JOURNAL OF THE ROYAL SOCIETY Interface

FirstCite® e-publishing

*J. R. Soc. Interface*
doi:10.1098/rsif.2011.0819
*Published online*

# Meta-DNA: synthetic biology via DNA nanostructures and hybridization reactions

**Harish Chandran[1,*], Nikhil Gopalkrishnan[1], Bernard Yurke[2] and John Reif[1,3]**

[1]*Department of Computer Science, Duke University, Durham, NC, USA*
[2]*Department of Materials Science and Engineering and Department of Electrical and Computer Engineering, Boise State University, Boise, ID, USA*
[3]*Adjunct Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia*

Can a wide range of complex biochemical behaviour arise from repeated applications of a highly reduced class of interactions? In particular, can the range of DNA manipulations achieved by protein enzymes be simulated via simple DNA hybridization chemistry? In this work, we develop a biochemical system which we call meta-DNA (abbreviated as mDNA), based on strands of DNA as the only component molecules. Various enzymatic manipulations of these mDNA molecules are simulated via toehold-mediated DNA strand displacement reactions. We provide a formal model to describe the required properties and operations of our mDNA, and show that our proposed DNA nanostructures and hybridization reactions provide these properties and functionality. Our meta-nucleotides are designed to form flexible linear assemblies (single-stranded mDNA (*ss*mDNA)) analogous to single-stranded DNA. We describe various isothermal hybridization reactions that manipulate our mDNA in powerful ways analogous to DNA–DNA reactions and the action of various enzymes on DNA. These operations on mDNA include (i) hybridization of *ss*mDNA into a double-stranded mDNA (*ds*mDNA) and heat denaturation of a *ds*mDNA into its component *ss*mDNA, (ii) strand displacement of one *ss*mDNA by another, (iii) restriction cuts on the backbones of *ss*mDNA and *ds*mDNA, (iv) polymerization reactions that extend *ss*mDNA on a template to form a complete *ds*mDNA, (v) synthesis of mDNA sequences via mDNA polymerase chain reaction, (vi) isothermal denaturation of a *ds*mDNA into its component *ss*mDNA, and (vii) an isothermal replicator reaction that exponentially amplifies *ss*mDNA strands and may be modified to allow for mutations.

**Keywords: DNA self-assembly; synthetic biology; DNA nanostructures**

## 1. INTRODUCTION

### *1.1. Synthetic biology using DNA nanosystems*

A major goal of synthetic biology is to produce synthetic biochemical systems which have functions similar to the biochemical functions of living organisms. Considerable related work has been done in the fields of artificial and synthetic life, see [1–3] for an overview of these fields. Two predominant approaches by researchers in this area have been to design protein-based or RNA-based biochemical systems. However, both proteins and RNA are very difficult to predictively design for given functions, and behaviour of the resulting protein-based or RNA-based biochemical systems can be very complex, making their engineering highly challenging. An alternative approach we propose here is to produce synthetic biochemical systems based on a very well understood, relatively robust molecule, such as DNA (DNA strands and DNA nanostructures). DNA–DNA interactions such as hybridization and strand displacement are reasonably well understood and a vast literature exists that studies, models, predicts and even controls such interactions (see [4] for a comprehensive review). The scale and complexity of experimental demonstrations have increased dramatically, including autocatalytic systems [5,6], scalable Boolean circuits [7,8] and neural network circuits [9]. The role of auto-catalysis in replication was studied in Plasson *et al.* [10]. These efforts enable us to programme intricate DNA systems using simple design rules.

In this work, we design synthetic biochemical systems, termed meta-DNA (mDNA) systems, consisting only of DNA nanostructures that capture the properties and structure of DNA in biological systems. Our work is reductive: we use simple DNA chemistry to emulate more complex enzyme-based DNA chemistry through use of toehold-mediated DNA strand displacement

*Authors for correspondence (harish@cs.duke.edu).

Electronic supplementary material is available at http://dx.doi.org/10.1098/rsif.2011.0819 or via http://rsif.royalsocietypublishing.org.

*Received* 23 November 2011
*Accepted* 12 December 2011

systems and DNA hairpin interactions. From a computer science perspective, our work can be thought of as using a lower level programming language to express programs encoded in a higher level programming language with the aim of a better abstract understanding of the phenomena. From a synthetic biology perspective, we are building novel biochemical systems to emulate useful, well-known natural biological systems and providing alternatives to enzymes. From an engineering perspective, our work is a minimalist approach to designing biochemical systems from simple, predictable yet powerful modules. Our systems are largely isothermal and autonomous which suggest that they may have applications for biochemical systems such as transport devices, molecular motors, detection, signalling and computing systems.

### *1.2. Organization of this paper*

In §2, we list the desirable properties of mDNA and give a design for a meta-nucleotide, single (*ss*mDNA) and double-stranded mDNA (*ds*mDNA), along with its secondary structure. In §3, we discuss the mDNA reactions of mDNA hybridization, mDNA denaturation, mDNA strand displacement, mDNA polymerization, mDNA restriction, mDNA helicase denaturation, exponential amplification using an mDNA replicator and mDNA synthesis. Section 4 discusses how mutations may be introduced in mDNA replications to perform directed evolution and addresses experimental challenges in implementing mDNA systems.

## 2. DESIRED PROPERTIES OF META-DNA

DNA is an ideal material to construct nanoscale structures, circuits and devices and has been used as scaffolding material for complex shapes, fuel for molecular motors and aptamers for various organic and inorganic molecules. The key properties of DNA that enable these functionalities are its programmability, predictable chemical interaction and secondary structure along with simple laboratory protocols for its manipulation. Synthetic DNA is also cheaply and readily available from a variety of commercial sources. At the most abstract level, fabricating structures and devices with DNA is akin to working with smart bricks that fit together in a specific predefined way and then putting them in a bag, shaking it and waiting for the bricks to self-assemble. We wish to abstract the structure and reactions of DNA and emulate them using only DNA–DNA interactions. In doing so, we would have circumvented DNA–enzyme chemistry with a synthetic biochemical system that uses only DNA hybridization. We list the desirable interactions of mDNA analogous to the key structural properties and biochemical reactions of DNA.

### *2.1. List of desirable interactions for meta-DNA*

— *ss*mDNA is a linear polymer of *meta-nucleotides*. Each meta-nucleotide (implemented by a DNA nanostructure, see figure 1*c*) contains both a segment of the *meta-backbone* of the molecule, which holds the mDNA together, and a *meta-base*. A meta-nucleotide is directional and has a 3′ (three prime) and 5′ (five prime) end. The active 5′ end of any meta-nucleotide should be able to bind to the active 3′ end of any other meta-nucleotide and vice versa via a meta-backbone link. Since each nucleotide has directionality a *ss*mDNA strand is asymmetric.

— Each meta-nucleotide is one of $2k$ types (where $k$ is a positive integer), $k$ meta-purines and $k$ meta-pyrimidines. There is a pairing between the meta-purines and the meta-pyrimidines, where each edge indicates an overwhelming complementary preference of meta-base linking between the meta-purine and meta-pyrimidine it connects.

— The meta-backbone link is a *strong* bond (implemented as a long sequence of hybridized DNA) while the meta-base link is a *weak* bond (implemented as a short sequence of hybridized DNA). Meta-nucleotides do not spontaneously form meta-backbone bonds in the absence of an appropriate DNA nanostructure catalyst.

— Two *ss*mDNA sequences that are reverse complementary (a complementary sequence with opposite directionality) to each other have the ability to *meta-hybridize* to form a *ds*mDNA. The strands of any *ds*mDNA are anti-parallel.

— The process of separation of a *ds*mDNA into its two constituent *ss*mDNA by the breaking of meta-base bonds is *mDNA denaturation*. When an ensemble of *ds*mDNA with the same sequence is heated to a temperature known as its melting temperature, half of the ensemble denatures into its constituent *ss*mDNA. The melting temperature depends on the sequence of meta-nucleotides of the mDNA and also on its length, with longer strands having a higher melting temperature.

— *ss*mDNA is flexible like a freely jointed chain and has small persistence length when compared with *ds*mDNA of comparable length which is like a worm-like chain.

— *Toehold-mediated DNA strand displacement* is the displacement of a single strand of DNA from a double helix by an incoming strand with a longer complementary region to the template strand. The incoming strand has a *toehold*, an empty single-stranded region on the template strand complementary to a subsequence of the incoming strand, to which it binds initially. It eventually displaces the outgoing strand via a kinetic process akin to a one-dimensional random walk. Strand displacement is a powerful phenomenon that allows one to design complex dynamic DNA nanosystems such as walkers, catalytic amplifiers, motors and circuits (see [4]). *mDNA strand displacement* is the displacement of a shorter *ss*mDNA from a template by a longer invading *ss*mDNA. To support mDNA strand displacement protocols, the weak meta-base bonds must be continuously broken and remade in a kinetic process called breathing. The rate of breathing should be positively correlated with temperature.

— *mDNA restriction* is the cleaving of the meta-backbone at a sequence specific recognition site by a set of DNA strands executing a sequence of DNA strand displacement reactions.

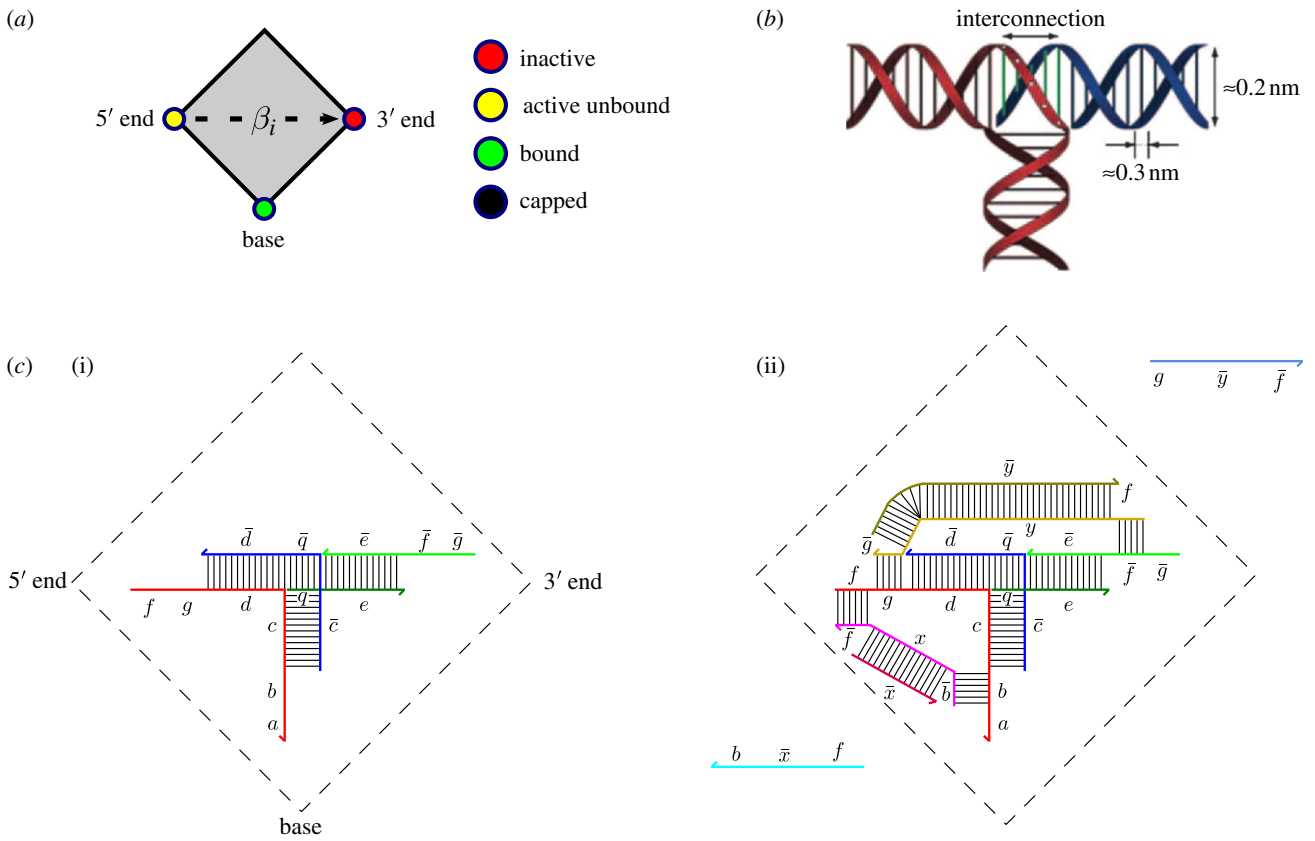



Figure 1. Meta-nucleotide structure and properties: the T-junction is used to achieve mDNA connectivity and geometry. A protected meta-nucleotide is used to prevent spontaneous aggregation of tiles into mDNA. (*a*) Abstraction of a meta-nucleotide. (*b*) Design of rigid T-junction self-assembled from DNA. Adapted from Hamada & Murata [11]. (*c*) Design of the meta-nucleotide: (i) an unprotected meta-nucleotide and (ii) a protected meta-nucleotide: the protection mechanism is designed to impose the requirement that the 5′ pad is activated before the 3′ pad.

— *mDNA polymerization* is the process of sequentially adding meta-nucleotides to the 3′ end of a primer *ss*mDNA strand meta-hybridized to a longer template *ss*mDNA strand. The meta-nucleotides added are complementary to the corresponding meta-nucleotides on the template.

### 2.2. Distinctions between meta-DNA and DNA

While mDNA structure and reactions are designed to closely simulate DNA structure and reactions, there are key distinctions. mDNA has a larger set of distinct bases allowing greater sequence diversity. While *ss*mDNA and *ds*mDNA are designed to simulate the physical properties of *ss*DNA and *ds*DNA, they have different sizes, flexibilities, persistence lengths and thermal stabilities. These differences imply that while DNA has been used to construct intricate two-dimensional and three-dimensional nanostructures, mDNA is unlikely to have the same capabilities. The kinetics of mDNA reactions are expected to be slower than DNA enzymatic reactions, however mDNA reactions consume less energy.

### 2.3. Abstract description of meta-DNA

We introduce an abstract tile-based model to represent mDNA reactions, similar to the abstract tile assembly model (TAM) [12]. While TAM is geared towards studying the computational power of self-assembly (see [13–20]), our tile-based model is used to represent the dynamic assembly processes and state changes that occur in mDNA reactions. We model a meta-nucleotide as an activatable tile [21] having three activatable pads: a 5′ pad, a 3′ pad and a base pad and represent it by a square tile as illustrated in figure 1*a*. The tile has directionality as indicated by an arrow from 5′ to 3′ which is imposed by the sequence in which the pads are activated, with 5′ always activated before 3′. Tiles bind to each other via symmetric pad interactions called *binding* or *linking*. Each binding has a *strength* associated with it (1, 2 or 3) that depends only on the type of pads involved in the binding. The strength of a binding models the energy required to break the bond. Base pads can only interact with other base pads through strength 1 bindings, and are called *meta-base bindings*. 5′ pads can only interact with 3′ pads through strength 3 bindings and these are called *meta-backbone bindings*. The pads exist in one of four states: inactive, active unbound, bound and capped. Inactive pads do not bind with other pads. Any active unbound 5′ pad of a meta-nucleotide can bind to any active unbound 3′ pad of another meta-nucleotide via a bond of strength 3 after which these pads go to the bound state. There are $2k$ different base pads, each corresponding to a meta-nucleotide type. These are split into two complementary perfectly matched sets. Let $\mathfrak{P} = \{\beta_1, \beta_2, \ldots, \beta_k, \bar{\beta}_1, \bar{\beta}_2, \ldots, \bar{\beta}_k\}$ be the set of pads.

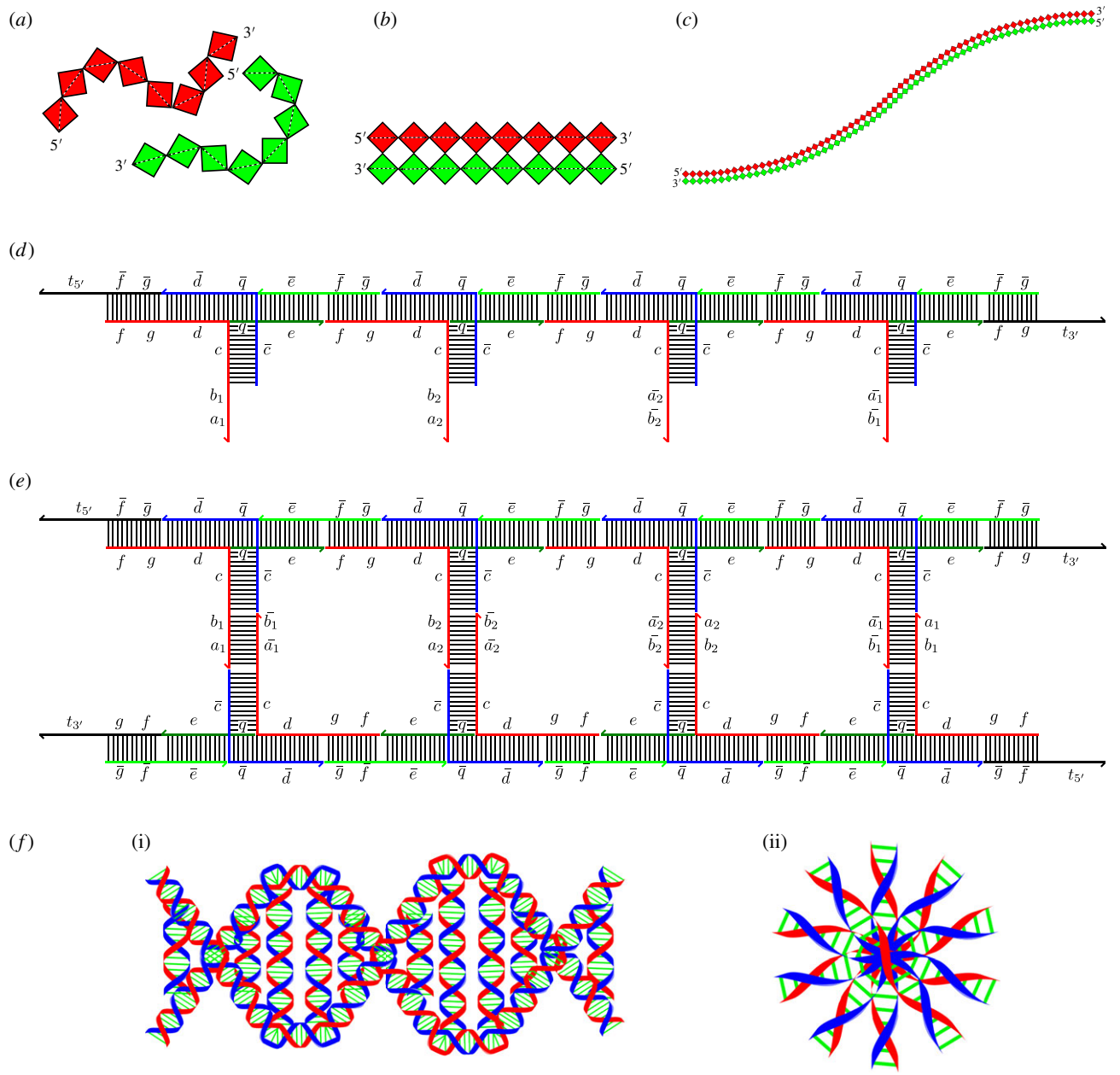



Figure 2. *ss*mDNA and *ds*mDNA: a two *ss*mDNA held together by base pad bindings. *ss*mDNA is flexible with a low persistence length, while *ds*mDNA is stiff and has a higher persistence length. These properties arise out of the geometric constraints imposed by the meta-nucleotide structure. (*a*) *ss*mDNA. (*b*) *ds*mDNA. (*c*) *ds*mDNA has a longer persistence length. (*d*) Internals of a *ss*mDNA. (*e*) Internals of a *ds*mDNA. (*f*) (i) Secondary structure of *ds*mDNA forming a double helix (ratio of bases per axial rotation is programmable via strand design). Not to scale. (ii) Axial view of *ss*mDNA. Only the rungs of the twisted ladder structure are shown.

For all $i$, pads $\beta_i$ and $\bar{\beta}_i$ bind with each other. The face label on any tile indicates the base type of the meta-nucleotide. Any active unbound base pad of a meta-nucleotide can bind in an anti-parallel manner to an active unbound complementary base pad of another meta-nucleotide by a strength 1 bond after which these pads go to the bound state.

A linear chain of these tiles held together by $5'$–$3'$ pad bindings forms a directional polymer of meta-nucleotides and hence is *ss*mDNA. The $3'$ base pad of the tile at the $3'$ end of the *ss*mDNA and the $5'$ base pad of the tile at the $5'$ end are always in the capped state. A *ds*mDNA is a dimer of two *ss*mDNA held together by base pad bindings. *ss*mDNA is flexible with a low persistence length, while *ds*mDNA is stiff and has a higher persistence length. These properties arise out of the geometric constraints imposed by the tile base structure of mDNA as illustrated in figure 2.

### 2.4. Strand design for meta-DNA

Recently, Hamada & Murata [11] reported a novel self-assembled rigid T-shaped interconnected junction where each arm is a DNA double helix (figure 1*b*). They synthesized tile-based structures such as one-dimensional linear ladders, one-dimensional ringed structures and two-dimensional lattices using the T-junction geometry. This suggests that their

T-junction motif is a useful widget for designing stable, rigid, well-behaved self-assembled objects. We use the T-junction as the key motif in achieving mDNA connectivity and geometry. Other DNA motifs can be imagined that give similar connectivity and geometry, but we choose the T-junction for its simplicity. Figure 1*c* shows an implementation of a meta-nucleotide tile as a DNA nanostructure. In figure 1*c*(i), we have the tile with no *protection* and in figure 1*c*(ii) we have the tile with *protection* strands. The purpose of the protecting strands is to prevent spontaneous aggregation of tiles into mDNA. The protection mechanism is designed to impose the requirement that the 5′ pad is activated before the 3′ pad (this property is used in the polymerization and replicator protocols). Each letter denotes a DNA sequence and a bar atop a letter indicates reverse complement of the sequence that the letter denotes. The red strand ($fgdcba$) contains the 5′ pad $fg$ and also the base pad $ba$ and the light green strand ($\bar{g}\bar{f}\bar{e}$) contains the 3′ pad $\bar{g}\bar{f}$. The blue strand ($\bar{c}\bar{q}\bar{d}$) and the dark green strand ($qe$) are bridging strands that hold the nanostructure together and give it the required geometry. We will have a detailed discussion of the secondary structure of the meta-nucleotide tile and mDNA in §2.5. The sequence of reactions that occur when a tile is deprotected are as follows. The pink strand $\bar{b}x\bar{f}$ protects (renders inactive) the 5′ pad. When $ba$ binds to its complement on another tile, $\bar{b}x\bar{f}$ is ripped away from the 5′ pad by the invasion of the strand $f\bar{x}b$ thus activating the 5′ pad. The strand $fy\bar{g}$ protects (renders inactive) the 3′ pad. When the 5′ pad binds to its complement on another tile, $fy\bar{g}$ is ripped away from the 3′ pad by the invasion of the strand $g\bar{y}\bar{f}$ thus activating the 3′ pad. Figure 3 illustrates all possible states a meta-nucleotide can exist in. We model *weak* base bonds and *strong* backbone bonds by making the sequences corresponding to the 5′ and 3′ pads much longer than the sequences corresponding to the base pads. The internal structure of both *ss*mDNA and *ds*mDNA are shown in figure 2*d*,*e*. Note the black protection strands on the tiles at the ends of the mDNA to implement capping. These protection strands can be ripped away from the mDNA by a complementary strand, allowing the *ss*mDNA to act as a primer.

### 2.5. Secondary structure of meta-DNA

Apart from simulating the reaction properties of DNA, we also wish to simulate its secondary structure, in particular:

— the flexible nature of *ss*DNA characterized by shorter persistence length;
— the rigid nature of *ds*DNA characterized by a longer persistence length; and
— the double helical structure of *ds*mDNA.

*ss*mDNA (figure 2*d*) is a linear polymer of T-junctions with consecutive base pad sections (the double helical structure $c \equiv \bar{c}$ where $\equiv$ symbolizes hybridization between two strands) not in the same plane, but slightly rotated so that they stick out of the plane of the paper. This rotation is controlled by carefully choosing the number of bases that make up the horizontal double helical section between consecutive vertical helical sections (figure 2*d*). This secondary structure for *ss*mDNA induces a helical twist for *ds*mDNA (imagined in figure 2*f*). We can think of the secondary structure of *ds*mDNA as a twisted one-dimensional ladder. The one-dimensional ladder design in Hamada & Murata [11] can be easily modified by adding or deleting a single base pair from the side rungs to induce a twist to get a double helix structure with approximately 10.5 meta-bases per turn of the double helix, mimicking the twist of *ds*mDNA. In particular, we choose the length of the repeating DNA double helical unit $efgdq \equiv \bar{q}\bar{d}\bar{g}\bar{f}\bar{e}$ to be either 41 (one less than the number of bases in four full turns of a DNA double helix) or 43 (one greater than the number of bases in four full turns of a DNA double helix). The following choice of lengths for the subsequences would potentially give us the required geometries and at the same time preserve the thermodynamic and kinetic properties that would allow our subsequent mDNA protocols to succeed ($|x|$ is the length of the DNA sequence represented by $x$): $|a_i| = |b_i| = |\bar{a}_i| = |\bar{b}_i| = 4, |c| = |\bar{c}| = |d| = |\bar{d}| = |q| = |\bar{q}| = 6$, $|f| = |\bar{f}| = |g| = |\bar{g}| = 12$ and $|e| = |\bar{e}| = 5$ or 7.

## 3. REACTIONS IN META-DNA

We set out protocols for mDNA that mimic DNA–DNA and DNA–enzyme interactions. We have two kinds of figures in the discussions that follow. The abstraction diagrams illustrate the protocols in the abstract activatable tile model while the internal structure diagrams illustrate the protocols in greater detail.

### 3.1. Assumptions for meta-DNA reactions

Before describing our protocols involving mDNA we give here our assumptions under which these reactions proceed. We also describe some simple rules of DNA chemistry which are repeatedly used like subroutines in our mDNA protocols.

— Our systems are maintained at only three different temperatures characterized by a parameter $\tau$. At room temperature or $\tau = 1$, both the meta-backbone bonds and the meta-base bonds are stable but breathing still occurs. Recall that breathing is the phenomenon of meta-base bonds spontaneously breaking and forming. At the melting temperature, $\tau = 2$, the meta-backbone bonds are stable but the base bonds are broken. At freezing temperature or $\tau = 0$, breathing does not occur.
— The pH, salt concentrations and other factors that affect hybridization are set to levels such that spontaneous hybridization between a DNA sequence and its complement can occur.
— Strand displacement, as defined earlier, always occurs and proceeds to completion. The strand that gets displaced out remains in the solution.
— We assume that each subsequence, denoted by a letter in the figures, only interacts with its perfect and full reverse complements and no other spurious interactions occur.

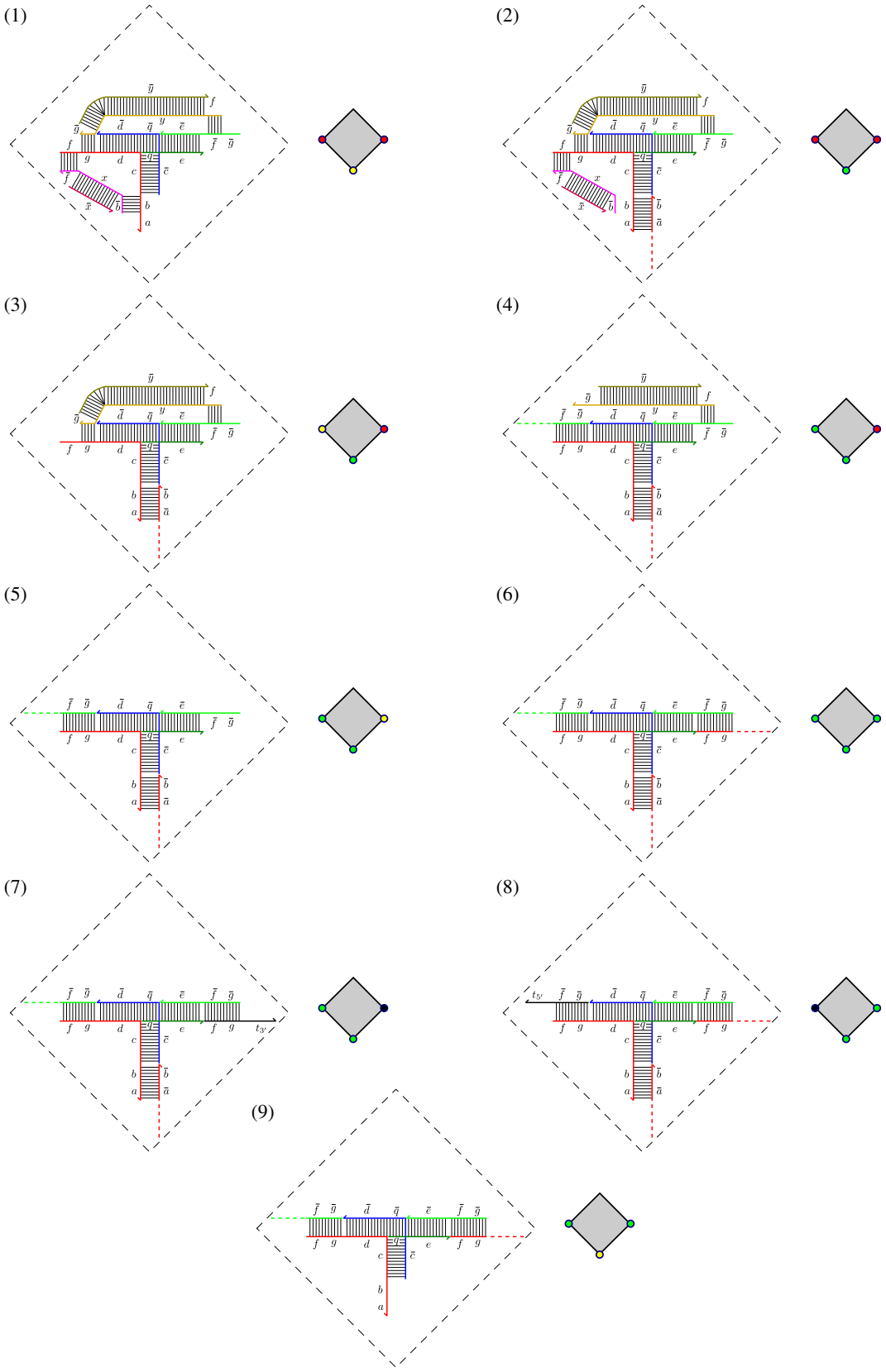



Figure 3. All possible states of a meta-nucleotide. Each mDNA reaction can be viewed in terms of state changes of the involved meta-nucleotides.

### *3.2. Meta-DNA hybridization and meta-DNA denaturation*

The simplest mDNA reactions are mDNA hybridization and mDNA denaturation, which are reverse reactions of each other. In mDNA hybridization, at temperature $\tau = 1$ two complementary *ss*mDNA strands bind via complementary base pad bindings to give *ds*mDNA (figure 4*a*). Heating *ds*mDNA to temperature $\tau = 2$ denatures the structure into its two component *ss*mDNA (figure 4*b*).

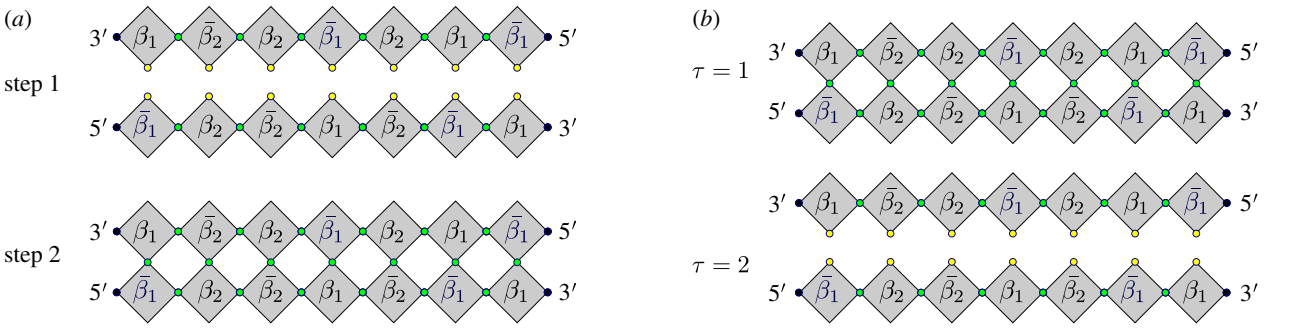



Figure 4. mDNA hybridization and mDNA denaturation by heating. (*a*) mDNA hybridization at $\tau = 1$. (*b*) mDNA denaturation at $\tau = 2$.

### *3.3. Meta-DNA strand displacement process*

mDNA strand displacement is defined as displacement of a *ss*mDNA from a meta-double helix by an incoming *ss*mDNA with a longer complementary region to the template strand. This reaction occurs at temperature $\tau = 1$. Figure 5 gives a high-level view of mDNA strand displacement using the activatable tile model. Two *ss*mDNA strands compete to hybridize with a single *ss*mDNA. The shorter of the two is completely hybridized to the template while the longer one comes in by gaining a toehold. Now, breathing of the meta-bases of the short strand gives an opportunity to the meta-bases of the competing incoming strand to hybridize with the template. Note that the intermediate steps are reversible. However, once the incoming *ss*mDNA strand completely displaces the outgoing *ss*mDNA strand, the reaction stops as the outgoing *ss*mDNA strand is extremely unlikely to come back in as it lacks a toehold.

Note that this reaction is made possible because the weak meta-base bonds can breathe at $\tau = 1$. Breathing in mDNA is expected to occur at a slower rate than in DNA because we require multiple bases to spontaneously denature for a single meta-base to denature. Also, contiguous bases in DNA are more localized than contiguous meta-bases in mDNA and hence the rate at which a meta-base occupies an empty spot on a complementary meta-base is also expected to be slower than for the corresponding process in DNA. Owing to these reasons, we would expect mDNA strand displacement to proceed slower than strand displacement in DNA.

### *3.4. Meta-DNA polymerization*

Polymerization in mDNA occurs by the extension of a *ss*mDNA, called a primer, by additions of free meta-nucleotides in the solution to the 3′ end via meta-backbone bonds. The meta-nucleotides added have base pads complementary to the corresponding base pads of the template strand. The protection strand $\bar{b}x\bar{f}$ on meta-nucleotides prevent their spontaneous aggregation while the protection strand $fy\bar{g}$ prevents de novo (in the absence of primer) polymerization (figure 1*a*).

Figure 6 gives an activatable tile model view of a single step in mDNA polymerization. Each base pad of the template strand that is not bound to its complement is in the active unbound state. The 3′ pad of the tile at the 3′ end of the primer is in the capped state (step 1).

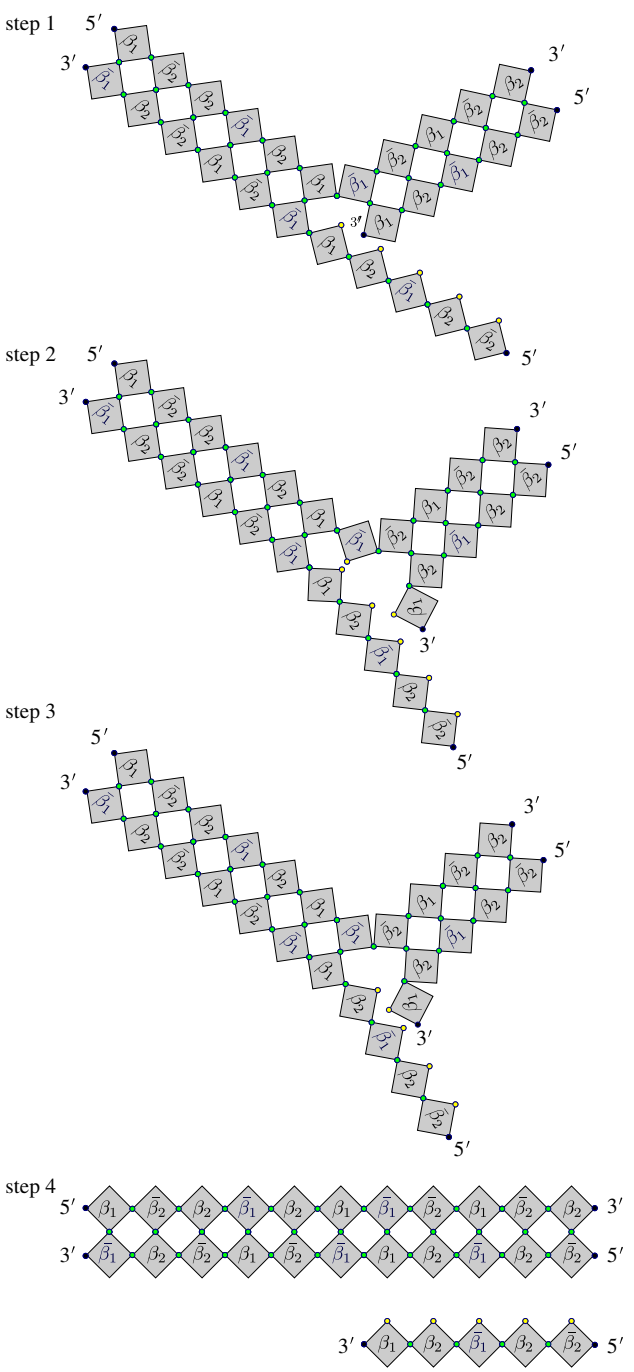



Figure 5. mDNA strand displacement (at $\tau = 1$) of a *ss*mDNA from a meta-double helix by an incoming *ss*mDNA with a longer complementary region to the template strand.

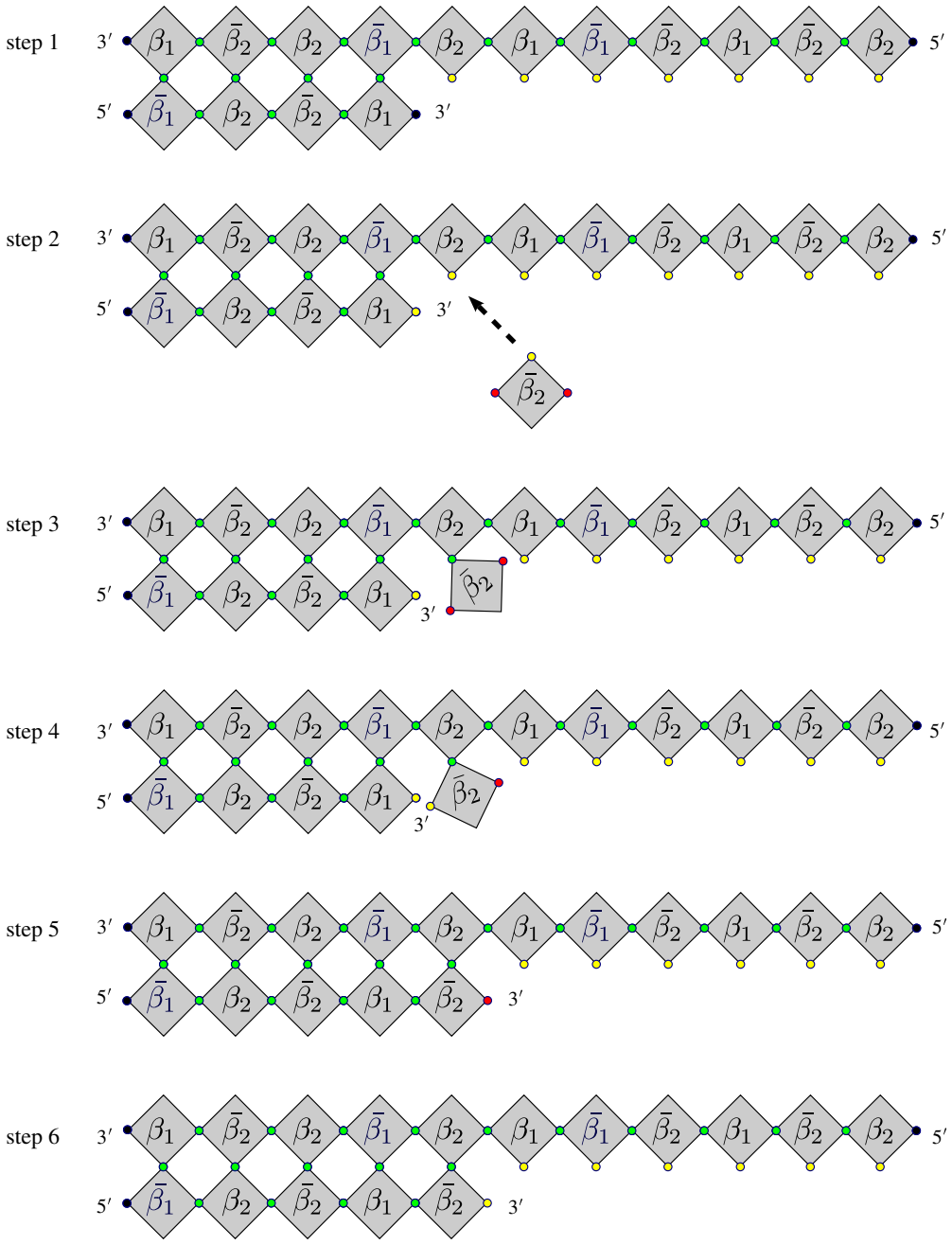



Figure 6. mDNA polymerization: polymerization in mDNA occurs by the extension of a *ss*mDNA, called a primer, by additions of free meta-nucleotides in the solution to the 3′ end via meta-backbone bonds. The meta-nucleotides added have base pads complementary to the corresponding base pads of the template strand.

It transitions to the active unbound state when its black capping strand is displaced by an initiator strand. The incoming meta-nucleotide (step 2) has its 5′ and 3′ pads in the inactive state, while its base pad is in the active unbound state. The complementary base pads bind (step 3), activating the 5′ pad of the incoming nucleotide (step 4). The 5′ pad then binds to the 3′ pad of the previous meta-nucleotide (step 5) causing the 3′ pad of the incoming nucleotide to transition to the active unbound state (step 6). This process occurs repeatedly till either no further free meta-nucleotides are available or the end of the template strand is reached. There is an alternate mechanism to stop the polymerization, which involves adding to the solution a black capper strand to transition the 3′ pad of growing 3′ end to a capped state, thus stopping further additions. Note that in a solution with an ensemble of these nanostructures, in general, the stopping point of mDNA polymerization cannot be carefully controlled and such an attempt would probably lead to various length subsequences of the fully complementary *ss*mDNA. However, we can exclude certain meta-base types from the meta-nucleotide mix in the solution ensuring that the polymerization halts when the complementary meta-base is encountered on the template strand. Since we have access to a larger alphabet of bases in our mDNA systems, we can set aside a set of bases for such purposes, analogous to *stop codons* in translation.

We will now examine this protocol in greater detail by looking at the internal strand structure (figures 7 and 8).

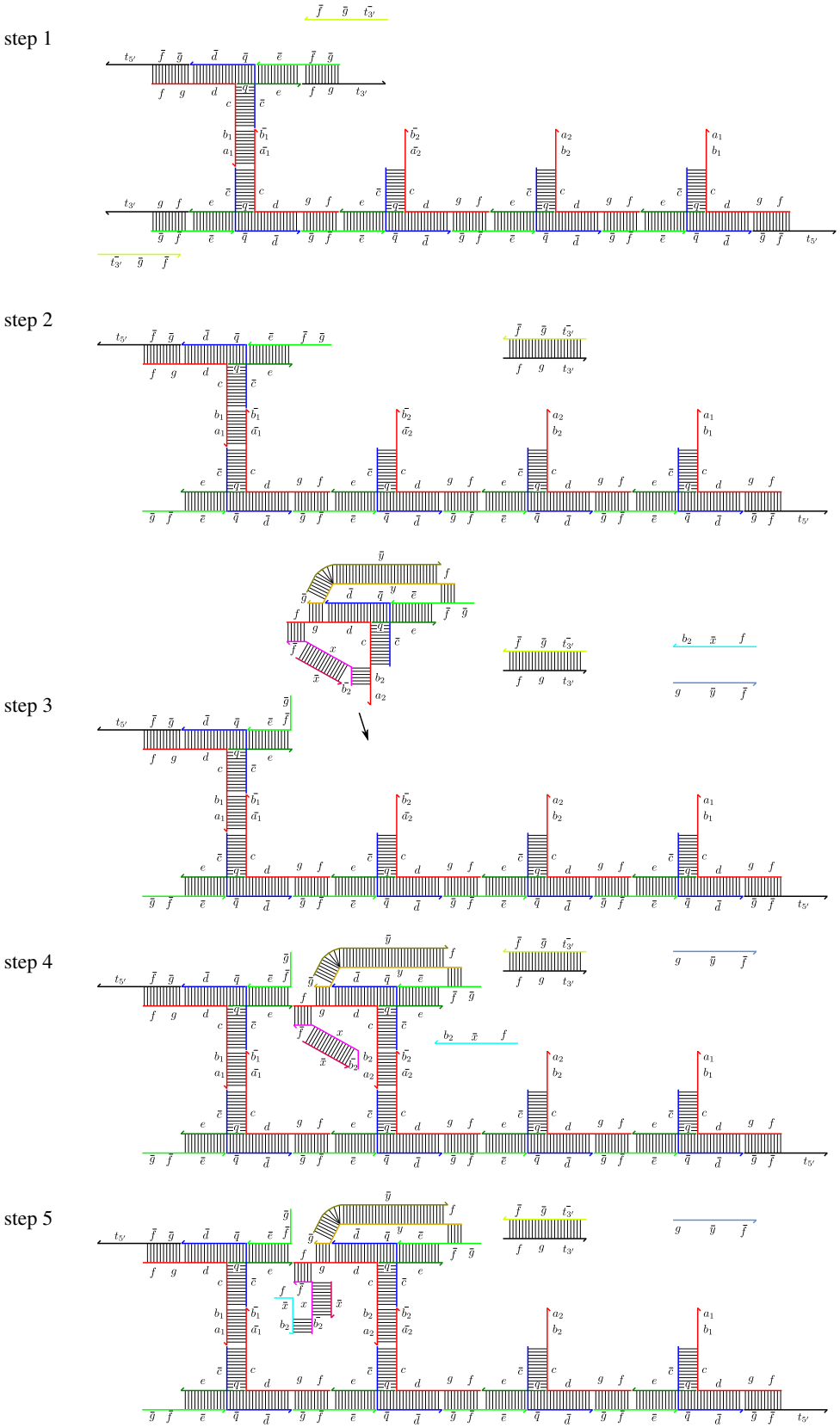



Figure 7. mDNA polymerization. Step 1: a primed template. Step 2: decapping of the primer. Step 3, 4: base binding. Step 5, 6: 5′ protection is stripped away.

The 3′ pad of the tile at the 3′ end of the primer is capped by $fgt_{3'}$ (step 1). The initiator strand $\bar{t}_{3'}\,\bar{g}\bar{f}$ binds to the capper strand, excising it, and thus activates the 3′ pad $\bar{g}\bar{f}$ (step 2). Now, the protected meta-nucleotide comes in (step 3) and binds to the complementary base sequence $\bar{a}_2\,\bar{b}_2$, displacing the strand $\bar{b}_2\,x\bar{f}$ (step 4). The toehold $\bar{b}_2$ is exposed on $\bar{b}_2\,x\bar{f}$ allowing $f\bar{x}b_2$ to bind (step 5). This strand displaces $\bar{b}_2\,x\bar{f}$ exposing $f$ on the 5′ end of the incoming meta-nucleotide (step 6). Now, the 5′ end of the incoming meta-nucleotide binds to the 3′ end of the

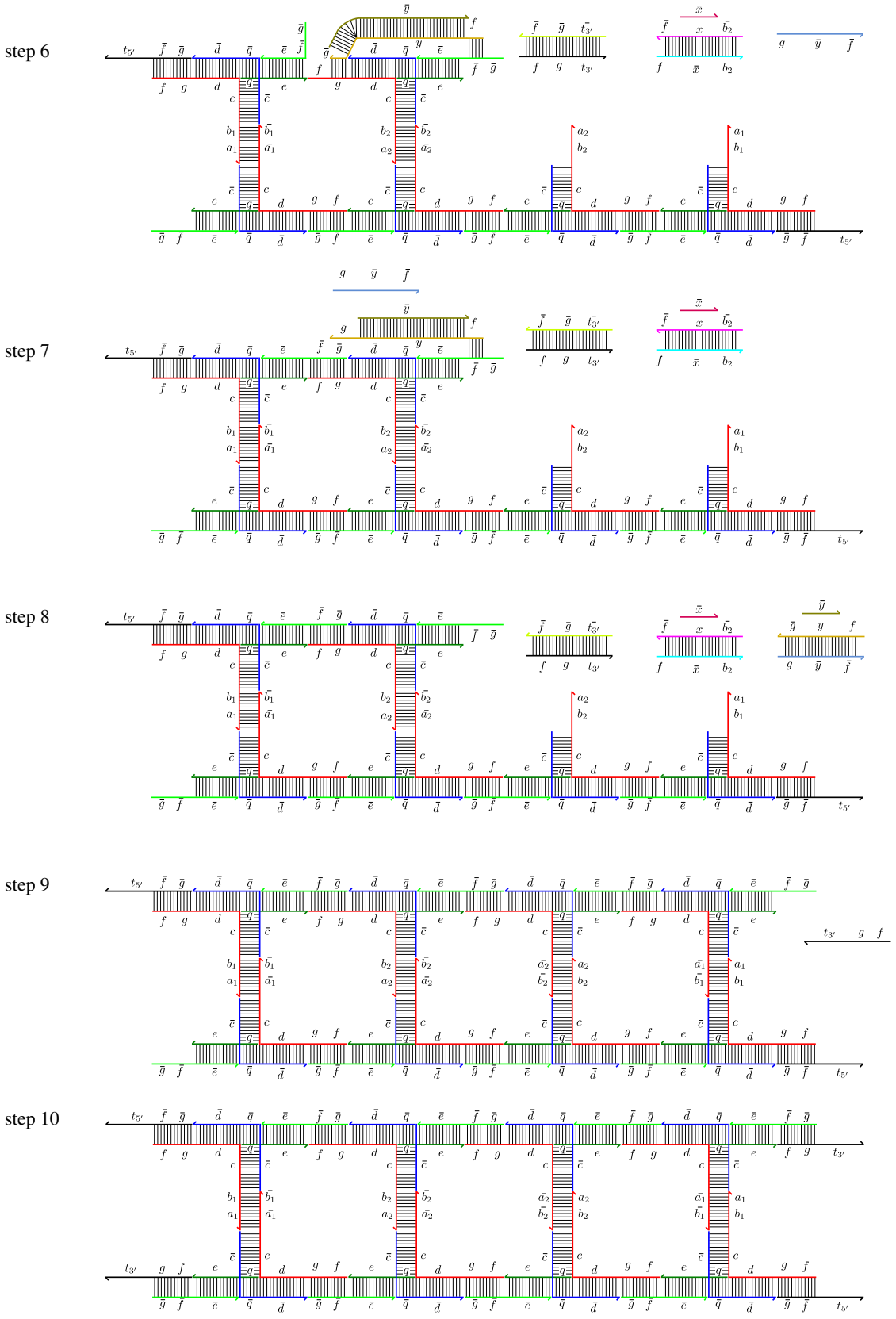



Figure 8. mDNA polymerization. Step 5, 6: 5′ protection is stripped away. Step 7: meta-backbone bond forms and displaces subsequence $\bar{g}$. Step 8: 3′ protection ripped away. Step 9: state after two more meta-nucleotide additions. Step 10: capping the 3′ end.

previous nucleotide, displacing the subsequence $\bar{g}$ of strand $fy\bar{g}$ (step 7). The strand $fy\bar{g}$ is stripped away by its complement through the now exposed toehold $\bar{g}$, activating the 3′ end ($\bar{g}\bar{f}$) of the incoming meta-nucleotide (step 8). This brings the 3′ end of the growing *ss*mDNA back to the same state as in step 2 and thus the reaction can repeat through further meta-nucleotide additions till the end of the template strand is reached (step 9). We introduce

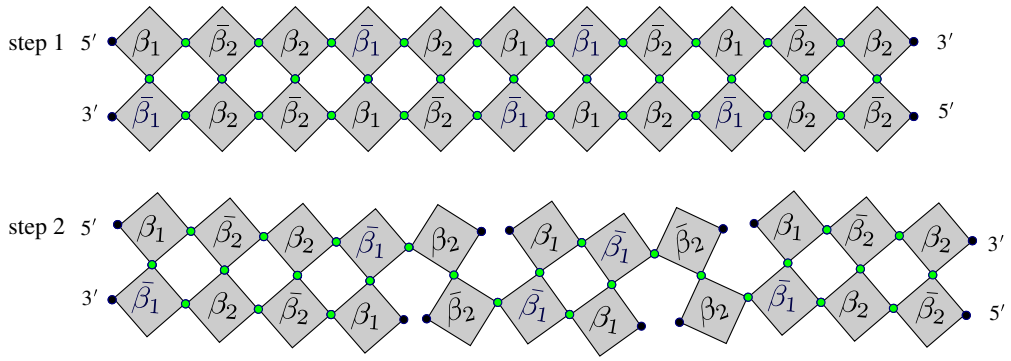



Figure 9. mDNA restriction at site $\beta_1$: we prefix a recognition sequence to the 5′ pad of $\beta_1$. This sequence is recognized by cleaving strands that break the meta-backbone bonds.

the 3′ capper sequence $fgt_{3'}$ after the polymerization reaction is allowed to run to completion. The capper sequence binds to the 3′ pad of the last meta-nucleotide added. This terminates the polymerization reaction. Note that in a solution with an ensemble of these nanostructures, we must wait for each copy of the reaction to proceed to step 9 before introducing the capper sequence or we must programme a stopper sequence, analogous to *stop codons* in translation. DNA polymerases like Φ29 possess exceptional strand displacing capability that aids in isothermal amplification of DNA. However, our mDNA polymerization protocol lacks this ability. Thus, if in the course of extending the primer an already bound *ss*mDNA is encountered then the mDNA polymerization process halts. A polymerization protocol naturally suggests an exponential amplification protocol for mDNA akin to the polymerase chain reaction (PCR). Indeed, the same abstract protocol for PCR can be implemented in mDNA systems, with mDNA polymerization replacing DNA polymerization. In §3.6, we show how mDNA PCR can be used in the synthesis of any specified mDNA sequence. In §3.8, we show how to achieve isothermal amplification of mDNA, in contrast to the temperature cycling required in mDNA PCR.

### 3.5. Meta-DNA restriction cuts

We can achieve site-specific mDNA restriction in both *ss*mDNA and *ds*mDNA by slightly modifying the internal structure of a nucleotide. This modification does not significantly affect the secondary structure of the mDNA and is compatible with all the other mDNA protocols described in this paper. We prefix a sequence $h_i$ to the 5′ pad of the meta-nucleotide with base pad $\beta_i$. This is the sequence that will be recognized by cleaving strands that break the meta-backbone bonds. We illustrate restriction for *ds*mDNA. The protocol for *ss*mDNA is very similar and can be thought of as a special case of the mDNA restriction of *ds*mDNA. The abstract activatable tile model of mDNA restriction is illustrated in figure 9. The 5′ end of the meta-nucleotide $\beta_1$ is recognized, cut and sent to the capped state. See the electronic supplementary material for more details.

### 3.6. Meta-DNA synthesis

We have so far assumed the existence of specified mDNA sequences without discussing how such sequences may be synthesized. Since the meta-backbone bonds are identical along the length of an mDNA, the problem of synthesizing a specific mDNA sequence is far from trivial. This problem is not specific to mDNA and also occurs in synthetic DNA synthesis. One of the key technological breakthroughs that have aided the progress of molecular biology are commercial protocols for fast and cheap chemical synthesis of short nucleic acid sequences. The commonly used scheme for synthesizing nucleic acids involves a solid-phase synthesis that sequentially couples nucleotides in the order specified by the required sequence of the nucleic acid. The coupling is interspersed with protection–deprotection of the growing chain to prevent incorrect building blocks from being incorporated. The purity of the final product depends on the efficiency of the protection–deprotection scheme, while the yield depends on the efficiency of coupling. Long (greater than 200 bases) nucleic acids cannot be synthesized with high purity and yield because both these quantities drop off geometrically and hence it is critical to have efficient protection–deprotection and coupling. A solid-phase protection–deprotection scheme could be adapted to synthesizing mDNA, but would be beset by the same type of errors and be subject to the same limitations as chemical DNA synthesis. Instead, we propose a novel protocol for synthesizing an mDNA sequence by performing mDNA PCR on a special template strand, followed by mDNA restriction.

Consider a special mDNA template sequence reverse complementary to the mDNA sequence being synthesized (figure 10, step 1). It has two distinguishing features: its meta-backbone linkages are all distinct sequences (illustrated as numbered white circles in figure 10) and it has a special end meta-nucleotide at its 5′ and 3′ ends. The distinct meta-backbone linkages are used to determine the order of the meta-nucleotides and thus the sequence of the mDNA being synthesized while the end meta-nucleotides act as sites where the primers for a PCR reaction will bind. The synthesis of the special mDNA template can be achieved using a simple one-pot annealing strategy commonly used to build DNA nanostructures. If the special mDNA template is very long and not enough distinct sequences can be found for the meta-backbone linkages, a hierarchical assembly technique (see [22]) can be used, allowing the reuse of linkage sequences while preserving the sequence of the mDNA template.

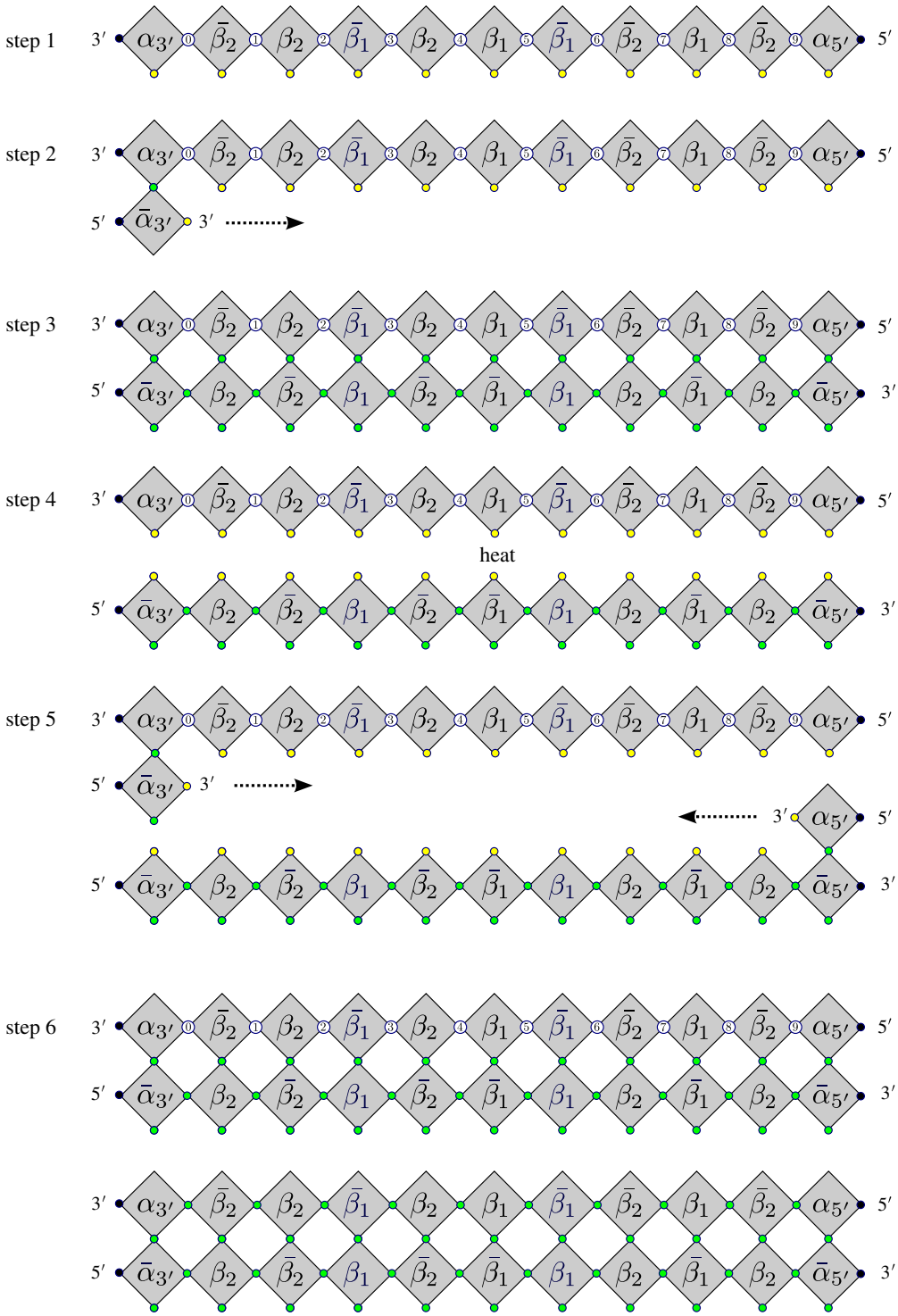



Figure 10. mDNA synthesis: the meta-nucleotide $\bar{\alpha}_{3'}$ sits on the special template strand and acts as a primer for mDNA polymerization (steps 2 and 3). Once the polymerization is complete, we raise the temperature of the system above the melting temperature of the *ds*mDNA formed and denature the constituent *ss*mDNA. Each of the two *ss*mDNA strands can now act as a template for polymerization through the primer meta-nucleotides producing two further copies of *ss*mDNA and so on.

The complete protocol is illustrated in figure 10 and uses the mDNA polymerization protocol described in §3.4. The meta-nucleotide $\bar{\alpha}_{3'}$ sits on the special template strand and acts as a primer for mDNA polymerization (steps 2 and 3). Once the polymerization is complete, we raise the temperature of the system above the melting temperature of the *ds*mDNA formed (but not hot enough to destroy meta-backbone linkages) and denature the constituent *ss*mDNA. Each of the two *ss*mDNA strands can now act as a template for polymerization through the primer meta-nucleotides $\bar{\alpha}_{3'}$ and $\alpha_{5'}$, producing two further copies of *ss*mDNA and so on. Once the mDNA PCR step is completed, we use mDNA restriction to snip out the special end meta-nucleotides $(\bar{\alpha}_{3'}, \bar{\alpha}_{5'}, \alpha_{3'}, \alpha_{5'})$. This exponential chain reaction produces both the *ss*mDNA and its reverse

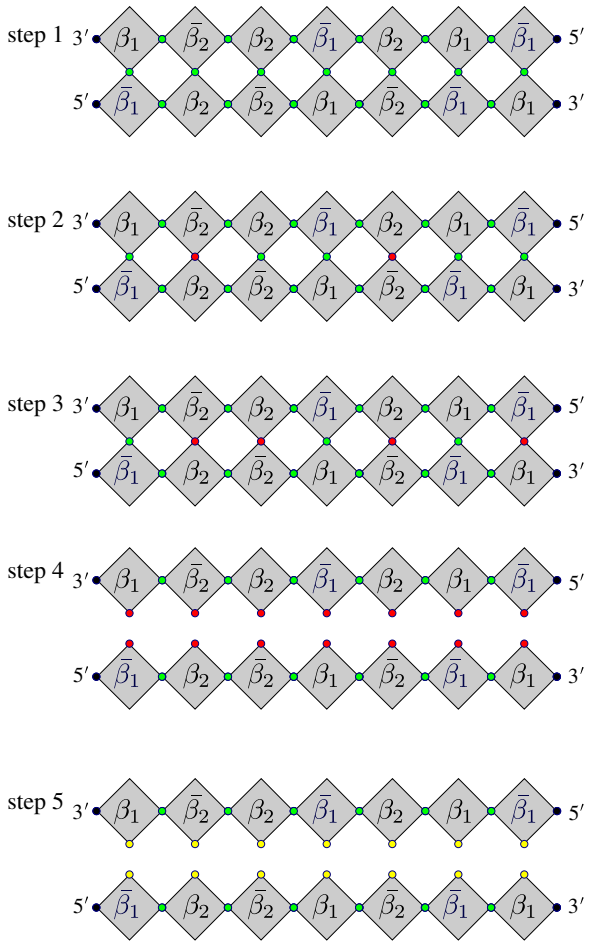



Figure 11. mDNA helicase denaturation: initially all the meta-base pads are in the bound state (step 1). Helicase mDNA activity breaks the bonds and sends the base pads to the inactive state (steps 2,3). When all the meta-base bonds are broken, the meta-strands float apart (step 4). We can reactivate the meta-base pads by transitioning them to the active unbound state (step 5).

complement. If we wish to synthesize only one of the two *ss*mDNA sequences, we simply retain the corresponding primer and leave the other primer out. However, this results in a linear rather than an exponential amplification protocol.

### 3.7. Helicase meta-DNA denaturation reaction

We have previously described mDNA denaturation using temperature. The same result can be achieved isothermally, which we call helicase mDNA denaturation. Consider the abstract view of helicase mDNA denaturation given in figure 11. Denaturation is the breaking of meta-base bonds of a *ds*mDNA. When all the meta-base bonds are broken, the meta-strands float apart. Initially all the meta-base pads are in the bound state (step 1). Helicase mDNA activity breaks the bonds and sends the base pads to the inactive state. This helicase mDNA reaction does not necessarily act contiguously. Some meta-base bonds are broken before others (steps 2 and 3). When all the meta-base pads are broken the meta-strands float apart (step 4). We can reactivate the meta-base pads by transitioning them to the active unbound state (step 5) at which point the *ss*mDNA can recombine to form a *ds*mDNA. See the electronic supplementary material for more details.

### 3.8. Exponential amplification in meta-DNA using a replicator

Driven by the important role that replication plays in biology, many self-replicating systems have been proposed, starting with Von Neumann. Early self-replicating systems were designed by von Kiedrowski [23] and Tjivikua *et al.* [24]. For a review of various artificial replicators, see [25–27]. A DNA-based artificial replicator was proposed by Zhang & Yurke [28]. Schulman & Winfree [29] study growth and evolution of simple crystals using DNA. Smith *et al.* [30] have independently developed abstractions for self-replication systems that can be thought of as tile-based and also rely on the idea of activation.

The major difference between this prior work on self-replicating systems and our approach is that mDNA allows for a wide variety of key operations far beyond merely the operation of replication, enabling a much more extensive and complex set of synthetic biochemical systems. This also introduces complexity in our replicator protocol, beyond what is necessary if the only function being designed is replication.

Our protocol for mDNA polymerization lacked meta-strand displacement capabilities and hence could not be used for isothermal PCR-like amplification. In this section, we describe a method to get isothermal exponential amplification using a replicator mechanism. The protocol is similar to mDNA polymerization, it involves linear contiguous extension of a primer by addition of meta-nucleotides, with the newly polymerized *ss*mDNA having the complementary sequence to that of the template. The key difference in the two protocols is a mechanism to isothermally dissociate the newly synthesized *ss*mDNA from the template. We achieve this using a new idea, not present in any previous protocol described in this work, of active DNA sequences sequestered in hairpins that are released by strand displacement reactions. The release of the 'hidden' DNA sequence inside the hairpin structure can be thought of as an activation step, setting off another strand displacement reaction. This idea of sequestering active sequences within hairpins has been demonstrated previously [5,31].

For the purposes of this protocol, we define a new state of the base pad in the activatable model of mDNA, which we call *semi active unbound* (shaded purple in figure 12*c*). If two base pads are in the semi active unbound state, they cannot bind to each other. However, a semi active unbound base pad can bind with an active unbound base pad, provided their sequences are complementary. There are also modifications (figure 12*a*,*b*) to how the other pad states are implemented, however the properties of the state do not change. Note the introduction of the hairpin structure $g_2\bar{p}\bar{a}lg_2$ at the 3′ end of a meta-nucleotide (figure 12*a*) which will be used to cleave the growing strand from the template (figure 12*c*), sending the base pad to the semi active unbound state.

Figure 13 illustrates the replicator in the abstract activatable tile model of mDNA. The top strand is the template and the bottom strand grows from a

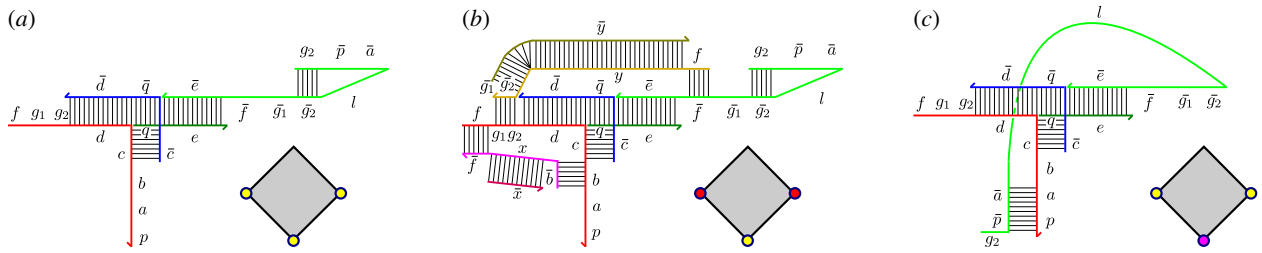



Figure 12. New states in mDNA replication: new states are required to represent the replicator protocol: (*a*) unprotected meta-nucleotide: the hairpin structure $g_2\bar{p}\bar{a}l\,\bar{g}_2$ is used to cleave the growing strand from the template, (*b*) protected meta-nucleotide to prevent spontaneous aggregation and (*c*) *semi active unbound* state (shaded purple). If two base pads are in the *semi active unbound* state, they cannot bind to each other. However, a *semi active unbound* base pad can bind with an *active unbound* base pad, provided their sequences are complementary.

primer. For initiating and terminating replication, we require the use of a pair of special tiles with complementary base pads, labelled $\beta_{3'}$ and $\bar{\beta}_{3'}$ in figure 13. These are capped at one end and occur at the terminal ends of the template and hence in each replicated mDNA. $\bar{\beta}_{3'}$ acts as the primer, initiating the replication. Consider the replication process after a few meta-nucleotides have been added (step 1). The growing strand is attached via a single base pad (at the $3'$ terminal meta-nucleotide) to the template. The base pads of the rest of the meta-nucleotides are in the semi active unbound state and hence cannot bind with each other. A new meta-nucleotide comes in (step 1) with its base pad in the active unbound state and binds to the template (step 2). This activates its $5'$ end (step 3) and allows the $3'$ end of the growing strand to bind (step 4). This activates the $3'$ end of the meta-nucleotide added and also cleaves the meta-base pad binding between the template and the previously attached meta-nucleotide of the growing strand, sending the cleaved meta-base pads to the semi active unbound state (step 5). The process (steps 1–5) repeats till the last meta-nucleotide $\beta_{3'}$ attaches. At this point, we want the two strands to separate. We achieve this by making the meta-base pad bond between $\beta_{3'}$ and $\bar{\beta}_{3'}$ relatively weak, allowing the strands to separate spontaneously owing to breathing of the base pad bond. This is implemented by choosing a very short length sequence for the base pads of $\beta_{3'}$ and $\bar{\beta}_{3'}$. Both the template and the newly synthesized *ss*mDNA can now act as templates for further replication and hence we can achieve exponential amplification. See the electronic supplementary material for more details. Note that we can easily introduce mutations in the replicator mechanism by designing DNA sequences for certain base pads that do not have any exact complementary base pads but rather several partial complements. This allows us to probabilistically evolve a diverse sequence population of *ss*mDNA. See §4.1 for a brief discussion.

## 4. DISCUSSION AND FUTURE DIRECTIONS

### 4.1. Directed evolution via mutations in meta-DNA replication

A major goal of synthetic biology is the construction of evolving replicating systems. In §3.8, we noted that we can easily introduce mutations in the replicator mechanism by designing DNA sequences for certain base pads that do not have any exact complementary base pads but rather several partial complements. This allows us to probabilistically evolve a diverse sequence population of *ss*mDNA. Incorporating environmental selection pressure in mDNA to evolve functional biosystems is a major open challenge. For example, we might evolve mDNA strands with higher melting temperatures by a sequence of replication reactions (with designed mutations) at increasing temperatures. A higher fraction of GC-rich meta-bases would get incorporated at higher temperatures and as a result the newly replicated mDNA strands would have higher melting temperatures. As temperature rises, the fidelity of meta-base binding improves leading to lower mutation rates allowing the exponential amplification of the GC-rich-evolved mDNA strands. Temperature changes can also be used to tune mutation rates in conjunction with other selection pressures. Taking inspiration from genetic regulatory networks, we also envision mDNA replicatory systems modulated by DNA computations. For instance, an mDNA sequence could serve as input for some DNA computation whose output disrupts replication by binding to the active regions of the mDNA template.

### 4.2. Experimental implementation of meta-DNA systems

The past few years have seen experimental implementations of increasingly complex DNA systems based purely on hybridization and strand displacement [7,32–34]. These systems show that intricate, large-scale DNA systems like those proposed in this paper may be successfully implemented. However, significant practical barriers need to be overcome before this happens. Most of our protocols rely on a supply of pure protected meta-nucleotides and synthesizing pure meta-nucleotides of various types in high concentration while avoiding spurious polymers is an experimental challenge. Standard annealing followed by gel electrophoresis purification techniques can be used to address this challenge, however there is scope for more algorithmic solutions. Another challenge is the characterization of specific *ss*mDNA and *ds*mDNA sequence synthesis. A standard approach is to associate distinct labels with each type of meta-nucleotide. The labels are distinguishable when viewed

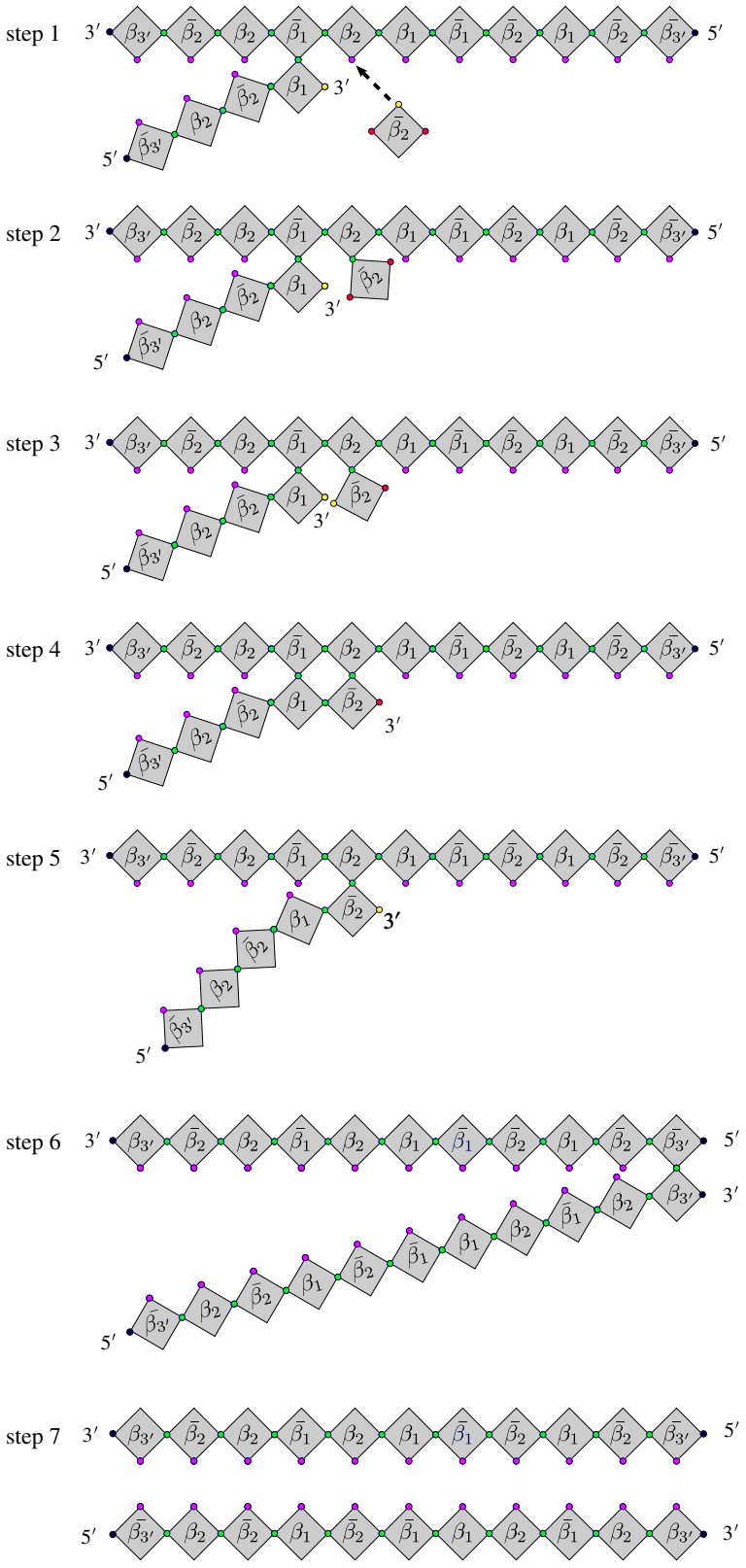



Figure 13. Replicator: a new meta-nucleotide comes in (step 1) with its base pad in the active unbound state and binds to the template (step 2). This activates its 5′ end (step 3) and allows the 3′ end of the growing strand to bind (step 4). This activates the 3′ end of the meta-nucleotide added and also cleaves the meta-base pad binding between the template and the previously attached meta-nucleotide of the growing strand, sending the cleaved meta-base pads to the semi active unbound state (step 5).

under an atomic force microscope (AFM). Standard labels include hairpins of tunable size, biotin–streptavidin pairs and small DNA-binding proteins. Overcoming these challenges would allow us to achieve laboratory implementations of *ss*mDNA, *ds*mDNA synthesis and mDNA polymerization.

Our designs are modular and tend to use a small set of reactions (hybridization, breathing and strand

displacement) executed by few types of nanostructures (T-junctions and hairpins). Hence, a first step towards experimental feasibility is the development of a modular model of mDNA protocols based on strand level kinetic simulation. We hope to use data from such simulations to enhance and fine tune our protocols and also perform *in vitro* experiments to validate them.

A key open question is the rate of mDNA reactions compared with the corresponding enzymatic reactions of DNA. mDNA protocols can be viewed as a mix of diffusion based and local strand displacement (the invading strand is locally tethered and hence is available in a high local concentration) events. Strand displacement events are typically slower than enzymatic reactions, however local strand displacement events happen much faster. Such localized systems are being studied with the help of models [35] and experiments [36] and hopefully will be much better understood in the near future, which will enable us to analytically predict rates of mDNA reactions.

## 5. CONCLUSION

In this work, we have outlined a synthetic biochemical system made purely from DNA strands that simulates the behaviour of various protein enzymes acting on DNA. The protocols described for the manipulation of mDNA rely solely on predictable Watson–Crick base pairing interactions. Thus, we suggest that by logically programming simple chemical interactions one can simulate complex biochemical behaviour. The first trickle of DNA and RNA devices has been deployed inside cells [37–39] to perform relatively simple tasks. We foresee more complex DNA devices and systems joining the trickle and leading to a flood of capabilities. We believe mDNA systems are a powerful paradigm for achieving novel functionality in a programmable manner and may play a role *in vivo* or within artificial cell-like systems.

We thank our anonymous reviewers for useful comments on the organization of the paper. This work was supported by NSF EMT grants CCF-0829797 and CCF-0829798.

## REFERENCES

1 Luisi, P. L. 2006 *The emergence of life: from chemical origins to synthetic biology*. Cambridge, UK: Cambridge University Press.
2 Bedau, M., McCaskill, J., Packard, N., Rasmussen, S., Adami, C., Green, D., Ikegami, T., Kaneko, K. & Ray, T. 2000 Open problems in artificial life. *Artif. Life* **6**, 363–376. (doi:10.1162/106454600300103683)
3 Packard, N. & Bedau, M. 2003 Artificial life. *Encycl. Cogn. Sci.* **1**, 209–215.
4 Zhang, D. & Seelig, G. 2011 Dynamic DNA nanotechnology using strand displacement reactions. *Nat. Chem.* **3**, 103–113. (doi:10.1038/nchem.957)
5 Yin, P., Choi, H., Calvert, C. & Pierce, N. 2008 Programming biomolecular self-assembly pathways. *Nature* **451**, 318–322. (doi:10.1038/nature06451)
6 Zhang, D., Turberfield, A., Yurke, B. & Winfree, E. 2007 Engineering entropy-driven reactions and networks catalyzed by DNA. *Science* **318**, 1121–1125. (doi:10.1126/science.1148532)
7 Qian, L. & Winfree, E. 2011 Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **332**, 1196–1201. (doi:10.1126/science.1200520)
8 Qian, L. & Winfree, E. 2011 A simple DNA gate motif for synthesizing large-scale circuits. *J. R. Soc. Interface* **8**, 1281–1297. (doi:10.1098/rsif.2010.0729)
9 Qian, L., Winfree, E. & Bruck, J. 2011 Neural network computation with DNA strand displacement cascades. *Nature* **475**, 368–372. (doi:10.1038/nature10262)
10 Plasson, R., Brandenburg, A., Jullien, L. & Bersini, H. 2011 Autocatalyses. *J. Phys. Chem. A* **115**, 8073–8085. (doi:10.1021/jp110079p)
11 Hamada, S. & Murata, S. 2009 Substrate-assisted assembly of interconnected single-duplex DNA nanostructures. *Angew. Chem. Int. Ed.* **48**, 6820–6823. (doi:10.1002/anie.200902662)
12 Rothemund, P. & Winfree, E. 2000 The program-size complexity of self-assembled squares. In *STOC '00 Proc. 32nd Annual ACM Symp. on Theory of Computing*, pp. 459–468. New York, NY: ACM. (doi:10.1145/335305.335358)
13 Soloveichik, D. & Winfree, E. 2007 Complexity of self-assembled shapes. *SIAM J. Comput.* **36**, 1544–1569. (doi:10.1137/S0097539704446712)
14 Adleman, L., Cheng, Q., Goel, A. & Huang, M.-D. 2001 Running time and program size for self-assembled squares. In *STOC '01 Proc. 33rd Annual ACM Symp. on Theory of Computing*. pp. 740–748. New York, NY: ACM. (doi:10.1145/380752.380881)
15 Aggarwal, G., Goldwasser, M., Kao, M.-Y. & Schweller, R. 2004 Complexities for generalized models of self-assembly. In *Proc. 15th Annual ACM-SIAM Symp. on Discrete Algorithms*, January, New Orleans, Louisiana. pp. 880–889. (doi:10.1137/S0097539704445202)
16 Kao, M.-Y. & Schweller, R. 2008 Randomized self-assembly for approximate shapes. In *35th Int. Colloq., ICALP 2008*, Reykjavik, Iceland, 7-11 July, Proceedings, Part I: Automata, Languages and Programming. *Lecture Notes in Computer Science*, **5125**, 370–384. (doi:10.1007/978-3-540-70575-8_31)
17 Cook, M., Fu, Y. & Schweller, R. 2011 Temperature 1 self-assembly: deterministic assembly in 3D and probabilistic assembly in 2D. In *Proc. 22nd Annual/ACM-SIAM Symp. on Discrete Algorithms,* San Francisco, CA, 23–25 January, pp. 459–468.
18 Doty, D. 2009 Randomized self-assembly for exact shapes. *Found. Comput. Sci.* (http://arxiv.org/abs/0901.1849)
19 Doty, D., Patitz, M. J. & Summers, S. M. 2009 Limitations of self-assembly at temperature 1. *DNA Computing* **5877**, 35–44. (http://arxiv.org/abs/0903.1857)
20 Chandran, H., Gopalkrishnan, N. & Reif, J. 2009 The tile complexity of linear assemblies. *Int. Colloq. Autom. Lang. Program.* **5555**, 235–253. (doi:10.1007/978-3-642-02927-1_21)
21 Majumder, U., LaBean, T. & Reif, J. 2007 Activatable tiles: compact, robust programmable assembly and other applications. In *DNA13 '07 Proc. 13th Int. Conf. on DNA Computing*, pp. 15–25. Berlin: Springer-Verlag.
22 Pistol, C. & Dwyer, C. 2007 Scalable, low-cost, hierarchical assembly of programmable DNA nanostructures. *Nanotechnology* **18**, 125 305–125 309. (doi:10.1088/0957-4484/18/12/125305)
23 von Kiedrowski, G. 1986 A self-replicating hexadeoxynucleotide. *Angew. Chem. Int. Ed.* **25**, 932–935. (doi:10.1002/anie.198609322)
24 Tjivikua, T., Ballester, P. & Rebek, J. 1990 A self-replicating system. *J. Am. Chem. Soc.* **112**, 1249–1250. (doi:10.1021/ja00159a057)

25 Vidonne, A. & Philp, D. 2009 Making molecules make themselves: the chemistry of artificial replicators. *Eur. J. Org. Chem.* **5**, 593–610. (doi:10.1002/ejoc.200800827)
26 von Kiedrowski, G. & Patzke, V. 2007 Self replicating systems. *ARKIVOK* 293–310.
27 Dadon, Z., Wagner, N. & Ashkenasy, G. 2008 The road to non-enzymatic molecular networks. *Angew. Chem. Int. Ed.* **47**, 6128–6136. (doi:10.1002/anie.200702552)
28 Zhang, D. Y. & Yurke, B. 2006 A DNA superstructure-based replicator without product inhibition. *Nat. Comput.* **5**, 183–202. (doi:10.1007/s11047-005-4465-x)
29 Schulman, R. & Winfree, E. 2008 How crystals that sense and respond to their environments could evolve. *Nat. Comput.* **7**, 219–237. (doi:10.1007/s11047-007-9046-8)
30 Smith, A., Turney, P. & Ewaschuk, R. 2002 JohnnyVon: self-replicating automata in continuous two-dimensional space. *Comput. Res. Repository.*
31 Dirks, R. & Pierce, N. 2004 Triggered amplification by hybridization chain reaction. *Proc. Natl Acad. Sci. USA* **101**, 15 275–15 278. (doi:10.1073/pnas.0407024101)
32 Rothemund, P. 2006 Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302. (doi:10.1038/nature04586)
33 Douglas, S., Dietz, H., Liedl, T., Hogberg, B., Graf, F. & Shih, W. 2009 Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459**, 414–418. (doi:10.1038/nature08016)
34 Gu, H., Chao, J., Xiao, S.-J. & Seeman, N. 2010 A proximity-based programmable DNA nanoscale assembly line. *Nature* **465**, 202–205. (doi:10.1038/nature09026)
35 Chandran, H., Gopalkrishnan, N., Phillips, A. & Reif, J. 2011 Localized hybridization circuits. *DNA* **6937**, 64–83. (doi:10.1007/978-3-642-23638-9_8)
36 Genot, A., Zhang, D., Bath, J. & Turberfield, A. 2011 Remote toehold: a mechanism for flexible control of DNA hybridization kinetics. *J. Am. Chem. Soc.* **133**, 2177–2182. (doi:10.1021/ja1073239)
37 Delebecque, C., Lindner, A., Silver, P. & Aldaye, F. 2011 Organization of intracellular reactions with rationally designed RNA assemblies. *Science* **22**, 470–474. (doi:10.1126/science.1206938)
38 Modi, S., Swetha Goswami, D., Gupta, G., Mayor, S. & Krishnan, Y. 2009 A DNA nanomachine that maps spatial and temporal pH changes inside living cells. *Nat. Nanotechnol.* **4**, 325–330. (doi:10.1038/nnano.2009.83)
39 Venkataraman, S., Dirks, R., Uedab, C. & Pierce, N. 2010 Selective cell death mediated by small conditional RNAs. *Proc. Natl Acad. Sci. USA* **107**, 16 777–16 782. (doi:10.1073/pnas.1006377107)

# 6 Supplementary Information

## 6.1 Internals of mDNA Restriction

In figure 14 the site $h_2$ associated with the meta-base $b_2a_2$ is recognized by the cleaver strand $\bar{g}\bar{f}\bar{h}_2r_2$. There is a single recognition site $h_2$ on each of the strands of the $ds$mDNA (step 1). The cleaver strands bind to the toehold $h_2$ on each strand (step 2) and break the meta-backbone bond by strand displacing the $3'$ pad $\bar{g}\bar{f}$ (step 3). We now introduce the $3'$ capper sequence $fgt_{3'}$ to cap the exposed $3'$ pads (step 4). We prevent interaction between the cleaver strand $\bar{g}\bar{f}\bar{h}_2r_2$ and the $3'$ capper sequence $fgt_{3'}$ by executing step 4 only after we are reasonably certain that step 3 is complete, otherwise these strands may bind to each other and slow down the rate of the restriction. This means that mDNA restriction is not autonomous. The strand $\bar{r}_2h_2fg$ is introduced to strip away the cleaver strand $\bar{g}\bar{f}\bar{h}_2r_2$ from the $5'$ pad by exploiting the toehold $r_2$ (step 5). This exposes the $5'$ pads (step 6) which are then capped by introducing the $5'$ capper sequence $\bar{g}\bar{f}t_{5'}$ (step 7), completing the process of mDNA restriction. Again, we prevent interaction between the strand $\bar{r}_2h_2fg$ and the $5'$ capper sequence $\bar{g}\bar{f}t_{5'}$ by executing step 7 only after we are reasonably certain that step 6 is complete, otherwise these strands may bind to each other and slow down the rate of the restriction. We do not require that these extraneous strands do not interact at all. Rather, it is sufficient that even after interacting among themselves there are a sufficient concentration of them to perform the tasks described in figures 14 and 15.

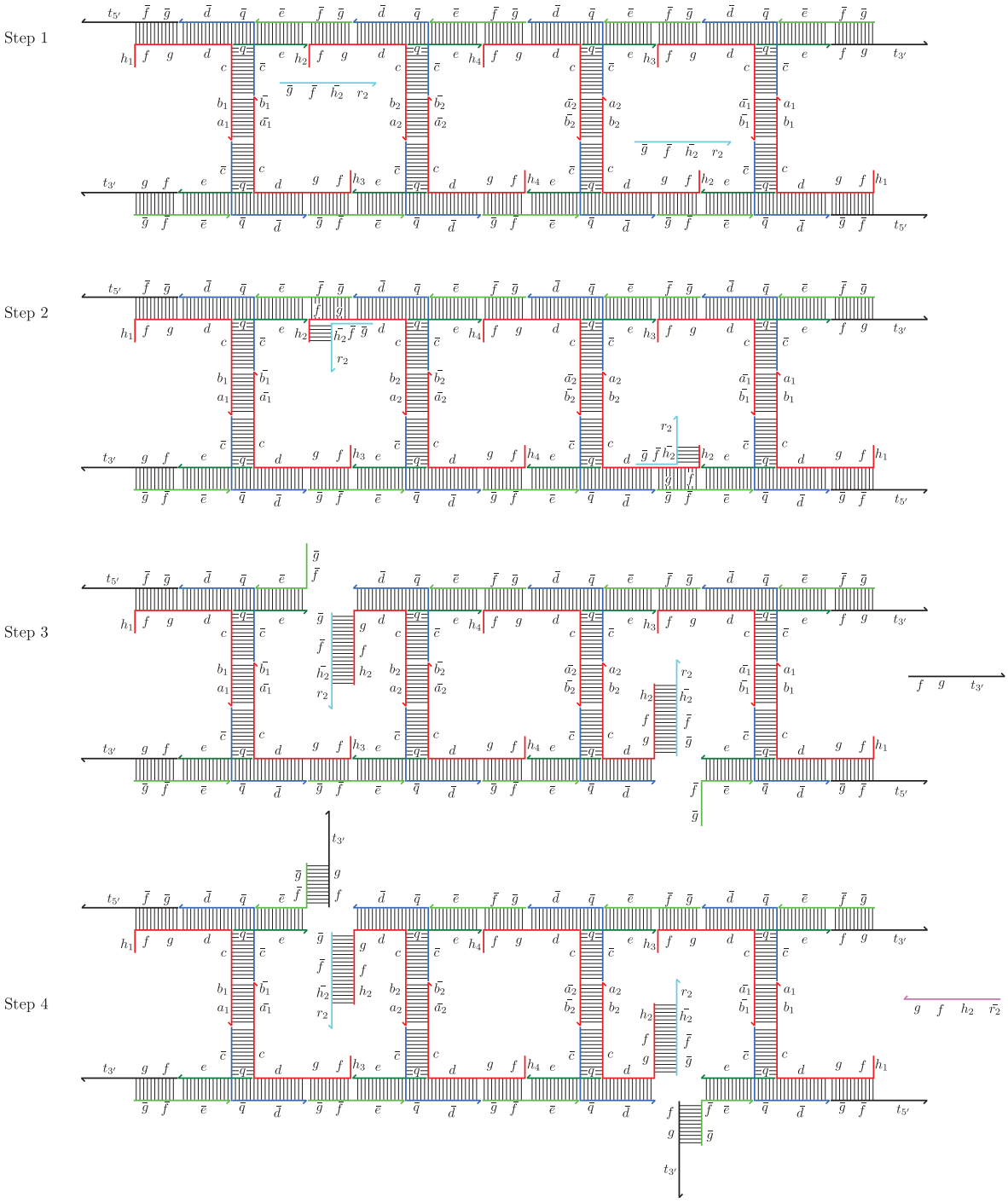



Figure 14: Restriction: Step 1: Recognition site $h_2$ indicates restriction location. Step 2, 3: The cleaver strands bind to the toehold $h_2$ on each strand and break the meta-backbone. Step 4: The $3'$ capping sequence $fgt_{3'}$ is added.

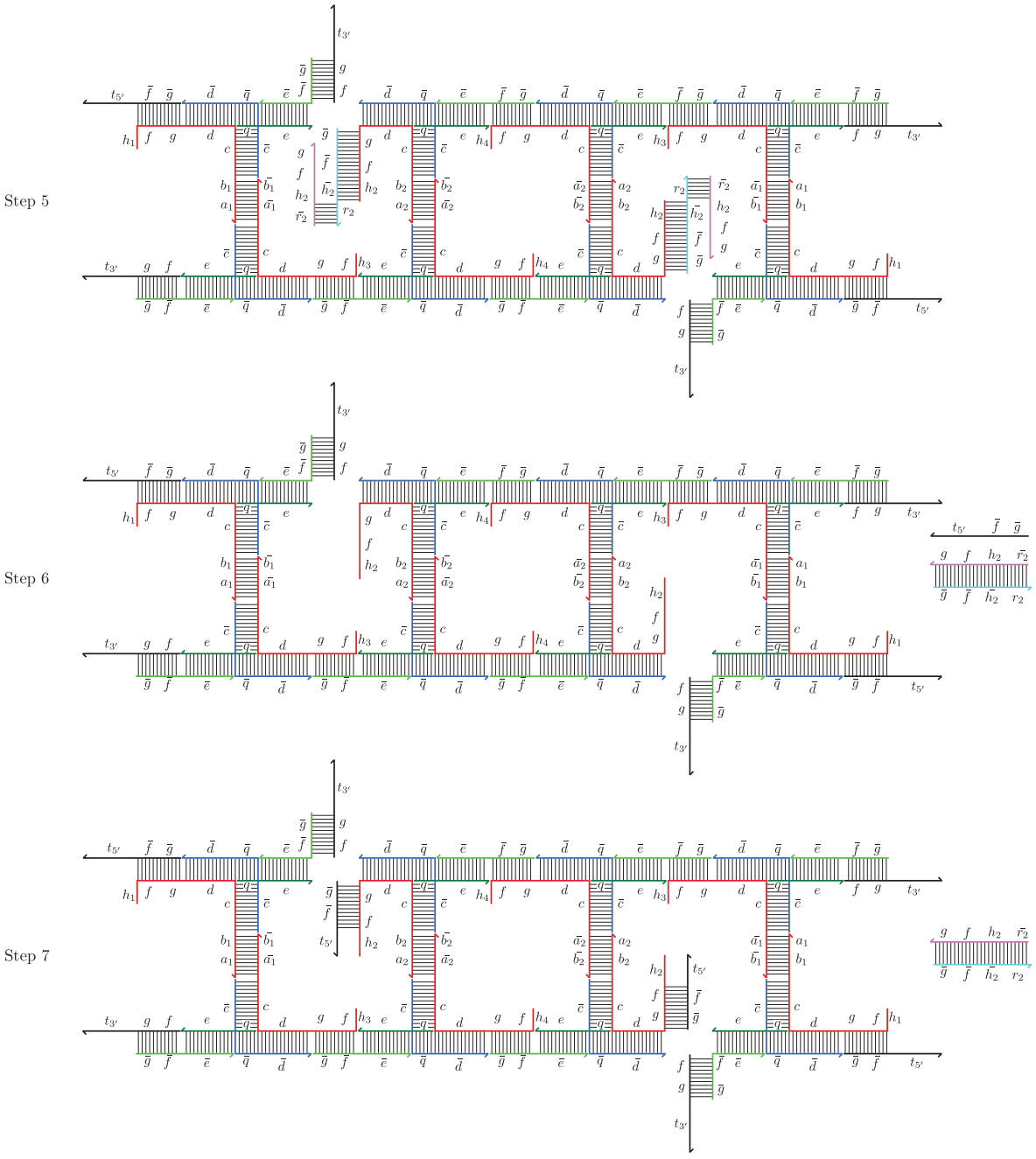
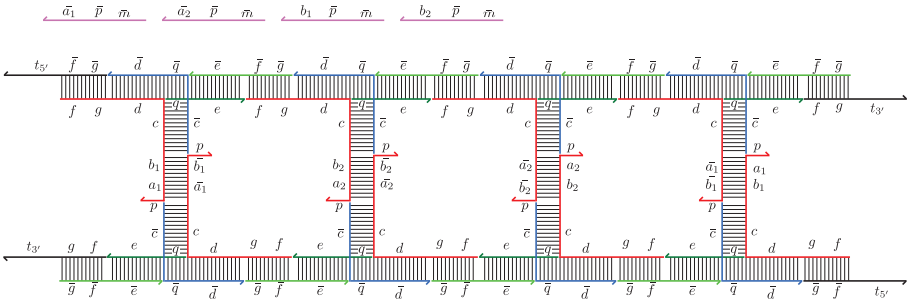



Figure 15: Restriction: Step 5, 6: The cleaver strand is stripped away by the pink strands. Step 7: 5′ ends are capped.

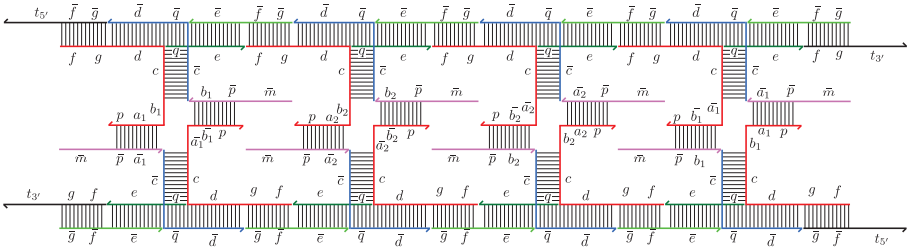
## 6.2 Internals of mDNA Helicase Denaturation

Helicase mDNA activity is performed by a host of strands (colored pink in figure 16), two for each type of meta-base bond that must be broken. In figure 16 there are two types of meta-base bonds and hence we have four strands to perform helicase mDNA activity (step 1). We introduce a slight modification in the internal strand structure by appending a sequence $p$ to the base pad of each meta-nucleotide. This modification does not significantly affect the secondary structure of mDNA and neither does it interfere with any other mDNA protocol described in this paper. This sequence ($p$) will act as a toehold for the strands involved in the helicase mDNA process. A pair of pink strands (for example $\bar{m}\bar{p}b_1$ and $\bar{m}\bar{p}\bar{a}_1$) invade the meta-base pad bond and break them (step 2). Note that half of the meta-base bond is broken by one of these strands and half by the other, ensuring symmetry. Once each meta-base bond is broken the meta-strands drift apart (step 3) and helicase mDNA activity is complete. We can strip off the protecting strand, for example $\bar{m}\bar{p}b_1$, by adding in its complement, say $\bar{b}_1pm$ (step 4). The *ss*mDNA can now recombine into *ds*mDNA. It is possible to repeat this process of mDNA denaturation and mDNA renaturation, however the protocol would not be autonomous.

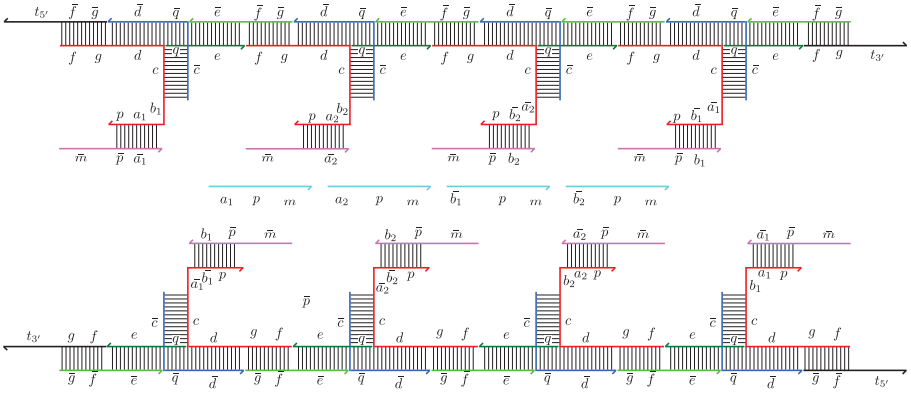
Step 1



Step 2



Step 3



Step 4

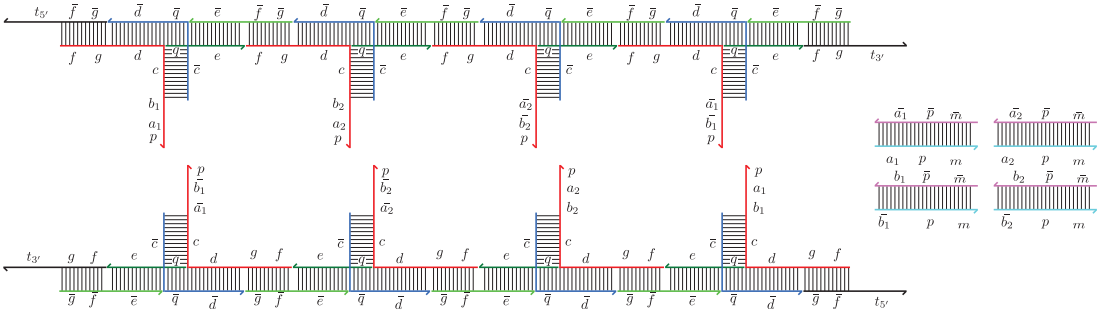


Figure 16: Denaturation: Step 1: The sequence $p$ will act as a toehold for the pink strands. Step 2: The pink strands symmetrically cleave meta-base bonds. Step 3: The strands float away. Step 4: The meta-nucleotides restored to *active unbound* state.

## 6.3 Internals of mDNA Replication

We have a four meta-base mDNA sequence, with the first and last bases being the special terminator bases (step 1). Note that the sequence $g = g_1 g_2$ and $\bar{g} = \bar{g}_2 \bar{g}_1$. A meta-nucleotide with an active unbound base pad (step 1) comes in and binds to the template via the sequence $b_2 a_2$ (step 2), strand displacing out the sequences $\bar{p} b_2$ (part of the light green strand) and $\bar{b}_2$ (part of the pink strand). The subsequence $\bar{p} b_2$ will be used in the separation of this newly incorporated meta-nucleotide from the template when a downstream meta-nucleotide is added in the next cycle. The strand $f \bar{x} b_2$ now strips away $\bar{b}_2 x \bar{f}$ via the toehold $\bar{b}_2$, activating the $5'$ pad of the incoming meta-nucleotide (step 3). This allows the $5'$ pad $f g_1 g_2$ to bind with the $3'$ pad $\bar{g}_2 \bar{g}_1 \bar{f}$, displacing $\bar{g}_2 \bar{g}_1$ and opening up the hairpin structure $\bar{p} \bar{a}_1 l$ (step 4). This allows the strand $f y \bar{g}_2 \bar{g}_1$ to be stripped away through the toehold $\bar{g}_2 \bar{g}_1$ by the strand $g_1 g_2 \bar{y} \bar{f}$, activating the $3'$ end (step 5). The released hairpin from step 4 can now cleave $a_1$ from $\bar{a}_1$ using the toehold $p$, which is half of the base pad binding of the upstream meta-nucleotide, while the other half $\bar{b}_1$ of the base pad binding is cleaved from $b_1$ via the toehold $p$ by the sequence $\bar{p} b_1$ (which is part of the previously incorporated meta-nucleotide) (step 6). Note that cleaving the bond between $b_1$ and $\bar{b}_1$ is actually a reversible process, and could have occurred after step 1. Only when combined with the cleaving of the bond between $a_1$ and $\bar{a}_1$ does it get biased towards the configuration depicted in step 6. This completes the addition of a single meta-nucleotide accompanied by cleaving of the upstream meta-nucleotide. The process repeats till the last meta-nucleotide is added at which point the short base sequence of $\bar{a}_{3'} \bar{b}_{3'}$ allows the *ss*mDNA to separate due to breathing. Both the separate *ss*mDNA are in a configuration that allows them to act as template strands for further replication, allowing exponential amplification of mDNA.

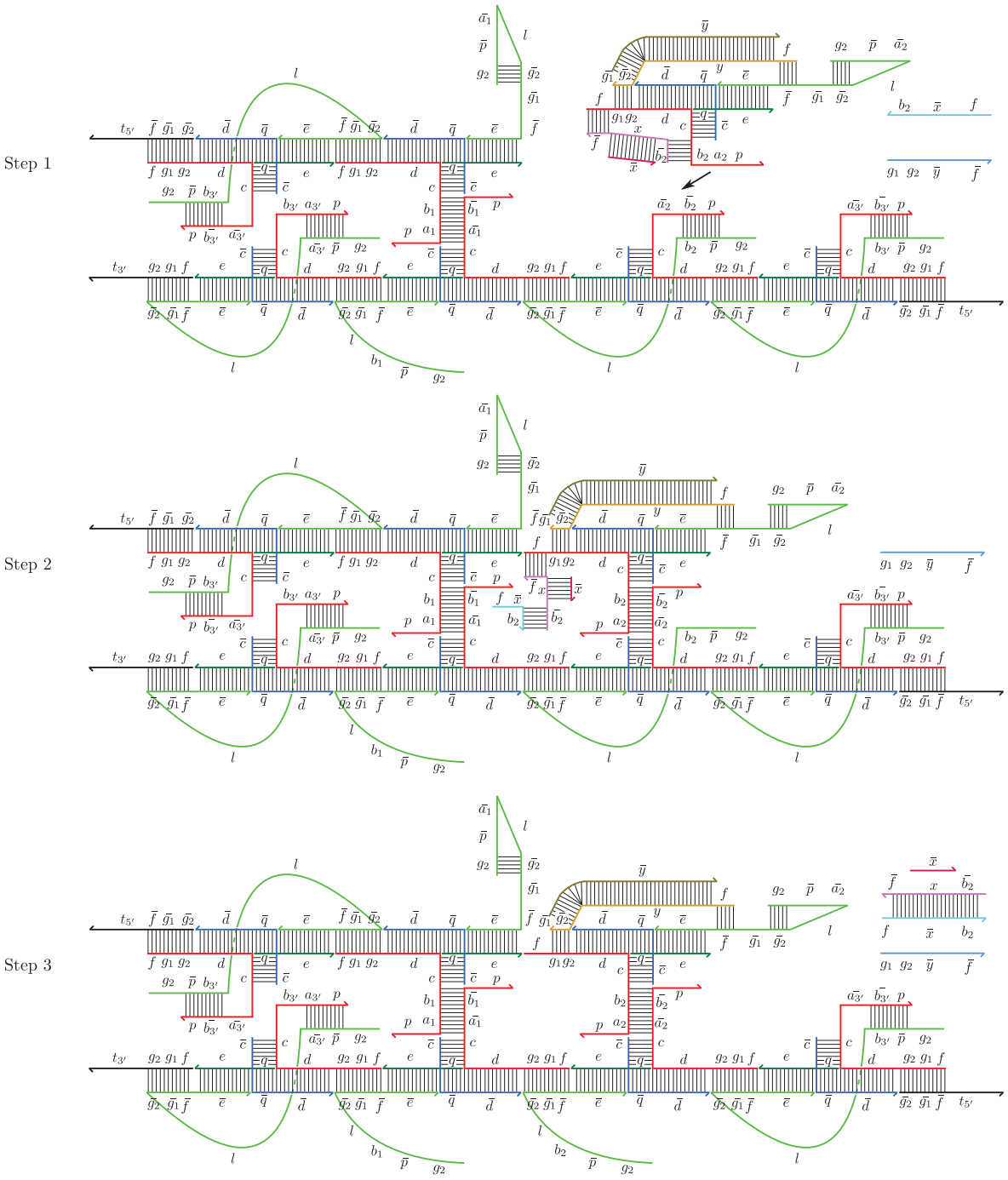



Figure 17: Replication: Step 1: A four meta-base mDNA sequence, with the first and last bases being the special terminator bases. Step 2: A meta-nucleotide with an active unbound base pad comes in and binds to the template via the sequence $b_2a_2$, strand displacing out the subsequences $\bar{p}b_2$ and $\bar{b}_2$. Step 3: The strand $f\bar{x}b_2$ now strips away $\bar{b}_2x\bar{f}$ via the toehold $\bar{b}_2$, activating the $5'$ pad of the incoming meta-nucleotide

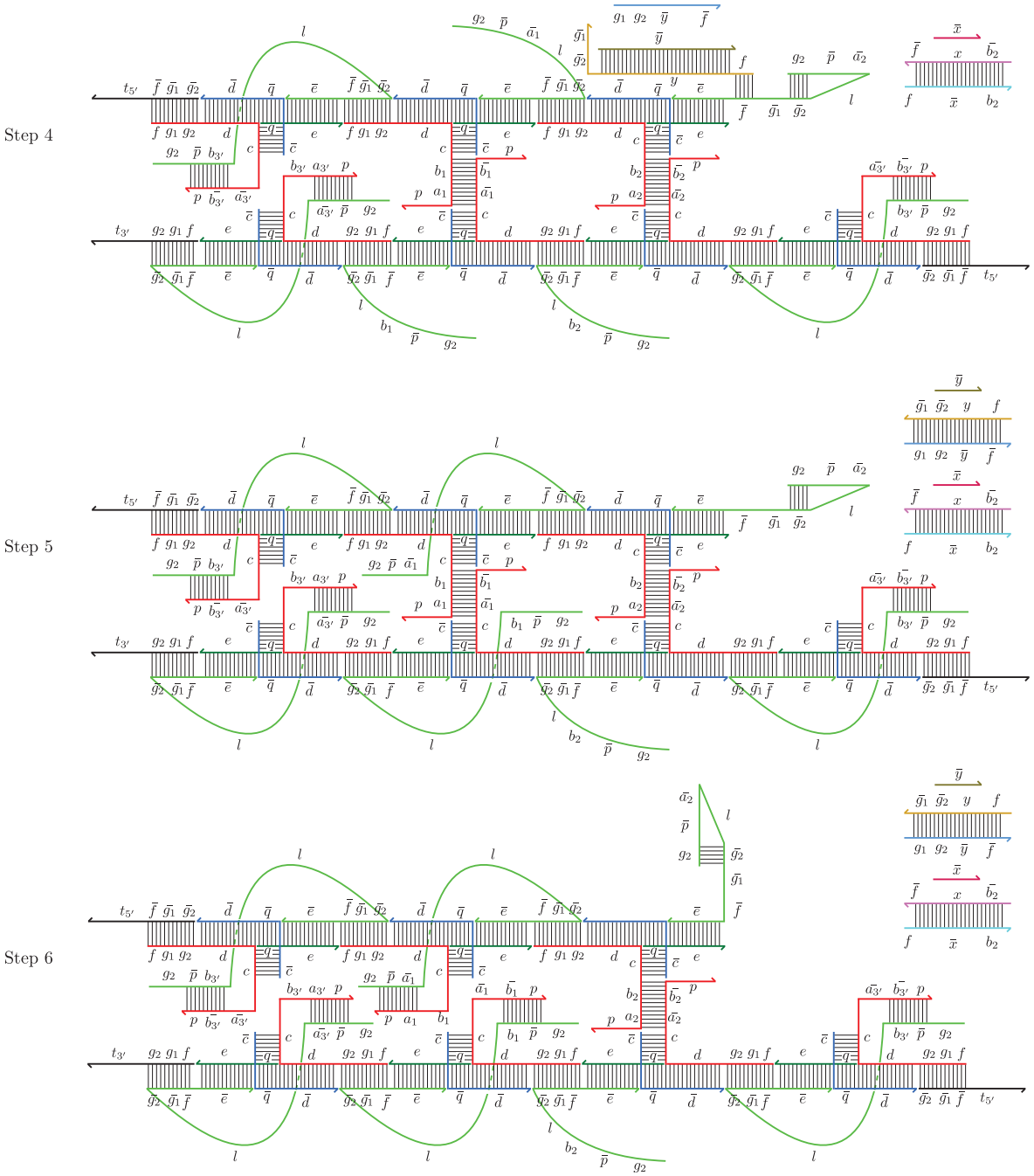



Figure 18: Replication: Step 4: the 5′ pad $fg_1g_2$ binds with the 3′ pad $\bar{g}_2\bar{g}_1\bar{f}$, displacing $\bar{g}_2\bar{g}_1$ and opening up the hairpin structure $\bar{p}\bar{a}_1 l$. Step 5: The strand $fy\bar{g}_2\bar{g}_1$ is stripped away through the toehold $\bar{g}_2\bar{g}_1$ by the strand $g_1g_2\bar{y}\bar{f}$, activating the 3′ end. Step 6: The released hairpin from step 4 can now cleave $a_1$ from $\bar{a}_1$ using the toehold $p$, which is half of the base pad binding of the upstream meta-nucleotide, while the other half $\bar{b}_1$ of the base pad binding is cleaved from $b_1$ via the toehold $p$ by the sequence $\bar{p}b_1$ (which is part of the previously incorporated meta-nucleotide)