

Chapter 4

Toward a Synthetic Polymer that Grows Exponentially Fast⁰

4.1 Abstract

The exponentially-fast-growing polymer introduced in the previous chapter is implemented here using DNA molecules. The result is the first synthetic linear polymer capable of growing in logarithmic time. Insertion and division are implemented by modifying the autonomous polymerization design [Venkataraman et al., 2007], whereby stable oligonucleotide complexes interact using four-way branch migration when a trigger Initiator complex is present in the solution. We experimentally verify the exponential kinetics of our system using spectrofluorimetry, we qualitatively compare the size of our polymers over time to those grown in a linear system via gel electrophoresis and we verify their shape via Atomic Force Microscopy.

⁰This work was coauthored by Nadine Dabby & Ho-Lin Chen*, and is in preparation [Dabby and Chen, 2013a] with the following contributions: experiments and analysis were performed by N.D. with supervision from H-L.C.; the manuscript was written with input from both authors.

4.2 Introduction

Material science and nanotechnology seek to achieve some of the formidable molecular tasks that biology takes for granted, such as the growth of complex structures in two or three dimensions in logarithmic time. One of the main thrusts in molecular programming is to use computer science as a tool for figuring out what can be achieved. While molecular systems that are Turing-complete have been demonstrated [Winfree, 1996], these systems still cannot achieve some of the feats biology has achieved. The need for new formalisms to describe what molecular systems [Woods et al., 2013] and macro-scale systems [Chirikjian, 1993, Goldstein and Mowry, 2004, Rosa et al., 2006, Griffith, 2004, White et al., 2004, Klavins et al., 2004, Jones and Mataric, 2003, Murata et al., 1994, Nagpal, 2008, Werfel and Nagpal, 2007, Arbuckle and Requicha, 2004, Rus and Vona, 2001, Butler et al., 2001, Yim et al., 1997, Yim et al., 2007, Groß and Dorigo, 2008, Walter et al., 2004] are capable of has spawned research in the area of active self assembly to describe the behaviors that have been and can be implemented by such systems [Yin et al., 2004, Tian et al., 2005, Bath et al., 2005, Pei et al., 2006, Green et al., 2008, Omabegho et al., 2009, Yin et al., 2008, Lund et al., 2010, Muscat et al., 2011, Dirks and Pierce, 2004, Venkataraman et al., 2007, Yin et al., 2008].

In the previous chapter, a model for active self assembly was constructed. This chapter presents the molecular implementation of two active behaviors (exponential growth and splitting or division of polymers) using DNA.

By encoding the order of the nucleotides in a DNA sequence, we can control the interaction of DNA strands. Sub-sequences of these strands are called domains and it is their binding (hybridization) and unbinding (disassociation) from complementary domains that determines what a system of strands can do. In DNA nanotechnology, dynamic systems of DNA molecules can be orchestrated by toeholds, the short sequences of DNA that are complementary to single-stranded domains in a target molecule [Yurke et al., 2000, Zhang and Winfree, 2009]. Toeholds serve as the inputs to dynamic DNA systems and initiate branch migration, a random walk process of bond breaking and formation that results in the exchange of one strand in the duplex for another single strand with the same sequence (described in Chapter 1).

Our DNA implementation is inspired by the Hybridization Chain Reaction (HCR) system de-

veloped by Dirks and Pierce [Dirks and Pierce, 2004]. Their construction, which triggers the polymerization of DNA monomers, uses two single-stranded DNA hairpins that each have the same 18 base-pair stem sequence and a toehold that is complementary to the other hairpin's loop sequence. These strands are kinetically trapped—they are unable to access their lowest energy state because of the large energy required to disrupt the hairpin conformation. As a result, they react with each other very slowly in the absence of an initiator. The initiator is a molecular trigger that consists of a domain that is complementary to one hairpin's toehold and another domain that is complementary to that hairpin's adjacent stem sequence. When the initiator is added to the solution of monomer hairpins, it binds to the toehold of the first hairpin and launches a strand displacement reaction that opens that hairpin. The newly exposed bases of the opened hairpin can then undergo a similar reaction with the second hairpin. The two hairpins will continue to polymerize until an equilibrium concentration of hairpin monomers is reached.

In a subsequent work, the HCR system was modified to employ four-way branch migration and create an autonomous polymerization motor [Venkataraman et al., 2007]. The metastable fuel hairpins from the Hybridization Chain Reaction system were modified to include an extra toehold, and the initiator strand was replaced by an initiator complex that is composed of an “anchor” strand and a “rickettsia” strand. Upon mixing, the first hairpin binds to the sticky ends of the anchor-rickettsia complex, initiating a four-way branch migration in which the rickettsia strand is passed from the anchor to the hairpin. The second hairpin then binds to the newly exposed sticky ends and the rickettsia strand is passed to the second hairpin. The rickettsia strand continues to be passed forward to newly added hairpins as the polymer grows behind it.

In Chapter 3, we explored the implications of modifying the Rickettsia system (described in Chapter 1) by adding an additional hairpin and an additional toehold within each loop. Here we first describe our molecular implementation of this modification (Section 4.3) and then describe our main results (Sections 4.4 and 4.6).

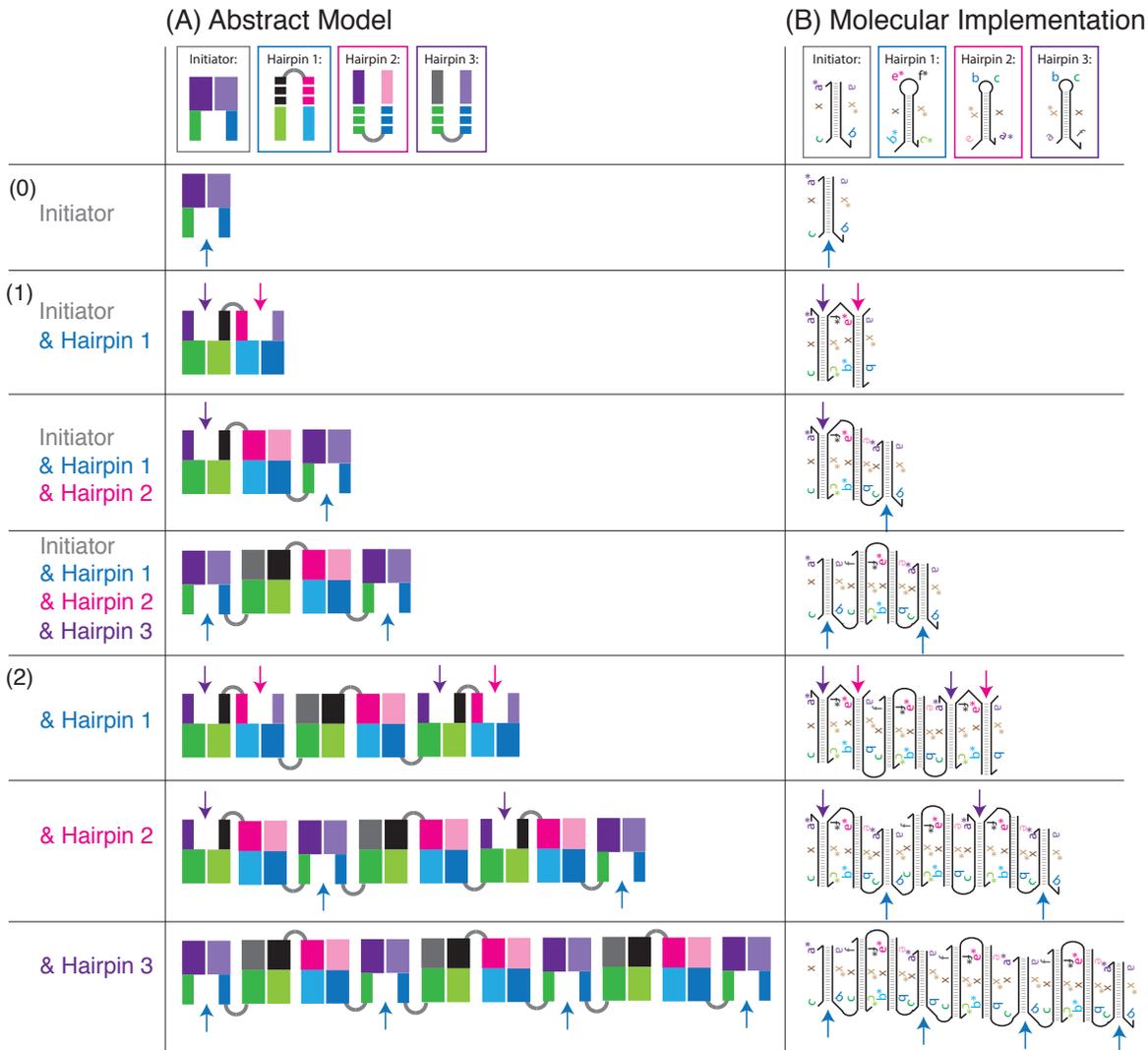


Figure 4.1: (Identical to Figure 3.4.) The schematic of our insertional polymer implementation shows the first two rounds of growth. (A) The abstract representation of our exponential growth polymer. (B) The molecular implementation of our polymer is color-coded the same way. DNA sequences in the oligonucleotides are color-coded by domain (purple, green, blue, brown, pink, and black). The boxes around each oligonucleotide in (B) correspond to the insertion arrows as follows: a blue arrow indicates an insertion site for Hairpin 1, a pink arrow indicates an insertion site for Hairpin 2, a purple arrow indicates an insertion site for Hairpin 3. Exponential growth occurs as follows: (0) The Initiator has one insertion site for Hairpin 1 (blue arrow). Insertion of Hairpin 1 is driven forward by the hybridization of 6 new base pairs. (1) After Hairpin 1 inserts into the Initiator, two new insertion sites are generated: one for Hairpin 2 (pink arrow) and one for Hairpin 3 (purple arrow). Hairpin 2 and Hairpin 3 are sequentially inserted (in solution insertion occurs asynchronously), each one generates a new insertion site for Hairpin 1 (blue arrows). After the first round of insertion, two insertion sites for Hairpin 1 are generated from what was initially (in round (0)) one site. (2) A second round of insertion is illustrated. After the second round of insertion, four new insertion sites for Hairpin 1 are generated.

4.3 Molecular Implementation

Figure 4.1 (identical to Figure 3.4 reproduced here for convenience) shows the molecular implementation of our exponential growth system from the previous chapter. Hairpin 1 (H1) and the Initiator (I) react first; this results in two new insertion sites: one that is complementary to Hairpin 2 (H2), and another that is complementary to Hairpin 3 (H3). Upon insertion of H2 and H3 into the growing polymer two new insertion sites that are complementary to H1 are regenerated. Thus for every initial H1 insertion site, after each round of insertions (of H1, H2 and H3), two new H1 insertion sites are created.

The initial reaction (insertion of H1 into the Initiator complex) is driven by the hybridization of six new base pairs. After that, each new hairpin that is inserted adds nine base pairs to the system. Some of these steps become reversible as the system reaches equilibrium. The free energy and reversibility of toehold-mediated four-way branch migration is explored in depth in Chapter 5.

In addition to the insertional monomers that grow the polymer, we introduce a new type of monomer, which we call a Divide complex, that upon insertion splits the polymer into two pieces, as we will discuss in Section 4.6.

Figure 4.2 is a legend for the set of DNA molecules used in this chapter. Each oligonucleotide complex (Initiator, Hairpin1, Hairpin 2, Hairpin 3, and Divide) is shown with color-coded motifs (purple, green, blue, brown, pink, and black) that correspond to the colored DNA subsequences (Figure 4.2A). The Initiator-ROX complex is a modified Initiator complex with a single fluorophore tag for gel electrophoresis experiments (Figure 4.2B). Hairpin 2RQ (H2RQ) is a modified Hairpin 2 molecule with a quencher and fluorophore pair on opposite ends of the molecule, used in the spectrofluorimetry experiments (Figure 4.2B). Hairpin 2L (H2L) and Hairpin 3L (H3L) are inactivated versions of Hairpins 2 and 3, in which the loops are replaced with a poly-T sequence (Figure 4.2C). The boxes around each oligonucleotide correspond to the insertion arrows as follows: a blue arrow indicates an insertion site for Hairpin 1, a pink arrow indicates an insertion site for Hairpin 2, Hairpin 2RQ or Hairpin 2L, a purple arrow indicates an insertion site for Hairpin 3 or Hairpin 3L, and a green arrow indicates an insertion site for the Divide complex (Figure 4.2D).

In each diagram, we utilize a domain abstraction for referring to stretches of consecutive nu-

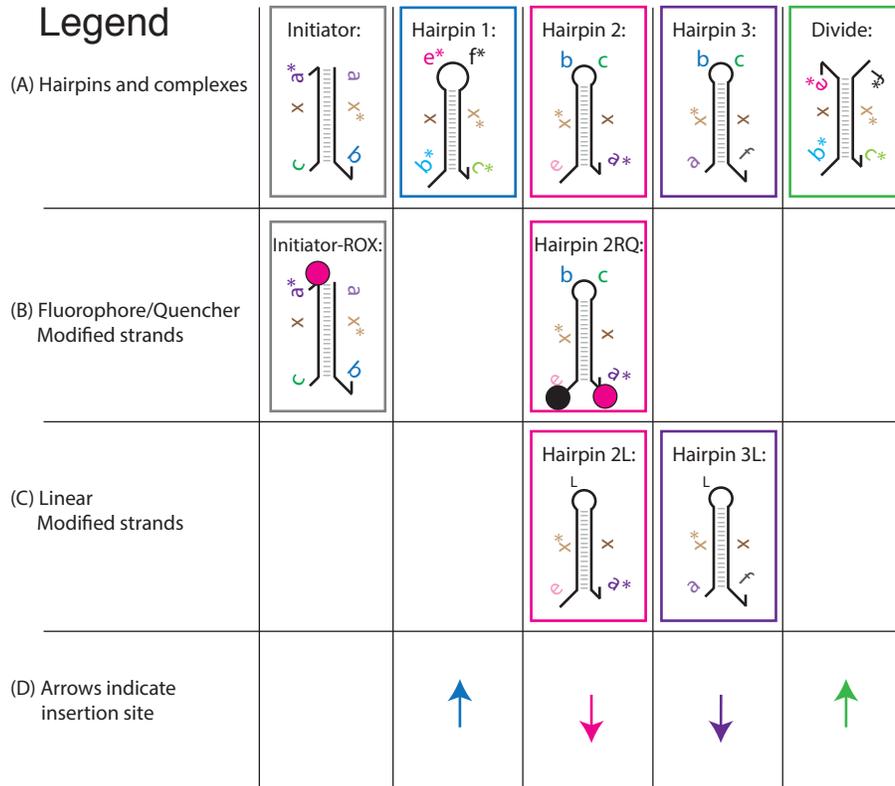


Figure 4.2: Legend of DNA hairpins and complexes. (A) Schematics of the Initiator complex, Hairpin 1, Hairpin 2, Hairpin 3, and Divide complex. Each oligonucleotide is shown with color-coded motifs that correspond to the DNA subsequences. (B) The Initiator-ROX complex is a modified Initiator complex with a single fluorophore tag for gel electrophoresis experiments. Hairpin 2RQ is a modified Hairpin 2 molecule with a quencher and fluorophore pair on opposite ends of the molecule, used in the spectrofluorimetry experiments. (C) Hairpin 2L and Hairpin 3L are inactivated versions of Hairpins 2 and 3, in which the loops are replaced with an inactive poly-T sequence. The boxes around each oligonucleotide in (A) (B) (C) correspond to the insertion arrows in (D) as follows: a blue arrow indicates an insertion site for Hairpin 1, a pink arrow indicates an insertion site for Hairpin 2, Hairpin 2RQ or Hairpin 2L, a purple arrow indicates an insertion site for Hairpin 3 or Hairpin 3L, a green arrow indicates an insertion site for the Divide complex.

cleotides that act as a unit in binding to complementary stretches of nucleotides. Domains are represented by Latin letters (Figure 4.1). Letters followed by an asterisk denote complementary domains, e.g.: x is complementary to x^* . Single-stranded molecules of DNA (henceforth strands) are comprised of concatenated domains. DNA complexes are composed of two or more noncovalently-bound strands. There are two types of toeholds in our system: long toeholds that indicate a stronger desired interaction (six bases in length) and short toeholds that indicate a weaker desired interaction (three bases in length).

In the next two sections, we confirm exponential growth by measuring the conversion of monomers into a product. We then qualitatively measure the size of products over time. Finally, we verify the predicted structure using Atomic Force Microscopy.

4.4 Exponential Growth Results

4.4.1 Exponential Growth Mechanism Controls

We tested each insertion step in the exponential growth mechanism by using the inactivated versions of Hairpins 2 and 3 (Figure 4.3). Hairpin 2L and Hairpin 3L were added to the Initiator and Hairpin 1 both individually (this results in exactly one insertion event) and together with the normal version of the other hairpin, which results in linear growth. We note that there is more product in lanes 14 (I, H1, H3L) and 15 (I, H1, H3) than there is in lanes 12 (I, H1, H2L) and 13 (I, H1, H2). The reactants in lanes 12 and 14 can only proceed through two steps of the polymerization reaction due to the inactivated strands. At equilibrium (after 6 hours) there is more dimerization between the Initiator-Hairpin 1 complex and Hairpin 3L than there is between the Initiator-Hairpin 1 complex and Hairpin 2L. Thus Hairpin 3 appears to have a greater affinity to the Initiator-Hairpin 1 complex than Hairpin 2. This observation implies that the two reactions have different rate constants, Hairpin 2 is either slower to react with its insertion site or faster to dissociate from its insertion site than Hairpin 3 (or both).

When all possible combinations of the reactants are made, the leaks in the system can be assessed using gel electrophoresis. A “leak” is an undesired molecular interaction. We tested four

different sequence designs (Section B.1), and chose the one with a low leak and fast reaction time. Lanes 9 (H1, H2) and 10 (H1, H3) in Figure 4.3 show that a small leak occurs between Hairpin 1 and Hairpin 2 and between Hairpin 1 and Hairpin 3. However, no leak occurs between any of the other species.

We quantify the leak via spectrofluorimetry experiments in Figure 4.8 [Yin et al., 2008]: we adjust the Initiator concentration $[I]$ by an additional term $[I]_{leak}$ to obtain an effective Initiator concentration $[I]_{effective} = [I] + [I]_{leak}$. We then fit the $[I]_{leak}$ parameter to our data and find that in the exponential system $[I]_{leak} = 0.04\times$ and in the linear system $[I]_{leak} = 0.01\times$. Reactions were started with the addition of Hairpin 1 in order to avoid the leak.

The reader may observe the presence of faint extra bands in the lanes that contain only individual hairpins. These are dimerized hairpins that form in small amounts from individual hairpins when the strands are annealed. We minimize their presence by snap cooling. Snap cooling the hairpins results in the same amount of dimerized monomers as gel purification (data not shown). All hairpins except for the Initiator were snap cooled prior to experiments. The Initiator is a gel-purified duplex composed of two molecules of DNA.

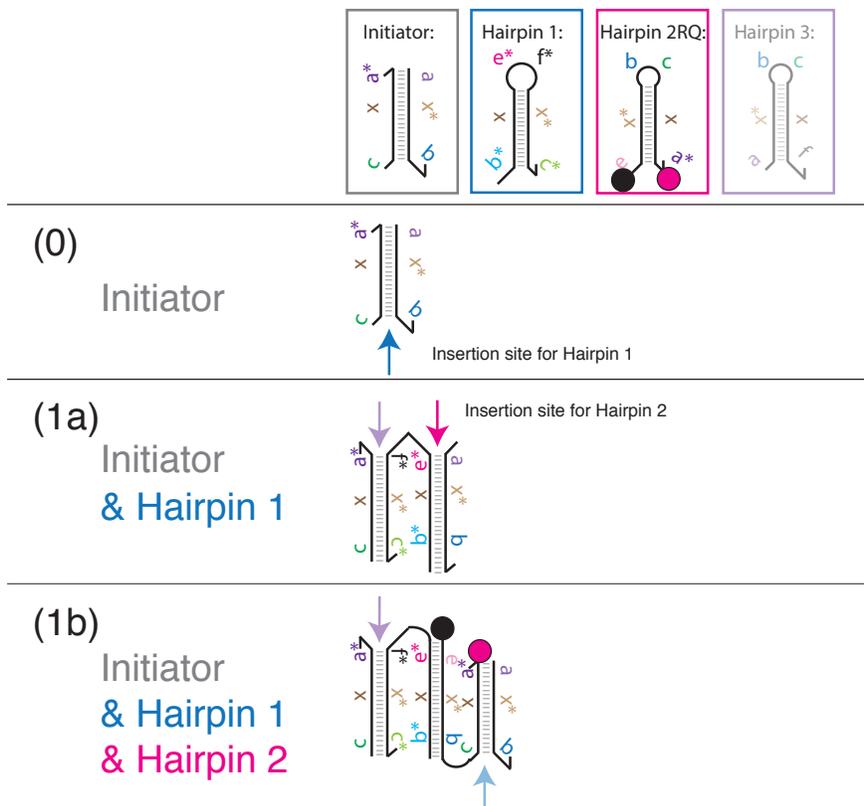


Figure 4.4: The experimental design for measuring conversion of monomer hairpins into the polymer. Above is a modified version of the schematic from Figure 4.1. The boxes around each oligonucleotide correspond to the insertion arrows in the mechanism below, which shows the incorporation of Hairpin 1 (1a) and Hairpin 2 (1b) into the polymer. Note that Hairpin 2RQ is a modified version of Hairpin 2 that includes a fluorophore-quencher pair. The fluorophore (pink circle) is quenched before Hairpin 2RQ reacts with the polymer. Upon Hairpin 2RQ's insertion into the polymer (1b), the fluorophore (pink circle) and quencher (black circle) are separated and the fluorophore emits light. We measure the kinetics of the incorporation of Hairpin 2RQ into the polymer by measuring the increase of fluorescence in the solution over time.

We probed both the linear and exponential polymerization over eight different Initiator concentration values. The time course of fluorescence intensity confirmed linear conversion of hairpins in the system with one inactivated strand (Figures 4.5, and 4.6), and exponential conversion of hairpins in the full system (Figures 4.7 and 4.8).

In a linear growth system, the total mass of polymer product, P , grows as a function of initial

Initiator concentration, I_0 , and time, t , as follows:

$$P = k \times I_0 \times t. \quad (4.1)$$

The time at which 10% of monomers are consumed, $t_{10\%}$, is

$$t_{10\%} = \frac{P_{10\%}}{k \times I_0}. \quad (4.2)$$

Thus, in a linear growth system, the time to 10% completion of polymer growth (10% conversion of hairpins) is inversely proportional to initial Initiator concentration. When plotted on a logarithmic concentration scale, the time to 10% conversion exponentially decays as a function of increasing initial Initiator concentration. This model fits our linear growth system data (Figure 4.6).

In an exponential growth system, the total mass of polymer product, P , grows as a function of initial Initiator concentration, I_0 , and time, t , as follows:

$$P = I_0 \times e^{(kt)}. \quad (4.3)$$

The time at which 10% of monomers are consumed, $t_{10\%}$, is

$$t_{10\%} = \frac{1}{k} \times (\ln(P_{10\%}) - \ln(I_0)). \quad (4.4)$$

Thus, in an exponential growth system, the time to 10% completion of polymer growth (10% conversion of hairpins) is a linear function of the logarithm of the initial Initiator concentration. When plotted on a logarithmic concentration scale, the time to 10% conversion linearly decreases with increasing initial Initiator concentration. This is what we observe in our exponential growth system data (Figure 4.8).

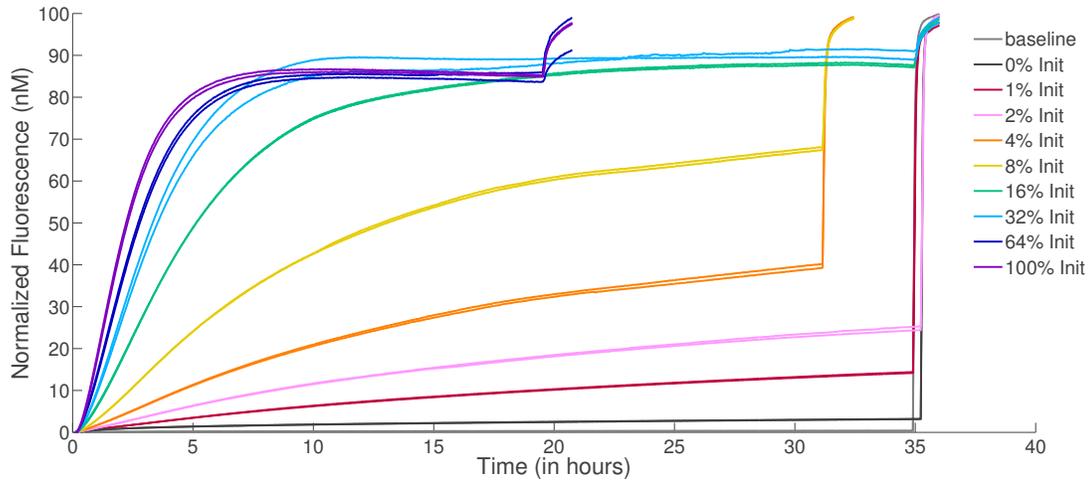


Figure 4.5: Linear polymer growth kinetics are observed in a fluorescence time course when inactivated Hairpin 3L is substituted for Hairpin 3. As Hairpin 2 is incorporated into the growing polymer, the system's fluorescence increases: this illustrates the conversion of hairpins into polymers. Plotted above are the kinetic traces of Hairpin 2RQ (all hairpins are present at 100 nM) with varying amounts of Initiator.

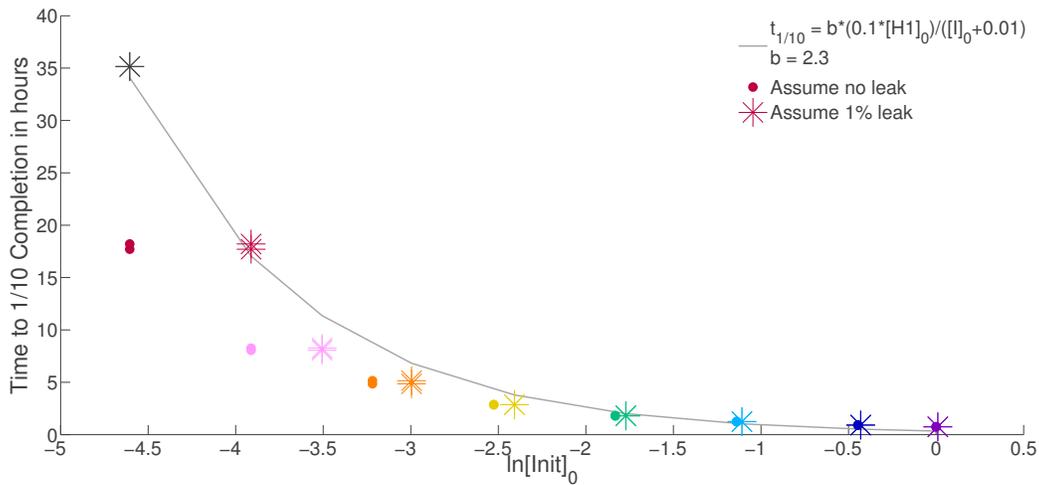


Figure 4.6: Linear polymer growth kinetics are observed when an inactivated version of Hairpin 3 is substituted for Hairpin 3. Above is the linear fit of the 10% completion time as a function of the relative concentration of Initiator to hairpins. Filled circles correspond to a system where we assume no leak. Asterisks indicate the same points but assume a leak equivalent to 1% of the Initiator concentration.

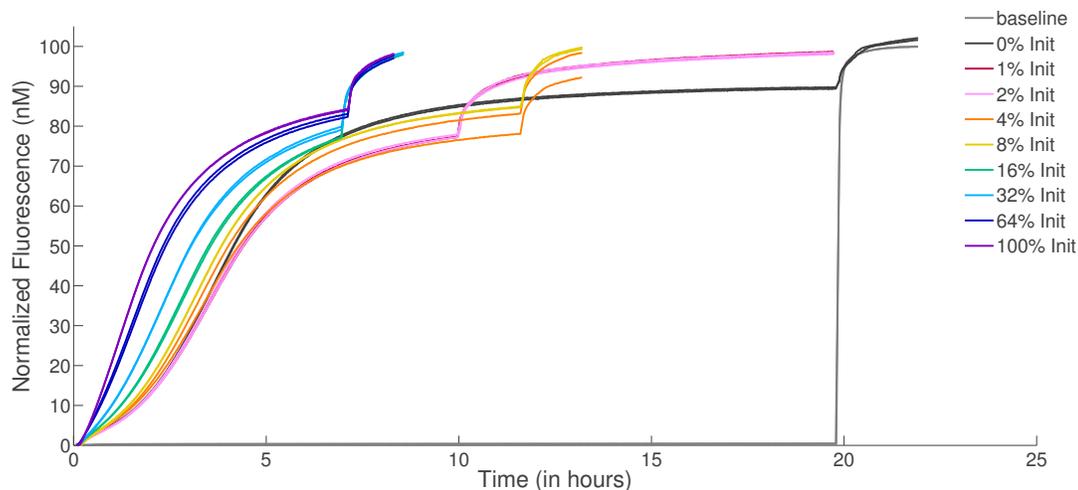


Figure 4.7: Exponential polymer growth kinetics examined via fluorescence. As Hairpin 2RQ is incorporated into the growing polymer, the system's fluorescence increases; this illustrates the conversion of hairpins into polymer. Plotted above are the kinetic traces of Hairpin 2RQ (all hairpins are present at 100 nM) with varying amounts of Initiator.

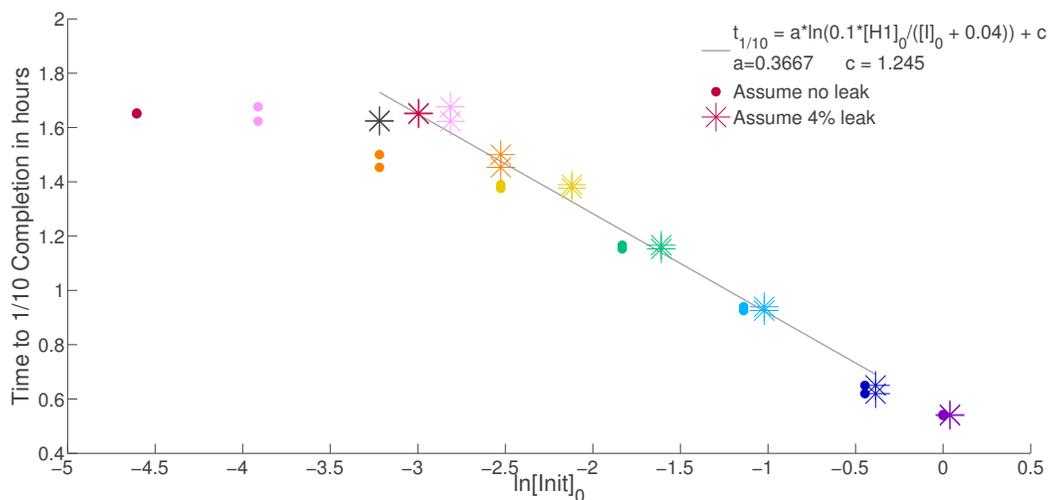


Figure 4.8: Linear fit of the 10% completion time as a function of the relative concentration of Initiator to hairpins. Filled circles correspond to a system where we assume no leak. Asterisks indicate the same points but assume a leak equivalent to 4% of the Initiator concentration. We assume leak is equivalent to added Initiator using a model from [Yin et al., 2008]. Using that assumption the data is shifted and the curves match. Thus the leak shown in Figure 4.3 does not significantly affect our data.

The polymers formed at each Initiator concentration were examined by gel electrophoresis in order to characterize their length distribution. Each Initiator molecule was tagged with one ROX

fluorophore. As the hairpins are successively added to a polymer, each polymer that is “properly initiated” retains exactly one fluorophore, thus the ROX fluorescence signal directly correlates to the number of polymers at a given size. The sizes were binned after post-staining with SYBR Gold, which allowed the DNA ladder to be visualized.

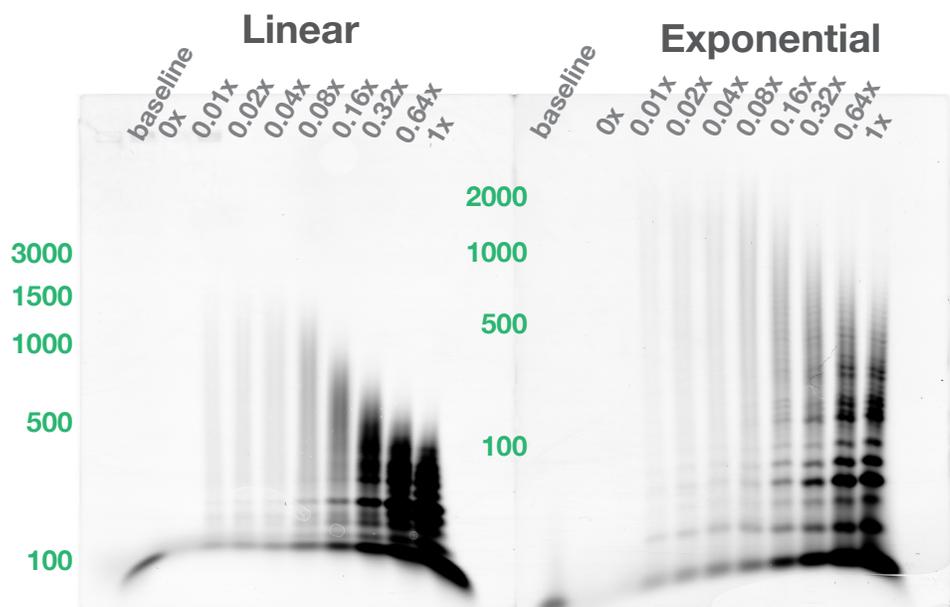


Figure 4.9: The average length of the polymer shrinks with increasing Initiator concentration (left to right) in both the linear and exponential systems. Rox fluorescence intensity imaged in this Super Fine Resolution Agarose gel shows the distribution of polymer lengths generated in the presence of Initiator concentrations $[I]_0 = \{0\%, 1\%, 2\%, 4\%, 8\%, 16\%, 32\%, 64\%, 100\%\}$ relative to hairpin concentrations after 6 hours of reaction with $1\mu\text{M}$ Hairpins. Lanes with Initiator concentrations smaller than 8% are difficult to resolve by eye in this gel as they contain less ROX-labeled Initiator. Figure B.5 shows an image of this gel after staining with SYBR Gold.

The mean length (in base pairs) of polymers decreases with increasing Initiator concentration above 4% of relative hairpin concentrations. (See Figure 4.9, 4.10, B.5 for gels and binned data of both linear and exponential systems). This is expected because high concentrations of Initiator outcompete existing insertion sites for free hairpins. In the case of Initiator concentration below 4% of relative Hairpin concentrations, the different amounts of leak in the systems are presumably responsible for the different distributions of polymer length between the linear and exponential

system. The smaller leak in the linear system (1%) would explain why the linear system produces longer polymers than the exponential system (which has a 4% leak).

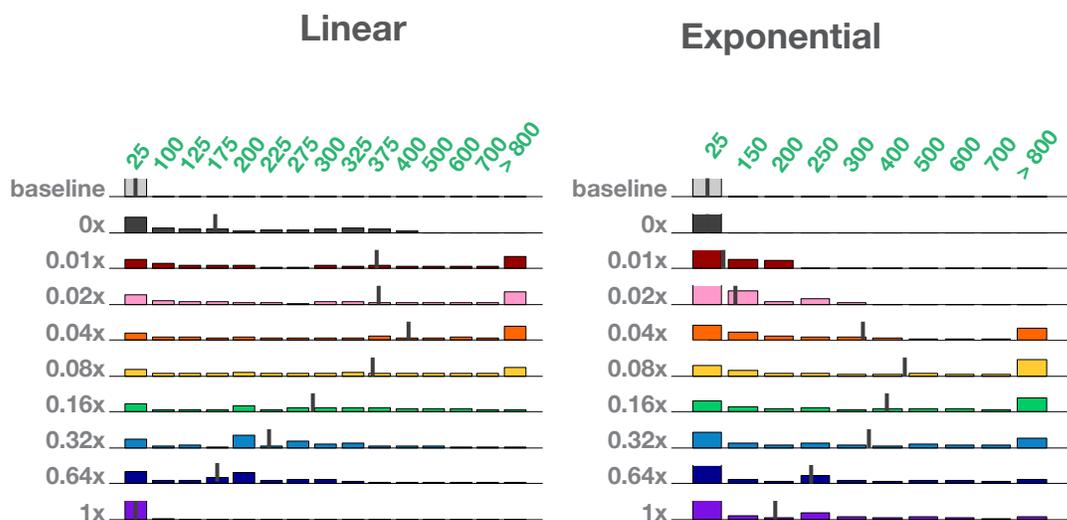


Figure 4.10: Normalized distributions of polymer length from the data in Figure 4.9 show a decrease in the mean length of polymer with increased Initiator concentration in both the linear and exponential systems. The upper limit of the y-axis for each distribution is 40% of the total concentration to allow the reader to clearly see the change in distribution. Vertical lines indicate the lower bound on the mean of the distribution, as calculated with all polymers larger than 800 base pairs being assigned a length of 800 base pairs. The mean of each distribution in which more than 8% Initiator was utilized decreases with increasing amounts of Initiator.

4.4.3 Imaging with Atomic Force Microscopy

Atomic Force Microscopy of the reaction product confirms the formation of unbranched polymers in the exponential system (Figure 4.11). In comparing images of both the polymer and the leak product, we find that the leak product is capable of growing much larger than the intended polymer, but the polymer grows faster. Others have shown that polymer growth in the absence of Initiator can provide an upper bound for how big the polymer can grow [Beck, 2011]. It is unclear whether the leak product is a linear polymer. It may be a highly pseudo-knotted structure.

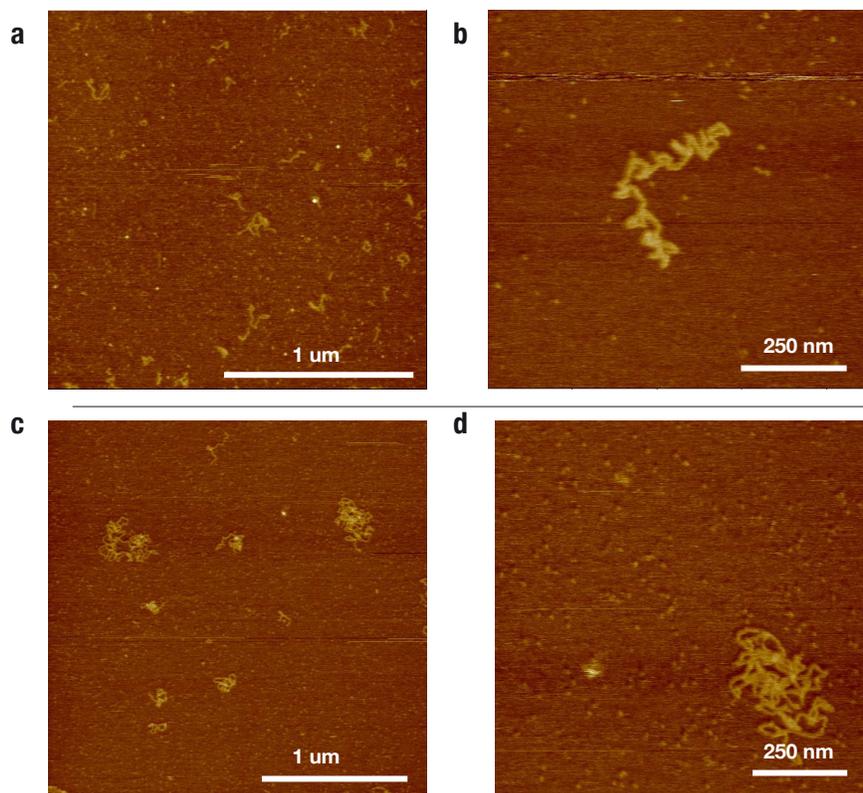


Figure 4.11: Atomic Force Microscopy images of exponentially grown polymer and leak product. a, Wide field image of polymer with 10% Initiator strand relative to the other hairpins after 5 hours (scale bar is 1 μm). b, One of these polymers (scale bar is 250 nm). c, Wide field image of the leak formed by hairpins in the absence of Initiator (scale bar is 1 μm). d, One of these leak products (scale bar is 250 nm).

4.4.4 Time Lapse Experiments

A qualitative difference between the exponential and linear systems is also observed when examining polymer size over time in Figures 4.12 and 4.13. (See Figures B.6 and B.8 for two additional exponential system time lapse gels and see Figures B.7 and B.9 for the SYBR Gold stained versions of all of these gels).

According to this preliminary gel data, the exponential system (Figure 4.13) generates longer polymer products sooner than the linear system (Figure 4.12). While it takes the linear system 25 min to produce a polymer of length 500 base pairs, the exponential system produces a 500-base pair

polymer within 10 min. The exponential system produces a detectable amount of 1000-base pair polymer within 20 minutes, at least four times faster than the linear system, which takes between 90 and 120 min to produce a 1000-base pair polymer. Although this is not proof of exponential growth, it is consistent with the expectation that exponential growth progresses more quickly than linear growth.

Figure 4.13 is particularly rich in data. In addition to showing that the polymers produced in the exponential system grow large quickly, the gel clearly shows that polymer growth occurs in quantized chunks of approximately 25 base pairs at a time. This is expected, as each hairpin contains between 54 and 57 nucleotides. The bands generated by the polymerization alternate between faint and dark within each lane. This corroborates our earlier claim that Hairpin 2 is slower to react with its insertion site than Hairpin 3. If the backward reaction rates for both of these reactions are equivalent, then this implies that the reaction between H2 and its insertion site is a slower step in the formation of polymers.

The exponential time lapse gel in Figure 4.13 and the replicate in Figure B.8 expose an issue. The signal of the bands relative to background fades from left to right. In the SYBR Gold-stained versions of these gels, as shown in Figure B.9, the lanes to the right show noticeably less total stained DNA than the other lanes. We suspect that this behavior is a result of the complexity of loading the gel: in order to ensure that the experiments are initiated and the gel is run exactly on time, the right half of the gel (higher time point reactions) is loaded approximately 30 minutes in advance of the shorter time lapse reactions. This may allow for the DNA in these wells to diffuse out of the wells in advance of running. Another concern is the fading of the bands at the top of the gel in the longer time lapse reactions. It is unclear why this fading occurs. More repetitions of these results will be necessary.

We hypothesize that the fluorescent loading dye bromophenol blue interferes with the fluorescence read-out of our properly initiated polymers. The gel in Figure 4.13 has a dark band in all lanes across the bottom of the gel. By comparison, this band becomes faint at intermediate times for the replicate in Figure B.8 and disappears at long time points in the replicate in Figure B.6. In the SYBR Gold-stained versions of the gels in Figure 4.13 and Figure B.8, as shown in Figure B.9, this band fades significantly. Since bromophenol blue does not fluoresce at the excitation spectra

of SYBR Gold, we can assume that only stained DNA is visible, and that if the dark lower bands in the gels were unused initiator, then there would be a larger amount of DNA at these lengths. Thus, a more complete analysis of these gels was precluded due to the interference of the fluorescent loading dye bromophenol blue and an improperly stained ladder in the linear system time lapse gel that makes it difficult to resolve at molecular weights above 1000 base pairs (Figure B.7).

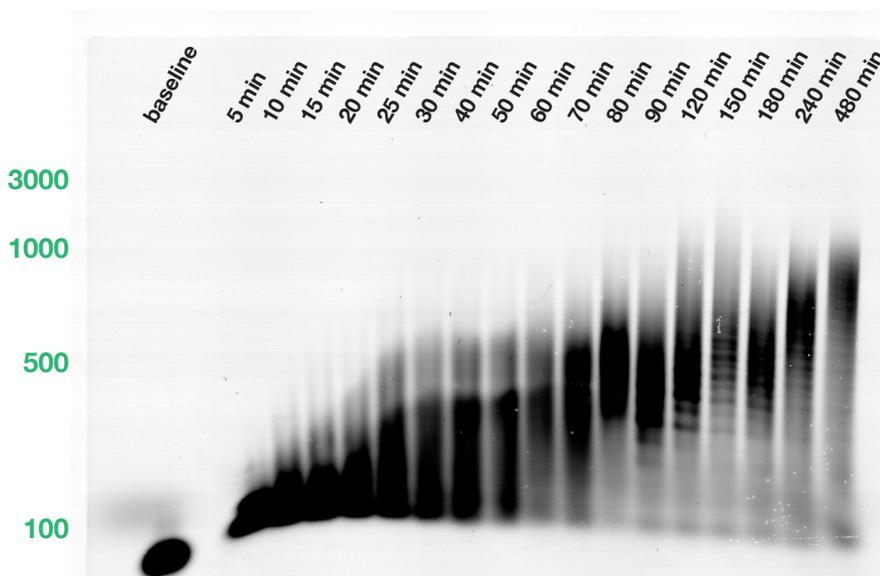


Figure 4.12: Gel time-lapse studies of linear polymer growth. Super Fine Resolution Agarose non-denaturing gels of the product of a polymerization reaction with 80 nM ROX-labeled Initiator, 1.5 μM Hairpin 1, and 1 μM of Hairpin 2 and Hairpin 3. ROX fluorescence was imaged prior to staining with SYBR Gold. (The SYBR Gold stained gel can be found in Figure B.7). A more complete analysis of this gel was precluded due to the interference of the fluorescent loading dye bromophenol blue as discussed in Section 4.4.4.

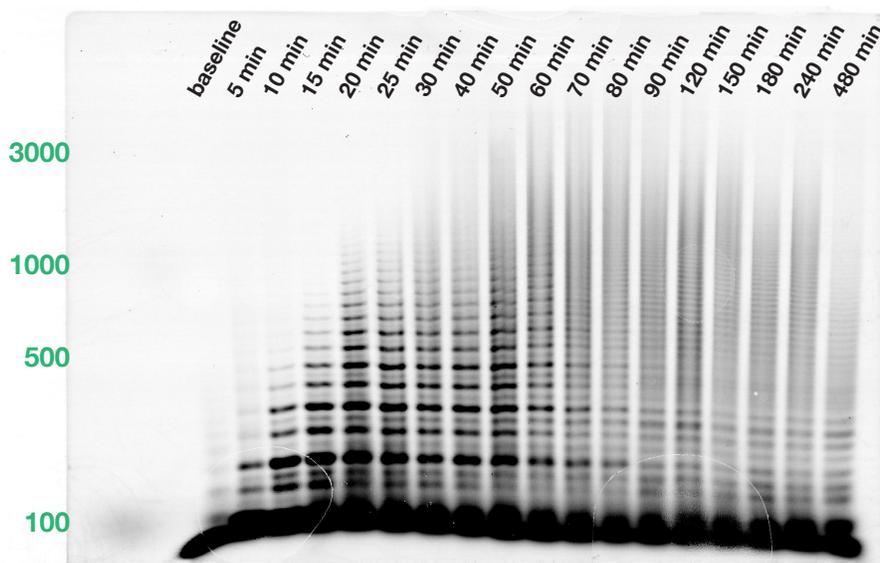


Figure 4.13: Gel time-lapse studies of exponential polymer growth. Super Fine Resolution Agarose non-denaturing gels of the product of a polymerization reaction with 80 nM ROX-labeled Initiator, 1.5 μM Hairpin 1, and 1 μM of Hairpin 2 and Hairpin 3. ROX fluorescence was imaged prior to staining with SYBR Gold. (The SYBR Gold stained gel can be found in Figure B.9). Two additional experimental runs of this experiment can be found in Figures B.6 and B.8. A more complete analysis of this gel was precluded due to the interference of the fluorescent loading dye bromophenol blue as discussed in Section 4.4.4.

4.5 Analysis

The exponential growth system described here can be modeled with the following chemical reactions:



where S2 and S3 are the insertion sites for H2 and H3 respectively, and P2 and P3 are double-stranded sections of polymer that are henceforth unreactive. The chemical reactions can be further

simplified by the following three assumptions:

1. Each Initiation site is equivalent.
2. The forward rates are the same for all three reactions. Thus, $k_1 = k_2 = k_3 = k$.
3. The reactions are irreversible. Thus, $k_{-1} = k_{-2} = k_{-3} = 0$.

The first assumption makes the set of reactions tractable. The second assumption comes from the next chapter, where we show that the forward rate of a four-way branch migration reaction is dependent on the length and sequences of interacting toeholds and consequently on the overall free energy changes in the system. In our implementation all toehold pairs share the property of being nine bases long with approximately equivalent GC content, making their free energy changes roughly equivalent. The final assumption is justified by the decreasing free energy of the system at each step. We note that this assumption may not hold as reactants are consumed by the system.

With the above assumptions the set of reactions can be reduced to:



We next require that $[S2] = [S3]$, henceforth they will both be replaced by a variable called $[S]$. Then we can simulate the rate of change of $[I]$, $[H1]$, $[H2]$, $[H3]$, and $[S]$ with the following differential equations:

$$\frac{d[I]}{dt} = -k[I][H1] + k[S][H2] + k[S][H3], \quad (4.11)$$

$$\frac{d[H1]}{dt} = -k[I][H1], \quad (4.12)$$

$$\frac{d[S]}{dt} = 2k[H1][I] - k[S][H2] - k[S][H3], \quad (4.13)$$

$$\frac{d[H2]}{dt} = -k[S][H2], \quad (4.14)$$

$$\frac{d[H3]}{dt} = -k[S][H3]. \quad (4.15)$$

The terms in the above equations (4.11, 4.13, 4.14 and 4.15) can be added together to establish

$$0 = 2 \times \frac{d[I]}{dt} + \frac{d[S]}{dt} + \frac{d[H2]}{dt} + \frac{d[H3]}{dt}, \quad (4.16)$$

or, equivalently,

$$\frac{d[S]}{dt} = -2 \times \frac{d[I]}{dt} - \frac{d[H2]}{dt} - \frac{d[H3]}{dt}. \quad (4.17)$$

Taking the integral of Equation 4.17 results in the solution $[S] = -2[I] - [H2] - [H3] + C$. Since the sum of $[S]$, $[I]$, $[H2]$, and $[H3]$ is a constant, we get:

$$[S] + 2[I] + [H2] + [H3] = C = 2[I]_0 + [H2]_0 + [H3]_0, \quad (4.18)$$

$$[S] = 2[I]_0 + [H2]_0 + [H3]_0 - 2[I] - [H2] - [H3]. \quad (4.19)$$

Assuming that $[H2] = [H3]$, and $[H2]_0 = [H3]_0$, this results in:

$$[S] = 2 \times ([I]_0 - [I] + [H2]_0 - [H2]), \quad (4.20)$$

and

$$[S] = 2 \times ([I]_0 - [I] + [H3]_0 - [H3]). \quad (4.21)$$

Finally, we use Equations 4.20 and 4.21 to simplify the set of differential equations (4.11–4.15) to:

$$\frac{d[I]}{dt} = -k[I][H1] + k[S][H2] + k[S][H3], \quad (4.22)$$

$$\frac{d[H1]}{dt} = -k[I][H1], \quad (4.23)$$

$$\frac{d[S]}{dt} = 2k[H1][I] - k[S][H2] - k[S][H3], \quad (4.24)$$

$$\frac{d[H2]}{dt} = -2k[H2]([I]_0 - [I] + [H2]_0 - [H2]), \quad (4.25)$$

$$\frac{d[H3]}{dt} = -2k[H3]([I]_0 - [I] + [H3]_0 - [H3]). \quad (4.26)$$

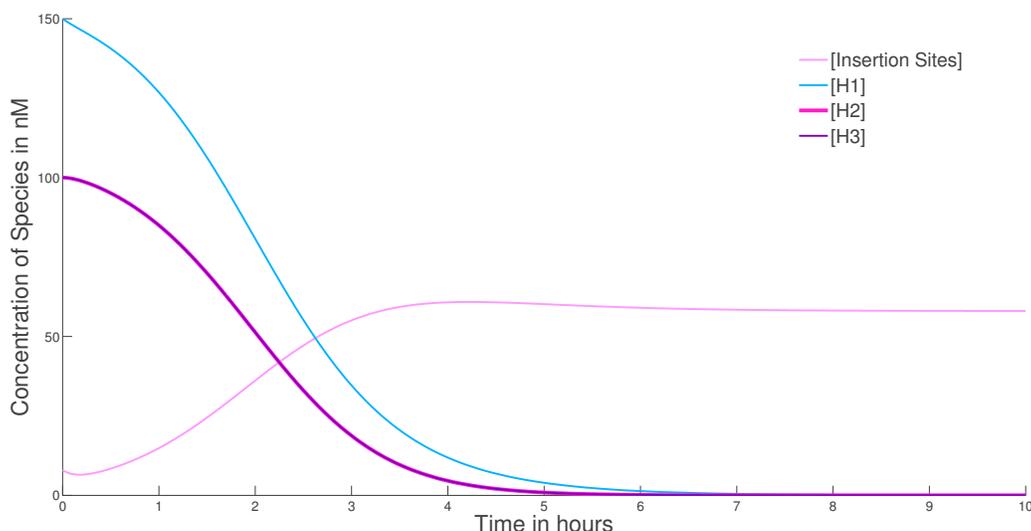


Figure 4.14: The ordinary differential equation model of the exponential growth system discussed in section 4.5 is simulated. The number of insertion sites initially decreases, as the sites interact with Hairpin 1, and then the sites begin to increase exponentially fast until a large fraction of the hairpins are consumed and the system reaches an equilibrium.

When simulated in Matlab, these equations result in a plot of the concentrations of I, H1, H2, and H3 over time. Figure 4.14 shows the simulation. The simulation starts with initial concentrations: $[I]_0 = 8 \text{ nM}$, $[H1]_0 = 150 \text{ nM}$, $[H2]_0 = [H3]_0 = 100 \text{ nM}$ as in our spectrofluorimetry experiments. We set $k = 5050 \text{ M}^{-1} \text{ sec}^{-1}$, as derived from the spectrofluorimetry data shown in Figure 4.8 and described in the next section.

The simulated concentration of H2 over time shows $[H2]_{total} - [H2]_{incorporated}$. The yellow trace in Figure 4.7 shows $[H2]_{incorporated}$ at the concentrations of molecules simulated. The simulated concentration of H2 is consistent with our measurements of H2 incorporation in the polymer. In the simulation, the number of insertion sites initially decreases, as the sites interact with Hairpin 1, and then the number of insertion sites increases exponentially fast until a large fraction of the hairpins are consumed and the system reaches an equilibrium. The growing number of insertion sites may serve as a proxy for the total concentration of polymer in our system, which we have not measured.

4.5.1 Parameter Fitting

The rate at which the number of H1 insertion sites [I] increases is

$$\frac{d[I]}{dt} = k[I][H1]. \quad (4.27)$$

We examine this rate at 10% completion time, because at that time the concentration of H1 is roughly constant, and the number of Insertion sites [I] is approximately 10% of $[H1]_0$. We substitute C for $k \times [H1]_0$ and $I(t_{10\%})$ for $0.1[H1]_0$ to get

$$\frac{d[I]}{dt} = C[I]. \quad (4.28)$$

Integrating this equation on both sides gives:

$$[I] = A \times e^{C \times t}, \quad (4.29)$$

where A is a constant determined by $[I]_0$. Thus

$$[I] = [I]_0 \times e^{C \times t}. \quad (4.30)$$

When $t = 0$, $[A] = [I]_0$. C is a measure of how quickly the number of H1 insertion sites double ($\frac{1}{C}$ is the slope of the plot comparing relative Initiator concentration to $\ln(t_{10\%})$). This value is derived as follows:

$$[I](t_{10\%}) = [I]_0 \times e^{(C \times t_{10\%})}, \quad (4.31)$$

and

$$0.1[H1]_0 = [I]_0 \times e^{(k \times [H1]_0 \times t_{10\%})}. \quad (4.32)$$

Now we can divide both sides by $[I]_0$ and take the natural log of both sides to get:

$$\ln\left(\frac{0.1[H1]_0}{I_0}\right) = (k \times [H1]_0 \times t_{10\%}). \quad (4.33)$$

We performed a linear fit on $\ln\left(\frac{0.1[H1]_0}{I_0}\right)$ and the 10% completion time on our spectrofluorimetry data (Figure 4.8). The slope of this line is $\frac{1}{C}$ or $\frac{1}{k \times [H1]_0}$ where k is the reaction rate constant. The slope of this line is 0.3667 hours or 1320 seconds. $[H1]_0$ in these experiments is 150 nM. Therefore:

$$k = \frac{1}{slope \times [H1]_0}, \quad (4.34)$$

$$k = \frac{1}{1320 \times 1.5 \times 10^{-7}}, \quad (4.35)$$

$$k = 5050 M^{-1} sec^{-1}. \quad (4.36)$$

This is the value used in our ordinary differential equation simulation discussed in Section 4.5.

4.6 Methods to Generate Other Behaviors

4.6.1 Division

Just as a polymer can grow in logarithmic time via parallel insertion, a population of polymers can be generated in logarithmic time using insertional division. Division is implemented by a complex that is identical in sequence to Hairpin 1 except that its loop has a break in it. When this complex inserts itself into a chain, the polymer splits into two. Figure 4.15 illustrates the general scheme and its implementation in DNA sequences.

We confirmed that monomer conversion is logarithmic in time at two different concentrations of Divide complexes (Figures B.13, B.15). We observe division of polymers when the Divide complex is added to the reactions six hours after initiation. We also observe short polymers when the Divide complexes are added to the solution at the beginning of the reaction, in which case they directly compete with exponential growth (Figure 4.16, 4.17, 4.18).

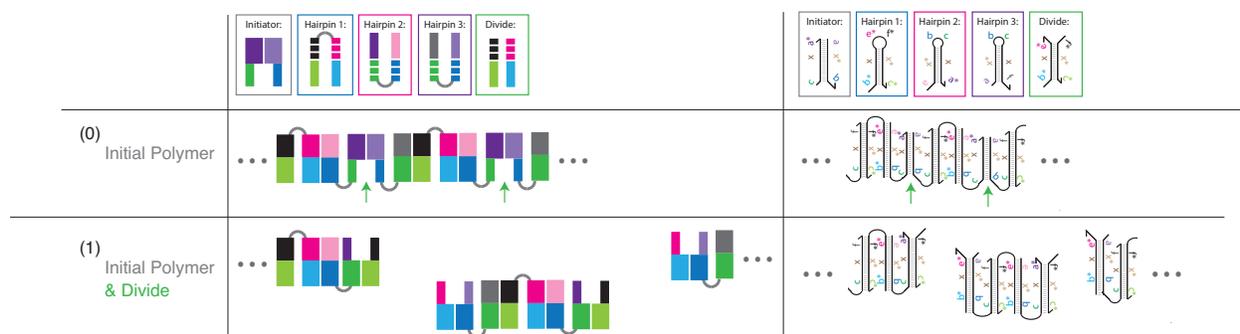


Figure 4.15: This figure depicts a system that implements division in a polymer. Each oligonucleotide is shown with color-coded motifs that correspond to the colored subsequences above. The boxes around each oligonucleotide correspond to the insertion arrows in the mechanism below, which shows the insertion of two Divide complexes. The Divide complex is identical to Hairpin 1, except that the hairpin is split between domains e^* and f^* . Note that the rest of the hairpins are the same as in Figure 4.1.

4.6.2 Treadmilling

When linear insertion is combined with end-point division, one behavior that emerges is “treadmilling”. Treadmilling is the condition in which there is growth at one end of a polymer while the other end is shrinking. Figure 4.19 shows a mechanism for treadmilling using the insertion system presented here. Note that we have not experimentally verified treadmilling. A successful implementation of this mechanism would require careful kinetic control over the insertion and division primitives. The next chapter addresses how such kinetics may be controlled via DNA sequence design.

4.7 Discussion

This work presents an advance in our ability to manipulate matter. It is part of a growing push in nanotechnology and material science toward fabricating smart materials that can be programmed to interact via molecular reactions, thus rendering them capable of being interfaced with biological compounds.

We have used molecular insertion to demonstrate the first synthetic linear polymer that grows in logarithmic time. We have presented a model in Chapter 3 that maps directly onto our molecular

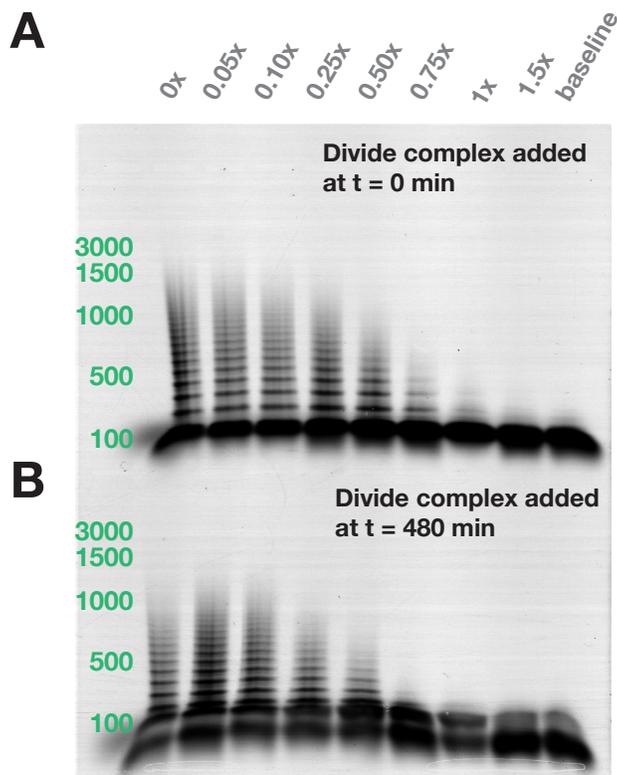


Figure 4.16: Polymer division. Super Fine Resolution Agarose non-denaturing gels of the product of a polymerization reaction with 80 nM ROX-labeled Initiator and 1 μM Hairpin 1, Hairpin 2, and Hairpin 3, to which Divide complex was added at concentrations $[D]_0 = \{0\%, 5\%, 10\%, 25\%, 50\%, 75\%, 100\%, 150\%\}$ relative to hairpin concentrations. (A) Divide complex was added with the hairpins at $t = 0$ min. (B) Divide complex was added after 6 hours of reaction. The size of polymers decreases with increased concentrations of Divide complex. See Figure B.11 for gels after staining with SYBR Gold.

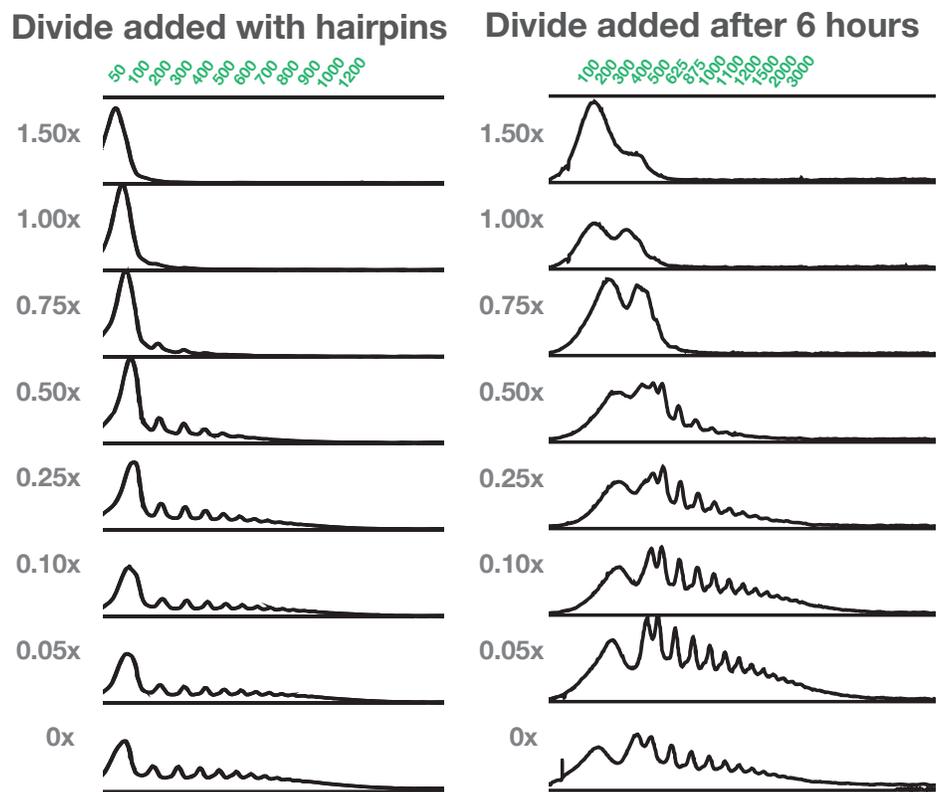


Figure 4.17: Division gel traces of the polymerization system with Divide complex added with hairpins (left) and six hours after hairpins are added (right) from Figure 4.16. Green numbers specify the size of each band, which are indicated as bumps in the gel traces.

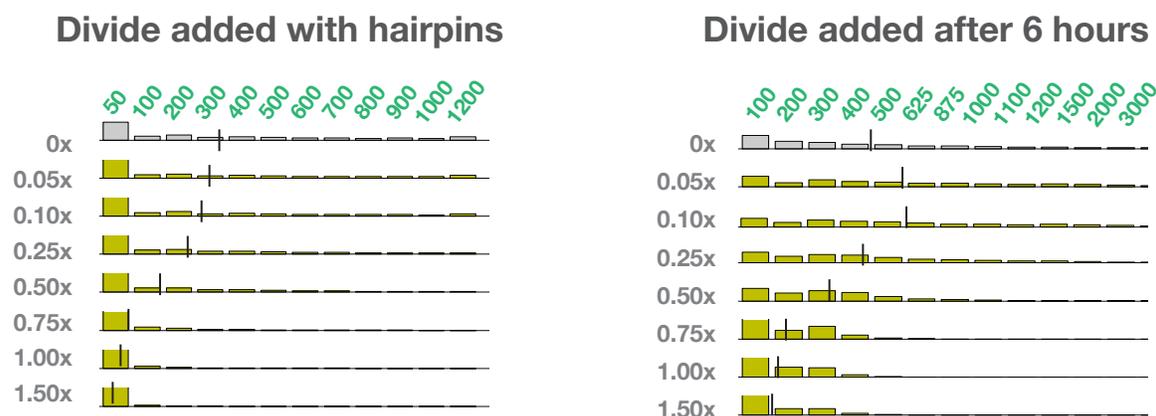


Figure 4.18: Division gel analysis of systems with Divide complex added with hairpins (left) and six hours after hairpins are added (right). The traces from Figure 4.17 were binned such that each bump in the trace was allotted to one bin. The y-axis for each distribution is 40% of the total concentration to allow the reader to clearly see the change in distribution. Vertical lines indicate the mean of the distribution. The mean of each distribution decreases with increased amounts of the Divide complex.

insertion system. This is a very powerful one-dimensional tool. It allows engineers to change the interconnections of molecules after a shape has been assembled. This is an important step toward fully reprogrammable molecular assembly. We have demonstrated three different types of behavior using a simple insertion primitive. We expect that different combinations of these actions can generate more behaviors.

Ours is not, however, the first exponentially fast growing structure ever synthesized. [Yin et al., 2008] constructed a binary molecular tree out of DNA. Their reaction begins with a root node, each node generates two child nodes in each generation of growth. [Yin et al., 2008] point out that, in the absence of steric effects, a linear increase in the number of node species will yield an exponential increase in the size of the binary tree. In practice, steric effects are always present. Our system is the first to implement parallel insertion and does not rely on adding layers to external edges for growth. This feature of our system allows the exponential growth phase of our system to last longer, as our system is not limited by cubic volume.

The next challenge will be to build reprogrammable molecular shapes in two and three dimen-

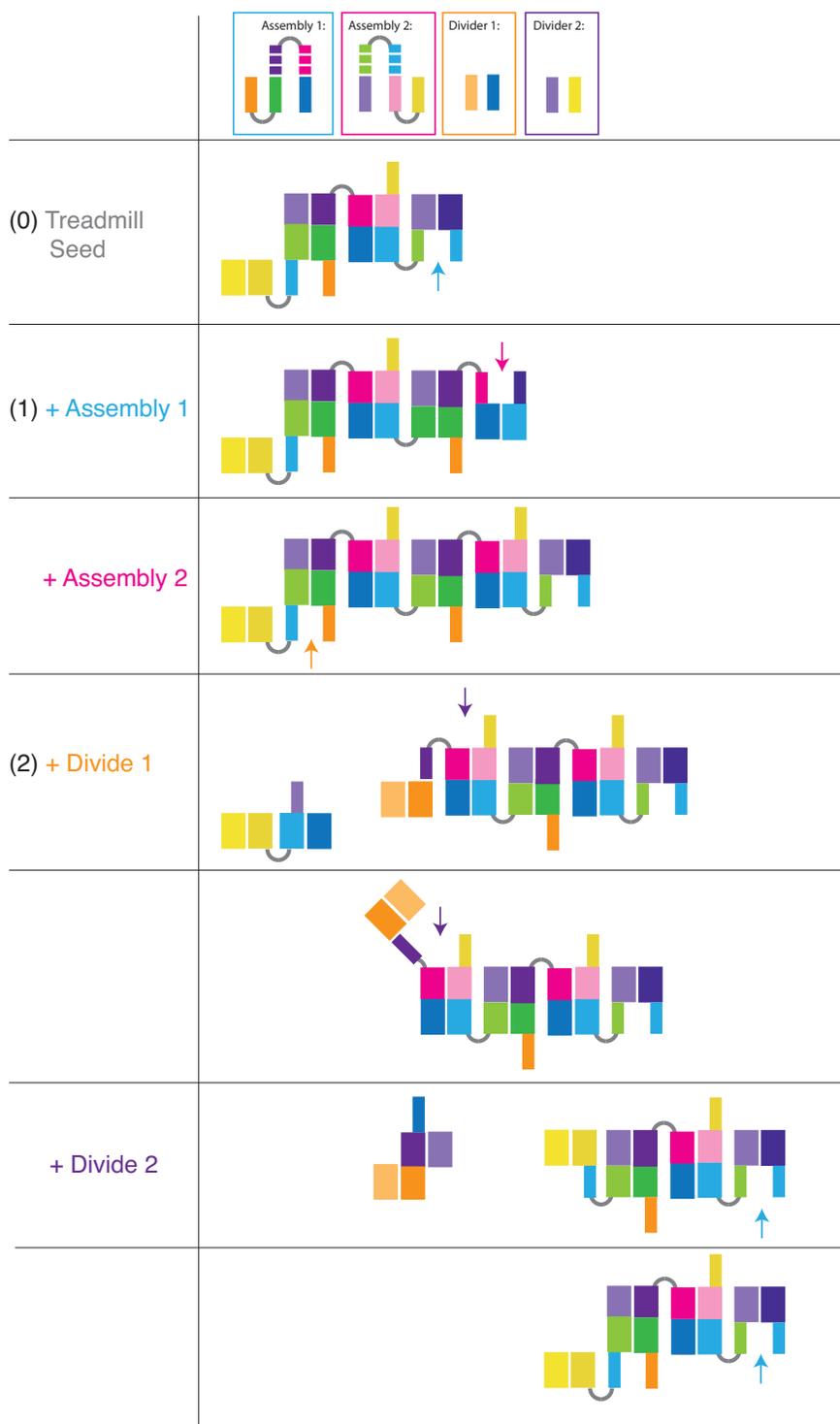


Figure 4.19: Combining the insertion and division behaviors can result in treadmilling, the growth of one end of a polymer while the other end is shrinking.

sions. Difficulties are likely to arise when scaling our current molecular system to these dimensions. Until we can precisely control the kinetics of hairpin insertion (to be discussed in the next chapter), we cannot guarantee the proper exponential growth of a shape in these higher dimensions. This is because our polymer is too flexible to accommodate insertions in multiple dimensions without the possibility of self-interactions forming a mis-shaped object. In order to generate a well-formed object using an elaboration of our system will require a more rigid structure.

A second limitation of our construction is the repeating DNA sequence utilized in the insertion and division primitives. In theory these structures can be programmed just like tiles in the tile assembly model, but in practice the repeating DNA sequence places a constraint on how many different actions can take place at a given site. The power of our system lies in its ability to grow a structure very quickly with only a few types of monomers by allowing subsets of molecules to move relative to each other. When a system like ours is scaled up its power would be limited, because Brownian motion drives these translocations only on small scales.

One may be able to extend this system by adding more complexity into the hairpin loops—additional structures or sequences that might accommodate other functionalities.

4.8 Materials and Methods

Experimental System. A typical fluorescence kinetics experiment contains Hairpin 2 labeled with a fluorophore and quencher pair on the 3' and 5' ends of the strand, respectively. Mixed together with H2 are I, H1 and either the inactivated or regular version of H3 for the linear and exponential systems respectively. H1 is added last to trigger the reaction. As H2 is integrated into the polymer, the quencher and fluorophore pair are separated from each other, yielding an increased fluorescence signal in the solution. At the end of the experiment, another strand of DNA is added into the solution in order to fully displace all unreacted hairpins (Figure 4.4). This “displacement” strand was added in $> 50\times$ excess to the concentration of H2RQ to ensure that the reaction quickly goes to completion. We use the final fluorescence level to normalize our fluorescence signals. Baseline reactions contain only I, H2 and H3, until the end of the experiment at which point the displace strand is added.

DNA Sequences and Design. The sequences presented in Supplementary Tables B.1 are based on those used in a previous insertional polymerization motor [Venkataraman et al., 2007]. These sequences were designed using the NUPACK web application [Zadeh et al., 2010, Zadeh et al., 2011] and our in-house DNA Design software package [Winfree, 2012] to minimize the presence of any unanticipated secondary structures that might interfere with the kinetics under investigation.

Buffer Conditions. DNA oligonucleotides were stored in $1\times$ SPSC buffer (50 mM Na₂HPO₄ pH 6.5, 1 M NaCl) at 4°C directly preceding experiments. All experiments and purifications were performed at 25°C.

Annealing. All annealing processes were performed with an Eppendorf Mastercycler Gradient thermocycler. The samples were brought down from 95°C to 16°C at a constant rate over the course of 90 min.

Snap Cooling. All Hairpins were snap cooled prior to experiments. This protocol entails heating the strand solution to 90°C for 5 min, then immediately putting solutions on ice for 45 min. This protocol encourages intramolecular hydrogen bonding of the hairpins.

Substrate Purification. DNA oligonucleotides used in this study were purchased from Integrated DNA Technologies (IDT), with standard desalting purification, except for strands with a quencher, fluorophore or a 5' toehold involved in the four-way branch migration, which were purchased with HPLC purification.

Concentrations of individual strand stocks were determined from the measured absorbance at 260 nm using a Nanodrop Biophotometer and using calculated extinction coefficients that account for hypochromicity effects in double-stranded DNA [Tataurov et al., 2008].

Initiator and Divide complexes were further purified by nondenaturing (ND) polyacrylamide gel electrophoresis (PAGE) as follows: Strands for each sample were prepared with nominally correct stoichiometry at 10 nM and annealed. The acrylamide (19:1 acrylamide:bis) was diluted from 40% acrylamide stock (Ambion). ND loading dye (containing Bromphenol Blue in 50% glycerol) was added to all samples, achieving a final glycerol concentration of 10% by volume. The samples were then run on 12% ND PAGE at 150 V for 6 hours. Gels were run at room temperature ($\approx 25^\circ\text{C}$). The band corresponding to the Initiator size was cut out and eluted in 1 mL of $1\times$ SPSC buffer for 2 days. Purified complexes were quantitated by measurement of absorbance at 260 nm

using an Eppendorf Biophotometer and calculated extinction coefficients as above.

Gel Assays. Combinatorial gels were run using 12% polyacrylamide and concentrations of all species at 100 nM. Solutions were left to react for 6 hours, then run in an XCell SureLock Mini-Cell Electrophoresis vertical gel box at 150V for 1 hour in TBE running buffer. After a gel was run, it was stained with SYBER Gold dye and imaged using an FLA-5100 fluorescent scanner (Fujifilm Life Science). Time Lapse, Final Value and Divide gels were run in 2% Super Fine Resolution Agarose (from AMRESCO) on a Thermo Scientific Owl Horizontal Gel box. In these experiments, the Initiator is tagged with a 3' ROX fluorophore on one strand. Thus each properly-initiated polymer has a single ROX tag. Time Lapse reactions contained the following concentrations of species [I] = 80nM, [H1] = 1.5 μ M, [H2] = 1 μ M, [H3] = 1 μ M. Final Values reactions contained the following concentrations of species [I] = 0 nM, 10 nM, 20 nM, 40 nM, 80 nM, 160 nM, 320 nM, 640 nM, 1 μ M; [H1] = 1.50 μ M; [H2] = 1 μ M; [H3] = 1 μ M. Divide reactions contained the following concentrations of species [I] = 80nM; [D] = 0 nM, 50 nM, 100 nM, 205 nM, 500nM, 750nM, 1 μ M, 1.5 μ M; [H1] = 1 μ M; [H2] = 1 μ M; [H3] = 1 μ M.

Atomic Force Microscopy. Atomic Force Microscopy images of polymer taken with 10% Initiator (10 nM) relative to Hairpin (100 nM). 50 μ L of 1 \times TAE 12.5 mM Mg⁺⁺ was deposited on mica (from Ted Pella), followed by 1 μ L of 5 mM Nickel Acetate and 2 μ L of 500 nM polymer sample after 5 hours of reaction. The sample was then imaged using a VEECO Nanoscope III with a vertical engage J-scanner.

Spectrofluorimetry Studies. Spectrofluorimetry studies were done using a SPEX Fluorolog-3 (Horiba) with external water bath and 1.6 mL synthetic quartz cells (Hellma 119-004F). The excitation was at 584 nm, while emission was at 604 nm. In all spectrofluorimetry experiments, the total reaction volume was 1.5 mL, the temperature was 25°C, and 2 nm band-pass slits were used for both excitation and emission monochrometers. Experiments were conducted with an integration time of 10 seconds at 60 second intervals. Prior to each experiment, all cuvettes were cleaned as follows: each cuvette was rinsed 15 times in Milli-Q water, 5 times in 70% ethanol, another 15 times in Milli-Q water, and finally once more in 70% ethanol and then Milli-Q water. For the slit size, concentrations, and integration times used, no measurable photobleaching was observed. Exponential and linear reactions contained the following concentrations of species [I] = 0 nM, 1

nM, 2 nM, 4 nM, 8nM, 16nM, 32 nM, 64 nM, 100 nM; [H1] = 150 nM; [H2] = 100 nM; [H3] = 100 nM. Divide reactions contained the following concentrations of species [I] = 10nM, 25 nM; [D] = 0 nM, 1 nM, 16 nM, 100 nM; [H1] = 150 nM; [H2] = 100 nM; [H3] = 100 nM.

Fluorescence Normalization. Fluorescence is normalized so that one normalized unit of fluorescence corresponds to 1 nM of unquenched fluorophore-labeled strand reporter 2. This normalization is based on the fluorescence levels of annealed samples with a minimal fluorescence measurement taken of the diluted Reporter complex before the experiment was initiated, and a maximal fluorescence value that is extracted from a biexponential fit of the data taken at the end of the experiment, after the displacement strand is added to displace all unreacted fluorophore-quencher pairs.