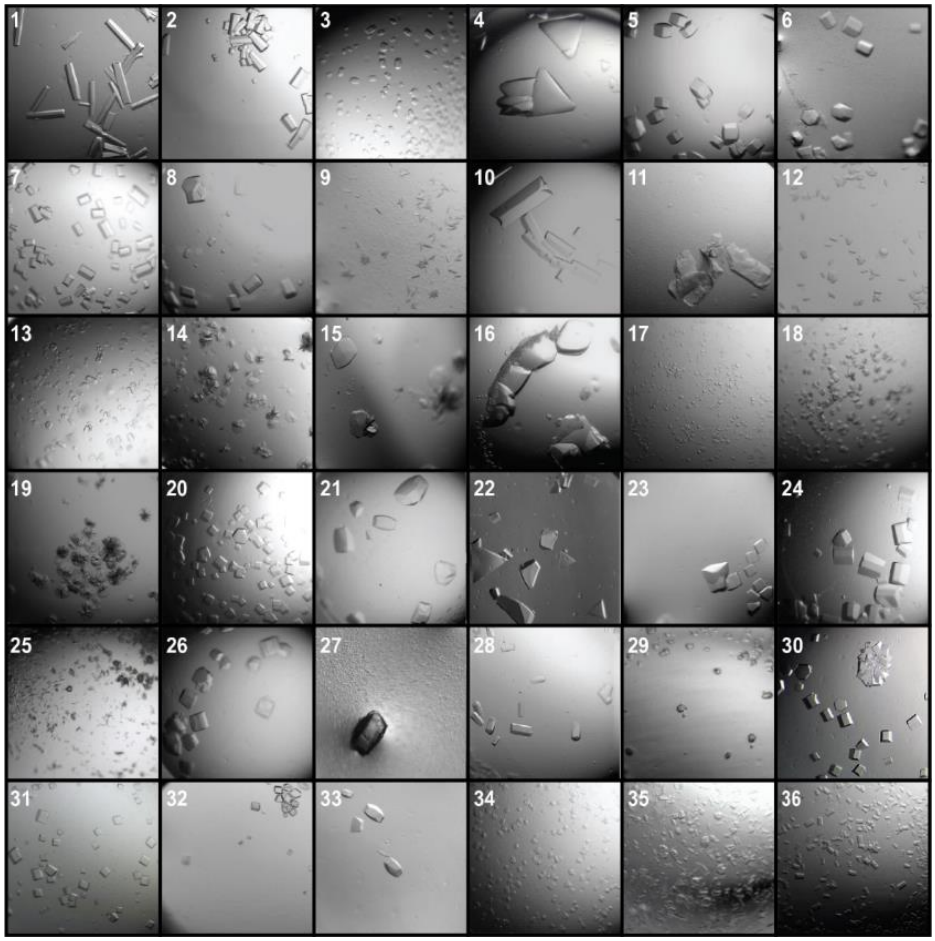


Figure 6.7. Crystal images of the junctions in the 4x6 system. Representative bright-field images from the crystallization screen of all 36 immobile Holliday junctions in the 4x6 system.

In parallel, we employed the previously reported J1 4x6 system using the same strategy to screen for similarities in crystallization efficacy, and provide additional experimental evidence for junctions that remained fatal in both constructs. With only half (18) of the junctions resulting in crystal formation (Figure 6.7), it was clear that the 4x5 system crystallized more robustly. Several of the junctions that were deemed fatal in the 4x5, also did not produce crystals in the 4x6 system including J11, 12, 13, 17, 18, and 27, suggesting that these junctions should be avoided in future designs that utilize the fixed isomer used in this work. In addition to the common fatal junctions, J3, 6, 9, 14,15,19, 21, 25, 29, 32, 34, and 35 also did not form crystals that were amenable to X-ray diffraction. In terms of resolution, we saw no appreciable improvements in diffraction quality (Table 6.2); however, we discovered that in five cases (J4, 5, 31, 33, and 36), the junction appeared to play a role in altering the symmetry from $P3_2$ (average cell dimensions a = b = 68.29 Å c = 55.68 Å) to $R3$ (average cell dimensions a = b = 114.9 Å c = 49.77 Å). The radically different cell constants between symmetries are a testament to the dramatic differences that can be induced sequence specifically by the junction. Of the five crystallizing in $R3$, junctions 4 and 36 crystallized exclusively in $R3$, and 5, 31, 33 exhibited the ability to crystallize in both $P3_2$ and $R3$ (Figure 6.8).

Table 6.2. Symmetry and resolution of the 36 immobile junctions in the 4x6 system

Junction	Space Group	Resolution
1	$P3_2$	3.05
2	$P3_2$	3.15
3	—	—
4	$R3$	3.15
5	$P3_2$	3.15
5	$R3$	3.10
6	—	—
7	$P3_2$	3.10
8	$P3_2$	3.00
9	—	—
10	$P3_2$	3.10
11	—	—
12	—	—
13	—	—
14	—	—
15	—	—
16	$P3_2$	3.20
17	—	—
18	—	—
19	—	—

Junction	Space Group	Resolution
20	$P3_2$	3.10
21	—	—
22	$P3_2$	3.10
23	$P3_2$	3.10
24	$P3_2$	3.10
25	—	—
26	$P3_2$	3.10
27	—	—
28	$P3_2$	3.10
29	—	—
30	$P3_2$	3.15
31	$P3_2$	4.20
31	$R3$	3.15
32	—	—
33	$P3_2$	3.10
33	$R3$	3.15
34	—	—
35	—	—
36	$R3$	3.05

Between the three junctions exhibiting both symmetries, the conditions resulting in $R3$ showed a preference for low concentrations of divalent salts and organic solvents, and $P3_2$ demonstrated a requirement for high salt concentrations and revealed drastically different lattices (to be discussed later). These differences could be the result of a preference for one angle over another due to the presence (or absence of salt). These differences are readily apparent when the structures are superimposed, with the junctions

that adopted the rhombohedral symmetry having significantly larger interduplex angles (Figure 6.9).

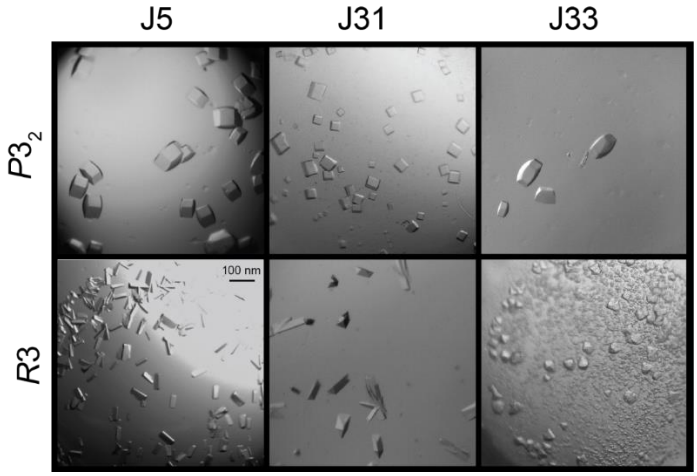


Figure 6.8. Junctions that switch symmetry. Bright field images for the three junctions that were found to have $R3$ and $P3_2$ symmetry in the original sequence 4x6 system

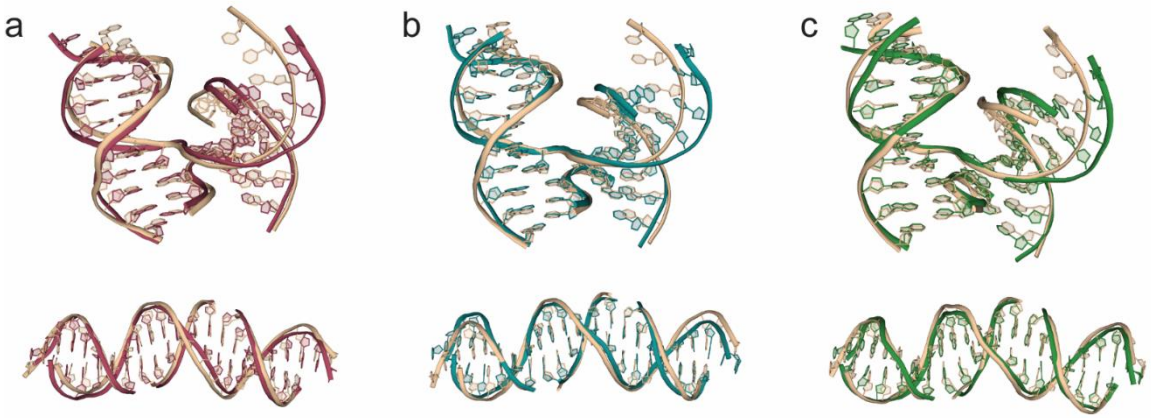


Figure 6.9. Overlay of the three junctions that switch symmetry. The two different symmetries observed in (A) J5 (B) J31 (C) J33 are superimposed with $P3_2$ always being tan and the $R3$ being red, blue, and green.

6.2.3 Junctions in a "Scrambled" 4x6 System

To investigate the potential role the junction flanking (outside the junction) in terms of crystal stability and resolution, and maintenance of symmetry, "scrambled" sequences with targeted base substitutions along each arm were investigated. The overall GC content of the flanking sequences was held constant while the junction remained

unperturbed. The criteria for the changes was simple, if the base was originally a purine, it was switched to pyrimidine as well as changing the base pairing (AT↔GC) (Figure 6.10). The primary motivation for this aspect of the work was to determine if the fatal junctions were strictly due to the junction sequence and not some interplay with the flanking sequences.

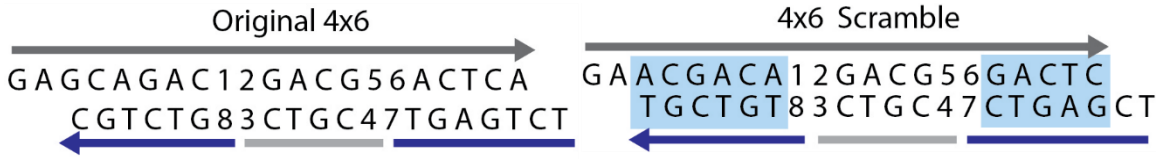


Figure 6.10. Original vs. scrambled 4x6 sequences. The bases that flank each junction were changed from purine to pyrimidine and vice versa, as well as changing the base pair to determine if the fatality of junctions was consistent.

When screening the junction variants of this 4x6 scrambled system across the 48 different buffers, it was found that once again close to half (20) result in crystals (Figure 6.11). There was significant overlap between the fatality of junctions within the original 4x6 and the 4x6 scramble with 13 out of the 16 being consistent and only J4, 20 and 28 proving fatal in the scrambled versions. Contrary to the buffer preference observed with the native $P3_2$ crystals, the scrambles exhibited an exclusive preference towards for low salt buffers, much like the native (unscrambled) sequences that contained *R*3 symmetry. However, of the 20 structures, only J1 and J2 conserved the $P3_2$ symmetry (Table 6.3).

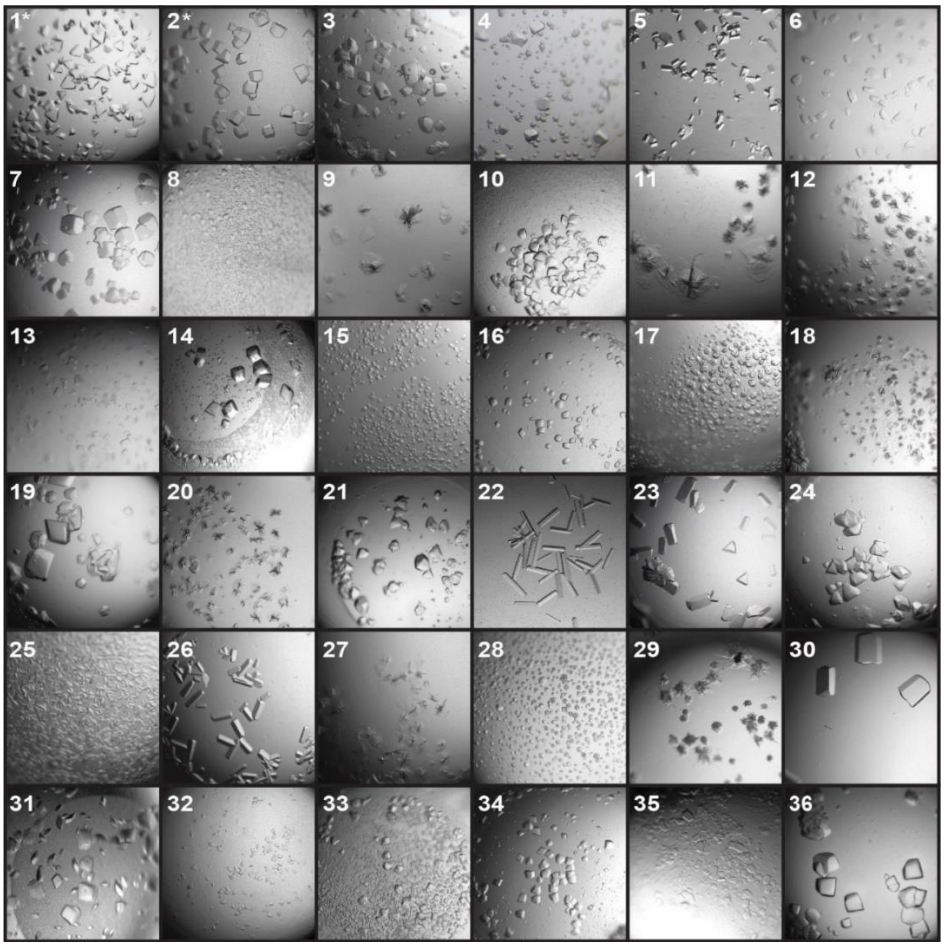


Figure 6.11. Crystal images of the junctions in the 4x6 scramble system. Representative bright-field images from the crystallization screen of all 36 immobile Holliday junctions in the 4x6 scramble system. The two $P3_2$ outliers are highlighted with an asterisk.

In addition, we observed modest improvements in resolution to better than the 3.05 Å maximum resolution with the original sequences, with resulting resolution as high as 2.7 Å in the J36 scramble junction structure. The marked improvements in resolution in the *R*3 scramble crystals appears to be a result of the lattice packing. It also apparent that flanking sequences can play an extraordinary role in crystal packing in concert with the junction. However, it is also clear there there was nearly a consensus overlap in fatality between the scramble and original sequences suggesting that it is not based on flanking sequence. Taking this into consideration it can be assumed that the junction

sequence is the primary factor determining the ability of the motif to crystallize regardless of symmetry.

Table 6.3. Symmetry and resolution of the 36 immobile junctions in the 4x6 scramble system

Junction	Space Group	Resolution	Junction	Space Group	Resolution
1	$P3_2$	3.05	19	*R*3	3.00
2	$P3_2$	3.05	20	—	—
3	*R*3	2.85	21	*R*3	3.05
4	—	—	22	*R*3	3.15
5	*R*3	3.10	23	*R*3	3.00
6	—	—	24	*R*3	2.90
7	*R*3	3.00	25	—	—
8	*R*3	3.05	26	*R*3	3.00
9	—	—	27	—	—
10	*R*3	2.80	28	—	—
11	—	—	29	—	—
12	—	—	30	*R*3	3.10
13	—	—	31	*R*3	2.95
14	*R*3	3.00	32	—	—
15	—	—	33	*R*3	3.10
16	*R*3	3.00	34	*R*3	3.00
17	—	—	35	—	—
18	—	—	36	*R*3	2.70

6.2.4 Summary of Crystal Screening and Fatal Junctions

When examining the results obtained from screening all 36 junctions in the 4x5, 4x6, and 4x6 scrambled systems, several important points are apparent. First, junction sequence absolutely influences 3D DNA constructs in a multitude of ways, and should be considered when designing these sequences. This effect is most obvious when looking at all of the "fatal" junctions in totality, as several were consistently identified in all three crystallization scenarios. Thus, it is reasonable to hypothesize that they have very different angles, which would in turn prevent formation using the isomer used here. The

results discussed above reveal several junctions that should be avoided including J11, 12,13,17,18, and J27 (Table 6.4). Conversely, several junctions performed well across all systems and were able to facilitate crystal formation across multiple symmetries, including J1, 5, 7, 10, 16, 22, 23, 24, 26, 31, 33, and 36. These robust junctions should be considered as a starting point in future designs.

Table 6.4. Fatality of junctions across all systems (indicated in red).

Junction	1	2	3	4	5	6	7	8	9	10	11	12
4x5												
4x6												
4x6 Scramble												

Junction	13	14	15	16	17	18	19	20	21	22	23	24
4x5												
4x6												
4x6 Scramble												

Junction	25	26	27	28	29	30	31	32	33	34	35	36
4x5												
4x6												
4x6 Scramble												

Additionally, it is now indisputable that the J1 sequence is not the sole junction that can facilitate crystallization across all self-assembling 3D designs in structural DNA nanotechnology; in fact several junction sequences have been identified as significantly superior. For every data set that was collected, both a junction and duplex model were solved to study various parameters, resulting in 134 novel models that can be further studied to fully understand how these junction sequences work. The full data collection and refinement statistics can be found in appendix A along with the corresponding PDB code for each solved structure. Many of the specific features and details revealed by the structures will be discussed in the remainder of the chapter.

One of the parameters that was analyzed using a junction model was the interduplex angle to determine what range of angles each junction was capable of adopting within each system and symmetry. The calculated IDA angles from each of the junction structures are reported in Table 6.5 and are broken down by system and symmetry to obtain averages. There was little variation observed in the two symmetries ($P3_221$ and $P3_2$) in the 4x5 system, with the angles both averaging ~56°. However, the $R3$ junctions had a significantly larger angle in both the 4x6 and 4x6 scramble than the corresponding $P3_2$ junctions in the same system with the $P3_221$ having an average angle of 54.6° and 58.05° in the original and scramble and the $R3$ having average angles of 58.37° and 61.00°.

Table 6.5 Calculated IDA across all structures.

4x5			
$P3_221$		$P3_2$	
Junction	Angle	Junction	Angle
1	56.84	3	55.20
5	58.18	6	55.12
7	54.69	8	56.81
10	58.56	9	55.94
14	55.32	15	57.81
19	55.19	16	54.53
20	56.04	21	58.44
23	58.51	22	56.46
26	55.98	24	59.52
56.59±1.50		25	55.61
		28	53.54
		29	57.36
		31	55.62
		32	55.46
		33	56.88
		34	52.89
		35	55.89
		36	55.77
		56.05±1.63	

4x6			
$P3_2$		$R3$	
Junction	Angle	Junction	Angle
1	54.34	4	61.72
2	55.71	5	57.34
5	52.48	31	59.36
7	56.25	33	55.49
8	53.25	36	57.92
10	55.15	58.37±2.33	
16	52.44		
20	52.56		
22	57.7		
23	54.61		
24	54.73		
26	55.35		
28	54.14		
30	54.66		
31	54.53		
33	55.67		
54.60±1.44			

4x6 Scramble			
$P3_2$		$R3$	
Junction	Angle	Junction	Angle
1	57.06	3	60.14
2	59.03	5	58.79
58.05±1.39		7	59.93
		8	61.28
		10	62.89
		14	62.1
		16	61.34
		19	60.24
		21	63.45
		22	60.2
		23	60.83
		24	60.47
		26	62.44
		30	60.95
		31	60.46
		33	61.6
		34	59.41
		36	61.54
		61.00±1.21	

6.3 Lattice Structure and Cavity Sizes

While the local geometry of the junctions all seemed to be similar, when visualizing the entire lattice, it was apparent that only modest differences in junction angle propagates drastically, and dictates the packing of the overall scaffold. Each system was found to crystallize in two distinct symmetries depending on the junction sequence utilized. Within the 4x5 system, both $P3_221$ (original J1 symmetry) and $P3_2$ symmetries were observed. In addition, both the 4x6 (originally $P3_2$ in the J1 version) and 4x6 scrambled systems, revealed both $P3_2$ and $R3$ symmetry. Lattices with $R3$ symmetry were more densely packed, but exhibit similar solvent channels along the six-fold axis as the $P3_2$. The packing is attributed to the layers lacking the planarity observed with $P3_2$ symmetry.

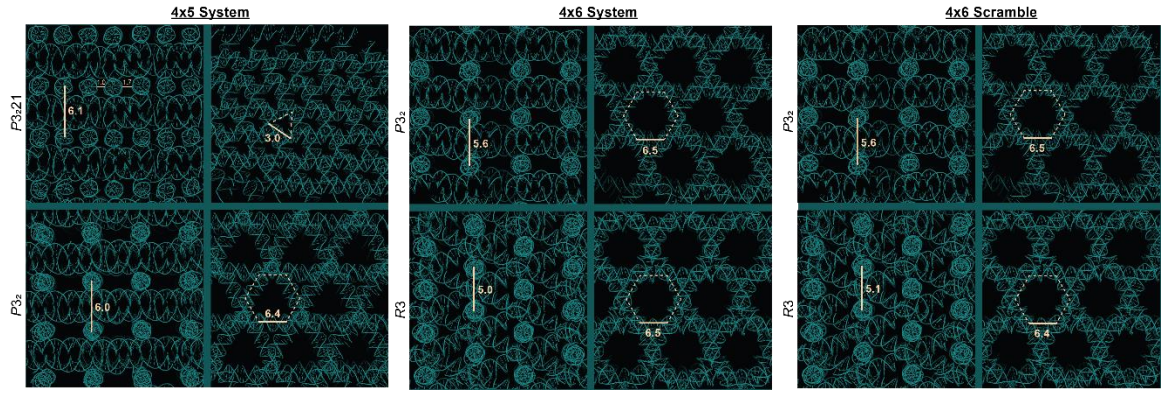


Figure 6.12. Cavity shape and volume. Average distances were calculated for each symmetry observed within each system to determine the volume of the cavities, corresponding to what size guest they could host. The height of each cavity is defined as the distance between three layers (C-axis of cell dimensions) while the edge of the corresponding shape, triangle for $P3_221$ and hexagon for $P3_2$ and $R3$, was calculated in Pymol to find averages.

The resulting lattices can be classified into six unique categories based on both the system and symmetry. Each lattice within each category was then analyzed to determine the average cavity volume. For all categories, the height of the defined cavity was chosen to be the height of three layers or height of the unit cell (Table 6.6) with each edge corresponding to a defined as a geometric shape based on the structure down the 6-fold axis (Figure 6.12). Each system and symmetry were classified as a hexagonal prism with the exception of 4x5- $P3_221$ which was deemed a triangular prism. The height and edge lengths were used to calculate the approximate cavity volumes for each system sand symmetry, and these void spaces which can later be used to determine what size of guest each is amenable to host.

Table 6.6. Cavity dimensions and volumes for each system.

System	4x5		4x6		4x6 Scramble	
Symmetry	$P3_221$	$P3_2$	$P3_2$	$R3$	$P3_2$	$R3$
Edge length (nm)	3	6.4	6.5	6.5	6.5	6.5
Height (nm)	6.1	6	5.6	5	5.6	5
Cavity Volume (nm^3)	24	638	615	549	615	549

6.4 Coordination of Ions

In addition to studying the structural aspects of the junction and duplex models, it became apparent that a variety of ions that were present within the crystallization buffers were coordinated at specific locations and sequences. While this phenomenon had been previously observed,[27] a comprehensive comparison of junction sequence and its ability to coordinate ions was necessary. When grouping each junction model by each system and symmetry and superimposing them, it became apparent that the position of the coordinated ion was highly conserved. Furthermore, the ions consistently coordinated

from junction bases at both crossover strands to create two discrete positions (Pos1 and Pos2) with Pos1 located at the junction of the S3 strand and Pos2 from S1 (Figure 6.13A-F).

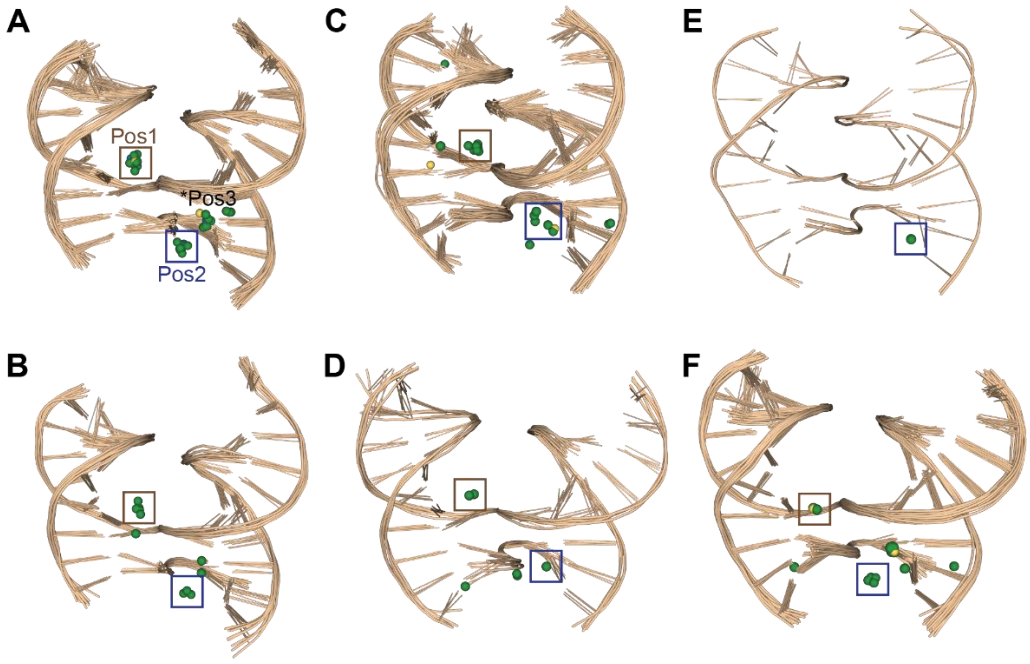


Figure 6.13. Global alignments for ion analysis. Alignments of all models within each given motif and symmetry emphasizing the conserved location of the ions along with a few outliers. (A) 4x5 $P3_221$, (B) 4x5 $P3_2$, (C) 4x6 $P3_2$, (D) 4x6 $R3$, (E) 4x6 scramble $P3_2$, and (F) 4x6 scramble R3. In all panels arsenic is green, cobalt is blue, and magnesium is yellow. Additionally, Pos1 is boxed in brown and Pos2 is boxed in blue.

An analysis of the combined flanking base and junction base at the two ion positions across each individual model highlighted the strong preference to coordinate an ion when a guanine was present (Table 6.7). Furthermore, when both the flanking and junction bases were guanines (GG), an ion was observed in nearly every occurrence. However, since the flanking sequences were maintained across the 36 junctions within each system, there is only a small sample of combinations available to study this effect, and additional studies are needed before a distinct pattern can be identified with

confidence.

Table 6.7. Summary of the ions at each position along with the sequence coordinating it.

4x5					
Junction	Symmetry	Pos1	Ion	Pos 2	Ion
1	$P3_221$	GG	As	TCG	—
5	$P3_221$	GG	As	CCG	—
7	$P3_221$	GG	—	ACG	—
10	$P3_221$	GG	As	GTG	As
14	$P3_221$	CG	—	GGG	—
19	$P3_221$	CG	As	GAG	As
20	$P3_221$	CG	As	CTG	As
23	$P3_221$	GG	As	GTG	—
26	$P3_221$	AG	—	GTG	—
2	$P3_2$	GG	—	CTG	—
3	$P3_2$	CG	As	GAG	As
6	$P3_2$	TG	As	AAG	As
8	$P3_2$	GG	As	GTG	As
9	$P3_2$	CG	As	CAG	As
15	$P3_2$	CG	—	CGG	As
16	$P3_2$	CG	—	CGG	As
21	$P3_2$	GG	As	CAG	As
22	$P3_2$	GG	—	GCG	—
24	$P3_2$	GG	As	GAG	As
25	$P3_2$	CG	Co	CGG	Co
28	$P3_2$	TG	As	GAG	As
29	$P3_2$	TG	As	CAG	As
30	$P3_2$	GG	—	CCG	—
31	$P3_2$	GG	As	GCG	As
32	$P3_2$	AG	As	GTG	As
33	$P3_2$	AG	—	CTG	—
34	$P3_2$	TG	Mg	GAG	Mg
35	$P3_2$	TG	—	CAG	As
36	$P3_2$	CG	As	GTG	As

4x6					
Junction	Symmetry	Pos1	Ion	Pos 2	Ion
1	$P3_2$	GG	As	TCC	As
2	$P3_2$	GG	As	CTC	—
5	$P3_2$	GG	—	CCC	—
7	$P3_2$	GG	—	ACC	—
8	$P3_2$	GG	As	GTC	As
10	$P3_2$	GG	—	GAC	—
16	$P3_2$	CG	—	GGC	—
20	$P3_2$	CG	—	CTC	—
22	$P3_2$	GG	As	GCC	Mg
23	$P3_2$	GG	As	GTC	As
24	$P3_2$	GG	As	GAC	—
26	$P3_2$	GG	As	ATC	As
28	$P3_2$	TG	As	GAC	?
30	$P3_2$	GG	As	CCC	As
31	$P3_2$	GG	—	GCC	—
33	$P3_2$	AG	As	CTC	—
4	$R3$	GG	As	CCC	—
5	$R3$	GG	—	CCC	As
31	$R3$	GG	As	GCC	—
33	$R3$	AG	—	CTC	—
36	$R3$	CG	—	GTC	—

4x6 Scramble					
Junction	Symmetry	Pos1	Ion	Pos 2	Ion
1	$P3_2$	GT	—	TCC	As
2	$P3_2$	GT	—	CGC	—
3	$R3$	CT	As	GAC	As
5	$R3$	GT	As	CCC	As
7	$R3$	GT	—	ACC	—
8	$R3$	GT	—	GTC	As
10	$R3$	GT	—	GAC	As
14	$R3$	CT	—	GGC	As
16	$R3$	CT	As	GGC	As
19	$R3$	CT	—	GAC	As
21	$R3$	CT	—	CAC	—
22	$R3$	GT	—	GCC	—
23	$R3$	GT	—	GTC	Mg
24	$R3$	GT	—	GAC	As
26	$R3$	GT	—	ATC	As
30	$R3$	GT	—	CCC	—
31	$R3$	GT	—	GCC	—
33	$R3$	AT	—	CTC	—
34	$R3$	TT	—	GAC	As
36	$R3$	CT	—	GTC	As

A reasonable explanation for the GG preference can be proposed when looking at the potential coordination sites from each base proximal to the ion. A representative junction model (4x5 J24) that contains a GG coordination site in pos1 and a GA site in pos2 (Figure 6.14A-C) highlights that guanine contains both the O6 and N7 atoms that

could coordinate an ion. This means that a GG sequence has four possible coordination atoms, increasing the chance that an ion could be stabilized at this location, whereas other sequences like GA only have three possible coordinating atoms because the adenine lacks a polar group.

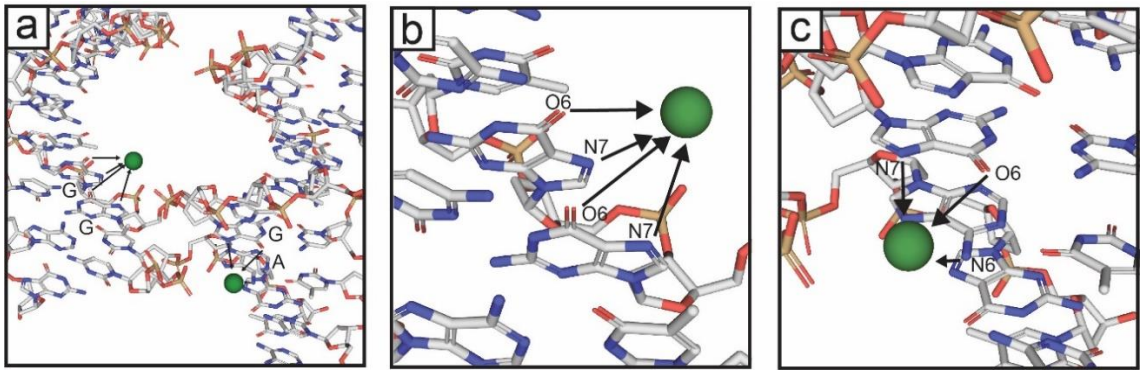


Figure 6.14. Possible candidates for ion coordination. (A) Representative junction with a GG and GA sequence at the junction with the possible contacts that coordinate the arsenic ion. (B) Zoom in of the GG junction with the atom that is within range of coordination labeled and (C) zoom in of the GA junction with the atoms within range of coordination labeled.

To demonstrate that there was no inherent bias for binding of arsenic, which could lead to questions of whether this ion positioning could be a potential artifact, three different ions —arsenic, magnesium, and copper— were all observed in pos1 and pos2, and the ion present was solely dependent on the crystallization buffer (Figure 6.15). The majority of the 48 buffers utilize cacodylic acid ($C_2H_7AsO_2$) as the buffering component, while others contain $MgCl_2$ to help screen electrostatic repulsion between two adjacent DNA backbones. Like magnesium, copper ions have also been shown to stabilize DNA structures[28, 29], however this element is only present in a select few buffers. Nevertheless, both magnesium and copper atoms were observed in the conserved positions in the absence of arsenic in a variety of cases for each system (Figure 6.15A-C).

Figure 6.15. Overlays of models with different ions. Representative junction models for the (A) 4x5, (B) 4x6, and (C) 4x6 scramble to highlight the conserved position of various ions across the systems. In each case, the ions are found within the crossovers and are coordinated to the bases that make up the junction and directly adjacent base, shown in brown. Arsenic ions are shown in green, magnesium is shown in yellow, and cobalt (found only in the 4x5) is shown in blue. The conserved positions, termed Pos1 and Pos2, are labeled in panel a.

6.5 Junction Study Conclusion

All 36 immobile junction sequences were screened across the 4x5 and 4x6 systems to determine if they had any effect on crystal symmetry, resolution, cell dimensions, and ability of the design to crystallize. As an initial metric, the ability to crystallize was studied and immediately demonstrated that the junction sequence is an import aspect to consider when designing a 3D crystal as several of the junctions were deemed "fatal." As a control to ensure that the lack of crystallization observed with these fatal junctions, a scrambled variation of the 4x6 system was also screened as a way to determine that the junction was the determining factor, and not the sequence that flanks it. All the crystals formed were then analyzed using X-ray diffraction, to obtain both the symmetry and resolution, and to ultimately reveal the entire crystal structure. The calculated IDAs revealed from each of the crystal structures will be compared to angles that are being determined using molecular dynamic simulations. Additionally, the ion binding sites are also being explored through MD simulations to compare the

experimentally obtained results that a GG sequence within the junction will site specifically coordinate the ion.

Overall, the 4x5 system was more robust than the 4x6 since only 7 junctions did not crystallize versus the 18 in the 4x6. Both a $P3_221$ and $P3_2$ symmetry was observed in the 4x5 system, with the majority being $P3_2$. This surprising result demonstrates that torsional strain (originally seen in the J1 4x5) in the central weaving strand can be relieved by switching to an appropriate junction sequence. The 4x6 system also exhibited two different symmetries, $P3_2$ (like the original J1 4x6) and a novel rhombohedral symmetry ($R3$). Out of the five 4x6 junctions that displayed $R3$ symmetry, three of them (J5, J31, and J33) could also adopt $P3_2$ symmetry in different crystals, indicating that junction angles have some particular sensitivity to ion composition in the buffer to exhibit one angle preference or the other. In addition to the original 4x6 system, a variation with scrambled flanking sequences was also screened. Unlike the original sequences, the vast majority of this scrambled version was found to crystallize with rhombohedral symmetry, suggesting it that flanking sequence can be an important factor that dictates the overall packing, but does not determine the ability of the motif to crystallize. All of these results are summarized in Figure 6.16.

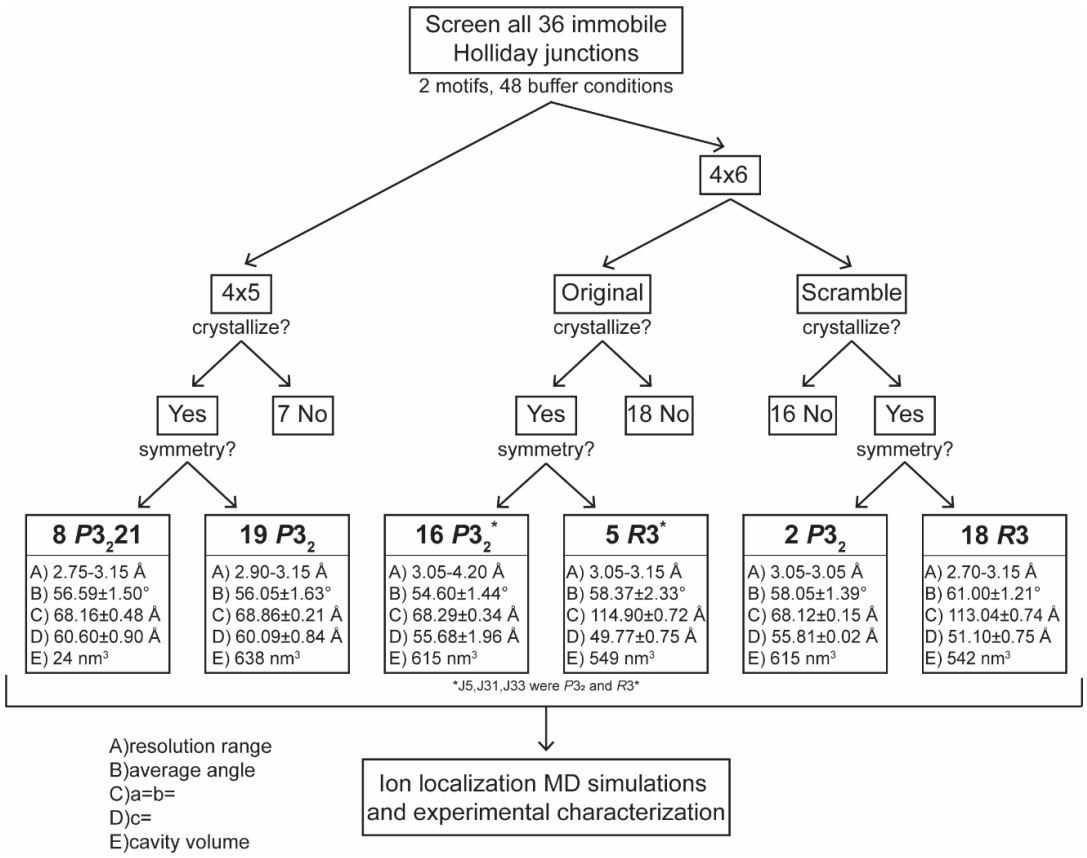


Figure 6.16. Flow chart summarizing the results of the junction study. All 36 immobile junctions were screened across 48 buffers to determine if they would crystallize, after which they were subjected to X-ray diffraction to obtain the symmetry of the lattice. The values that were obtained or calculated for the resolution range, average IDA, unit cell dimensions, and cavity volumes, are listed for each category of system and symmetry.

The results from this comprehensive study of the 36 immobile Holliday junction sequences across two different crystal systems can, in principle, be extended to any 3D DNA structure. Our work definitively demonstrated that the J1 junction should no longer be considered the "foundational junction" for structural DNA nanotechnology. While it crystallized in all cases, it led to an *undesired* symmetry within the 4x5 system, and alternative junctions lead to higher resolution datasets in a multitude of cases. Additionally, it showed that several junctions, including J11,12,13,17,18, and 27, are

"fatal" with the isomer assigned in this work, and should definitively be avoided in future designs. The results further allowed for new approaches for controlling crystal symmetry, lattice architecture, and routes toward improving resolution. The angles obtained here across various symmetries should also serve as an important guide towards the design of future crystal systems. Also, the work uncovered several additional parameters that must been examined in much greater detail in future experiments, including the correlation between flanking and junction sequence, and how it in turn controls both the symmetry and the ability to coordinate ions. All of these factors taken together provide a powerful new route for building lattices capable of hosting proteins for structural determination because the modular cavities need to be large enough to cater to the size of the guest, and the higher resolutions achieved here will aid in revealing essential atomic detail.

6.6 References

1. Holliday, R., A mechanism for gene conversion in fungi (Reprinted). *Genetics Research* 2007, *89* (5-6), 285-307.

2. Duckett, D. R.; Murchie, A. I. H.; Diekmann, S.; Vonkitzing, E.; Kemper, B.; Lilley, D. M. J., THE STRUCTURE OF THE HOLLIDAY JUNCTION, AND ITS RESOLUTION. *Cell* 1988, *55* (1), 79-89.

3. Clegg, R. M.; Murchie, A. I. H.; Zechel, A.; Carlberg, C.; Diekmann, S.; Lilley, D. M. J., FLUORESCENCE RESONANCE ENERGY-TRANSFER ANALYSIS OF THE STRUCTURE OF THE 4-WAY DNA JUNCTION. *Biochemistry* 1992, *31* (20), 4846-4856.

4. Zhang, X. W.; Lee, W.; Fan, X. D., Bio-switchable optofluidic lasers based on DNA Holliday junctions. *Lab on a Chip* 2012, *12* (19), 3673-3675.

5. Karymov, M. A.; Chinnaraj, M.; Bogdanov, A.; Srinivasan, A. R.; Zheng, G. H.; Olson, W. K.; Lyubchenko, Y. L., Structure, Dynamics, and Branch Migration of a DNA Holliday Junction: A Single-Molecule Fluorescence and Modeling Study. *Biophysical Journal* 2008, *95* (9), 4372-4383.

6. Azaro, M. A.; Landy, A., The isomeric preference of Holliday junctions influences resolution bias by lambda integrase. *Embo Journal* 1997, *16* (12), 3744-3755.

7. Shida, T.; Iwasaki, H.; Shinagawa, H.; Kyogoku, Y., Characterization and comparison of synthetic immobile and mobile Holliday junctions. *Journal of Biochemistry* 1996, *119* (4), 653-658.

8. Yadav, R. K.; Yadava, U., Molecular dynamics simulation of hydrated d(CGGGTACCCG)(4) as a four-way DNA Holliday junction and comparison with the crystallographic structure. *Molecular Simulation* 2016, *42* (1), 25-30.

9. Wang, W. J.; Nocka, L. M.; Wiemann, B. Z.; Hinckley, D. M.; Mukerji, I.; Starr, F. W., Holliday Junction Thermodynamics and Structure: Coarse-Grained Simulations and Experiments. *Scientific Reports* 2016, *6*.

10. Hyeon, C.; Lee, J.; Yoon, J.; Hohng, S.; Thirumalai, D., Hidden complexity in the isomerization dynamics of Holliday junctions. *Nature Chemistry* 2012, *4* (11), 907-914.

11. Yu, J.; Ha, T. J.; Schulten, K., Conformational model of the Holliday junction transition deduced from molecular dynamics simulations. *Nucleic Acids Research* 2004, *32* (22), 6683-6695.

12. Overmars, F. J. J.; Lanzotti, V.; Galeone, A.; Pepe, A.; Mayol, L.; Pikkemaat, J. A.; Altona, C., Design and NMR study of an immobile DNA four-way junction containing 38 nucleotides. *European Journal of Biochemistry* 1997, *249* (2), 576-583.

13. Ho, P. S.; Eichman, B. F., The crystal structures of DNA Holliday junctions. *Current Opinion in Structural Biology* 2001, *11* (3), 302-308.

14. Hays, F. A.; Watson, J.; Ho, P. S., Caution! DNA crossing: crystal structures of Holliday junctions. *Journal of Biological Chemistry* 2003, *278* (50), 49663-49666.

15. Hays, F. A.; Vargason, J. M.; Ho, P. S., Effect of sequence on the conformation of DNA Holliday junctions. *Biochemistry* 2003, *42* (32), 9586-9597.

16. Eichman, B. F.; Ortiz-Lombardia, M.; Aymami, J.; Coll, M.; Ho, P. S., The inherent properties of DNA four-way junctions: Comparing the crystal structures of Holliday junctions. *Journal of Molecular Biology* 2002, *320* (5), 1037-1051.

17. Hays, F. A.; Teegarden, A.; Jones, Z. J. R.; Harms, M.; Raup, D.; Watson, J.; Cavaliere, E.; Ho, P. S., How sequence defines structure: A crystallographic map of DNA structure and conformation. *Proceedings of the National Academy of Sciences of the United States of America* 2005, *102* (20), 7157-7162.

18. Khuu, P. A.; Voth, A. R.; Hays, F. A.; Ho, P. S., The stacked-X DNA Holliday junction and protein recognition. *Journal of Molecular Recognition* 2006, *19* (3), 234-242.

19. Ho, P. S., Structure of the Holliday junction: applications beyond recombination. *Biochemical Society Transactions* 2017, *45*, 1149-1158.

20. Watson, J.; Hays, F. A.; Ho, P. S., Definitions and analysis of DNA Holliday junction geometry. *Nucleic Acids Research* 2004, *32* (10), 3017-3027.

21. Zheng, J. P.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R. J.; Constantinou, P. E.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* 2009, *461* (7260), 74-77.

22. Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* 2016, *138* (38), 12690-12690.

23. Simmons, C. R.; Zhang, F.; MacCulloch, T.; Fahmi, N.; Stephanopoulos, N.; Liu, Y.; Seeman, N. C.; Yan, H., Tuning the Cavity Size and Chirality of Self-Assembling 3D DNA Crystals. *Journal of the American Chemical Society* 2017, *139* (32), 11254-11260.

24. Zhang, F.; Simmons, C. R.; Gates, J.; Liu, Y.; Yan, H., Self-Assembly of a 3D DNA Crystal Structure with Rationally Designed Six-Fold Symmetry. *Angewandte Chemie-International Edition* 2018, *57* (38), 12504-12507.

25. Carlstrom, G.; Chazin, W. J., Sequence dependence and direct measurement of crossover isomer distribution in model holliday junctions using NMR spectroscopy. *Biochemistry* 1996, *35* (11), 3534-3544.

26. McKinney, S. A.; Declais, A. C.; Lilley, D. M. J.; Ha, T., Structural dynamics of individual Holliday junctions. *Nature Structural Biology* 2003, *10* (2), 93-97.

27. van Buuren, B. N. M.; Hermann, T.; Wijmenga, S. S.; Westhof, E., Brownian-dynamics simulations of metal-ion binding to four-way junctions. *Nucleic Acids Research* 2002, *30* (2), 507-514.

28. M.Govindaraju; Shekar, H. S.; S.B.Sateesha; Raju, P. V.; Rao, K. R. S.; Rao, K. S. J.; A.J.Rajamma, Copper interactions with DNA of chromatin and its role in neurodegenerative disorders. *Journal of pharmaceutical analysis* 2013, (5), 354-359.

29. Paris, C.; Geinguenaud, F.; Gouyette, C.; Liquier, J.; Lacoste, J., Mechanism of Copper Mediated Triple Helix Formation at Neutral pH in Drosophila Satellite Repeats. *Biophysical journal* 2007, *92* (7), 2498-2506.

CHAPTER 7

A SELF-ASSEMBLED RHOMBOHEDRAL DNA CRYSTAL SCAFFOLD WITH TUNABLE CAVITY SIZES AND HIGH-RESOLUTION STRUCTURAL DETAIL

7.1 Application of the Systematic Junction Study

The original goal of the Yan laboratory was to modify Seeman's tensegrity triangle design to create a tensegrity square. This system was designed to expand the central strand comprising the triangle to contain 4 repeats of 7 bases (4x7), thus adding a fourth edge to the motif. The initial design showed promise, with rapid annealing of the component strands leading to well-ordered crystals in a 2D array with apparent square cavities, as evidenced by atomic force microscopy (AFM) (Figure 7.1).[2] While crystals for the tensegrity square (4x7) were eventually obtained, they exhibited an amorphous morphology, and appeared poly-crystalline leading to powder-like diffraction patterns that were not able to be indexed (Figure 7.2A). While the crystals did not perform well, decreasing the number of repeating bases in the central strand from seven to five (4x5) resulted in crystals readily which diffracted to 3.1 Å.[2]

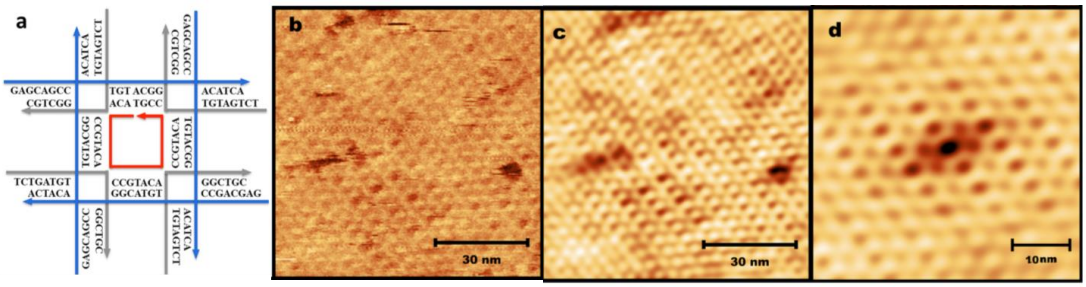


Figure 7.1. J1 Tensegrity Square. (A) The original design of the tensegrity square with the J1 Holliday Junction Sequences and (B) AFM image of the lattice alongside the (C,D) Fourier flattened images where the cavities show the desired square shape.[2]

The year following the publication of the 4x5 system, an additional base was added to each repeat of the central strand to create the 4x6 system, as discussed in

Chapter 5. The 4x6 system was in turn followed by the six-fold symmetric crystal design,[3] at which point all reported DNA crystal systems utilized the "J1" junction sequence. As outlined in Chapter 6, a comprehensive study of the 36 different immobile junction sequences was carried out in both the 4x5 and 4x6 systems, and the initial results from the study revealed that the original J1 was not necessarily the ideal junction that it was once thought to be. Additionally, it became apparent that altering the junction sequence could not only lead to better resolution than those obtained using J1, but could affect how robustly the sequences crystallized, demonstrating that the junction should be seriously considered when designing a 3D DNA crystal. Early results highlighted that J10 allowed for robust crystallization across both systems, due to its ability to crystallize in the majority of the 48 buffers screened, while also diffracting to higher resolution. J10 was also able to crystallize in a variety of space groups ($P3_221$, $P3_2$, and $R3$). At this point, the original J1 4x7 design was revisited modified with the J10 sequence, which immediately led to well defined crystals that diffracted to 2.7 Å (PDB code:6U40), which is higher than any other reported structure.

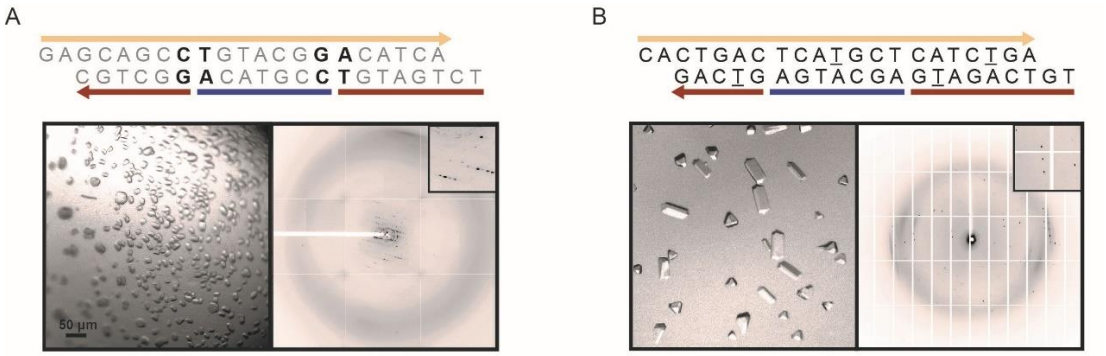


Figure 7.2. Junction variations of the 4x7. Sequences, bright-field images, and diffraction patterns for the (A) J1, which indicate "powder-like" diffraction and were not indexable and (B) J10 4x7 with R3 symmetry and a resolution of 2.7 Å. Representative Bragg reflections are inset.

This chapter will discuss the overall design and packing of the novel 4x7 system containing unique rhombohedral (*R*3) symmetry. With the unprecedented resolution, additional details in these self-assembling crystals were revealed, rendering it an ideal system for exploring further design parameters such as: (1) the effect of both length and sequence of the sticky ends; (2) determining if an observed polar contact between layers was responsible for the dense packing and resulting symmetry; and (3) altering the cavity sizes through expansion of the contents of the asymmetric unit by one helical turn.

7.2 SEQUENCE DESIGN AND CRYSTALLIZATION

Like the previous symmetric 4xN systems, the 4x7 design is comprised of three component strands: the central weaving strand (S1; blue) that contains four identical repeats of seven bases, four copies of the linear strand (S2; tan) as well as four copies of the crossover strands (S3; brown). Based on the design, the tethering of every four 2-turn duplexes (each tailed by 2-bp sticky ends) was mediated solely by the central 4x7 strand.

These building blocks can then assemble via sticky-end cohesion that is dictated by the 2bp (CA/GT) overhangs on the 5' end of the S2 and S3 strands (Figure 7.3 A).

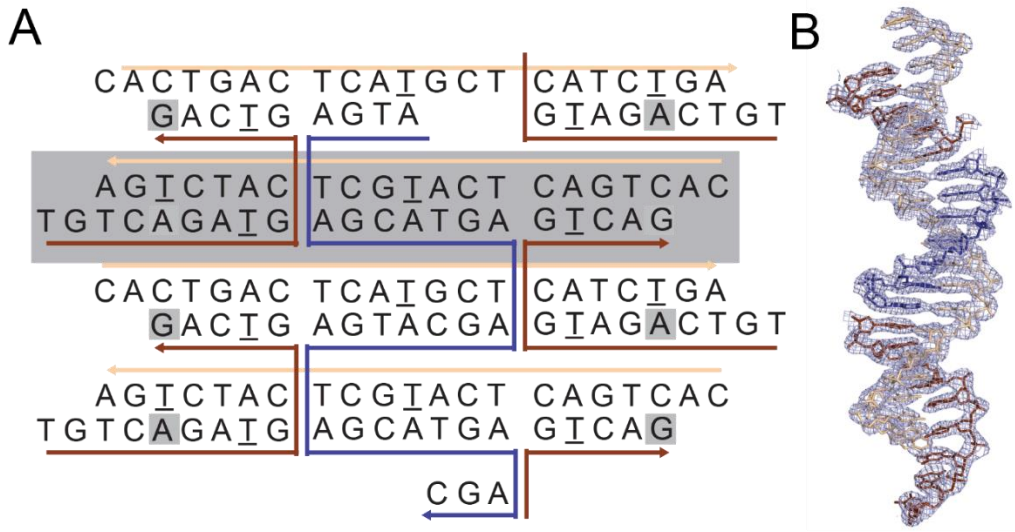


Figure 7.3. Design and structure of the 2-turn model. (A) 2D topology demonstrating the assembly of the three component strands: S1-blue, S2-tan, and S3-brown. Positions of the bromines used to determine the initial maps are underlined and one ASU is outlined in a gray rectangle. (B)The asymmetric unit (ASU) with the helical contents of the 4x7 motif contained in electron density at 2.7 Å resolution. All main features of the design are readily observable in density. The $2F_o - F_c$ map is contoured at $\sigma = 1.5$

Unlike the previous systems, the 4x7 crystals had dramatically different rhombohedral symmetry (*R*3). This novel symmetry meant that a molecular replacement solution could not be obtained; as a result, the design necessitated using a heavy atom bromine derivative located at the C5 position on specified thymines (two per turn; 4 per duplex) in order to obtain initial electron density maps. The density that resulted from each bromine site was readily apparent in the maps (Figure 7.4) and served as the substructure used for model building (PDB code:6V6R). The base stacking between adjacent base pairs was readily observable in the maps and yielded additional details that were attributable to the higher resolution in both the native (Figure 7.3B) and derivative structures.

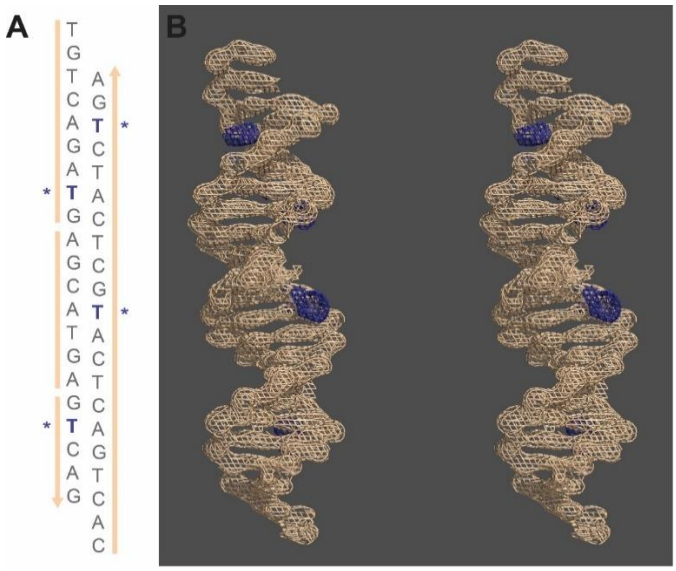


Figure 7.4. Structure of the bromine derivatized crystals. (A) Sequence for the contents of the ASU with the four bromo-dU positions highlighted in blue with an asterisk; (B) Stereoview of the anomalous difference maps for the bromine derivatized structure with $2F_o - F_c$ density (tan) contoured at σ=1.8 and with $F_o - F_c$ density (blue) contoured with σ=3.0 for each of the corresponding bromine positions. Note: The third and fourth peaks in the model can be visualized on the back side of the duplex.

7.3 PACKING OF THE OVERALL STRUCTURE

While the overall packing followed the same basic design principles as the previously reported 4xN systems (i.e. a central stand tethering four stacked duplexes that assemble in 3D space via sticky ends), the 4x7 system central building block yielded layers that did not remain in plane with one another. While layers 1 and 4 remained parallel, layers 2 and 3 were tilted. This phenomenon is displayed in Figure 7.5 with each layer containing three duplexes extending through the sticky ends to make the orientations of each layer obvious.

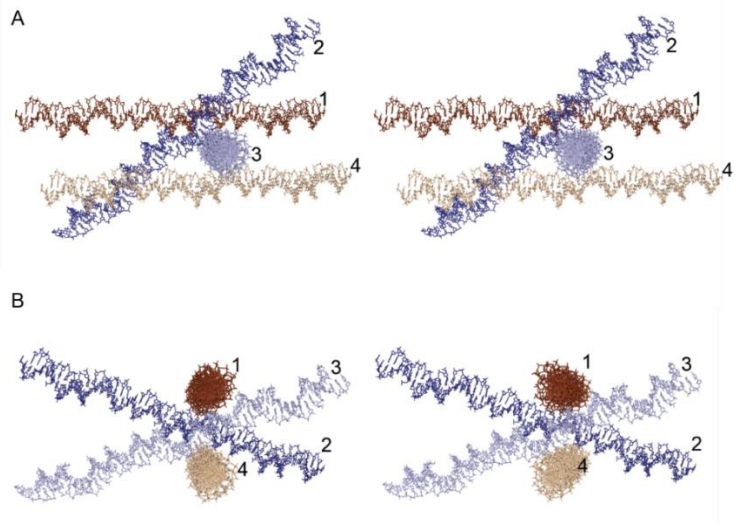


Figure 7.5. Central building block of the 4x7. Two views (A, B), rotated 90° from each other, shown in stereo view to demonstrate the non-parallel packing of layers 2 and 3 of the central block. To make the effect more apparent, each layer displays 3 duplexes.

The angles between alternating layers were calculated to be ~65˚, with a distance of 4.5 nm between the "top" (layer 1) to "bottom" (layer 4), packing above and below junction points between crisscrossing layers 2 and 3 (Figure 7.6A). Due to the angled duplexes, a close polar contact was observed between layers 1 and 3, and 2 and 4. This interaction between alternating layers occurred between the 3'-hydroxyl from the guanine (G14) in layer 1 and two phosphate backbone oxygens from adenine (A5) in layer 3, with each base belonging to strand S3 of the system. This interaction contained both a strong and weak electrostatic contact with distances of 2.45 Å and 3.8 Å, respectively (Figure 7.5 B). While these H-bond distances were found to be on either side of the documented range, the strong and weak polar interactions are not without precedent.[5, 6]

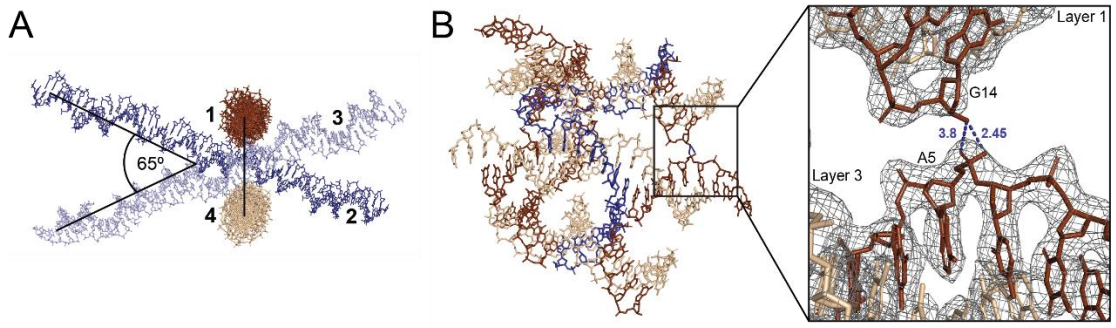


Figure 7.6. Angled layers and polar contacts of the 4x7. (A) The angle between the non-parallel layers (2 and 3) was calculated to be 65° along with a height of 4.5 nm between the ASU. (B) A polar contact formed between alternating layers between A5 and G14 was found with distances of 3.8 and 2.45 Å.

In order to determine if this polar contact was responsible for the R3 packing, or simply an artifact of that particular DNA sequence, a "mutation" at that site was made (G14→T, and A5→C). The resulting dataset diffracted to 2.95 Å (PDB code:6UEF) and revealed a similar interaction between atoms with distances of 2.5 Å and 4.3 Å (a length no longer capable of forming a hydrogen bond), but with only the single short contact. The overall packing was unperturbed, despite one of the contacts no longer being present, suggesting that the polar contact was not contributory to the rhombohedral symmetry. Additionally, this contact did not exist when three turns per duplex were used, which will be discussed in further detail later.

7.4 ROLE OF STICKY END SEQUENCE AND LENGTH

In addition to the native and polar contact mutation designs, the role of both sticky end (SE) length and sequence was also investigated to determine if they affect the resolution, symmetry, and the ability to crystallize. Sticky end lengths ranging from 1 to 3 base pairs of varying A-T and G-C content were considered, all of which resulted in crystals (Figure

7.7). Each of these sticky end modifications was named based on the number of bases that participate, followed by the sequence at the 5' end of the S2 strand.

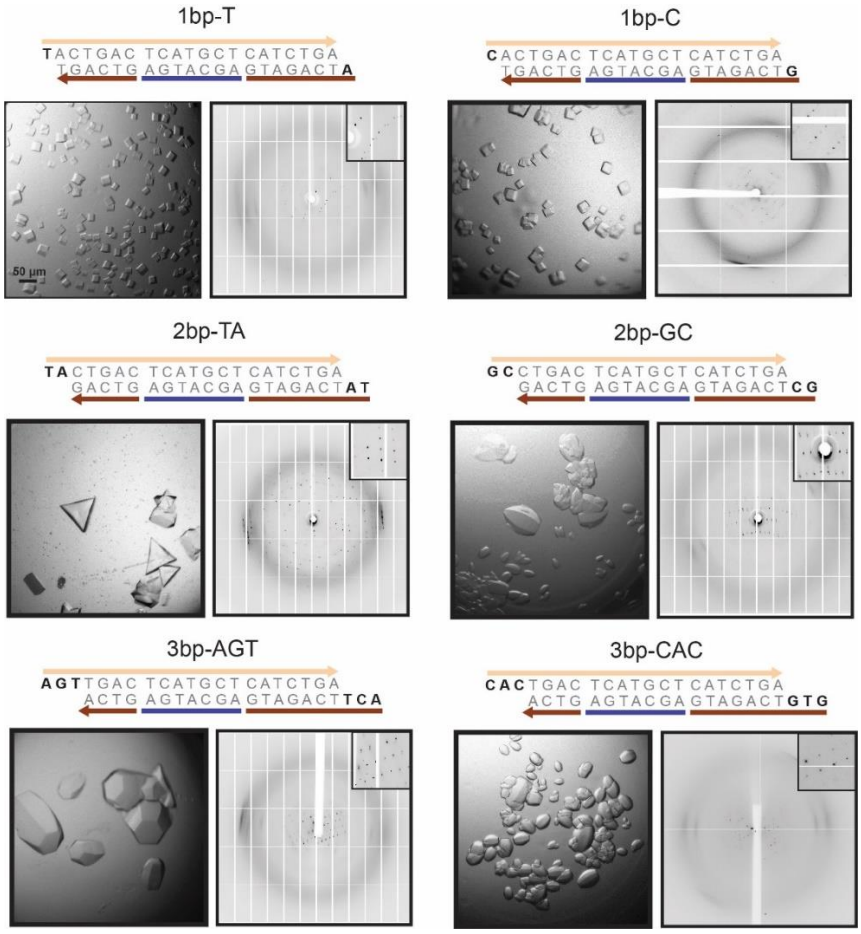


Figure 7.7. Sticky end variants of the 4x7 system. Each variant is shown with its sequence, representative bright field images, and diffraction pattern. Each variation's unique SE sequence is shown in bold font with all bright field images at the same scale, and a representative section of the diffraction pattern inset.

Seeman recently reported that altering the sticky end sequences in the tensegrity triangle led to a modest improvement in resolution, but had no effect on the overall symmetry in nearly all cases.[7] Unlike what was observed in the tensegrity triangle, however, only one (2bp-TA) out of the six 4x7 SE variants retained the native *R*3 symmetry. Additionally, only the 2bp-TA modification (PDB code:6UDN) led to a higher resolution (2.6 Å) compared to the native 2bp-CA. It should also be mentioned that the

TA variant exhibited a comparable single polar interaction (2.8 Å), between the same atoms across layers, as the native 2-turn CA design. Both 1bp and 3bp variants as well as the 2bp-GC all diffracted to lower resolution (between 4.0 - 5.0 Å) with representative diffraction images displayed in Figure 7.7. While all the data sets were able to be indexed allowing for space group identification (1bp-T had *P*1 symmetry while the remaining four systems—1bp-C, 2bp-GC, 3bp-CAC and 3bp-AGT—all exhibited *P*6 symmetry), none of them resulted in datasets sufficient to solve a structure. This result suggests that the sticky end sequence is a key design parameter that must be considered for novel DNA crystals in future studies.

7.5 EXPANDING FROM A 2- TURN TO A 3- TURN MOTIF

To date, all the published crystal designs, with the exception of the tensegrity systems, were composed of several 2-turn duplexes tethered by a central strand. With the goal of using the cavities that form between layers of these duplexes to scaffold biomolecules such as proteins, the ability to control the cavity size and volume is vital, therefore, we sought to alter the cavities by adding an extra helical turn (resulting in three total turns) to each stacked duplex. With each full helical turn being comprised of 10.5 bases, three full turns results in a non-integer number (31.5 bp). Because of this, duplexes that were either slightly under- and over-twisted (31 and 32 bases, respectively) each were designed and screened for crystallization. Additionally, the extra turn could be designed either on the flanking region between a junction and the sticky ends (Figure 7.7A, B), or between the two Holliday junctions (Figure 7.7C, D). The original sequence from the 2-turn motif was conserved, including the J10 junction, and the sequence of the

third turn was randomly generated in Tiamat[8] while roughly maintaining the overall GC content (45-48%).

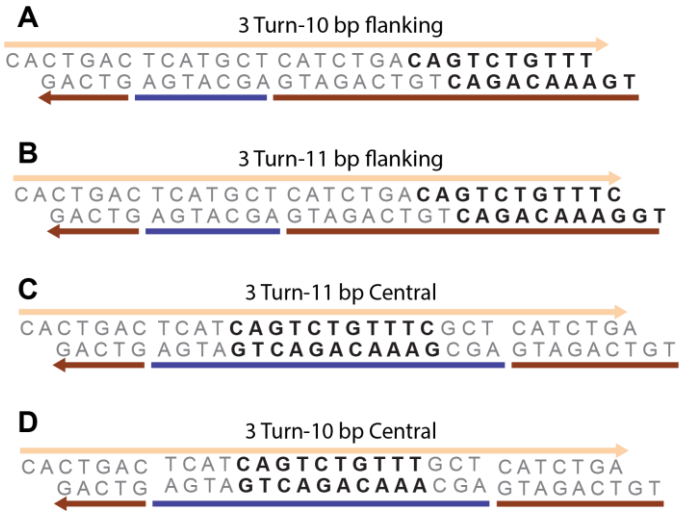


Figure 7.8. Expanding to a 3-turn design. All four variations explored when adding an additional helical turn, the color of component strands is consistent with previous figures (S1-blue, S2-tan, S3-brown). (A) contains 10 additional bases and (B) has 11 additional bases flanking the junctions while (C) has 11 additional bases (D) has 10 additional bases between the two junctions.

Each of the four designs was thoroughly screened and optimized a multitude of times to obtain quality crystals. Both of the flanking versions along with the TA sticky end variation of 11bp flanking crystallized readily and were sent for data collection (Figure 7.9 A-C); however, neither of the central (interjunction) versions crystallized to any appreciable degree for data collection (Figure 7.9 D, E). Unfortunately, the 10bp-flanking crystals diffracted poorly and were unable to be indexed, but the 11bp-flanking maintained the *R*3 symmetry and diffracted to 4.5 Å (PDB code: 6UAL), which is considerably lower resolution than the 2-turn system. However, when compared to the 3-turn tensegrity triangle resolution (6 Å) reported by Seeman[1], the 11bp-flanking 3-turn system was far superior.

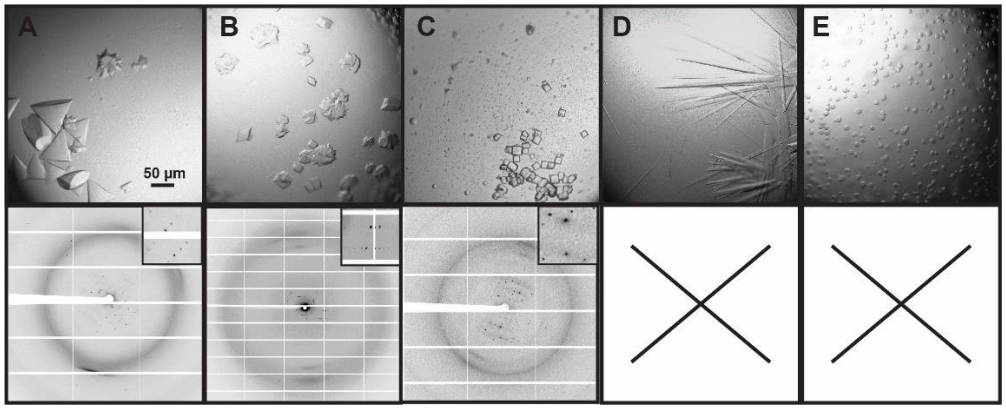


Figure 7.9. 3-turn variation images and corresponding diffraction frames. All of the 3-turn variations are shown including (A) flanking-11bp, (B) flanking 10-bp, (C) TA sticky end variation of 11bp flanking, (D) central-11bp, and (E) central-10 bp.

Since the 3-turn design maintained the same symmetry as the 2-turn system, a derivative dataset was not necessary to determine the overall structure and the 2-turn model was able to be used for molecular replacement before building in the additional helical turn. The overall packing was very similar to the 2-turn system; however, the atomic detail and base stacking observed in the 2-turn system was no longer apparent due to the lower resolution (Figure 7.10A). The lower resolution is most likely attributable to the higher solvent content resulting from the larger cavities.

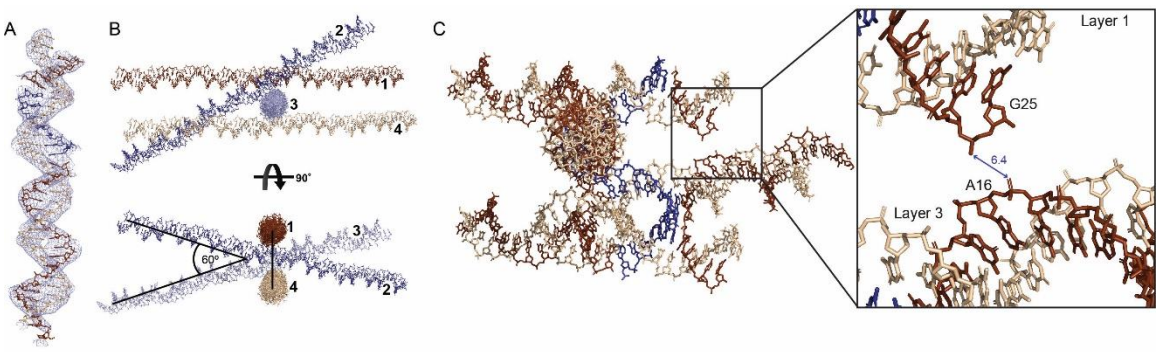


Figure 7.10. Designing larger solvent cavities by expansion of the unit cell. (A) The 3- turn model built into its electron density (B) Each of the 4 component layers (numbered 1-4) two symmetry related duplexes are connected via 2-base sticky ends are shown. Angles between respective layers were calculated as ~60˚ (C) The closest distance observed in this structure was 6.4 Å, and clearly too long to form any hydrogen bond. However, the bases involved here are consistent with those shown to have a polar contact in the 2-turn system.

Consistent with the 2-turn model, the structure yielded stacks of helices scaffolded by the central strand with the extra 11 bases on alternating sides. Unlike the 2-turn system, however, in the 3-turn crystal there was no indication of a tight interaction between layers, with the closest atoms approaching one another at a distance of 6.4 Å between the G25 and A16 at the termini of alternating layers. Moreover, even though the 2-bp SE modification (CA to TA) improved the resolution in the native system, the same modification applied to the 3-turn system resulted in poorer diffraction.

7.6 ANALYSIS OF CAVITY SIZE IN THE 2- AND 3- TURN SYSTEMS

A direct comparison of the 2- and 3-turn systems shows that the extra helical turn requires the overall packing to adapt slightly to compensate for the additional 11 bases. When looking at the central building block, two notable parameters changed: (1) the distance between the top and bottom layers, which was 45 Å in the 2-turn model, was slightly expanded to 48 Å in the 3-turn system (Figure 7.11); and (2) the angles between each layer of the 3-turn system were 60˚, unlike the ~65˚ angle in the 2-turn structure.

These angles were calculated by drawing vectors from crossover points to the end of each neighboring duplex roughly along the helix axis. While neither of the differences was significant enough to alter the overall packing of the lattice, they do demonstrate that the addition of extra helical turns can alter local geometry, or the model could be less accurate due to the lower resolution. Thus, these alterations should be taken into consideration in the future when deciding if the system is amenable to a particular application.

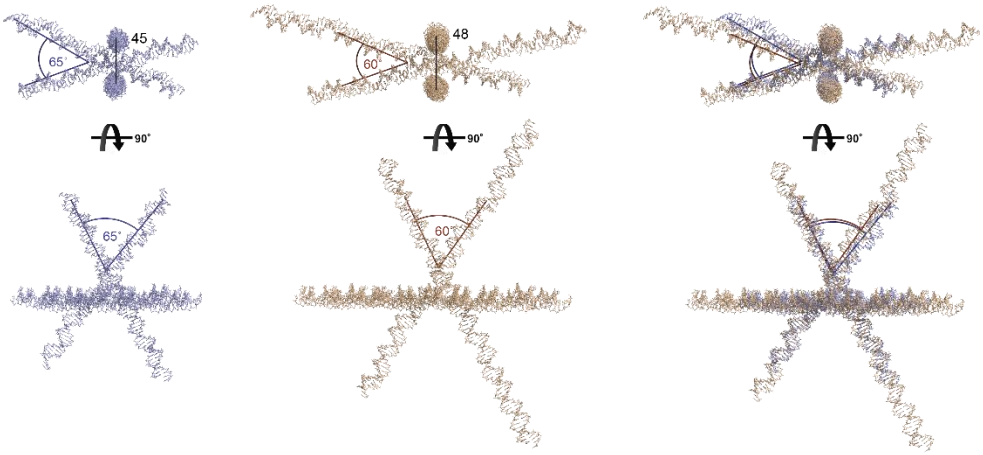


Figure 7.11. 2-turn and 3-turn block unit overlay. Direct comparison of the central building block of the 2-turn (light blue) and the 3-turn (tan) systems, highlighting the difference in angle between layers and height of the layers.

With the ultimate goal of using these crystals to scaffold guests, the size and shape of the cavities in both the 2- and 3-turn systems was analyzed to determine how suitable they would be as hosts for various biomolecules. The 2-turn crystal yielded an array of cavities that were ~4.7 nm along the cross-section, and ~6.6 nm along each edge of the hexagonal solvent channel (Figure 7.12 A) while the 3-turn cavities edges measured ~10 nm, with a cross-section diameter of 8 nm (Figure 7.12 B). Each individual cavity is also defined as having the height of four duplexes stacked, which can

be viewed at a 90˚ rotation of the six-fold axis. In the 2-turn system, this distance was ~4.5 nm and ~4.8 nm in the 3-turn system. It is possible that these differences could be attributed to the disorder inherent to the lower resolution (4.5 Å) of the crystal, along with the additional flexibility of each duplex due to the third helical turn.

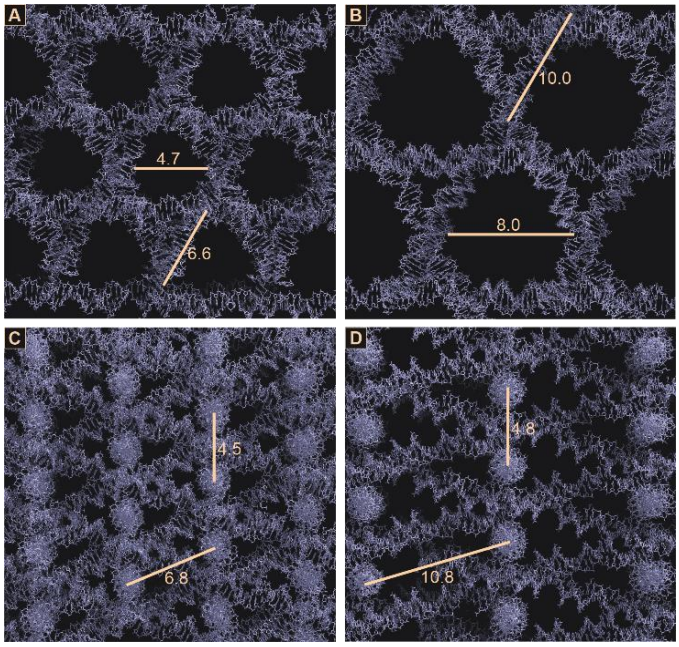


Figure 7.12. Tuning the sizes of the cavities within the crystal scaffolds of the 2- and 3-turn 4x7 systems. (A) and (B) Views of the crystal packing along the six-fold symmetry axis for the 2-turn and 3-turn systems, respectively. Distances measured across the center of each cavity and along each edge demonstrate the sizes of the periodic cavities in the array and are indicated by tan scale bars. As dictated by the design, the expanded cavities in the 3-turn show size increases that correspond almost exactly to the expected length of the additional a helical turn in the motif (3.4 nm). (C) and (D) Views oriented via 90-degree rotation along the six-fold symmetry axis for the 2-turn and 3-turn systems, respectively. The cavities can be envisioned as a hexagonal container where the volume of the cylinder is calculated down the six-fold axis through 3 layers (4.5 and 4.8 nm) with the edges and diameter used as shown. All lengths are measured in nanometers.

For both constructs, the cavity volumes were calculated using a hexagonal prism with distances that correspond to the edges along the six-fold axis, and the heights corresponding to the distances between layers 1 and 4 of each 4-duplex block. Using

these parameters, the calculated cavity volumes in the 2-turn and 3-turn crystals are 500 and 1250 nm^3, respectively. The larger volume of the 3-turn cavities (approximately 2.5 fold) could potentially accommodate guest molecules such as β-glucosidase (76.6 kDa) which has an approximate volume of 1100 nm^3, while the 2-turn system would be more appropriate for smaller guests such as thrombin (34.3 kDa, ~400 nm^3).[9,10] While even the 3-turn system is still not amenable as a host for much larger proteins, the design principles outlined in this work could be extended to other 3D self-assembling systems to provide even larger cavities. . Additionally, the higher resolution achieved with this rhombohedral system is essential to using these crystals scaffolds for structural determination of guest molecules, since additional atomic detail could be observed.

7.7 CONCLUSION

In this work, the J10 sequence that initially crystallized in the junction study was employed in the original 4x7 tensegrity square design that contained J1 and diffracted poorly. Switching to J10 allowed for the structural determination of a novel rhombohedral (*R*3) DNA crystal scaffold that diffracted to an unprecedentedly high resolution. Additionally, the effect of sticky end length and sequence was investigated, leading to the conclusion that choosing the proper sticky ends are vital for proper crystallization and good diffraction. It was found that by mutating the original 2bp-CA to a 2bp-TA sticky end, it led to a modestly improved resolution (2.7 to 2.6 Å), while all other variations resulted in poor crystal quality. Additionally, the cavity size was successfully expanded by ~50% by including an extra 11 bases to the flanking region

outside of the junctions. This novel design provides another option in the ever-growing selection of self-assembling DNA crystal lattices that can be used to host various molecules in the future, while also elucidating important new design rules for parameters such as preferred sticky end sequence and modular cavity sizes.

7.8 REFERENCES

1. Zheng, J. P.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R. J.; Constantinou, P. E.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **2009,** *461* (7260), 74-77.

2. Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* **2016,** *138* (38), 12690-12690.

3. Simmons, C. R.; Zhang, F.; MacCulloch, T.; Fahmi, N.; Stephanopoulos, N.; Liu, Y.; Seeman, N. C.; Yan, H., Tuning the Cavity Size and Chirality of Self-Assembling 3D DNA Crystals. *Journal of the American Chemical Society* **2017,** *139* (32), 11254-11260.

4. Zhang, F.; Simmons, C. R.; Gates, J.; Liu, Y.; Yan, H., Self-Assembly of a 3D DNA Crystal Structure with Rationally Designed Six-Fold Symmetry. *Angewandte Chemie-International Edition* **2018,** *57* (38), 12504-12507.

5. Dannenberg; J, J., An Introduction to Hydrogen Bonding By George A. Jeffrey *Journal of the American Chemical Society*: Oxford University Press: New York and Oxford., 1998; Vol. 120, pp 5604-5604.

6. Oltrogge, L. M.; Boxer, S. G., Short Hydrogen Bonds and Proton Delocalization in Green Fluorescent Protein (GFP). *Acs Central Science* **2015,** *1* (3), 148-156.

7. Ohayon, Y. P.; Hernandez, C.; Chandrasekaran, A. R.; Wang, X. Y.; Abdallah, H. O.; Jong, M. A.; Mohsen, M. G.; Sha, R. J.; Birktoft, J. J.; Lukeman, P. S.; Chaikin, P. M.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., Designing Higher Resolution Self-Assembled 3D DNA Crystals via Strand Terminus Modifications. *Acs Nano* **2019,** *13* (7), 7957-7965.

8. Williams; S.; Lund, K.; Lin, C.; Wonka, P.; Lindsay, S.; Yan, H. In *T iamat:A Three-Dimensional Editing Tool for Complex DNA Structures, DNA Computing: 14th International Meeting on DNA Computing, Prague, Czech Republic*, Springer-Verlag: *Prague, Czech Republic, 2008*; pp 90–101.

9. Bode, W.; Mayr, I.; Baumann, U.; Huber, R.; Stone, S. R.; Hofsteenge, J., the refined 1.9 a crystal-structure of human alpha-thrombin - interaction with d-phe-pro-arg chloromethylketone and significance of the tyr-pro-pro-trp insertion segment. *Embo Journal* **1989,** *8* (11), 3467-3475.

10. Kao, M. R.; Kuo, H. W.; Lee, C. C.; Huang, K. Y.; Huang, T. Y.; Li, C. W.; Chen, C. W.; Wang, A. H. J.; Yu, S. M.; Ho, T. H. D., Chaetomella raphigera beta-glucosidase D2-BGL has intriguing structural features and a high substrate affinity that renders it an efficient cellulase supplement for lignocellulosic biomass hydrolysis. *Biotechnology for Biofuels* **2019,** *12* (1).

CHAPTER 8

USING THE 3D SELF-ASSEMBLING CRYSTALS AS A SCAFFOLD TO HOST BIOMOLECULAR GUESTS

8.1 Introduction

Structural DNA nanotechnology was conceptualized by Nadrian Seeman in 1980, when he sought to utilize a 3D DNA crystalline lattice as a way to scaffold proteins (guests) that could not otherwise crystallize. However, his proposal was elusive in that the first rationally designed self-assembled DNA crystal, the "tensegrity" triangle, was not reported until 2009, and the goal of using it as a means to determine the structure of a guest molecule has yet to be achieved.[1] However, since the first crystal structure was published, a significant amount of work has been done to determine the design rules for these unique crystal systems.

Following the tensegrity triangle, three additional systems with a "4xN" central weaving strand have been published, each with a unique number of interjunction bases: (1) the 4x5 design, which contained small aperiodic cavities that are not amenable to host any full size protein[2]; (2) the 4x6 system, which relieved the torsional strain that caused the 4x5 system's aperiodicity, and in so doing resulted in large homogeneous arrays of cavities in the scaffold[3]; and (3) the 4x7 design, which resulted in a rhombohedral lattice with unprecedented resolution, and demonstrated the tunability of the sizes of the cavities changing the number of helical turns in each duplex.[4] A fourth study from the Yan laboratory introduced a rationally design 6-fold symmetry crystal that was comprised of only two component strands, but was otherwise unrelated to any of the systems described in this dissertation.[5]

With additional design rules and improved resolution, achieving the goal of guest structural solution seems more attainable. This chapter will look at the steps that have been taken during my doctoral studies towards ultimately accomplishing the foundational goal of structural DNA nanotechnology. This work includes the incorporation of several different DNA-binding peptides and proteins that have been implemented through means of soaking, co-crystallization, and conjugation techniques. Additionally, as a proof of concept, the structures of several small molecular minor groove binders (MGB) were solved using the DNA crystals as a scaffold.

8.2 Scaffolding Proteins at Discrete Locations

8.2.1 Methodology to incorporating proteins

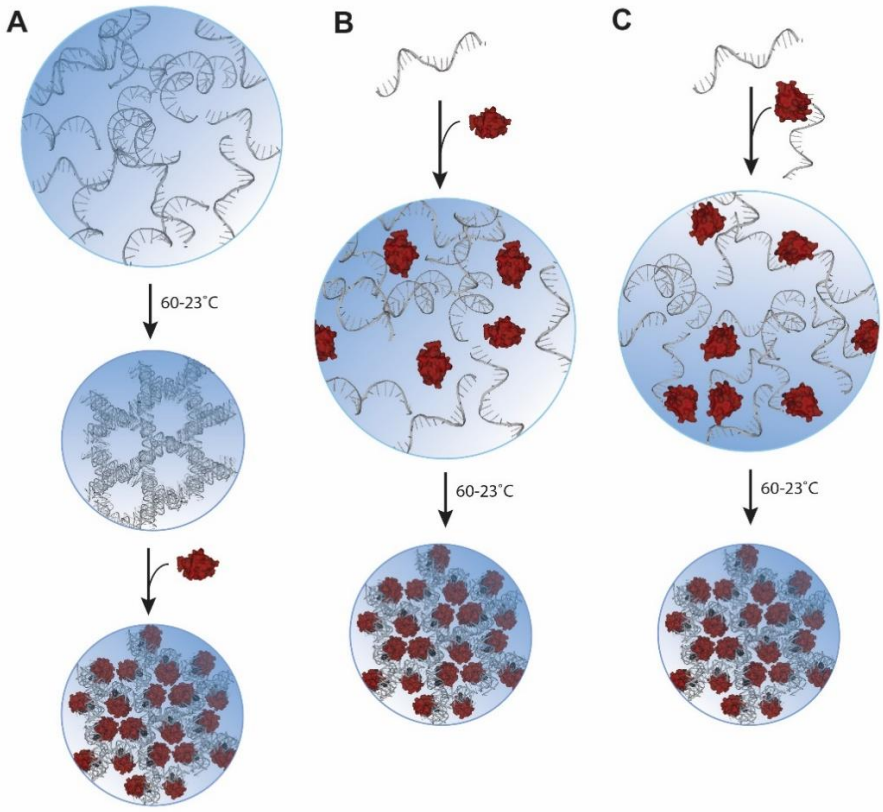


Figure 8.1. Methodology for the incorporation of proteins. Schematic showing the three approaches to scaffold the protein including (A) soaking (B) co-crystallization (C) and conjugation.

There are three possible ways to encapsulate a protein within the DNA lattice: post-crystallization soaking (Figure 8.1A) , co-crystallization (Figure 8.1B) , and covalently linking it to a component strand that "pre-tethers" it to the scaffold (Figure 8.1C). Post-crystallization soaking circumvents the possibility that the protein of interest (POI) might not be thermally stable at the starting incubation temperature of 60 °C required for formation of the DNA crystal (which would cause it to unfold and mostly likely precipitate). However, this method suffers from several disadvantages. First, many of the crystallization screening buffers contain viscous additives (e.g. polyethylene glycol) in the buffer, potentially making diffusion of the protein difficult. Hampering the ability of the protein to reach its cognate target sequence with 100% occupancy makes the ability to fully resolve the guest molecule within complete electron density accounting for the entire molecule challenging. Furthermore, if the channels in the crystal are of similar size to the protein, the guest molecules bound to cavities on the outer edge of the crystal could block access to the interior. A final issue with soaking is that you often observe non-specific binding, where positively charged patches on the surface of the protein target could electrostatically interact with the negatively charged backbone of the DNA at undesired sites.

Co-crystallization is a second option which introduces the protein to the mixture containing the component strands *before* it goes through its crystallization process of annealing and vapor diffusion. Like the soaking approach, there is no way to guarantee 100% occupancy even though the protein has more time to find its binding site during crystal formation. Moreover, this approach often leads to aggregation in the stock solution causing the DNA and proteins to precipitate before the mixture is even aliquoted

into each well. Some methods for remedying this issue, include lowering the overall concentration of the DNA and protein or introducing salts to help screen the charge-charge interactions. It is also possible that the precipitated components will resolubilize during the annealing process due to the heat. However, unlike the soaking method, when thermally unstable proteins are subjected to the thermal gradient they could entirely denature.

A final approach towards introducing the protein to the crystal scaffold is as a bioconjugate. Covalent conjugation provides a significant advantage because it guarantees 100% occupancy since the protein is tethered to one of the component strands, ensuring every time the strand binds to form a crystal, the protein is also present. Unfortunately, this approach is more laborious because it requires additional steps to obtain the DNA-peptide/protein conjugate, which can often lead to low yields of the product, making it problematic to obtain the large amount of material needed for crystallization screening. Another issue with achieving a sufficient yield is that the conjugate contains an anionic and a cationic portion, and intramolecular aggregation can occur before it is in solution at a reasonable concentration (~300 μM). Further, there are additional experimental design aspects that must be considered such as the length of the linker that is used to make the conjugate. This could introduce an issue whereby too long of a linker allows for the tethered protein to sample more than one binding site, and may lead to less than 100% occupancy, while too short of a linker would prohibit the protein from reaching its intended binding site at all. Lastly, as was the case with the co-crystallization method, denaturation of the guest molecule due to the thermal gradient is

also an experimental concern. These three approaches all offer a unique set of benefits and limitations and were therefore pursued in parallel.

8.2.2 Inspiration from Transcription Factors

To use the soaking or co-crystallization approach, the POI must intrinsically bind to DNA with sequence specificity. Nature provides the perfect example of this property with transcription factors (TFs), which are proteins that bind to DNA upstream of a gene to regulate its transcription activity. The binding region of these proteins is typically between 4 to 12 nucleotides long, and they bind with affinities ranging from pico- to micromolar.[6] While most transcription factors bind in the major groove, which is ~22 Å wide, some are small enough to bind in the minor groove, which is ~12 Å (Figure 8.2A).

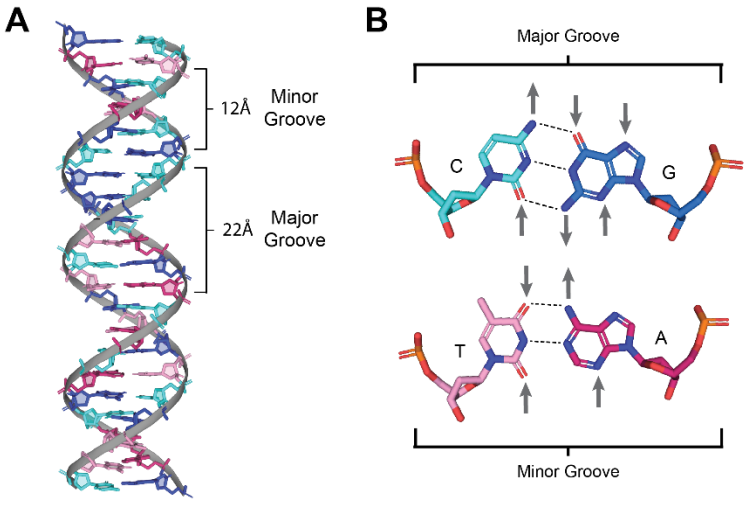


Figure 8.2. Binding modes of transcription factors. (A) Location and dimensions of the major and minor grooves used for bindings (B) Available donor (outward arrow) and acceptor groups (inward arrow) found in the major and minor grooves for the Watson-Crick base pairs.

The sequence specificity of DNA binding proteins comes from a hydrogen bond donor/acceptor pattern (Figure 8.2B) created by the functional groups of the specific

bases in the groove and the adjacent phosphate and the side chains of the amino acids. Additionally, most TFs contain positively charged amino acids such as lysine and arginine to assist with the DNA binding. Several major groove binding TF classes include helix-turn-helix, leucine zippers, zinc fingers, homeodomains, and forkhead domains, while minor groove binding TFs include AT-hooks and SRY proteins (Figure 8.3).

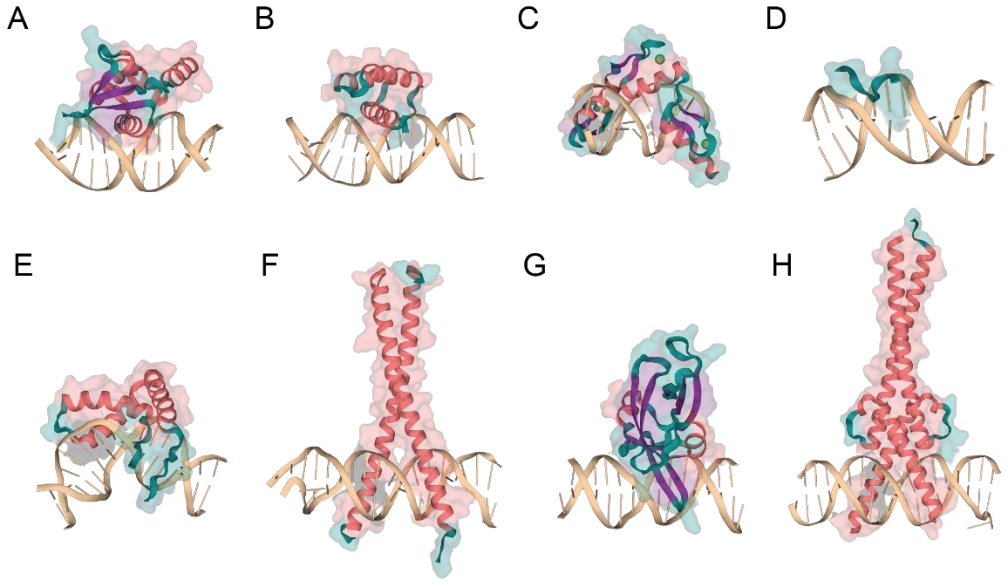


Figure 8.3. Examples of transcription factors. Various classes of transcription factors shown in cartoon views with the binding DNA shown in tan, α-helical regions shown in pink, β-sheets shown in purple, and random loop regions shown in teal. (A) forkhead domain, adapted from 6EL8 (B) homeodomain, adapted from 2HDD (C) zinc finger, adapted from 1AAY (D) AT-hook, adapted from 3UXW (E) Sox HMG, adapted from 3U2B (F) leucine zipper, adapted from 1YSA (G) LEC transcription factor with B3 binding domain, adapted from 6J9C (H) helix-loop-helix, adapted from 1NKP

When choosing a transcription factor to incorporate into the DNA crystals, several features must be considered. The first is the obvious size restriction for both the protein and its binding sequence. The protein must be relatively compact, so as to fit into the cavity sizes of the known crystal systems (approximately less than 5 nm per side) and

ideally has a binding sequence that fits between the junction and the sticky ends. However, if the binding sequence is longer than this region, it can bridge over the sticky end, as long as the phosphate that is missing at the nick site does not directly interact with the protein. The solubility and isoelectric point (pI) of the protein must also be taken into consideration because it needs to be compatible with the crystallization buffers, and must be stable at high concentrations (minimum of 120 μM) to be added in at least a 1:1 stoichiometric ratio with the binding site. The melting temperature must also be taken into consideration if the protein is to be introduced through a co-crystallization or conjugation strategy since it will need to be stable at the higher temperatures during annealing.

8.2.3 Incorporation of an engrailed homeodomain

The considerations listed above are required when choosing a TF as the guest. Homeodomains were determined to be the ideal target because they are reasonably small (~60 residues) with a simple overall structure. These proteins consist of three short helical regions, one of which is responsible for binding within the major groove, and an N-terminal minor groove binding arm. While helix 3 (Figure 8.3A) is responsible for the overall binding, it was found that the N-terminal arm is important to achieve the highest binding affinity and specificity.[7-9] Several studies have shown that a mutation in residue 50 (Q50K) changes the binding sequence from the wild type TAATTA to TAATCC with higher specificity and binding affinity.[10-12]

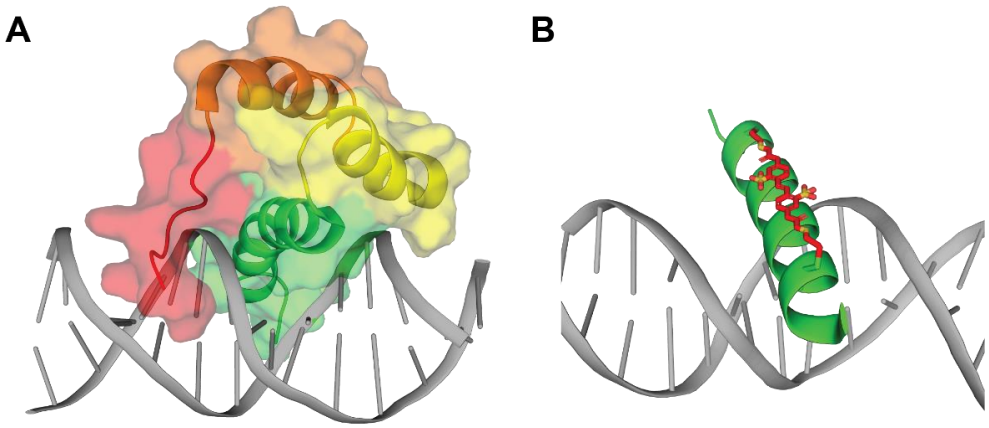


Figure 8.4. Homeodomain binding and truncation. (A) Overall structure of the homeodomain class with the N-terminal minor groove binding arm shown in red, helix 1 and 2 shown in orange and yellow, and the helix responsible for binding (helix 3) shown in green. Adapted from 2ME6. (B) Cartoon representation of the truncated homeodomain, HDH3, with a crosslinker with *i*, *i*+11 spacing to stabilize the α-helix

In 2005, a truncated variation (HDH3) was reported, which consisted of the 18 amino acids that made up the binding α-helix (Figure 8.4B). However, without the full protein, these residues did not adopt their native conformation, and instead form a random coil, disallowing binding. To overcome this limitation, two nonbinding residues (originally both isoleucine) at an *i*, *i*+11 spacing (i.e. three turns of the helix apart), were mutated to cysteines, allowing them to be crosslinked with an azobenzene (Figure 8.5A). This crosslinker served to stabilize the peptide in its native α-helical conformation, allowing for its binding with a $K_d = 7.5 \pm 1.3$ nM.[13] This high binding affinity, and the ease of synthesis for a 18-residue peptide, made HDH3 an ideal starting target for incorporation into the crystals. The peptide was synthesized using traditional SPPS methods, and a fluorophore was incorporated on the N-terminus to observe the binding based upon the uniformity of the coloring of the crystal. Following synthesis of the peptide, it was cleaved from the resin and purified using RP-HPLC prior to crosslinking.

The crosslinker (with the initially photoswitchable azobenzene replaced with a non-isomerizable stilbene) was then mixed with the peptide at a 1:1 ratio in PBS and allowed to react overnight before the crosslinked and unmodified peptide were separated using RP-HPLC. While the crosslinking step never led to a high yield of usable peptide (≤ 60%), adding the crosslinker at a higher ratio often led to two crosslinkers attached to the same peptide. The final product was characterized using MALDI-TOF MS (Figure 8.5B).

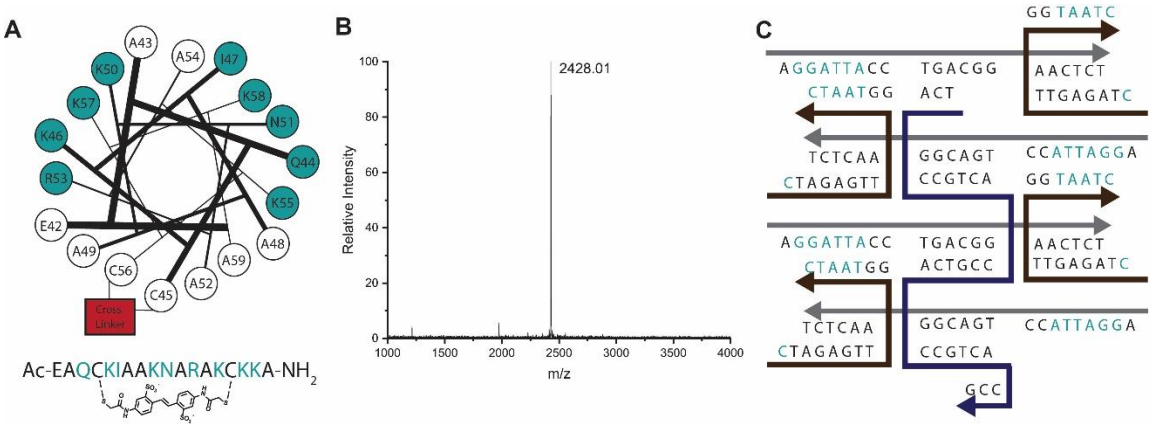


Figure 8.5. HDH3 peptide sequence, characterization, and its corresponding DNA lattice for binding. (A) Helical diagram for the full sequence of HDH3 with the stilbene crosslinker; the binding residues are highlighted in teal. (B) MALDI-TOF MS spectrum confirming the identity of the synthesized and crosslinked peptide (expected mass of 2393). (C) Modified 4x6 system with the TAATCC binding sequence spanning the sticky end shown in teal.

The next step was to modify the DNA strands, so they contained the Q50K binding sequence (TAATCC). The 4x6 system was chosen as the host scaffold because it produced a lattice with larger periodic cavities. With a binding region of six nucleotides, the only place that it would fit was on the duplex outside of the two junctions and it was placed over sticky ends so its binding would not be sterically hindered by the junctions (Figure 8.5C). It was also confirmed that the new sequences still readily crystallized.

As a first test, a portion of the HDH3 peptide was labeled with rhodamine, and soaked into the crystals to monitor its uptake. The crystals turned pink (whereas crystals lacking the binding site did not selectively take up the protein), indicating that the peptide had successfully diffused throughout the crystal and bound to its site. The protein-laden scaffold subsequently diffracted to 3.2 Å (Figure 8.6A, B). Unfortunately, the resulting molecular replacement structure had no additional density attributable to the peptide (Figure 8.6C, D). This approach was repeated several times, but showed no improvement. The lack of density could have been due to several factors, including a significant lack of occupancy at the binding sites, the peptide binding non-specifically to the lattice, or the peptide not being completely static and immobilized on the duplex.

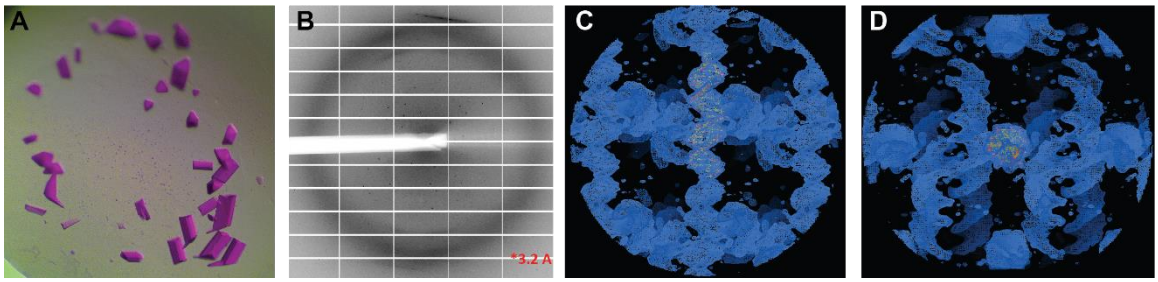


Figure 8.6. Incorporating HDH3 using a soaking technique. (A) Representative image of HDH3 binding crystals after soaking with rhodamine-labeled peptide and (B) diffraction pattern of peptide-soaked crystals. (C, D) Molecular replacement solutions of the peptide-soaked crystals to demonstrate the lack of additional density corresponding to the peptide.

As discussed earlier, one way to ensure 100% occupancy was to incorporate the peptide through a DNA-peptide conjugate. For this purpose, an adapted variation of the HDH3 peptide was synthesized, with the addition of a GSG linker and N-terminal azidolysine that could be utilized for a copper free click reaction with a DBCO labeled DNA handle (Figure 8.7A). Prior to choosing a conjugation site, the distance to the binding region was first determined to ensure that the linker chosen was sufficient to

allow the peptide to bind without strain, but close enough that it could not sample more than one binding site. Furthermore, since employing a copper free click which is not regiospecific, both the 1,3- and 1,5- cycloaddition products had to be considered. Eventually the 5' end of the S3 strand was chosen because its position was approximately 2 nm away from the azidolysine residue when docked into the binding site using PyMOL,[14] with a total linker length of 2.5 nm (Figure 8.7C, D). Several attempts to purify each of the click isomers individually were made, since a slight separation occurs during RP-HPLC purification (Figure 8.7B); however, this approach was unsustainable due to the extremely low yield, with no noticeable improvement in the electron density. While the S3-HDH3 conjugate system readily crystallized, the scaffolds typically diffracted to much lower resolutions (~4.5 Å) than the native 4x6 crystals (Figure 8.7E, F). Despite the peptide being tethered near the binding site, only a small portion of additional density was observed (Figure 8.7G).

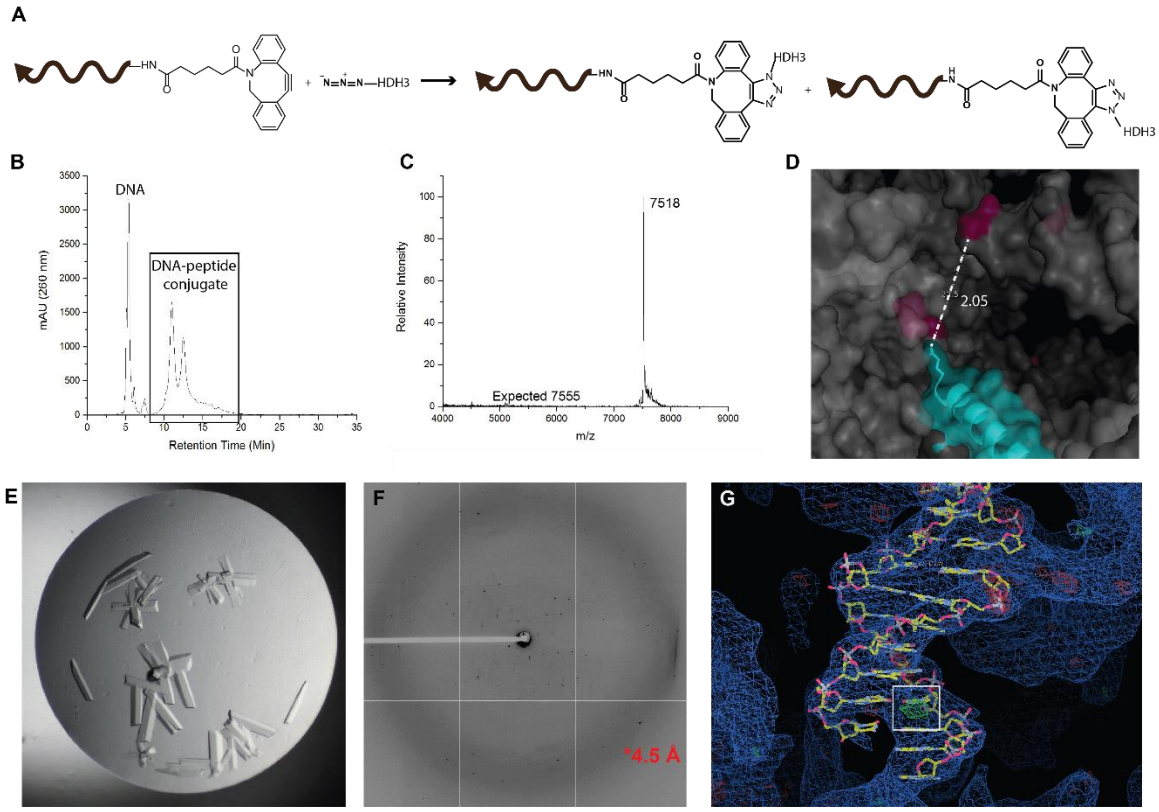


Figure 8.7. Incorporation of HDH3 through conjugation. (A) Conjugation scheme using a copper free click to generating the 1,3- and 1,5- cycloaddition products. (B) RP-HPLC chromatogram showing three peaks, the first corresponding to unmodified DNA, and the next two (~10-15min.) corresponding to the two isomer conjugates. (C) Mass spectrum of the S3-HDH3 conjugates. (D) Measurement showing the distance from the 5' of S3 to the approximate location of azidolysine residue that was added. (E) Representative image of the 4x6 HDH3 conjugate crystals. (F) Diffraction pattern of the conjugate crystals with a resolution of 4.5 Å. (G) Molecular replacement solution with the small portion of additional electron density (boxed off in white) that can be attributed to the presence of the peptide.

While the truncated peptide (HDH3) was the initial target due to the ease of synthesis, it was known that the peptide itself had a thousandfold lower binding affinity (K_d = 7.5 nM) than the full-length protein (K_d = 8.8 pM). Furthermore, it is often observed that the lower the solvent content within a given crystal, the better the resulting resolution. For these reasons, attaching the full-length protein was also explored.

Initially the protein was expressed in *E. coli* using a traditional Gibson assembly approach.[15] The designed plasmid contained a glutathione S-transferase (GST) tag to increase the solubility of the protein and allow for facile purification. Following expression the cells were lysed, and both the supernatant and pellet were analyzed by SDS-PAGE to determine whether the protein was soluble, and properly overexpressed. The soluble GST-HDD protein was then purified using a GST column, with PBS pH 7.3 as a running buffer and Tris pH 8.0 supplemented with 10 mM glutathione as the elution buffer. Each fraction was again analyzed using gel electrophoresis (Figure 8.8A), and pure fractions were collected and incubated with thrombin in order to cleave the GST tag. The cleaved HDD was then purified by once again by affinity chromatography to bind the uncleaved GST-HDD, and the fractions containing the flowthrough were pooled and concentrated (Figure 8.8B).

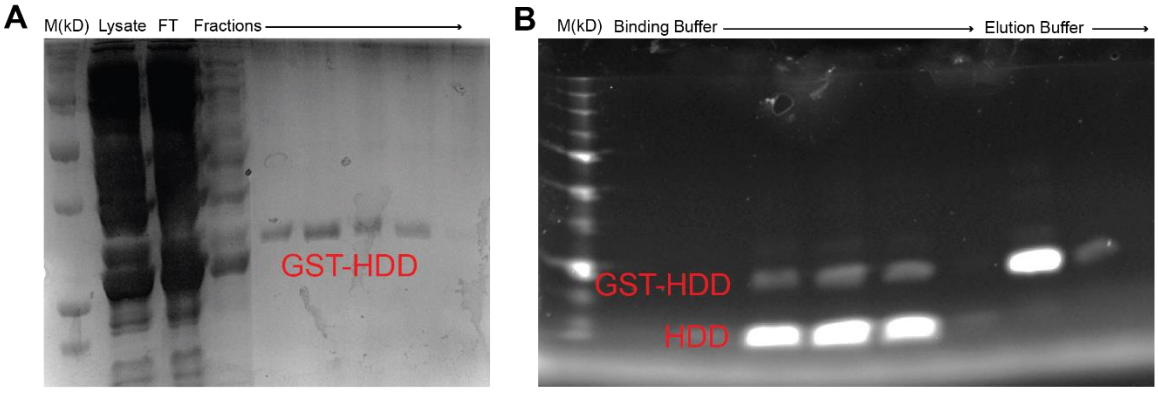


Figure 8.8. Expression of Engrailed Homeodomain. (A) SDS-PAGE analysis run after purification of GST-HDD from cell lysate stained using Coomassie blue for imaging and (B) SDS-PAGE following the cleavage of the GST tag, run on a stain free gel.

The yield of the purified protein was relatively low, and unfortunately not adequate for crystallization screening. Additionally, the expressed protein contained no available functional groups for site-specific modification to allow for DNA conjugation,

so it could only be used for co-crystallization and soaking approaches. Since the full protein did not require any post-translational modifications, and was only ~70 residues (which is near the limit for microwave-assisted solid phase synthesis) a method was developed to synthesize the protein using SPPS. Furthermore, a mutation was introduced during the synthesis to incorporate a cysteine that could be used for site-specific conjugation (Figure 8.8a). To successfully synthesize the protein, coupling times were doubled from the standard 2 minutes to 4 minutes, and each residue was coupled twice. Following synthesis and purification, the synthetic protein was characterized using MALDI-TOF MS (Figure 8.9B), and subsequently conjugated to an amine-modified DNA strand via a succinimidyl 3-(2-pyridyldithio)propionate (SPDP) linker (Figure 8.9C). Unfortunately, the DNA-protein conjugate could not be used for crystallization because it irreversibly aggregated during the concentration step. Additionally, trying to soak the synthetic protein alone into already formed crystals also led to precipitation (Figure 8.9D). Circular dichroism (CD) spectroscopy was also used to characterize the synthesized protein, and it properly displayed its expected α-helical character. Additionally, the synthetic protein was also able to properly bind to a 21-bp duplex containing the TAATCC sequence (Figure 8.9E). Co-crystallization was then used for

incorporation of the protein into the crystal because it was soluble up to a 2x (2:1) ratio (protein: binding sequence).

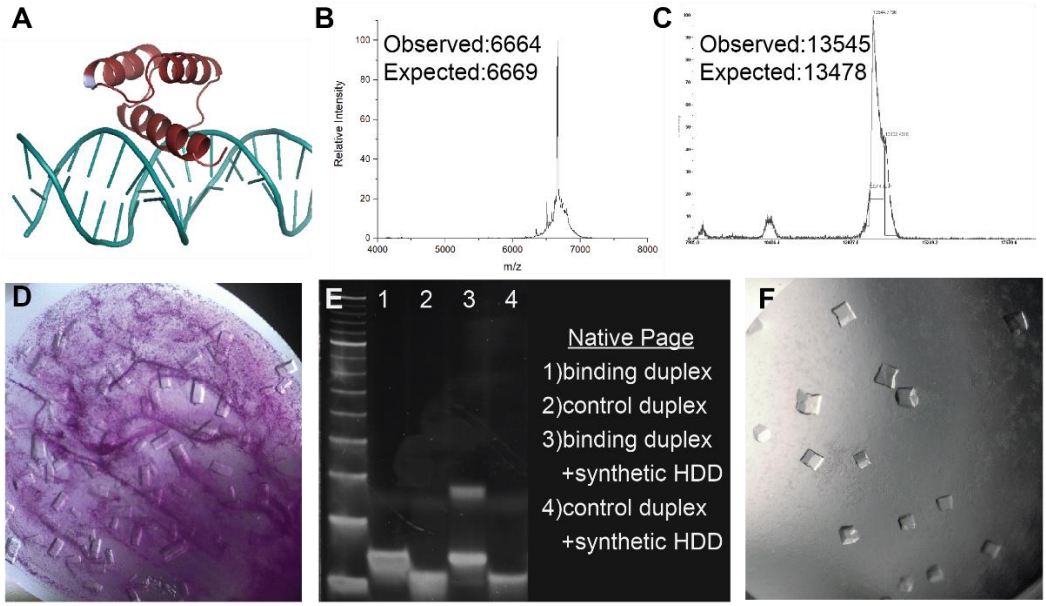


Figure 8.9. Synthesis, characterization, and incorporation of synthetic HDD. (A) location of the cysteine mutation (in the loop between helix 1 and 2) in Q50K HDD is highlighted in purple (B) MS of synthetic HDD and (C) DNA-HDD conjugate. (D) Image of precipitated rhodamine-labeled HDD during a soaking experiment (E) Native PAGE showing an upward band shift in lane 3, corresponding to the DNA-protein complex and a (F) representative image of the HDD co-crystals at a 2x binding ratio.

With little success in obtaining extra density corresponding to the HDH3 peptide or the HDD protein in the 4x6 system, alternative targets were explored. An additional truncated homeodomain was reported in 2018, which consisted of not only the binding helix, but also the N-terminal minor groove binding arm.[16] This peptide, called BPIB, incorporated aminoisobutyric (AiB) acid, an unnatural amino acid that is known to propagate helical structure due to the additional steric hindrance imparted by the second methyl group. This in turn stabilizes helix 3, which was found to have no secondary structure without the AiB; furthermore, each AiB residue was strategically introduced at

a location that was not responsible for binding. The recognition helix was ~12.6 Å from the N-terminal arm, so a known unstructured peptide sequence, IKMEFID, of comparable distance was inserted between the two. This truncated peptide was then tested against several binding sequences, with the best affinity being 23.6 ± 6.81 nM with a CAAT sequence.

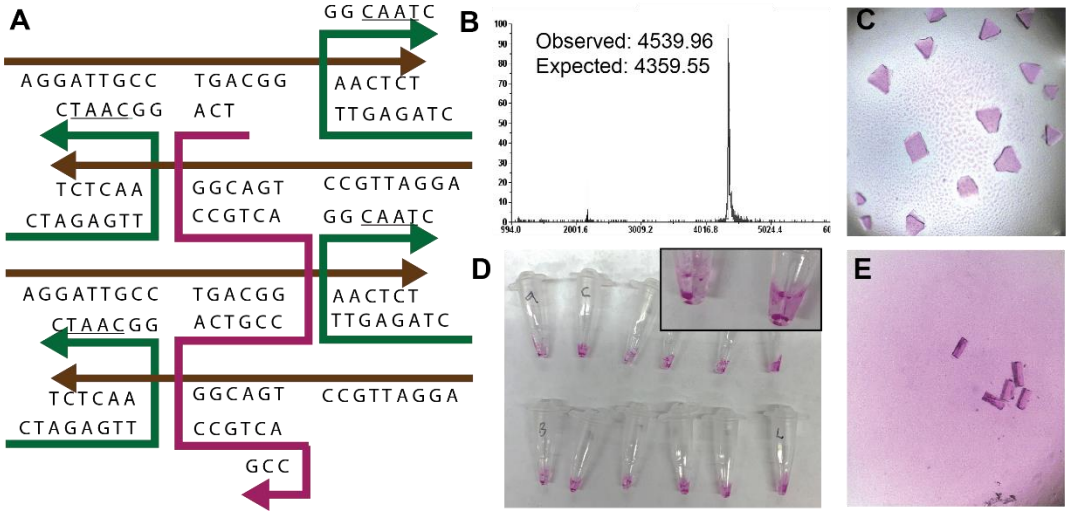


Figure 8.10. Incorporation of BPIB into the 4x6 system. (A) 4x6 system with the CAAT binding site modification (B) MALDI-MS spectrum of the rhodamine labeled BPIB that was synthesized (C) co-crystallization at a 1:1 ratio of binding site to protein (D) Representative image of testing the 48 crystallization buffers for BPIB solubility at 3.5x binding ratio (~750 µM) (E) soaking in BPIB at a 2x binding ratio.

The binding sequence was incorporated into the 4x6 system near the 3' end of S3 (Figure 8.10A) and crystals formed readily in a variety of buffers. The reported BPIB sequence (LRKPRSIY-IKMEFID-SBTQBKIWBQNBRSK) was synthesized, labeled with rhodamine, and characterized by MALDI-TOF MS (Figure 8.10B). Two approaches were taken with BPIB, which included both soaking and co-crystallization. While promising crystals were produced at a 1x binding ratio through co-crystallization

and 2x binding ratio through soaking (Figure 8.10C, E), no density was observed using these approaches. When trying to push the ratio of binding site to protein, it became apparent that BPIB was not very soluble. To confirm this, a mixture of DNA duplex and BPIB at a 3.5x binding ratio was added to each of the 48 crystallization screening buffers, incubated for 5 minutes, and centrifuged to determine the amount of precipitation which was significant (Figure 8.10D). The BPIB co-crystals routinely diffracted to ~3.0 Å indicating that the binding of the protein did not perturb the lattice.

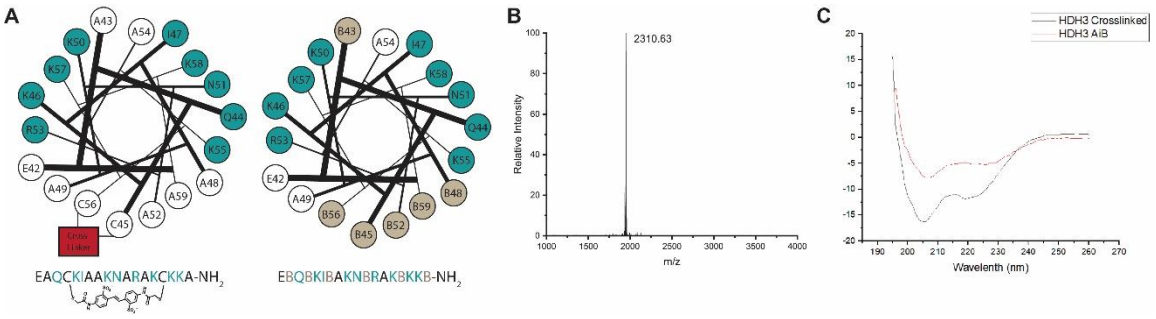


Figure 8.11. Redesigning HDH3 with AiB. (A) comparison of the published HDH3 structure and the modified AiB HDH3 with its corresponding mass spectrum (B). CD spectrum showing the α-helical character of the AiB version (C).

Considering the clear stability of the BPIB crystals, and the high solubility of HDH3, the truncated peptide was redesigned to incorporate AiB residues in order to avoid the crosslinking step while still inducing helical character (Figure 8.11A). The peptide was characterized with both mass spectrometry and CD spectroscopy to ensure its successful synthesis, and to confirm that the chosen AiB locations were sufficient to induce helical structure (8.11B, C). Additionally, at this point the 4x7 structure had been solved, and contained cavities sizes that were amenable to host HDH3 while also diffracting to much higher resolutions than the 4x6 system. With the rationale that the higher the resolution, the more likely extra density would be observed even if the

occupancy was not optimal, the TAATCC binding sequence was inserted into the 4x7 motif (Figure 8.12A).

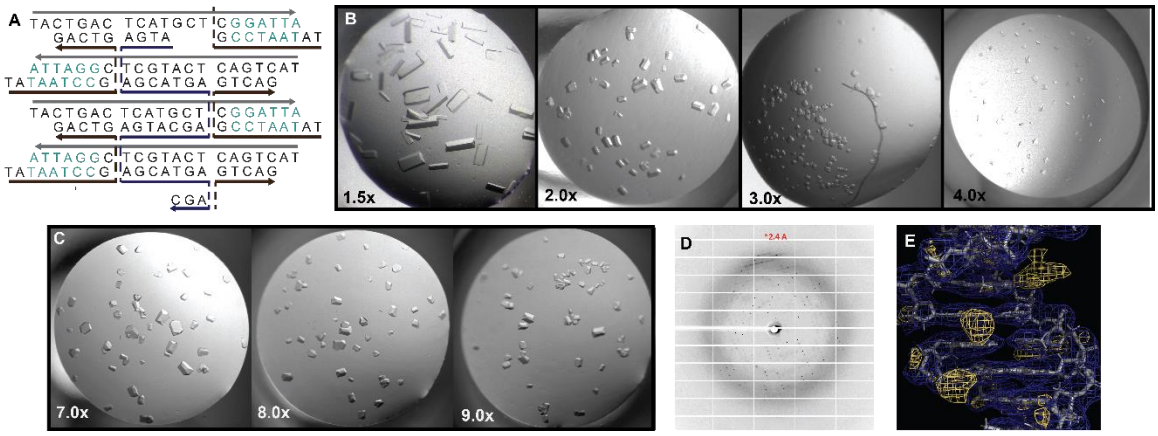


Figure 8.12. AiB HDH3 in the 4x7 system. (A) Modified sequences for the 4x7 with the HDH3 binding sequence shown in teal (B) Representative photos from full crystallization screens with increasing binding ratios, showing the decreasing crystal size. (C) representative photos from setups that included the buffer into the mixture before adding the protein (D) Diffraction pattern of the 4x7 AiB HDH3 crystals showing increased resolution (E) Initial molecular replacement solution showing the additional electron density that corresponds to the peptide at the proper site in the major groove shown in yellow.

Initially, the new sequences were screened to find optimal co-crystals at ratios of the AiB-HDH3 ranging from 1-4x (Figure 8.12B). During this screen, it was observed that the higher ratios produced smaller crystals, and the DNA/peptide stock started to precipitate around 4x. Traditionally, the 4x7 system is set up with the DNA strands incorporated at concentrations of 30:120:120 µM (S1:S2:S3), but to try to stop precipitation, the co-crystals at a 4x binding ratio were re-screened with concentrations of 20:80:80 µM. To further enhance the binding ratio, the buffer was added directly to the DNA stock before the addition of the peptide. With this methodology, the binding ratio was able to produce crystals suitable for harvesting with up to a 9x binding ratio (Figure

8.10c). Upon diffraction, it was once again observed that the incorporation of the AiB containing peptide led to improved diffraction over the native crystals, or crystals with the chemically crosslinked peptide (Figure 8.12D). Furthermore, inspection of the electron diffraction maps, positive difference density (shown in yellow) was observed at the proper location for the intended binding site in the major groove (Figure 8.12E), a result that suggests great promise towards achieving a host-guest crystal structure.

8.2.4 Switching to a soluble consensus homeodomain

An alternative homeodomain, produced using a consensus design (sequence averaging), was explored concurrently with HDD. This consensus homeodomain (CHD) was expected to be more stable than the engrailed homeodomain sequence, and was shown to bind to the same 6-nt sequence.[17] With the most promising results coming from the peptide variations of the HDD protein, similar principles were used to create three truncated consensus homeodomains (tCHD). Much like with BPIB, the N-terminal minor groove binding arm was included in each of the truncated variations, with the first variation simply being the N-terminal arm, the linker, and the native recognition helix with no mutations. The second and third variants also contained the same N-terminal arm and linker, but the first mutated two non-binding residues with *i, i*+11 spacing to cysteines with one that could be crosslinked, and the other incorporated four AiB residues to stabilize the helical structure.

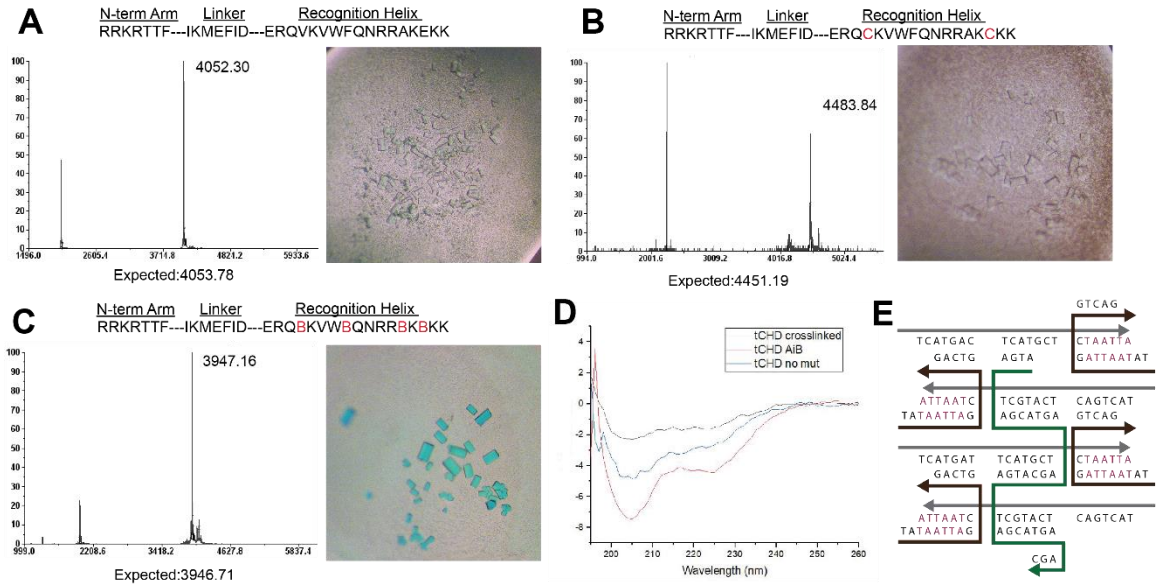


Figure 8.13. Truncated CHD variations and characterization. The three designed truncated versions that were tested are shown with their sequence, MALD-TOF MS spectra, and Cy5-labeled soaking experiments with crystals containing the binding sequence (A-C). Each of the variants was tested using CD spectroscopy to determine if any α-helical character was present in the recognition helix with only AiB showing promise. (E) Sequence mutations of the 4x7 system with the bases responsible for binding shown in purple.

Following SPPS, each variation was characterized using MALDI-TOF MS and subsequently labeled with Cy5. Crystals containing the binding sequence in the 4x7 system (Figure 8.13E) were produced and soaking experiments for all three versions were carried out. While no peptide was observed to diffuse and bind throughout the crystals for the no mutation and crosslinked variants, the AiB-tCHD soaked crystals readily turned blue, indicating that the protein had diffused throughout (Figure 8.13A-C). Furthermore, the secondary structures for each variant were analyzed using CD spectroscopy, with only the AiB variant showing the characteristic minima for α-helices at 208 and 222 nm; therefore only the AiB-tCHD peptide was explored during crystallographic studies.

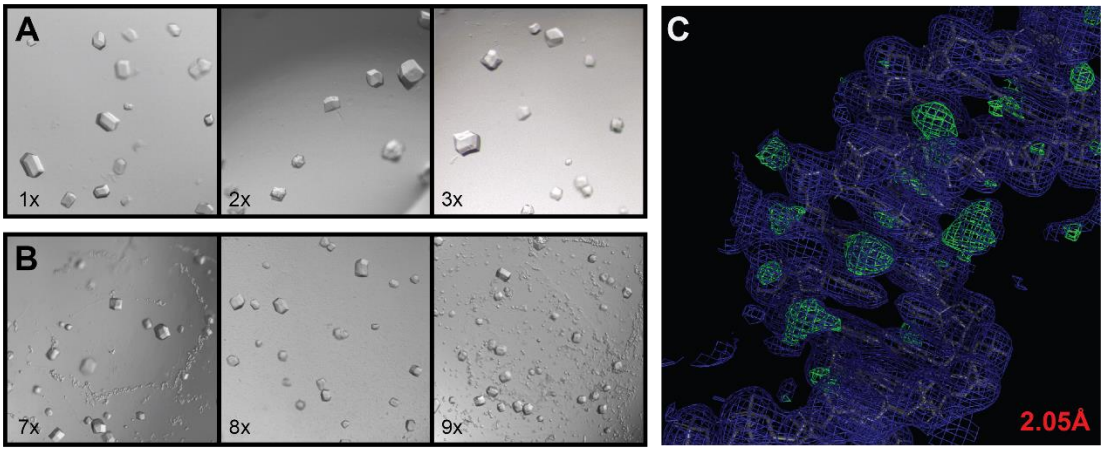


Figure 8.14. Co-crystallization of AiB-tCHD results. (A) Representative images for the crystals obtained during traditional full screens of the 4x7 AiB-tCHD cocrystals at binding ratios of 1-3x. (B) Representative images for the crystals obtained while incorporating the buffer into the master mix of 4x7 AiB-tCHD cocrystals at binding ratios of 7-9x.

Similar to the route taken with AiB HDH3, full screens for co-crystals were carried out with increasing binding ratios. With this approach, it was observed that this method was only successful at lower ratios (1-3x), before the sample precipitated in the master mix (Figure 8.14A). While some crystals still formed with a precipitated mix, it is impossible to know the true concentration and binding ratio in each of the drops since the mixture is no longer homogeneous. In addition, subsequent molecular replacement solutions revealed no additional density. The binding ratio was further optimized by adding the high salt buffer to the DNA mixture prior to adding the protein, which ultimately resulted in crystals at a 9x binding ratio (Figure 8.14B), and eventually diffracted to 2.05 Å, a resolution never before seen with any self-assembled crystal. Also, the initial results show positive (extra) density at the anticipated location in both the major and minor grooves (Figure 8.14C), presumably corresponding to the recognition helix and the N-terminal binding arm, respectively. These initial experiments demonstrate that the designed truncated peptide likely binds to the DNA with high

affinity and is a promising route towards solving *de novo* protein structures in the DNA lattice.

8.2.5 Incorporation of the minor groove binding AT-hook

While most TFs bind in the major groove, several are known to bind solely in the minor groove. One example of this is an AT-hook, which binds at sequence specifically at an AATT region, lacks any formal secondary structure and contains the sequence (K/R)XRGRP (Figure 8.15A).[18, 19] These short sequences have also been shown to help stabilize the B-form DNA duplex instead of inducing kinks like some major groove binding TFs do. Three copies of this sequence are often found in non-histone chromosomal high mobility groups (HMG), separated by linkers of 11-23 amino acids each. Each of the three DNA binding domains (DBD-1, DBD-2, DBD-3) can bind individually, and can thus be truncated down to each region; however, to obtain nanomolar binding affinities at least two copies must work cooperatively.[20]

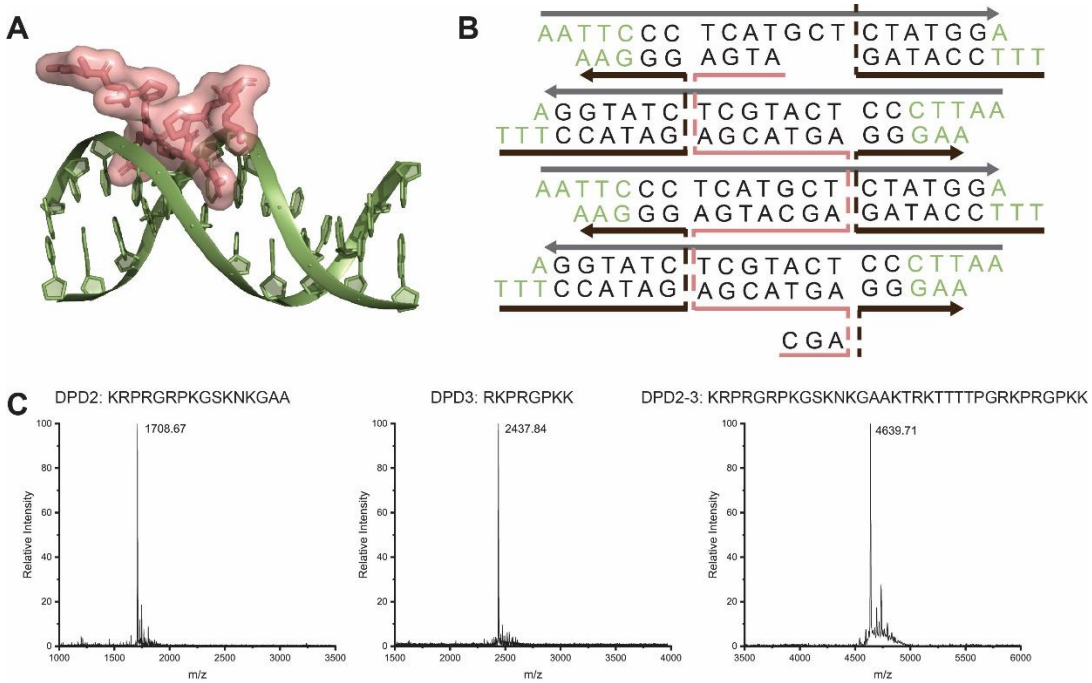


Figure 8.15. AT-hook in the 4x7 system. (A) Representative image of a minor groove binding AT hook peptide modified from PDB code 3UXW. (B) 4x7 design with modified sequences to include the AT binding sequence AAATTC (C) MALDI-TOF mass spectra of the three different AT hooks, DPD-2 (expected mass:1707), DPD-3(expected mass:2436), and DPD2-3(expected mass:4643).

As the name suggests, these motifs prefer to bind to AT-rich regions with the highest affinity being a consecutive sequence of five A-T base pairs.[21] With the best results coming from protein incorporation in the 4x7 system, the sequences were redesigned to contain the AAATTC binding sequence (Figure 8.15B). The only crystal of the AT hook binding domain (DPD3) was published in 2012, and it showed that the binding is due to hydrogen bonding between the amine of the arginine with the thymine bases. Additionally, the structures showed that the minor groove widened (from ~9.5 to ~12.5 Å) to accommodate the peptide and the duplex was bent to a 24° angle.[19] With little structural data available, both the second and third binding domains were synthesized individually, along with a third peptide that contained both regions (figure 8.15C). While full screens with each of the three peptides were carried out, no promising crystals past a 2x binding ratio were produced. However, upon optimizing a buffer containing 50 mM cacodylate pH 6.5, 100 mM $MgCl_2$, 2.0 mM $CoH_{18}N_6$, 5% isopropanol, the three versions crystallized at a 5x ratio (Figure 8.16AD). Unlike the homeodomains, the AT-hook crystals preferred a low salt buffer so adding the buffer to the stock prior to adding the peptides did not help with solubility; however, a lower overall degree of aggregation was observed in the DNA-peptide mixture.

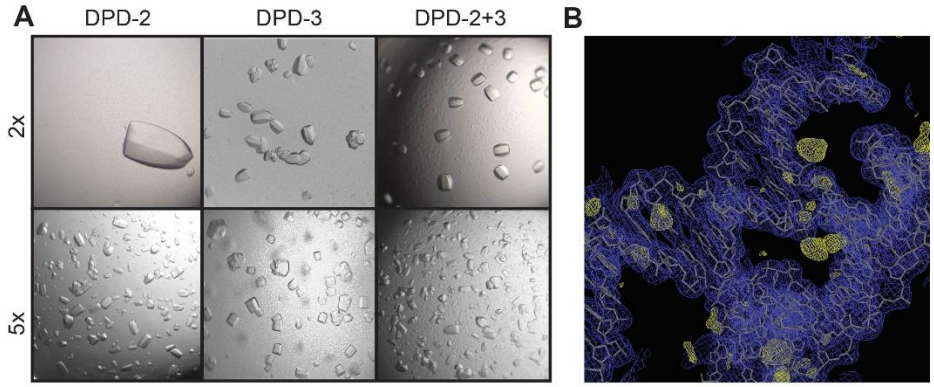


Figure 8.16. AT-hook DPD variant co-crystal results. Representative images of the cocrystals produced at both 2x and 5x binding rations for the three different variants (A) along with the molecular replacement solution highlighting the additional density (yellow) of DPD2 around the junction (B).

While all three variations produced crystals, both individual binding domains (DPD2 & DPD3) outperformed the peptide containing both binding domains which diffracted to lower resolution and did not show additional density. Curiously, the resulting density from the molecular replacement solution showed that both variants contained unexpected density surrounding the Holliday junction instead of the designed binding site (Figure 8.16B). While there are no crystal structures demonstrating this phenomenon to date, several studies utilized DNA foot printing to demonstrate that this is a common binding mode of an AT-hook.[22, 23] This preliminary data suggests that AT hook peptides also hold great promise towards coordination and structure determination of a novel structure.

8.3 Proof of Concept: A look at minor groove binders

the ultimate goal is to scaffold biomolecules with unknown structures such as novel proteins, small molecules with existing crystal structures were explored as a proof

of concept that the crystals could precisely scaffold some molecular guest. Toward this end, a group of commercially available DNA-binding molecules were chosen, all of which bind AT-rich regions in the minor groove. These molecules can all be classified as polyamines, and those selected for the work included netropsin, a common antibiotic precursor, and DAPI and Hoechst which are both used as fluorescent DNA stains in microscopy. Each of these molecules have known binding modes to the sequence AATT and have a curved shape that mimics the curvature of DNA (Figure 8.17A).[24-26] The 4x5 and 4x6 systems were both modified to contain either binding sequence at two discrete locations (Figure 8.17B, C). Each of the three MGBs were solved in at least one of the positions and netropsin was shown to bind at both positions, simultaneously.

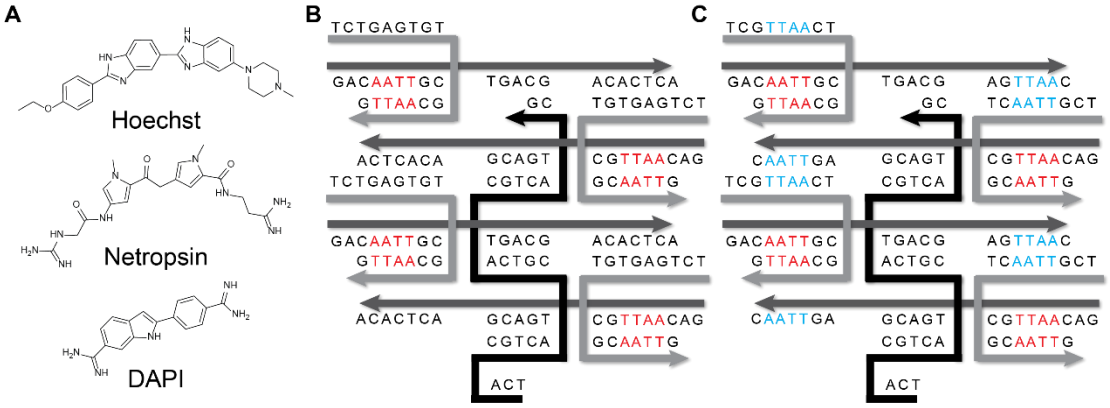


Figure 8.17. MGBs and the modified 4x5 system. Chemical structure of Hoechst, netropsin, and DAPI to illustrate their curved nature and polyamine composition that is responsible for the binding (A). The 4x5 system with the modified AATT binding site (red) at position 1 (B) and position 2, blue, (C) that were used to scaffold the binders individually or simultaneously.

Like proteins, these small molecules can be incorporated into the DNA crystal lattices by soaking or co-crystallization. Both DAPI and Hoechst are fluorescent, with emission wavelengths of ~460 nm when bound to DNA, allowing for an easy visualization of the molecules diffusing through the solvent channels and binding

efficiently and uniformly throughout the crystal (Figure 8.18A). Unlike proteins, these molecules cannot denature during the annealing process, making them ideal candidates for co-crystallization. With this approach being the most promising with the peptide targets, each of the binders was introduced to the DNA mixture at a 4x ratio of MGB to the available binding site(s) prior to crystallization. Much like the peptide co-crystals, the presence of the MGBs led to higher resolution data sets, indicating that they can actually stabilize the lattice. Following molecular replacement, continuous electron density for each respective molecule was observed at each designed location in the minor groove (Figure 8.17B-D). Each of the resulting structures of the MGBs within the DNA crystal structures corresponded perfectly with reported structures with duplex DNA, including all known molecule-DNA contacts.[24-32]

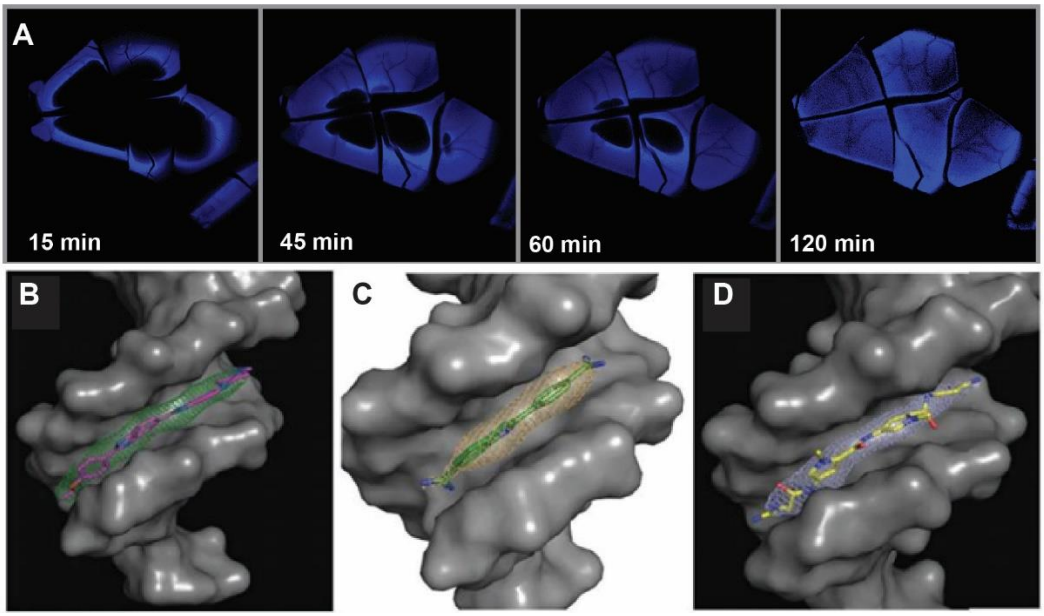


Figure 8.18. Incorporating MGBs via soaking and co-crystallization. Time course study of DAPI being soaked into the crystal with its binding sequence over 2 hours (A) and the structures of the MGBs within their corresponding electron density for Hoechst (B), DAPI (C), and netropsin (D).

8.4 Conclusion

The work carried out in this chapter demonstrates great strides towards achieving the original goal of DNA nanotechnology and shows that while this goal was not as simple as originally conceived, but it is not impossible. Several DNA-binding proteins and peptides were incorporated into crystals that had been modified to contain their intended DNA binding sequences. Three different approaches were tested simultaneously because they each come with unique shortcomings and advantages. While the soaking approach seemed promising due to uniformly colored crystals, it always yielded the least promising structural results. The conjugation approach was intuitively superior due to the increased chance that the peptide/protein would be present at a 100% occupancy; however, it was often found the yield of the conjugation reaction was only high enough to produce a limited amount of crystals, and it was extremely challenging to produce an ample amount of material required for large scale crystal screening. Additionally, the conjugates often were insoluble at the high concentrations (~300 μM) that are required for crystallization. The preliminary results from the DNA-peptide conjugate approach indicate that this route should not be discarded, but at the moment, the resolutions obtained were consistently lower than the crystal scaffolds alone, with only a limited amount of observable density, thus far.

At this point, the co-crystallization approach appears to hold the most promise. Several peptides (AiB HDH3, AT-Hook, and AiB tCHD) have been optimized to incorporate the peptides at high ratios (~6-9x of the binding sites), and have consistently shown additional electron density in the designed locations. It is also noteworthy that the incorporation of the non-canonical aminoisobutyric acid (AiB) residue into the peptides

to stabilize the alpha-helical secondary structure has consistently led to higher resolution structure, which is vital to the accurate description of a *de novo* crystal structure. This approach was also used to successfully scaffold several small molecule MGBs, and is the first example of using DNA scaffolds to host a guest molecule that is resolved in continuous electron density. As such, this last result validates the foundational goal of DNA nanotechnology—to scaffold molecules and solve their structure—and provides a tantalizing path for future studies with peptides, proteins, or other guest species.

8.5 References

1. Zheng, J. P.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R. J.; Constantinou, P. E.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **2009,** *461* (7260), 74-77.

2. Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* **2016,** *138* (38), 12690-12690.

3. Simmons, C. R.; Zhang, F.; MacCulloch, T.; Fahmi, N.; Stephanopoulos, N.; Liu, Y.; Seeman, N. C.; Yan, H., Tuning the Cavity Size and Chirality of Self-Assembling 3D DNA Crystals. *Journal of the American Chemical Society* **2017,** *139* (32), 11254-11260.

4. Simmons, C. R.; MacCulloch, T.; Zhang, F.; Liu, Y.; Stephanopoulos, N.; Yan, H., A Self-Assembled Rhombohedral DNA Crystal Scaffold with Tunable Cavity Sizes and High-Resolution Structural Detail. *Angewandte Chemie International Edition* **2020,** *n/a* (n/a).

5. Zhang, F.; Simmons, C. R.; Gates, J.; Liu, Y.; Yan, H., Self-Assembly of a 3D DNA Crystal Structure with Rationally Designed Six-Fold Symmetry. *Angewandte Chemie-International Edition* **2018,** *57* (38), 12504-12507.

6. Jung, C.; Bandilla, P.; Von Reutern, M.; Schnepf, M.; Rieder, S.; Unnerstall, U.; Gaul, U., True equilibrium measurement of transcription factor-DNA binding affinities using automated polarization microscopy. *Nature communications* **2018,** *9* (1), 1605-11.

7. Ades, S. E.; Sauer, R. T., Specificity of Minor-Groove and Major-Groove Interactions in a Homeodomain-DNA Complex. *Biochemistry (Easton)* **1995,** *34* (44), 14601-14608.

8. Noyes, M. B.; Christensen, R. G.; Wakabayashi, A.; Stormo, G. D.; Brodsky, M. H.; Wolfe, S. A., Analysis of Homeodomain Specificities Allows the Family-wide Prediction of Preferred Recognition Sites. *Cell (Cambridge)* **2008,** *133* (7), 1277-1289.

9. Zhigang, S.; Veronica Ebu, I.; Haicheng, L.; Lekha, P.; Katrina, M. C.; Tom, C.; Gaetano, T. M.; Cory, A., Design of a "MinimAl" Homeodomain: The N-Terminal Arm Modulates DNA Binding Affinity and Stabilizes Homeodomain Structure. *Proceedings of the National Academy of Sciences - PNAS* **1994,** *91* (18), 8373-8377.

10. Tucker-Kellogg, L.; Rould, M. A.; Chambers, K. A.; Ades, S. E.; Sauer, R. T.; Pabo, C. O., Engrailed (Gln50→Lys) homeodomain–DNA complex at 1.9 Å resolution: structural basis for enhanced affinity and altered specificity. *Structure (London)* **1997,** *5* (8), 1047-1054.

11. Fraenkel, E.; Rould, M. A.; Chambers, K. A.; Pabo, C. O., Engrailed Homeodomain-DNA Complex at 2.2 A Resolution: A Detailed View of the Interface and Comparison with Other Engrailed Structures. *Journal of molecular biology* **1998,** *284* (2), 351-361.

12. Ades, S. E.; Sauer, R. T., Differential DNA-binding specificity of the engrailed homeodomain: The role of residue 50. *Biochemistry (Easton)* **1994,** *33* (31), 9187-9194.

13. Guerrero, L.; Smart, O. S.; Woolley, G. A.; Allemann, R. K., Photocontrol of DNA Binding Specificity of a Miniature Engrailed Homeodomain. *Journal of the American Chemical Society* **2005,** *127* (44), 15624-15629.

14. Delano, W. L. *The PYMOL Molecular Graphics System*, DeLano Scientific: San Carlos, CA, 2002.

15. Chen, C., A Practical Guide for Introducing Gibson Assembly®Gene Constructs into Chemically Competent Cells. *BioTechniques* **2015,** *58* (3).

16. Ghosh, B.; Boila, L. D.; Choudhury, S.; Mondal, P.; Bhattacharjee, S.; Pal, S. K.; Sengupta, A.; Roy, S., A Potent Conformation-Constrained Synthetic Peptide Mimic of a Homeodomain Selectively Regulates Target Genes in Cells. *ACS chemical biology* **2018,** *13* (8), 2003-2009.

17. Tripp, K. W.; Sternke, M.; Majumdar, A.; Barrick, D., Creating a Homeodomain with High Stability and DNA Binding Affinity by Sequence Averaging. *Journal of the American Chemical Society* **2017,** *139* (14), 5051-5060.

18. Bewley, C. A.; Gronenborn, A. M.; Clore, G. M., Minor groove-binding architectural proteins: Structure, function, and DNA recognition. *Annual review of biophysics and biomolecular structure* **1998,** *27* (1), 105-131.

19. Fonfría-Subirós, E.; Acosta-Reyes, F.; Saperas, N.; Pous, J.; Subirana, J. A.; Campos, J. L., Crystal structure of a complex of DNA with one AT-hook of HMGA1. *PloS one* **2012,** *7* (5), e37120-e37120.

20. Huth, J.; Bewley, C.; Nissen, M.; Evans, J.; Reeves, R.; Gronenborn, A.; Clore, G., The solution structure of an HMG-l(Y)-DNA

complex defines a new architectural minor

groove binding motif. *Nature Structural Biology* **1997,** *4*, 657-665.

21. Claus, P.; Schulze, E.; Wisniewski, J., Insect Proteins Homologous to Mammalian High Mobility Group Proteins I/Y. *Journal of biological chemistry* **1994,** *269*, 33042-33048.

22. Zlatanova, J.; Holde, K., Binding to four-way junction DNA: a common property of architectural proteins? *The FASEB journal* **1998,** *12* (6), 421-431.

23. Hill, D. A.; Reeves, R., Competition between HMG-I(Y), HMG-1 and histone H1 on four-way junction DNA. *Nucleic acids research* **1997,** *25* (17), 3523-3531.

24. Alniss, H. Y., Thermodynamics of DNA Minor Groove Binders. *Journal of medicinal chemistry* **2019,** *62* (2), 385-402.

25. Nunn, C. M.; Garman, E.; Neidle, S., Crystal Structure of the DNA Decamer d(CGCAATTGCG) Complexed with the Minor Groove Binding Drug Netropsin. *Biochemistry (Easton)* **1997,** *36* (16), 4792-4799.

26. Lewis, E. A.; Munde, M.; Wang, S.; Rettig, M.; Le, V.; Machha, V.; Wilson, W. D., Complexity in the binding of minor groove agents: netropsin has two thermodynamically different DNA binding modes at a single site. *Nucleic acids research* **2011,** *39* (22), 9649-9658.

27. Khan, G. S.; Shah, A.; Zia ur, R.; Barker, D., Chemistry of DNA minor groove binding agents. *Journal of photochemistry and photobiology. B, Biology* **2012,** *115*, 105-118.

28. Luis, A. M.; Kenneth, J. B., Origins of Netropsin Binding Affinity and Specificity: Correlations of Thermodynamic and Structural Data. *Proceedings of the National Academy of Sciences - PNAS* **1987,** *84* (13), 4359-4363.

29. Padroni, G.; Withers, J. M.; Taladriz-Sender, A.; Reichenbach, L. F.; Parkinson, J. A.; Burley, G. A., Sequence-Selective Minor Groove Recognition of a DNA

Duplex Containing Synthetic Genetic Components. *Journal of the American Chemical Society* **2019,** *141* (24), 9555-9563.

30. Parolin, C.; Zanotti, G.; Palu, G., A Model for the Sequence-Dependent DNA Binding of 4′,6-Diamidino-2-Phenylindole (DAPI). *Biochemical and biophysical research communications* **1995,** *208* (1), 332-338.

31. Wilson, W. D.; Tanious, F. A.; Barton, H. J.; Jones, R. L.; Fox, K.; Wydra, R. L.; Strekowski, L., DNA sequence dependent binding modes of 4',6-diamidino-2-phenylindole (DAPI). *Biochemistry (Easton)* **1990,** *29* (36), 8452-8461.

32. Guo, P.; Farahat, A. A.; Paul, A.; Harika, N. K.; Boykin, D. W.; Wilson, W. D., Compound Shape Effects in Minor Groove Binding Affinity and Specificity for Mixed Sequence DNA. *Journal of the American Chemical Society* **2018,** *140* (44), 14761-14769.

CHAPTER 9

SUMMARY AND OUTLOOK

9.1 DNA-peptide hybrid materials: summary and outlook

DNA makes an ideal building block for the construction of a large variety of nanostructures; however, they often lack functionality and are not always compatible for biological environments.[1-5] One way to overcome these shortcomings is to incorporate peptides to make DNA peptide hybrid materials, by exploiting electrostatic interactions or by conjugating the peptide to a DNA handle. The work carried out in the dissertation demonstrates several possible applications for these DNA-peptide hybrid materials, along with the synthetic approaches taken to produce them.

Chapter 2 outlined a facile way to integrate functional peptide sequences at high densities while also stabilizing the DNA nanostructures in low salt conditions. This was done by utilizing a highly cationic peptide, polylysine, to coat the DNA at a 1:1 nitrogen to phosphate ratio to act as a counterion, as previously demonstrated by the Shih laboratory.[6] Expanding on the simple K10 coating molecule, the endosomal escape peptide (aurein 1.2) was added to both the N- and C-termini allowing for ~48 copies to be displayed from the six-helix bundle (6HB). Upon monitoring the uptake efficiency, kinetics, and final location of three different samples (plain 6HB, K10-6HB, EE-K10-6HB) in three different hepatic cells lines several findings can be reported. The uptake efficiency of the three different samples did not vary, indicating that the coating molecules have no effect on cell delivery. However, there was a distinct correlation between the amount of nanostructure found intercellularly and the size of the cell, with the larger cells containing more. In all three cell lines, only the EE-K10-6HB was able to

achieve cytosolic delivery, while the plain 6HB and K10-6HB both remained in the endosomes (Figure 9.1A). This report was the first instance of a functional peptide being able to facilitate endosomal escape of DNA nanostructures; however, upon adding serum into the media all three samples became trapped in the endosomes. Further studies showed that the functional peptide coating was inactivated due to the formation of a protein corona.

Chapter 3 surveyed several applications of DNA-peptide conjugates along with the synthetic route taken to produce them with one and two DNA handles, as well as a DNA-peptide co-block polymer. In the case with a single DNA handle, three different unique POCs were scaffolded onto a 6HB, simultaneously to mimic the 3 CDR loops traditionally observed in an antibody (Figure 9.1B). Overall, ten POCs were synthesized, each with a different DNA handle sequence conjugated using a copper free click to a novel peptide (containing an AzK) that was discovered to bind to the POI (transferrin) at micromolar affinity by employing a peptide microarray. 120 different combinations with three of the ten POCs were assembled, and their binding affinity was tested using SPR. Unfortunately, only several of the constructs displayed momentary binding at 25 nM.

In order to add two DNA handles to a single peptide, two orthogonal conjugation reactions were required. Both a SPaaC and a CuAAC were used with one DNA handle containing a DBCO moiety and the second containing an azide, both of which were added to the DNA using NHS ester amine couplings. To increase the yield of the second reaction, the DNA handles were designed to be partially complementary to add proximity. The DNA-peptide-DNA (DPD) conjugatess were employed for two different aspects, the first used an enzymatically cleavable peptide sequence to link two different

double crossover tiles that could be cleaved in the presence of a matrix metalloproteinase. The second used a complementary strand to constrain the peptide in a loop-like fashion. This application was used to produce two DPDs, one containing an RGDS peptide and the other containing the sequence PHSRN, both of which are found in the protein fibronectin, and work synergistically for cell adhesion. After incorporating a 5' phosphorylation site on one of the two DPDs, the two were ligated to produce a DPDPD alternating copolymer.

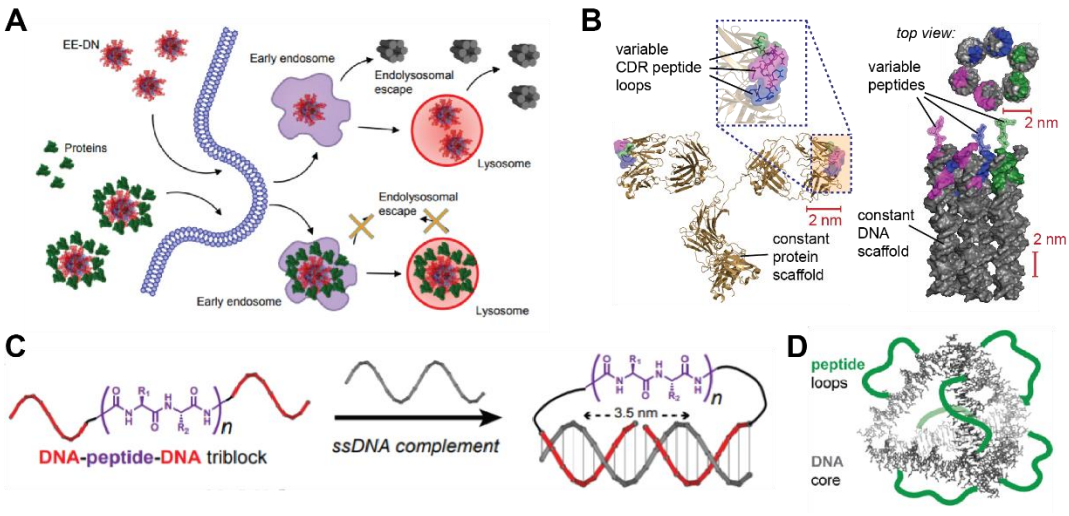


Figure 9.1. DNA-peptide hybrid summary and outlooks. (A) Summary of functionalized coating on 6HB study. (B) Cartoon depiction of a POC based synbody. (C) Schematic of using a DPD to constrain the peptide portion into a loop and (D) A possible future application of using the DNA-peptide copolymers to display multiple peptides loops off of a DNA structure.

The synthetic routes taken to produce the various DNA-peptide hybrid materials have laid the groundwork for future studies. In the case with the electrostatic coatings, the same approach could be utilized to functionalize DNA nanostructures with any desired peptide sequence such as a targeting peptide to help increase uptake efficiency. Additionally, these coating molecules can be modified to incorporate some form of non-fouling polymer to prohibit the formation of a protein corona around the DNA

nanostructures. It is possible that they could be further stabilized by using a crosslinking agent such as glutaraldehyde or incorporating cysteines on the termini to form disulfide bridges.

The DNA-peptide conjugates can further be employed for the same proposed application, to create synbodies, after further optimization of both the conjugation chemistry, linker length, and the shape and size of the DNA nanostructure. Additionally, these POCs, can be used to anchor any desired peptide sequence site, specifically on a nanostructure. By adding a second DNA handle, the peptide portion of the DPD can be used as a way to incorporate functionality to the nanostructure by acting as a loop capable of binding a particular target. This approach offers an advantage over a singular DNA handle as most peptides are shown to have better binding affinity when cyclized. Future studies can use the DPDs to optimize peptide conformation by changing the distance and placement of the two different handles, making the peptide more constrained or more flexible (Figure 9.1C). Furthermore, the synthetic approach of using multiple DPDs to create a templated copolymer via ligation can be extended to produce longer polymers with additional peptides (Figure 9.1D).

9.2 DNA Crystal Scaffolds

The advent of multiple DNA crystal scaffolds is fairly recent, therefore the design rules necessary to produce a desired lattice with desired cavity sizes have not been comprehensively elucidated. However, the work that has been described in the past few chapters shows that significant progress has been made towards understanding an expanding number of design rules that dictate the 3D packing of self-assembled DNA crystals. Several parameters were explored, including the number of interjunction bases,

the sequence of the Holliday junction and its flanking duplexes, and controlling the size of the cavities within the crystal.

The The effect of interjunction distances was explored throughout the entire 4xN series, resulting in different cavity shapes and sizes by increasing the number or repeating bases from five to six, and ultimately to seven.[7-9] The original 4x5 system was crystallized with $P3_221$ symmetry, and had small irregular cavities (Figure 9.2A) after its structure was solved at 3.1 Å. It was initially thought that the aperiodic cavities were solely a result of the torsional strain in the central weaving strand (20 bp) caused by a single base deficiency (21 bp/turn) which would prevent a full 720° rotation. As a result, the addition of an extra base was explored, resulting in the 4x6 system. The additional base did indeed relieve the torsional strain, and regular cavities with larger volumes were obtained (Figure 9.2B) with the overall structure containing $P3_2$ symmetry at 3.1 Å resolution. In addition to controlling the cavity size, the 4x6 system was also explored as a way to control the handedness (D-DNA to L-DNA) of the crystals using the L-DNA enantiomer synthesized from left-handed phosphoramidites. The crystals were also shown to be resistant to nuclease degradation, which could be used to protect any future guest that it might host.

The 4x7 system, which was originally designed with a J1 sequence, successfully crystallized after the immobile junction sequence was changed to J10. This additional interjunction base resulted in a novel rhombohedral symmetry ($R3$) at an unprecedented

resolution (2.7 Å) with cavities that were similar to the 4x6 along the six-fold axis, but densely packed due to the layers no longer being planar (Figure 9.2C)

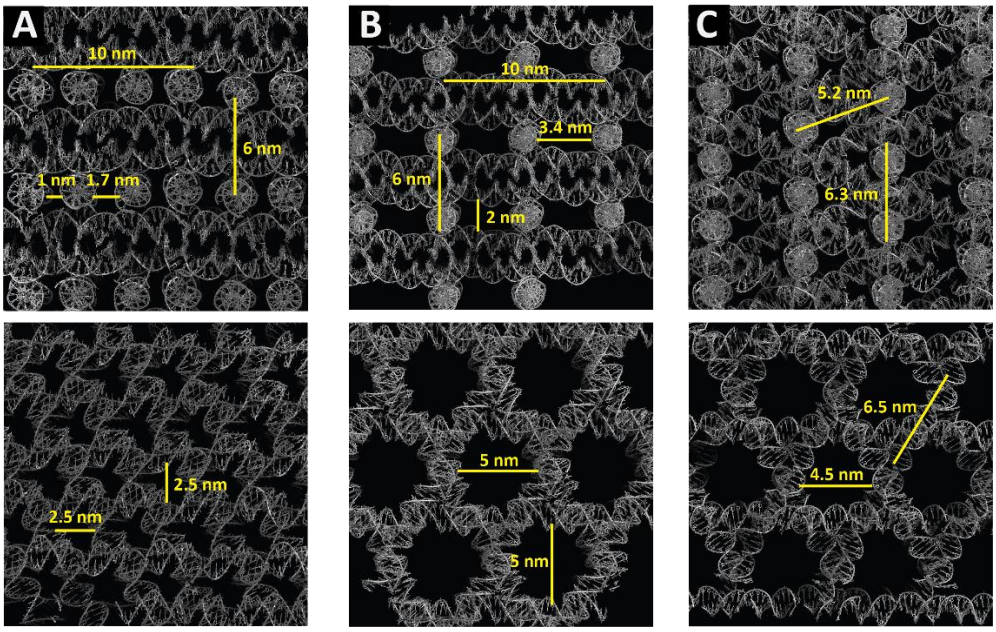


Figure 9.2 Effect of changing interjunction bases. Two views, each rotated 90° from one another to show the different cavity shapes and sizes produced by simply changing interjunction base to produce the 4x5 (A), 4x6 (B) and 4x7 systems (C).

The 4x7 system was used to explore two other design parameters that were hypothesized to affect the crystal lattice, the sticky end sequence and number of helical turns within each duplex. Sticky end lengths of 1 to 3 nucleotides were all explored with varying GC content, with all modifications leading to poor diffraction quality that did not contain *R*3 symmetry, with the exception of the 2bp-TA which modestly improved the resolution to 2.6 Å. Four variants that increased the length of each duplex to contain three turns were explored with both 10 or 11 extra bases being introduced both in between the two HJs and the regions that flank it. The only design that led to quality crystals (4.5 Å) was the 11 -bp flanking region system, which resulted in identical

symmetry to the 2-turn system but with larger cavities, and the expected unit cell parameters. These results indicate that it could be possible to modify any of the known crystal systems with additional helical turns to produce larger cavities amenable to larger guest proteins.

The ability to finally produce reliable 4x7 crystals arose from the change of junction sequence; however, there are 36 possible immobile HJ sequences. Aside from the 4x7, all other published constructs use the J1 sequence in their designs. All 36 junction sequences were screened across both the 4x5 and 4x6 systems to determine whether or not they had any effect on the ability to crystallize, resolution, and packing. In the 4x5 system, seven junctions were deemed "fatal" due to their inability to crystallize, whereas eight retained the same symmetry ($P3_221$) as the original J1 lattice. The remaining 19 surprisingly contained $P3_2$ symmetry with nearly identical cell parameters as the J1 4x6 lattice, indicating that the torsional stress could potentially be relieved by slight perturbations of the junction sequence. Only 18 out of the 38 junctions crystallized in the 4x6 system: 16 as $P3_2$, 5 as $R3$, and three (J5, J31, and J33) in both symmetries. Because such a large number of junctions were fatal in the 4x6 system, an additional "scrambled" version containing unique flanking sequences was explored to ensure that the junction sequence alone prevented crystallization. In the scrambled version, 16 junctions did not crystallize, and the overwhelming majority exhibited $R3$ symmetry, which starkly contrasted with the original sequence having the propensity for $P3_2$. These results strongly indicated that crystallization and resulting symmetry could be dictated by a combination of both Holliday junction and the sequence that flanks it (Figure 9.3).

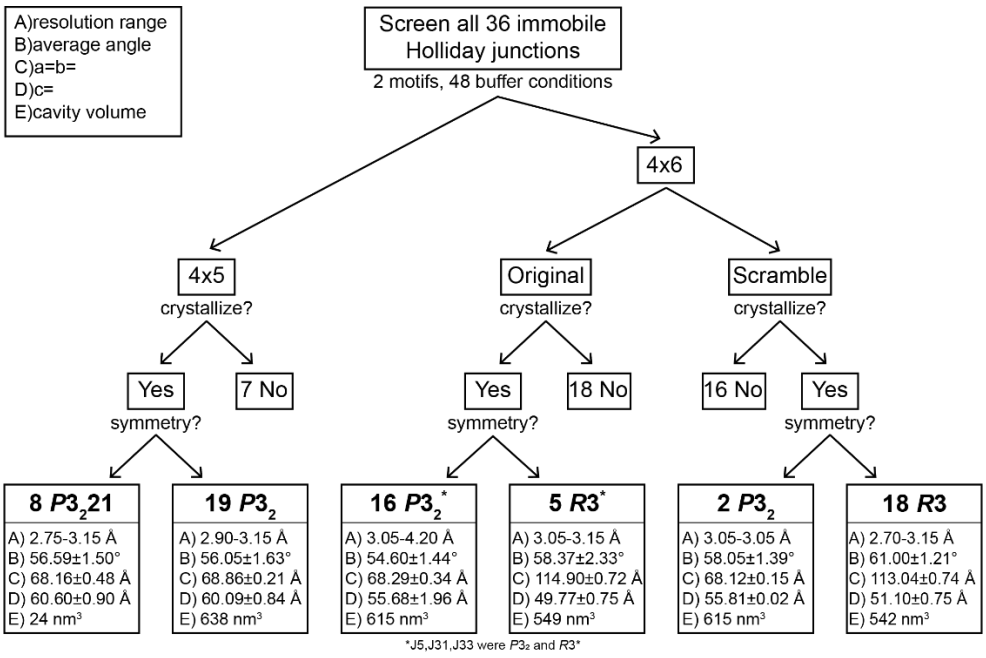


Figure 9.3. Summary of the 36 Holliday junctions in the 4x5 and 4x6 systems. Flowchart summarizing the findings of the HJ study including the number of junctions that crystallized in each system, along with their resulting symmetry. The results for the resolution, junction angles, unit cell dimensions, and cavity volumes are broken down for each system and symmetry.

In addition to examining multiple design parameters of DNA crystal systems, efforts were made towards incorporation of guests molecules for their structural determination. As model guests, several transcription factors were employed because they are known to bind to specific DNA sequences. Three different approaches were explored with each of the target proteins/peptides in parallel; soaking, co-crystallization, and chemical conjugation. In every case, the co-crystals have been the most promising, but complete density resulting from a fully occupied and rigid guest has not been achieved. However, by incorporating small molecule minor groove binders (MGB) such

as netropsin through co-crystallization, electron density maps were obtained as a result of the sequence-specific positioning of the molecule at its intended location.

Because symmetry could result from a combination of junction and flanking sequence, experiments were designed to identify what nucleotides in the flanking sequence were responsible for dictating the switch in symmetry from $P3_2$ to $R3$ in the 4x6 crystals. The J10 sequence was chosen because it retained only the $P3_2$ in the original 4x6 system, but switched to $R3$ in the 4x6 scramble crystals. Each position (number of bases from the junction) was modified by changing a purine to a pyrimidine of the opposite base pair (i.e. A↔C and T↔G) at each site, with six positions on the 5' end of S2 and five positions on the 3' end, resulting in six new sequences (Figure 9.4).

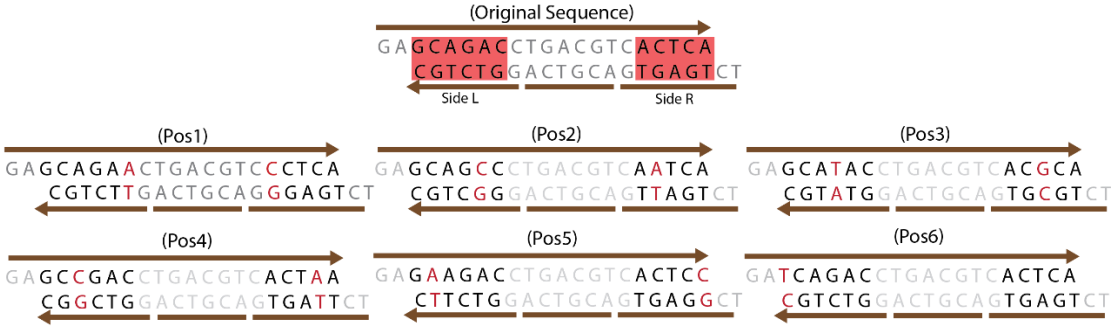


Figure 9.4. Position variations with Pos1 indicated as the base immediately adjacent to the junction. The original sequence has the two flanking regions boxed in red. The six variations with the bases maintained throughout the experiments are grayed out, with the modified bases shown in red.

Each variant diffracted to resolutions ranging from (2.65 to 4 Å), but only the mutations in Pos1 switched symmetry from $P3_2$ to $R3$. To determine whether variations of both sites, or just one or the other influenced symmetry, two additional sequence variants were produced (Figure 9.5A), and it was determined that Pos1 on side L (L1) contained $R3$ symmetry, but R1 did not. In addition, when the nucleotide on S2 was moved to S3 and vice versa on either side (Figure 9.5B), only the left side affected symmetry.

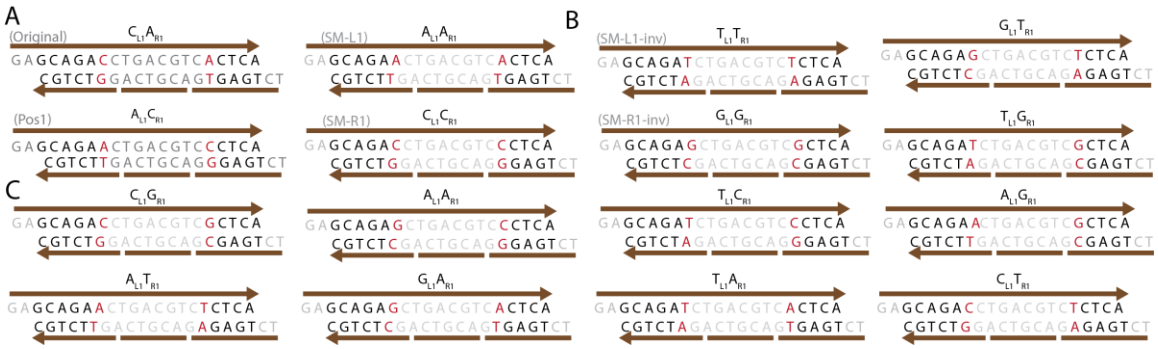


Figure 9.5. All 16 Pos1 variants. All possible base combinations when considering only pos1 (HJ adjacent base) including (A) the original sequences, the pos1 variant and the single mutations of that, the (B) inverted variants of those where the base on S2 is moved to S3, and the (C) remaining 8 options.

There are 16 possible combinations of nucleotides at Pos1. To determine if symmetry was exclusively attributable to the AT base pair at position L1, the 10 remaining possible combinations of nucleotides at Pos1 were explored (Figure 9.5C). To date, 15 of the 16 variants have been measured with eight exhibiting $P3_2$ symmetry, each containing a G or C base at L1, and seven $R3$ with either an A or T. Based upon this, the remaining sequence variant is expected to follow this trend.

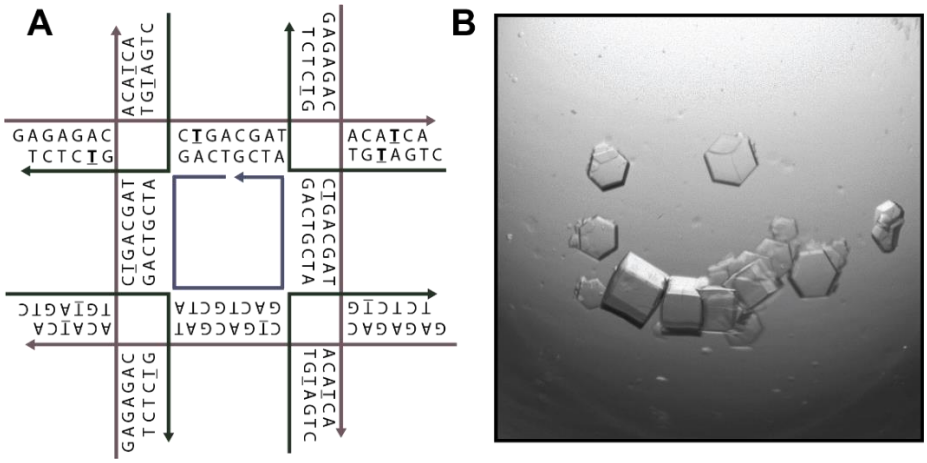


Figure 9.6. 4x8 design and crystals. (A) Overall sequences of the unsolved 4x8 design to finish the 4xN series with 1bp SE sequences with the bromo-dUs used for the initial maps indicated. (B) Representative bright field image of the 4x8 crystals.

The various symmetries observed within the 4x5,[9] 4x6,[8] and 4x7[7] systems demonstrates that changing the number of bases in the repeating sequence of the central weaving strand has a significant impact. One additional design comprised of 4 repeats of 8 bases (4x8) following all the same design parameters as its counterparts is being explored (Figure 9.6A). After screening, large hexagonal prism crystals were produced and diffracted to a 3.1Å resolution with $P6_5$ symmetry (Figure 9.6B). A molecular replacement solution was not possible using any of the other systems, therefore, heavy atoms were required to obtain initial diffraction maps. The sequences were redesigned to incorporate 2 bromo-dUs where thymines originally existed; however, to date, no data with an adequate anomalous signal have been obtained at a high enough resolution to solve the structure.

9.3 Concluding Remarks

The work included in this dissertation demonstrates how DNA nanostructures can be used for a variety of applications. The first few chapters gave an overview how

functionality could be added by expanding the nanostructures to be composed of both DNA and peptides. These hybrid materials were produced by exploiting an electrostatic interaction or through conjugation. The remaining chapters explored both the design rules of self-assembling DNA crystals, as well as the work done to achieve the original goal of DNA nanotechnology. Both of these sub-fields in DNA nanotechnology (hybrid materials and crystals), hold significant potential for a variety of applications, and the work presented here can provide some of the groundwork that can be used as a basis for future research.

9.4 References

1. Seeman, N. C., DNA in a material world. *Nature* **2003,** *421* (6921), 427-431.

2. Seeman, N. C., Biochemistry and structural DNA nanotechnology: An evolving symbiotic relationship. *Biochemistry* **2003,** *42* (24), 7259-7269.

3. Seeman, N. C.; Sleiman, H. F., DNA nanotechnology. *Nature Reviews Materials* **2018,** *3* (1).

4. Pinheiro, A. V.; Han, D. R.; Shih, W. M.; Yan, H., Challenges and opportunities for structural DNA nanotechnology. *Nature Nanotechnology* **2011,** *6* (12), 763-772.

5. Hu, Y.; Niemeyer, C. M., From DNA Nanotechnology to Material Systems Engineering. *Advanced Materials* **2019,** *31* (26).

6. Ponnuswamy, N.; Bastings, M. M. C.; Nathwani, B.; Ryu, J. H.; Chou, L. Y. T.; Vinther, M.; Li, W. A.; Anastassacos, F. M.; Mooney, D. J.; Shih, W. M., Oligolysine-based coating protects DNA nanostructures from low-salt denaturation and nuclease degradation. *Nature Communications* **2017,** *8*.

7. Simmons, C. R.; MacCulloch, T.; Zhang, F.; Liu, Y.; Stephanopoulos, N.; Yan, H., A Self-Assembled Rhombohedral DNA Crystal Scaffold with Tunable Cavity Sizes and High-Resolution Structural Detail. *Angewandte Chemie International Edition* **2020,** *n/a* (n/a).

8. Simmons, C. R.; Zhang, F.; MacCulloch, T.; Fahmi, N.; Stephanopoulos, N.; Liu, Y.; Seeman, N. C.; Yan, H., Tuning the Cavity Size and Chirality of Self-Assembling 3D DNA Crystals. *Journal of the American Chemical Society* **2017,** *139* (32), 11254-11260.

9. Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* **2016,** *138* (38), 12690-12690.

REFERENCES

Chapter 1

1. Feynman, R. P., There's plenty of room at the bottom [data storage]. *Journal of microelectromechanical systems* **1992,** *1* (1), 60-66.

2. Kaiser, D., Richard Feynman: A Life in Science. John Gribbin , Mary Gribbin. *Isis* **2001,** *92* (1), 207-208.

3. Yurchenco, P. D.; Birk, D. E.; Mecham, R. P., *Extracellular matrix assembly and structure*. Academic Press: San Diego, California, 1994.

4. Bishop, K. J. M.; Wilmer, C. E.; Soh, S.; Grzybowski, B. A., Nanoscale Forces and Their Uses in Self-Assembly. *Small (Weinheim an der Bergstrasse, Germany)* **2009,** *5* (14), 1600-1630.

5. Yadav, S.; Sharma, A. K.; Kumar, P., Nanoscale Self-Assembly for Therapeutic Delivery. *Frontiers in bioengineering and biotechnology* **2020,** *8*, 127-127.

6. Ninham, B. W., *Molecular forces and self assembly in colloid, nano sciences and biology*. Cambridge University Press: Cambridge, U.K, 2010.

7. Lee, Y. S., *Self-Assembly and Nanotechnology Systems: Design, Characterization, and Applications*. John Wiley & Sons, Incorporated: New York, 2011.

8. Bensaude-Vincent, B., Self-Assembly, Self-Organization: Nanotechnology and Vitalism. *Nanoethics* **2009,** *3* (1), 31-42.

9. Merrifield, R. B., Solid Phase Synthesis (Nobel Lecture). *Angewandte Chemie (International Edition in English)* **1985,** *24(10),* 799-810.

10. Ito, H.; Ike, Y.; Ikuta, S.; Itakura, K., Solid phase synthesis of polynucleotides. VI. Farther studies on polystyrene copolymers for the solid support. *Nucleic acids research* **1982,** *10* (5), 1755-1769.

11. Eastaugh, B.; Sternal-Johnson, C., Recontruction of M.C.Escher Depth. Exploriment Design, 2011.

12. Pinheiro, A. V.; Han, D. R.; Shih, W. M.; Yan, H., Challenges and opportunities for structural DNA nanotechnology. *Nature Nanotechnology* **2011,** *6* (12), 763-772.

13. Seeman, N. C., DNA Nanotechnology at 40. *Nano letters* **2020,** *20* (3), 1477-1478.

14. Pray, L., Discovery of DNA Structure and Function: Watson and Crick. *Nature Education* **2008,** *1:1(100).*

15. Watson, J. D.; Crick, F. H., Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. J.D. Watson and F.H.C. Crick. Published in Nature, number 4356 April 25, 1953. *Nature (London)* **1974,** *248* (5451), 765-765.

16. Qiu, H.; Dewan, J. C.; Seeman, N. C., A DNA decamer with a sticky end: the crystal structure of d-CGACGATCGT. *Journal of molecular biology* **1997,** *267* (4), 881-898.

17. Seeman, N. C., Biochemistry and structural DNA nanotechnology: An evolving symbiotic relationship. *Biochemistry* **2003,** *42* (24), 7259-7269.

18. Kowalczykowski, S. C.; Dixon, D. A.; Eggleston, A. K.; Lauder, S. D.; Rehrauer, W. M., Biochemistry of homologous recombination in Escherichia coli. *Microbiological reviews* **1994,** *58* (3), 401-465.

19. Holliday, R., A mechanism for gene conversion in fungi (Reprinted). *Genetics Research* **2007,** *89* (5-6), 285-307.

20. Clegg, R. M.; Murchie, A. I. H.; Zechel, A.; Carlberg, C.; Diekmann, S.; Lilley, D. M. J., FLUORESCENCE RESONANCE ENERGY-TRANSFER ANALYSIS OF THE STRUCTURE OF THE 4-WAY DNA JUNCTION. *Biochemistry* **1992,** *31* (20), 4846-4856.

21. Duckett, D. R.; Murchie, A. I. H.; Diekmann, S.; Vonkitzing, E.; Kemper, B.; Lilley, D. M. J., THE STRUCTURE OF THE HOLLIDAY JUNCTION, AND ITS RESOLUTION. *Cell* **1988,** *55* (1), 79-89.

22. Kallenbach, N. R.; Ma, R. I.; Seeman, N. C., AN IMMOBILE NUCLEIC-ACID JUNCTION CONSTRUCTED FROM OLIGONUCLEOTIDES. *Nature* **1983,** *305* (5937), 829-831.

23. Seeman, N. C., NUCLEIC-ACID JUNCTIONS AND LATTICES. *Journal of Theoretical Biology* **1982,** *99* (2), 237-247.

24. Zadegan, R. M.; Norton, M. L., Structural DNA Nanotechnology: From Design to Applications. *International Journal of Molecular Sciences* **2012,** *13* (6), 7149-7162.

25. Zhang, F.; Nangreave, J.; Liu, Y.; Yan, H., Structural DNA Nanotechnology: State of the Art and Future Perspective. *Journal of the American Chemical Society* **2014,** *136* (32), 11198-11211.

26. Ke, Y.; Castro, C.; Choi, J. H., Structural DNA Nanotechnology: Artificial Nanostructures for Biomedical Research. *Annual review of biomedical engineering* **2018,** *20* (1), 375-401.

27. Zhang, C. A.; He, Y.; Su, M.; Ko, S. H.; Ye, T.; Leng, Y. J.; Sun, X. P.; Ribbe, A. E.; Jiang, W.; Mao, C. D., DNA self-assembly: from 2D to 3D. *Faraday Discussions* **2009,** *143*, 221-233.

28. Zhang, F.; Jiang, S. X.; Wu, S. Y.; Li, Y. L.; Mao, C. D.; Liu, Y.; Yan, H., Complex wireframe DNA origami nanostructures with multi-arm junction vertices. *Nature Nanotechnology* **2015,** *10* (9), 779-+.

29. He, Y.; Tian, Y.; Ribbe, A. E.; Mao, C. D., Highly connected two-dimensional crystals of DNA six-point-stars. *Journal of the American Chemical Society* **2006,** *128* (50), 15978-15979.

30. Ma Ri Fau - Kallenbach, N. R.; Kallenbach Nr Fau - Sheardy, R. D.; Sheardy Rd Fau - Petrillo, M. L.; Petrillo Ml Fau - Seeman, N. C.; Seeman, N. C., Three-arm nucleic acid junctions are flexible. (0305-1048 (Print)).

31. Seeman, N. C., DNA nanotechnology: Novel DNA constructions. *Annual Review of Biophysics and Biomolecular Structure* **1998,** *27*, 225-248.

32. Chen, J. H.; Seeman, N. C., SYNTHESIS FROM DNA OF A MOLECULE WITH THE CONNECTIVITY OF A CUBE. *Nature* **1991,** *350* (6319), 631-633.

33. Fu, T. J.; Seeman, N. C., DNA DOUBLE-CROSSOVER MOLECULES. *Biochemistry* **1993,** *32* (13), 3211-3220.

34. Winfree, E.; Liu, F. R.; Wenzler, L. A.; Seeman, N. C., Design and self-assembly of two-dimensional DNA crystals. *Nature* **1998,** *394* (6693), 539-544.

35. Rothemund, P. W. K., Folding DNA to create nanoscale shapes and patterns. *Nature* **2006,** *440* (7082), 297-302.

36. Douglas, S. M.; Dietz, H.; Liedl, T.; Hogberg, B.; Graf, F.; Shih, W. M., Self-assembly of DNA into nanoscale three-dimensional shapes (vol 459, pg 414, 2009). *Nature* **2009,** *459* (7250), 1154-1154.

37. Douglas, S. M.; Dietz, H.; Liedl, T.; Hogberg, B.; Graf, F.; Shih, W. M., Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **2009,** *459* (7245), 414-418.

38. Han, D. R.; Pal, S.; Nangreave, J.; Deng, Z. T.; Liu, Y.; Yan, H., DNA Origami with Complex Curvatures in Three-Dimensional Space. *Science* **2011,** *332* (6027), 342-346.

39. Douglas, S. M.; Bachelet, I.; Church, G. M., A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science* **2012,** *335* (6070), 831-834.

40. Ke, Y. G.; Ong, L. L.; Shih, W. M.; Yin, P., Three-Dimensional Structures Self-Assembled from DNA Bricks. *Science* **2012,** *338* (6111), 1177-1183.

41. Zheng, J. P.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R. J.; Constantinou, P. E.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **2009,** *461* (7260), 74-77.

42. Dongran, H.; Suchetan, P.; Jeanette, N.; Zhengtao, D.; Yan, L.; Hao, Y., DNA Origami with Complex Curvatures in Three-Dimensional Space. *Science (American Association for the Advancement of Science)* **2011,** *332* (6027), 342-346.

43. Seeman, N. C., DNA engineering and its application to nanotechnology. *Trends in Biotechnology* **1999,** *17* (11), 437-443.

44. Seeman, N. C.; Sleiman, H. F., DNA nanotechnology. *Nature Reviews Materials* **2018,** *3* (1).

45. Zahid, M.; Kim, B.; Hussain, R.; Amin, R.; Park, S. H., DNA nanotechnology: a future perspective. *Nanoscale research letters* **2013,** *8* (1), 1-13.

46. Aldaye, F. A.; Palmer, A. L.; Sleiman, H. F., Assembling materials with DNA as the guide. *Science* **2008,** *321* (5897), 1795-1799.

47. Aldaye, F. A.; Sleiman, H. F., Supramolecular DNA nanotechnology. *Pure and Applied Chemistry* **2009,** *81* (12), 2157-2181.

48. Zhang, Q.; Jiang, Q.; Li, N.; Dai, L. R.; Liu, Q.; Song, L. L.; Wang, J. Y.; Li, Y. Q.; Tian, J.; Ding, B. Q.; Du, Y., DNA Origami as an In Vivo Drug Delivery Vehicle for Cancer Therapy. *Acs Nano* **2014,** *8* (7), 6633-6643.

49. Madhanagopal, B. R.; Zhang, S. Q.; Demirel, E.; Wady, H.; Chandrasekaran, A. R., DNA Nanocarriers: Programmed to Deliver. *Trends in Biochemical Sciences* **2018,** *43* (12), 997-1013.

50. Hu, Y.; Niemeyer, C. M., From DNA Nanotechnology to Material Systems Engineering. *Advanced Materials* **2019,** *31* (26).

51. Liu, M.; Fu, J.; Hejesen, C.; Yang, Y.; Woodbury, N. W.; Gothelf, K.; Liu, Y.; Yan, H., A DNA tweezer-actuated enzyme nanoreactor. *Nature communications* **2013,** *4* (1), 2127-2127.

52. Liu, M.; Fu, J.; Qi, X.; Wootten, S.; Woodbury, N. W.; Liu, Y.; Yan, H., A Three-Enzyme Pathway with an Optimised Geometric Arrangement to Facilitate Substrate Transfer. *Chembiochem : a European journal of chemical biology* **2016,** *17* (12), 1097-1101.

53. Yang, Y.; Wang, J.; Shigematsu, H.; Xu, W.; Shih, W. M.; Rothman, J. E.; Lin, C., Self-assembly of size-controlled liposomes on DNA nanotemplates. *Nature chemistry* **2016,** *8* (5), 476-483.

54. Fröhlich, E., The role of surface charge in cellular uptake and cytotoxicity of medical nanoparticles. *International journal of nanomedicine* **2012,** *7*, 5577-5591.

55. Gradišar, H.; Božič, S.; Doles, T.; Vengust, D.; Hafner-Bratkovič, I.; Mertelj, A.; Webb, B.; Šali, A.; Klavžar, S.; Jerala, R., Design of a single-chain polypeptide tetrahedron assembled from coiled-coil segments. *Nature chemical biology* **2013,** *9* (6), 362-366.

56. Habibi, N.; Kamaly, N.; Memic, A.; Shafiee, H., Self-assembled peptide-based nanostructures: Smart nanomaterials toward targeted drug delivery. *Nano today* **2016,** *11* (1), 41-60.

57. Jennifer, E. P.; Christos, C.; Todd, O. Y., Nanohedra: Using Symmetry to Design Self Assembling Protein Cages, Layers, Crystals, and Filaments. *Proceedings of the National Academy of Sciences - PNAS* **2001,** *98* (5), 2217-2221.

58. Doles, T.; Božič, S.; Gradišar, H.; Jerala, R., Functional self-assembling polypeptide bionanomaterials. *Biochemical Society transactions* **2012,** *40* (4), 629-634.

59. Vassilev, T. L.; Kazatchkine, M. D.; Van Huyen, J. P. D.; Mekrache, M.; Bonnin, E.; Mani, J. C.; Lecroubier, C.; Korinth, D.; Baruch, D.; Schriever, F.; Kaveri, S. V., Inhibition of cell adhesion by antibodies to Arg-Gly-Asp (RGD) in normal immunoglobulin for therapeutic use (intravenous immunoglobulin, IVIg). *Blood* **1999,** *93* (11), 3624-3631.

60. Van Agthoven, J. F.; Xiong, J.-P.; Alonso, J. L.; Rui, X.; Adair, B. D.; Goodman, S. L.; Arnaout, M. A., Structural basis for pure antagonism of integrin αVβ3 by a high-affinity form of fibronectin. *Nature structural & molecular biology* **2014,** *21* (4), 383-388.

61. Assa-Munt, N.; Jia, X.; Laakkonen, P.; Ruoslahti, E., Solution structures and integrin binding activities of an RGD peptide with two isomers. *Biochemistry* **2001,** *40* (8), 2373-2378.

62. Stephanopoulos, N.; Freeman, R.; North, H. A.; Sur, S.; Jeong, S. J.; Tantakitti, F.; Kessler, J. A.; Stupp, S. I., Bioactive DNA-Peptide Nanotubes Enhance the Differentiation of Neural Stem Cells Into Neurons. *Nano Letters* **2015,** *15* (1), 603-609.

63. Freeman, R.; Stephanopoulos, N.; Alvarez, Z.; Lewis, J. A.; Sur, S.; Serrano, C. M.; Boekhoven, J.; Lee, S. S.; Stupp, S. I., Instructing cells with programmable peptide DNA hybrids. *Nature Communications* **2017,** *8*.

64. Liskamp, R.; Rijkers, D.; Bakker, S., *Bioactive Macrocyclic Peptides and Peptide Mimics*. Wiley-VCH: 2008.

65. Finke, A.; Busskamp, H.; Manea, M.; Marx, A., Designer Extracellular Matrix Based on DNA-Peptide Networks Generated by Polymerase Chain Reaction. *Angewandte Chemie-International Edition* **2016,** *55* (34), 10136-10140.

66. Ruoslahti, E., Fibronectin and its receptors. *Annual Review of Biochemistry* **1988,** *57*, 375-413.

67. Bogdanowich-Knipp, S. J.; Jois, D. S. S.; Siahaan, T. J., The effect of conformation on the solution stability of linear vs. cyclic RGD peptides. *Journal of Peptide Research* **1999,** *53* (5), 523-529.

68. Kapp, T. G.; Rechenmacher, F.; Neubauer, S.; Maltsev, O. V.; Cavalcanti-Adam, E. A.; Zarka, R.; Reuning, U.; Notni, J.; Wester, H.-J.; Mas-Moruno, C.; Spatz, J.; Geiger, B.; Kessler, H., A comprehensive evaluation of the activity and selectivity profile of ligands for RGD-binding integrins. *Scientific reports* **2017,** *7* (1), 39805-39805.

69. Roxin, A.; Zheng, G., Flexible or fixed: a comparative review of

linear and cyclic cancer-targeting peptides. *Future Med. Chem.* **2012,** *4(12)*, 1601-1618.

70. Patel, P. R.; Kiser, R. C.; Lu, Y. Y.; Fong, E.; Ho, W. C.; Tirrell, D. A.; Grubbs, R. H., Synthesis and Cell Adhesive Properties of Linear and Cyclic RGD Functionalized Polynorbornene Thin Films. *Biomacromolecules* **2012,** *13* (8), 2546-2553.

71. Koivunen, E.; Wang, B.; Ruoslahti, E., Phage Libraries Displaying Cyclic Peptides with Different Ring Sizes: Ligand Specificities of the RGD-Directed Integrins. *Nature biotechnology* **1995,** *13* (3), 265-270.

72. Subik, K.; Lee, J.-F.; Baxter, L.; Strzepek, T.; Costello, D.; Crowley, P.; Xing, L.; Hung, M.-C.; Bonfiglio, T.; Hicks, D. G.; Tang, P., The expression patterns of ER, PR, HER2, CK5/6, EGFR, KI-67 and AR by immunohistochemical analysis in breast cancer cell lines. *Breast cancer : basic and clinical research* **2010,** *4* (1), 35-41.

73. Geng, L.; Wang, Z.; Jia, X.; Han, Q.; Xiang, Z.; Li, D.; Yang, X.; Zhang, D.; Bu, X.; Wang, W.; Hu, Z.; Fang, Q., HER2 targeting peptides screening and applications in tumor imaging and drug delivery. *Theranostics* **2016,** *6* (8), 1261-1273.

74. Tian, Y.; Zhou, M. X.; Shi, H. G.; Gao, S. J.; Xie, G. C.; Zhu, M.; Wu, M.; Chen, J.; Niu, Z. W., Integration of Cell-Penetrating Peptides with Rod-like Bionanoparticles: Virus-Inspired Gene-Silencing Technology. *Nano Letters* **2018,** *18* (9), 5453-5460.

75. Lebleu, B.; Moulton, H. M.; Abes, R.; Ivanova, G. D.; Abes, S.; Stein, D. A.; Iversen, P. L.; Arzumanov, A. A.; Gait, M. J., Cell penetrating peptide conjugates of steric block oligonucleotides. *Advanced Drug Delivery Reviews* **2008,** *60* (4-5), 517-529.

76. Vives, E.; Schmidt, J.; Pelegrin, A., Cell-penetrating and cell-targeting peptides in drug delivery. *Biochimica Et Biophysica Acta-Reviews on Cancer* **2008,** *1786* (2), 126-138.

77. Wright, L.; Rothbard, J.; Wender, P., Guanidinium rich peptide transporters and drug delivery. *Current Protein Pept Sci* **2003,** *4(2)*, 105-124.

78. El-Sayed, A.; Futaki, S.; Harashima, H., Delivery of Macromolecules Using Arginine-Rich Cell-Penetrating Peptides: Ways to Overcome Endosomal Entrapment. *Aaps Journal* **2009,** *11* (1), 13-22.

79. Neundorf, I.; Rennert, R.; Hoyer, J.; Schramm, F.; Löbner, K.; Kitanovic, I.; Wölfl, S., Fusion of a short HA2-derived peptide sequence to cell-penetrating peptides improves cytosolic uptake, but enhances cytotoxic activity. *Pharmaceuticals (Basel, Switzerland)* **2009,** *2* (2), 49-65.

80. Sugita, T.; Yoshikawa, T.; Mukai, Y.; Yamanada, N.; Imai, S.; Nagano, K.; Yoshida, Y.; Shibata, H.; Yoshioka, Y.; Nakagawa, S.; Kamada, H.; Tsunoda, S. i.; Tsutsumi, Y., Comparative study on transduction and toxicity of protein transduction domains. *British journal of pharmacology* **2008,** *153* (6), 1143-1152.

81. Li, M.; Tao, Y.; Shu, Y. L.; LaRochelle, J. R.; Steinauer, A.; Thompson, D.; Schepartz, A.; Chen, Z. Y.; Liu, D. R., Discovery and Characterization of a Peptide That Enhances Endosomal Escape of Delivered Proteins in Vitro and in Vivo. *Journal of the American Chemical Society* **2015,** *137* (44), 14084-14093.

82. Strohl, W. R., Current progress in innovative engineered antibodies. *Protein & cell* **2018,** *9* (1), 86-120.

83. Williams, B. A. R.; Diehnelt, C. W.; Belcher, P.; Greving, M.; Woodbury, N. W.; Johnston, S. A.; Chaput, J. C., Creating Protein Affinity Reagents by Combining Peptide Ligands on Synthetic DNA Scaffolds. *Journal of the American Chemical Society* **2009,** *131* (47), 17233-17241.

84. Gupta, N.; Lainson, J. C.; Belcher, P. E.; Shen, L. H.; Mason, H. S.; Johnston, S. A.; Diehnelt, C. W., Cross-Reactive Synbody Affinity Ligands for Capturing Diverse Noroviruses. *Analytical Chemistry* **2017,** *89* (13), 7174-7181.

85. Kye, M.; Lim, Y. B., Reciprocal Self-Assembly of Peptide-DNA Conjugates into a Programmable Sub-10-nm Supramolecular Deoxyribonucleoprotein. *Angewandte Chemie-International Edition* **2016,** *55* (39), 12003-12007.

86. Lou, C. G.; Martos-Maldonado, M. C.; Madsen, C. S.; Thomsen, R. P.; Midtgaard, S. R.; Christensen, N. J.; Kjems, J.; Thulstrup, P. W.; Wengel, J.; Jensen, K. J., Peptide-oligonucleotide conjugates as nanoscale building blocks for assembly of an artificial three-helix protein mimic. *Nature Communications* **2016,** *7*.

87. Patutina, O. A.; Bichenkova, E. V.; Miroshnichenko, S. K.; Mironova, N. L.; Trivoluzzi, L. T.; Burusco, K. K.; Bryce, R. A.; Vlassov, V. V.; Zenkova, M. A., miRNases: Novel peptide-oligonucleotide bioconjugates that silence miR-21 in lymphosarcoma cells. *Biomaterials* **2017,** *122*, 163-178.

88. Zubin, E. M.; Romanova, E. A.; Volkov, E. M.; Tashlitsky, V. N.; Korshunova, G. A.; Shabarova, Z. A.; Oretskaya, T. S., Oligonucleotide-peptide conjugates as potential antisense agents. *Febs Letters* **1999,** *456* (1), 59-62.

89. Arar, K.; Aubertin, A. M.; Roche, A. C.; Monsigny, M.; Mayer, R., SYNTHESIS AND ANTIVIRAL ACTIVITY OF PEPTIDE-OLIGONUCLEOTIDE CONJUGATES PREPARED BY USING N-ALPHA-(BROMOACETYL)PEPTIDES. *Bioconjugate Chemistry* **1995,** *6* (5), 573-577.

90. Astakhova, K.; Ray, R.; Taskova, M.; Uhd, J.; Carstens, A.; Morris, K., "Clicking" Gene Therapeutics: A Successful Union of Chemistry and Biomedicine for New Solutions. *Molecular pharmaceutics* **2018,** *15* (8), 2892-2899.

91. Spruijt, E.; Tusk, S. E.; Bayley, H., DNA scaffolds support stable and uniform peptide nanopores. *Nature Nanotechnology* **2018,** *13* (8), 739-+.

92. Bruick, R. K.; Dawson, P. E.; Kent, S. B.; Usman, N.; Joyce, G. F., Template-directed ligation of peptides to oligonucleotides. *Chemistry & Biology* **1996,** *3* (1), 49-56.

93. Dovgan, I.; Koniev, O.; Kolodych, S.; Wagner, A., Antibody–Oligonucleotide Conjugates as Therapeutic, Imaging, and Detection Agents. *Bioconjugate chemistry* **2019,** *30* (10), 2483-2501.

94. Stephanopoulos, N.; Ortony, J. H.; Stupp, S. I., Self-assembly for the synthesis of functional biomaterials. *Acta Materialia* **2013,** *61* (3), 912-930.

95. Stephanopoulos, N., Peptide–Oligonucleotide Hybrid Molecules for Bioactive Nanomaterials. *Bioconjugate chemistry* **2019,** *30* (7), 1915-1922.

96. Nazari, Z. E.; Herrero, J. G.; Fojan, P.; Gurevich, L., Formation of Conductive DNA-Based Nanowires via Conjugation of dsDNA with Cationic Peptide. *Nanomaterials* **2017,** *7* (6).

97. Tung, C.-H.; Stein, S., Preparation and Applications of Peptide−Oligonucleotide Conjugates. *Bioconjugate chemistry* **2000,** *11* (5), 605-618.

98. Charbgoo, F.; Alibolandi, M.; Taghdisi, S. M.; Abnous, K.; Soltani, F.; Ramezani, M., MUC1 aptamer-targeted DNA micelles for dual tumor therapy using doxorubicin and KLA peptide. *Nanomedicine-Nanotechnology Biology and Medicine* **2018,** *14* (3), 685-697.

99. Ponnuswamy, N.; Bastings, M. M. C.; Nathwani, B.; Ryu, J. H.; Chou, L. Y. T.; Vinther, M.; Li, W. A.; Anastassacos, F. M.; Mooney, D. J.; Shih, W. M., Oligolysine-based coating protects DNA nanostructures from low-salt denaturation and nuclease degradation. *Nature Communications* **2017,** *8*.

100. Wang, S.-T.; Gray, M. A.; Xuan, S.; Lin, Y.; Byrnes, J.; Nguyen, A. I.; Todorova, N.; Stevens, M. M.; Bertozzi, C. R.; Zuckermann, R. N.; Gang, O., DNA origami protection and molecular interfacing through engineered sequence-defined peptoids. *Proceedings of the National Academy of Sciences - PNAS* **2020,** *117* (12), 6339-6348.

101. Flory, J. D.; Shinde, S.; Lin, S.; Liu, Y.; Yan, H.; Ghirlanda, G.; Fromme, P., PNA-Peptide Assembly in a 3D DNA Nanocage at Room Temperature. *Journal of the American Chemical Society* **2013,** *135* (18), 6985-6993.

102. Janssen, B. M. G.; Lempens, E. H. M.; Olijve, L. L. C.; Voets, I. K.; Dongen, v. J. L. J.; Greef, d. T. F. A.; Merkx, M., Reversible blocking of antibodies using bivalent peptide-DNA conjugates allows protease-activatable targeting. *Chemical science (Cambridge)* **2013,** *4* (4), 1442-1450.

103. Marczynke, M.; Gröger, K.; Seitz, O., Selective Binders of the Tandem Src Homology 2 Domains in Syk and Zap70 Protein Kinases by DNA-Programmed Spatial Screening. *Bioconjugate chemistry* **2017,** *28* (9), 2384-2392.

104. Eberhard, H.; Diezmann, F.; Seitz, O., DNA as a Molecular Ruler: Interrogation of a Tandem SH2 Domain with Self-Assembled, Bivalent DNA–Peptide Complexes. *Angewandte Chemie (International ed.)* **2011,** *50* (18), 4146-4150.

105. MacCulloch, T.; Buchberger, A.; Stephanopoulos, N., Emerging applications of peptide-oligonucleotide conjugates: bioactive scaffolds, self-assembling systems, and hybrid nanomaterials. *Organic & Biomolecular Chemistry* **2019,** *17* (7), 1668-1682.

106. Jiang, T.; Meyer, T. A.; Modlin, C.; Zuo, X.; Conticello, V. P.; Ke, Y., Structurally Ordered Nanowire Formation from Co-Assembly of DNA Origami and Collagen-Mimetic Peptides. *Journal of the American Chemical Society* **2017,** *139* (40), 14025-14028.

107. Anastassacos, F. M.; Zhao, Z.; Zeng, Y.; Shih, W. M., Glutaraldehyde Cross-Linking of Oligolysines Coating DNA Origami Greatly Reduces Susceptibility to Nuclease Degradation. *Journal of the American Chemical Society* **2020,** *142* (7), 3311-3315.

108. Roodhuizen, J. A. L.; Hendrikx, P.; Hilbers, P. A. J.; de Greef, T. F. A.; Markvoort, A. J., Counterion-Dependent Mechanisms of DNA Origami Nanostructure Stabilization Revealed by Atomistic Molecular Simulation. *Acs Nano* **2019,** *13* (9), 10798-10809.

109. Pazos, E.; Mosquera, J.; Vazquez, E.; Mascarenas, J., DNA Recognition by Synthetic Constructs. *ChemBioChem* **2011,** *12*, 1958-1973.

110. Bergmann, F.; Bannwarth, W., Solid phase synthesis of directly linked peptide-oligodeoxynucleotide hybrids using standard synthesis protocols. *Tetrahedron letters* **1995,** *36* (11), 1839-1842.

111. Eritja, R.; Pons, A.; Escarceller, M.; Giralt, E.; Albericio, F., SYNTHESIS OF DEFINED PEPTIDE-OLIGONUCLEOTIDE HYBRIDS CONTAINING A NUCLEAR TRANSPORT SIGNAL SEQUENCE. *Tetrahedron* **1991,** *47* (24), 4113-4120.

112. Lu, K.; Duan, Q. P.; Ma, L.; Zhao, D. X., Chemical Strategies for the Synthesis of Peptide-Oligonucleotide Conjugates. *Bioconjugate Chemistry* **2010,** *21* (2), 187-202.

113. Kye, M.; Lim, Y. B., Synthesis and purification of self-assembling peptide-oligonucleotide conjugates by solid-phase peptide fragment condensation. *Journal of Peptide Science* **2018,** *24* (8-9).

114. Gogoi, K.; Mane, M. V.; Kunte, S. S.; Kumar, V. A., A versatile method for the preparation of conjugates of peptides with DNA/PNA/analog by employing chemo-selective click reaction in water. *Nucleic Acids Research* **2007,** *35* (21).

115. Kong, D. H.; Lei, Y.; Yeung, W.; Hili, R., Enzymatic Synthesis of Sequence-Defined Synthetic Nucleic Acid Polymers with Diverse Functional Groups. *Angewandte Chemie-International Edition* **2016,** *55* (42), 13164-13168.

116. Williams, B. A. R.; Lund, K.; Liu, Y.; Yan, H.; Chaput, J. C., Self-assembled peptide nanoarrays: An approach to studying protein-protein interactions. *Angewandte Chemie-International Edition* **2007,** *46* (17), 3051-3054.

117. Harrison, J. G.; Balasubramanian, S., Synthesis and hybridization analysis of a small library of peptide-oligonucleotide conjugates. *Nucleic Acids Research* **1998,** *26* (13), 3136-3145.

118. Merg, A. D.; Thaner, R. V.; Mokashi-Punekar, S.; Nguyen, S. T.; Rosi, N. L., Triblock peptide-oligonucleotide chimeras (POCs): Programmable biomolecules for the assembly of morphologically tunable and responsive hybrid materials. *Chemical communications (Cambridge, England)* **2017,** *53* (90), 12221-12224.

119. Abraham, J. N.; Gour, N.; Bolisetty, S.; Mezzenga, R.; Nardin, C., Controlled aggregation of peptide-DNA hybrids into amyloid-like fibrils. *European Polymer Journal* **2015,** *65*, 268-275.

120. Carter, J.; LaBean, T., Coupling Strategies for the Synthesis of Peptide-Oligonucleotide Conjugates for Patterned Synthetic Biomineralization. *Journal of Nucleic Acids* **2011**.

121. Hein, C. D.; Liu, X.-M.; Wang, D., Click Chemistry, A Powerful Tool for Pharmaceutical Sciences. *Pharmaceutical research* **2008,** *25* (10), 2216-2230.

122. Buchberger, A.; Simmons, C. R.; Fahmi, N. E.; Freeman, R.; Stephanopoulos, N., Hierarchical Assembly of Nucleic Acid/Coiled-Coil Peptide Nanostructures. *Journal of the American Chemical Society* **2020,** *142* (3), 1406-1416.

123. Forget, D.; Boturyn, D.; Defrancq, E.; Lhomme, J.; Dumy, P., Highly efficient synthesis of peptide-oligonucleotide conjugates: Chemoselective oxime and thiazolidine formation. *Chemistry-a European Journal* **2001,** *7* (18), 3976-3984.

124. Zatsepin, T. S.; Stetsenko, D. A.; Arzumanov, A. A.; Romanova, E. A.; Gait, M. J.; Oretskaya, T. S., Synthesis of peptide-oligonucleotide conjugates with single and multiple peptides attached to 2 '-aldehydes through thiazolidine, oxime, and hydrazine linkages. *Bioconjugate Chemistry* **2002,** *13* (4), 822-830.

125. Serva, S.; Lagunavicius, A., Direct Conjugation of Peptides and 5-Hydroxymethylcytosine in DNA. *Bioconjugate Chemistry* **2015,** *26* (6), 1008-1012.

126. Rogers, F. A.; Manoharan, M.; Rabinovitch, P.; Ward, D. C.; Glazer, P. M., Peptide conjugates for chromosomal gene targeting by triplex-forming oligonucleotides. *Nucleic Acids Research* **2004,** *32* (22), 6595-6604.

127. Stanojevic, D.; Verdine, G., Deconstruction of GCN4/GCRE into a monomeric peptide-DNA complex. *Nat Struct Mol Biol* **1995,** *2*, 450-457.

128. Chotera, A.; Sadihov, H.; Cohen-Luria, R.; Monnard, P. A.; Ashkenasy, G., Functional Assemblies Emerging in Complex Mixtures of Peptides and Nucleic Acid-Peptide Chimeras. *Chemistry-a European Journal* **2018,** *24* (40), 10128-10135.

129. Pazos, E.; Portela, C.; Penas, C.; Vázquez, M. E.; Mascareñas, J. L., Peptide-DNA conjugates as tailored bivalent binders of the oncoprotein c-Jun. *Organic & biomolecular chemistry* **2015,** *13* (19), 5385-5390.

130. Taskova, M.; Mantsiou, A.; Astakhova, K., Synthetic Nucleic Acid Analogues in Gene Therapy: An Update for Peptide-Oligonucleotide Conjugates. *Chembiochem* **2017,** *18* (17), 1671-1682.

Chapter 2

1. Bastings, M. M. C.; Anastassacos, F. M.; Ponnuswamy, N.; Leifer, F. G.; Cuneo, G.; Lin, C. X.; Ingber, D. E.; Ryu, J. H.; Shih, W. M., Modulation of the

Cellular Uptake of DNA Origami through Control over Mass and Shape. *Nano Letters* **2018,** *18* (6), 3557-3564.

2. Raniolo, S.; Croce, S.; Thomsen, R. P.; Okholm, A. H.; Unida, V.; Iacovelli, F.; Manetto, A.; Kjems, J.; Desideri, A.; Biocca, S., Cellular uptake of covalent and non-covalent DNA nanostructures with different sizes and geometries. *Nanoscale* **2019,** *11* (22), 10808-10818.

3. Mishra, S.; Feng, Y.; Endo, M.; Sugiyama, H., Advances in DNA Origami–Cell Interfaces. *ChemBioChem* **2019,** *21* (1-2), 33-44.

4. Stephanopoulos, N., Strategies for Stabilizing DNA Nanostructures to Biological Conditions. *ChemBioChem* **2019,** *20* (17), 2191-2197.

5. Shi, J. J.; Kantoff, P. W.; Wooster, R.; Farokhzad, O. C., Cancer nanomedicine: progress, challenges and opportunities. *Nature Reviews Cancer* **2017,** *17* (1), 20-37.

6. Ahrens, E. T.; Bulte, J. W. M., Tracking immune cells in vivo using magnetic resonance imaging. *Nature Reviews Immunology* **2013,** *13* (10), 755-763.

7. Lammers, T.; Ferrari, M., The success of nanomedicine. *Nano Today* **2020,** *31*.

8. Pelaz, B.; Alexiou, C. H.; Alvarez -Puebla, R. A.; Alves, F.; Andrews, A. M.; Ashraf, S.; Balogh, L. P.; Ballerini, L.; Bestetti, A.; Brendel, C.; Bosi, S.; Carril, M.; Chan, W. C. W.; Chen, C. Y.; Chen, X. D.; Chen, X. Y.; Cheng, Z.; Cui, D. X.; Du, J. Z.; Dullin, C.; Escudero, A.; Feliu, N.; Gao, M. Y.; George, M.; Gogotsi, Y.; Grunweller, A.; Gu, Z. W.; Halas, N. J.; Hampp, N.; Hartmann, R. K.; Hersam, M. C.; Hunziker, P.; Jian, J.; Jiang, X. Y.; Jungebluth, P.; Kadhiresan, P.; Kataoka, K.; Khademhosseini, A.; Kopecek, J.; Kotov, N. A.; Krug, H. F.; Lee, D. S.; Lehr, C. M.; Leong, K. W.; Liang, X. J.; Lim, M. L.; Liz-Marzan, L. M.; Ma, X. M.; Macchiarini, P.; Meng, H.; Mohwald, H.; Mulvaney, P.; Nel, A. E.; Nie, S. M.; Nordlander, P.; Okano, T.; Oliveira, J.; Park, T. H.; Penner, R. M.; Prato, M.; Puntes, V.; Rotello, V. M.; Samarakoon, A.; Schaak, R. E.; Shen, Y. Q.; Sjoqvist, S.; Skirtach, A. G.; Soliman, M. G.; Stevens, M. M.; Sung, H. W.; Tang, B. Z.; Tietze, R.; Udugama, B. N.; VanEpps, J. S.; Weil, T.; Weiss, P. S.; Willner, I.; Wu, Y. Z.; Yang, L. L.; Yue, Z.; Zhang, Q.; Zhang, X. E.; Zhao, Y. L.; Zhou, X.; Parak, W. J., Diverse Applications of Nanomedicine. *Acs Nano* **2017,** *11* (3), 2313-2381.

9. Abedini, A.; Bakar, A. A. A.; Larki, F.; Menon, P. S.; Islam, M. S.; Shaari, S., Recent Advances in Shape-Controlled Synthesis of Noble Metal Nanoparticles by Radiolysis Route. *Nanoscale Research Letters* **2016,** *11*, 1-13.

10. da Silva, A. G. M.; Rodrigues, T. S.; Slater, T. J. A.; Lewis, E. A.; Alves, R. S.; Fajardo, H. V.; Balzer, R.; da Silva, A. H. M.; de Freitas, I. C.; Oliveira, D. C.; Assaf, J. M.; Probst, L. F. D.; Haigh, S. J.; Camargo, P. H. C., Controlling Size, Morphology, and Surface Composition of AgAu Nanodendrites in 15 s for Improved Environmental

Catalysis under Low Metal Loadings. *Acs Applied Materials & Interfaces* **2015,** *7* (46), 25624-25632.

11. Groeneveld, E.; de Mello Donegá, C., The Challenge of Colloidal Nanoparticle Synthesis. *Nanoparticles: Workhorses of Nanoscience* **2014**, 145-189.

12. Cheng, Y. H.; He, C. L.; Riviere, J. E.; Monteiro-Riviere, N. A.; Lin, Z. M., Meta-Analysis of Nanoparticle Delivery to Tumors Using a Physiologically Based Pharmacokinetic Modeling and Simulation Approach. *Acs Nano* **2020,** *14* (3), 3075-3095.

13. Wilhelm, S.; Tavares, A. J.; Dai, Q.; Ohta, S.; Audet, J.; Dvorak, H. F.; Chan, W. C. W., Analysis of nanoparticle delivery to tumours. *Nature Reviews Materials* **2016,** *1* (5).

14. Venditto, V. J.; Szoka, F. C., Cancer nanomedicines: So many papers and so few drugs! *Advanced Drug Delivery Reviews* **2013,** *65* (1), 80-88.

15. Park, K., The beginning of the end of the nanomedicine hype. *Journal of Controlled Release* **2019,** *305*, 221-222.

16. Gause, K. T.; Wheatley, A. K.; Cui, J. W.; Yan, Y.; Kent, S. J.; Caruso, F., Immunological Principles Guiding the Rational Design of Particles for Vaccine Delivery. *Acs Nano* **2017,** *11* (1), 54-68.

17. Anchordoquy, T. J.; Barenholz, Y.; Boraschi, D.; Chorny, M.; Decuzzi, P.; Dobrovolskaia, M. A.; Farhangrazi, Z. S.; Farrell, D.; Gabizon, A.; Ghandehari, H.; Godin, B.; La-Beck, N. M.; Ljubimova, J.; Moghimi, S. M.; Pagliaro, L.; Park, J.-H.; Peer, D.; Ruoslahti, E.; Serkova, N. J.; Simberg, D., Mechanisms and Barriers in Cancer Nanomedicine: Addressing Challenges, Looking for Solutions. *ACS nano* **2017,** *11* (1), 12-18.

18. Cooper, G. M., *The cell : a molecular approach*. 4th ed. ed.; ASM Press: Washington, D.C, 2007.

19. Poteryaev, D.; Datta, S.; Ackema, K.; Zerial, M.; Spang, A., Identification of the Switch in Early-to-Late Endosome Transition. *Cell (Cambridge)* **2010,** *141* (3), 497-508.

20. Endocytosis. *Nature reviews. Molecular cell biology* **2007,** *8* (6), 427.

21. Huotari, J.; Helenius, A., Endosome maturation. *Embo Journal* **2011,** *30* (17), 3481-3500.

22. Pei, D. H.; Buyanova, M., Overcoming Endosomal Entrapment in Drug Delivery. *Bioconjugate Chemistry* **2019,** *30* (2), 273-283.

23. Yang, X.; Fan, B.; Gao, W.; Li, L. P.; Li, T. T.; Sun, J. H.; Peng, X. Y.; Li, X. Y.; Wang, Z. J.; Wang, B. Q.; Zhang, R. P.; Xie, J., Enhanced endosomal escape by photothermal activation for improved small interfering RNA delivery and antitumor effect. *International Journal of Nanomedicine* **2018,** *13*, 4333-4344.

24. Ahmed, S.; Fujitab, S.; Matsumura, K., Enhanced protein internalization and efficient endosomal escape using polyampholyte-modified liposomes and freeze concentration. *Nanoscale* **2016,** *8* (35), 15888-15901.

25. Gao, Y. J.; Li, Y. F.; Li, Y. S.; Yuan, L.; Zhou, Y. X.; Li, J. W.; Zhao, L.; Zhang, C.; Li, X. R.; Liu, Y., PSMA-mediated endosome escape-accelerating polymeric micelles for targeted therapy of prostate cancer and the real time tracing of their intracellular trafficking. *Nanoscale* **2015,** *7* (2), 597-612.

26. Huang, G. J.; Chen, Q.; Wu, W. T.; Wang, J. W.; Chu, P. K.; Bai, H. Z.; Tang, G. P., Reconstructed chitosan with alkylamine for enhanced gene delivery by promoting endosomal escape. *Carbohydrate Polymers* **2020,** *227*.

27. Fröhlich, E., The role of surface charge in cellular uptake and cytotoxicity of medical nanoparticles. *International journal of nanomedicine* **2012,** *7*, 5577-5591.

28. Cristina, M.; Manuela, M.; Gheorghita, M.; Vasile, O., Understanding the Factors Influencing Chitosan-Based Nanoparticles-Protein Corona Interaction and Drug Delivery Applications. *Molecules (Basel, Switzerland)* **2020,** *25* (20), 4758.

29. Raouane, M.; Desmaële, D.; Urbinati, G.; Massaad-Massade, L.; Couvreur, P., Lipid Conjugated Oligonucleotides: A Useful Strategy for Delivery. *Bioconjugate chemistry* **2012,** *23* (6), 1091-1104.

30. Gissot, A.; Camplo, M.; Grinstaff, M. W.; Barthélémy, P., Nucleoside, nucleotide and oligonucleotide based amphiphiles: A successful marriage of nucleic acids with lipids. *Organic & biomolecular chemistry* **2008,** *6* (8), 1324-1333.

31. Shmushkovich, T.; Monopoli, K. R.; Homsy, D.; Leyfer, D.; Betancur-Boissel, M.; Khvorova, A.; Wolfson, A. D., Functional features defining the efficacy of cholesterol-conjugated, self-deliverable, chemically modified siRNAs. *Nucleic acids research* **2018,** *46* (20), 10905-10916.

32. Jeong, J. H.; Kim, S. W.; Park, T. G., Novel Intracellular Delivery System of Antisense Oligonucleotide by Self-Assembled Hybrid Micelles Composed of DNA/PEG Conjugate and Cationic Fusogenic Peptide. *Bioconjugate chemistry* **2003,** *14* (2), 473-479.

33. Kundu, A.; Nandi, S.; Nandi, A. K., Nucleic acid based polymer and nanoparticle conjugates: Synthesis, properties and applications. *Progress in materials science* **2017,** *88*, 136-185.

34. Nielsen, P. E., Peptide nucleic acid: a versatile tool in genetic diagnostics and molecular biology. *Current Opinion in Biotechnology* **2001,** *12* (1), 16-20.

35. Whitehouse, W. L.; Noble, J. E.; Ryadnov, M. G.; Howorka, S., Cholesterol Anchors Enable Efficient Binding and Intracellular Uptake of DNA Nanostructures. *Bioconjugate Chemistry* **2019,** *30* (7), 1836-1844.

36. Ponnuswamy, N.; Bastings, M. M. C.; Nathwani, B.; Ryu, J. H.; Chou, L. Y. T.; Vinther, M.; Li, W. A.; Anastassacos, F. M.; Mooney, D. J.; Shih, W. M., Oligolysine-based coating protects DNA nanostructures from low-salt denaturation and nuclease degradation. *Nature Communications* **2017,** *8*.

37. Anastassacos, F. M.; Zhao, Z.; Zeng, Y.; Shih, W. M., Glutaraldehyde Cross-Linking of Oligolysines Coating DNA Origami Greatly Reduces Susceptibility to Nuclease Degradation. *Journal of the American Chemical Society* **2020,** *142* (7), 3311-3315.

38. Dervan, P. B., A Personal Perspective on Chemical Biology: Before the Beginning. *Israel journal of chemistry* **2019,** *59* (1-2), 71-83.

39. Alniss, H. Y., Thermodynamics of DNA Minor Groove Binders. *Journal of medicinal chemistry* **2019,** *62* (2), 385-402.

40. Pazos, E.; Mosquera, J.; Vazquez, E.; Mascarenas, J., DNA Recognition by Synthetic Constructs. *ChemBioChem* **2011,** *12*, 1958-1973.

41. Roodhuizen, J. A. L.; Hendrikx, P.; Hilbers, P. A. J.; de Greef, T. F. A.; Markvoort, A. J., Counterion-Dependent Mechanisms of DNA Origami Nanostructure Stabilization Revealed by Atomistic Molecular Simulation. *Acs Nano* **2019,** *13* (9), 10798-10809.

42. Neundorf, I.; Rennert, R.; Hoyer, J.; Schramm, F.; Löbner, K.; Kitanovic, I.; Wölfl, S., Fusion of a short HA2-derived peptide sequence to cell-penetrating peptides improves cytosolic uptake, but enhances cytotoxic activity. *Pharmaceuticals (Basel, Switzerland)* **2009,** *2* (2), 49-65.

43. Sugita, T.; Yoshikawa, T.; Mukai, Y.; Yamanada, N.; Imai, S.; Nagano, K.; Yoshida, Y.; Shibata, H.; Yoshioka, Y.; Nakagawa, S.; Kamada, H.; Tsunoda, S. i.; Tsutsumi, Y., Comparative study on transduction and toxicity of protein transduction domains. *British journal of pharmacology* **2008,** *153* (6), 1143-1152.

44. Li, M.; Tao, Y.; Shu, Y. L.; LaRochelle, J. R.; Steinauer, A.; Thompson, D.; Schepartz, A.; Chen, Z. Y.; Liu, D. R., Discovery and Characterization of a Peptide That Enhances Endosomal Escape of Delivered Proteins in Vitro and in Vivo. *Journal of the American Chemical Society* **2015,** *137* (44), 14084-14093.

45. Amblard, M.; Fehrentz, J.-A.; Martinez, J.; Subra, G., Methods and protocols of modern solid phase peptide synthesis. *Molecular biotechnology* **2006,** *33* (3), 239-254.

46. Burns, J. R.; Seifert, A.; Fertig, N.; Howorka, S., A biomimetic DNA-based channel for the ligand-controlled transport of charged molecular cargo across a biological membrane. *Nature Nanotechnology* **2016,** *11* (2), 152-156.

47. Zhang, Y. N.; Poon, W.; Tavares, A. J.; McGilvray, I. D.; Chan, W. C. W., Nanoparticle-liver interactions: Cellular uptake and hepatobiliary elimination. *Journal of Controlled Release* **2016,** *240*, 332-348.

48. Frtús, A.; Smolková, B.; Uzhytchak, M.; Lunova, M.; Jirsa, M.; Kubinová, Š.; Dejneka, A.; Lunov, O., Analyzing the mechanisms of iron oxide nanoparticles interactions with cells: A road from failure to success in clinical applications. *Journal of controlled release* **2020,** *328*, 59-77.

49. Tsoi, K. M.; MacParland, S. A.; Ma, X. Z.; Spetzler, V. N.; Echeverri, J.; Ouyang, B.; Fadel, S. M.; Sykes, E. A.; Goldaracena, N.; Kaths, J. M.; Conneely, J. B.; Alman, B. A.; Selzner, M.; Ostrowski, M. A.; Adeyi, O. A.; Zilman, A.; McGilvray, I. D.; Chan, W. C. W., Mechanism of hard-nanomaterial clearance by the liver. *Nature Materials* **2016,** *15* (11), 1212-1221.

50. Lunova, M.; Prokhorov, A.; Jirsa, M.; Hof, M.; Olzynska, A.; Jurkiewicz, P.; Kubinova, S.; Lunov, O.; Dejneka, A., Nanoparticle core stability and surface functionalization drive the mTOR signaling pathway in hepatocellular cell lines. *Scientific Reports* **2017,** *7*.

51. Lunova, M.; Zablotskii, V.; Dempsey, N. M.; Devillers, T.; Jirsa, M.; Sykova, E.; Kubinova, S.; Lunov, O.; Dejneka, A., Modulation of collective cell behaviour by geometrical constraints. *Integrative Biology* **2016,** *8* (11), 1099-1110.

52. Bade, N. D.; Kamien, R. D.; Assoian, R. K.; Stebe, K. J., Curvature and Rho activation differentially control the alignment of cells and stress fibers. *Science Advances* **2017,** *3* (9).

53. Amodeo, A. A.; Skotheim, J. M., Cell-size control. *Cold Spring Harbor perspectives in biology* **2016,** *8* (4), a019083-a019083.

54. Khetan, J.; Shahinuzzaman, M.; Barua, S.; Barua, D., Quantitative Analysis of the Correlation between Cell Size and Cellular Uptake of Particles. *Biophysical Journal* **2019,** *116* (2), 347-359.

55. Ni, R.; Feng, R.; Chau, Y., Synthetic Approaches for Nucleic Acid Delivery: Choosing the Right Carriers. *Life (Basel, Switzerland)* **2019,** *9* (3), 59.

56. Auvinen, H.; Zhang, H.; Nonappa; Kopilow, A.; Niemelä, E.; Nummelin, S.; Correia, A.; Santos, H.; Linko, V.; Kostiainen , M., Protein Coating of DNA

Nanostructures for Enhanced Stability and Immunocompatibility. *advanced healthcare materials* **2017,** *6*.

57. Juliano, R. L., Intracellular Trafficking and Endosomal Release of Oligonucleotides: What We Know and What We Don't. *Nucleic acid therapeutics* **2018,** *28* (3), 166-177.

58. Del Pino, P.; Pelaz, B.; Zhang, Q.; Maffre, P.; Nienhaus, G. U.; Parak, W. J., Protein corona formation around nanoparticles - from the past to the future. *Materials Horizons* **2014,** *1* (3), 301-313.

59. Ke, P. C.; Lin, S.; Parak, W. J.; Davis, T. P.; Caruso, F., A Decade of the Protein Corona. *Acs Nano* **2017,** *11* (12), 11773-11776.

60. Shang, L.; Nienhaus, G. U., In Situ Characterization of Protein Adsorption onto Nanoparticles by Fluorescence Correlation Spectroscopy. *Accounts of Chemical Research* **2017,** *50* (2), 387-395.

61. Geng, L.; Wang, Z.; Jia, X.; Han, Q.; Xiang, Z.; Li, D.; Yang, X.; Zhang, D.; Bu, X.; Wang, W.; Hu, Z.; Fang, Q., HER2 targeting peptides screening and applications in tumor imaging and drug delivery. *Theranostics* **2016,** *6* (8), 1261-1273.

62. Neve, R.; Chin, K.; Fridlyand, J.; Yeh, J.; Gazdar, A.; Gray, J., A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **2006,** *10(6)*, 515-27.

63. Subik, K.; Lee, J.-F.; Baxter, L.; Strzepek, T.; Costello, D.; Crowley, P.; Xing, L.; Hung, M.-C.; Bonfiglio, T.; Hicks, D. G.; Tang, P., The expression patterns of ER, PR, HER2, CK5/6, EGFR, KI-67 and AR by immunohistochemical analysis in breast cancer cell lines. *Breast cancer : basic and clinical research* **2010,** *4* (1), 35-41.

64. Wang, S.-T.; Gray, M. A.; Xuan, S.; Lin, Y.; Byrnes, J.; Nguyen, A. I.; Todorova, N.; Stevens, M. M.; Bertozzi, C. R.; Zuckermann, R. N.; Gang, O., DNA origami protection and molecular interfacing through engineered sequence-defined peptoids. *Proceedings of the National Academy of Sciences - PNAS* **2020,** *117* (12), 6339-6348.

Chapter 3

1. MacCulloch, T.; Buchberger, A.; Stephanopoulos, N., Emerging applications of peptide-oligonucleotide conjugates: bioactive scaffolds, self-assembling systems, and hybrid nanomaterials. *Organic & Biomolecular Chemistry* **2019,** *17* (7), 1668-1682.

2. Venkatesan, N.; Kim, B. H., Peptide conjugates of oligonucleotides: Synthesis and applications. *Chemical Reviews* **2006,** *106* (9), 3712-3761.

3. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O., SEQUENCE-SELECTIVE RECOGNITION OF DNA BY STRAND DISPLACEMENT WITH A THYMINE-SUBSTITUTED POLYAMIDE. *Science* **1991,** *254* (5037), 1497-1500.

4. Lonnberg, H., Solid-Phase Synthesis of Oligonucleotide Conjugates Useful for Delivery and Targeting of Potential Nucleic Acid Therapeutics. *Bioconjugate Chemistry* **2009,** *20* (6), 1065-1094.

5. Lu, K.; Duan, Q.-P.; Ma, L.; Zhao, D.-X., Chemical Strategies for the Synthesis of Peptide−Oligonucleotide Conjugates. *Bioconjugate chemistry* **2010,** *21* (2), 187-202.

6. Kolb, H. C.; Finn, M. G.; Sharpless, K. B., Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angewandte Chemie International Edition* **2001,** *40* (11), 2004-2021.

7. Service, R. F., Chemistry: Click chemistry clicks along. *Science (American Association for the Advancement of Science)* **2008,** *320* (5878), 868-869.

8. Hein, C. D.; Liu, X.-M.; Wang, D., Click Chemistry, A Powerful Tool for Pharmaceutical Sciences. *Pharmaceutical research* **2008,** *25* (10), 2216-2230.

9. Northrop, B. H.; Frayne, S. H.; Choudhary, U., Thiol-maleimide "click" chemistry: Evaluating the influence of solvent, initiator, and thiol on the reaction mechanism, kinetics, and selectivity. *Polymer chemistry* **2015,** *6* (18), 3415-3430.

10. Besanceney-Webler, C.; Jiang, H.; Zheng, T.; Feng, L.; Soriano del Amo, D.; Wang, W.; Klivansky, L.; Marlow, F.; Liu, Y.; Wu, P., Increasing the Efficacy of Bioorthogonal Click Reactions for Bioconjugation: A Comparative Study. *Angewandte Chemie International Edition* **2011,** *50* (35), 8051-8056.

11. Liang, L.; Astruc, D., The copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) “click” reaction and its applications. An overview. *Coordination chemistry reviews* **2011,** *255* (23), 2933-2945.

12. Presolski, S. I., Ligand-Accelerated Copper(I)-Catalyzed Azide-Alkyne Cycloaddition: Kinetic Studies and Reaction Mechanism. ProQuest Dissertations Publishing: 2011.

13. Presolski, S.; Hong, V. P.; Finn, M. G., Copper-Catalyzed Azide–Alkyne Click

Chemistry for Bioconjugation. *Current Protocols Chemical Biology* **2011,** *3*, 153-162.

14. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R., A Strain-Promoted [3 + 2] Azide−Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *Journal of the American Chemical Society* **2004,** *126* (46), 15046-15047.

15. Sapra, P.; Shor, B., Monoclonal antibody-based therapies in cancer: Advances and challenges. *Pharmacology & therapeutics (Oxford)* **2013,** *138* (3), 452-469.

16. Chiu, M. L.; Goulet, D. R.; Teplyakov, A.; Gilliland, G. L., Antibody Structure and Function: The Basis for Engineering Therapeutics. *Antibodies (Basel)* **2019,** *8* (4), 55.

17. Strohl, W. R., Current progress in innovative engineered antibodies. *Protein & cell* **2018,** *9* (1), 86-120.

18. Miersch, S.; Sidhu, S. S., Synthetic antibodies: Concepts, potential and practical considerations. *Methods (San Diego, Calif.)* **2012,** *57* (4), 486-498.

19. Deyev, S. M.; Lebedenko, E. N., Modern Technologies for Creating Synthetic Antibodies for Clinical Application. *Actanaturae* **2009,** *1* (1), 32-50.

20. Owens, B., Faster, deeper, smaller-the rise of antibody-like scaffolds (vol 35, pg 602, 2017). *Nature Biotechnology* **2017,** *35* (11), 1004-1004.

21. Williams, B. A. R.; Diehnelt, C. W.; Belcher, P.; Greving, M.; Woodbury, N. W.; Johnston, S. A.; Chaput, J. C., Creating Protein Affinity Reagents by Combining Peptide Ligands on Synthetic DNA Scaffolds. *Journal of the American Chemical Society* **2009,** *131* (47), 17233-17241.

22. Miller, J. L., Iron deficiency anemia: A common and curable disease. *Cold Spring Harbor perspectives in biology* **2013,** *5* (7), a011866-a011866.

23. Ulbrich, K.; Hekmatara, T.; Herbert, E.; Kreuter, J., Transferrin- and transferrin-receptor-antibody-modified nanoparticles enable drug delivery across the blood–brain barrier (BBB). *European journal of pharmaceutics and biopharmaceutics* **2009,** *71* (2), 251-256.

24. Scognamiglio, I.; Di Martino, M. T.; Campani, V.; Virgilio, A.; Galeone, A.; Gullà, A.; Gallo Cantafio, M. E.; Misso, G.; Tagliaferri, P.; Tassone, P.; Caraglia, M.; De Rosa, G., Transferrin-Conjugated SNALPs Encapsulating 2′-O-Methylated miR-34a for the Treatment of Multiple Myeloma. *BioMed research international* **2014,** *2014*, 217365-7.

25. Peng, Y.; Rizal, F. H.; Sudheer, S.; Harry, M. T. C.; Sung Ha, P.; Thomas, H. L.; John, H. R., Programming DNA Tube Circumferences. *Science (American Association for the Advancement of Science)* **2008,** *321* (5890), 824-826.

26. Lutolf, M. P.; Hubbell, J. A., Synthetic biomaterials as instructive extracellular microenvironments for morphogenesis in tissue engineering. *Nature Biotechnology* **2005,** *23* (1), 47-55.

27. Chau, Y.; Luo, Y.; Cheung, A. C. Y.; Nagai, Y.; Zhang, S. G.; Kobler, J. B.; Zeitels, S. M.; Langer, R., Incorporation of a matrix metalloproteinase-sensitive substrate into self-assembling peptides - A model for biofunctional scaffolds. *Biomaterials* **2008,** *29* (11), 1713-1719.

28. Chau, Y.; Tan, F. E.; Langer, R., Synthesis and characterization of dextran-peptide-methotrexate conjugates for tumor targeting via mediation by matrix metalloproteinase II and matrix metalloproteinase IX. *Bioconjugate Chemistry* **2004,** *15* (4), 931-941.

29. Aimes, R. T.; Quigley, J. P., MATRIX METALLOPROTEINASE-2 IS AN INTERSTITIAL COLLAGENASE - INHIBITOR-FREE ENZYME CATALYZES THE CLEAVAGE OF COLLAGEN FIBRILS AND SOLUBLE NATIVE TYPE-I COLLAGEN GENERATING THE SPECIFIC 3/4-LENGTH AND 1/4-LENGTH FRAGMENTS. *Journal of Biological Chemistry* **1995,** *270* (11), 5872-5876.

30. Patterson, J.; Hubbell, J. A., Enhanced proteolytic degradation of molecularly engineered PEG hydrogels in response to MMP-1 and MMP-2. *Biomaterials* **2010,** *31* (30), 7836-7845.

31. Bergmann, F.; Bannwarth, W., Solid phase synthesis of directly linked peptide-oligodeoxynucleotide hybrids using standard synthesis protocols. *Tetrahedron letters* **1995,** *36* (11), 1839-1842.

32. Hayashi, G.; Yanase, M.; Nakatsuka, Y.; Okamoto, A., Simultaneous and Traceless Ligation of Peptide Fragments on DNA Scaffold. *Biomacromolecules* **2019,** *20* (3), 1246-1253.

33. Horton, D. A.; Bourne, G. T.; Smythe, M. L., Exploring privileged structures: The combinatorial synthesis of cyclic peptides. *Molecular diversity* **2000,** *5* (4), 289-304.

34. Joo, S. H., Cyclic peptides as therapeutic agents and biochemical tools. *Biomolecules & therapeutics* **2012,** *20* (1), 19-26.

35. Lee, A. C.-L.; Harris, J. L.; Khanna, K. K.; Hong, J.-H., A comprehensive review on current advances in peptide drug development and design. *International journal of molecular sciences* **2019,** *20* (10), 2383.

36. Liskamp, R.; Rijkers, D.; Bakker, S., *Bioactive Macrocyclic Peptides and Peptide Mimics*. Wiley-VCH: 2008.

37. Finke, A.; Busskamp, H.; Manea, M.; Marx, A., Designer Extracellular Matrix Based on DNA-Peptide Networks Generated by Polymerase Chain Reaction. *Angewandte Chemie-International Edition* **2016,** *55* (34), 10136-10140.

38. Campbell, I. D.; Humphries, M. J., Integrin structure, activation, and interactions. *Cold Spring Harbor perspectives in biology* **2011,** *3* (3), 1-14.

39. Assa-Munt, N.; Jia, X.; Laakkonen, P.; Ruoslahti, E., Solution structures and integrin binding activities of an RGD peptide with two isomers. *Biochemistry* **2001,** *40* (8), 2373-2378.

40. Bogdanowich-Knipp, S. J.; Jois, D. S. S.; Siahaan, T. J., The effect of conformation on the solution stability of linear vs. cyclic RGD peptides. *Journal of Peptide Research* **1999,** *53* (5), 523-529.

41. Bogdanowich-Knipp, S. J.; Chakrabarti, S.; Williams, T. D.; Dillmall, R. K.; Siahaan, T. J., Solution stability of linear vs. cyclic RGD peptides. *Journal of Peptide Research* **1999,** *53* (5), 530-541.

42. Krammer, A.; Craig, D.; Thomas, W. E.; Schulten, K.; Vogel, V., A structural model for force regulated integrin binding to fibronectin's RGD-synergy site. *Matrix biology* **2002,** *21* (2), 139-147.

43. Leahy, D. J.; Aukhil, I.; Erickson, H. P., 2.0 Å Crystal Structure of a Four-Domain Segment of Human Fibronectin Encompassing the RGD Loop and Synergy Region. *Cell (Cambridge)* **1996,** *84* (1), 155-164.

44. Aota, S.-i.; Nomizu, M.; Yamada, K. M., The Short Amino Acid Sequence Pro-His-Ser-Arg-Asn in Human

Fibronectin Enhances Cell-adhesive Function. *The Journal of Biological Chemistry* **1994,** *269*, 24756-24761.

45. Liu, Y.; Ke, Y.; Yan, H.; Rinker, S.; Chhabra, R., Self-assembled DNA nanostructures for distance-dependent multivalent ligand-protein binding. *Nature nanotechnology* **2008,** *3* (7), 418-422.

46. Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. â. G., Analysis and Optimization of Copper-Catalyzed Azideâ□“Alkyne Cycloaddition for Bioconjugation. *Angewandte Chemie International Edition* **2009,** *48* (52), 9879-9883.

Chapter 4

1. Seeman, N. C., NUCLEIC-ACID JUNCTIONS AND LATTICES. *Journal of Theoretical Biology* **1982,** *99* (2), 237-247.

2. Duckett, D. R.; Murchie, A. I. H.; Diekmann, S.; Vonkitzing, E.; Kemper, B.; Lilley, D. M. J., THE STRUCTURE OF THE HOLLIDAY JUNCTION, AND ITS RESOLUTION. *Cell* **1988,** *55* (1), 79-89.

3. Watson, J. D.; Crick, F. H., Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. J.D. Watson and F.H.C. Crick. Published in Nature, number 4356 April 25, 1953. *Nature (London)* **1974,** *248* (5451), 765-765.

4. Kallenbach, N. R.; Ma, R. I.; Seeman, N. C., AN IMMOBILE NUCLEIC-ACID JUNCTION CONSTRUCTED FROM OLIGONUCLEOTIDES. *Nature* **1983,** *305* (5937), 829-831.

5. Eis, P. S.; Millar, D. P., CONFORMATIONAL DISTRIBUTIONS OF A 4-WAY DNA JUNCTION REVEALED BY TIME-RESOLVED FLUORESCENCE RESONANCE ENERGY-TRANSFER. *Biochemistry* **1993,** *32* (50), 13852-13860.

6. Zhang, S. W.; Seeman, N. C., SYMMETRICAL HOLLIDAY JUNCTION CROSSOVER ISOMERS. *Journal of Molecular Biology* **1994,** *238* (5), 658-668.

7. Fu, T.-J.; Tse-Dinh, Y.-C.; Seeman, N. C., Holliday Junction Crossover Topology. *Journal of molecular biology* **1994,** *236* (1), 91-105.

8. Miick, S. M.; Fee, R. S.; Millar, D. P.; Chazin, W. J., Crossover isomer bias is the primary sequence-dependent property of immobilized Holliday junctions. *Proceedings of the National Academy of Sciences of the United States of America* **1997,** *94* (17), 9080-9084.

9. Ho, P. S.; Eichman, B. F., The crystal structures of DNA Holliday junctions. *Current Opinion in Structural Biology* **2001,** *11* (3), 302-308.

10. Nadrian, C. S., Art as a Stimulus for Structural DNA Nanotechnology. *Leonardo (Oxford)* **2014,** *47* (2), 142-149.

11. Seeman, N. C.; Sleiman, H. F., DNA nanotechnology. *Nature Reviews Materials* **2018,** *3* (1).

12. Stanley, N. C.; Annie, C. Y. C.; Herbert, W. B.; Robert, B. H., Construction of Biologically Functional Bacterial Plasmids In Vitro. *Proceedings of the National Academy of Sciences - PNAS* **1973,** *70* (11), 3240-3244.

13. Winfree, E.; Liu, F. R.; Wenzler, L. A.; Seeman, N. C., Design and self-assembly of two-dimensional DNA crystals. *Nature* **1998,** *394* (6693), 539-544.

14. Liu, D.; Wang, M.; Deng, Z.; Walulu, R.; Mao, C., Tensegrity: Construction of Rigid DNA Triangles with Flexible Four-Arm DNA Junctions. *Journal of the American Chemical Society* **2004,** *126* (8), 2324-2325.

15. Holliday, R., A mechanism for gene conversion in fungi (Reprinted). *Genetics Research* **2007,** *89* (5-6), 285-307.

16. Ho, P. S., Structure of the Holliday junction: applications beyond recombination. *Biochemical Society Transactions* **2017,** *45*, 1149-1158.

17. Shrestha, P.; Emura, T.; Koirala, D.; Cui, Y. X.; Hidaka, K.; Maximuck, W. J.; Endo, M.; Sugiyama, H.; Mao, H. B., Mechanical properties of DNA origami

nanoassemblies are determined by Holliday junction mechanophores. *Nucleic Acids Research* **2016,** *44* (14), 6574-6582.

18. Gelbin, A.; Schneider, B.; Clowney, L.; Hsieh, S.-H.; Olson, W. K.; Berman, H. M., Geometric Parameters in Nucleic Acids: Sugar and Phosphate Constituents. *Journal of the American Chemical Society* **1996,** *118* (3), 519-529.

19. Seeman, N. C., DNA Nanotechnology at 40. *Nano letters* **2020,** *20* (3), 1477-1478.

20. Hu, Y.; Niemeyer, C. M., From DNA Nanotechnology to Material Systems Engineering. *Advanced Materials* **2019,** *31* (26).

21. Veneziano, R.; Ratanalert, S.; Zhang, K. M.; Zhang, F.; Yan, H.; Chiu, W.; Bathe, M., DNA NANOTECHNOLOGY Designer nanoscale DNA assemblies programmed from the top down. *Science* **2016,** *352* (6293).

22. Seeman, N. C., DNA nanotechnology: Novel DNA constructions. *Annual Review of Biophysics and Biomolecular Structure* **1998,** *27*, 225-248.

23. Seeman, N. C., DNA in a material world. *Nature* **2003,** *421* (6921), 427-431.

24. Deng, Z. X.; Lee, S. H.; Mao, C. D., DNA as nanoscale building blocks. *Journal of Nanoscience and Nanotechnology* **2005,** *5* (12), 1954-1963.

25. Pinheiro, A. V.; Han, D. R.; Shih, W. M.; Yan, H., Challenges and opportunities for structural DNA nanotechnology. *Nature Nanotechnology* **2011,** *6* (12), 763-772.

26. Zhang, C. A.; He, Y.; Su, M.; Ko, S. H.; Ye, T.; Leng, Y. J.; Sun, X. P.; Ribbe, A. E.; Jiang, W.; Mao, C. D., DNA self-assembly: from 2D to 3D. *Faraday Discussions* **2009,** *143*, 221-233.

27. Weiss, P. S., A conversation with Prof. Ned Seeman: Founder of DNA nanotechnology. *ACS nano* **2008,** *2* (6), 1089-1096.

28. Zheng, J. P.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R. J.; Constantinou, P. E.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **2009,** *461* (7260), 74-77.

29. Ohayon, Y. P.; Hernandez, C.; Chandrasekaran, A. R.; Wang, X. Y.; Abdallah, H. O.; Jong, M. A.; Mohsen, M. G.; Sha, R. J.; Birktoft, J. J.; Lukeman, P. S.; Chaikin, P. M.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., Designing Higher Resolution Self-Assembled 3D DNA Crystals via Strand Terminus Modifications. *Acs Nano* **2019,** *13* (7), 7957-7965.

30. Zhao, J. M.; Chandrasekaran, A. R.; Li, Q.; Li, X.; Sha, R. J.; Seeman, N. C.; Mao, C. D., Post-Assembly Stabilization of Rationally Designed DNA Crystals. *Angewandte Chemie-International Edition* **2015,** *54* (34), 9936-9939.

31. Sha, R. J.; Birktoft, J. J.; Nguyen, N.; Chandrasekaran, A. R.; Zheng, J. P.; Zhao, X. S.; Mao, C. D.; Seeman, N. C., Self-Assembled DNA Crystals: The Impact on Resolution of 5 '-Phosphates and the DNA Source. *Nano Letters* **2013,** *13* (2), 793-797.

32. Stahl, E.; Praetorius, F.; Mann, C. C. D.; Hopfner, K. P.; Dietz, H., Impact of Heterogeneity and Lattice Bond Strength on DNA Triangle Crystal Growth. *Acs Nano* **2016,** *10* (10), 9156-9164.

33. Hao, Y. D.; Kristiansen, M.; Sha, R. J.; Birktoft, J. J.; Hernandez, C.; Mao, C. D.; Seeman, N. C., A device that operates within a self-assembled 3D DNA crystal. *Nature Chemistry* **2017,** *9* (8), 824-827.

34. Rusling, D. A.; Chandrasekaran, A. R.; Ohayon, Y. P.; Brown, T.; Fox, K. R.; Sha, R. J.; Mao, C. D.; Seeman, N. C., Functionalizing Designer DNACrystals with a Triple-Helical Veneer. *Angewandte Chemie-International Edition* **2014,** *53* (15), 3979-3982.

35. Wang, T.; Sha, R. J.; Birktoft, J.; Zheng, J. P.; Mao, C. D.; Seeman, N. C., A DNA Crystal Designed to Contain Two Molecules per Asymmetric Unit. *Journal of the American Chemical Society* **2010,** *132* (44), 15471-15473.

36. Zadeh, J. N.; Steenberg, C. D.; Bois, J. S.; Wolfe, B. R.; Pierce, M. B.; Khan, A. R.; Dirks, R. M.; Pierce, N. A., NUPACK: analysis and design of nucleic acid systems. *J Comput Chem* **2011,** *37*, 170-173.

37. Rhodes, G., *Crystallography made crystal clear a guide for users of macromolecular models*. 3rd ed. ed.; Elsevier/Academic Press: Amsterdam ;, 2006.

38. Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* **2016,** *138* (38), 12690-12690.

39. Zhang, F.; Simmons, C. R.; Gates, J.; Liu, Y.; Yan, H., Self-Assembly of a 3D DNA Crystal Structure with Rationally Designed Six-Fold Symmetry. *Angewandte Chemie-International Edition* **2018,** *57* (38), 12504-12507.

40. Liu, D.; Wang, M. S.; Deng, Z. X.; Walulu, R.; Mao, C. D., Tensegrity: Construction of rigid DNA triangles with flexible four-arm DNA junctions. *Journal of the American Chemical Society* **2004,** *126* (8), 2324-2325.

41. Paukstelis, P. J., Three-dimensional DNA crystals as molecular sieves. *Journal of the American Chemical Society* **2006,** *128* (21), 6794-6795.

42. Alexis, T. B., The Impact of Nanoscience on Heterogeneous Catalysis. *Science (American Association for the Advancement of Science)* **2003,** *299* (5613), 1688-1691.

43. Geng, C.; Paukstelis, P. J., DNA Crystals as Vehicles for Biocatalysis. *Journal of the American Chemical Society* **2014,** *136* (22), 7817-7820.

44. Stebe, K. J.; Lewandowski, E.; Ghosh, M., Materials science. Oriented assembly of metamaterials. *Science (American Association for the Advancement of Science)* **2009,** *325* (5937), 159-160.

45. Robert, J. M.; Byeongdu, L.; Matthew, R. J.; Nadine, H.; George, C. S.; Chad, A. M., Nanoparticle Superlattice Engineering with DNA. *Science (American Association for the Advancement of Science)* **2011,** *334* (6053), 204-208.

46. Jones, M. R.; Osberg, K. D.; Macfarlane, R. J.; Langille, M. R.; Mirkin, C. A., Templated Techniques for the Synthesis and Assembly of Plasmonic Nanostructures. *Chemical reviews* **2011,** *111* (6), 3736-3827.

47. Zhao, J. M.; Zhao, Y.; Li, Z.; Wang, Y.; Sha, R. J.; Seeman, N. C.; Mao, C. D., Modulating Self-Assembly of DNA Crystals with Rationally Designed Agents. *Angewandte Chemie-International Edition* **2018,** *57* (50), 16529-16532.

48. Melinger, J. S.; Sha, R.; Mao, C.; Seeman, N. C.; Ancona, M. G., Fluorescence and Energy Transfer in Dye-Labeled DNA Crystals. *The journal of physical chemistry. B* **2016,** *120* (48), 12287-12292.

49. Gottesfeld, J. M.; Melander, C.; Suto, R. K.; Raviol, H.; Luger, K.; Dervan, P. B., Sequence-specific Recognition of DNA in the Nucleosome by Pyrrole-Imidazole Polyamides. *Journal of molecular biology* **2001,** *309* (3), 615-629.

50. Kang, J. S.; Meier, J. L.; Dervan, P. B., Design of Sequence-Specific DNA Binding Molecules for DNA Methyltransferase Inhibition. *Journal of the American Chemical Society* **2014,** *136* (9), 3687-3694.

Chapter 5

1. Watson, J. D.; Crick, F. H., Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. J.D. Watson and F.H.C. Crick. Published in Nature, number 4356 April 25, 1953. *Nature (London)* **1974,** *248* (5451), 765-765.

2. Mandal, P. K.; Collie, G. W.; Kauffmann, B.; Huc, I., Racemic DNA Crystallography. *Angewandte Chemie-International Edition* **2014,** *53* (52), 14424-14427.

3. Hauser, N. C.; Martinez, R.; Jacob, A.; Rupp, S.; Hoheisel, J. D.; Matysiak, S., Utilising the left-helical conformation of L-DNA for analysing different marker types on a single universal microarray platform. *Nucleic Acids Research* **2006,** *34* (18), 5101-5111.

4. Urata, H.; Shinohara, K.; Ogura, E.; Ueda, Y.; Akagi, M., MIRROR-IMAGE DNA. *Journal of the American Chemical Society* **1991,** *113* (21), 8174-8175.

5. Urata, H.; Ogura, E.; Shinohara, K.; Ueda, Y.; Akagi, M., SYNTHESIS AND PROPERTIES OF MIRROR-IMAGE DNA. *Nucleic Acids Research* **1992,** *20* (13), 3325-3332.

6. Ashley, G. W., MODELING, SYNTHESIS, AND HYBRIDIZATION PROPERTIES OF (L)-RIBONUCLEIC ACID. *Journal of the American Chemical Society* **1992,** *114* (25), 9731-9736.

7. Damha, m. J.; giannaris, p. A.; marfey, p., antisense l/d-oligodeoxynucleotide chimeras - nuclease stability, base-pairing properties, and activity at directing ribonuclease-h. *Biochemistry* **1994,** *33* (25), 7877-7885.

8. Duckett, d. R.; murchie, a. I. H.; diekmann, s.; vonkitzing, e.; kemper, b.; lilley, d. M. J., the structure of the holliday junction, and its resolution. *Cell* **1988,** *55* (1), 79-89.

9. Seeman, N. C., DNA structural engineering using immobile junctions. *Current Opinion in Structural Biology* **1991,** *1* (4), 653-661.

10. Liu, Y. L.; West, S. C., Timeline - Happy Hollidays: 40th anniversary of the Holliday junction. *Nature Reviews Molecular Cell Biology* **2004,** *5* (11), 937-U21.

11. Holliday, R., A mechanism for gene conversion in fungi (Reprinted). *Genetics Research* **2007,** *89* (5-6), 285-307.

12. Delano, W. L. *The PYMOL Molecular Graphics System*, DeLano Scientific: San Carlos, CA, 2002.

13. Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* **2016,** *138* (38), 12690-12690.

Chapter 6

1. Holliday, R., A mechanism for gene conversion in fungi (Reprinted). *Genetics Research* 2007, *89* (5-6), 285-307.

2. Duckett, D. R.; Murchie, A. I. H.; Diekmann, S.; Vonkitzing, E.; Kemper, B.; Lilley, D. M. J., THE STRUCTURE OF THE HOLLIDAY JUNCTION, AND ITS RESOLUTION. *Cell* 1988, *55* (1), 79-89.

3. Clegg, R. M.; Murchie, A. I. H.; Zechel, A.; Carlberg, C.; Diekmann, S.; Lilley, D. M. J., FLUORESCENCE RESONANCE ENERGY-TRANSFER ANALYSIS OF

THE STRUCTURE OF THE 4-WAY DNA JUNCTION. *Biochemistry* 1992, *31* (20), 4846-4856.

4. Zhang, X. W.; Lee, W.; Fan, X. D., Bio-switchable optofluidic lasers based on DNA Holliday junctions. *Lab on a Chip* 2012, *12* (19), 3673-3675.

5. Karymov, M. A.; Chinnaraj, M.; Bogdanov, A.; Srinivasan, A. R.; Zheng, G. H.; Olson, W. K.; Lyubchenko, Y. L., Structure, Dynamics, and Branch Migration of a DNA Holliday Junction: A Single-Molecule Fluorescence and Modeling Study. *Biophysical Journal* 2008, *95* (9), 4372-4383.

6. Azaro, M. A.; Landy, A., The isomeric preference of Holliday junctions influences resolution bias by lambda integrase. *Embo Journal* 1997, *16* (12), 3744-3755.

7. Shida, T.; Iwasaki, H.; Shinagawa, H.; Kyogoku, Y., Characterization and comparison of synthetic immobile and mobile Holliday junctions. *Journal of Biochemistry* 1996, *119* (4), 653-658.

8. Yadav, R. K.; Yadava, U., Molecular dynamics simulation of hydrated d(CGGGTACCCG)(4) as a four-way DNA Holliday junction and comparison with the crystallographic structure. *Molecular Simulation* 2016, *42* (1), 25-30.

9. Wang, W. J.; Nocka, L. M.; Wiemann, B. Z.; Hinckley, D. M.; Mukerji, I.; Starr, F. W., Holliday Junction Thermodynamics and Structure: Coarse-Grained Simulations and Experiments. *Scientific Reports* 2016, *6*.

10. Hyeon, C.; Lee, J.; Yoon, J.; Hohng, S.; Thirumalai, D., Hidden complexity in the isomerization dynamics of Holliday junctions. *Nature Chemistry* 2012, *4* (11), 907-914.

11. Yu, J.; Ha, T. J.; Schulten, K., Conformational model of the Holliday junction transition deduced from molecular dynamics simulations. *Nucleic Acids Research* 2004, *32* (22), 6683-6695.

12. Overmars, F. J. J.; Lanzotti, V.; Galeone, A.; Pepe, A.; Mayol, L.; Pikkemaat, J. A.; Altona, C., Design and NMR study of an immobile DNA four-way junction containing 38 nucleotides. *European Journal of Biochemistry* 1997, *249* (2), 576-583.

13. Ho, P. S.; Eichman, B. F., The crystal structures of DNA Holliday junctions. *Current Opinion in Structural Biology* 2001, *11* (3), 302-308.

14. Hays, F. A.; Watson, J.; Ho, P. S., Caution! DNA crossing: crystal structures of Holliday junctions. *Journal of Biological Chemistry* 2003, *278* (50), 49663-49666.

15. Hays, F. A.; Vargason, J. M.; Ho, P. S., Effect of sequence on the conformation of DNA Holliday junctions. *Biochemistry* 2003, *42* (32), 9586-9597.

16. Eichman, B. F.; Ortiz-Lombardia, M.; Aymami, J.; Coll, M.; Ho, P. S., The inherent properties of DNA four-way junctions: Comparing the crystal structures of Holliday junctions. *Journal of Molecular Biology* 2002, *320* (5), 1037-1051.

17. Hays, F. A.; Teegarden, A.; Jones, Z. J. R.; Harms, M.; Raup, D.; Watson, J.; Cavaliere, E.; Ho, P. S., How sequence defines structure: A crystallographic map of DNA structure and conformation. *Proceedings of the National Academy of Sciences of the United States of America* 2005, *102* (20), 7157-7162.

18. Khuu, P. A.; Voth, A. R.; Hays, F. A.; Ho, P. S., The stacked-X DNA Holliday junction and protein recognition. *Journal of Molecular Recognition* 2006, *19* (3), 234-242.

19. Ho, P. S., Structure of the Holliday junction: applications beyond recombination. *Biochemical Society Transactions* 2017, *45*, 1149-1158.

20. Watson, J.; Hays, F. A.; Ho, P. S., Definitions and analysis of DNA Holliday junction geometry. *Nucleic Acids Research* 2004, *32* (10), 3017-3027.

21. Zheng, J. P.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R. J.; Constantinou, P. E.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* 2009, *461* (7260), 74-77.

22. Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* 2016, *138* (38), 12690-12690.

23. Simmons, C. R.; Zhang, F.; MacCulloch, T.; Fahmi, N.; Stephanopoulos, N.; Liu, Y.; Seeman, N. C.; Yan, H., Tuning the Cavity Size and Chirality of Self-Assembling 3D DNA Crystals. *Journal of the American Chemical Society* 2017, *139* (32), 11254-11260.

24. Zhang, F.; Simmons, C. R.; Gates, J.; Liu, Y.; Yan, H., Self-Assembly of a 3D DNA Crystal Structure with Rationally Designed Six-Fold Symmetry. *Angewandte Chemie-International Edition* 2018, *57* (38), 12504-12507.

25. Carlstrom, G.; Chazin, W. J., Sequence dependence and direct measurement of crossover isomer distribution in model holliday junctions using NMR spectroscopy. *Biochemistry* 1996, *35* (11), 3534-3544.

26. McKinney, S. A.; Declais, A. C.; Lilley, D. M. J.; Ha, T., Structural dynamics of individual Holliday junctions. *Nature Structural Biology* 2003, *10* (2), 93-97.

27. van Buuren, B. N. M.; Hermann, T.; Wijmenga, S. S.; Westhof, E., Brownian-dynamics simulations of metal-ion binding to four-way junctions. *Nucleic Acids Research* 2002, *30* (2), 507-514.

28. M.Govindaraju; Shekar, H. S.; S.B.Sateesha; Raju, P. V.; Rao, K. R. S.; Rao, K. S. J.; A.J.Rajamma, Copper interactions with DNA of chromatin and its role in neurodegenerative disorders. *Journal of pharmaceutical analysis* 2013, (5), 354-359.

29. Paris, C.; Geinguenaud, F.; Gouyette, C.; Liquier, J.; Lacoste, J., Mechanism of Copper Mediated Triple Helix Formation at Neutral pH in Drosophila Satellite Repeats. *Biophysical journal* 2007, *92* (7), 2498-2506.

Chapter 7

1. Zheng, J. P.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R. J.; Constantinou, P. E.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **2009,** *461* (7260), 74-77.

2. Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* **2016,** *138* (38), 12690-12690.

3. Simmons, C. R.; Zhang, F.; MacCulloch, T.; Fahmi, N.; Stephanopoulos, N.; Liu, Y.; Seeman, N. C.; Yan, H., Tuning the Cavity Size and Chirality of Self-Assembling 3D DNA Crystals. *Journal of the American Chemical Society* **2017,** *139* (32), 11254-11260.

4. Zhang, F.; Simmons, C. R.; Gates, J.; Liu, Y.; Yan, H., Self-Assembly of a 3D DNA Crystal Structure with Rationally Designed Six-Fold Symmetry. *Angewandte Chemie-International Edition* **2018,** *57* (38), 12504-12507.

5. Dannenberg; J, J., An Introduction to Hydrogen Bonding By George A. Jeffrey *Journal of the American Chemical Society*: Oxford University Press: New York and Oxford., 1998; Vol. 120, pp 5604-5604.

6. Oltrogge, L. M.; Boxer, S. G., Short Hydrogen Bonds and Proton Delocalization in Green Fluorescent Protein (GFP). *Acs Central Science* **2015,** *1* (3), 148-156.

7. Ohayon, Y. P.; Hernandez, C.; Chandrasekaran, A. R.; Wang, X. Y.; Abdallah, H. O.; Jong, M. A.; Mohsen, M. G.; Sha, R. J.; Birktoft, J. J.; Lukeman, P. S.; Chaikin, P. M.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., Designing Higher Resolution Self-Assembled 3D DNA Crystals via Strand Terminus Modifications. *Acs Nano* **2019,** *13* (7), 7957-7965.

8. Williams; S.; Lund, K.; Lin, C.; Wonka, P.; Lindsay, S.; Yan, H. In *T iamat:A Three-Dimensional Editing Tool for Complex DNA Structures, DNA Computing: 14th*

International Meeting on DNA Computing, Prague, Czech Republic, Springer-Verlag: *Prague, Czech Republic, 2008*; pp 90–101.

9. Bode, W.; Mayr, I.; Baumann, U.; Huber, R.; Stone, S. R.; Hofsteenge, J., the refined 1.9 a crystal-structure of human alpha-thrombin - interaction with d-phe-pro-arg chloromethylketone and significance of the tyr-pro-pro-trp insertion segment. *Embo Journal* **1989,** *8* (11), 3467-3475.

10. Kao, M. R.; Kuo, H. W.; Lee, C. C.; Huang, K. Y.; Huang, T. Y.; Li, C. W.; Chen, C. W.; Wang, A. H. J.; Yu, S. M.; Ho, T. H. D., Chaetomella raphigera beta-glucosidase D2-BGL has intriguing structural features and a high substrate affinity that renders it an efficient cellulase supplement for lignocellulosic biomass hydrolysis. *Biotechnology for Biofuels* **2019,** *12* (1).

Chapter 8

1. Zheng, J. P.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R. J.; Constantinou, P. E.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **2009,** *461* (7260), 74-77.

2. Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* **2016,** *138* (38), 12690-12690.

3. Simmons, C. R.; Zhang, F.; MacCulloch, T.; Fahmi, N.; Stephanopoulos, N.; Liu, Y.; Seeman, N. C.; Yan, H., Tuning the Cavity Size and Chirality of Self-Assembling 3D DNA Crystals. *Journal of the American Chemical Society* **2017,** *139* (32), 11254-11260.

4. Simmons, C. R.; MacCulloch, T.; Zhang, F.; Liu, Y.; Stephanopoulos, N.; Yan, H., A Self-Assembled Rhombohedral DNA Crystal Scaffold with Tunable Cavity Sizes and High-Resolution Structural Detail. *Angewandte Chemie International Edition* **2020,** *n/a* (n/a).

5. Zhang, F.; Simmons, C. R.; Gates, J.; Liu, Y.; Yan, H., Self-Assembly of a 3D DNA Crystal Structure with Rationally Designed Six-Fold Symmetry. *Angewandte Chemie-International Edition* **2018,** *57* (38), 12504-12507.

6. Jung, C.; Bandilla, P.; Von Reutern, M.; Schnepf, M.; Rieder, S.; Unnerstall, U.; Gaul, U., True equilibrium measurement of transcription factor-DNA binding affinities using automated polarization microscopy. *Nature communications* **2018,** *9* (1), 1605-11.

7. Ades, S. E.; Sauer, R. T., Specificity of Minor-Groove and Major-Groove Interactions in a Homeodomain-DNA Complex. *Biochemistry (Easton)* **1995,** *34* (44), 14601-14608.

8. Noyes, M. B.; Christensen, R. G.; Wakabayashi, A.; Stormo, G. D.; Brodsky, M. H.; Wolfe, S. A., Analysis of Homeodomain Specificities Allows the Family-wide Prediction of Preferred Recognition Sites. *Cell (Cambridge)* **2008,** *133* (7), 1277-1289.

9. Zhigang, S.; Veronica Ebu, I.; Haicheng, L.; Lekha, P.; Katrina, M. C.; Tom, C.; Gaetano, T. M.; Cory, A., Design of a "MinimAl" Homeodomain: The N-Terminal Arm Modulates DNA Binding Affinity and Stabilizes Homeodomain Structure. *Proceedings of the National Academy of Sciences - PNAS* **1994,** *91* (18), 8373-8377.

10. Tucker-Kellogg, L.; Rould, M. A.; Chambers, K. A.; Ades, S. E.; Sauer, R. T.; Pabo, C. O., Engrailed (Gln50→Lys) homeodomain–DNA complex at 1.9 Å resolution: structural basis for enhanced affinity and altered specificity. *Structure (London)* **1997,** *5* (8), 1047-1054.

11. Fraenkel, E.; Rould, M. A.; Chambers, K. A.; Pabo, C. O., Engrailed Homeodomain-DNA Complex at 2.2 A Resolution: A Detailed View of the Interface and Comparison with Other Engrailed Structures. *Journal of molecular biology* **1998,** *284* (2), 351-361.

12. Ades, S. E.; Sauer, R. T., Differential DNA-binding specificity of the engrailed homeodomain: The role of residue 50. *Biochemistry (Easton)* **1994,** *33* (31), 9187-9194.

13. Guerrero, L.; Smart, O. S.; Woolley, G. A.; Allemann, R. K., Photocontrol of DNA Binding Specificity of a Miniature Engrailed Homeodomain. *Journal of the American Chemical Society* **2005,** *127* (44), 15624-15629.

14. Delano, W. L. *The PYMOL Molecular Graphics System*, DeLano Scientific: San Carlos, CA, 2002.

15. Chen, C., A Practical Guide for Introducing Gibson Assembly®Gene Constructs into Chemically Competent Cells. *BioTechniques* **2015,** *58* (3).

16. Ghosh, B.; Boila, L. D.; Choudhury, S.; Mondal, P.; Bhattacharjee, S.; Pal, S. K.; Sengupta, A.; Roy, S., A Potent Conformation-Constrained Synthetic Peptide Mimic of a Homeodomain Selectively Regulates Target Genes in Cells. *ACS chemical biology* **2018,** *13* (8), 2003-2009.

17. Tripp, K. W.; Sternke, M.; Majumdar, A.; Barrick, D., Creating a Homeodomain with High Stability and DNA Binding Affinity by Sequence Averaging. *Journal of the American Chemical Society* **2017,** *139* (14), 5051-5060.

18. Bewley, C. A.; Gronenborn, A. M.; Clore, G. M., Minor groove-binding architectural proteins: Structure, function, and DNA recognition. *Annual review of biophysics and biomolecular structure* **1998,** *27* (1), 105-131.

19. Fonfría-Subirós, E.; Acosta-Reyes, F.; Saperas, N.; Pous, J.; Subirana, J. A.; Campos, J. L., Crystal structure of a complex of DNA with one AT-hook of HMGA1. *PloS one* **2012,** *7* (5), e37120-e37120.

20. Huth, J.; Bewley, C.; Nissen, M.; Evans, J.; Reeves, R.; Gronenborn, A.; Clore, G., The solution structure of an HMG-l(Y)-DNA

complex defines a new architectural minor

groove binding motif. *Nature Structural Biology* **1997,** *4*, 657-665.

21. Claus, P.; Schulze, E.; Wisniewski, J., Insect Proteins Homologous to Mammalian High Mobility Group Proteins I/Y. *Journal of biological chemistry* **1994,** *269*, 33042-33048.

22. Zlatanova, J.; Holde, K., Binding to four-way junction DNA: a common property of architectural proteins? *The FASEB journal* **1998,** *12* (6), 421-431.

23. Hill, D. A.; Reeves, R., Competition between HMG-I(Y), HMG-1 and histone H1 on four-way junction DNA. *Nucleic acids research* **1997,** *25* (17), 3523-3531.

24. Alniss, H. Y., Thermodynamics of DNA Minor Groove Binders. *Journal of medicinal chemistry* **2019,** *62* (2), 385-402.

25. Nunn, C. M.; Garman, E.; Neidle, S., Crystal Structure of the DNA Decamer d(CGCAATTGCG) Complexed with the Minor Groove Binding Drug Netropsin. *Biochemistry (Easton)* **1997,** *36* (16), 4792-4799.

26. Lewis, E. A.; Munde, M.; Wang, S.; Rettig, M.; Le, V.; Machha, V.; Wilson, W. D., Complexity in the binding of minor groove agents: netropsin has two thermodynamically different DNA binding modes at a single site. *Nucleic acids research* **2011,** *39* (22), 9649-9658.

27. Khan, G. S.; Shah, A.; Zia ur, R.; Barker, D., Chemistry of DNA minor groove binding agents. *Journal of photochemistry and photobiology. B, Biology* **2012,** *115*, 105-118.

28. Luis, A. M.; Kenneth, J. B., Origins of Netropsin Binding Affinity and Specificity: Correlations of Thermodynamic and Structural Data. *Proceedings of the National Academy of Sciences - PNAS* **1987,** *84* (13), 4359-4363.

29. Padroni, G.; Withers, J. M.; Taladriz-Sender, A.; Reichenbach, L. F.; Parkinson, J. A.; Burley, G. A., Sequence-Selective Minor Groove Recognition of a DNA Duplex Containing Synthetic Genetic Components. *Journal of the American Chemical Society* **2019,** *141* (24), 9555-9563.

30. Parolin, C.; Zanotti, G.; Palu, G., A Model for the Sequence-Dependent DNA Binding of 4′,6-Diamidino-2-Phenylindole (DAPI). *Biochemical and biophysical research communications* **1995,** *208* (1), 332-338.

31. Wilson, W. D.; Tanious, F. A.; Barton, H. J.; Jones, R. L.; Fox, K.; Wydra, R. L.; Strekowski, L., DNA sequence dependent binding modes of 4',6-diamidino-2-phenylindole (DAPI). *Biochemistry (Easton)* **1990,** *29* (36), 8452-8461.

32. Guo, P.; Farahat, A. A.; Paul, A.; Harika, N. K.; Boykin, D. W.; Wilson, W. D., Compound Shape Effects in Minor Groove Binding Affinity and Specificity for Mixed Sequence DNA. *Journal of the American Chemical Society* **2018,** *140* (44), 14761-14769.

Chapter 9

1. Seeman, N. C., DNA in a material world. *Nature* **2003,** *421* (6921), 427-431.

2. Seeman, N. C., Biochemistry and structural DNA nanotechnology: An evolving symbiotic relationship. *Biochemistry* **2003,** *42* (24), 7259-7269.

3. Seeman, N. C.; Sleiman, H. F., DNA nanotechnology. *Nature Reviews Materials* **2018,** *3* (1).

4. Pinheiro, A. V.; Han, D. R.; Shih, W. M.; Yan, H., Challenges and opportunities for structural DNA nanotechnology. *Nature Nanotechnology* **2011,** *6* (12), 763-772.

5. Hu, Y.; Niemeyer, C. M., From DNA Nanotechnology to Material Systems Engineering. *Advanced Materials* **2019,** *31* (26).

6. Ponnuswamy, N.; Bastings, M. M. C.; Nathwani, B.; Ryu, J. H.; Chou, L. Y. T.; Vinther, M.; Li, W. A.; Anastassacos, F. M.; Mooney, D. J.; Shih, W. M., Oligolysine-based coating protects DNA nanostructures from low-salt denaturation and nuclease degradation. *Nature Communications* **2017,** *8*.

7. Simmons, C. R.; MacCulloch, T.; Zhang, F.; Liu, Y.; Stephanopoulos, N.; Yan, H., A Self-Assembled Rhombohedral DNA Crystal Scaffold with Tunable Cavity Sizes and High-Resolution Structural Detail. *Angewandte Chemie International Edition* **2020,** *n/a* (n/a).

8. Simmons, C. R.; Zhang, F.; MacCulloch, T.; Fahmi, N.; Stephanopoulos, N.; Liu, Y.; Seeman, N. C.; Yan, H., Tuning the Cavity Size and Chirality of Self-Assembling 3D DNA Crystals. *Journal of the American Chemical Society* **2017,** *139* (32), 11254-11260.

9. Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* **2016,** *138* (38), 12690-12690.

APPENDIX A

DATA COLLECTION AND REFINEMENT STATISTICS FOR CHAPTERS 6-8

Table A.1. Data collection and refinement statistics for chapter 6 (4x6 system).

	D-DNA SAD	D-DNA Native	L-DNA SAD	L-DNA Native
Data Collection				
Resolution (Å)	50-3.1	50-3.05	50-3.0	50-3.0
Space group	$P3_2$	$P3_2$	$P3_1$	$P3_1$
Cell dimensions				
a, b, c (Å)	68.5,68.5,56.8	68.5,68.5,55.8	67.7,67.7,54.3	68.6,68.6,55.7
α, β, γ (°)	90,90,120	90,90,120	90,90,120	90,90,120
R_{merge}	0.084(0.25)	0.108(0.674)	0.186(0.721)	0.110(0.467)
$I/\sigma I$	27.6(2.6)	23.9(1.9)	41.1(2.6)	39.0(2.6)
Completeness (%)	98.0(92.1)	89.25(57.7)	86.8(58.6)	86.1(53.7)
Redundancy	3.7(3.2)	7.6(6.7)	10.4(9.3)	10.7(9.5)
Refinement				
# Reflections	—	4710	—	4949
R_{work}/R_{free}	—	21.13/23.66	—	23.76/24.57
# atoms DNA	—	851	—	851
RMS Deviations	—		—	
bond lengths (Å)	—	0.07	—	0.016
bond angles (°)	—	0.923	—	1.332

Table A.2. Data collection and refinement statistics for Ch.7 (4x5 Junction Variations).

Variation	**J1**	**J3**	**J5**	**J6**	**J7**
PDB Code: Junction	6X8C	6XDV	6XDW	6XDX	6XDY
PDB Code: Duplex	5KEK	6WQG	6WRB	6X8B	6WSN
Data Collection					
Beamline	NSLS X25	APS 19-ID	APS 19-BM	APS 19-ID	APS 19-ID
Space group	$P3_221$	$P3_2$	$P3_221$	$P3_2$	$P3_221$
Resolution (Å)	3.1	3	3.15	2.9	3.05
Cell dimensions					
a, b, c (Å)	67.9,67.9,59.3	68.9, 68.9,60.7	67.9,67.9,59.5	68.9,68.9,59.4	67.8,67.8,60.5
α, β, γ (°)	90,90,120	90,90,120	90,90,120	90, 90, 120	90, 90, 120
Wavelength (Å)	0.98	0.92	1	1	1
Total observations	29781	64132	35589	64266	53942
No. unique reflections	3041	6381	2891	6565	3237
R_{pim}	2.6(26.8)	3.2(21.2)	4.4(20.9)	4.4(35.4)	2.7 (27.0)
$CC_{1/2}$	1.001(0.888)	1.00(.938)	1.008(.882))	0.997(0.77)	1.007 (0.865)
$I/\sigma I$	40.08(2.96)	45.828(1.85)	17.38(1.72)	42.1(1.0)	55.438 (1.22)
Completeness (%)	99.4(93.5)	99.1(90.4)	98.0(82.8)	93.8(64.5)	99.6 (99.4)
Redundancy	9.8(7.0)	10.1(8.1)	12.3(7.3)	9.8 (8.0)	16.7 (12.4)
Refinement: Junction					
R_{work}/R_{free}	23.84/26.71	22.67/25.24	22.40/25.99	22.40/25.99	23.97/25.13
No. atoms					
DNA	855	855	855	855	855
ligand/ion	1	3	1	1	0
R.m.s deviations					
Bond lengths (Å)	0.004	0.007	0.012	0.012	0.004
Bond angles (°)	0.604	0.765	1.231	1.231	0.611
Refinement: Duplex					
R_{work}/R_{free}	20.42/25.97	23.30/25.05	25.08/26.33	23.78/25.05	23.13/52.95
No. atoms					
DNA	853	856	855	855	855
ligand/ion	2	3	2	5	0
R.m.s deviations					
Bond lengths (Å)	0.0139	0.006	0.005	0.005	0.004
Bond angles (°)	1.317	0.82	0.609	0.727	0.582

*The value for the highest-resolution shell is shown in parentheses

Variation	J8	J9	J10	J14	J15
PDB Code: Junction	6XDZ	6XEI	6XEJ	6XEK	6XEL
PDB Code: Duplex	6WSO	6WSP	6WSQ	6WSR	6WSS
Data Collection					
Beamline	APS 19-ID	APS 19-ID	APS 19-ID	APS 19-ID	APS 19-ID
Space group	$P3_2$	$P3_2$	$P3_221$	$P3_221$	$P3_2$
Resolution (Å)	3.1	3.05	3.05	2.85	3
Cell dimensions					
a, b, c (Å)	68.9,68.9,59.8	68.7,68.7,60.8	68.8,68.8,62.0	68.9,68.9,62.1	68.8,68.8,60.9
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Wavelength (Å)	0.92	0.92	0.98	0.92	0.92
Total observations	50597	57115	49240	80590	62192
No. unique reflections	5407	5886	4339	4213	6272
R_{pim}	3.2(27.9)	3.1(25.1)	2.7 (32.7)	1.9 (17.1)	2.9 (20.5)
$CC_{1/2}$	0.857 (.856)	0.972(.901)	0.992 (0.817)	0.917(0.916)	0.896 (.925)
$I/\sigma I$	42.19 (1.571)	33.2(1.556)	58.70(1.96)	66.19 (3.2)	40.88 (1.9)
Completeness (%)	94.0 (62.2)	95.9 (66.8)	98.5 (100.0)	99.8 (100.0)	97.3 (73.7)
Redundancy	9.4 (7.6)	9.7(7.7)	11.3 (11.7)	19.1 (19.4)	9.9 (8.4)
Refinement: Junction					
R_{work}/R_{free}	22.16/23.11	24.55/26.74	23.20/24.55	24.10/26.27	25.19/28.79
No. atoms					
DNA	855	855	855	855	855
ligand/ion	2	3	2	2	1
R.m.s deviations					
Bond lengths (Å)	0.005	0.005	0.011	0.011	0.005
Bond angles (°)	0.641	0.669	1.259	1.843	0.608
Refinement: Duplex					
R_{work}/R_{free}	19.87/22.23	26.14/28.89	24.30/29.19	23.42/24.47	24.83/27.21
No. atoms					
DNA	855	855	855	855	855
ligand/ion	3	3	1	2	2
R.m.s deviations					
Bond lengths (Å)	0.01	0.005	0.007	0.015	0.005
Bond angles (°)	0.962	0.71	0.859	1.422	0.633

*The value for the highest-resolution shell is shown in parentheses

Variation	J16	J19	J20	J21	J22

PDB Code: Junction	6XEM	6XFC	6XFD	6XFE	6XFF
PDB Code: Duplex	6WST	6WSU	6WSV	6WSW	6WSX
Data Collection					
Beamline	APS 19-ID	APS 19-ID	APS 19-ID	ALS 5.0.2	ALS 5.0.2
Space group	$P3_2$	$P3_221$	$P3_221$	$P3_2$	$P3_2$
Resolution (Å)	3.05	2.75	3.1	3.1	3.1
Cell dimensions					
a, b, c (Å)	69.0,69.0,61.3	68.5,68.5,60.8	67.6.67.6,60.4	69.0,69.0,59.4	69.4,69.4,59.4
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90,90,120	90,90,120
Wavelength (Å)	0.92	0.98	0.98	0.92	0.92
Total observations	54438	50811	54504	46129	44428
No. unique reflections	5999	4478	3074	4917	4852
R_{pim}	4.0 (26.8)	3.6 (52.6)	3.3 (21.4)	2.2 (29.2)	4.3(34.3)
$CC_{1/2}$	1.072 (.925)	0.957 (.652)	0.994 (0.946)	.990 (.825)	.974(.718)
$I/\sigma I$	33.0 (2.27)	54.91 (1.435)	56.70 (1.5)	35.4 (1.1667)	31.9 (1.163)
Completeness (%)	96.9 (75.9)	98.6 (99.6)	99.3 (95.5)	87.3 (50.6)	85.1 (51.2)
Redundancy	9.1 (7.0)	11.3 (9.3)	17.7 (11.8)	9.4(6.4)	9.2 (6.4)
Refinement: Junction					
R_{work}/R_{free}	22.49/25.39	21.82/24.23	22.31/28.58	20.85/23.34	22.91/26.69
No. atoms					
DNA	855	855	855	856	855
ligand/ion	1	3	2	2	0
R.m.s deviations					
Bond lengths (Å)	0.006	0.006	0.004	0.005	0.006
Bond angles (°)	0.766	0.771	0.604	0.662	0.808
Refinement: Duplex					
R_{work}/R_{free}	21.29/26.26	21.09/23.28	24.34/28.26	23.39/27.79	24.61/25.63
No. atoms					
DNA	855	855	855	854	855
ligand/ion	3	4	0	3	0
R.m.s deviations					
Bond lengths (Å)	0.01	0.009	0.009	0.004	0.011
Bond angles (°)	0.994	0.992	1.448	0.66	1.072

*The value for the highest-resolution shell is shown in parentheses

Variation	J23	J24	J25	J26	J28	J29

PDB Code: Junction	6XFG	6XFW	6XGM	6XFX	6XFY	6XGZ
PDB Code: Duplex	6WSY	6WSZ	6WT0	6WRJ	6WRI	6WT1
Data Collection						
Beamline	ALS 5.0.2	ALS 5.0.2	APS 19-ID	ALS 5.0.2	ALS 5.0.2	ALS 5.0.2
Space group	$P3_221$	$P3_2$	$P3_2$	$P3_221$	$P3_2$	$P3_2$
Resolution (Å)	3.05	3.1	3.1	3.1	3.05	3.1
Cell dimensions						
a, b, c (Å)	68.5,68.5,60.2	68.5,68.5,60.2	69.0,69.0,59.5	67.6,67.6,60.6	69.0,69.0,60.6	68.7,68.7,58.1
α, β, γ (°)	90,90,120	90,90,120	90,90,120	90,90,120	90,90,120	90,90,120
Wavelength (Å)	1	1	1	1	1	1
Total observations	61900	47136	44780	49300	56694	43300
No. unique reflections	3313	5010	5321	2902	5829	4640
R_{pim}	1.3(23.2)	2.6 (25.2)	8.1 (45.9)	2.0(8.3)	2.1(31.6)	5.3(18.2)
$CC_{1/2}$	.999(.905)	0.977 (0.864)	0.961 (0.822)	1.009(.991)	.988(.855)	1.164(.943)
$I/\sigma I$	51.48(2.0)	36.08 (1.27)	28.94 (1.51)	36.17 (3.36)	29.37(1.357)	22.38(1.769)
Completeness (%)	99.9(99.4)	85.1 (47.9)	93.1 (65.7)	96.3(66.9)	95.2 (64.6)	84.3(51.3)
Redundancy	18.7 (15.9)	9.4 (7.1)	8.4 (5.0)	17.0 (13.1)	9.7 (7.2)	9.3 (7.1)
Refinement: Junction						
R_{work}/R_{free}	21.46/24.05	24.31/25.47	21.90/23.46	22.65/27.07	22.69/24.79	24.41/26.22
No. atoms						
DNA	855	855	855	855	855	855
ligand/ion	2	2	2	0	3	2
R.m.s deviations						
Bond lengths (Å)	0.006	0.004	0.008	0.005	0.005	0.005
Bond angles (°)	0.829	0.667	1.602	0.724	0.664	0.673
Refinement: Duplex						
R_{work}/R_{free}	21.78/25.45	23.79/26.17	24.73/26.23	23.49/28.15	19.75/24.75	21.78/25.54
No. atoms						
DNA	855	855	855	855	855	855
ligand/ion	2	3	2	0	2	2
R.m.s deviations						
Bond lengths (Å)	0.007	0.753	0.005	0.005	0.009	0.008
Bond angles (°)	0.692	0.602	0.668	0.692	0.996	1.467

*The value for the highest-resolution shell is shown in parentheses

Variation	**J31**	**J32**	**J33**	**J34**	**J35**	**J36**

PDB Code: Junction	6XG0	6XGJ	6XGN	6XGO	6XGK	6XGL
PDB Code: Duplex	6WRC	6WR9	6WR7	6WRA	6WR5	6WR3
Data Collection						
Beamline	ALS 5.0.2	ALS 5.0.2	ALS 5.0.2	APS 19-ID	ALS 5.0.2	ALS 5.0.2
Space group	$P3_2$	$P3_2$	$P3_2$	$P3_2$	$P3_2$	$P3_2$
Resolution (Å)	3.15	3.05	3.1	3.0	3.05	3.15
Cell dimensions						
a, b, c (Å)	68.9,68.9,59.0	69.0,69.0,60.8	68.4,68.4,61.30	68.8,68.8,59.7	68.7, 68.7, 60.1	68.8,68.8,60.6
α, β, γ (°)	90,90,120	90,90,120	90,90,120	90, 90, 120	90,90,120	90, 90, 120
Wavelength (Å)	1	0.92	1	1	1	0.92
Total observations	42070	51714	40438	36218	52784	41936
No. unique reflections	4650	5625	4430	6287	5471	4722
R_{pim}	3.8(16.6)	2.3(17.7)	2.6 (27.9)	5.9 (34.3)	2.3(12.9)	4.2(15.4)
$CC_{1/2}$	.981(.947)	1.015(.953)	.998 (0.605)	0.911 (.915)	.985(.970)	.996(.930)
$I/\sigma I$	35.59(2.11)	29.12 (1.739)	30.84 (1.16)	9.578(2.114)	36.06(3.381)	26.66 (1.7)
Completeness (%)	88.3 (62.2)	93.4 (72.8)	78.1 (32.3)	99.4 (95.6)	91.9(62.1)	87.3(55.0)
Redundancy	9.0 (5.6)	9.2 (6.3)	9.1 (6.2)	5.8 (4.7)	9.6 (8.0)	8.9 (5.6)
Refinement: Junction						
R_{work}/R_{free}	23.88/26.14	22.44/25.31	20.09/22.85	21.12/23.34	19.65/21.08	21.72/23.96
No. atoms						
DNA	855	855	855	855	855	855
ligand/ion	2	3	0	2	2	3
R.m.s deviations						
Bond lengths (Å)	0.004	0.008	0.008	0.011	0.015	0.007
Bond angles (°)	0.657	0.894	1.322	1.76	1.463	0.806
Refinement: Duplex						
R_{work}/R_{free}	24.82/26.72	22.28/24.63	20.62/24.21	20.64/23.51	25.04/26.36	24.75/25.52
No. atoms						
DNA	855	855	855	855	856	855
ligand/ion	0	3	0	2	2	3
R.m.s deviations						
Bond lengths (Å)	0.006	0.009	0.007	0.011	0.005	0.005
Bond angles (°)	0.74	0.985	1.426	1.732	0.638	0.71

*The value for the highest-resolution shell is shown in parentheses

Table A.3. Data collection and refinement statistics for Ch.7 (4x6 Junction Variations).

Variation	J1	J2	J5	J7	J8
PDB Code: Junction	6XNA	7JFT	7JFU	7JFV	6XO5
PDB Code: Duplex	5VY6	7JPB	7JPA	7JPC	7JP9
Data Collection					
Beamline	ALS 8.2.2	NSLS 17-ID2	ALS 5.0.2	ALS 5.0.2	ALS 8.2.2
Resolution (Å)	3.05	3.15	3.15	3.1	3
Space group	$P3_2$	$P3_2$	$P3_2$	$P3_2$	$P3_2$
Cell dimensions					
a, b, c (Å)	68.44,68.44,55.68	69.11,69.11,56.40	68.17,68.17,55.46	68.01,68.01,54.15	68.30,68.30,54.28
α, β, γ (°)	90,90,120	90,90,120	90,90,120	90,90,120	90,90,120
Wavelength (Å)	1	0.98	1	0.92	1
Total observations	36158	27398	29992	17407	
No. unique reflections	4784	4138	3637	3716	3906
R_{pim}	4.5 (37.2)	4.8 (16.2)	3.6 (11.3)	3.7 (27.4)	4.0(27.9)
$CC_{1/2}$	0.945 (0.883)	0.991 (0.960)	1.066 (0.928)	0.978 (0.875)	0.983(0.87)
$I/\sigma I$	23.9 (1.9)	15.97 (2.33)	30.58 (3.0)	25.96 (1.29)	20.5 (0.79)
Completeness (%)	86.6 (55.7)	80.1 (41.5)	73.2 (36.3)	73.2 (34.4)	80.0 (44.3)
Redundancy	7.6 (6.7)	6.6 (3.2)	8.2 (6.2)	4.7(3.7)	4.6 (2.4)
Refinement: Junction					
R_{work}/R_{free}	20.17/23.42	22.42/24.68	21.32/26.32	19.27/23.44	21.19/24.68
No. atoms					
DNA	855	855	855	855	855
ligand/ion	4	2	0	0	1
R.m.s deviations					
Bond lengths (Å)	0.004	0.006	0.006	0.006	0.007
Bond angles (°)	0.646	0.773	0.759	0.776	1.551
Refinement: Duplex					
R_{work}/R_{free}	21.13/23.66	22.36/25.21	20.1/21.62	18.00/21.32	25.69/27.33
No. atoms					
DNA	851	855	855	855	855
ligand/ion	0	2	0	0	2
R.m.s deviations					
Bond lengths (Å)	0.07	0.005	0.007	0.007	0.006
Bond angles (°)	0.923	0.755	1.244	1.169	0.734
*The value for the highest-resolution shell is shown in parentheses					

Variation	J10	J16	J20	J22	J23

PDB Code: Junction	7JFW	7JFX	7JH8	7JH9	7JHA
PDB Code: Duplex	7JP8	7JP7	7JP6	7JP5	7JON
Data Collection					
Beamline	ALS 5.0.2	APS 19-ID	APS 19-BM	APS 19-BM	ALS 8.2.2
Resolution (Å)	3.1	3.2	3.1	3.1	3.1
Space group	$P3_2$	$P3_2$	$P3_2$	$P3_2$	$P3_2$
Cell dimensions					
a, b, c (Å)	67.74,67.74,53.48	68.15.68.15,53.79	68.07.68.07,56.06	68.55,68.55,55.36	68.63,68.63,55.96
α, β, γ (°)	90,90,120	90,90,120	90,90,120	90,90,120	90,90,120
Wavelength (Å)	0.98	0.92	1	1.1	0.98
Total observations	23461	30431	34237	22626	29383
No. unique reflections	4054	3636	4830	4420	4641
R_{pim}	4.6 (26.5)	2.9 (10.3)	2.6 (14.8)	6.6 (46.2)	5.0 (22.9)
$CC_{1/2}$	0.98 (0.892)	0.879 (0.948)	0.993 (0.958)	0.998	1.009(0.908)
$I/\sigma I$	23.96 (1.57)	21.66 (3.92)	32.41 (2.64)	20.81 (1.6)	15.43 (1.6)
Completeness (%)	82.1 (47.6)	79.0 (43.9)	90.3 (62.5)	83.2 (48.9)	86.8 (56.6)
Redundancy	5.8 (4.9)	8.4 (4.2)	7.1 (5.6)	5.1 (4.1)	6.3 (4.7)
Refinement: Junction					
R_{work}/R_{free}	19.22/21.73	24.25/28.00	24.61/27.24	23.81/25.80	51.5/22.6
No. atoms					
DNA	855	855	858	855	855
ligand/ion	0	0	0	2	2
R.m.s deviations					
Bond lengths (Å)	0.006	0.006	0.006	0.005	0.006
Bond angles (°)	0.809	1.119	0.663	0.694	0.798
Refinement: Duplex					
R_{work}/R_{free}	22.82/24.73	24.23/25.96	26.05/27.00	24.43/26.86	22.72/24.41
No. atoms					
DNA	855	855	855	855	855
ligand/ion	1	3	1	2	5
R.m.s deviations					
Bond lengths (Å)	0.005	0.005	0.006	0.005	0.005
Bond angles (°)	0.701	0.802	0.691	0.798	0.729

*The value for the highest-resolution shell is shown in parentheses

Variation	**J4**	**J5**	**J31**	**J33**	**J36**

PDB Code: Junction	7JHR	7JHS	7JHT	7JHU	7JHV
PDB Code: Duplex	7HRY	7JRZ	7JS0	7JS1	7JS2
Data Collection					
Beamline	ALS 5.0.2	APS 19-ID	ALS 5.0.2	APS 19-ID	ALS 5.0.2
Resolution (Å)	3.15	3.1	3.15	3.15	3.05
Space group	R3	R3	R3	R3	R3
Cell dimensions					
a, b, c (Å)	115.15,115.15,48.66	114.78,114.78,49.62	116.08,116.08,49.43	113.87,113.87,50.81	114.61,114.61,50.35
α, β, γ (°)	90,90,120	90,90,120	90,90,120	90,90,120	90,90,120
Space group	H3	H3	H3	H3	H3
Cell dimensions					
a, b, c (Å)	64.95	68.22	68.26	67.86	68.11
α, β, γ (°)	113.9	114.37	114.55	113.99	114.17
Wavelength (Å)	1	0.92	1	0.92	1
Total observations	28824	38645	28107	25502	45010
No. unique reflections	3297	4170	3347	3676	4633
R_{pim}	2.0 (10.9)	3.8 (21.6)	4.0 (18.7)	5.8 (21.0)	4.2 (24.1)
$CC_{1/2}$	1.019 (0.967)	0.972 (0.955)	0.948 (0.899)	0.905 (0.867)	0.924 (0.908)
$I/\sigma I$	35.85 (2.71)	11.78 (1.5)	27.18 (1.83)	40.30 (2.03)	40.82 (1.69)
Completeness (%)	79.4 (26.6)	94.0 (65.8)	81.1 (30.9)	86.4 (35.3)	99.4 (95.4)
Redundancy	8.7 (6.7)	9.3 (6.0)	8.4 (5.2)	6.9 (3.7)	9.7(8.2)
Refinement: Junction					
R_{work}/R_{free}	24.54/27.75	25.92/26.87	22.17/25.84	24.36/27.27	20.79/23.61
No. atoms					
DNA	855	855	855	849	855
ligand/ion	2	3	1	0	0
R.m.s deviations					
Bond lengths (Å)	0.004	0.005	0.005	0.005	0.007
Bond angles (°)	0.635	0.74	0.679	0.768	0.797
Refinement: Duplex					
R_{work}/R_{free}	25.06/28.32	25.97/29.21	20.14/21.68	22.09/26.33	18.81/22.12
No. atoms					
DNA	855	855	855	855	855
ligand/ion	1	4	2	0	0
R.m.s deviations					
Bond lengths (Å)	0.005	0.006	0.012	0.006	0.014
Bond angles (°)	0.661	0.674	1.211	0.767	1.246

Table A.4. Data collection and refinement statistics for Ch.7 (4x6 Scramble Junction Variations).

Variation	J1	J2
PDB Code: Junction	7JK0	7JJZ
PDB Code: Duplex	7JKD	7JKE
Data Collection		
Beamline	ALS 5.0.2	ALS 5.0.2
Resolution (Å)	3.05	3.05
Space group	$P3_2$	$P3_2$
Cell dimensions		
a, b, c (Å)	67.97,67.97,55.83	68.27,68.27,55.79
α, β, γ (°)	90,90,120	90,90,120
Wavelength (Å)	1	1
Total observations	23080	36297
No. unique reflections	4289	4231
R_{pim}	3.2 (21.3)	3.7 (30.7)
$CC_{1/2}$	1.027 (0.951)	0.981 (0.785)
$I/\sigma I$	30.28(2.31)	21.33 (1.125)
Completeness (%)	78.6 (49.1)	77.7 (46.1)
Redundancy	5.4 (5.1)	8.6 (7.5)
Refinement: Junction		
R_{work}/R_{free}	22.46/24.75	23.81/26.46
No. atoms		
DNA	855	855
ligand/ion	1	0
R.m.s deviations		
Bond lengths (Å)	0.004	0.005
Bond angles (°)	0.631	0.73
Refinement: Duplex		
R_{work}/R_{free}	23.62/27.00	25.15/27.49
No. atoms		
DNA	855	855
ligand/ion	1	3
R.m.s deviations		
Bond lengths (Å)	0.004	0.004
Bond angles (°)	0.597	0.563

Variation	J3	J5	J7	J8	J10
PDB Code: Junction	7JJY	7JJX	7JJW	7JJ6	7JJ5

PDB Code: Duplex	7JKG	7JKH	7JKI	7JKJ	7JKK
Data Collection					
Beamline	ALS 5.0.2	ALS 5.0.2	APS 19-ID	ALS 5.0.2	ALS 5.0.2
Resolution (Å)	2.85	3.1	3	3.05	2.8
Space group	R3	R3	R3	R3	R3
Cell dimensions					
a, b, c (Å)	112.21,112.21,50.99	113.26,113.26,49.90	113.72,113.72,52.05	113.17,113.17,50.78	113.08,113.08,52.13
α, β, γ (°)	90,90,120	90,90,120	90,90,120	90,90,120	90,90,120
Space group	H3	H3	H3	H3	H3
Cell dimensions					
a, b, c (Å)	66.94	67.33	67.44	66.97	67.66
α, β, γ (°)	113.9	114.14	113.69	114.05	113.63
Wavelength (Å)	1	1	1	1	1
Total observations	57963	33207	45718	36594	62932
Unique reflections	5596	3640	4711	4101	6150
R_{pim}	2.8 (35.5)	2.0 (17.3)	4.6 (26.2)	5.4 (27.8)	3.8 (38.8)
$CC_{1/2}$	0.986 (0.82)	1.012 (0.902)	0.951 (0.894)	0.933 (0.807)	1.069 (0.893)
$I/\sigma I$	44 (1.5)	36.13(2.04)	51.26 (2.36)	22.17 (1.15)	57.96 (1.35)
Completeness (%)	100 (100)	84.7 (39.3)	95.5 (89.6)	91.7 (54.1)	99.7 (100)
Redundancy	10.4 (9.2)	9.1 (7.4)	9.7 (7.1)	8.9 (6.0)	10.2 (9.4)
Refinement: Junction					
R_{work}/R_{free}	20.93/24.29	20.07/23.03	21.04/25.18	19.86/23.74	22.70/24.30
No. atoms					
DNA	855	855	855	855	855
ligand/ion	2	2	2	2	1
R.m.s deviations					
Bond lengths (Å)	0.008	0.006	0.005	0.008	0.005
Bond angles (°)	0.946	0.755	0.689	0.885	0.784
Refinement: Duplex					
R_{work}/R_{free}	20.86/23.42	24.45/25.41	21.44/25.79	22.85/24.08	23.02/25.68
No. atoms					
DNA	855	855	855	855	855
ligand/ion	4	0	2	2	4
R.m.s deviations					
Bond lengths (Å)	0.01	0.005	0.005	0.004	0.005
Bond angles (°)	1.016	0.739	0.709	0.723	0.729

Variation	**J14**	**J16**	**J19**	**J21**	**J22**
PDB Code: Junction	7JJ4	7JJ3	7JJ2	7JIQ	7JIP

PDB Code: Duplex	7JL9	7JLA	7JLB	7JLC	7JLD
Data Collection					
Beamline	ALS 5.0.2	BNL AMX	ALS 5.0.2	APS 19-ID	ALS 5.0.2
Resolution (Å)	3	3	3	3.05	3.15
Space group	R3	R3	R3	R3	R3
Cell dimensions					
a, b, c (Å)	112.75,112.75,51.06	113.31,133.31,52.18	111.97,111.97,50.99	113.38,113.38,51.31	113.02,113.02,49.34
α, β, γ (°)	90,90,120	90,90,120	90,90,120	90,90,120	90,90,120
Space group	H3	H3	H3	H3	H3
Cell dimensions					
a, b, c (Å)	66.82	59.48	67.61	67.63	66.87
α, β, γ (°)	113.89	111.95	113.81	113.84	114.16
Wavelength (Å)	1	1	1	0.92	1
Total observations	45460	48142	44318	47219	33416
Unique reflections	4615	4978	4591	4622	3586
R_{pim}	2.1 (29.9)	2.9 (23.6)	2.4 (21.7)	2.5 (23.7)	2.5 (11.0)
$CC_{1/2}$	1.001 (0.833)	1.015 (0.825)	0.979 (0.900)	0.943 (0.933)	0.976 (0.974)
$I/\sigma I$	43.75 (1.0)	33.71 (1.6)	41.57 (1.6)	47.89 (1.57)	35.41 (2.96)
Completeness (%)	97.5 (76.7)	99.5 (92.8)	96.8 (78.8)	99.2 (94.7)	89 (52)
Redundancy	9.9 (7.4)	9.7 (7.5)	9.7 (7.8)	10.2 (8.2)	9.3 (6.3)
Refinement: Junction					
R_{work}/R_{free}	22.60/26.75	19.98/22.59	22.79/26.80	18.97/20.93	21.40/23.61
No. atoms					
DNA	855	854	855	855	855
ligand/ion	2	5	2	0	1
R.m.s deviations					
Bond lengths (Å)	0.005	0.007	0.004	0.01	0.005
Bond angles (°)	0.701	0.905	0.657	10639	0.658
Refinement: Duplex					
R_{work}/R_{free}	22.41/25.48	21.80/25.36	23.27/27.56	20.66/23.58	19.99/22.67
No. atoms					
DNA	855	855	855	855	855
ligand/ion	2	2	2	0	2
R.m.s deviations					
Bond lengths (Å)	0.005	0.005	0.005	0.01	0.006
Bond angles (°)	0.801	0.737	0.803	1.704	0.806

Variation	**J23**	**J24**	**J26**	**J30**
PDB Code: Junction	7JIO	7JIN	7JIM	7JI9

PDB Code: Duplex	7JLE	7JLF	7JNJ	7JSB
Data Collection				
Beamline	APS 19-ID	APS 19-ID	BNL FMX	APS 19-ID
Resolution (Å)	3	2.9	3	3.1
Space group	R3	R3	R3	R3
Cell dimensions				
a, b, c (Å)	114.26,114.26,51.47	112.59,112.59,51.86	112.09,112.09,51.13	114.70.114.70,50.46
α, β, γ (°)	90,90,120	90,90,120	90,90,120	90,90,120
Space group	H3	H3	H3	H3
Cell dimensions				
a, b, c (Å)	68.34	65.86	67.08	69.2
α, β, γ (°)	113.93	113.7	113.8	114.34
Wavelength (Å)	1	1	1	1
Total observations	50842	53206	37394	41400
Unique reflections	4894	5331	4461	4290
R_{pim}	3.4 (22.5)	5.4 (52.2)	5.3 (47.5)	5.3 (22.0)
$CC_{1/2}$	1.003 (0.92)	1.064 (0.671)	0.931 (0.525)	0.955 (0.916)
$I/\sigma I$	47 (2.71)	49.12 (1.13)	15.5 (0.5)	43.42 (2.22)
Completeness (%)	99.7 (100)	99.0(96.4)	94.9 (64.2)	95.8 (83.9)
Redundancy	10.4 (9.3)	10.0 (7.6)	8.4 (5.3)	9.7 (6.7)
Refinement: Junction				
R_{work}/R_{free}	18.68/22.49	22.13/24.17	22.59/26.13	20.20/22.75
No. atoms				
DNA	855	855	855	855
ligand/ion	4	1	2	1
R.m.s deviations				
Bond lengths (Å)	0.01	0.006	0.005	0.006
Bond angles (°)	1.836	0.778	0.68	0.727
Refinement: Duplex				
R_{work}/R_{free}	20.37/22.83	22.49/23.2	20.82/24.27	22.72/24.15
No. atoms				
DNA	855	855	855	858
ligand/ion	3	2	5	1
R.m.s deviations				
Bond lengths (Å)	0.008	0.006	0.008	0.005
Bond angles (°)	0.881	0.825	0.822	0.748

Variation	**J31**	**J33**	**J34**	**J36**
PDB Code: Junction	7JI8	7JI7	7JI6	7JI5

PDB Code: Duplex	7JSC	7JNK	7JNL	7JLM
Data Collection				
Beamline	APS 19-ID	ALS 5.0.2	ALS 5.0.2	APS 19-ID
Resolution (Å)	2.95	3.1	3	2.7
Space group	R3	R3	R3	R3
Cell dimensions				
a, b, c (Å)	113.11,113.11,50.39	113.57,113.57,51.80	111.80.111.80,51.30	112.71,112.71,50.62
α, β, γ (°)	90,90,120	90,90,120	90,90,120	90,90,120
Space group	H3	H3	H3	H3
Cell dimensions				
a, b, c (Å)	67.39	67.67	66.5	67.14
α, β, γ (°)	114	113.78	113.68	113.91
Wavelength (Å)	0.92	1	1	0.92
Total observations	50184	43588	44343	67034
Unique reflections	4983	4441	4591	6544
R_{pim}	2.9 (39.2)	3.9 (19.1)	2.4 (19.5)	2.5 (35.4)
$CC_{1/2}$	1.018 (0.786)	0.919 (0.936)	1.05 (0.947)	0.924 (0.746)
$I/\sigma I$	52.72 (1.24)	25.6 (1.57)	50.73 (2.03)	54.97 (1.39)
Completeness (%)	98.5 (87.6)	98.7 (85.6)	96.3 (70.3)	99.7 (98.8)
Redundancy	10.1 (7.3)	9.8 (6.9)	9.7 (8.4)	10.2 (8.8)
Refinement: Junction				
R_{work}/R_{free}	25.46/28.14	21.84/23.14	20.81/24.16	22.91/26.77
No. atoms				
DNA	855	855	855	855
ligand/ion	1	2	1	1
R.m.s deviations				
Bond lengths (Å)	0.005	0.005	0.007	0.006
Bond angles (°)	0.731	0.769	0.882	0.781
Refinement: Duplex				
R_{work}/R_{free}	25.05/28.22	23.47/24.44	23.23/28.00	22.36/24.35
No. atoms				
DNA	858	858	858	858
ligand/ion	1	2	4	2
R.m.s deviations				
Bond lengths (Å)	0.007	0.006	0.007	0.006
Bond angles (°)	0.811	0.676	0.764	0.737

Table A.5. PDB Accession Codes for all junction variations.

4x5	4x6 (P32)	4x6 (R3)	4x6 Scramble

	Duplex	Junction	Duplex	Junction	Duplex	Junction	Duplex	Junction
1	5KEK	6X8C	5VY6	6XNA	—	—	7JKD	7JK0
2	-—	—	7JPB	7JFT	—	—	7JKE	7JJZ
3	6WQG	6XDV	—	—	—	—	7JKG	7JJY
4	—	—	—	—	7JRY	7JHR	—	—
5	6WRB	6XDW	7JPA	7JFU	7JRZ	7JHS	7JKH	7JJX
6	6X8B	6XDX	—	—	—	—	—	—
7	6WSN	6XDY	7JPC	7JFV	—	—	7JKI	7JJW
8	6WSO	6XDZ	7JP9	6XO5	—	—	7JKJ	7JJ6
9	6WSP	6XEI	—	—	—	—	—	—
10	6WSQ	6XEJ	7JP8	7JFW	—	—	7JKK	7JJ5
11	—	—	—	—	—	—	—	—
12	—	—	—	—	—	—	—	—
13	—	—	—	—	—	—	—	—
14	6WSR	6XEK	—	—	—	—	7JL9	7JJ4
15	6WSS	6XEL	—	—	—	—	—	—
16	6WST	6XEM	7JP7	7JFX	—	—	7JLA	7JJ3
17	—	—	—	—	—	—	—	—
18	—	—	—	—	—	—	—	—
19	6WSU	6XFC	—	—	—	—	7JLB	7JJ2
20	6WSV	6XFD	7JP6	7JH8	—	—	—	—
21	6WSW	6XFE	—	—	—	—	7JLC	7JIQ
22	6WSX	6XFF	7JP5	7JH9	—	—	7JLD	7JIP
23	6WSY	6XFG	7JON	7JHA	—	—	7JLE	7JIO
24	6WSZ	6XFW	7JOL	7JHB	—	—	7JLF	7JIN
25	6WT0	6XGM	—	—	—	—	—	—
26	6WRJ	6XFX	7JOK	7JHC	—	—	7JNJ	7JIM
27	—	—	—	—	—	—	—	—
28	6WRI	6XFY	7JOJ	6XO6	—	—	—	—
29	6WT1	6XFZ	—	—	—	—	—	—
30	—	—	7JOI	6XO7	—	—	7JSB	7JI9
31	6WRC	6XG0	7JOH	6XO8	7JS0	7JHT	7SJC	7JI8
32	6WR9	6XGJ	—	—	—	—	—	—
33	6WR7	6XGN	7JOG	6XO9	7JS1	7JHU	7JNK	7JI7
34	6WRA	6XGO	—	—	—	—	7JNL	7JI6
35	6WR5	6XGK	—	—	—	—	—	—
36	6WR3	6XGL	—	—	7JS2	7JHV	7JNM	7JI5

Table A.6. Data collection and refinement statistics for Ch.8 (4x7 system).

Variation	Anomalous	Native	PC Variant	2bp TA	3-Turn
PDB Code	6V6R	6U40	6UEF	6UDN	6UAL
Data Collection					

Beamline	ALS 5.0.2	NSLS AMX 17-ID-1	ALS 5.0.2	ALS 5.0.2	NSLS AMX 17-ID-1
Resolution (Å)	50-2.7	50-2.7	50-2.95	50-2.6	50-4.5
Space group	R3	R3	R3	R3	R3
Cell dimensions					
a, b, c (Å)	68.49	68.33	68.53	68.64	104.53
α, β, γ (°)	115.39	115.41	115.55	115.37	117.74
Space group	H3	H3	H3	H3	H3
Cell dimensions					
a, b, c (Å)	115.8,115.8,44.8	116.6,116.6,45.0	115.9,115.9,44.0	116.1,116.1,45.0	179.3,179.3,48.1
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Wavelength (Å)	0.92	1	1	1	1
Total observations	29434	61013	46133	69622	27803
No. unique reflections	12116	12514	4634	6933	3195
R_{pim}	0.049(0.53)	0.037(0.311)	0.030(0.285)	0.032(0.334)	0.034(0.294)
$CC_{1/2}$	0.93(0.621)	0.988(0.825)	1.004(0.886)	0.986(0.754)	1.0(0.867)
$I/\sigma I$	34.41(1.13)	38.74(1.8)	52.48(2.21)	55.64(1.69)	50.86(1.39)
Completeness (%)	98.5(92.5)	99.7(100)	99.8(100)	99.9(100)	95.1(73.0)
Redundancy	2.4(1.9)	4.9(3.8)	10.0(9.1)	10.0(9.8)	8.7(7.1)
Refinement					
Resolution (Å)	50-2.7	50-2.7	50-2.95	50-2.6	50-4.5
R_{work}/R_{free}	0.2146/0.2786	0.2388/0.2545	0.2352/0.2559	0.2339/0.2556	0.1974/0.2427
No. atoms					
DNA	1307	1334	1334	1335	2033
ligand/ion	2	2	0	2	0
R.m.s deviations					
Bond lengths (Å)	0.005	0.005	0.005	0.006	0.005
Bond angles (°)	0.753	0.776	0.791	0.847	0.679

*The value for the highest-resolution shell is shown in parentheses

Table A.7. Data collection and refinement statistics for Ch.8 (4x7 SE Variants).

Variation	2bp GC	1bp C	1bp T	3bp CAC	3bp AGT	3-Turn TA
Data Collection						
Beamline	ALS 5.0.2	ALS 5.0.2	ALS 5.0.2	APS 19-BM	APS 19-ID	NSLS AMX 17-ID-1
Space group	P6	P6	P1	P6	P6	P_12_1
Resolution (Å)	50-4.2	50-4.6	50-4.4	50-4.0	50-5.0	50-4.7
Cell dimensions						
a, b, c (Å)	68.85,68.85,196.61	68.68,68.68,177.39	69.51,70.66,112.57	68.64,68.64,174.43	68.59,68.59,194.29	86.70,43.20,101.90
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 90	90, 90, 120	90, 90, 120	90, 110, 90
Wavelength	1	1	1	1	1	1
Total observations	56710	43160	41838	51259	42911	23141
No. unique reflections	3508	2385	12431	3967	2199	3696
R_{pim}	0.051(0.296)	0.029(0.233)	0.061(0.427)	0.028(0.432)	0.051(0.408)	0.061(0.467)
$CC_{1/2}$	0.909(0.831)	1.014(0.852)	0.928(0.792)	0.986(0.779)	1.058(0.648)	0.994(0.916)
$I/\sigma I$	31.88(1.375)	44.37(1.7)	22.46(1.18)	36(1.93)	55.29(1.08)	4.30(120)
Completeness (%)	91.0(74.1)	91.6(69.4)	95.7(83.2)	99.6(99.5)	96.1(83.0)	98.0(87.5)
Redundancy	16.2(14.6)	18.1(17.3)	3.4(3.1)	12.9(10.3)	19.5(18.9)	6.3(5.4)

*The value for the highest-resolution shell is shown in parentheses

APPENDIX B

PUBLISHED PORTIONS OF THIS DISSERTATION

Portions of the data presented in this dissertation have been previously published, submitted for publication, or will be submitted for publication in following journals. The published materials were included with permission from all co-authors.

Chapter 2

B. Smolková, T. MacCulloch, A. Frtús, M. Uzhytchak, M. Lunova, M. Hof, P. Jurkiewicz, A. Dejneka, N. Stephanopoulos*, O. Lunov*, "Effect of the protein corona on endosomal escape of functionalized DNA nanostructures" *ACS Applied Materials*, 2021, In Review.

Chapter 3

T. MacCulloch, N. Stephanopoulos*, "Proximity-enhanced synthesis of alternating DNA-peptide block copolymers", 2021, *ChemComm*, In preparation

Chapter 5

C. Simmons, F. Zhang, T. MacCulloch, N.E. Fahmi, N. Stephanopoulos, Y. Liu, N. Seeman, H. Yan*, "Tuning the Cavity Size and Chirality of Self-Assembling 3D DNA Crystals" J. Am. Chem. Soc. 2017, 139, 11254-11260.

Chapter 6

C.R. Simmons‡, T. MacCulloch‡, M. Krepl, M. Matthies, A. Buchberger, I. Crawford, J. Sponer, P. Sulc, Y. Liu, N. Stephanopoulos*, H. Yan*, "A Comprehensive Structural and Computational Toolbox of Immobile Holliday Junctions for DNA-directed Self-assembly" (manuscript in preparation; ‡ co-first authors)

Chapter 7

C.R. Simmons‡, T. MacCulloch‡, F. Zhang, Y. Liu, N. Stephanopoulos*, H. Yan*, "A Self-Assembled Rhoombohedral DNA Crystal Scaffold with Tunable Cavity Sizes and High-Resolution Structural Detail" *Angew. Chem. Int. Ed*. 2020, 59, 18619-18626. (‡co-first authors)

APPENDIX C

PERMISSIONS

Chapter 1:

Figure 1.1B reprinted from Pinheiro, A. V.; Han, D. R.; Shih, W. M.; Yan, H., Challenges and opportunities for structural DNA nanotechnology. *Nature Nanotechnology* **2011,** *6* (12), 763-772 with permission from Springer Nature. .

Figure 1.3B reprinted from Seeman, N. C., DNA nanotechnology: Novel DNA constructions. *Annual Review of Biophysics and Biomolecular Structure* **1998,** *27*, 225-248 with permission from

Figure 1.3C reprinted Chen, J. H.; Seeman, N. C., SYNTHESIS FROM DNA OF A MOLECULE WITH THE CONNECTIVITY OF A CUBE. *Nature* **1991,** *350* (6319), 631-633 with permission from Springer Nature.

Figure 1.3E reprinted from Winfree, E.; Liu, F. R.; Wenzler, L. A.; Seeman, N. C., Design and self-assembly of two-dimensional DNA crystals. *Nature* **1998,** *394* (6693), 539-544 with permission from Springer Nature.

Figure 1.4A reprinted from Rothemund, P. W. K., Folding DNA to create nanoscale shapes and patterns. *Nature* **2006,** *440* (7082), 297-302 with permission from Springer Nature.

Figure 1.4B reprinted from Douglas, S. M.; Dietz, H.; Liedl, T.; Hogberg, B.; Graf, F.; Shih, W. M., Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **2009,** *459* (7245), 414-418 with permission from Springer Nature.

Figure 1.4C reprinted from Han, D. R.; Pal, S.; Nangreave, J.; Deng, Z. T.; Liu, Y.; Yan, H., DNA Origami with Complex Curvatures in Three-Dimensional Space. *Science* **2011,** *332* (6027), 342-346 from AAAS.

Figure 1.4D reprinted from Douglas, S. M.; Bachelet, I.; Church, G. M., A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science* **2012,** *335* (6070), 831-834 from AAAS.

Figure 1.4E reprinted from Ke, Y. G.; Ong, L. L.; Shih, W. M.; Yin, P., Three-Dimensional Structures Self-Assembled from DNA Bricks. *Science* **2012,** *338* (6111), 1177-1183 from AAAS.

Figure 1.9A reprinted from Lou, C. G.; Martos-Maldonado, M. C.; Madsen, C. S.; Thomsen, R. P.; Midtgaard, S. R.; Christensen, N. J.; Kjems, J.; Thulstrup, P. W.; Wengel, J.; Jensen, K. J., Peptide-oligonucleotide conjugates as nanoscale building blocks for assembly of an artificial three-helix protein mimic. *Nature Communications* **2016,** *7* from Springer Nature.

Figure 1.9B reprinted from Williams, B. A. R.; Diehnelt, C. W.; Belcher, P.; Greving, M.; Woodbury, N. W.; Johnston, S. A.; Chaput, J. C., Creating Protein Affinity Reagents by Combining Peptide Ligands on Synthetic DNA Scaffolds. *Journal of the American Chemical Society* **2009,** *131* (47), 17233-17241 from American Chemical Society.

Figure 1.9C reprinted from Stephanopoulos, N.; Freeman, R.; North, H. A.; Sur, S.; Jeong, S. J.; Tantakitti, F.; Kessler, J. A.; Stupp, S. I., Bioactive DNA-Peptide Nanotubes Enhance the Differentiation of Neural Stem Cells Into Neurons. *Nano Letters* **2015,** *15* (1), 603-609 from American Chemical Society.

Figure 1.9D reprinted from Kye, M.; Lim, Y. B., Reciprocal Self-Assembly of Peptide-DNA Conjugates into a Programmable Sub-10-nm Supramolecular Deoxyribonucleoprotein. *Angewandte Chemie-International Edition* **2016,** *55* (39), 12003-12007 with permission from Wiley.

Chapter 4:

Figure 4.2 reprinted from Zheng, J. P.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R. J.; Constantinou, P. E.; Ginell, S. L.; Mao, C. D.; Seeman, N. C., From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **2009,** *461* (7260), 74-77 from Springer Nature.

Figure 4.5 reprinted from Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* **2016,** *138* (38), 12690-12690 from American Chemical Society.

Figure 4.7 reprinted from Zhang, F.; Simmons, C. R.; Gates, J.; Liu, Y.; Yan, H., Self-Assembly of a 3D DNA Crystal Structure with Rationally Designed Six-Fold Symmetry. *Angewandte Chemie-International Edition* **2018,** *57* (38), 12504-12507 with permission from Wiley.

Chapter 5:

Figure 5.1-Figure5.6 reprinted from Simmons, C. R.; Zhang, F.; MacCulloch, T.; Fahmi, N.; Stephanopoulos, N.; Liu, Y.; Seeman, N. C.; Yan, H., Tuning the Cavity Size and Chirality of Self-Assembling 3D DNA Crystals. *Journal of the American Chemical Society* **2017,** *139* (32), 11254-11260 from American Chemical Society.

Chapter 7:

Figure 7.1 reprinted from reprinted from Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* **2016,** *138* (38), 12690-12690 from American Chemical Society.

Figure7.2-7.12 reprinted from C.R. Simmons‡, T. MacCulloch‡, F. Zhang, Y. Liu, N. Stephanopoulos*, H. Yan*, "A Self-Assembled Rhombohedral DNA Crystal Scaffold with Tunable Cavity Sizes and High-Resolution Structural Detail" *Angew. Chem. Int. Ed.* 2020, 59, 18619-18626 with permission from Wiley.

APPENDIX D

MATERIALS AND METHODOLOGY

DNA purification. All oligos were purchased from Integrated DNA technologies (Coralville, Iowa) and purified using a 14% urea based denaturing polyacrylamide gel which used 1x TBE (89 mM Tris,89 mM boric acid, 2mM EDTA running buffer at 45 volts for 2 hours. The desired band was detected using UV shadowing, cut out, and excised using an elution buffer comprised of ammonium acetate, magnesium acetate, and EDTA. The DNA was crashed out using 100% ethanol and pelleted by centrifugation at 16000 RPM for 5 minutes. The pellet was resuspended in nanopure water and the concentration was obtained using a Thermofisher nanodrop.

Peptide Synthesis. All peptides were synthesized on a CEM Liberty Blue microwave assisted synthesizer at a 0.1 or 0.05 mmol scale using a rink amide resin and standard Fmoc chemistry. Briefly, a 20% piperidine solution was used for deprotection, 0.5 M diisopropylcarbodiimide was used as an activator, and a solution of 1 M oxyma with 0.1 M diisopropylethylamine was used as an activator base. Amino acids were added to the resin at a concentration of 0.2 M and coupled for four minutes up to the 15 amino acids, afterwhich they were double coupled. Further more any residues after number 40 were double coupled for 6 minutes each. The peptide was cleaved from the resin at room temperature over 3 hours using 95:2.5:2.5 mixture of trifluoroacetic acid (TFA):triisopropyl silane (TIS):water or over 4 hours using 90:5:2.5:2.5 mixture of TFA:DODT:TIS:water when a cysteine residue was present. The crude peptide was precipitated into cold ethyl ether, centrifuged to form a pellet, and resuspended in water with 0.1 M TFA. A Waters HPLC was utilized to purify the peptide over a gradient of 0-80% acetonitrile with 0.1% TFA. Fractions with an absorbance (230 nm) reaching a threshold of 300 mAU were collected and analyzed using matrix assisted laser desorption

ionization time-of-flight (MALDI-TOF) mass spectrometry. All pure fractions with the corresponding peptide mass were pulled together and lyophilized.

MALDI MS Characterization. All samples were characterized using a Bruker Microflex LRF MALDI mass spectrometer. Peptides were analysized in positive reflector mode using α-cyanohydroxycinnaminic acid as a matrix with a laser intensity of ~30%. Note that some larger peptides were analyzed in linear positive mode. DNA and DNA-peptide conjugates used hydroxypicolinic acid or 6-aza-2-thiothymine (0.7 M) each with ammonium ascorbate added at a 70 mM concentration both suspended in a 50:50 mixture of water:acetonitrile) as a matrix. They were shot using positive linear mode at a laser intensity of ~80-100%, in cases that the sample would not fly, it was switched to negative linear mode.

Chapter 2

Fabrication and characterization of DNs All oligonucleotides were obtained from Integrated DNA Technologies (Coralville, Iowa) and purified using 14% urea-based denaturing polyacrylamide gel electrophoresis (PAGE). One strand was labeled with AlexaFluor-488 for imaging in the agarose gels and in microscopy experiments. Each strand was added to a mixture at 10 μM in 1x tris-acetic acid-EDTA (TAE) buffer with 12.5 mM MgCl2 and annealed from 95 to 4° C over 2 hours. The successful formation of the 6-helix bundle was confirmed using agarose gel electrophoresis.

DN coating and characterization The DNs (1 μM) were mixed with the desired K10 at a 1:1 N:P ratio and incubated at room temperature for a minimum of 2 hours. All coated DNs used for cell experiments utilized the pHrodo-labeled K10 at 20 mol% of the total K10 concentration. All coated DNs run on agarose gels utilized the fluorescein labeled

K10 at 20 mol% of the total K10 concentration. In order to determine the optimal N:P ratio for complete coating of the DNs, the structures were electrophoresed using 1.5% agarose gels at 65V for 60 minutes and imaged using the fluorescein labeled K10.

Chapter 3

Synthesis of Peptide-Oligo Conjugates using Copper Free Click. The amine modified oligo was put up into 1x PBS (pH 7.5) at a concentration of 1 μM. NHS-Sulfo-DBCO ester dissolved in DMSO was added to the oligo at a 5 molar excess and left at room temperature for four hours to react, after which a second aliquot of the NHS-sulfo-DBCO was added to react overnight. The DBCO modified DNA was separated from the unreacted DNA using reverse phase chromatography on an Agilent 1220 series HPLC on a Zorbax Eclipse C18 column using 50 mM triethylammonium acetate and methanol as the running buffers. A gradient of 0-70% methanol was applied over 45 minutes while monitoring both 260 and 309 nm, corresponding to the absorbance of the DNA and the DBCO respectively. The peak with an absorbance for both components was collected and transferee back into water using a 3kd molecular weight cut off filter. The DBCO modified DNA was then mixed with the azidolysine containing peptide in a 1:4 ratio in 1x PBS (pH 7.5) and allowed to react overnight. The DNA-peptide was purified out from the DNA-DBCO using the same HPLC method as the DNA-DBCO. All peaks were characterized using MALDI ms to identify the desired peak.

Synthesis of a DNA-peptide-DNA Conjugate using Proximity Aided Copper Catalyzed Click An amine modified DNA was reacted NHS-PEG_4-azide and purified out using the same method as the DBCO-DNA. The azide-DNA was identified using MALDI MS. The peptide oligonucleotide conjugate ,obtained from the copper free click,

and the azide modified DNA strand were annealed at a 1:1 stoichiometry in 2x PBS (pH 7.5) at a concentration of 15 μM using a thermal gradient of 95 to 4 °C over 1 hour. A copper catalyzed click was utilized to conjugate the DNA-peptide (containing a propargylalanine) to the azide modified DNA. A mixture was made with final concentrations of 4 μM duplex, 1 mM aminoguanidine, 0.4 μM (1:5) Copper:THPTA and 20 mM PBS. 40 mM sodium ascorbate was added to the mixture to reduce the oxidation state of the copper and the reaction was carried out for 3 hours at room temperature at which point it was quenched with 250 mM EDTA.

Ligation to create the DPDPD Directly following the CuACC and addition of the EDTA, the fully annealed duplex was buffer exchanged into water, afterwhich 1x ligation buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM ATP, and 10 mM DTT) was added. 2 units of T4 Ligase was added and allowed to react for overnight at 16°C. The ligase was heat deactivated by incubated at 65°C for 10 minutes.

Denaturing PAGE Formation of the DPD and DPDPD were monitored using 8% urea based denaturing polyacrylamide gel electrophoresis. The gel was run in TBE buffer for 80 minutes at 45 V and imaged after staining with SybrGold.

MMP Degradation Human derived MMP-8 was added to the DX tiles linked by the DNA-MMP-DNA and incubated for 48 hours at 37°C for complete degradation. Each individual tile, one containing the azide, one containing a DP, was run alongside the tiles linked by the DPD with and without MMP incubation on a 6% native PAGE. The gel was run using TAE buffer at 200 mAmps for 3 hours.

Crystallization. All oligonucleotides were purchased from Integrated DNA Technologies (Coralville, Iowa) and purified using urea based denaturing polyacrylamide

gel electrophoresis. Following purification, the DNA was resuspended in nanopure H_2O and washed 5 times using 3 kD molecular weight cut off filters (Amicon) to remove any remaining salts. A mixture of the three component strands (S1:S2:S3) was made with final concentrations of 30:120:120 μM for each design and junction. Each mixture was screened using broad matrix screen of 48 different conditions (shown below), adapted from a discontinued Sigma-Aldrich kit, using sitting drop-vapor diffusion. Each reservoir contained 0.5 mL of the crystallization buffer while each well contained 4 μL of the DNA stock and 2 μL of the corresponding buffer from the reservoir. A thermal gradient of j60 to 25ºC with a 0.3 ºC/hour ramp was applied to the trays using a Torrey Pines chilling incubator. Prior to harvesting, the crystals were cryo-protected using a 30% glycerol solution of the mother liquor. All crystals were then harvested using cryo-loops (Hampton Research) and cryo-cooled by immediate submersion in liquid nitrogen. Data were collected in a nitrogen cold-stream (100 K) at the various beamlines.

Chapter 5

Denaturing PAGE for DNase Degradation A duplex consisting of the right handed S2 and its complement was annealed from 95°C to 4°C over 30 minutes in 50 mM HEPES (pH 7.5) with 80 mM $MgCl_2$ and 2.5 mM spermine at 100 μM. 10 μL of the stock was incubated at 37°C with 1 μL (2 units) of DNase I for 0, 5, 10, and 15 minutes. Another 10 μL of the stock was incubated at 37°C with 1 μL of heat inactivated (75°C for 10 minutes) DNase I for 15 minutes. Each sample underwent a 10-fold dilution with H_2O and was run on a 15 % native polyacrylamide gel, along with a sample with no added DNase, and a 10-base pair double stranded DNA ladder.

Time-Course Degradation Study DNA crystallization was performed as outlined above for both the right handed and left handed 4 X 6 motifs in 50 mM HEPES (pH 7.5) with 80 mM $MgCl_2$ and 2.5 mM spermine. Crystals were imaged using a LEICA S6D microscope and application suite at time 0 min. 1 µL (2 Units) of DNase 1 (New England Biolabs) was added to the drop and it was resealed and incubated at 37°C. The tray was removed and the drops were imaged again at time intervals of 10 min., 20 min., 40 min., 1 hr., 2 hr., 4 hr., 6 hr, 8 hr., and overnight. This experiment was repeated with 2 other crystallization buffers; 50 mM cacodylate (pH 6.0) with 10 mM $MgCl_2$, 2.5 mM spermine, and 2.5 M NaCl and 50 mM cacodylate (pH 6.5) with 5 mM $CoH_{18}N_6$ and 2.5 mM KCl.

Chapter 7

Crystal Setup All oligonucleotides were purchased from Integrated DNA technologies (IDT; Coralville, IA). Brominated strands were purified by HPLC from IDT, and all others were purified using 14% denaturing PAGE. Following purification, each strand was resuspended in nanopore water to a final concentration of 300 µM. DNA stocks were made in a 1:4:4 (S1:S2:S3) with the central strand at 30 µM. Crystallization screens were performed as described previously[1-3], and crystal "hits" from the broad matrix screen were prepared for crystal optimization. Crystallization of each construct in this study was carried out using sitting drop vapor diffusion with 500 µL of reservoir buffer and a total drop volume of 6 µL with a 2:1 ratio of DNA stock to buffer. The "native" and 3-turn crystals were prepared in a buffer composed of 50 mM sodium cacodylate pH 7.0 with 20 mM $MgCl_2$, 1.0 mM spermine, 1.0 mM $CoH_{18}N_6$, and 15% ethanol while the bromine derivatized crystals were grown in 50 mM sodium cacodylate pH 6.5 with 100 mM $MgCl_2$,

1.0 mM $CoH_{18}N_6$, and 10% ethanol. The TA sticky end motif was crystallized in 50 mM cacodylate pH 6.0 with 20mM $MgCl_2$, 1.0 mM spermine, and 15% ethanol, and the PC variant used 50 mM TRIS pH 8.0 with 200 mM $MgCl_2$ and 15% ethanol. The crystal trays were heated in a chilling incubator (Torrey Pines) to 60°C for 1 hour and then the temperature was ramped using a linear gradient to 25°C at a rate of 0.3°C/hr. A stock of artificial mother liquor supplemented with 30% glycerol was added to the drop, and the crystals were harvested using cryo-loops (Hampton Research) and cryo-cooled by immediate submersion in liquid nitrogen. Data were collected in a nitrogen cold-stream (100 K) at the corresponding light sources indicated in Table S2.

Data Collection and Refinement The resulting diffraction data were processed in HKL2000[4], and indexed in both the rhombohedral (R3) and hexagonal (H3) settings as reported in Table 1. The structures described here are made in reference to the R3 cell parameters. The substructure of the bromine derivatized crystals with initial positions, density modification, solvent flattening and map calculations was performed in the SHELXC/D/E[5] suite in hkl2map[6]. Initial bromine positions were further confirmed with HySS and AutoSol from the PHENIX[7] package. Initial phasing, refinement, and model building was carried out using SOLVE and RESOLVE[8] from AutoBuild. Manual iterations of model building were performed in Coot, and the initial model was treated as a single rigid body during refinement with subsequent rounds using restrained refinement in REFMAC[9] from CCP4[10], and real space, XYZ coordinate calculation in phenix.refine. Atom occupancies and B-factor calculations were then refined, followed by simulated annealing to conclude refinement using an R_{free} set containing 5% of the unique reflections for each structure described here. The native, point mutation, and three-turn

structures were then determined using molecular replacement using the derivative structure as the initial search using PHASER[11] with model building and refinement executed as described above. All coordinates and structure factors were deposited in the Protein Data Bank (PDB) with accession codes 6V6R, 6U40, 6UEF, 6UDN, and 6UAL. All structure figures contained in this report were prepared using PyMOL

Chapter 8

Co-crystallization Screen All three component strands (S1:S2:S3) were added to nanopure water in the correct ratio (1:3:3). The peptide/small molecule/protein was added to the mixture, at a predetermined ratio to the binding site that was located on the S2 and S3 strands (X * 120μM) directly before aliquoting the stock into each of the 48 wells to minimize any precipitation that may occur. The stock was mixed by pipetting up and down and briefly spun down. In cases where precipitation was observed the total DNA concentration was lowered with S1 being added at 20 uM and S2 and S3 at 80 20 uM. After determining buffers that would facilitate co-crystallization of the DNA and proteins, the binding was steadily increased until the crystals would no longer form. For the homeodomains, the buffer was added directly to the DNA stock prior to adding the proteins to prevent precipitation at binding ratios greater than 4x.

1. Simmons, C. R.; Zhang, F.; Birktoft, J. J.; Qi, X. D.; Han, D. R.; Liu, Y.; Sha, R. J.; Abdallah, H.; Hernandez, C.; Ohayon, Y.; Seeman, N. C.; Yan, H., Construction and Structure Determination of a Three-dimensional DNA Crystal (vol 138, pg 10047, 2016). *Journal of the American Chemical Society* **2016,** *138* (38), 12690-12690.

2. Simmons, C. R.; Zhang, F.; MacCulloch, T.; Fahmi, N.; Stephanopoulos, N.; Liu, Y.; Seeman, N. C.; Yan, H., Tuning the Cavity Size and Chirality of Self-Assembling 3D DNA Crystals. *Journal of the American Chemical Society* **2017,** *139* (32), 11254-11260.

3. Zhang, F.; Simmons, C. R.; Gates, J.; Liu, Y.; Yan, H., Self-Assembly of a 3D DNA Crystal Structure with Rationally Designed Six-Fold Symmetry. *Angewandte Chemie-International Edition* **2018,** *57* (38), 12504-12507.

4. Otwinowski, Z.; Minor, W., Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A* **1997,** *276*, 307-326.

5. Sheldrick, G. M., Macromolecular phasing with SHELXE. *Zeitschrift Fur Kristallographie* **2002,** *217* (12), 644-650.

6. Pape, T.; Schneider, T. R., HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *Journal of Applied Crystallography* **2004,** *37*, 843-844.

7. Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C., PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallographica Section D-Biological Crystallography* **2002,** *58*, 1948-1954.

8. Terwilliger, T. C., SOLVE and RESOLVE: Automated structure solution and density modification. *Macromolecular Crystallography, Pt D* **2003,** *374*, 22-37.

9. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J., Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D-Structural Biology* **1997,** *53*, 240-255.

10. Bailey, S., THE CCP4 SUITE - PROGRAMS FOR PROTEIN CRYSTALLOGRAPHY. *Acta Crystallographica Section D-Biological Crystallography* **1994,** *50*, 760-763.

11. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *Journal of Applied Crystallography* **2007,** *40*, 658-674.