2022

Investigation of DNA Hybridization in Localized Systems in Close Proximity

Ashley M. Sewsankar
University of Central Florida

INVESTIGATION OF DNA HYBRIDIZATION IN LOCALIZED SYSTEMS IN CLOSE PROXIMITY

by

ASHLEY SEWSANKAR

A thesis submitted in fulfillment of requirements
For the Honors in the Major Program in Chemistry
In the Department of Chemistry
in the College of Sciences
and in the Burnett Honors College
at the University of Central Florida
Orlando, Florida

Spring Term, 2022

Thesis Chair: Dmitry Kolpashchikov, Ph.D

Abstract

Hybridization of two or more DNA or RNA strands is well documented for the process taking place with all strands free in solution or when one strand is immobilized on a substrate. This study contributes to the investigation of the hybridization process when two single DNA strands (ssDNA) are in close proximity. We took advantage of an X sensor in which hybridization of four DNA strands enables the formation of a DNA four-way junction (crossover or X) structure. We immobilized multiple layers of crossover structures to study its hybridization being triggered by short ssDNA coming from solution and further investigate how many layers of these structures can hybridize by the addition of only one ssDNA (called input). Using a molecular beacon as reporter, we combined crossover DNA strands that recognize the reporter sequence at one side and at the other, the sequence of its input or downward crossover layer. Fluorescent signal was detected by separation of the molecular beacon's fluorophore and quencher, as it hybridizes with the system of layers. Immobilization of the X structures into the scaffold proved to increase their communication, in comparison to being free in solution. This evidence gives us significant information for the communication of hybridized layers in a localized system, showing a promising standard for development of multilayered logic gates. The potential of these crossover DNA strands using X structure include applications in the future of biological systems, nanotechnology, and target DNA recognition for its ability to quickly recognize a signal and propagate it through extended DNA nanostructure in a controlled manner.

ACKNOWLEDGMENTS

The author is grateful to Thesis Chair Dr. Dmitry Kolpashchikov, for his knowledge, research, critique and words of encouragement. Special thanks to Andrea Bardales, who has been an amazing mentor, inspiration and collaborator on this project. Mrs. Bardales has been a phenomenal supervisor and teacher who has guided me on the path to becoming a better researcher. She was always there to explain and provide illustrations at critical points, as well as providing her developments in designing our DNA nanostructures. Great appreciation to the Kolpashchikov lab group for their support. Tatiana Molden for her work and progress in applying logic to DNA nanodevices for treatment against cancer that gives us a reason to continue to develop this field. Thank you to the Burnett Honors Colleges for providing me with resources and opportunity. Thank you to the College of Sciences and Burnett Honors College for selection of the author to be awarded the 2022 Honors Undergraduate Thesis Scholarship. The author would like to give a special dedication to her late-great niece. Nicole Bindraban, who passed away in 2020 from cancer, when I first began my research journey, your happy spirit taught me a lot about life.

Table of Contents

List of Figures

List of Tables

Purpose

DNA nanotechnology shows great promise and opportunity to diagnose and treat the currently untreatable, such as cancer [1], genetic disorders [2], and various other diseases. The ubiquity of DNA is what gives it its prospective application as a nanodevice [3], diagnostic tool [4], and target detection tool for treatment [1]. In developing oligonucleotide-based probes (X sensors), that can be used in detecting target DNA after amplification, such as in one hour assay detecting the presence of the ZIKV-specific nucleic acid sequences [4]. Currently used Nucleic Acid Amplification Tests (NAATs) mostly rely on the reverse transcription-polymerase chain reaction (RT-PCR), which is time-consuming and requires sophisticated equipment and highly trained personnel, thus making it difficult to transition the diagnostics from the laboratory to point-of-care tests (POCT) [4]. Alternatively, an enzyme-DNA circuit can recognize NADH, which is produced as an output of an enzymatic system, as input, and release a DNA oligonucleotide as an output, which can be processed by a downstream DNA computing system as input [5]. The importance of detecting input DNA oligonucleotides (short strands of DNA) and thus producing a cascade through DNA computing system is to further perform actions resulting in, for example, detection of target nucleic acid-based biomarkers in cancer cells and perform programmed actions such as AND, NOT, and XOR. The simplicity of the sequencing of DNA led to the discovery of the use of DNA as a computing device that can be manipulated to indicate the presence of an analyte with a specific target sequence and hybridize with a single crossover DNA layer in a unit of multiple crossover layers. We are studying the hybridization of layers that are short oligonucleotides arranged in a four-way junction that mimic the construct in Figure 1. Our system is based on the X sensors (Fig. 1) that forms a fluorescent complex with a self-hybridized MB probe upon hybridization to an input oligonucleotide. Response to

hybridization through layers that are constructed in this way is recorded using a universal molecular beacon that can fluoresce due to the separation of its fluorophore (or reporter) and quencher. In the absence of complementary strands and input, the molecular beacon probe remains a self-hybridized hairpin structure. TaqMan qPCR, a linear probe, >15nt in length maximize the possibility that a single binding site is targeted within a genome made of billions of base pairs, but since these probes are stable under physiological conditions they are not reliable for SNV detection [6]. MB probes have better selectivity to detect SNVs by fluorescence, due to self-hybridizing, they require specificity in hybridization to separate fluorophore from quencher and specific nucleotide recognition to separate from itself. With addition of binding arms coming together in a four-way junction (see Figure 1.) and thus complexing with a MB probe, allowing for only a target analyte complementary to the opposite side of the four-way junction to be detected. Signaling is sensitive to the structure of the DNA strands and allows a strict way to understand the stability of layers of hybridized DNA in solution. Since separation of fluorophore and quencher can only be definitively achieved through these binding arms coming together. Due to its small size, massive parallelism, low power consumption and huge storage capacity [3], DNA is considered ideal and has the potential to be more advantageous than electronic computers. Being that there is biocompatibility to allow the analysis of a biological input that can be used under physiological conditions for diagnosis and treatment of diseased cells. The goal was to develop a fixed scaffold for the layers to transmit hybridization energy to the MB. This should promote stability and result in propagation of signal for up to five layers.

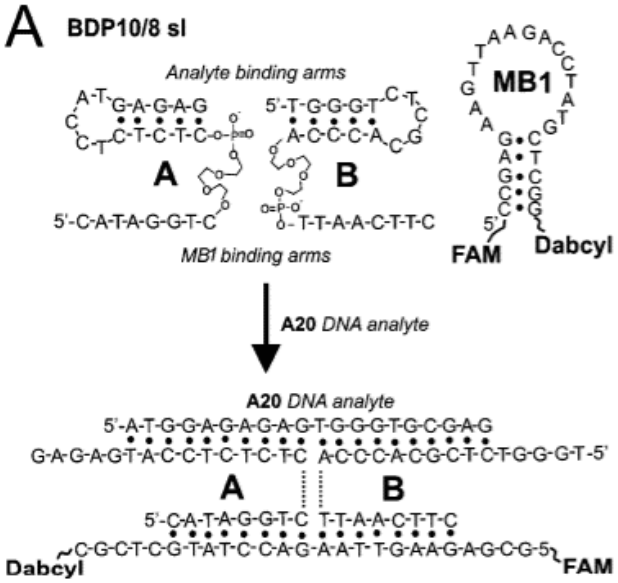

Figure 1. **Binary structure binding Input and Molecular Beacon.** Primary and secondary structure of BDPs used in study [7] A) Structure of BDP10/8sl in the absence (top) or presence (bottom) of A20 DNA analyte (input). Tri-ethylene glycol linkers are depicted by the dashed lines. Molecular beacon used has a Fam and Dabacyl (quencher). Source: Figure from (Kolpashchikov, 2006) [7].

In the integration of crossover DNA strands, challenges arose in communication between more than three layers of oligonucleotides. While two and three layers of communicating logic gates have been designed in multiple variations, universal, scalable crossover layers have not yet been achieved and therefore constitutes one of the major challenges in the development of DNA nano processors. The problems of slow communication between DNA gates as well as inter-gate crosstalk are the two major factors that impede DNA logic gate integration in circuit. The tile design, shown in Figure 3A., strongly suggests that localizing the gates on a DNA scaffold can enable the building of long chains of communicating gates. Parameters involved in studying the hybridization between DNA are dependent on concentration, association rate constant, dissociation rate constant, and melting temperature [8]. Through testing different parameters, the optimization of signaling can be achieved. Our reasoning for organizing DNA into 2-D layers is the benefits of higher hybridization rates (or the ability for DNA or RNA to bind to its complementary strand), a reduced level of undesired background DNA association, and the

possibility to transmit a signal over longer distances [9]. The hybridization of DNA in the formation of a four-way junction is a useful way to create close-knit DNA strands, among its other applications in studying its ability to be mapped topologically in 3D helices. When integrating crossover layers the signal will be able to travel through multiple crossover layers and the molecular beacon will then report the presence of an input with high output fluorescent intensity and an increased signal-to-background ratio when supported by a scaffold.

Background (Literature Review)

DNA consists of an anionic sugar-phosphate backbone and inside its cylindrical helix are neutrally charged nucleobases. The nucleobases themselves are aromatic nitrogen heterocycles comprised of monocyclic pyrimidines and bicyclic purines, which are polar along their edges, carrying multiple substituents. The hybridization of DNA double helix formation correlates to the thermodynamics of base stacking and base pairing. DNAs base stacking capabilities, involve purines and pyrimidines that are attracted to each other through van der Waals interactions [10]. Other driving forces for base stacking are the hydrophobic effects of purines and pyrimidines on one another. The backbone of DNA consists of phosphodiester bonds that are anionic and are addressed by the addition of magnesium within the buffer solution. Double-stranded DNA hybridizes due to noncovalent bonding from hydrogen bonding between purines that have two hydrogen bonds and pyrimidines that have three. Water molecules can interact with themselves as well as with DNA bases via dispersion, electrostatic, as well as hydrogen bonding interactions [11]. Individually, each of these forces is strong, but together, these factors counterbalance each other producing a composite effect that is experimentally found to be rather mild-stacking at 37 °C for a single pair of bases, typically stabilized by a free energy of no more than 1 kcal/mol [11]. The Watson & Crick pair with two or three such bonds, add approximately 0.5 ± 1.8 kcal/mol of stabilization per base pair of DNA [10] and with that thermodynamic studies at physiological temperatures also suggest that G|C pairs in DNA are only ~ -0.5 kcal/mol stable compared to A|T pairs, which are $\sim +0.1$ kcal/mol unstable, and the thermodynamic stabilities of DNA helices appear to be dominated instead by stacking free energies [12]. The experimental consensus is that the stacking free energy between two nucleobases is rather moderate, of the order of 1 kcal/mol depending on the identities of the stacked bases [11]. Following rules on the

thermodynamic nature of base pairing and base stacking capabilities, we can achieve an organization of crossover layers like that shown in Figure 3A.

Oligonucleotide hybridization requires steps that include nucleation, zippering, and relaxation. A nucleation step involves bringing two random coil ssDNAs together to form an intermediate state, in which a few base pairs have formed, and this intermediate proceeds further to form the fully hybridized dsDNA molecule [13]. Nucleation is an entropy-decreasing process ($\Delta S‡ < 0$), which makes a positive contribution to the free energy barrier through $\Delta G‡ = \Delta H‡ - T\Delta S‡$, where $\Delta H‡$ is the activation enthalpy [13]. Conversely, the melting of a hairpin structure is an entropy-increasing process ($\Delta S‡ > 0$), which contributes negatively to the free energy barrier describing the relaxation of a hairpin structure, such as the MB probe [13]. Therefore, one would expect to see a switch in the rate-limiting step from the relaxation of the hairpin at a low temperature to its nucleation at higher temperature [13]. According to the current theory, the renaturation process should have three distinct steps namely (a) formation of nonspecific contact (b) nucleation or correct contact formation and (c) zipping [14]. In the first step, the reacting c-ssDNAs collide with each other via three-dimensional (3D) diffusion-controlled routes, resulting in the formation of Watson-Crick (WC) base pairs at random nonspecific contacts between the reacting c-ssDNAs [14]. Such nonspecific WC contacts randomly translocate along c-ssDNAs either via thermally driven one-dimensional (1D) slithering dynamics or internal displacement mechanisms until finding the correct-contact and initiate the nucleation process which is then followed by spontaneous zippering of c-ssDNAs [14]. A random sequence tailored for a unique molecular recognition event can be prompted by flanking either side with a short concatemer (i.e., a sequence consisting of a repetitive motif) [15]. Such principles could be useful in addressing problems that require optimization in the selectivity and sensitivity of DNA

reassociation and/or hybridization, including the assembly of nanoscale structures [15]. Studying the mechanism on the hybridization of ssDNA oligonucleotides are important to understanding how the structure forms and stabilizes in solution. Understanding of Gibbs Free Energy, entropy, the theory of renaturation, and zippering is important to understand the thermodynamics for developing properly sequenced logic gates.

DNA functions as a storage for genetic information, being that there are three billion base pairs that encode around one hundred thousand different kinds of proteins [16]. The encoding capability of DNA, harness the potential to be applied in nanodevices for theragnostic purposes upon recognition of a target DNA or RNA. For example, Molden et. al demonstrates that DNA nanodevices have the capability to detect a target sequence being cleaved due to its high selectivity with an RNA cancer marker, resulting in triggering of a deoxyribozyme that can cleave a housekeeping gene mRNA [1]. This action is possible with function that is programmed with the same planning with which we program electronic computers. Another example in Mela et. al, DNA origami nanostructures functionalized with aptamers as a vehicle for delivering active antimicrobial components in a target-specific and efficient manner [17]. The chain of the connected logic modules can be easily scaled up in a modular fashion by feeding the output of an upstream gate to the next downstream gate in the chain. These actions that can be programmed into crossover DNA-based units, proving that there is versatility in nanodevices.

In this way, binary arms in our DNA construct can be replaced with catalytic function. Binary arms upon hybridization to a complementary analyte, the two strands could also be used to form a catalytically active Dz core, which cleaves a fluorophore- and quencher-labeled substrate (F-sub), thus producing a fluorescent output [6]. The analyte binding site is on one side and MB binding site is on another, separated by the 4WJ demonstrated by the adaptor strands.

The geometry of the binding arms includes a tri-ethylene glycol linkers and hexa-ethylene glycol linkers are the stem that conjoins the binding sites for an input and a molecular beacon, depicted in Figure 1. as dashed lines in the binding arms A and B, 5' to 3', splitting our sequence for that A or B segment of the layer in half [7]. The fragments of DNA hybridize using at least 15-20 nucleotides. Only the two adaptor oligonucleotides without chemical conjugation need to be tailored for each new analyte sequence [18]. In the presence of the target nucleic acid sequence, the four strands associate to form an X-shaped four-way DNA junction (4WJ), depicted in Figure 1. The MB probe fluorescently reports the formation of this quadripartite complex. Herein, the two most costly components of the electrochemical detection, electrode-attached sequence and the MeB-conjugated sequence, in a generic format. Therefore, a single electrode can be used for the analysis of different analytes [18].

Structural designs that were tested in our study include a rail design (Figure 3B.) and tile design (Figure 3A.) that differ in scaffold sequencing, the scaffold for the rail design lack a TTT linker region, present in the Tile design. The X sensor uses a four-way junction to stabilize thermodynamically within a localized scaffold and could then be applied to the differentiation of single nucleotide variations (SNVs). The adaptor strands are highly selective toward the analyte because of their sensitivity to single mismatched base pairing, this is reflected in reduced fluorescent signaling. Noting that nanostructures could process a target that has SNV's and as a result, has potential in detecting cancer markers that have mutations in sequencing. An example of a logic gate in study done by Molden et. al., DNA crossover tiles (X tiles), depicted in Figure 2., have negative AND (NAND) Boolean logic [2]. Four-way strand association of NAND FS, BS and 1a, 1b is required to produce fluorescent response and achieve NAND gate function [2]. When two inputs are present the NAND gate produces low output and opposite when inputs are

absent. These X tiles were further integrated into an entire processor that was able to analyze three inputs and produce output based on their absence/presence. The bridge fragments are only partially complementary to the arms of binary or split Dz (BiDz) sensor and are fully complementary to inputs 1 and 2, making input/bridge hybridization more favorable [2]. In the absence of inputs, the NAND1 population should be roughly evenly split between the associations, in which either Bridge a or Bridge b stabilizes the catalytic Dz core, shown in Figure 2. [2]. The RNA-cleaving function can be tailored for suppressing mRNA in gene therapy applications, while the MB probe can be used for monitoring the gate's outputs in a reusable format [2]. Bringing the interacting DNA sequences in close proximity helped reduce nonspecific associations and increase hybridization rates, accelerating response, though more time is required in the integration of another NAND gate resulting in a need for optimization [2].

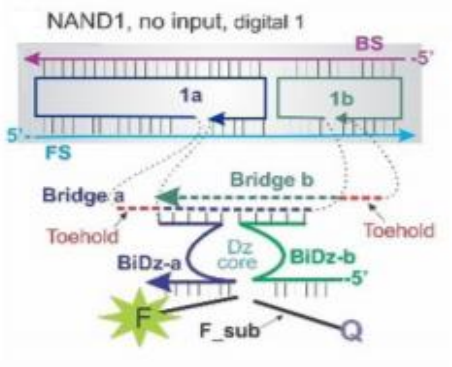

Figure 2. NAND1 Logic Gate. A tile-integrated, reusable NAND gate of an original design. The NAND1 gate is an association of four oligonucleotide strands: front strand (FS), 1a, 1b, and the back strand (BS). High output (digital 1) is produced when the fragments Bridge a or Bridge b (blue and green dashed lines) hybridize to BiDz-a and BiDz-b (blue and green solid lines). The bridge fragments including toeholds (red) are complementary to Inputs 1 and 2, respectively. High output (digital 1) is generated in the absence of inputs (Molden et al., 2020) [1].

Approach

1.1 Hybridization in structural designs

By studying a single unit made up of DNA strands in 4WJ assembled to a scaffold, we can better understand the interaction of DNA in close proximity. By reading fluorescence intensity where data is collected for X sensor with MB, Input absent, and Input present. Keeping a running record of results between different types of scaffolds (e.g. Rail or Tile) to improve our design. Using a Molecular Beacon probe (pictured in Figure 1.), we can observe fluorescence due to the proximity of the Fluorophore and Quencher that make up the molecular beacon to obtain information on the hybridization between the multilayered structure. One side of the DNA unit can recognize the DNA analyte while at the other side the molecular beacon notifies the presence of the input. Due to it's ability to easily test the hybridization of DNA strands, we are able to add and remove layers as needed to study the formation and capabilities of the construct and extrapolate the understanding for construction of DNA based logic computations. The ultimate goal is to create a DNA computer that can be used to recognize multiple targets that are for example cancer cells specific.

1.2 Description of DNA design of study

Layers of a DNA logic unit are sequenced to have a binding site (the A and B binding arms) that is complementary to the molecular beacon that reads fluorescence for the unit, and lowermost layer that complements the input that is sought to be detected, as seen in Figure 3. Including a scaffold surrounding the layers of strands, that results in stability and the complex of layers to hybridize and remain linear. The bases that complement the rails for each structure have binding of 10 base pairs this helps to prevent the helical nature of DNA that forms a turn every 10.5 base pairs as it hybridizes to for dsDNA.

Formation of the layers in a 4-way junction has optimum stability in separation of the fluorophore and quencher and results in high fluorescent signaling, in either propagation of signal through layers or single layer target detection. The crossover DNA strands are joined with polyethylene glycol linkers to form a single unit with binding arms A and B together. This will be further examined in a Rail Design (shown in Figure 3B.) to study localization, and results of hybridization rates among crossover DNA layers linked and non-linked. In a way this is like having TTT linkers that precede layer 1 in the AB-Tile 6L, simplified in Figure 3A. The structure is not only made up of DNA sequences for functionality, but also can accurately detect the presence of a target DNA. Molecular beacons are engineered to be compatible with each of the layer's, respectively. The formation of the structure shows a response in the molecular beacon and gives us an idea of whether an input has hybridized based on the high fluorescence produced in the separation of fluorophore and quencher. Whether we are testing with or without an input sequence present can be assessed by the fluorescent signal, without an input the fluorescence is low indicating the layer has not formed.

The proper conditions for DNA include a neutral pH, 22^{o}C, and can be further assessed using the melting temperature of the strand. To regulate the melting temperature the G-C content can be increased and the Gibbs free energy of the strand complex is assessed.. Effective ways in testing the strand configuration are by using an application such as NUPACK as a guide showing the proper binding of layer complexes and help organize structure to have almost no secondary structure formation among DNA crossover strands.

1.3 Designs utilized to study hybridization.

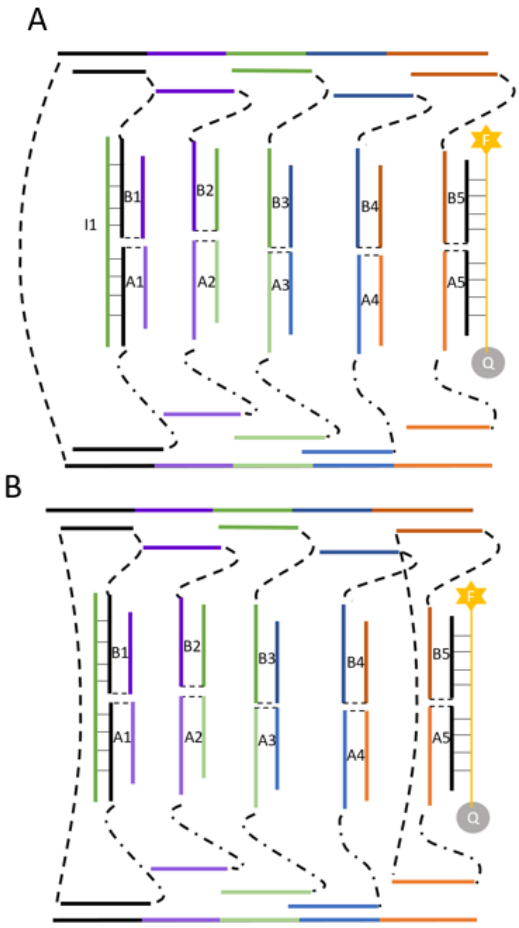

Figure 3. Tile and Rail design. A) Tile Design. The base pairing of the input (16-18 nt) is interchangeable among layers with unique sequences, across from it the MB probe (~21 nt), that has Fam (yellow) and Quencher (gray) components. In the tile design, TTT linkers precede the first layer shown by dotted line, using isp18 spacers sequenced to a TTT linker (~15 nt) that runs along to each spacers joining the two rails. B) Rail Design. Accurate representation of 4WJ hybridization of layers bound to Rail in which layers 5 and 1 are linked with PEG linkers that mimic TTT linker from the Tile design. Oligonucleotides of the same color depict where the strands hybridize to their complementary strand.

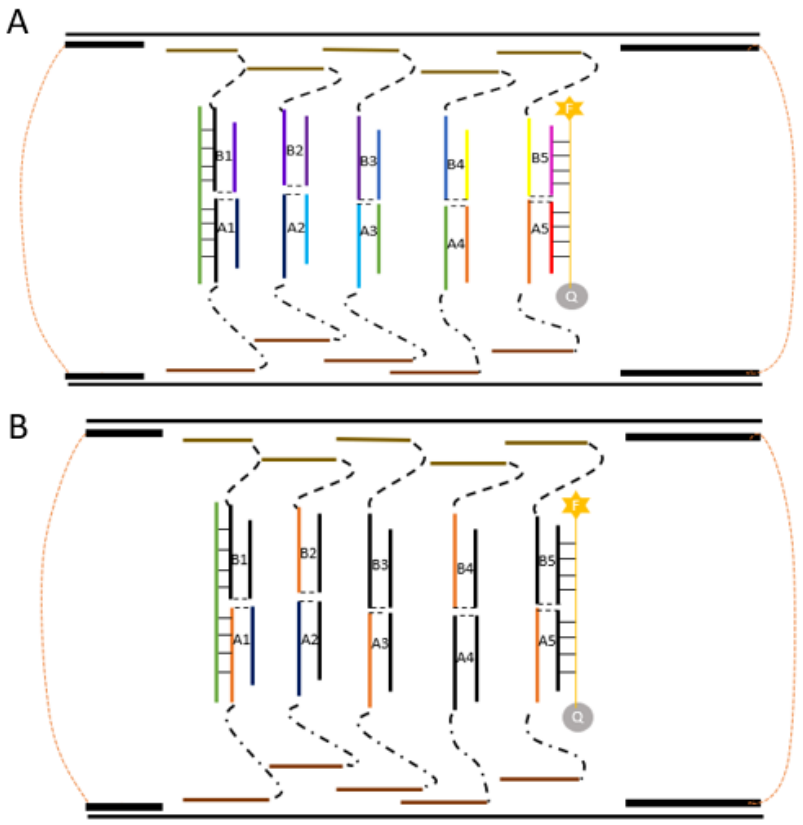

Figure 4. Rail design with staples and ALAS. A) Rail Design with staples. The rail was extended by 10nt on either end and complementary oligonucleotide staples were added, with TTT linkers (shown in orange) holding both ends of their staple. B) ALAS design. Layers were reduced by 1-3nt at the input/layer binding sites of alternating layers (shown in orange).

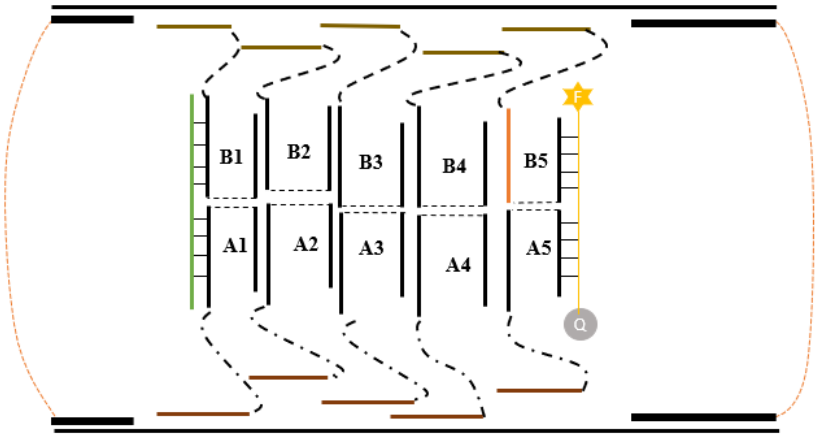

Figure 5. Optimized Rail design with staples. A) Optimized Rail Design with staples. To find the best performing layer combinations, reduced layers were tested to compare SBR. The results proved that reducing B5's layer/input binding site had the best SBR.

Materials and Methods

1. Materials and Instrument setting: DNAse/protease-free water was purchased from Fisher Scientific Inc. (Pittsburg, PA) and used for all buffers and for the stock solutions of oligonucleotides. $MgCl_2$ 1M solution was purchased from Teknova (Hollister,CA), Trizma HCl (pH 7.4) and Triton X100 were purchased from Sigma-Aldrich (St. Louis, MO), SeaKem LE agarose was purchased from Lonza (Rockland, ME). All oligonucleotides and molecular beacons were custom-made by Integrated DNA Technologies, Inc. (Coraville, IA) and concentrated solutions were prepared by resuspension with water at room temperature and stored at -20°C until use. The concentrations of oligonucleotides in stock solutions were determined from Lambert Beer equation, which absorptions readings at 260 nm were measured with a Thermo Scientific Nanodrop One UV-Vis Spectrophotometer, while the corresponding extinction coefficients were determined by using OligoAnalyzer 3.1 software (Integrated DNA Technologies, Inc.). Fluorescent experiments were performed using Perkin Elmer LS55 Fluorescent Spectrometer (Waltham, MA), excitation and emission were 485 nm and 517 nm, respectively, both slits at 15 nm and scan speed of 200 nm/min. Gel electrophoresis experiments were performed using BioRad electrophoresis equipment (Hercules, CA), ad visualized using BioRad Gel Doc XR+.

2. Individual characterization of Layers: A 1.5X master mix containing all DNA strands corresponding to scaffold and logic gates, and Buffer mix (Trizma-HCl buffer, $MgCl_2$ and Triton X100). The buffer used is prepared to hold a final concentration Tris-HCl of 50mM, 0.03% Triton and 50mM $MgCl_2$ on the sample for fluorescent experiments. Fluorescence is read at a wavelength of 485nm and emission 517 nm

3. DNA Tile assembling: A 2X master mix containing all DNA strands corresponding to scaffold and logic gates, and Buffer mix (Trizma-HCl buffer, $MgCl_2$ and Triton X100) were annealed at

95°C for 2 min and slowly cooled at 22°C. The correct size of DNA tiles formed was confirmed by 4% agarose gel electrophoresis (run at 70V for 2h). Gels were visualized using GelRed nucleic acid stain (Biotium). The size of DNA tiles was compared to the fragments of Invitrogen Ultra Low Range DNA ladder (Thermo Fisher Scientific).

4. Fluorescent experiments: Forty microliter aliquots of a master mix containing a molecular beacon and the DNA tile construct or only molecular beacon (negative control) in Buffer mix were dispensed to individual microcentrifuge tubes. Input stocks solutions were added to a final concentration of 200nM. For input absent samples, water was used instead of input stocks. All samples were incubated at 22°C for 20 min before reading of fluorescence emission at 517 nm (excitation 485 nm).

5. Signal-to-Background: Background (I-) describes the MB and input absent construct. Input absent includes the binding arms of the tested layers and it's hybridizing with the tile or rail, as well as the MB that depicts the formation of those structures in the form of fluorescence. Signal (I+) is defined by the input binding or target sequence that hybridizes with the construct at the layer being studied and propagates through layer(s) for recognition and fluorescence to be presented by the MB. The equation used for signal-to-background is $\frac{(I+)-MB}{(I-)-MB}$. Signal MB in the equation refers to only the molecular beacon.

Table 1. Oligonucleotide sequences used.

Names	Strand Sequences
(YES) Logic gates[a]	
XA1	5′ GAG TC TTC /iSp9/ tca cca cac/iSp9/ **C TAG CAT AGT**
XB1	5′ **CGA GGTA GCG** /iSp9/gac aga cgt/iSp9/ GGC TTG AG
XA2	5′ TCT AGC GC /iSp9/ gaa ga ctc/iSp9/ **GA CAT ACA G GA**
XB2	5′ **TCC GAT ACG AT** /iSp9/ct caa gcc /iSp9/ GAG AGT TC
XA3	5′ GTGT CCA T /iSp9/ gc gct aga /iSp9/ **ACA ACG GCA T**
XB3	5′-**G CAG TAC GTG** /iSp9/ ga act ctc /iSp9/ CGA AAT CC
XA4	5′ GAT CTA TTG /iSp9/ a tgg acac /iSp9/ **GG TGC TCA GGC**
XB4	5′- **TCC GTA GTT GC**/iSp9/ g gat ttcg /iSp9/ TAT GTT AAC
XA5	CAC TGA CA (27.6)/iSp9/ caa tag gtc/iSp9/**GA AGG ACT GAG**
XB5	**TCC GAA GTA GG** /iSp9/ gtt aac ata /iSp9/ CAA CAC AAG
5AX-vi	CAC TGA CA /iSp9/ CAA TAG ATC /iSp9/ **GA AGG ACT GAG**
5B1-vi	**TCC GAA GTA G** /iSp9/ GTT AAC ATA /iSp9/ CAA CAC
XB4-7	**TCC GTA GTT GC** /iSp9/ G GAT TTCG /iSp9/ TAT GTT AAC
XA3-6	*GT CCA T* /iSp9/ *GC GCT AGA* /iSp9/ **ACA ACG GCA T**
XB2-vi	**TCC GAT ACG AT** /iSp9/ CT CAA GCC /iSp9/ GAG AG
XA1-vi	GTC TTC /iSp9/ TCA CCA CAC/iSp9/ **CTA GCA TAG T**
X_L1_Rail-1	5′ GAG TC TTC /iSp18/ tca cca cac /iSp18/ CTA GCA TAG T /iSp18//iSp18/ CGA GGTA GCG /iSp18/ gac aga cgt /iSp18/ GGC TTG AG
X_L5_Rail-1	CAC TGA CA /iSp18/ caa tag atc **/iSp18/ GA AGG ACT GAG /iSp18**//iSp18/ **TCC GAA GTA G** /iSp18/ gtt aac ata /iSp18/ CAA CAC AAG
Scaffolds	
AB tile-6L	GTTC ATC TGTC C TAC TTC GGA GC AAC TAC GGA CAC GTA CTG C AT CGT ATC GGA CGC TACC TCG /iSp18/ ttt ttt ttt ttt ttt /iSp18/ ACT ATG CTA G TC C TGT ATG TC A TGC CGT TGT GCC TGA GCA CC CTC AGT CCT TC TGT GTC CCT C
1AB tile-6L_/iSp18/	GTTC ATC TGTC C TAC TTC GGA GC AAC TAC GGA CAC GTA CTG C AT CGT ATC GGA CGC TACC TCG /iSp18/ ACT ATG CTA G TC C TGT ATG TC A TGC CGT TGT GCC TGA GCA CC CTC AGT CCT TC TGT GTC CCT C
X_Rail-1_a	ACTATGCTAG TCCTGTATGTC ATGCCGTTGT GCCTGAGCACC CTCAGTCCTTC /iSp18/ TGTGTATT CTCTCTC
X_Rail-1_b	GTC AGA TCA AAT GTC /iSp18/ CTACTTCGGA GCAACTACGGA CACGTACTGC ATCGTATCGGA CGCTACCTCG
X_Rail-10_a	ACT ATG CTA G TCC TGT ATG TC A TGC CGT TGT GCC TGA GCA CC CTC AGT CCT TC TGT GTC CCT C
X_Rail-10_b	GT TCA TCT GTC C TAC TTC GGA GC AAC TAC GGA CAC GTA CTG C AT CGT ATC GGA CGC TACC TCG
Inputs	
X_I1	GTG TGG TGA ACG TCT GTC
X_I2	GAG TC TTC GGC TTG AG
X_I3	TCT AGC GC GAG AGT TC
X_I4	GTGT CCA T C GAA ATC C
X_I5	GAT CTA TTG TAT GTT AAC
Molecular Beacons[b]	
MB NOT	FAM-**CTCAAG CC GAA GA CTC** CTTGAG-DABCYL
UMB_XA20	FAM- CCGG A T**GA ACT CTC GCG CTA GA**A ACC GG-BHQ
MB5	FAM-CAT **GG GAT TTC GA TGG ACA C**CATG-DABCYL
UMB5	FAM-5′-CGC**G TTA ACA TA CAA TAG AT C**GCG-BQ1
MB Inha	FAM-CGCT **CTT GTG TTG TG TCA GTG** AGCG-BHQ

[a] Formatting: For each strand composing the crossover logic gate, the portion complementary to the molecular beacon is in uppercase and the portion complementary to input is in lowercase; sequences complementary to scaffold strand is in bold

uppercase. The term /iSp9/ or /iSp18/ refers to a tri-ethylene glycol or hexa-ethylene glycol intra-strand moiety inserted to covalently bound the phosphate DNA backbone.

[b] Formatting: A fluorescein derivate, represented as FAM, is attached to 5' end of each molecular beacon used; a molecular quencher, Dabcyl or BHQ(Black Hole quencher) is attached to the 3' end of each molecular beacon. The portion of molecular beacon that self-complement for hairpin structure formation is in underlined uppercase and the portion complementary to crossover layer is in bold uppercase.

Results and Discussion

The analysis of DNA hybridization was explored through experimentation by testing the propagation of signal among crossover "A" and "B" that are free in solution. The first experiment performed was to understand how each layer behaves by itself. Each layer has a complementary input and MB, that will hybridize with their input binding site and MB binding sites with the formation of the 4WJ. In the second experiment, layers were added free in solution, we envision that these layers are free floating in solution. The only way for them to form in 4WJ is when the complement of their input/layer binding site is present. To prevent cross communication and study the hybridization of layer complexes (e.g. 5^{th}, 5-4^{th}, 5-4-3^{rd}…), input was added for the lowermost numbered layer. For example, 5-4-3^{rd} would form in the presence of input 3. Requiring layer 3, "A" and "B" fragments to first be brought together in 4WJ and then subsequent layers to follow. Thus layer 5 is formed in complex with layer 4 and 3 to open and hybridize the self-hybridized MB. Using the molecular beacon for layer 5, we study the propagation of signal through up to five layers within modified scaffolds of longer DNA strands. Our experimentation focuses on the optimization of signal and reduction of background, relying on the formation and hybridization of 4WJ structures. In solution we observe the detection of target DNA (input) for up to 5 layers present in solution and the communication of the 4WJ layers or hybridization energy produced through propagation to the MB.

Layers were then placed into a scaffold in subsequent experiments, the scaffold is made up of two DNA strands of up to 50-60 nucleotides. To these rails, first a TTT linker was added to both ends shown in Figure 3A., from results of this linkage. We further linked both ends, to create a localized area in which hybridization can occur. Linking both ends was first done by extending the rails 10 base pairs and using TTT linkers attached to the ends of short (10nt) fragments of

DNA referred to as staples. We further optimize the layers in final enclosed stapled scaffold to achieve increase in signal-to-background, which shows the communication of the 4WJ layers.

1.1 Characterization of Individual Layers

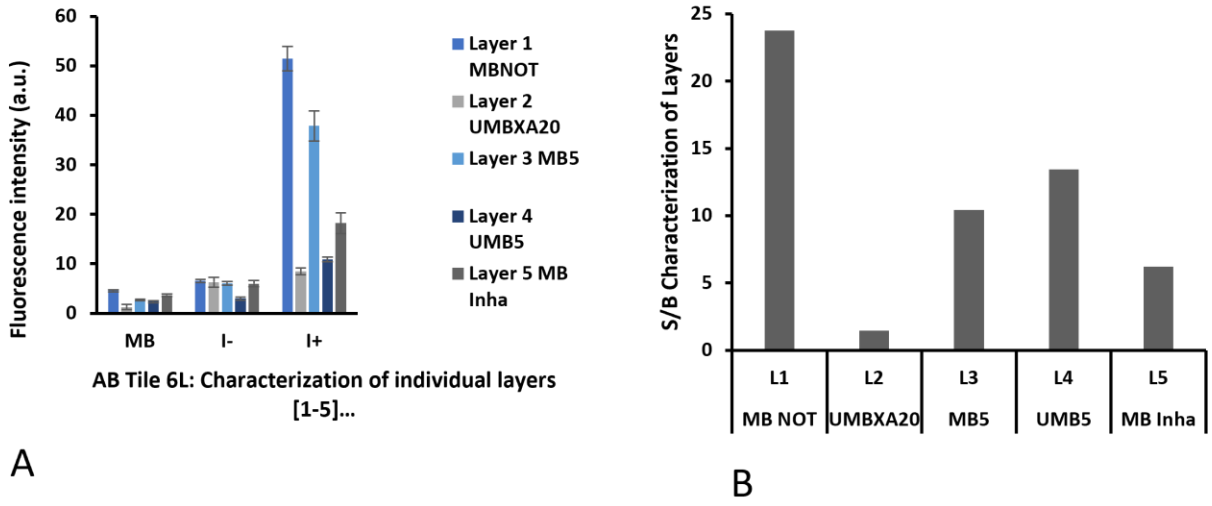

Figure 5. Characterization of individual layers results. A) Preliminary data for Characterization of individual layers for Layers 1-5, using their respective molecular beacons B) Signal-to-background for the characterization of individual layers derived from preliminary data.

To achieve high fluorescence intensity, the crossover strands require the recognition of input signal. It's important to clarify that MB recognition region is different for each layer, which accounts for the differences in the signal produced by their respective output reporter (Figure 5A). Observe in Figure 5a. the background produced when layer and MB are present. The layer is separate in A and B without the presence of input. The individual layers each show signal-to-background that recognizes its input binding site complement or target sequence. Ordering the layers from least to greatest signal-to-background ratio; L1, L4, L3, L5, and L2. It's important to note that hierarchy of layers only reports signal through the layers that come before it. For example, L3 molecular beacon reports inputs binding to L3, L2, and L1, when all layers leading up to the molecular beacon are bound. Rearrangement of the structure to include the layer with the greatest signal-to-background for its target may be a potential alteration to the design, for example using the MB for layer 1 would require altering the input binding site and input

sequences for subsequent layers. Since the input shows efficient hybridization through one layer, what happens when more crossover strands are added between input and molecular beacon. Is the 4WJ complex still able to form?

1.2 Tile-Free AB-6L (Layers 5th-1st) vs. Tile-present AB-6L (Layers 5th-1st).

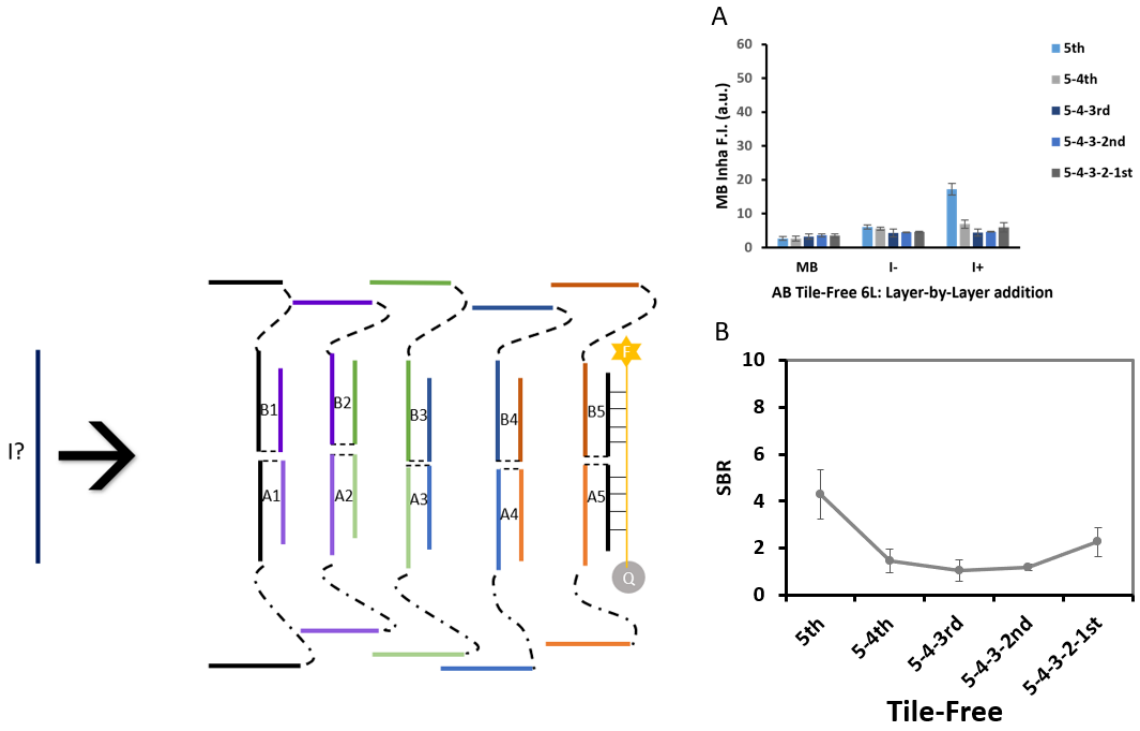

Figure 6. AB Tile-free Layer-by-Layer addition results. The structure shown to the left is of A and B strands that have complementarity to an input from the lower most layer, away from the MB binding site, any input is added to test the formation of a 4WJ layer. A) AB Tile-Free 6L, Layer -by-Layer addition from 5th-1st, showing fluorescence of MB, input absent, and input present for each layer combination. B) Tile-Free signal-to-background ratio, depicts the performance of each layer combination using the equation $\frac{(I+)-MB}{(I-)-MB}$.

When the layers are free in solution, their formation is dependent on the input to carry hybridization energy, detectable by the MB. In Figure 6 the diagram shows the layers that are dependent on the input to have complex formation. Based on the SBR, with a threshold of 3 times the background, the only layer that resulted in the formation of 4WJ is Layer 5. This is expected due to the known SBR for layer 5's characterization. An interesting result is the

complex formation of all layers present in solution, the SBR is almost 3. Hybridization energy produced by the formation of the complex is promising and unexpected. Since the 5th layer 4WJ forms with the presence of input 1 all the layers must first form for production of detectable signal. This could mean that the complex is stable on its own.

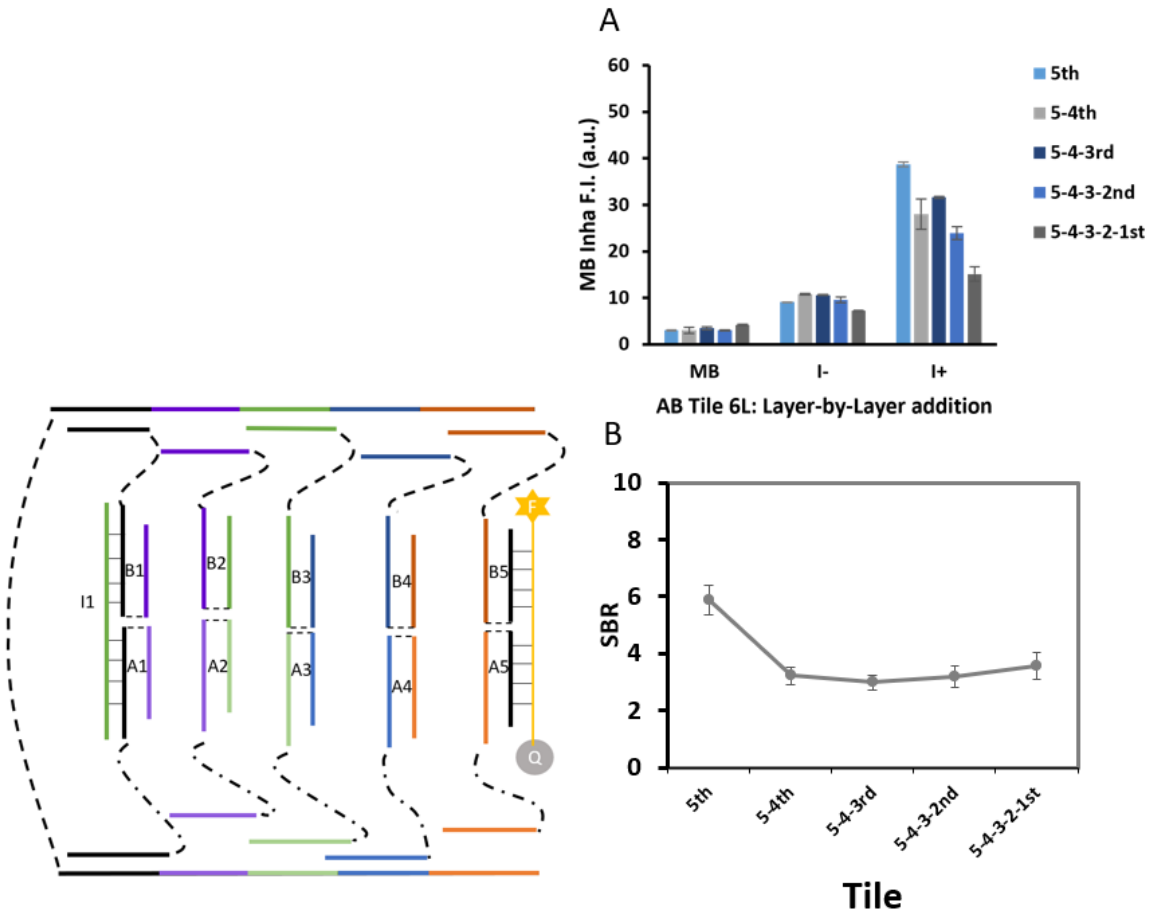

Figure 7. Tile-present preliminary data and SBR of Layer-by-Layer addition. A) AB Tile 6L: Layer by Layer addition, depicts the Input absent (I-), Input present (I+), and MB. B) Signal-to-Background of AB Tile 6L.

Having a linkage of the rails before the layer 1 binding site on the rail, also known as Tile, increases the stability of the structure by creating a localized space for layers to hybridize. When the structure is stable and hybridizes with a single input the highest fluorescence is achieved due

to propagation of hybridization energy among the layers. Due to the localization of each crossover layer, there is an increase of input-independent interactions among them, leading to the increase in background fluorescence.

When the tile is present an increase in fluorescence is achieved that is above the threshold of 3, or 3 times the background for all layer additions (Figure 7B). When no tile is present the SBR remains low having an unpredictable response, since the layers are free as A and B fragments (Figure 6B). The results for Tile-free are all below the threshold besides the fifth layer that is stable on its own. With tile the 4WJ layers can communicate signal above the threshold and have higher fluorescent values for each layer combination.

The tile scaffold in the design proved to be effective in increasing the SBR (seen in comparison of Figure 6B and 7B.). The SBR when the tile is present increased by 1-fold for the 5^{th} layer, expected from stabilization within rails, 2-fold increase for propagation through two layers, greater than 2.5-fold for propagation through three and four layers, and 1.5-fold for propagation through all five layers. These results give us enough evidence to justify the use of a scaffold.

1.3 Rail System (Layers 5th-1st)

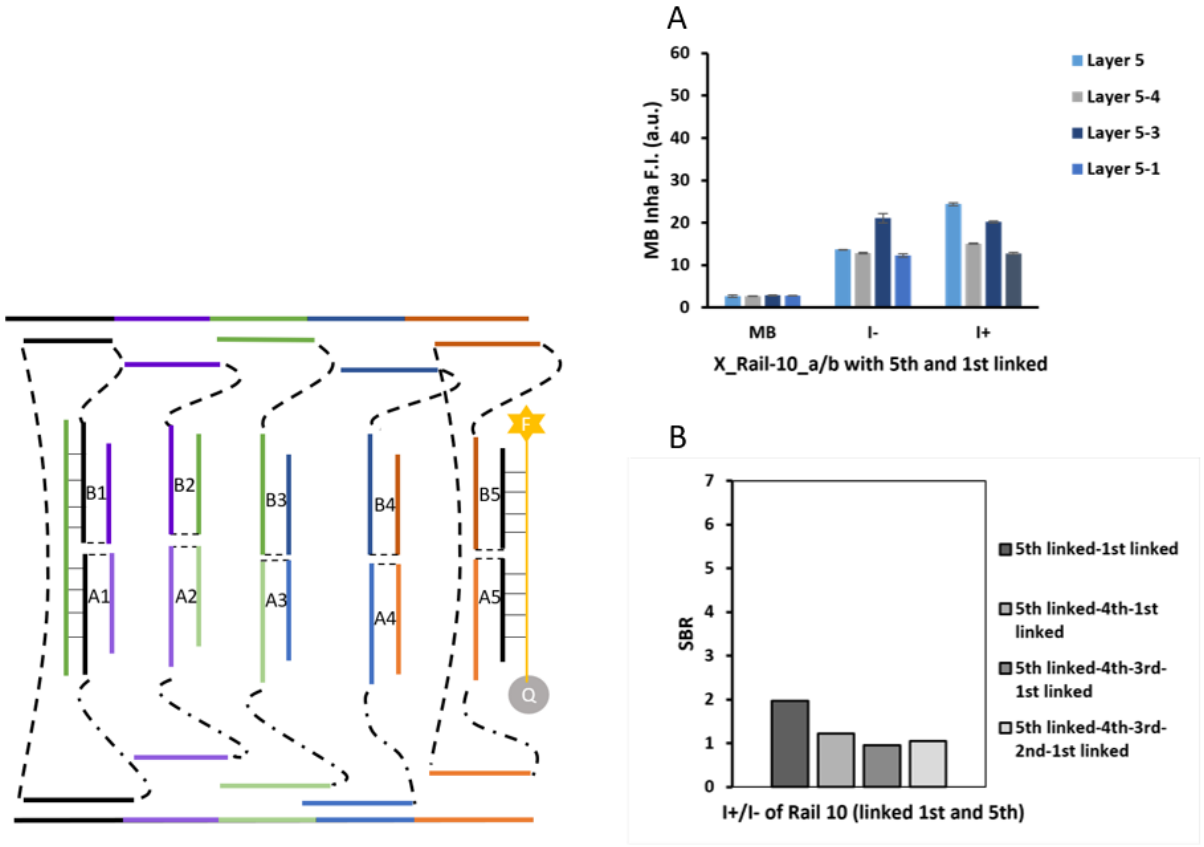

Figure 8. Rail design 5th and 1st linked Layer-by-Layer addition and SBR. A) Rail design with 5th and 1st linked by PEG linkers B) Signal-to-Background data for Rail design with PEG linked 5th and 1st. Above the graph 4WJ layers of 5th and 1st linked are depicted within rails taken from Figure 3b.

The Rail design had reduced signaling in the preliminary data (Figure 8A.) and in signal-to-background (Figure 8B.), as compared to the Tile design. Our goal was to have an attachment at both ends of the structure to create a fixed area to which the layers can bind within. Further to reduce the time for the strands to find their complement in solution, by having them localized within the scaffold. Reduced signaling is attributed to the cross-hybridization and bulky structure with the inclusion of layer one. Another explanation is having the 5th linked layer and 1st linked layer present there is a stretching of the rail at both ends of its structure. There shouldn't be interaction of the 1st layer until the 2nd layer is bound. The inputs corresponding to the layer used to perform this experiment are I5, I4, I3, and I1. This structural formation of the whole complex

when input 1 is introduced creates a square chip that the other three layers will hybridize within. Filling in those gaps gave us a SBR that is explained in the Layer-by-Layer addition, as more strands are added the background increases due to hybridization events that have resulted in layer 5's A and B fragments coming together from binding the scaffold and being linked. An interesting note is that with the addition of layer's 4 and 3, the linked 5^{th} layer fluoresces without presence of input, this data doesn't meet the goal of having a scaffold. In that the background was too high with presence of layer 1 as a part of the scaffold and disrupts the overall crossover formation.

1.4 Rails with staples.

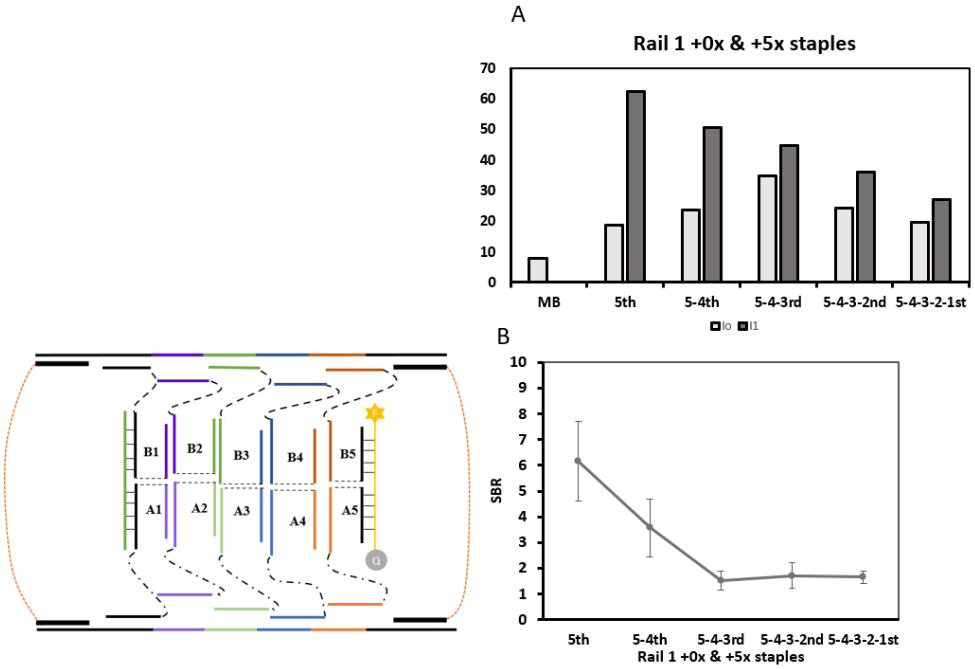

Figure 9. Rails with staples. A) Layer-by-Layer addition was performed for the structure in Figure 3a. The staples attach to the rail that was extended (~10nt) to allow for a fixed, rigid structure to which 4WJ layers can hybridize.

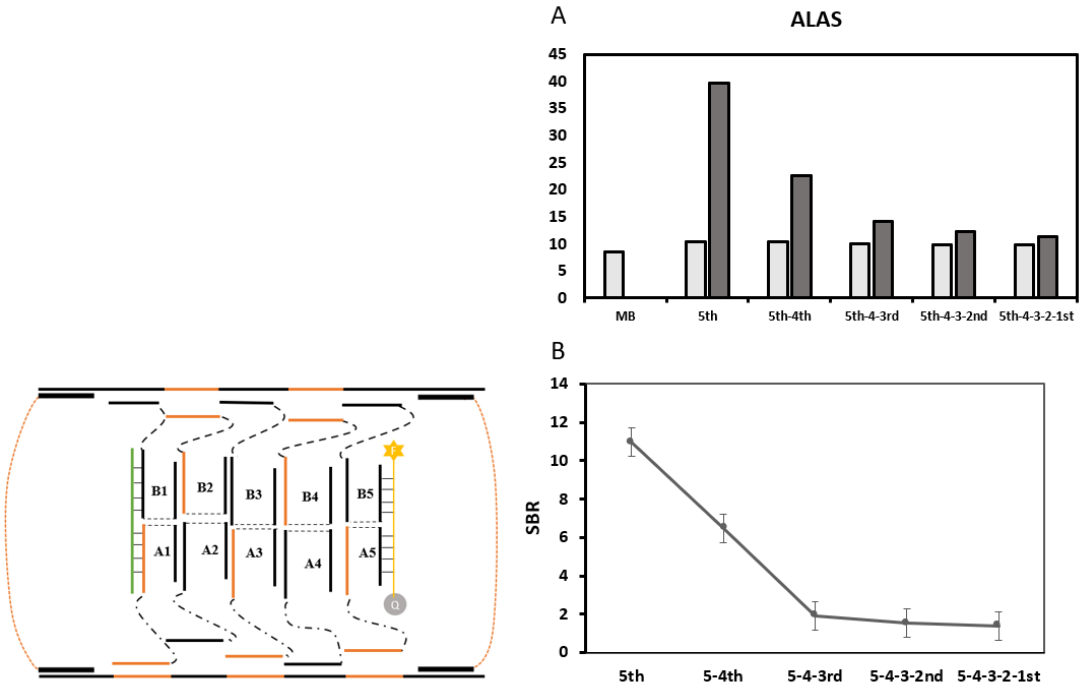

Figure 10. Alternating Long and Short oligonucleotides at the Input binding site. A) ALAS design layer-by-layer addition. B) Signal-to-Background ratio of ALAS design. Structure above graphs taken from Figure 4b.

The rail scaffold design was modified by the addition of 10 nucleotides at either end of the scaffold, plus the addition of staples linked with PEG linkers. Instead of the layers seen in the previous design from Figure 8. This data can be compared to that of the Tile design in Figure 6B., that had a signal of 40 to 10 background for the 5^{th} layer, the addition of strands in the stapled design has signal of 60 to 15 background (Figure 9A.). We can infer that the increase in signal compared to background is due to the stability of this structure, since we linked the end of the rails binding to the fifth layer it is easily able to form in solution. Refer to tile design (Figure 3A. or Figure 7), in which TTT linked end is before the 1st layer.

With the concept of a chip developed, we further justified the reduction of nucleotides in an alternating fashion. By reducing nucleotides, the data shown in Figure 9a. has a consistent background for all layer-by-layer additions. In other words, reduction of nucleotides each alternating layer/input binding site, resulted in a leveling of the background fluorescence for layer-by-layer addition. The results show that the signal-to-background reaches the threshold of 3 for the first two layers. Though the signal decreases by two-fold, the signal to background is doubled, here we achieved having a reduction in the background that is predictable. That can be used when developing 5-layered gates to achieve a level background for each combination of layers in a complex.

1.5 Optimizing ALAS

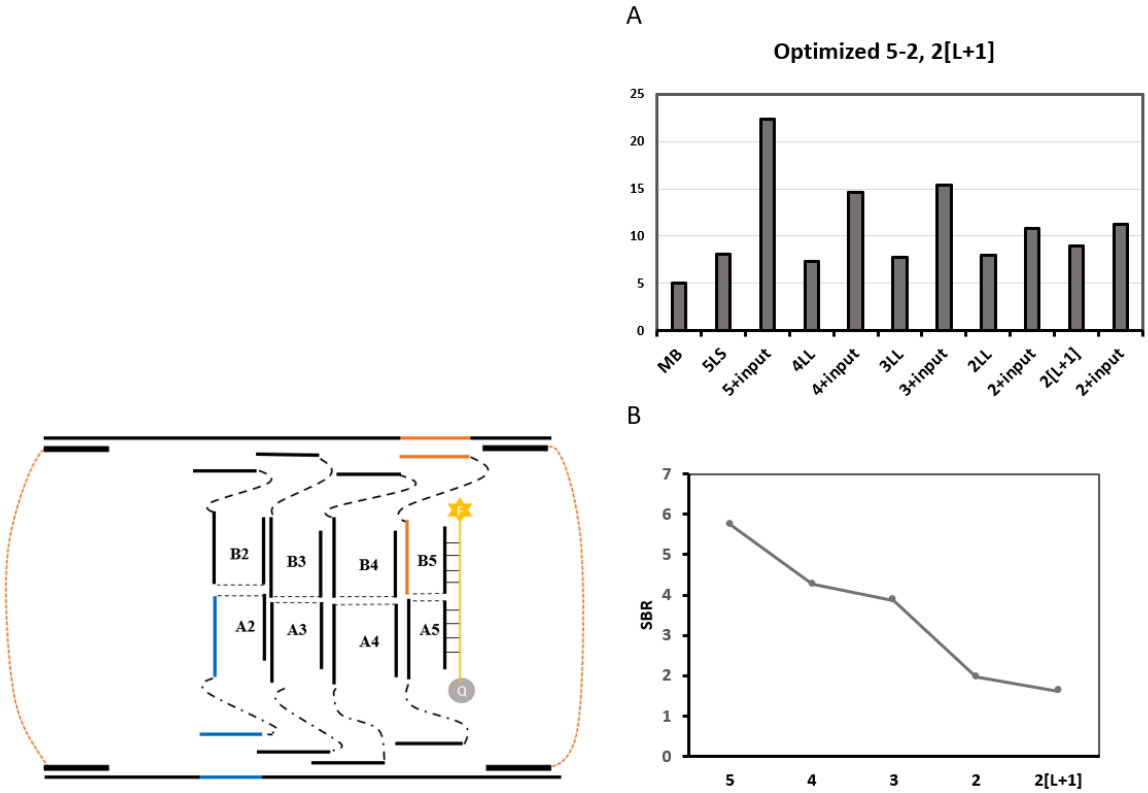

Figure 11. Optimized layers from ALAS based on improving SBR. A) Fluorescence for 5LS to 2 and 2 [L+1]. B) Signal-to-background ratio for ALAS optimized design. Figure above graphs refers to Figure 5. with alteration of Layer 2 for further analysis of a single nucleotide addition to layer 2.

To obtain these results, each layer was tested with corresponding shortened layer, for example the 5th layer was compared by LS and LL oligonucleotide sequences and the one with the better SBR performance was utilized and tested with the 4th layer LS and LL, and so on. In this experiment, we can conclude that this is the best performing layer combination, so far, for having the lowest consistent background response in the preliminary data. The results here could be due to shifting the hybridization to the side of the quencher, since there are less nucleotides binding the fluorophore end of the MB. From this data we see depreciation of signal and expect there to be less SBR for the addition of Layer 1 than that of Layer 2. Here an extra step of adding a nucleotide to the quencher side of the MB resulted in reduction of SBR (Figure 11B).

Conclusion

The results of this study prove that by using a four-way junction to bind target and X sensors, the proper structure forms and hybridizes with its target, that is then detected by a reporter showing high fluorescent intensity. The Tile and Stapled Rail design are scaffolds arranged to localize the structure of DNA oligonucleotides in four-way junction. Using layers arranged in the four-way junction we achieved hybridization of layers and propagation of signal through those layers that improved with a scaffold. For each design we found the same behavior, that the more layers added, the less fluorescent intensity detected. Starting by characterization of layers to give us an idea of how each layer behaves and show that each layer's molecular beacon can report variable signal based on the presence and absence of input. The Tile and Tile-free study proves that the scaffold can help to get better hybridization of layer 5 and subsequent layers due to it localizing the 4WJ layers that allows better signal propagation and structure formation. The Rail design lacks the TTT linker, and this has an effect on the stability of layers, giving a reduced signal. But proves that there are more ways to optimize logic gates in using a linked layer configuration at both ends of the design.

The experiments performed show proof that the absence and presence of input results in formation of 4WJ-hybridizing nanostructures. We chose to proceed with the stapled rail design for its rigid structure to keep 4WJ layers localized. Finally reducing the number of nucleotides that bind the input/layer in the ALAS design led to the discovery that a uniform background can be achieved and applied in designing 5 layered nanostructures. More research should be done with base pair reduction to confirm the reliability of separating the MB from the fluorophore side, seen in optimizing ALAS.

Future Perspective

The current understanding of DNA nanodevices using crossover DNA is inadequate in integration of multiple layers, due to the resulting depreciation in signal with addition of 4WJ layers. The scaffold used in the AB tile design is effective in curving the rate of depreciation in fluorescent signaling and therefore is an effective optimization to the design. By linking both ends of a scaffold, 4WJ layers were localized in a rigid structure for communicating hybridization. Our goal is to have these nanostructures reproduced and linked, to have 5 layer-repeating units that communicate signal and hybridize efficiently. To these hybridizing crossover layers, forms of logic like that shown in the NAND gate [1] and other forms of Boolean logic can be applied. Following in the footsteps of Boolean logic on a computer processing unit. As the stability enhances these oligonucleotides have endless possibilities in the detection of target genes for various diseases and performing an action on those targeted strands.

References

[1] T. A. N. C. T. a. K. D. M. Molden, "Cut and Paste for Cancer Treatment: A DNA Nanodevice that Cuts Out an RNA Marker Sequence to Activate a Therapeutic Function," *Angew. Chem. Int. Ed.,* pp. 21190-21194, 2020.

[2] T. G. M. &. K. D. Molden, "Manufacturing Reusable NAND Logic Gates and Their Initial Circuits for DNA Nanoprocessors," *Chem. Eur. J.,* vol. 27, pp. 2421-2426, 2021.

[3] E. &. J. A. Jamalzadeh, "High Accuracy Multi-input DNA Logic Gate Using The Spatially Localized DNA Structures," *IEEE Xplore,* 2020.

[4] C. e. a. Lynch, "Selective Determination of Isothermally Amplified Zika Virus RNA USing a Universal DNA-Hairpin Probe in Less than 1 Hour," *Analytical Chemistry,* vol. 91, no. 21, pp. 13458-13464, 2019.

[5] S. G. Y. G. N. K. D. a. K. E. Mailloux, "Bridging the Two Worlds: A Universal Interface between Enzymatic and DNA Computing Systems," *Angew. Chem.,* pp. 127, 6662-6666, 2015.

[6] D. Kolpashchikov, "Evolution of Hybridization Probes to DNA Machines and Robots," *Acc. Chem. Res.,* pp. 1949-1956, 2019.

[7] D. M. Kolpashchikov, "A Binary DNA Probe for Highly Specific Nucleic Acid Recognition," *J. AM. CHEM. SOC. ,* pp. 10625-10628, 2006.

[8] A. W. H. W. M. S. R. G. E. T. J. &. H. S. Munir, "Modeling Hybridization Kinetics of Gene Probes in a DNA Biochip Using FEMLAB," *Microarrays,* vol. 6, no. 9, pp. 1-12, 2017.

[9] A. S. S. &. K. D. Lake, "Molecular Logic Gates Connected through DNA Four-Way Junctions," *Angew. Chem. Int. Ed. ,* vol. 49, pp. 4459-4462, 2010.

[10] E. M. J. &. G. K. Kool, "Mimicking the Structure and Function of DNA: Insight into DNA Stability and Replication," *Angew, Chem, Int. Ed.,* vol. 39, pp. 990-1009, 2000.

[11] C. H. Mak, "Unraveling Base Stacking Driving Forces in DNA," *J. Phys. Chem. B,* vol. 120, pp. 6010-6020, 2016.

[12] R. &. M. C. Li, "A Deep Dive into DNA Base Pairing Interactions Under Water," *J. Phys. Chem. B. ,* vol. 124, pp. 5559-5570, 2020.

[13] Y. &. Z. X. Yin, "Kinetics and Dynamics of DNA Hybridization," *Acc. Chem. Res.,* vol. 44, no. 11, p. 1172–1181, 2011.

[14] G. &. M. R. Niranjani, "Theory on the Mechanism of DNA Renaturation: Stochastic Nucleation and Zipping," *PLOS ONE,* pp. 1-28, 2016.

[15] E. S. D. &. d. P. J. Sambriski, "Uncovering pathways in DNA oligonucleotide hybridization via transition state analysis," *PNAS,* vol. 106, no. 43, pp. 18125-18130, 2009.

[16] National Research Council, "Committee on Mapping and Sequencing the Human Genome. Mapping and Sequencing the Human Genome.," *National Academy of Sciences,* vol. 2, pp. 1-128, 1988.

[17] I. V.-R. P. M. S. C. G. B. D. H. R. S. H. E. M. &. K. C. Mela, "DNA Nanostructures for Targeted Antimicrobial Delivery," *Angew. Chem. Int. Ed. ,* vol. 59, pp. 12698-12702, 2020.

[18] S.-C. D. H.-Y. C. M.-C. K. D. M. Sun, "Multi-labeled electrochemical sensor for cost-effiecient detection of single nucleotide substitutions in folded nucleic acids," *Sensors & Actuators: B. Chemical 287*, pp. 569-275, 2019.

[19] K. V. M. N. A. L. S. M. A. F. E. H. E. J. L. B. R. &. S. B. Afonin, "Activation of different split functionalities on re-association of RNA-DNA hybrids," *NATURE NANOTECHNOLOGY*, vol. 8, pp. 296-304, 2013.