Abstract

Experimental demonstration of the Turing universal tile based algorithmic DNA self-assembly has been limited by significant assembly errors. An important class of errors, called co-ordinated growth errors, occur when an incorrect tile binds to a growing assembly even when some of its pads are mismatched with its neighbors. Activatable DNA tiles, introduced originally by [12], employ a protection/deprotection strategy to strictly enforce the direction of tiling assembly growth that prevents these errors, ensuring the robustness of the assembly process. Tiles are initially inactive, meaning that each tile’s pads are protected and cannot bind with other tiles. After an activation event, the tile transitions to an active state and its pads are exposed, allowing further growth. In this paper we demonstrate the formation of three different DNA nano-assemblies using activatable tiles. In the first assembly, we use two tiles A and B that have complementary sticky ends such that they can form a linear co-polymeric chain ...ABABABAB... Additionally, the sticky ends of A can be rendered inactive (non-sticky) by attachment of a protection strand to tile A. This protection strand can be subsequently removed from tile A (thus activating it) by the polymerization extension of a primer strand bound to the protection strand. Upon the introduction of the polymerase, tile A gets deprotected enabling the formation of the linear co-polymeric chain. The second type of assembly, consists of three tiles A,B,C, each of which is a DNA hairpin. On activation of each tile, the assembly is of the form of a ribbon ...ABCABCABC... The stem loop of each tile acts as a protection strand, and constitutes an inactive state of the tile. On activation, the stem loop is open, and can activate another tile. Our third system, is a system which shows catalytic assembly of the hairpin shaped tiles of system 2. These initial results point to the possibility of more complex activatable systems that can alleviate some of the errors encountered in tile based DNA self-assembly.

1 Introduction

DNA nanotechnology is a rapidly emerging discipline that uses the molecular recognition properties of DNA to create designed, artificial structures out of DNA. Precise nanoscale objects can be programmatical created using DNA strands and this has resulted in a myriad of nanostructures (see (author?) [27, 11, 28, 20, 9, 15, 10, 5, 3, 33] for some illustrative examples). But more importantly, dynamic behavior of these DNA nanostructures can also be controlled via the action of other DNA strands (see [31, 19, 21, 23, 29, 7, 24]) and enzymes that act on the DNA strands (see [30, 11, 18]).

An important class of DNA nanostructures are tile-based assemblies which are formed via addition of DNA tiles to a growing lattice. Each DNA tile is designed to have multiple single-stranded components called sticky ends that can bind to sticky ends of other tiles in the lattice. Tile based lattices can be used as an addressable nanosubstrate and 3D lattices have the potential to bind proteins at lattice points aiding in their crystallization and subsequent structure determination via X-ray diffraction. [27] demonstrated the first lattices out of DNA tiles. Since then many lattices, including 3D lattices, have been demonstrated (see [28, 9, 33]). By carefully programming the sticky ends of various

One-Time, Directed and Catalytic Activation of 1-D DNA Tiles

Sudhanshu Garg, Harish Chandran, Nikhil Gopalkrishnan and John Reif
Department of Computer Science, Duke University, Durham, USA
Email: sgarg@cs.duke.edu

October 31, 2012
tiles, complex lattices can be created (see [13, 16]). In fact, [17] showed in theory that any computable function can be implemented via a set of tiles with carefully designed sticky ends. Unfortunately experimental demonstration of complex tilings have been limited by significant assembly errors. A major source of errors are *co-ordinated growth errors* which occur when a tile partially satisfies the programming at a lattice point, i.e., it’s sticky ends match with only some of its neighbors but that is sufficient for it to be wrongly incorporated into the assembly.

[12] described a novel protection/deprotection strategy to strictly enforce the direction of tiling assembly growth to ensure the robustness of the assembly process. In their system tiles are initially inactive, meaning that each tile’s output sticky ends are protected and cannot bind with other tiles. Only after other tiles bind to the tile’s input sticky ends, the tile transitions to an active state and its output sticky ends are exposed, allowing further growth, and the formation of an assembly. In this paper we demonstrate the formation of three different DNA nano-assemblies using activatable tiles. Our first assembly is formed from tiles that are activated simultaneously, second assembly is formed from tiles that activate sequentially, while the third assembly is formed from tiles that activate both sequentially and catalytically. The success achieved in engineering such assemblies point to the possibility of more complex activatable systems that can possibly alleviate some of the errors encountered in tile based DNA self-assembly.

### 2 The Problem of Activatable Self-Assembly

Self-assembled DNA tiling systems are error prone. Many theoretical solutions have been suggested to remedy some of the errors that arise in DNA tilings (see [author?] [26, 2, 14]). Figure 1 illustrates the mechanism of a class of errors known as co-ordinated growth errors that occur quite frequently in DNA tilings. Consider self-assembly of the set of 8 tiles in Figure 1(a) at temperature \( \tau = 2 \). The tiles themselves are illustrated as gray rectangles with distinguishing face labels and pad types that are shown near the appropriate edge of each tile. Null pads are omitted and the binding strength of each pad is indicated by the number of rectangular bars that accompanies it. Let us assume that the seed tile is \( T_1 \). The partial assembly illustrated in Figure 1(b) is self-assembled purely via strength 2 bonds without any errors. At this juncture, the assembly can proceed to completion without any errors by the sequential \( \tau \) stable incorporation of tiles \( T_5 \) and \( T_6 \) using two strength 1 bonds each as depicted in Figure 1(c). Alternatively, tile \( T_7 \) can unstably bind to the assembly using only its West pad (\( f \)). Note that this interaction causes a mismatch indicated in red. Though the half-life of this fleeting interaction is small and in most cases \( T_7 \) detaches from the assembly, it is possible that during this interaction tile \( T_8 \) can \( \tau \) stably attach to the assembly using two strength 1 interactions. In doing so it \( \tau \) stabilizes the previously unstable \( T_7 \) using two strength 1 interactions. The entire assembly is now stable but carries a mismatched pad indicated in red in Figure 1(d).

This error could have been avoided if the East pad (\( i \)) of \( T_7 \) was unavailable for binding with the West pad (\( i \)) of \( T_8 \). Activatable tiles provide this capability by making pads unavailable for binding (inactive) until the appropriate moment. In the activatable version of the example described above, the East pads of \( T_5 \) and \( T_7 \) would have been protected (inactive) initially. The East pad of each tile is activated only after detecting the successful binding of both their respective West and South pads. Thus when \( T_5 \) binds \( \tau \) stably using both its West (\( f \)) and South (\( d \)) pads, its East pad (\( g \)) is activated and made available for binding with the West pad (\( g \)) of \( T_6 \). When \( T_7 \) binds unstably using only its West pad (\( f \)), its East pad (\( i \)) remains inactive preventing \( T_8 \) to \( \tau \) stabilize the erroneous assembly. [12] provide strand level designs to achieve this behavior via the use of a strand displacing polymerase.

In section 3 we look at the simpler case of one dimensional activatable assembly. In particular we demonstrate how to protect and subsequently activate pads for a one dimensional tiling system with two tile types.
3 Linear (1D) Assembly formed on Simultaneous Activation of Tiles

This section, demonstrates a 1D assembly formed from tiles which are all activated simultaneously. In the initial configuration of the system, there are two tile types present, both inactive. On the addition of a catalyst (*Bst Polymerase*), all instances of both tile types become active. The assembly formation proceeds thereafter. This experiment serves as a proof of concept, demonstrating that polymerases can be used in order to activate a set of tiles. Note that this model is not an exact solution to the idea proposed by [12], wherein the tiles should instead be activated in sequence.

3.1 Design of Linear Activating Assemblies

Here, we describe the simplified system designed to demonstrate activated assemblies. Our system consists of two tiles, Tile A and Tile B each having two sticky ends (pads). The sticky ends are programmed such that the two sticky ends of Tile A can hybridize to the two sticky ends of Tile B. This allows two copies tile A to attach to each copy of tile B and vice versa resulting in a linear co-polymer of alternating tile As and tile Bs. The tiles themselves are DNA duplexes with overhangs that act as sticky ends. The sticky ends of each tile
can be protected via hybridization with a protector complex rendering the tiles inactive. The protector complex contains a primer that can be extended by a strand displacing polymerase which results in deprotection of the tile. This system is illustrated in Figure 2.

![Figure 2: Co-Polymerization of Tile A and Tile B after Activation](image)

Initially the protected versions of tile A and B, labeled Inactive Tile A and Inactive Tile B are synthesized. Inactive Tile A is composed of two parts, Active Tile A and protector complex A. Active Tile A, comprised of the strands acbe and ce, is a double stranded DNA with overhangs that serve as sticky ends. Protector complex A consists of the strand ag1f and the primer f. Similarly Inactive Tile B is composed of two parts, Active Tile B and protector complex B. Active Tile B, comprised of the strands acbe and ce, is a double stranded DNA with overhangs that serve as sticky ends. Protector complex B consists of the strand bg2f and the primer f. Inactive Tile A and Inactive Tile B do not react with each other in solution.

In the presence of the strand displacing Bst DNA polymerase, the primers on both the inactive tiles are extended to form waste products Waste A and Waste B. In doing so, the sticky end a of Inactive Tile A and sticky end b of Inactive Tile B are exposed. This activation event results in the formation of Active Tile A and Active Tile B and they co-polymerize into a linear polymer with alternating A and B tiles. The 3' ends of the four DNA strands composing the tiles are augmented with a special domain e that is designed to have minimum interactions with any strand in the system. This ensures that the polymerase does not spuriously extend these strands.

We modified the system described above to undergo dimerization by omitting one of the sticky ends. Specifically, the sticky end b of the Active Tile B complex is omitted and this modified tile is labeled Blunt Tile B. Active Tile A, formed by the activation event described earlier, reacts with Blunt Tile B to form a dimer. This process is illustrated in Figure 3.

![Figure 3: Dimerization of Tile A and Tile B after Activation](image)

### 3.2 Experimental Demonstration of Linear Activating Assemblies

This section provides details of the actual experimental setup and data obtained from the experiments. Once the domain level design of the system was complete, the actual DNA sequences assigned to each domain was designed by hand, taking care to minimize sequence symmetry. The length of the protecting strand bg2f was designed to be greater
than the length of the protecting strand \( \bar{a}g_1 \bar{f} \) to better separate the two complexes in analytical gels. These sequences were then checked for any spurious secondary structures via the aid of online DNA folding servers of (author?) \(^3\). The optimized sequences were then ordered at 100 nmol synthesis scale from IDT DNA with standard desalting. The DNA strands were then PAGE purified and brought up to a working stock concentration of 20 \( \mu \)M. The inactive tile complexes were formed by mixing the constituent strands at equimolar ratio of 1 \( \mu \)M in 20 \( \mu \)L reaction volume, the only exception being the primer strand \( f \) which was added in 2 \( \times \) excess as lower primer concentrations resulted in decreased yield of inactive tiles. The mixture was then heated to 90 \(^\circ\)C in a buffer of 1 \( \times \) TAE with 12.5 mM Mg ions and cooled to room temperature in three hours.

1600 units of Bst DNA Polymerase, Large Fragment at concentration 8000 units/ml was ordered from New England Biolabs. Bst DNA Polymerase Large Fragment is a portion of the Bacillus stearothermophilus strand displacing DNA Polymerase protein, one unit of which can incorporate 10 nmol of dNTP’s into acid insoluble material in 30 minutes at 65 \(^\circ\)C. To test the activatable system, 4 picomol each of Inactive Tile A and Inactive Tile B with one unit of Bst DNA Polymerase were incubated in a reaction buffer of 20 mM Tris-HCl 10 mM (NH\(_4\))\(_2\)SO\(_4\), 10 mM KCl, 2 mM MgSO\(_4\) and 0.1 % Triton X-100 with 100 \( \mu \)g/ml BSA and 200 \( \mu \)M each of dATP, dCTP, dGTP and dTTP for 3 hours at room temperature. As a positive control, Active Tile A and Active Tile B were prepared in separate tubes and 4 picomol each were incubated at room temperature for 3 hours. As a negative control, 4 picomol each of Inactive Tile A and Inactive Tile B were incubated in polymerase buffer in the absence of Bst DNA Polymerase at room temperature for 3 hours. In another control experiment, 4 picomol each of the protector complex \((\bar{a}g_1f + f\) and \(bg_2 f + f\)) were incubated in separate test tubes with Bst DNA Polymerase in polymerase buffer at room temperature for 3 hours.

Figure 4 is a non-denaturing PAGE gel image showing the results of these experiments. Lane 2 and 3 contain Active Tile A and Active Tile B while lanes 4 and 5 contain Inactive Tile A and Inactive Tile B respectively. Lane 6 and 7 constitute a control experiment showing the extension of the primer on the protector complex A to produce Waste A; lane 6 contains protector complex A incubated in polymerase buffer while lane 7 contains protector complex A incubated in polymerase buffer with Bst DNA Polymerase to produce Waste A. Lanes 8 and 9 show an equivalent control experiment with protector complex B. Lanes 10 and 11 show the activation of Inactive Tile A and Inactive Tile B producing Active Tile A & waste A and Active Tile B & Waste B respectively. Lanes 12 and 13 show the negative and positive controls respectively; lane 12 shows Inactive Tile A and Inactive Tile B incubated in polymerase buffer while lane 13 shows Active Tile A and Active Tile B incubated in polymerase buffer. Both these lanes do not contain the Bst DNA Polymerase. Lane 12 shows no reaction between the inactive tiles while lane 13 shows the formation of linear polymers by active tiles. Finally, lane 14 shows the working of the system where Inactive Tile A and Inactive Tile B are incubated in polymerase buffer with Bst DNA Polymerase. We can see that the tiles are activated leading to the formation Waste A and Waste B and the activated tiles form a linear polymer. Figure 5 is a non-denaturing PAGE gel image showing the results of a similar experiment in the modified system with Blunt Tile B resulting in the formation of a dimer.

4 Directed Linear Activating Assemblies

4.1 Introduction

The scheme described in section 3 is an example of a one-time switch mechanism. On turning the switch on (by addition of either polymerase or primer), all tiles are activated, followed by subsequent formation of the assembly. In contrast, the scheme depicted in this section activates a tile only when the correct input has bound to it. Incase an incorrect tile binds as an input, the assembly does not proceed further. The assembly begins with the addition of an initiator, that only activates instances of the tiles that have bound to the initiator. The initiator is analogous to the seed tile
in the abstract Tile Assembly Model (aTAM)\textsuperscript{[25]}. Subsequent bindings activate subsequent tiles, and the assembly grows until it is feasible for it to do so. It is feasible for the assembly to proceed as long as the concentration of the reactants are above a certain threshold (after which it becomes too slow). Initial work \textsuperscript{[4]}, namely the Hybridization Chain Reaction by Pierce et. al., has demonstrated that directed assembly can be performed reliably using DNA molecules. Other similar work includes the work of \textsuperscript{[6]}, where they demonstrate the use of DNA protection molecules, which act analogously to glues in our activatable tiles system. They suggest two mechanisms PTM and LTM to protect input/output ends of tiles. The Protected Tile Mechanism (PTM) protects the input ends, but not the output ends. The output ends can thus engage in spurious attachment and cause errors in assembly. The Layered Tile Mechanism (LTM), protects both input and output ends, but the protection for the output ends is constrained to be 6nt in length (3nt + 3nt across two strands), so that the protection layer can float away once the correct set of inputs are bound. This is a constraint, which can cause both sequence limitations, and reduce the range of temperatures at which reactions have to occur. Also, they mention that experimental work using this mechanism, even at low concentrations, is extremely erroneous, because it is hard to get a high yield of accurately layered tiles in a one-pot reaction. In a one-pot reaction, all the necessary constituents are added initially, and a set of subsequent steps, such as annealing, isothermal incubation etc. is performed. This is in contrast to a multi-pot reaction, where different constituents may be added sequentially, or prepared separately and then added together. Work by Yin et. al. \textsuperscript{[29]} also demonstrates directed assembly using hairpins. They have defined a finite automata which shows the direction that the assembly is expected to take. However, their work also faces the same limitations as seen in strand displacement mechanisms, namely short toehold lengths and temperature limitations.

4.2 Design

Figure \ref{fig:activatable} shows an example of a simple activatable system, with directed assembly. The system consists of three hairpins, named \textit{Tiles A, B} and \textit{C}, named after their stem domains \textit{a}, \textit{b}, and \textit{c}. Each hairpin has a stem length of 21 nt, and a hairpin loop of 42 nt. The stem domains are initially inactive, since they are double stranded. The absence of overhangs in the tiles is to ensure that no spurious interactions should take place. When added together, the tiles do not initiate assembly.

On addition of the initiator, which is \textit{Seed C} (or just domain \textit{c}), \textit{tile A} becomes active. \textit{Seed C} binds to the hairpin of \textit{tile A}, and the polymerase activity activates \textit{tile A}. The mechanism of activation is exactly how the initiator activated \textit{tile A}, by binding to the hairpin region. This can be seen in figure \ref{fig:activatable}.

Note that the addition of the initiator only activated \textit{tile A} unlike the earlier system in section 3.1. \textit{Tile A} can now bind to \textit{tile B}, and the polymerase activity activates \textit{tile B}. The mechanism of activation is exactly how the initiator activated \textit{tile A}, by binding to the hairpin region. This can be seen in figure \ref{fig:activatable}.

\textit{Tile B} in turn activates \textit{tile C}, and \textit{tile C} then returns to activate another copy of \textit{tile A}. This has been designed to be cyclic, and theoretically can loop until the concentration of tiles or dNTPs is too low. The expected end product can be seen in

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{activatable_system.png}
\caption{An example of a simple activatable system, with directed assembly.}
\label{fig:activatable}
\end{figure}
the figure 6. It is of the form of a DNA ribbon, which has a central seam, and helices on either end, of length 63bp.

4.3 Experiments

There are 3 hairpins in the system. Each hairpin is 84 nt long, and is composed of 4 domains, each of 21 nt. The domains $a,b,c$ are the primer domains, which get activated sequentially. The domains $t_1,t_2,t_3$ are 21nt domains, designed in order to have least cross talk with other domains. They play no role in the assembly, other than to increase the size of the hairpin loop making it more elastic, so that the primer can bind without causing rigidity. The persistence length of DNA is 50 nm $\sim$150bp, and any double stranded region below this length acts as a rigid rod.

The results of the experiments can be seen in figure 2. Stepwise reactions have been performed to verify the activity of one, two and three hairpins, each step also involving a single seed and DNA polymerase. As can be seen in lane 2, the polymerase succeeds in opening the stem, activating tile B in the process. Gel analytics estimate the base pair mobility of this band to be about 73bp. This is in accordance with the expected size of activated tile B (It has 63bp dsDNA, and 21bp ssDNA).

Next, we verified that an individual hairpin acts as a primer after activation, while it does not do the same before activation. The lanes 4-8, include both positive and negative controls, and show that we do achieve this expected behaviour. A dimer, composed of two hairpins can be seen in lanes 4,5,6 of the same, and gel analytics reveal that the base pair mobility is about 132 bp. As before, this result is in accordance with our expected length (There are 126bp of dsDNA and 21bp of ssDNA).

Here, we introduce the 3rd and final tile to the mix. In the presence of activated tile B, tile C should get activated, and proceed to completing a cycle. These reactions are run at 37C, which is lower than the melting temperature of each of these tiles (52C). Hence, we do get the formation of trimers, as expected. However, our final goal is to achieve long chain polymers, which should increase linearly with incubation time.

The figure on the left ??, shows the state of the system on addition of 1x Seed, while the figure on the right shows the state of the system on adding 0.1x Seed. In the presence of 1x seed, the number of copies of the seed is the same as the hairpins. This results in each hairpin getting initiated by a seed, and not by another hairpin, resulting in most structures being trimers. In the presence of lower quantities of seeds however, a hairpin can be initiated by both a seed, or by another hairpin, which results in longer polymers. This however, was not obtained from the gel images. On quantitative analysis between the gel images on the left and right of figure 2, most of the product formed is a trimer, and this is proportional to the quantity of the seed being added in. Hence, the assembly does not proceed past the size of 3. There are however, a fewer higher order structures being formed. The fact that the assembly stops at 3 is counter-intuitive, and we try to examine why this might be so.

4.3.1 Longer Ribbon Formation

The objective of the design, being to form long ribbons, we ran an experiment to see if structures larger than 3 arm junctions were being formed. A 3-arm junction is formed in the order A-B-C, or in B-C-A as in our experiment. In order to form a hairpin ribbon, of the form B-C-A-B, we put in 2X Hairpin B, keeping hairpins A and C at 1X, and the Seed A at 1X. Once all the seed produces 3-arm junctions, 1x concentration of Hairpin B will be left, and the 3-arm junction should open the hairpin. The results of this experiment can be seen in figure 8. On the 5% and 6% gel, the darkest band, is a 3-arm junction, with progressively lighter bands that signify longer ribbons. Figure a) shows a 10% gel, while b) shows a 5% gel, to evaluate what the longer polymers are. The following analysis in table 1 reveals an approximate relative percentage of n-hairpin ribbons that have possibly formed, and the number of base pairs in those nanostructures. The values in green represent gel band movement that has correlated with what was expected, while the red values are inconsistent with what was predicted.
4.4 Discussions and Conclusions

We have achieved linear assemblies of sizes 1-9 tiles. The experimental goal of our project, was to be able to achieve error free assemblies of sizes as long as possible. However, we have encountered certain experimental issues, that have prohibited the synthesis of larger assemblies. We hypothesize the loss in growth of the assemblies, to two reasons. 1) After 3 tiles, the addition of a 4th tile seems to be lacking. The primer is 21 nt long, and each of the arms of the tiles are 63 bp. For the primer to attach to another hairpin, it will need to be unprotected by the steric hindrances it faces from the 3 other arms. Also, since the primer complement is present in the looped region of the hairpin, it’ll be even more sterically unstable for the hairpin to move into a position favoured by the primer to attach to. 2) Not unlike the steric hindrance faced by the primer, it is also sterically unfavourable for the polymerase subunit to attach to the 3’ end of the primer, and extend another copy of the hairpin. As the length of the assembly increases, the steric hindrances faced in the addition of a new hairpin, and in the involvement of the polymerase increase, leading to a decrease in the amount of synthesized assembly.

Steric hindrances can be reduced by introduce “spacer” domains, which will allow the arms of the assembly, to move a little freely around. Note the arms of the assembly are analogous to stiff rods, since the persistence length of double stranded DNA ranges from 100-150bp [8]. These spacer domains might improve the assembly, but only up to a certain length, after which they will face steric hindrances too.

Another explanation for forming mostly three tile assemblies, is that it is possible that the primer formed after the third assembly acts catalytically, and strand displaces the initiator. More details on this are in section 5.

Based on these observations, the use of hairpin tiles may not be an optimal approach to the synthesis of long assemblies. However, they do show the potential of working in a sequential manner, since our negative controls have been seen to have almost no spurious growth (lane 7, figure ??).

There are other solutions to the directed assembly problem, which do not involve hairpin tiles. One example of such a solution is present in figure 9, which is simply the same design, with the hairpins opened. Another solution for this is in figure ?? This solution is a mixture of both polymerase based strand displacement, and enzyme free strand displacement mechanisms. This also leads us to conclude that it is possible to design and implement more of these activatable systems, which are one handle to control self assembly.

5 Catalytic Assembly of Activatable Tiles

5.1 Introduction

As we see in the previous section, sequential assembly of tiles is a useful tool to direct growth in a particular orientation, at the nanoscale. We observe the shape of assembly that we form in the previous section, using three tiles, is a 3-arm junction. Note that after the addition of the third tile, the primer is made of domain c, which is complementary to the initial arm that the synthesis of assembly started from. This leads us to observe that the primer could infact be displacing the original initiator via strand displacement, and the polymerase would continue this displacement, leading to a closed 3-arm junction. The strand so synthesized, would now be able to start another sequence of assembling three tiles. It would then get displaced again, and sequence another three tiles, and continue this process, until the system is out of any
This protocol, could be extended to 4-arm, 5-arm ..., n-arm junctions in theory. As we have seen in the previous section, forming an assembly of even 4-arm junctions, is limited, due to the presence of steric hindrances. We hence experiment on 3-arm junctions currently. We also vary the types of catalysts we use, which gives us a range of experiments to evaluate against one another, and calculate the relative yield.

5.2 Design

The design protocol as shown in figure 10 is an example of a catalytic assembly of a 3-arm junction. The idea follows from the design in section 4.2, where an \( n \)-hairpin ribbon is formed in \( n \) steps. In this protocol, the primer gets "kicked off" after the formation of a 3-hairpin ribbon, and is involved in the catalysis of another 3-hairpin ribbon. It can keep catalyzing 3-hairpin ribbons, until it is feasible to do so.

The pathway for the system is identical to the pathway as described in the directed assembly model. An initiator activates tile A, which activates tile B. Tile B activates tile C. Tile C in this system, has a choice. It can either choose to activate another copy of tile A, or it can displace the extended initiator from the current ribbon, and hence form a 3-arm junction.

After forming a 3-hairpin ribbon, there is an open sticky end \( c \) of tile C. This sticky end can attach to two different strands, and hence there are two pathways for tile C to choose from as shown in figure 11. To make the correct catalytic pathway more likely for tile C, we provide \( c \) a toehold on another strand in the same nanostructure. Being a localized hybridization interaction, \( c \) should find this toehold faster than the alternate, i.e. a diffusion based interaction to find a copy of tile A. Another reason for giving tile C a region to hybridize, is to allow the polymerase to bind onto the 3-hairpin ribbon. This design is mentioned in section 5.3.2. Polymerases sit only on dsDNA, and hence this region is one where \( Bst \) Polymerase can attach and strand displace the extended initiator.

5.2.1 Initiator Strand Design

The initiator \( c_{0.5} \) shown in figure 10 initially is 11 nt long. The subscript 0.5 denotes that this initiator is half the length of the original initiator \( c \). The 2\textsuperscript{nd} half of \( c \), i.e., the half strand of \( c \) containing the 3’ end, has been retained. In the next step of the reaction, the polymerase extends this initiator to 53 nt. This extended initiator acts as the catalyst. It is made up of 3 domains \( c_{0.5}, t_1, a \). One of the domains \( a \) can interact with another tile B, and initiate a parallel pathway, which does not result in catalysis. These two pathways are shown in figure 11. The initiator gets sequestered, and is absorbed into a sink. Hence, the catalytic reaction has a pathway that needs to be avoided.

Note that the catalyst has a longer toehold region in the correct pathway than in the incorrect pathway. Hence, we surmise that the catalyst is more likely to bind to hairpin A instead of hairpin B. Following this intuition, we designed a set of experiments in order to verify the catalytic activity.

5.3 Experiments

5.3.1 Varying toehold lengths

We experimented with three different toehold \( c' \) lengths of 0, 11 and 21 nt. The catalysts used are as shown in figure 12 along with the toehold lengths, and the domains used. The longer the toehold, the easier it will be for the primer in the 3rd tile, to attach to, followed by subsequent strand displacement by \( Bst \) Polymerase. Prior to strand displacement by the polymerase, we expect toehold mediated strand displacement. This will also give the \( Bst \) Polymerase a larger duplex region to attach to. The gel images for each of the toehold lengths is as shown in figure 13.

Catalytic Activity was not obtained with any of the three different catalysts. Each of the three catalysts only show a linear increase in the amount of product formed, as their concentrations are increased. The experiments were also run over different time intervals (See supplement ??), but there was no detectable change (based on image analysis of the electrophoresis gel) in activity past 6 hours incubation time.
5.3.2 Eliminating the Incorrect Pathway

To block such unintended interactions, we modified the catalyst as shown in figure 14. We reduce the length of the domain $a_0$ to half its length (10 nt), and call it $a_{0.5}$. 10 nt is a small enough toehold to dissociate from any spurious interactions this domain may have at 37° C. Hence, as seen in figure 14b), the incorrect pathway now reduces to spurious association and dissociation of the primer from the hairpin. We also inhibit the polymerase from spuriously extending the primer on the 3' end. This inhibition of activity, is represented by a black mark in the figure, and the way to accomplish this is by deactivating the hydroxyl group at the 3' end. There are several 3' modifications, such as 3' phosphorylation of the DNA backbone, or ddNTPs\[\], which are commercially available \[22\] and can be used to prevent polymerase activity.

Note that the catalyst uses different mechanisms to cause strand displacement in the tiles. For the first tile, it uses toehold mediated strand displacement \[32\]. The toehold is present on the hairpin loop region of the tile C. On opening the first tile, the primer is exposed, which is free to bind to the next tile, and open it via the strand displacement activity of the *Bst Polymerase*. The same mechanism is used in the opening of the third and final tile. The primer of the final tile, has a toehold on the first tile, and it is the polymerase that acts to release the catalyst from the 3-hairpin ribbon.

5.4 Discussions and Conclusions

The next steps would be to experimentally demonstrate the working of the new type of catalyst introduced into the system. We also have not conclusively verified the reasoning behind the polymerase unable to act catalytically. There is also an alternate design for a system, without any hairpins, as shown in figure 9, and we are considering that design, to demonstrate proof of concept. However, the lure of hairpin based systems, is that hairpins by themselves are quite stable, and the probability of spurious nanostructures being formed is extremely low. Also, in non-hairpin systems, the self-assembled product tends to dissociate from each other, while in hairpin based systems, growth continues from the self-assembled product, which in turn causes the assembly to grow larger. Given these factors, we are modifying the current design, with a goal of a functioning catalytic system.
Figure 4: PAGE analysis of polymerization of activatable tiles: 10% Non-Denaturing PAGE Gel. Lane 1: 10 Basepair ladder, Lane 2: Active Tile A, Lane 3: Active Tile B, Lane 4: Inactive Tile A, Lane 5: Inactive Tile B, Lane 6: Protector Complex A, Lane 7: Protector Complex A + Polymerase, Lane 8: Protector Complex B, Lane 9: Protector Complex B + Polymerase, Lane 10: Inactive Tile A + Polymerase, Lane 11: Inactive Tile B + Polymerase, Lane 12: Inactive Tile A + Inactive Tile B, Lane 13: Active Tile A + Active Tile B, Lane 14: Inactive Tile A + Inactive Tile B + Polymerase, Lane 15: 10 Basepair ladder. Gel and running buffer contain 1 X TAE with 12.5 mM Mg. The gel was run for 7 hours at 125 V and then stained and destained using Ethidium Bromide for 45 minutes each. Image was obtained using Alpha Imager.
Figure 5: PAGE analysis of dimerization of activatable tiles: 10 % Non-Denaturing PAGE Gel. Lane 1: 10 Basepair ladder, Lane 2: Active Tile A, Lane 3: Blunt Tile B, Lane 4: Inactive Tile A, Lane 5: Protector Complex A, Lane 6: Protector Complex A + Polymerase, Lane 7: Inactive Tile A + Polymerase, Lane 8: Inactive Tile A + Blunt Tile B, Lane 9: Active Tile A + Blunt Tile B, Lane 10: Inactive Tile A + Blunt Tile B + Polymerase. Lane 11: 10 Basepair ladder Gel and running buffer contain 1 X TAE with 12.5 mM Mg. The gel was run for 7 hours at 125 V and then stained and destained using Ethidium Bromide for 45 minutes each. Image was obtained using Alpha Imager.
Figure 6: Directed Assembly of Tiles

Figure 7: Directed (sequential) Assembly of 3 Tiles
Figure 8: Testing the formation of a Four Hairpin Ribbon
Figure 9: Future Work - Another Design

Figure 10: Catalytic Formation of 3-Hairpin Ribbons

Catalytic Assembly of Tiles
Correct Catalytic Pathway

Incorrect Catalytic Pathway

Figure 11: Two Catalytic Pathways

Figure 12: Catalyst Activity Mechanism, 3 Catalysts Used
Figure 13: Gel Images for each of the 3 catalysts
Catalyst Released
Polymerase Activity Not needed in this step

Indicates No Possible Primer Extension on 3' end

Catalyst New Mechanism

Figure 14: Catalyst New Mechanism
References


