Evolutionary Processes in Replicating DNA Species

Dissertation

zur Erlangung des Grades
Doktor der Naturwissenschaft (Dr. rer. nat.)
an der Fakultät für Physik
der Ludwig-Maximilians-Universität München

vorgelegt von
Georg Urtel
aus Hallein

München, 2016

Erstgutacher: Prof. Dr. Dieter Braun
Zweitgutachter: Prof. Dr. Erwin Frey
Datum der mündlichen Prüfung: 25.07.2016

Zusammenfassung

Eine wichtige Triebfeder der Darwinschen Evolution verschiedener Spezies sind deren gegenseitige Wechselwirkungen innerhalb eines Ökosystems. Das Studium dieser Ökosysteme erlaubt es daher, deren zugrunde liegende evolutionäre Mechanismen zu verstehen. Dies wird allerdings oft durch ihre Komplexität oder die lange Lebensdauer mancher Spezies erschwert. Modellsysteme können dabei helfen diese Probleme zu überwinden indem sie die Kontrolle entscheidender Parameter wie etwa der Konzentration von Ressourcen oder Degradationsraten erlauben. In dieser Arbeit werden DNA-Moleküle die sich mit Hilfe von Enzymen replizieren als Modellsystem genutzt um zwei evolutionäre Phänomene zu beleuchten: Kooperation in der frühen Evolution und zyklische Inhibierung verschiedener Spezies.

Im ersten Teil dieser Arbeit wird die Kooperation zwischen DNA-Spezies untersucht. Dazu wurden zwei Spezies benutzt die eine Hairpin-Sekundärstruktur bildeten, weshalb sie in einer Polymerase-Kettenreaktion nur langsam replizierten. Eine Selektion von schnelleren Replikatoren durch Verdünnungen führte zum Aussterben dieser Hairpins. Hatten sie allerdings die gleiche Loop-Sequenz, kooperierten sie und erzeugten eine Kreuzung welche keine Sekundärstruktur aufwies und daher die Verdünnungen überlebte. Das Umkehren der Reaktion erlaubte es die Hairpins wieder zu erzeugen und zeigte, dass deren Information erhalten blieb. Die Experimente demonstrierten wie Fitness und Komplexität in einem Oligonukleotidpool durch eine Kreuzung gesteigert werden können.

Der zweite Abschnitt beschäftigt sich mit der zyklischen Inhibierung von Spezies in einem Schere-Stein-Papier Spiel. Dabei verfolgen drei Spezies eine vererbbare Strategie, welche zur Inhibierung einer anderen Spezies führt: Papier wickelt Stein ein, Stein schleift Schere und Schere zerschneidet Papier. Eine Möglichkeit dies auf molekularer Ebene umzusetzen ist die DNA-Toolbox, welche es erlaubt biochemische Reaktionsnetzwerke zu erstellen, in denen die Spezies DNA-Moleküle sind welche sich mithilfe von Enzymen replizieren. Zweidimensionale Simulationen mit einem Reaktions-Diffusions-Modell der DNA-Toolbox zeigten phasenverschobene Oszillationen und Spiralwellen. Durch räumliche Heterogenitäten konnten die dynamischen Muster beeinflusst werden. Für die experimentelle Umsetzung wurde ein Sequenz-Screening durchgeführt, dessen vorläufige Ergebnisse hier präsentiert werden. Die Erstellung des vollen Systems würde es ermöglichen die Bedingungen für die zeitlichen und räumlichen Oszillationen experimentell zu bestimmen und deren Robustheit gegenüber Störungen zu testen.

Die DNA-Toolbox erlaubt es auch andere biologische Prozesse zu untersuchen, wie etwa die Regulierung genetischer Netzwerke. Im dritten Abschnitt wurde sie genutzt um einen experimentell erzeugten Morphogen-Gradienten zu interpretieren. Dazu reagierte eine autokatalytische DNA-Spezies mit dem Morphogen-Gradienten in einem Reaktions-Diffusions-Prozess. Es bildete sich eine Wellenfront, welche durch den Gradienten verlangsamt und lokalisiert wurde. Messungen der Position, Geschwindigkeit und Breite der Wellenfront wurden durchgeführt. Ähnliche Prozesse sind in der Embryogenese von Lebewesen zu finden, bei der der Morphogen-Gradient an genetische Netzwerke gekoppelt ist.

Zusammengefasst zeigt sich, dass DNA und Enzyme nützliche Modellsysteme sind um komplexe biologische Phänomene zu untersuchen. Der *bottom-up* Ansatz erlaubt es unterschiedliche Aspekte wie Kooperation oder Embryogenese mit ähnlichen Techniken zu erforschen.

Abstract

The interaction of species in ecosystems is one of the main driving forces of Darwinian evolution. Observing the development of ecosystems allows to draw conclusions about the underlying evolutionary mechanisms. However, such observations are typically difficult, because of the complexity of ecosystems and the long lifetime of some species. Simple model systems that allow the precise control of key parameters, e.g. resource concentrations and degradation rates, can overcome this problems. Here, DNA molecules that replicate with the help of an enzymatic machinery were chosen as model systems to look at evolution from two different perspectives: cooperation of species in the early earth and cyclical inhibition of species.

In the first part of the thesis, cooperation between the DNA species was studied. Two different species were designed to have hairpin secondary structures that slowed down replication during a polymerase chain reaction. Selection for fast replicators by serial dilutions resulted in an extinction of the hairpin species. However, when the two hairpin species only differed in their stems, but had the same loop sequence, they cooperated and formed crossbreeds. The crossbreeds lacked the secondary structure and survived the dilutions. Reversing the reaction resulted in reemergence of the hairpin species, demonstrating that the information was preserved. The presented results show how fitness and complexity of oligonucleotide pools on the early earth could have been increased by a simple crossbreeding step.

In the second part, cyclical inhibition of species in a rock-paper-scissors game was studied. In such a game, three species play an inheritable strategy, where each species inhibits another species: paper wraps rock, rock crushes scissors, and scissors cut paper. One possibility to implement such a game on the molecular level is the DNA-toolbox: a framework to build biochemical reaction-networks where short DNA molecules are replicated by a polymerase and a nickase. Two-dimensional simulations with a reaction-diffusion model of the DNA-toolbox resulted in phase-shifted oscillations in time and characteristic chiral waves. Addition of spatial heterogenities to the simulations allowed to modify the the dynamic patterns and even could arrest them. For an experimental realization, a sequence screen to find suitable DNA species was performed and preliminary results are presented. Accomplishing the formation of the full system offers the opportunity to experimentally test the conditions needed for spatio-temporal oscillations and to probe their robustness to perturbations.

The established DNA-toolbox can be used to study other complex biological processes, e.g. regulation of genetic networks. In the third part of the thesis, the DNA-toolbox is applied to interpret an experimentally created morphogen gradient. A self-replicating species reacts with a morphogen gradient in a reaction-diffusion process. This resulted in the formation of a traveling wavefront that localized in the gradient. Measurements of the positions, velocities and widths of the wavefronts are presented and analyzed. In biology, similar processes are found in embryogenesis, where a morphogen gradient is coupled to a genetic network.

Taken together, model systems built of DNA and enzymes are useful tools to study complex biological phenomena. The *bottom-up* approach allows to tackle diverse topics like cooperation or embryogenesis with similar techniques.

Contents

Part I

Extincting Molecules Preserve their Information via Cooperation

1 Motivation and Goals

The central dogma of molecular biology states that DNA is the long term information storage within cells, which gets transcribed to RNA, which gets translated to proteins by ribosomes [1]. In this processes as well as in DNA replication, a lot of proteins are involved. This leads to a chicken and egg problem for an explanation for the origin of life. Proteins need DNA and DNA needs proteins. What came first? A possible answer was given when it was found that RNA cannot only be used to store information, but also catalyzes reactions [2, 3]. Thus, it is a reasonable assumption that life originated from replicators which were made from RNA or similar biopolymers, acting as information storage and catalysts [4–6]. In order to make the transition from simple chemistry to the precursors of life, several stepping stones have to be reached. Over the past years, a lot of experiments and theoretical studies showed that such a transition is possible:

1. There are several known ways of creating nucleotides and nucleobases from prebiotically plausible molecules [7, 8].

2. Concentration of this monomers can be increased through adsorption to surfaces, wet-dry cycles, eutectic ice phases or thermal traps, allowing the formation of oligomers [9–14].

3. A way to pass on sequence information without the help of ribozymes has to emerge [15–17].

4. RNA molecules have to develop catalytic activities, which allow a more efficient transmission of information [18–21].

Ribozymes like the polymerase described by *Attwater et al.* need binding sites in order to base-pair with their RNA substrates [21]. An exponential template directed replication of a ribozyme has, in general, to be a two step process: First, the templates information is copied to its complement, which is then used to create the copy of the initial template in a second step (an exception would be palindromic sequences, which would drastically restrict the sequence space for functional RNA)[22]. In order for such a replication process to take place, the initial template as well as its complement need a binding site for the ribozyme that replicates them, otherwise replication will be a linear process.

The spontaneous emergence of two different binding sites which can be targeted by a single ribozyme seems to be an unlikely event. However, the binding sites do not necessarily have to be different. If a template contains the complementary sequence of its binding site, the product created can be used as template as well. As a consequence, the strand has to form a hairpin, which blocks the binding site and slows down replication by inhibiting binding of the polymerase. The question is, whether there can be a transition from such a slow replicating, but simple hairpin to a more complex, faster replicator, which has two binding sites.

To answer this question experimentally, DNA species were used which are chemically very similar to RNA. Instead of a ribozyme that facilitates their replication, *polymerase chain reaction* (PCR) is used. The enzymatically driven replication allows to associate changes in fitness to the details of primer

binding sites and to the secondary structures of the used templates. The thermal cycling and thus repeated annealing and melting of oligonucleotides is a prebiotically plausible mechanism [23–25]. To select for fast replicators serial dilution transfers were made, where a fraction of the DNA species was added to fresh replication medium (polymerase, nucleotides, primer, buffer). A crossbreeding step facilitated through cooperative sequence mixing of two related hairpin species allowed the creation of faster replicators that lack the secondary structure. This crossbreed species replicated faster and preserved the sequence information of the hairpins, which went extinct under sufficiently strong dilution conditions.

2 Theoretical Considerations

Linear and exponential replication have completely different replication behaviors. If degradation of the molecular species is added to the model, only exponential replicators can survive.

Linear Replication

In template directed replication processes, linear replication is observed if a replicate cannot be used as a template. In PCR experiments this was realized using linear DNA strands which have only one primer (see Section 3). If there is any degradation process present in the system, this type of replication will always result in extinction because the templates cannot be replaced. If the last template is degraded, it is just a matter of time until the last replicates vanish. This is shown with a simple calculation, where the time evolution of the total concentration c of template and product is described by:

$$\frac{dc}{dt} = \frac{dc_P}{dt} + \frac{dc_T}{dt}$$

where c_P is the product concentration and c_T is the template concentration. With a replication rate $r > 0$ and a degradation rate $\gamma > 0$, the time evolution can be written as

$$\frac{dc_P}{dt} = -\gamma \cdot c_P + r \cdot c_T \text{ and } \frac{dc_T}{dt} = -\gamma \cdot c_T$$

The equation for the template concentration has a trivial solution $c_T = c_{T,t=0} \cdot e^{-\gamma \cdot t}$. Pluging all this into the inital equation gives

$$\frac{dc}{dt} = -\gamma \cdot c_P + (r-\gamma) \cdot c_T = -\gamma \cdot c_P + (r-\gamma) \cdot c_{T,t=0} \cdot e^{-\gamma \cdot t}$$

For large times t, the last term will vanish and thus exponential degradation of products will be the dominating process in the system:

$$\lim_{t \to \infty} \frac{dc}{dt} = -\gamma \cdot c_P$$

The dynamics of the system can be interesting, when $r > \gamma$ and the initial template concentration is sufficiently high. In this case, there can be an intial increase in concentration for an extended period of

time. However, after a while the template concentration will decrease more and more until replication becomes very slow and degradation dominates the system.

The linear growth process was described with a simple model and was in good agreement with experiments (Section 3.2). The concentration after $j+1$ PCR cycles is given by $c_{j+1} = c_j + c_T \cdot \varepsilon_{linear}$ with the growth parameter ε_{linear}. The degradation in the experiments is not continuous, but a stepwise process. In the model this is implemented by dividing the concentrations by the dilution factor d after a certain number of cycles.

Exponential Replication

For exponential replication, the products are templates at the same time and the time evolution is described by the simple differential equation for the concentration c: $\frac{dc}{dt} = -\gamma \cdot c + r \cdot c$. Thus, the trivial solution $c = c_{t=0} \cdot e^{(r-\gamma)\cdot t}$ implies that a species survives when the replication rate exceeds the degradation rate $r \geq \gamma$. Another way to describe the evolutionary success of a species is to use a fitness value. This approach is more appropriate, since the degradation is not a continuous process in the experiments, but done by stepwise dilutions. The time t' between two dilutions is defined as the time for a generation and the fitness is the ratio of the concentrations from one generation to the next $F = c(t+t')/c(t)$. To prevent extinction, a species fitness has to fulfill $F \geq 1$. The stepwise dilution is described by a dilution factor d, i.e. the concentration is reduced to c/d. With the PCR efficiency $\varepsilon = r \cdot \tau$ which is a measure for the concentration increase after one cycle of length τ. The fitness then becomes:

$$F = \frac{e^{\varepsilon \cdot m}}{d} \tag{1}$$

where m is the number of cycles between two dilutions.

Like for the linear replication, a simple model is used to describe the replication process for small ε. The update rule for the growth process is then $c_{j+1} = c_j \cdot (1+\varepsilon)$ and the dilution is implemented as described above for the linear model.

3 Experimental Realization

As a model system, DNA molecules with a hairpin structure were used, which replicate with PCR. The sequence design allowed the formation of crossbreeds, which lack the secondary structure and replicate faster.

3.1 Materials and Methods

All oligonucleotides were purchased from biomers.net GmbH (Ulm, Germany) with HPLC purification. A list of the used sequences can be found in Appendix A, Table 3. Primers were labeled with the fluorescent dyes Cy3 and Cy5 on their 5'-ends to determine the concentrations of single species on polyacrylamide (PAA) gels. Primers were added with a concentration of 250 nM per primer. The gels

were 12.5 % PAA gels (19:1) with 50 % urea (from Carl Roth GmbH + Co. KG, Karlsruhe) and were run at 400 Volts for typically 15 minutes. Longer runtimes were sometimes used for better separation of longer strands, but primers were dragged out of the gel in this case. Before gel loading, the samples were mixed with a 2x loading buffer (95 % formamide, 5 % glycerol, 50 µM EDTA and 50 mg/l Orange G), heated to 95 °C for three minutes and quenched in iced water. Images of gels were taken in a lab-built geldoc. To obtain concentrations of replicated DNA from the gels, a band of interest was integrated and divided by the intensity of the complete lane including the primers. The fraction multiplied with the primer concentration gives the concentration of a specific band.

For the PCR, three different polymerases were used, with three different buffers: Phusion High Fidelity DNA Polymerase was used in a final concentration of 8 U/ml in 1x HF Buffer (all from New England Biolabs). It was used for all experiments, except stated otherwise. Deep Vent (exo-) DNA Polymerase with a final concentration of 20 U/ml was used in ThermoPol Reaction Buffer. A Taq-polymerase was used with the QIAGEN Fast Cycling PCR Kit. NTP concentration was 0.2 mM in all reactions and 1x EvaGreen (from Biotum) was added. The DNA species were always added as double strands. For thermal cycling and fluorescent readouts a C1000 thermal cycler with a CFX96 (Touch) Real-Time PCR Detection System from Biorad was used. PCR protocols were 2-step protocols and consisted of initial denaturation at 95 °C for 1-3 minutes, denaturation steps for 10 s at 95 °C and annealing for 30 s for 67 °C (the temperature is indicated in the text or graphs if it deviates). The total time for a temperature cycle and the readout was measured to be $\tau = 72s$. The lid temperature was set to 105 °C to prevent condensation of the reaction mix. The errors for fitness and efficiency values were calculated assuming a pipetting error of 10 %. Error bars in figures were standard deviations from experiments which were conducted more often.

Serial Dilution Transfers

To challenge the replicators in the experiments, serial dilution transfers were performed. After a certain replication time t' the samples were cooled down to 1 °C in the cycler or quenched on ice. A part of the reaction mixture was transferred to fresh resources (polymerase, primers, NTPs) and used for further rounds of replication (Figure 1a). To achieve $d = 5$ as dilution factor, 1 µl of the product was added to 4 µl of fresh medium. The replicated DNA which was not transferred was used for gel electrophoresis. Thus, all shown concentrations from experiments with dilutions correspond to values right before the dilution. To measure how accurate dilution factors were, 5 µl of a PCR mix with labeled primers was prepared and cycled for 30 cycles. 1 µl of the product was then transferred to fresh resources, mixed and transferred two more times (Figure 1b). It was found that the desired dilution factor deviates from the measured dilution factor by 20 %. The deviation was caused by the low volume and the high viscosity of the transfered reaction solution.

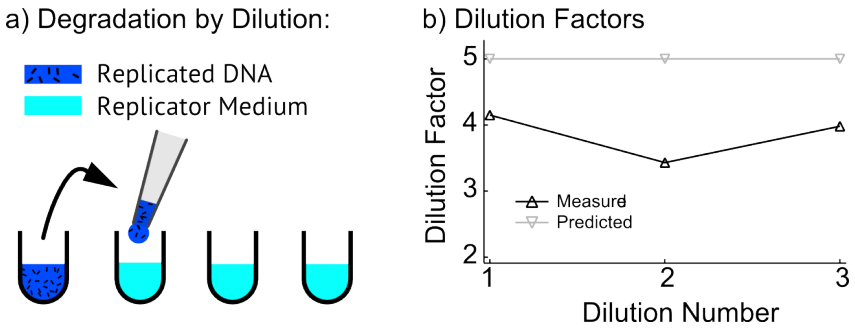


Figure 1: Serial Dilution Transfer: a) After several PCR cycles, the mixture was diluted. A portion of the replicated DNA was transferred to fresh resources. b) A PCR product was diluted in three subsequent steps. The measured dilution factor deviated from the predicted dilution factor by 20 %.

3.2 Three Different Replication Modes

In this section, three different types of replication will be introduced, which depend on the sequence design and the number of primers added to the system.

1. Linear replication can be observed when a DNA strand without secondary structures is supplied with one primer. The replicate produced by the polymerase lacks a binding site for the primer and cannot be used as template (Figure 2a).

2. Hairpins can replicate exponentially with one primer, since a new primer binding site is created in the replicate. Thus, the product can be used as template. However, the secondary structure leads to inhibited primer binding and slow replication (Figure 2b).

3. Fast exponential replication takes place when a DNA strand without secondary structure is supplied with two primers. The replicates can then be used as templates and there is no inhibition due to secondary structures (Figure 2c).

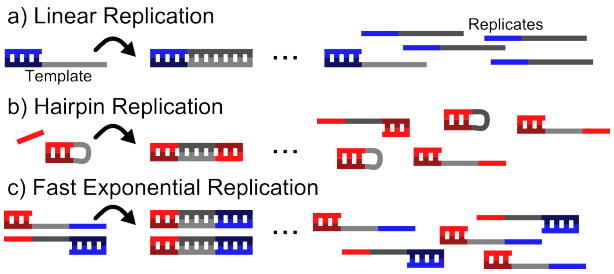


Figure 2: Three Modes of Replication: a) When a primer is elongated on a linear template, a replicate is produced. After a while, the replicates accumulate, but cannot participate in further reactions. The template concentration remains unchanged. b) When a hairpin is replicated, a new primer binding site is produced and the replicate can be used as template. This allows exponential replication, but the secondary structure slows down the process. c) Fast exponential replication takes place when a linear template is supplied with two primers.

Linear Replication

To explore linear replication, a linear DNA strand was used as a template and supplied with a single primer (sequences can be found in Tables 3 and 4 in Appendix A) and PCR was conducted for several cycles. As described in Section 2 the increase in concentration depended on the template concentration (Figure 3a). In every cycle 18 % of the templates were used for replication. If the linear replication was challenged by exponential dilution steps, where after six PCR cycles a dilution with a factor of $d = 5$ was made, the population decreased exponentially (Figure 3b). Weaker dilutions or more rounds of replication between them might allow the linear species to survive longer and even to temporarily reach higher concentrations, but extinction cannot be avoided. However, such a replicator only has one primer binding site, keeping it rather simple. If in an RNA world a machinery emerges which can replicate other strands by a mechanism that needs such a binding site, linear replicators have the advantage of having an accessible binding site. Assuming a higher mutation rate for such ribozymes compared to the Phusion High Fidelity DNA Polymerase which has a 3'-5'-proof-reading function, the change of some base pairs might lead to the evolution of hairpin motifs, which already allow exponential replication, as described in the next section.

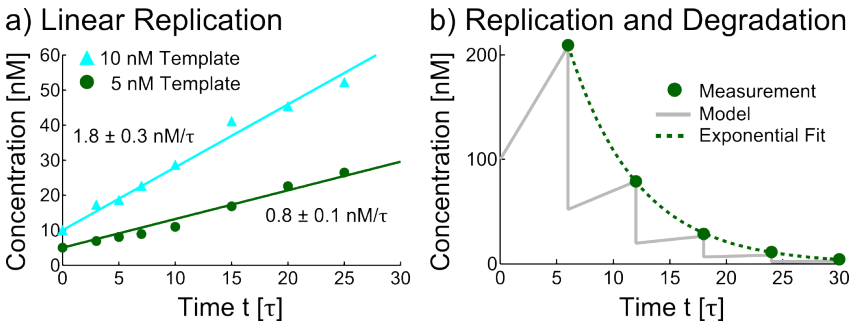


Figure 3: Linear Replicators: a) The concentration increase for linear replicators depended linear on the template concentration. The indicated values are the slopes of the fits and $\tau = 72\,s$ is the time for one cycle. b) When the reaction solution was diluted, linear replicators went extinct exponentially. A look at the model (gray curve, model described in Section 2) shows that the slope decreased after every dilution, because the template concentration decreased. For the model $d = 4$ and $\varepsilon_{linear} = 0.18$ was used with an initial template concentration of 100 nM.

Exponential Hairpin Replication

Hairpin motifs in replicators can emerge from mutations of initially linear species as described above or can emerge from elsewhere, e.g. because they are protected from cleavage and thus have a lower effective degradation rate [16]. The structure has two characteristics: A double stranded stem and a single stranded loop. In a PCR reaction, the stem inhibits the binding site and prevents primer binding, but when replication is successful a new binding site is created on the replicate, making it a template. This creation of new templates allows exponential replication and gives the hairpin replicators the possibility

to survive dilutions when the fitness $F \geq 1$.

A short hairpin with a 18 bp stem and a 48 nt loop wasused to observe exponential growth with a single primer (Figure 4a). Experiments with a dilution factor of $d = 5$ after 6 PCR cycles led to extinction of the hairpin (Figure 4b). However, starting from lower initial concentrations compared to the linear replicator in Figure 3b, the hairpin still had a higher concentration after 30 PCR cycles. The graph also shows, that the model with $d = 3.8$ and $\varepsilon = 0.14$ is in good agreement with the data and the fitness $F = 0.62^{+0.05}_{-0.07}$ obtained from an exponential fit. Independent measurements of the efficiency without dilutions gave ε=0.12 (Figure 4a).

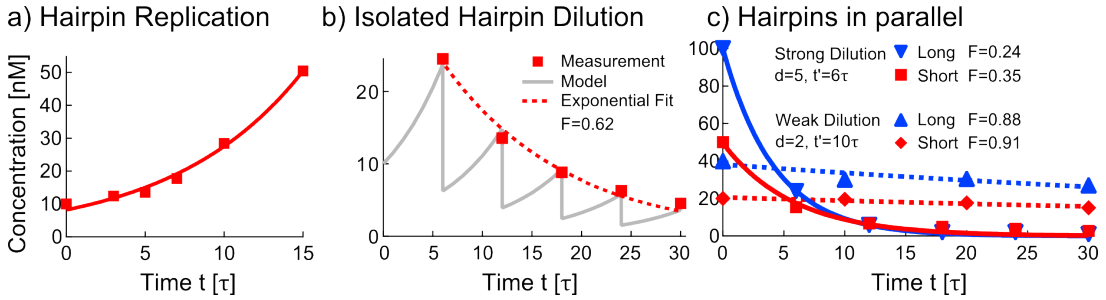


Figure 4: Hairpin Replicators: a) A hairpin supplied with one primer already replicated exponentially. The time for one cycles was $\tau = 72s$ and the fitted efficiency was ε=0.12. b) The short hairpin species went extinct under strong dilution conditions ($d = 5, t' = 6\tau$). The result is in good agreement with the simple model described in Section 2 (grey curve). The fitness value was extracted from an exponential fit (red curve). c) Parallel replication of the long hairpin (blue) and the short hairpin (red) led to extinction under strong (solid line) and weak dilutions (dashed line).

Furthermore, a long hairpin with a 38 bp stem and a 48 nt loop was designed to observe parallel replication of two species without any interactions. However, it was found that the short hairpin replicated more slowly in the presence of the long hairpin and only reached a fitness of $F = 0.35^{+0.07}_{-0.06}$ under the same replication and dilution conditions as described above. For the long hairpin such a dependency was not found. Upon mixing the hairpins and using the same dilution conditions as in Figure 4b, both hairpins went extinct (Figure 4c, solid lines). The long hairpin had an even lower fitness of $F = 0.25^{+0.05}_{-0.04}$.

Using weaker dilution conditions with a dilution factor of $d = 2$ after 10 PCR cycles also led to extinction of both species, but they vanished much slower (Figure 4c, dashed lines). The fits gave fitness values of $F = 0.88$ for the long hairpin and $F = 0.91$ for the short hairpin. From equation (1) follows that $\varepsilon = ln(F \cdot d)/m$. With this, for the long hairpin the efficiency is $\varepsilon = 0.04$ for the strong dilution and $\varepsilon = 0.06$ for the weak dilution. This is in good agreement with independent measurements without dilution, where $\varepsilon = 0.05$ was obtained. However, for this values the dilution rate was assumed to be exactly the projected rate without any errors. For the short hairpin the efficiencies vary more: $\varepsilon = 0.09$ for the strong dilution and $\varepsilon = 0.06$ for the weak dilution. This is lower than the values of $\varepsilon = 0.12$ and $\varepsilon = 0.14$ mentioned above. However, this is in agreement with the observation that the presence of the long hairpins lowers the fitness of the short hairpin. Especially in the case of weak dilutions the

concentrations of long hairpins remained relatively high, possibly leading to a stronger inhibition of the small hairpins.

Fast Exponential Replication

As described, hairpin replication was slowed down by the secondary structure, whereas linear DNA strands supplied with one primer went extinct due to their vanishing templates. This was prevented by adding a second primer, which allowed to create new templates. This led to uninhibited, exponential replication until product inhibition and primer depletion slowed down the process (Figure 5a). Figure 5b shows that under the same dilution conditions used in Figures 3b and 4b the fitness was $F = 1.81 \pm 0.08 > 1$. For the simple model described in Section 2, the values $d = 2.9$ and $\varepsilon = 0.275$ were used. However, the dilution factor strongly deviated from the intended dilution factor (see Section 3.1). The fast exponential replicator was the only species which survived the given dilution conditions and exceeded the fitness of the fastest hairpin replicator by a factor of three. However, the replicator was more complex and needed two binding sites to initiate replication.

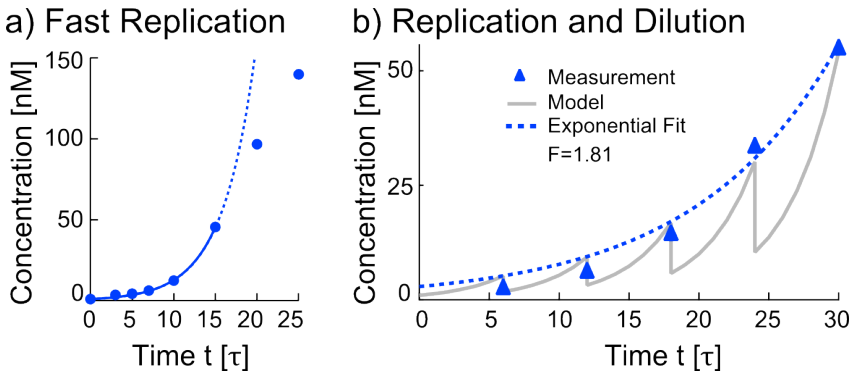


Figure 5: Fast Exponential Replication: a) Due to product inhibition and primer depletion, the fast exponential replicator slowed down after roughly 15 cycles ($c_{t=0} = 1\,nM$) and the exponential fit deviated from the data. The fit gave an initial efficiency of $\varepsilon = 0.290$ b) Even under strong dilution conditions ($d = 5, t' = 6\tau$) the replicator survived. The fitness was $F = 1.81 \pm 0.08 > 1$ and the parameters for the model (gray curve) were $d = 2.9$ and $\varepsilon = 0.275$.

3.3 Crossbreeding of Hairpins

The fitness advantage of the fast exponential replicator is obvious, but for a origin of life scenario the hairpin molecule seems to be a better choice due to its lower complexity, which only needs one binding site to initiate replication. However, two cooperating hairpin replicators which differ in their stem sequence, but have the same loop information, can create a fast exponential replicator through a crossbreeding step (Figure 6a). This cooperative process is initiated by the incomplete replication of one of the hairpins, for example due to a temperature increase which causes dehybridization from the template. A halfway elongated primer is produced, which is further elongated in the next round of replication.

The similar loop sequences allow to finish the replication using the other hairpin species as a template. In this case, a new crossbreed species emerges, which has an intermediate length (long hairpin 124 nt, short hairpin 84 nt, crossbreed 104 nt) and is able to conduct fast exponential replication as described in Section 3.2.

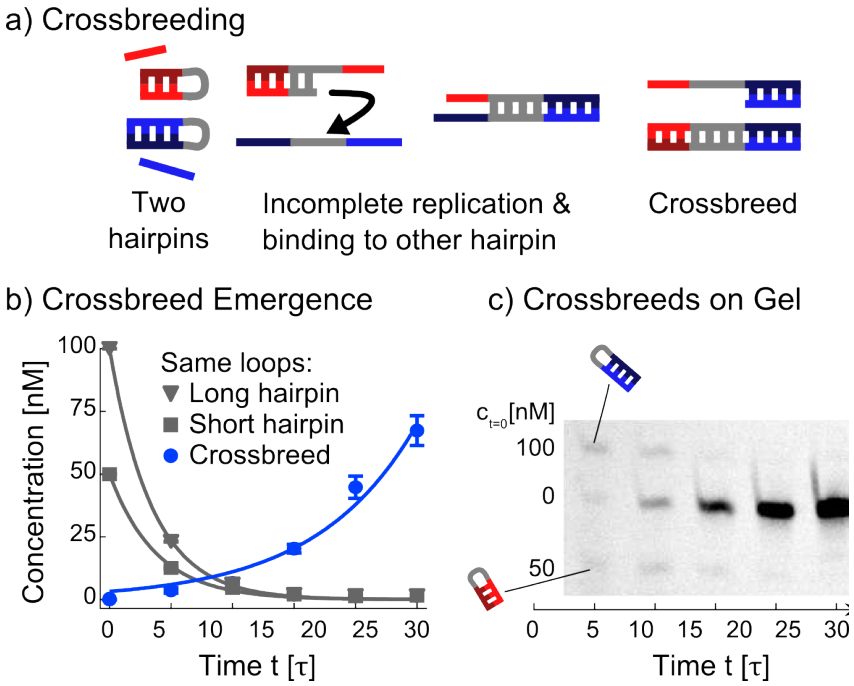


Figure 6: Crossbreeding: a) Two hairpins with the same loop form a crossbreed when one of them is incompletely replicated. In the next round of replication, the halfway elongated primer can use the other hairpin as template. Further elongation leads to the crossbreed, a fast exponential replicator without secondary structure. b) Starting with hairpins only, crossbreeds emerged after a few cycles and took over the entire population, while the hairpins vanished. Lines are exponential fits. c) A gel showing the emergence of the crossbreed. The majority of the initial hairpin population is unlabeled, thus they seemed faint compared to the crossbreed bands.

Under strong dilution conditions ($d = 5, t' = 6\tau$) the hairpin replicators still vanished and reached a fitness of $F = 0.27 \pm 0.05$ for the small hairpin and $F = 0.25 \pm 0.04$ for the long hairpin (Figure 6b,c). However, the crossbreeds preserved the sequence information of both hairpins in a rearranged manner from extinction. The fast replicators took over the entire population and survived the dilution. Since the actual initial concentration of the crossbreeds was 0 nM, it was used as a fit parameter for the exponential fit. The fitness of $F = 1.85$ was in good agreement with the fitness found in the experiments with the isolated crossbreed species in Section 3.2. The fitted initial concentration was $c_{t=0} = 3.3^{+0.7}_{-0.6}$. The generation of new crossbreeds was more efficiently done through replication than through the crossbreeding step. Especially the fast vanishing of the hairpins allowed to assume that after the first dilution step crossbreeding was only a minor contribution to the increase in crossbreed concentration.

Sequence Preservation in the Crossbreeds and Reversibility

In PCR, the emergence of faster replicating templates is usually associated with parasitic artifacts like primer dimers [26]. Primer dimers are created when primers hybridize with a small part of their 3'-ends and a replicate is created which has approximately the length of two primers. Other parasites can be produced when primers bind not on the primer binding site, but somewhere else. This can happen because of improper primer design, low annealing temperatures or high magnesium concentrations. Both described species are parasitic, can replicate very fast and potentially become more abundant than the actual target template. However, in contrast to the crossbreed, they decrease the sequence information similar to Spiegelman's monster [27].

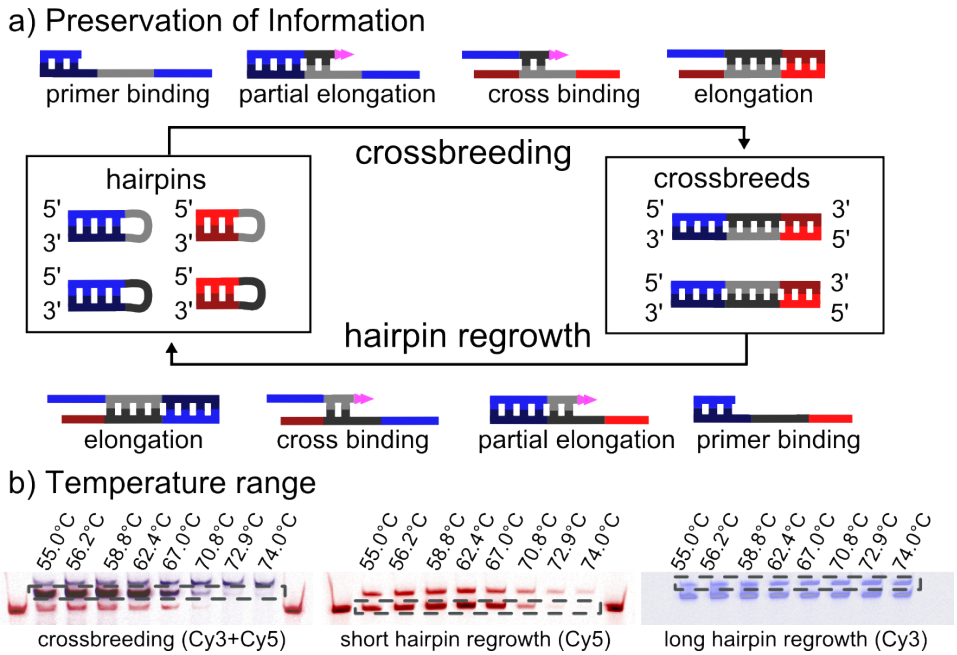


Figure 7: Preservation of Information: a) The sequence information of the four hairpin molecules (left box) is preserved in a rearranged manner in the crossbreeds (right box). The hairpin and crossbreed species are connected by similar reaction paths. The crossbreeding path is facilitated by the partial elongation of a primer and the crossbinding of the unfinished product to another hairpin species. Further elongation creates the crossbreed species. This process is reversible and hairpins can regrow from an initial crossbreed population. The key mechanism is again the unfinished elongation of a primer followed by crossbinding. b) The mechanisms worked over a wide range of annealing temperatures, as shown on a PAGE-gel. For the left gel, the long hairpin and the short hairpin were mixed with their primers, which have Cy3- and Cy5-labels. The image is an overlay of both channels and shows the emerging crossbreeds (dashed box) which incorporated both primers. The middle and right gels show that hairpins indeed regrew (dashed boxes) if both crossbreeds were only supplied with one primer. The first and the last lanes of the left and middle gels show a Cy5-labeled 84-mer as reference.

The sequence of a hairpin can be separated in three parts: The stem sequence which binds the primer, the loop sequence and the second stem sequence, which is complementary to the initial stem part (Fig-

ure 7a). When such a hairpin is replicated, the product only differs in the loop sequence, which is complementary and thus contains the same information. The second hairpin species has a different stem, thus in total four different sequences build up the ecosystem in the beginning. Crossbreeding can occur in several ways, depending on the hairpin sequence on which the replication is started, leading to four different crossbreed sequences which also follow the same patterns: The stem sequence of one hairpin is followed by the loop sequence (or its complement) and the stem sequence of the other hairpin. It is noteworthy that the crossbreeding step is reversible and starting with crossbreeds allows the formation of hairpins. This is shown in Figure 7a for a specific example, but depending on the template used for the interrupted replication all four types of hairpins (two species and their respective complementary strands) can be created.

The process worked over a broad temperature range and seemed to be limited by the binding temperatures of the primers. For the experiments shown in Figure 7b either the hairpin species with both primers (left gel) or the crossbreed species with only one primer (middle, right gel) were cycled with varying annealing temperatures. In all cases products were created which were in agreement with the suggested mechanisms. Those products are highlighted with dashed boxes in Figure 7b. The crossbreeding process worked at least for temperatures up to 70.8°C. For higher temperatures the crossbreeding process became very weak, but so did the replication of the short hairpin. Thus, the process seems to depend on the primer binding, which is crucial to start the process as shown in Figure 7a. The hairpin regrowth of the short hairpin worked over the same temperature range. When the primer for the long hairpin was added instead, the longer hairpin emerged on the gel. This process worked over the full temperature range, indicating that the strength of primer binding plays an important role, since the longer hairpin also has a longer primer, which binds more strongly.

Finite-Element Simulation of Crossbreeds

On the early earth, replicators not only had to survive dilution and degradation, but were also challenged by diffusion. To describe this additional challenge, a one-dimensional finite-element simulation was performed within the COMSOL software package. The growth of the species is described by logistic equations to include the product inhibition and limitation of resources observed for the fast exponential replicator in Figure 5a. Contrary to the experiments and the simple model used so far, the dilution is not a stepwise process, but continuous. In an RNA world setting this can be a continuous flow through a system or degradation by UV light or chemical processes:

$$\frac{dc_i}{dt} = r_i \cdot c_i \cdot \left(1 - \frac{c_i}{K}\right) - \gamma \cdot c_i + D_i \cdot \frac{d^2 c_i}{dt^2} \text{ with } i = \{\text{short hairpin, long hairpin}\}$$

with the concentrations c_i, growth rate r_i, carrying capacity K, degradation rate γ and diffusion coefficient D_i. The equation for the crossbreeds is the same, except an additional term to include the crossbreeding step with the factor δ.

$$\frac{dc_{CB}}{dt} = r_{CB} \cdot c_{CB} \cdot \left(1 - \frac{c_{CB}}{K}\right) - \gamma \cdot c_{CB} + D_{CB} \cdot \frac{d^2 c_{CB}}{dt^2} + \delta \cdot c_{Short\,Hairpin} \cdot c_{Long\,Hairpin}$$

with the index *CB* referring to the crossbreeds. The values of the parameters can be found in Table 5 (in Appendix B) and were chosen as follows: The diffusion constants were computed for a constant temperature of 75 °C following reference [28]. The degradation rate was calculated to resemble the stepwise dilutions $\gamma = ln(d)/t'$. For strong dilution conditions with $d = 5, t' = 6\tau = 432\,s$ this gives $\gamma = 3.72 \cdot 10^{-3}\,1/s$. The rate constants r_i for the hairpins were chosen to match the measured efficiencies by $r_i = \varepsilon_i/\tau$. For the crossbreeds, the growth rates and the carrying capacity were measured from a logistic fit to an experiment without dilution. The rate $r_{CB} = 6.25 \cdot 10^{-3}\,1/s$ was higher than for an exponential fit, but this was expected due to the counteracting carrying capacity K at high concentrations. The carrying capacity was measured to be $K = 207.1\,\text{nM}$, but divided by the dilution factor $d = 5$ for the model. Since the crossbreed was the only species which grows fast enough to become significantly inhibited by primer depletion and product inhibition, the same carrying capacity was used for the hairpins. The crossbreeding rate δ was varied to fit the data. The model was compared to the results from the experiments shown in Figure 6b and showed good agreement (Figure 8a). The data obtained from gels always gave the value before a dilution step, whereas the finite-element model had continuous degradation implemented. Thus, the simulated values were multiplied with the dilution factor $d = 5$ to correct for this. Starting with two spatially separated hairpin populations only overlapping in a small region, the hairpins vanished fast (Figure 8b). However, the overlap was enough to create the crossbreeds which took over the population and replicated until it reached saturation. In the center, the vanishing hairpins and the emerging crossbreed had the same concentration at 0.1 nM, which is far away from the single molecule level.

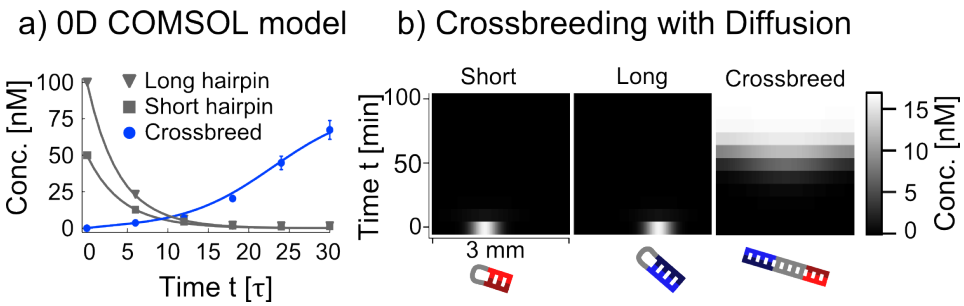


Figure 8: Simulations of Crossbreeding: a) Comparison of the results of the simulation and the experiment show good agreement in 0 dimensions (0D). The values from the simulation were multiplied with the dilution factor $d = 5$ to take into account that the experimental values were measured before the dilution. b) Starting with to different hairpin populations in the simulation leads to fast extinction. However, a small overlap of the Gaussian profiles was enough to create the crossbreed, which took over the entire population.

4 A Molecular Bet-Hedger

In nature, environments can always change and so does the selection pressure. It is known from many species that one genotype can lead to several different phenotypes, allowing to spread the risk. For ex-

ample, the bacterium Bacillus subtilis has such a bet-hedging behavior [29]. When resources become limited, some individuals switch to alternative metabolites, whereas others start an irreversible sporulation process. Such a twofold strategy allows the genotype to survive with a higher chance, even if one of the phenotypes vanishes temporarily. The switching between phenotypes can be purely stochastic [30] or epigenetic, like in the case of B. subtilis.

Here, a DNA strand is described that switches between fast replicating, low diffusing phenotypes and slow replicating, fast diffusing phenotypes and many intermediate states. Such a bet-hedging approach has recently been described in theory [31]. The molecular bet-hedger was designed to use a similar mechanism as the crossbreed, but the mechanism did not involve a second species but used repetetive sequence patterns for the crossbreeding. The geometrical properties of the bet-hedgers are similar to the short hairpins described in Section 3.2: The bet-hedger is a hairpin with a 18 bp stem and a 48 nt loop, which consisted of three repeats of a 16 nt pattern: 5'-**TAGTTATGTCGATACGAG**GATCGTCAAAG-AATGT||GATCGTCAAAGAATGT||GATCGTCAAAGAATGT**CTCGTATCGACATAACTA**-3' (stem is bold, pattern separated with '||', see Figure 9a). The bet-hedging mechanism is facilitated by incomplete replication and the emergence of incompletely elongated primers. Interruption during loop replication can occur in the first, second or third repeat of the 16 nt pattern. In the next PCR cycle, these partially elongated primers can bind a template for further elongation (Figure 9b). The repetitive patterns allow frame shifted hybridization, such that elongation can lead to shorter and longer loops. Such a frameshift through slippage is also known from microsatellites [32]. It is expected, that a possible steady state length distribution depends on the selection pressure. Longer hairpins have a smaller diffusion coefficient, whereas the effects on the replication rate are more complex:

1. It was observed in experiments that a longer hairpin loop allowed faster replication. This can be explained by the lower local concentration of the complementary stem parts. Short hairpins are likely to form their self-inhibitory secondary structure faster than long ones. They are also more stable (Figure 9c).

2. The time to finish a base-by-base replication is proportional to the length of the template, giving shorter sequences an advantage [28].

3. Longer DNA strands form more stable double strands, causing increased product inhibition.

Considering these three influences on the replication rate lead to the assumption that there might be an optimal length for a given combination of amplitude and frequency of the temperature cycles. Adjusting the temperature cycles can shift this sweet spot. When denaturation temperature is lowered, long double strands are not melted and short sequences have an advantage (Figure 9c). A short elongation time might also favor the shorter sequences, since the replication of longer sequences will be more likely interrupted. However, this interruption triggers the formation of incompletely elongated strands and might result in even longer phenotypes.

The sequence length has an influence on the replication rate, but also on the diffusion constant (Figure 9b). In well mixed environments, this should not have any effects on the outcome of an experiment.

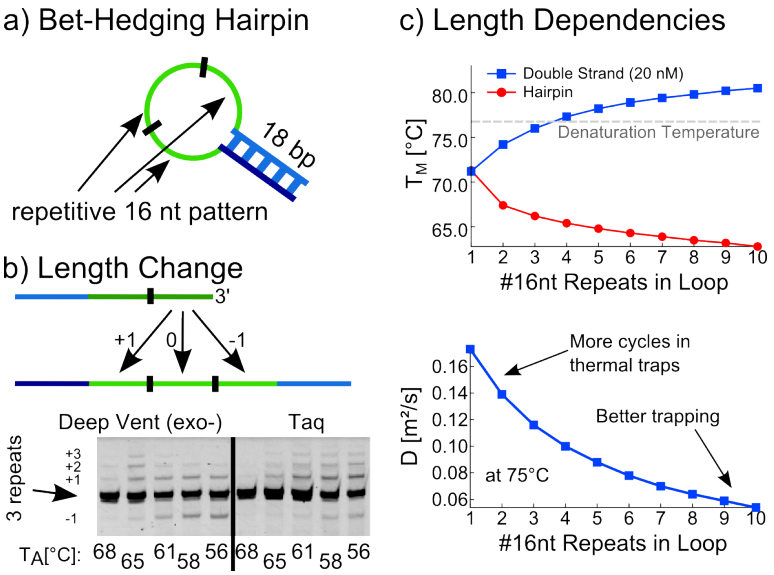


Figure 9: A Bet-Hedging DNA Ppecies: a) The initial bet-hedger is a hairpin with a 18 bp stem. The loop sequence consists of three repetitive 16 nt sequences. b) An incompletely replicated primer (top strand, interruption in second repeat) has several possibilities to bind to the initial bet-hedger, which acts as template here. After elongation towards 3'-direction of the primer a 16 nt repeat can be added (+1) or deleted (-1). It is also possible to obtain the same length as the template (0). The gel shows that the bet-hedger came in different lengths indeed after several rounds of replication. The distribution was obtained starting with 10 nM of the 3 pattern motif hairpin. PCR was conducted with annealing for 30 seconds with the temperatures indicated for 30 cycles. The mechanism worked for at least two different polymerases. c) Top graph shows predicted values for melting temperatures for the bet-hedger with different lengths [33]. The melting temperature for dsDNA increases with the number of 16 nt repeats, whereas it decreases for secondary structures in ssDNA. Choosing a lower denaturation temperature can shift the selection pressure towards shorter sequences. Bottom graph shows diffusion coefficients calculated according to reference [28]. In a thermal trap shorter sequences are cycled more often, whereas longer sequences are trapped more efficient [28].

However, using PCR in a gel or a thermal trap should allow to observe more interesting outcomes. In gels, the diffusion of DNA is slower and the absence of convection allows to prevent mixing of the reaction medium. Our initial idea was to make experiments with one-dimensional gels, starting with templates on one side and observe a reaction diffusion front towards the other side (similar to [34]). Short hairpins might diffuse fast, but replication is rather slow. Long hairpins will locally replicate fast, but are stuck and deplete primers. A molecular bet-hedger might have the chance to diffuse as a small hairpin and create the fast replicating phenotype at locations with more resources. To this end, an inverted microscope was built (described in [35, 36]), which has a heatable sapphire stage to allow temperature oscillations. Quantitative PCR was observed with this instrument, but due to problems with the gels the project was stopped.

Another environment where the molecular bet-hedger might show interesting dynamics are thermal

traps [23]. In such settings, a horizontal temperature gradient is applied through an upright positioned pore. This leads to thermal convection and thus laminar flow between the hot and the cold side, allowing PCR [24, 37]. In addition, thermophoresis has to be taken into account, which is the directed movement of molecules along a temperature gradient [38]. DNA molecules have the tendency to move to the cold side, where the gravitation driven convection flow is pointing downwards and traps the oligonucleotides. Thermophoresis is length dependend and longer DNA molecules are trapped more efficiently [10]. Moreover, if a flux through the pore is applied from bottom to top, shorter oligonucleotides are more likely to be flushed out of the trap, whereas the longer molecules tend to be on the cold side, where convection counteracts the influx. This allows replication and an influx-dependent length selectivity [28]. When the influx is low, shorter molecules can undergo more cycles due to their faster diffusion, which gives them more opportunities to replicate. When the influx is increased, the higher chance of being at the hot side can lead to an increased outflux, giving longer sequences a better chance of survival. When the influx velocity is changed, specialized molecules might vanish. Bet-hedging DNA molecules can use both strategies and have a higher chance of survival.

5 Conclusion

The presented approach allows to observe replication, selection and a mutation step in a test tube. An initial population of two different, slow replicating hairpin species created fast replicating crossbreeds. This crossbreeding reaction required that the hairpin species only differ in their stem sequence, but have the same loops. The crossbreeds took over the population and outgrew the hairpins. However, the crossbreeding reaction was reversible and the information of the hairpins was preserved.

The possibility to change back and forth between two types of replication can be seen as risk-spreading strategy, which is usually associated with bet-hedgers [29, 30]. Bet-hedgers are species with a single genotype, which have different phenotypes. A hairpin with a repetitive loop structure showed this property. The crossbreeding mechanism enabled it to change the length of the loop. This is expected to change the diffusion constant and the replication rate, which is expected to give an advantage depending on the environment.

The used materials and methods are easy to handle and widely used in many laboratories. Despite this simplicity, the experiments allow to extrapolate the results to an RNA world, where RNA has to store and replicate information. The RNA replicating ribozymes that are known so far require a binding site on the template to initiate replication. Assuming ribozymes that can target several orthogonal binding sites or primers and a random pool of possible templates, it seems unlikely that a template which has two binding sites is in the mix. For example, for a primer length of 10 nt, the chances of finding a perfectly matching template in a random oligonucleotide pool are roughly 4^{-10} whereas the estimation gives a probability of 4^{-20} for two primers and two binding sites. Of course this is only a rough estimation and does not include the possibility of binding with one or two mismatches or wobble pairs.

Simulations suggest that oligonucleotide pools on early earth might have been rich in hairpin struc-

tures [16], which would allow exponential replication from the beginning. Assuming that an early replication machinery was more error-prone than its today's enzymatic pendants, also starting from a linear template can result in hairpin formation due to mutations. Under weak dilution or degradation, hairpins can accumulate and create an enlarged pool of oligonucleotides. Mutations in the loop can allow the formation of crossbreeds, which can survive even stronger dilutions. In the presented experiments, the loops were not designed to have relevant secondary structures, but in an RNA world this might allow the replicators to catalyze other reactions.

Part II

The DNA-Toolbox as a Framework for a Rock-Paper-Scissors Game

6 Motivation and Goals

The children's game rock-paper-scissors (RPS) is an easy game with simple rules that shows cyclical dominance. Two players independently choose one of the three name giving strategies, without knowing the decision of the other player. The strategies are compared and the winner is determined by the following rules: rock beats scissors, scissors beat paper and paper beats rock. If both players have the same strategy the game is repeated. There is no strategy which guaranties a win.

When populations are involved in such a game and the strategies are determined by the genotype, oscillations can occur [39]. Starting from a population with high rock-fraction, scissors will struggle and be suppressed. This makes it easy for paper to grow, replace rock and become the dominant species. However, the decreasing rock-fraction allows scissors to recover and outcompete paper. From scissors, the population switches back to rock dominance and the cycle starts again. Such dynamics have been observed in lizard populations, where three inheritable male mating strategies are present [40]. A single oscillation in the lizard population took six years, a timescale which can be reduced to hours in engineered systems.

An example is given by the *repressilator*, which is an *in vivo* transcriptional oscillator in *Escherichia coli* [41]. The design principle is a network of three genes, each encoding a protein that represses expression of another gene in a cyclical fashion. A fluorescent marker is expressed with one of the proteins, which allowed to follow the oscillations over time.

To this temporal orchestration observed in RPS games comes a spatial component when the system is observed in two dimensions (2D) and the species are allowed to move randomly [42, 43]. Due to the localized interactions, spiral waves emerge and the species 'chase' each other in the same order as they do in time: Rock follows scissor, scissors follows paper and paper follows rock. In an experiment with *E.coli*, it was shown that the 'chasing' of neighboring colonies can be observed [44]. Pattern formation can also be monitored in chemical systems like the Belousov-Zhabotinsky reaction [45], but the chemical parameters are less accessible for an experimenter.

To implement a rock-paper-scissors game on the molecular level *in vitro*, the DNA-toolbox seems to be suitable [46, 47]. It is a framework that allows the *de novo* creation of biochemical reaction networks similar to the system of the Winfree group [48, 49]. It uses DNA with various modifications and three enzymes: A polymerase, a nickase and an exonuclease. In such networks, the dynamic entities are short DNA strands, which are continuously built up and degraded by the enzymes. Their production is facilitated by templates that bind an 'input' strand and create an 'output' strand with the help of the enzymatic machinery. The network can be programmed by smart sequence design and several bistable as well as oscillating systems have been engineered [47, 50, 51]. Interestingly, it also allows to observe dynamic pattern formation like wavefronts or spiral waves [34, 52]. In this part of the dissertation, I will describe how the DNA-toolbox works and the first steps towards a rock-paper-scissors system. This project is done in collaboration with André Estévez-Torres (Laboratory Jean Perrin, UPMC, Paris, France).

7 The DNA-Toolbox

The DNA-toolbox allows to create programmable, biochemical reaction networks based on DNA and an enzymatic machinery [46, 47, 50]. The species needed to create a synthetic ecosystem are represented by short DNA strands, typically 11 or 12 nucleotides in length. By reversible hybridization to DNA templates and the use of a polymerase and a nickase, new DNA strands can be created from the initial strand. Using different types of species and templates allows to create reaction networks. An exonuclease that degrades oligonucleotides starting from the 5'-end is added to keep them dynamic. The DNA-toolbox does not require temperature oscillations, but works close to the melting temperature of the DNA species.

7.1 Main Mechanisms: From Input to Output and Inhibition

To create networks, various DNA species have to be linked to each other. This is done by templates, constructed with an 'input' site on the 3'-end and an 'output' site on the 5'-end (Figure 10a).

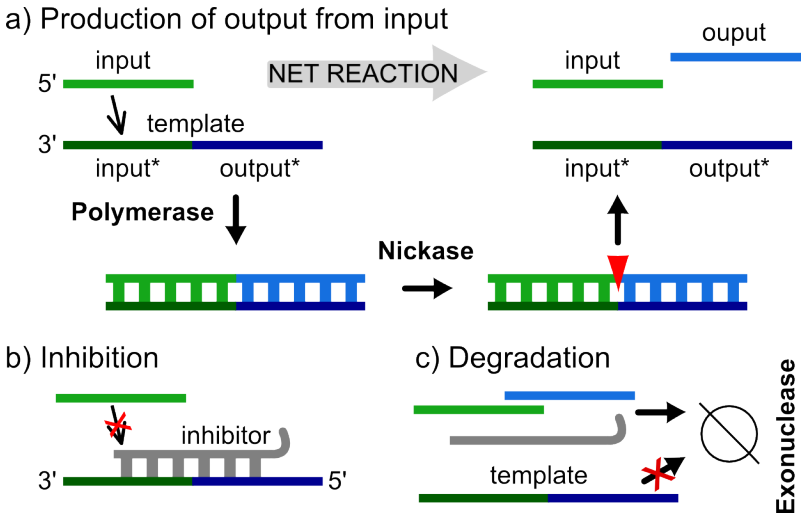


Figure 10: The main mechanisms of the DNA-Toolbox: a) To create an output DNA from an input DNA (net reaction), a template and two enzymes are involved. The template contains the complementary sequence of the input (input*) and allows hybridization of the input, which will be elongated by a polymerase according to the template. The input/template complex contains a nickase recognition site and the elongated input strand is nicked (red triangle). Input and output can unbind and participate in other reactions. b) This mechanism is prevented by inhibitors, which are DNA strands that bind to a template and block a part of the input site. Inputs have to compete with inhibitors, slowing down output production. c) To keep the system dynamic and prevent saturation, an exonuclease is added to the system. It degrades inputs, outputs and inhibitors. Templates are chemically protected from degradation.

If an input binds to the template, the polymerase elongates the sequence according to the output site of the template and consumes NTPs. Such a double stranded complex is rather stable and very likely to

stay in this conformation until the nickase cleaves the elongated input strand, creating the initial input and a new output strand. Thus, the input is designed in such a way that a 5 bp recognition site for the nickase is part of the sequence (5'-GAGTC-3'). The nickase cleaves only the elongated input strand 4 bases beyond the recognition site, but does not affect the template. Since the used polymerase has a strand displacement property, also longer output strands that bind stronger can be released from the template. The strength of a template can be tuned by removing one or two nucleotides from the 3'-end which will weaken the binding of the input, but will not inhibit elongation by the polymerase.

The second key mechanisms of the DNA-toolbox is inhibition to prevent the described process (Figure 10b). Here, a DNA strand binds to a template, but is neither elongated nor nicked. This prevents binding of an input to the template and inhibits the production of an output. To prevent elongation of the inhibitor itself, it has mismatches with the template at its 3'-end. Cleaving is prevented by not including the full recognition site of the nickase to the inhibitor part that binds to the input site of a template. Since the nickase cleaves the DNA strand 4 bases beyond the recognition site, the inhibitor can cover the complete recognition site on the templates output site. Inhibitor length can vary and is just limited by possible nickase recognition sites. Stronger binding increases the inhibition.

The third main mechanism of the DNA-toolbox is the degradation of the DNA species and the inhibitors by an exonuclease that breaks down DNA into very short pieces (1-mers, 2-mers, 3-mers), starting from the 5'-end. Templates have modifications on their 5'-end to prevent their degradation, which would lead to a breakdown of the reaction network (see 7.2).

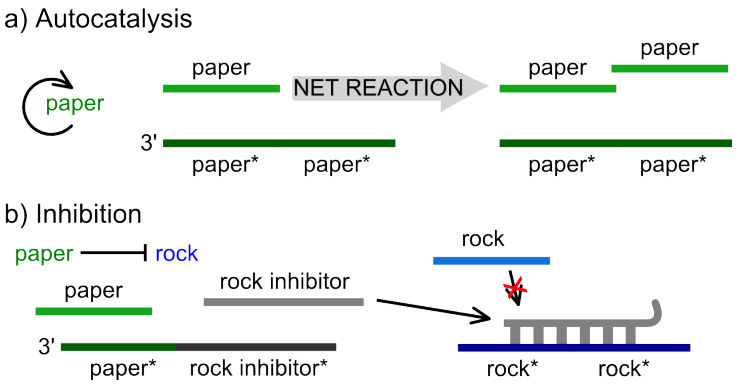


Figure 11: Autocatalysis and Inhibition: a) For autocatalysis, input and output sites of a template are the same. Here, both are of the species paper, which can thus grow exponentially. b) Inhibitors can also be generated dynamically. Here, a template is shown that produces a rock inhibitor from paper. If the inhibitor binds to the autocatalytic template of the rock species, the input site is blocked. This leads to slower autocatalysis and might end in rock depletion due to the exonuclease. The inhibitory strand can be seen as a species-specific toxin.

Depending on the sequence of the input and output sites of a template, three template types can be distinguished:

1. Autocatalytic templates: Input and output have the same sequence. The net reaction then becomes:

input → 2x input (Figure 11a).

2. Inhibitor templates: Binding of an input leads to synthesis of an inhibitor, for example to inhibit an autocatalytic template (Figure 11b). The inhibitor binds strongly to the inhibitor template because of its length, but the strand displacement of the polymerase helps to release it after the nicking. The net reaction is: input → input + inhibitor.

3. Activating templates: Input and output sequences differ and the reaction is of the type input → input+output (or since the output can be further used as input in other reactions a more general statement: species1 →species1+species2). However, this type of template is not used in this thesis.

Selfstart and Parasites

Two side effects associated with the DNA-toolbox are selfstart and parasites. Selfstart is an effect that is observed when an autocatalytic template is incubated with the full enzymatic machinery and NTPs, but in the absence of an input. In such a setting, replication can emerge with a delay compared to a reaction triggered by an initial input (Figure 12a). This behavior has been observed by other groups and does not seem to be a contamination, but intrinsic to all processes with exponential amplification reaction (EXPAR) [46, 53]. Selfstart is sequence dependend and not predictable. Depending on the type of the reaction network this might even be useful. Especially in a rock-paper-scissors network where oscillations are expected, selfstart can help a depleted species to recover faster.

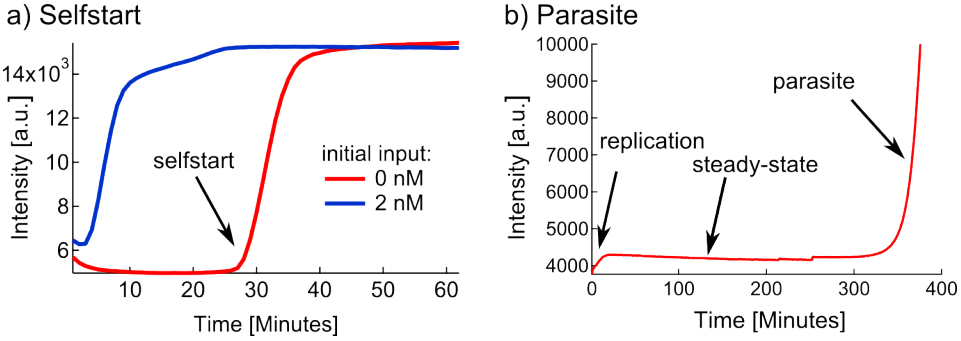


Figure 12: Selfstart and Parasites: a) If a template is incubated in the reaction solution, input is not necessary to start replication. In the shown experiment, the sample without initial input (red curve) had a delay of ~25 minutes compared to the reaction with input. EvaGreen was used to observe replication. b) An autocatalytic replication process reached a steady state after a short time. 300 minutes later, the EvaGreen signal showed an additional increase. This emerging artifact with unknown structure is termed *parasite*.

Parasites are unwanted side effects in which a self-amplifying mutant emerges that outgrows the system (Figure 12b). It can be best observed in experiments with EvaGreen, which allows to detect DNA over a wide range of concentrations. After an initial growth phase of a species, a steady state is reached

in which replication and degradation rates are equal and the EvaGreen signal stays constant. If a parasite is created, the signal increases again and reaches much higher values [46]. However, the exact structure of the parasite is unknown.

7.2 Template Modifications

In order to enhance the functionality of the DNA-toolbox and to prevent unwanted side reactions, several modifications have been shown to be useful (see [46] and Table 1). As stated above, it is necessary to prevent the exonuclease from attacking the templates. One way to do this is to use two or three phosphothioate modifications at the template's 5'-end. The exonuclease is not able to cleave this backbone modification and the template concentration is kept constant. However, it was observed that this leads to decreased exonuclease activity, presumably because the exonuclease binds to the protected strand and thus cannot degrade other oligonucleotides (see supplementary information of [54]). Another approach is to use templates with a 5'-biotinylation and add streptavidin. The streptavidin is used as a spacer, which prevents the exonuclease from binding to the 5'-end. However, this has another side effect: The streptavidin seems to sterically hinder the polymerase from finishing the elongation process. This can be easily overcome by adding two spacer-bases to the templates 5'-end, which will not be copied by the polymerase.

MODIFICATION	LOCATION	FUNCTION
phosphothioate	5'-end	prevent exonuclease digestion
biotin (+streptavidin)	5'-end	prevent exonuclease digestion
dU as dT replacement	Nickase target sequence on output site of templates	prevent non-functional nickase binding
phosphate	3'-end	prevent elongation
dye	3'-end	prevent elongation, reporting through N-quenching [55]

Table 1: Modifications of the oligonucleotides

The target sequence of the nickase is 5 nt long and has to be part of the input sites of templates. Thus, for autocatalytic templates where input and output sites have the same sequence, a second target sequence for the nickase is present, which can inhibit the nickase [46]. To prevent non-functional binding of the nickase to this site, a thymine (T) in the target sequence is replaced by a uracil (U). The polymerase does not distinguish between T and U, but the nickase is not inhibited anymore.

It is also of great importance to prevent elongation of the templates by the polymerase. This is done by adding a phosphate or a dye to the 3'-end.

The main function of dyes is to report binding of molecules by N-quenching [55]. If an input molecule binds to the input site of such a labeled template, the dye is quenched. However, this does not work for all dyes and all sequences. Quenching is usually the strongest when the dye is attached to a C, because the complementary base G is the most efficient quencher. Compared to intercalating dyes,

labeled templates offer the advantage of being sequence specific. However, the dynamic range depends on the template concentration, which can make problems for low template concentrations. Reporter molecules are a way to avoid this problems.

7.3 Reporters

To detect and distinguish several species, reporter molecules can be used instead of labeled templates (Figure 13a). This allows to change the template concentration in any desired way without running into the detection problem. The design principle is simple: For any given input, the reporter sequence is just the complementary strand with several modifications (modifications are listed in Table 1). On the 3'-end is a dye, which has the N-quenching property, leading to a signal decrease when input is bound. To avoid degradation of the reporter, it has a 5'-end biotin modification to bind streptavidin. Since reporters for inputs always contain a target site for the nickase, a dU modification is used to prevent nickase binding. Decoupling the reporting from the templates additionally allows cheaper screening for templates (e.g. varying the length of the input site or trying different inhibitors as an output). Instead of fluorescent labels the templates simply need a phosphate at the 3'-end.

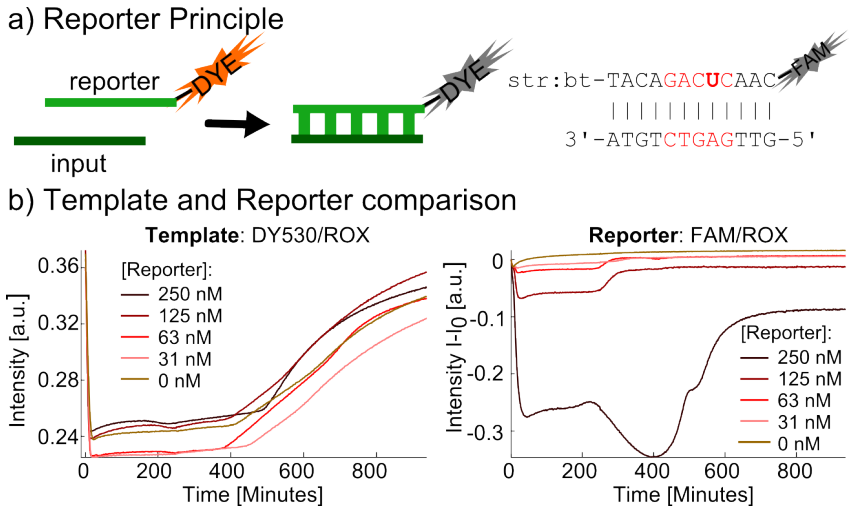


Figure 13: Reporters and Turnovers: a) Reporters are fluorescently labeled oligonucleotides that bind to a specific target molecule. Binding results in quenching of the dye on the 3'-end. On the right side is an example of a reporter for autocatalytic species. The reporter has a target sequence for the nickase (red) and needs a uracil-modification (bold, see 7.2). Best results are obtained when the label is attached to a C, because G is a good quencher. The 5'-end has a biotin-modification to bind streptavidin and prevent exonuclease degradation. b) An experiment with DY530-labeled templates and reporters shows that turnover experiments in which NTPs are limited did not depend on the reporter concentration (left graph, DY530-channel). However, rather high reporter concentrations were needed to see the full dynamics (right graph, FAM-channel). To allow better comparison of the reporters, the initial intensity was subtracted from the signal for each reporter concentration. Both signal channels are divided by the reference signal from ROX.

It is important to test that the reporters do not change the dynamics of the system. Thus, turnover experiments were performed in which an autocatalytic species grew with a reduced concentration of NTP, which is the material and energy source for the polymerase (Figure 13b). In such an experiment three phases can be distinguished:

1. In the beginning, replication leads to an increase in the concentration of the autocatalysts.
2. After a while the concentration reaches a steady state where growth and degradation are balanced. In this case, the signal of the template label or reporter does not change. However, constant signal might be reached before the maximum concentration is reached, if all reporters are quenched already.
3. When NTPs are depleted, growth stops and the exonuclease degrades all unprotected DNA strands. Thus, the signal returns to its initial value (except signal is lost due to bleaching).

Turnovers for an autocatalytic species that uses a DY530-labeled template were conducted with the addition of different concentrations of a FAM-labeled reporter. The curves in the DY530-channel do not seem to depend on the reporter concentration, suggesting that the dynamics are unchanged. However, the FAM-channel did not show the typical turnover shape for all reporter concentrations. Rather high concentrations had to be used, which can be explained by the lower melting temperatures of the reporters compared to the templates. This might also explain why the signal of the reporter increases faster compared to the DY530-channel in the third phase of the turnover, where the degradation dominates the dynamics.

The values of the reporter were not coming back to 0 for all samples. This seemed to be a bleaching effect and was also observed when the reporter is incubated for a longer time without any reaction.

7.4 A Rock-Paper-Scissors Network

Using the DNA-toolbox allows to create a rock-paper-scissors network. A single oscillation has already been observed in an experiment, but the system stalled in a state and stopped oscillating after that [56]. To create a rock-paper-scissors network, six templates are needed (Figure 14a):

- 3 autocatalytic templates that allow the species rock, paper and scissors to replicate.
- 3 inhibitor templates to allow interaction between the three species and create the cyclical dominance.

The network design connects the three autocatalytic species with inhibitory templates to create a circle: Rock inhibits scissors inhibits paper inhibits rock (Figure 14a). This approach is the same as for the reppressilator from *Elowitz et al.*, a transcriptional oscillator that works *in vivo* [41]. To check if such a network built with the DNA-toolbox allows oscillations, a simulation was performed. To this end, DACCAD was used, which is a program made to test reaction networks designed with the DNA-toolbox

in silico [57]. The program uses a model, in which the binding of DNA is described with kinetic rate equations and the enzymatic processes are modeled with Michaelis-Menten kinetics. For the simulation the standard parameters of the program were used for the enzymes and oligonucleotides. The initial species concentrations were 1 nM, 2 nM or 3 nM, all template concentrations are 10 nM. After a transient phase in the first 200 minutes, the simulation shows oscillations (Figure 14b).

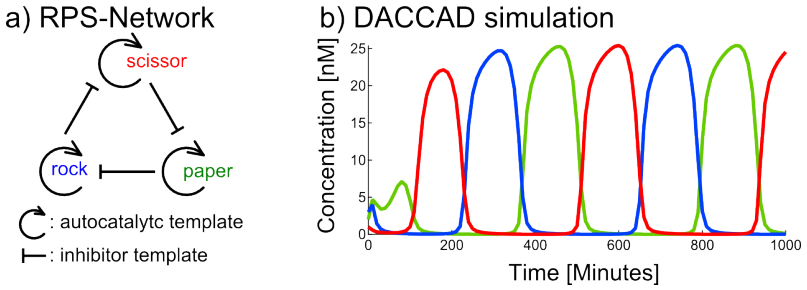


Figure 14: Autocatalytic species form a rock-paper-scissors network through inhibition: a) Three orthogonal, autocatalytic species form a rock-paper-scissors network when connected with inhibitor templates. b) Dynamics of such a system simulated with DACCAD [57].

8 Simulations in 2 dimensions

One of the main motivations for this project is to create wave patterns in a 2D system. The DACCAD model, which was used in the 0D case (Figure 14b), is very complex. For the 2-dimensional simulations, a simpler approach was chosen, which was also used by Padirac et al. [50]. This simple model was adapted to the rock-paper-scissors game and a diffusion term was added to the equations to allow exploring spatiotemporal phenomena in finite element simulation with the COMSOL software package.

8.1 The Simple Model

The simple model was developed by Padirac et al. to describe a bistable system in which two autocatalytic species can inhibit each other [50]. The original model is very similar to the model used for the rock-paper-scissors model. The main differences are the diffusion term, which allows to describe spatial behavior and the coupling of the equations, which has to be readjusted to resemble a rock-paper-scissors game. The equation for an autocatalytic rock species is:

$$\frac{dr}{dt} = \frac{t_r \cdot r}{1 + r + \lambda_r \cdot i_r} - r \cdot \gamma + D_r \Delta r \tag{2}$$

The dynamic variable $r = [r]/K_r$ is made dimensionless by dividing the rock concentration by its dissociation constant with its template. $t_r = k_r \cdot [t_r]/K_r$ is a measure for the template concentration, which is multiplied by k_r, the rate of output creation from an input bound to a template and divided by the dissociation constant. At the same time, it is the maximum growth rate in the limit of large r and in the absence

of inhibition. The degradation term is described with the parameter γ. D_r is the diffusion coefficient for the rock species and Δ is the Laplace operator. $i_r = [i_r]/K_{i_r}$ is a dimensionless measure for the rock inhibitor concentration and λ_r is a measure for the inhibitors strength. The autocatalytic equations for scissors and paper are the same as for rock (replace 'r' with 'p' or 's').

The three equations (r=rock, p=paper, s=scissor) for the autocatalytic species are coupled by three equations for the respective inhibitors (compare to Figure 14a)

$$\frac{di_r}{dt} = \frac{t_{i_r} \cdot p}{1+p} - i_r \cdot \gamma + D_{i_r} \Delta i_r \tag{3}$$

$$\frac{di_p}{dt} = \frac{t_{i_p} \cdot s}{1+s} - i_p \cdot \gamma + D_{i_p} \Delta i_p \tag{4}$$

$$\frac{di_s}{dt} = \frac{t_{i_s} \cdot r}{1+r} - i_s \cdot \gamma + D_{i_s} \Delta i_s \tag{5}$$

The variables and parameters are the same or formed in the same way as for equation (2) and most of them were chosen similar to the values from Padirac et al. [50, 52]. The diffusion coefficients are similar to the values from *Zadorin et al.* [34] where diffusion is controlled by adding hydrodynamic drags to the templates, but are varied to even lower values in this thesis. The degradation rate was adjusted to get oscillations comparable to the DACCAD model in Figure 14b. The interaction parameters λ_i were chosen to be the same for all species and were adjusted to show oscillations similar to the 0 dimensional simulation, thus many parameters have the same values:

PARAMETER	VALUE
t_r,t_p,t_s,t_{i_r},t_{i_p},t_{i_s}	0.6 1/min
λ_r,λ_p,λ_s	30
D_r,D_p,D_s ,D_{i_r},D_{i_p},D_{i_s}	1000 µm²/min-15860 µm²/min
γ	0.3 1/min

Table 2: Parameters for the simple model

Testing the Model in 0 and 1 dimension

Several experiments are done *in silico* to explore the rock-paper-scissors network with the simple model.

First, the finite-element simulation was tested in 0 dimensions to see if oscillations are observed on a similar timescale as the in DACCAD simulation in Figure 14b. To allow a better comparison, the maxima of both simulations were normalized to 1. Figure 15a shows that the oscillations of both simulations were in good agreement with each other for a longer period of time, except for a phase shift caused by the time it took to stabilize the oscillations. The periods of the oscillations from the finite-element model became longer and longer, but none of the simulations in two dimensions exceeded 1000 minutes. Pattern formation can already be observed within this time (see Section 8.2).

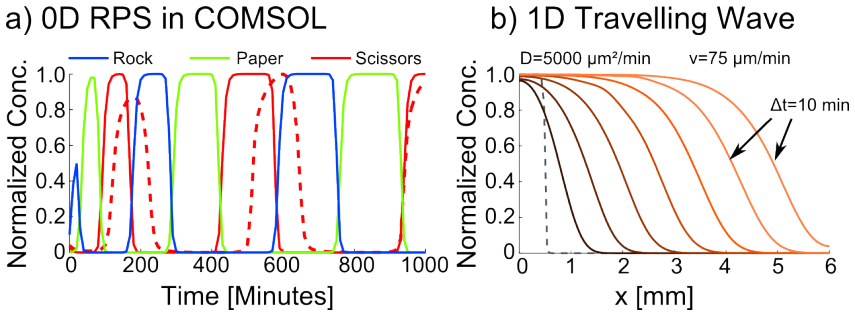


Figure 15: Finite-element model in 0D and 1D: a) The simple RPS finite-element model with the parameters from Table 2 showed oscillations. The scissors time trace of the DACCAD model shown in Figure 14 was normalized and added for comparison (dashed). b) In a system with only one autocatalytic species starting from a rectangular distribution (dashed line), a traveling wavefront formed and moved from left to right with 75 µm/min. The species diffusion constant was 5000 µm²/min and the time interval between two curves was 10 minutes.

Since the 0D case neglects the spatial component, a 1D simulation was conducted. To allow comparison to experimental values [34, 52], only one species was simulated (Figure 15b). Starting from an initial rectangular distribution, the shape changes to a sigmoidal and the front moves from left to right. The velocity was measured to be 75 µm/min for a diffusion coefficient D=5000 µm²/min and the reaction parameters from Table 2. The velocities reported in 1D channels were 32 µm/min-65 µm/min [34] and 80 µm/min-400 µm/min in a 2D system [52]. To compare the model to the experimental results, parameters similar to the 2D experiment which resulted in a velocity v=80 µm/min were chosen for a simulation (t_r=0.528 1/min, γ=0.25 1/min, D=15860 µm²/min). The result of the simulation was a front with a velocity 123 µm/min. This deviation is acceptable for such a simple model and the velocity is in the range of experimentally observed velocities.

Experiments show that the velocity strongly depends on the diffusion coefficient and the replication rate, both of which are tunable globally or for each species separately [34]. The replication rate depends on the polymerase and template concentration, whereas the diffusion of all species could be influenced by crowding agents like polyethylene-glycol (PEG). The species are not freely diffusing all the time, but are also found in bound states with their template. Thus, an effective diffusion coefficient must be considered in experiments which can be influenced by attaching hydrodynamic drags to the templates. The reported diffusion coefficient for freely diffusing species are $18 \cdot 10^3$ µm²/min, whereas the lowest measured effective diffusion coefficient is 5000 µm²/min.

8.2 Formation of Spatiotemporal Patterns

When the simple model described in Section 8.1 was applied to a 2-dimensional system, the formation of spatiotemporal patterns and spiral waves was observed (Figure 16a). Since the equations are fully deterministic and lack stochastic terms, spontaneous symmetry breaking cannot be observed. Therefore, the species were initially distributed in six Gaussian peaks with a low overall concentration in the system.

Since the color scale used is linear and normalized to the maximum in all figures, the patterns became visible only after roughly 100 minutes. Integrating the intensity over the whole surface showed that coexistence of species is enhanced compared to the results shown in Figure 15a. In the 0D case, the species almost completely vanished before they recovered, whereas in the 2D situation the species minimum amount was approximately 2 times lower than the maximum. This is in qualitative agreement with other theoretical predictions on RPS-systems from *Reichenbach et al.* [42], where it was observed that low mobility enhances biodiversity. In their simulations, stochastic effects were included that led to faster extinction of species for large mobilities, which corresponded to a well-mixed environment and thus to the 0D case.

Changing the diffusion coefficient to lower values led to the formation of finer patterns (Figure 16b). The diffusion coefficient of D=5000 µm²/min was the lowest value measured in the publication of *Zadorin et al.* [34] and already led to qualitatively different patterns. The value was achieved experimentally by adding hydrodynamic drags to the templates, which slowed down the diffusion of the species when they were bound to the templates. In the simulations, the values were even further lowered to D=1000 µm²/min, which might be experimentally feasible when crowding agents are used in addition.

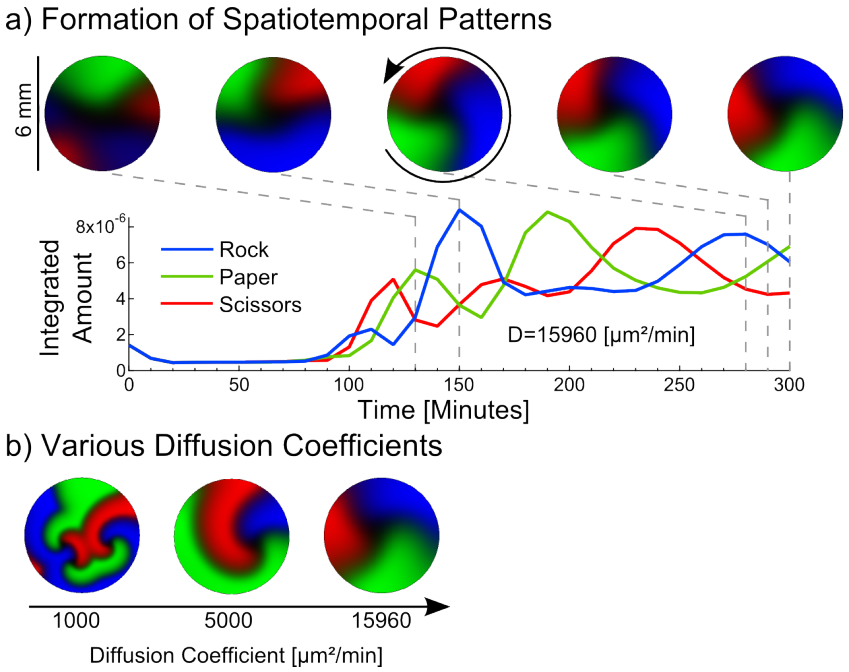


Figure 16: Spatiotemporal Patterns: a) Applying the RPS simulation in 2D led to the formation of patterns (top). Rock (blue) invades areas that were dominated by scissors (red). In turn, scissors replaced paper (green) that itself invaded rock areas. This led to a counter-clockwise rotating spiral, indicated by the arrow. Integrating over the surface area shows that the total amount of the species oscillated as well (graph), but the population size of each species was more stable than in the 0D case. Thus, the spatial dimension helped stabilizing biodiversity. The gray dashed lines indicate the times at which the shown images were taken. b) With increasing diffusion coefficient, the patterns became more coarse. All images were taken after 300 minutes.

Simulations of Inhomogeneous Template Distribution

For the simulations described so far, the template distribution was assumed to be homogeneous. As described in Section 7.2, templates can be biotinylated and bound to streptavidin to prevent exonuclease degradation. If the streptavidin is attached to a surface, the templates can be spatially localized (Figure 17a). This allows to influence the spatiotemporal behavior. Since for the rock-paper-scissors network six templates are required, there are manifold possibilities to arrange them. Here, only the distribution of inhibitor templates in very simple geometries was investigated. Thus, the inhibitor growth rates t_{i_r},t_{i_p},t_{i_s} were dispersed, whereas all other parameters were the same as in Table 2. For simplicity, the diffusion coefficient D=5000 µm²/min was assumed to be constant for all inputs and inhibitors. However, in an experiment the actual diffusion coefficient should depend on the location. Inputs and inhibitors would be temporarily immobilized when they are bound to a template which is attached to the surface.

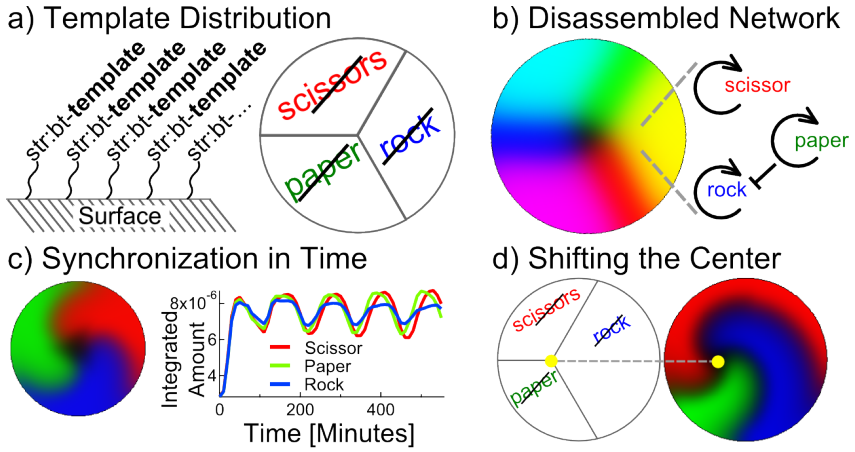


Figure 17: Inhomogeneous Template Distribution: a) Left: Streptavidin linked to a surface allows to immobilize the biotinylated templates. Right: A distribution of inhibitor templates (crossed 'rock' means rock inhibitor templates etc.). b) When the respective inhibitor templates were only present in the restricted areas shown in a), the networks disassembled and a static pattern emerged. As an example the local network in the yellow area is shown, where rock was suppressed, but scissor and paper replicated. The cyan and magenta colored areas also allowed coexistence of two species. The diffusion of inhibitors created zones in which only one species survived (red, green, blue) and a void in the center (black). c) When the maximum growth rate of a specific inhibitor template was 0.6 1/min in its respective areas and 0.3 1/min elsewhere, the moving pattern was stabilized and the integrated amount of the species oscillated in phase. d) When the growth rates were the same as in c), the center of the spiral (yellow point) was shifted with the pattern. All images in this figure were taken after 300 minutes.

Since the dispersed templates already break symmetry, the starting conditions for the three species were chosen to be homogeneous. This numerical approach already allowed observing three different effects:

- Disassembling of the reaction network into sub-networks (Figure 17b).
- Shift of the center of the rotating spiral (Figure 17c).

- Synchronized oscillations in time for the total species amount (Figure 17d).

The maximum growth rates of the inhibitors were set to 0.6 1/min in the regions with the attached templates. If there were no additional inhibitor templates, the RPS-network disassembled locally and created a pattern which was stable in time. Figure 17b shows such a sub-network: In a region with the template that produces the inhibitor for rock, the species scissor and paper were not affected and replicated, whereas rock went extinct. The same principle created two different sub-networks in the areas with templates for paper- and scissors-inhibitors. Between two adjacent areas, transition zones were observed in which only one species was present, because of diffusion of the inhibitors. In the center this led to a spot without any species.

In contrast to the static pattern just described, dynamic patterns were observed when the inhibitor templates were not only restricted to their areas, but were also present elsewhere in lower quantity. In an experiment, this could either be realized by attaching the templates also to the other areas or by just adding it to the reaction mixture. In the model, the maximum growth rates of the inhibitors were set to 0.6 1/min in the areas with attached templates and 0.3 1/min everywhere else.

In Figure 17c, the resulting spiral is shown, which rotated counterclockwise. A plot of the integrated amounts of each species over time shows synchronized oscillations. When a species had to transit a zone with high concentration of its inhibitor template, the overall amount of the species decreased. Because of the geometrical symmetry of the system, this happened for all species at the same time and led to the in-phase oscillations over time. The center of the surface was again an area in which all species were inhibited.

When the areas with the attached templates were shifted, the center point of the spiral shifted as well. The maximum growth rates of the inhibitors were again set to 0.6 1/min and 0.3 1/min as just described, but in the locations as shown in Figure 17d. This shifted the center of the spiral to the point where the three areas touch.

Using more complex patterns, e.g. by including distributions of autocatalytic template or gradients of the templates might allow the formation of other variable static and dynamic patterns.

9 Sequence Screen

To find sequences that can be used for the RPS system, a sequence screen was conducted.

9.1 Materials and Methods

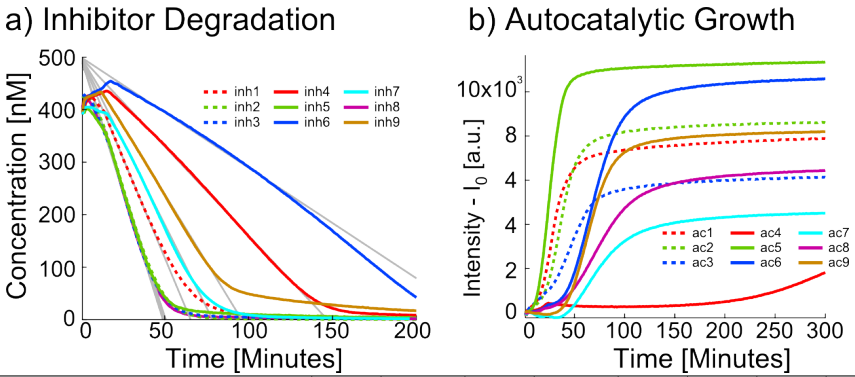
All oligonucleotides were ordered from biomers.net GmbH and a list of the sequences can be found in Appendix C. Streptavidin, the nickase Nt.BstNBI, Bst DNA polymerase (Large Fragment), ThermoPol Reaction Buffer (Detergent-free) and deoxynucleotides (dNTP) were ordered at New England Biolabs. The polymerase was used at a concentration of 16 U/ml and the nickase concentration was 20 U/ml. Netropsin and Synperonic F 108 were ordered from Sigma-Aldrich. The exonuclease ttRecJ was pro-

vided by André Estévez-Torres and used at a concentration of 12.5 nM if not stated otherwise. The reference dye ROX was purchased from Affymetrix, the intercalating dyes EvaGreen from Biotum and SybrGold from Life Technologies. For incubation and fluorescent readouts a C1000 thermal cycler with a CFX96 (Touch) Real-Time PCR Detection System from Biorad was used. If not stated otherwise, the temperature was set to 42 °C. To prevent condensation on the lid, the lid was heated to 65 °C or the sample was covered with an oil layer. Pipetting was done on ice, since the enzymes show activity at room temperature.

The buffer used was composed of 1x ThermoPol Reaction Buffer (Detergent-free) and substituted with: 25 mM Tris-HCl (pH 8), 5 mM $MgCl_2$, 50 mM NaCl, 0.1 % Symperonic F 108 and 2 µM Netropsin, 3 mM dithiothreitol, 0.5 mg/ml BSA. Nucleotide concentration was 40 µM or 400 µM. When biotinylated templates or reporters were used with streptavidin, samples were taken from a stock solution of oligonucleotide and streptavidin mixed in a 1:1 ratio.

9.2 Screen with Unmodified Oligonucleotides

Two sets of orthogonal autocatalysts were designed by a software provided by Nathanaël Aubert. The first set (set 1) contained three species, the second set (set 2) contained six species. Autocatalytic inputs, inhibitors and autocatalytic templates without any modifications were ordered for an initial screening (see Table 6, Appendix C for sequences). Degradation of the inputs, inhibitors and autocatalytic growth were measured and compared.



species:	1	2	3	4	5	6	7	8	9
input degradation [nM/min]	1.6	10.2	2.8	1.34	2.4	3.3	8.7	6.6	3.7
inhibitor degradation [nM/min]	7.2	9.4	9.7	3.4	9.5	2.1	6.1	9.7	5.0
minute of maximal signal change	26	34	48	327	23	65	56	69	61

Figure 18: Sequence Screen: a) 500 nM inhibitor strands of the two autocatalytic sets (dashed, solid) were incubated with the intercalating dye SybrGold and 1.6 U/ml exonuclease. Lines were fitted (gray) and the fit used to scale the data. This was done to correct an initial signal increase, which seems to be an artifact of the thermocycler. The fits also allowed obtaining maximum degradation rates. b) 100 nM of each autocatalytic template were incubated with 2 nM of its input strand with low polymerase concentration and the intercalating dye EvaGreen to observe growth. The appended table shows the measured degradation rates for inputs and inhibitors as wells as the time at maximal signal change during growth.

For the degradation assays, 500 nM of the input and inhibitor strands were incubated with exonuclease and the intercalating dye SybrGold. Figure 18a shows the results for the inhibitors (inh1-inh9). A baseline correction to the raw data was made and lines were fitted to the initial degradation phase. The fit was used to correct for an initial signal increase, which might be an artifact from the thermocycler. Furthermore, the slopes of the fits gave the maximal degradation rates (see Table below Figure 18). It was observed that the initial artifact can be prevented by covering the samples with an oil layer. However, this did not change the measured values significantly. Degradation rates were obtained for input strands (inp1-inp9) in the same way.

To compare the growth of the autocatalysts, 100 nM of each template were incubated with 2 nM of the respective input, nickase and low polymerase concentrations. The signal increase over time stems from EvaGreen and is shown in Figure 18b. To obtain a numerical value, the minute of the maximal signal increase was determined for every curve and the results are shown in the Table below the figure.

Set 1 (species 1-3) was neglected due to the rather large variation in the input degradation. In set 2 (species 4-9) species 4 was neglected because of the slow growth of the autocatalyst, whereas species 5 was too fast and neglected as well. The species 6-9 were used for further studies. Having four different species is a redundancy that allows to assemble the reaction network in twelve different ways, whereas having three species would allow only two ways (1 inhibits 2 inhibits 3 inhibits 1 or 1 inhibits 3 inhibits 2 inhibits 1).

9.3 Testing a Subset of Species with Modifications

For the species 6-9 templates with full modifications were ordered. They had biotin at the 5'-end with the additional spacer sequence AA, the fluorophore DY530 on the 3'-end and the uracile modification in the nickase recognition site on the output side. Turnover experiments, which have already been described in Section 7.3, were conducted with the species 6, 7, 8 and 9 (Figure 19a). To this end, 100 nM of template from a stock containing streptavidin were incubated with 2 nM of its respective inputs. In addition to the signal from DY530, the intercalating dye EvaGreen was added to have a control channel. The samples showed the typical turnover behavior in the EvaGreen channel (signal increase, plateau phase, signal depletion after NTPs were consumed). The DY530 channel showed slightly different results: For the species 6 and 9, the fluorophore was quenched when the input concentration increased. The signal for species 8 was rather hard to interpret, because the signal did not look like a turnover at all. The template of species 7 showed signal increase when more input was produced and the shape looked similar to the EvaGreen control channel. Comparing the plateau phase with the steady state phase after degradation showed that the species 6 and 9 gave the strongest signal change, whereas species 7 showed roughly 50 % less change. A control sample without EvaGreen only showed this initial signal increase, but stayed flat afterwards.

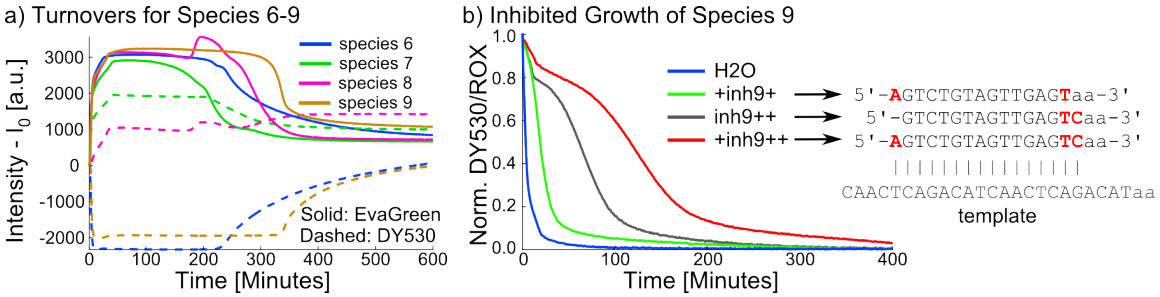


Figure 19: Turnovers and Inhibition: a) Turnovers in the EvaGreen channel (solid) show the typical behavior for all 4 species: growth, steady state and degradation. However, the DY530-fluorophore (dashed) only showed clear quenching in the case of species 6 and 9. The signal of species 8 behaved strange, whereas the signal of species 7 was enhanced upon input replication. b) A growth experiment with species 9 (60 nM template, 2 nM input) and different inhibitors (500 nM) showed delayed replication due to the inhibitors. As expected, the longest inhibitor +inh9++ showed the strongest effect. This experiment was done without exonuclease.

Varying Inhibitor Length

The interaction between rock, paper and scissor is facilitated by inhibitors. However, it was found that the inhibitors used in Figure 18a were too weak to significantly inhibit the self-replication once the species reached higher concentrations. Thus, stronger inhibitors have been tested for species 6 and 9. The following notation is used for the inhibitors: The inhibitor used in Figure 18a for species 9 is called 'inh9' and binds to the autocatalytic template with 13 bases. 'inh9++' has two additional binding bases on its 3'-end, '+inh9+' has an additional binding base on each end and '+inh9++' has one additional binding base on its 5'-end and two on its 3'-end. In an experiment with autocatalytic species 9 and a high initial concentration of the different inhibitors, it became clear that the longer inhibitor was clearly stronger (Figure 19b). The inhibitors were strong enough to delay growth when they were added in the beginning, but were not able to inhibit the system when they were injected to an autocatalytic species after growth had reached the steady-state. Instead of further increasing the strength of the inhibitors, the autocatalytic templates were weakened in the next step.

Using Weaker Autocatalysts with Reporters

To weaken the input binding of the autocatalytic species 6 and 9, new templates were ordered that had one or two bases missing at their 3'-end. They were labeled with the dye JOE on their 3'-ends, but quenching was too weak to be used as a readout. Thus, reporter molecules were used to monitor the concentration changes.

For species 6, a BMN5-labeled reporter was designed, whereas the reporter for species 9 had a FAM-label. Best results for fluorophore quenching were obtained when it was quenched with a G. This becomes clear in Figure 19a, where the inputs 6 and 9, both with a G at their 5'-end, show clear turnover curves in the DY530-channel. Thus, species 7 had been slightly changed by adding a G to its 5'-end,

making it a 13-mer with a BMN3-labeled reporter of 12 bases in length. The template used for this new species bound the 13-mer only with 11 bases, just like the other species.

To check for crosstalks between the three reporter molecules, an experiment was performed in which 150 nM of all reporters were incubated with every possible combination of templates (and the corresponding input) to observe turnover curves (Figure 20). An additional tube contained the reporters and the reaction mixture, but no template. This sample was used as a reference to correct for bleaching of the fluorophores over time. The results showed a clear difference between the presence and absence of a species for all combinations. There was a slight crosstalk between species 6 and 7' that led to a signal change in the reporter for 7' when species 6 is present and species 7' absent. However, this signal change could be clearly distinguished from the typical turnover shape.

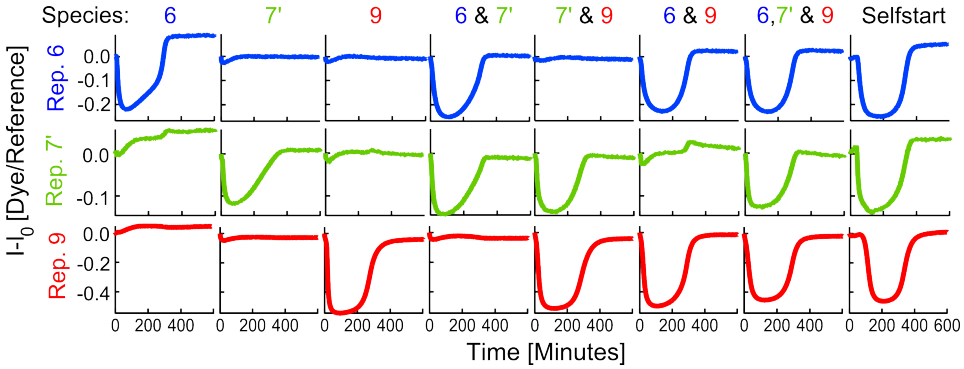


Figure 20: Turnovers with three Species and three Reporters: Different combinations of templates and inputs were incubated in the presence of three reporters (150 nM each) and the three enzymes with 40 µM NTPs. Each column represents three different readouts in a single tube and in the top the combinations of species in a tube are indicated. Here, 50 nM template were used which bind the 2.5 nM initial input with 11 bases on the input site. To correct for bleaching, the time traces were divided by signal from a reference tube which contained the reporters, but no templates. Clear turnovers were only observed when a species was present. However, there was a slight crosstalk coming from the reporter for species 7' when species 6 is present and 7' absent. In an experiment where all templates without inputs were incubated the expected, delayed selfstart was observed.

To observe selfstart, another tube was added that contained all templates, but no inputs. As expected, turnovers were observed, but they were delayed in comparison to the samples with initial input.

This set of species was chosen for further experiments to create the RPS-system. To get an oscillating system, it is important that inhibition is strong enough to reduce the concentration of a species after it reaches a steady state, but also reversible to allow growth at later times. To this end, a turnover experiment with species 9 was performed, where the signal change of autocatalytic replication was monitored in real time. When the system reached a steady state, the strong inhibitor +inh9++ was injected with a pipette. Since this slightly diluted the reaction mixture, water was injected instead of inhibitors in a control sample. To obtain the desired behavior, the exonuclease concentration had to be increased by a factor of 5 to 62.5 nM compared to previous experiments.

This experiment was done with three templates of different lengths on the input site. The choice of

the template had an enormous effect on the outcome of the experiment: The inhibition of weak templates (10 nt on the input side) was so strong that there was no replication observed after injection of the inhibitor (Figure 21a). Strong templates (12 nt on the input side) showed no change of the species concentration when inhibitor was injected (Figure 21c,d).

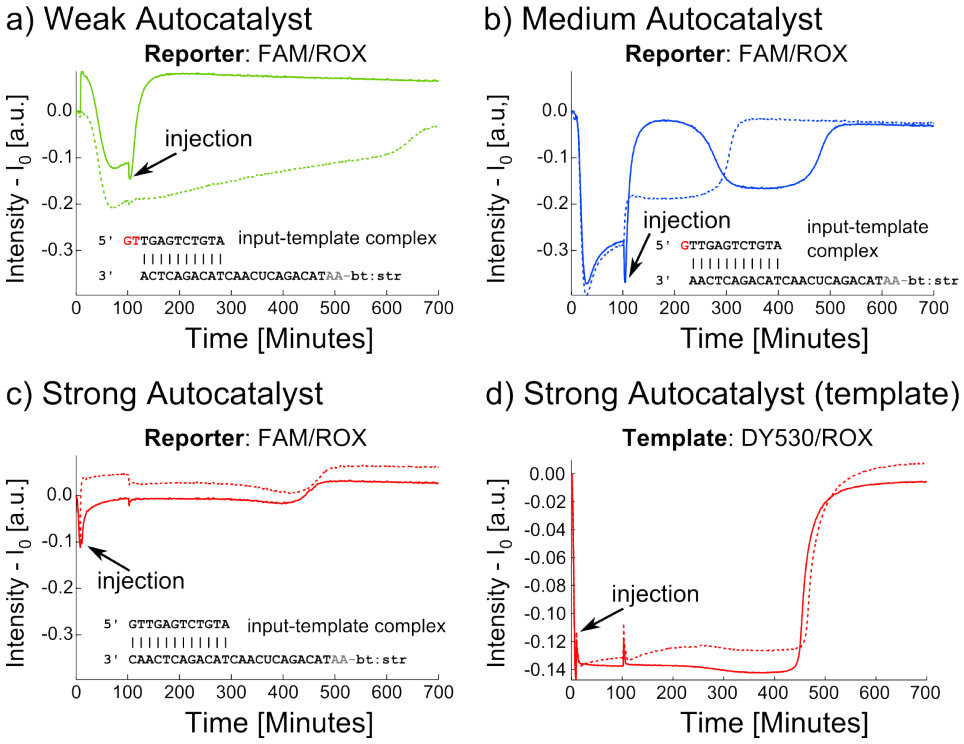


Figure 21: Inhibitor Injections: Species 9 was incubated with 60 nM autocatalytic template of different strengths and 250 nM reporter. Strength was modulated by the number of bases the species bound on the input site (unbound bases are shown in red). Autocatalysis took place until a steady state was reached and 2 µl of 1 µM inhibitor +inh9++ (solid curves) was injected. Injections were made after 10 minutes for the strong autocatalysts and after 103 minutes for the other templates. In control experiments water were injected (dashed curves), which had no impact except drifts in signal. a) When the weak autocatalytic template was used, the inhibition caused the signal to go back to its initial value, indicating degradation of the autocatalytic species. The system stayed in this state until the experiment was over. b) For the template with medium strength, the desired dynamics were observed: The inhibitor led to vanishing of species 9, but after some time the species recovered and concentration increased until NTPs were depleted. The control sample shows a typical turnover curve, except for a signal shift caused by the injection. c,d) The signal change of the reporters was very low for the strong templates. However, the template itself was labeled with DY530 which allows to follow the dynamics over time. The injection of inhibitor did not seem to have any pronounced effects.

The intermediate template (11 nt on the input side) showed degradation of inputs after injection of the inhibitor (Figure 21b). Unlike for the weak template, the species recovered, reached a new steady state and was degraded after the NTPs were depleted.

All water injections showed the typical turnover shaped curves, which were slower compared to

previous experiments. This was most likely because of the higher exonuclease concentration and the concentration changes the injections caused regarding the enzymes and the buffer.

Using weaker templates had the additional advantage that reporters seemed to work better and lower concentrations could be used. This can be explained with the weaker binding on the templates input site, which led to more free binding partners for the reporters.

10 Conclusion

The simulations obtained in 0D and 2D showed that the DNA-toolbox is a good candidate for the creation of a rock-paper-scissors network and observing phenomena like temporal oscillations and dynamic pattern formation. Inhomogeneously distributed templates allowed to change the spatiotemporal behavior, for example from dynamic patterns to static patterns. A 1D simulation with experimental parameters from *Padirac et al. [52]* led to velocities comparable to the experimental values.

On the experimental side, three autocatalytic species with similar growth dynamics were established, which were monitored in real-time with external reporter molecules without disturbing the system. For reversible inhibition, it turned out that the strengths of the autocatalysis and inhibition had to be balanced properly. The time it took for a species to recover from an inhibitor injection was three hours, suggesting that it is feasible to observe full oscillations on the timescale of several hours. The recovery from inhibition is crucial to obtain oscillation without stalling in one of the three states, but has to be accomplished for the other two species as well.

Establishing the full rock-paper-scissors network with the DNA-toolbox offers the opportunity to test under which conditions oscillations emerge and how stable they are. Depending on the template and enzyme concentrations, the system is expected to stall in a system with a dominant species or to reach a coexistence state without oscillations. Since the enzyme activities and the binding kinetics change with the temperature, the system can also be perturbed easily without changing its composition. This gives the opportunity to see how robust the system is.

In contrast to the *repressilator*, which is an *in vivo* oscillator in *E.coli* [41], the volumes and concentrations are high. Thus, the *in vitro* approach is expected to allow to study RPS dynamics in a deterministic regime.

Part III

Wavefronts in Morphogen Gradients Created with the DNA-Toolbox.

11 Motivation and Goals

Morphogen gradients are of great importance for the creation of protein patterns in cell polarization or embryogenesis. An example is provided by the development of the fruit fly *Drosophila* [58]. RNA localized at the anterior pole of the egg leads to the expression of the protein bicoid, which diffuses and creates a gradient. Bicoid acts as transcriptional activator and regulates protein expression within the genetic network of *Drosophila,* leading to the patterns that then relate to different parts of the body. In this process, the genetic network has to interpret the flat gradient in a way, that gives rise to sharply distinguished zones.

Theoretical results of *Rulands et al.* [59] suggest that the genetic networks do not have to be complex to create stable localized wave fronts. In their work, they study 1-dimensional, bistable systems with self replicating particles which diffuse, replicate and degrade. The replication process can be activated externally by a spatially inhomogeneously distributed morphogen. In the system forms a stabilized wavefront, which separates an area with high particle concentration from an area with low particle concentration. In their study, they show under which conditions such a localized wavefront can form and how it behaves when perturbed.

In this part of the thesis, an experimental approach to create such localized wavefronts is presented. The simple setup of the DNA-toolbox allows to study a minimal bistable system in which an autocatalytic species encounters a gradient of a suppressor. A traveling wavefront forms and moves from regions with low suppressor concentration to higher suppressor concentrations, but gets slowed down in the process [60, 61]. This leads to the formation of a localized wavefront, i.e. a steady state in which the system is divided into a part in which autocatalysis takes place and into another part where it is completely suppressed. Using fluorescence microscopy allows to follow the dynamics of the wavefront and measure its width. The presented measurements were performed in the *Laboratory Jean Perrin* (UPMC, Paris, France) and supervised by André Estévez-Torres.

12 Theoretical Considerations

To describe 1-dimensional reaction-diffusion phenomena, the following equation is used [62]:

$$\frac{\partial c}{\partial t} = f(c) + D\frac{\partial^2 c}{\partial x} \tag{6}$$

with the concentration c, the diffusion D and the function *f(c)* which describes the reaction process. This equation allows traveling wave solutions with velocity v such that $c(x) = c(x - vt)$. For the case that $f(c) = k \cdot c \cdot (1 - c)$ with the positive growth rate k equation 6 is termed Fisher-Kolmogoroff equation. In the 0-dimensional case this leads to a sigmoidal growth curve for the concentration of a species, which is a good approximation to experimental results (Figure 22).

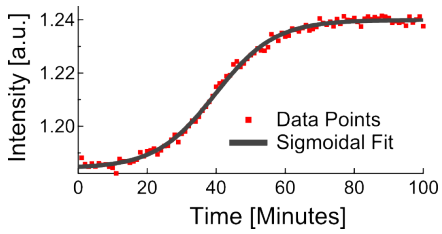


Figure 22: Species Growth with Sigmoid Fit: The shown species replicated from an initial concentration of 0.5 nM until a steady state was reached. This process was described with a sigmoidal fit. For the fluorescence signal the intercalating dye EvaGreen was used and divided by a reference signal from DY530.

For traveling wavefronts it is expected that the velocity scales with the square-root of k, or $v^2 \sim k$. For the DNA-toolbox, it is known that the replication rate depends linear on the polymerase concentration c_{pol} and thus $k \sim c_{pol}$ or:

$$v^2 \sim c_{pol} \tag{7}$$

which has been observed experimentally [34]. In this part of the thesis, a suppression mechanism will be introduced that interferes with the autocatalytic growth process of the DNA-toolbox. Thus, $k = k(c_{pol}, S)$, where S is the concentration of the suppressor that acts as a morphogen. For a wavefront traveling in the presence of such a morphogen gradient, it is expected that the front slows down with increasing suppressor concentration and eventually stops.

13 Experimental Realization

The autocatalytic mechanism of the DNA-toolbox has been described in Section 7.1 and only differs slightly compared to Part II. The main differences in this study were the use of the nicking enzyme Nb.BsmI instead of Nt.BstNBI and different DNA sequences. An autocatalytic DNA species binds with 10 out of 12 bases to a DNA template and gets elongated by a polymerase (Figure 23a). The product is a strand, which contains the input sequence two times and is bound to the template. The nicking enzyme targets this product and cleaves it such that the two fragments have the same sequence, which is the input sequence. The nicking site of the nickase is in the recognition site, whereas it was 4 bases away from the recognition site in Part II of this thesis.

The morphogen is a so called pseudotemplate (pT) that binds the autocatalytic species that leads to its elongation with a short sequence to create a pseudospecies (Figure 23b) [61]. The short sequence added is not part of the recognition site of the nicking enzyme and thus is not cleaved off. When the pseudospecies binds to an autocatalytic template, it cannot be elongated and thus autocatalysis is weakened or even completely suppressed. To create a one-dimensional gradient of this suppressing morphogen, a PDMS adapter was used that has a channel to connect a pipette tip and a capillary (Figure 23c, left). The capillary was filled with the reaction mixture that contained everything that is required for an autocatalytic species

to replicate and then plugged with one end into the PDMS adapter. The other end of the capillary was put into a droplet of the same reaction mixture supplemented with a higher pseudotemplate concentration and a reference dye to track the gradient. Then the pipette was pressed and released several times. Gradients were monitored using the reference dye and showed stability for at least 1000 minutes (Figure 23c, right).

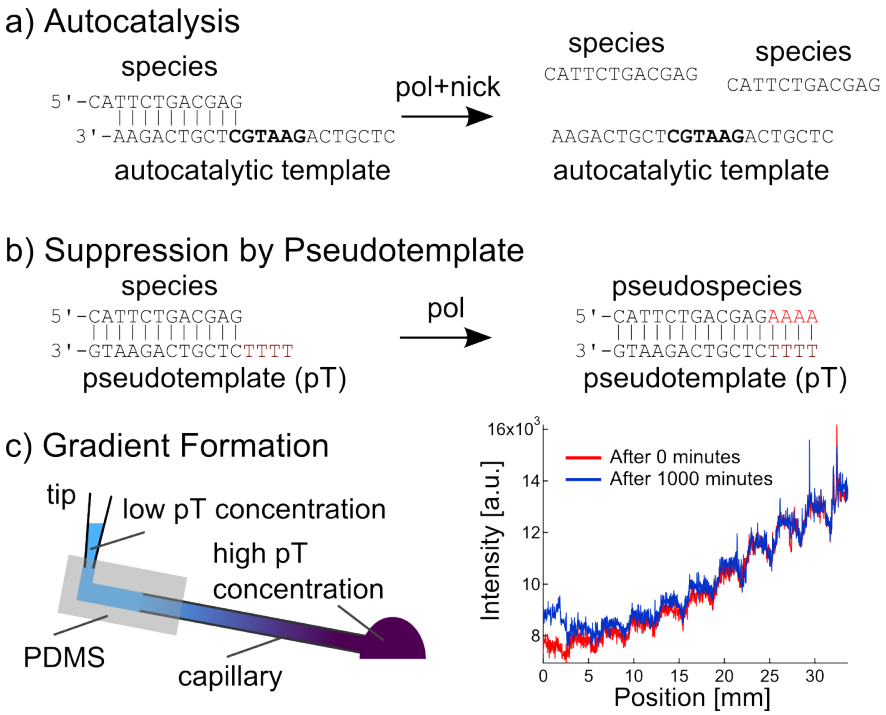


Figure 23: Pseudotemplates and Gradients: a) The autocatalytic process is similar to the one described in Part II. A species binds to the template, gets elongated by the polymerase and nicked by a nicking enzyme. The nicking enzyme recognizes the bases around the nicking site (bold). b) Pseudotemplates (pT) allow the elongation of a species for several bases (red). The bases are chosen such that the newly created pseudospecies is not nicked and cannot be further elongated on autocatalytic templates. This is a suppression mechanism that turns molecules of the species into nonfunctional pseudospecies. c) Gradients are created by attaching a capillary filled with a low pT concentration to a pipette with a PDMS adapter. The other side of the capillary is placed in a mix that contains a high pT concentration and the piston of the pipette is slowly moved up and down. The graph shows that the gradient of a reference dye is stable for at least 1000 minutes. The repetitive pattern stems from an inhomogeneous illumination.

13.1 Materials and Methods

All oligonucleotides were ordered from biomers.net GmbH. The autocatalytic species CATTCTGACGAG (all sequences are written from 5' to 3') used the labeled template CY3.5-C*T*C*G*TCA**GAATG**-**C**TCGTCAGAA for replication (* denotes phosphothioates that prevent degradation by the exonuclease, recognition site for the nickase is in bold). Two different pseudotemplates (pT) were used which differed in the attached bases to the species and in the protection against the exonuclease: 1) T*T*T*T*CTCGT-

CAGAATG used phosphothioates, 2) bt-AAAAAACTCGTCAGAATG used a 5'-end biotyinlation (bt) and two adenins as a spacer and was used with streptavidin (compare Section 7.2).

Streptavidin, the nickase Nb.BsmI, Bst DNA polymerase (Large Fragment), ThermoPol Reaction Buffer (Detergent-free), and deoxynucleotides (dNTP) were ordered from New England Biolabs. The polymerase was used at a concentration of 16 U/ml if not mentioned otherwise and the nickase concentration was 500 U/ml. Netropsin and Synperonic F 108 were ordered from Sigma-Aldrich. The exonuclease ttRecJ was used at a concentration of 25 nM if not stated otherwise. EvaGreen was used at a final concentration of 1x as suggested by the supplier. For 0-dimensional measurements, a short oligonucleotide with the dye DY530 was added as a reference at a concentration of 10 nM.

The buffer composition was as follows: ThermoPol Reaction Buffer (Detergent-free) supplied with 5 mM $MgSO_4$, 50 mM NaCl, 0.1 % Symperonic F 108, 1 µM Netropsin, 3 mM dithiothreitol and 0.5 mg/ml BSA. NTP concentration was 400-800 µM. For the autocatalytic template 25 nM were used and 1 nM for the species if not indicated otherwise. Pseudotemplate concentrations were changed over a wide range and gradients were monitored with the dye cascade blue.

For 0-dimensional experiments a C1000 thermal cycler with a CFX96 (Touch) Real-Time PCR Detection System from Biorad was used. For 1-dimensional experiments, a Zeiss Axio Observer Z1 microscope with a CoolLED pE-2 ber-coupled illuminator, DAPI, GFP, and RFP filter sets, a Marzhauser XY motorized stage, a Tokai Hit thermo plate, a Andor iXon Ultra 897 EMCCD camera, a shutter and a 2.5x objective was used. The temperature was set to 45°C. The pipetting was done on ice for as long as possible, since the enzymes show activity at room temperature. Preparation was done at room temperature as soon as the capillary filling started. In the 1-dimensional experiments 50 mm x 4 mm x 0.2 mm capillaries were used. To create gradients, the capillaries were filled with the reaction solution with low pT concentration and then with one end connected to a PDMS adapter attached to a pipette. The other end was dipped into 65 µl of the reaction solution containing the high pT concentration and the pipette was pressed and released 15 times slowly with the volume adjusted to 12.5 µl. The high pT concentration got diluted during this process by approximately a factor 2.

When the gradient was created, the capillary was placed on a glass slide, thermally connected with mineral oil and the ends were sealed with epoxy droplets. For experiments, where the pT concentration of the gradient was measured quantitatively, two additional capillaries were placed on the slide, containing the reaction solutions with low pT and high pT (with the reference dye cascade blue) to have two reference points for the evaluation. When the glass slide was prepared, it was placed on the stage of the inverted microscope and also thermally connected with mineral oil. The plate was heated to 45°C when all other adjustments (fixing the glass slide with scotch tape, detecting the capillaries, fixing a curtain to reduce background) were finished. Prior to the measurement, the system was allowed to equilibrate for 2 minutes. Images were acquired every 10 minutes. To correct for inhomogeneous illumination in the EvaGreen- and Cy3.5-channels all images of a capillary were divided by the image of this capillary after 10 minutes.

13.2 Pseudotemplates in 0 dimensions

To observe the effects of the pseudotemplate, experiments in 0 dimensions were conducted with an initial species concentration of 10 nM, 25 nM autocatalytic template and a range of pseudotemplate concentrations. To prevent the degradation by the exonuclease, the used pseudotemplates were protected by 5'-biotinylation and bound streptavidin. The autocatalytic template had a Cy3.5-label on its 5'-end that was quenched when the autocatalytic species bound to it and thus allowed to monitor the replication dynamics over time (see Figure 24a-e). All positive controls with 0 nM suppressor showed replication and reached a constant signal within an hour. Increasing the pseudotemplate concentration slowed the replication down such that this state was reached at later times or it completely prevented replication. The polymerase concentration also significantly affected the kinetics of the process. For a polymerase concentration of 16 U/ml, the growth of the autocatalyst was strongly delayed already for a pseudotemplate concentration of 15 nM and completely suppressed for 25 nM (Figure 24a).

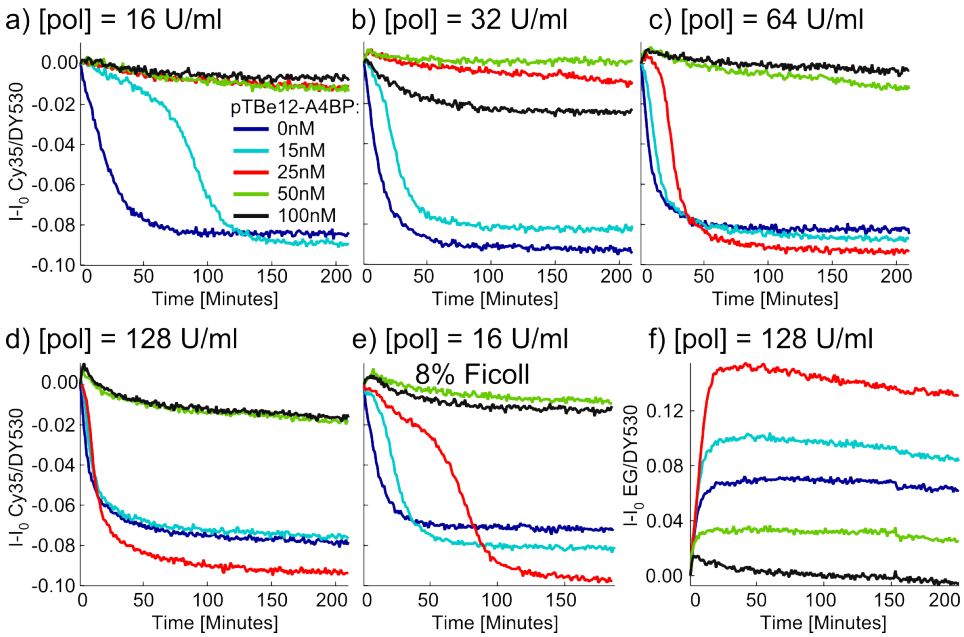


Figure 24: Pseudotemplate and Polymerase Range: For all graphs, the same range of pseudotemplates was used, which is shown in the legend of a). As expected, high pseudotemplate concentrations suppressed the growth of the species that had a starting concentration of 10 nM. a)-d) The threshold concentration of pseudotemplate changed with the polymerase concentration (pol). e) A similar effect was observed when ficoll was added, which either helped growth or weakened the suppression mechanism. f) and d) show the same experiment, but in different fluorescent channels. When an intercalating dye was added, higher pseudotemplate concentrations led to more signal. This was observed in the 1-dimensional experiments as well.

Doubling the polymerase amount changed the delay only for 15 nM pseudotemplate, but there was still no replication at the higher suppressor concentrations (Figure 24b). Replication despite the presence of 25 nM suppressor was observed for even higher polymerase concentrations of 64 U/ml and 128 U/ml

(Figure 24c,d). Not only the polymerase concentration had an effect on the threshold concentration of the suppressor, but also ficoll, which was used to change the viscosity (Figure 24e). In an experiment with 16 U/ml polymerase and 8% ficoll, replication was observed up to 25 nM of the suppressor. However, ficoll is very viscous and thus the pipetting error was probably larger in this experiment. In addition to the quenching of the Cy3.5-label, growth was also observed in the color channel of EvaGreen. Figure 24f shows the same experiment as Figure 24d but in the channel of the intercalating dye, which was also used for the 1-dimensional experiments. Just as in the case of the Cy3.5-channel, a steady state was reached, but the signal in this state varied from curve to curve. This has two main reasons: First of all, the Cy3.5-labeled templates had a rather low concentration, which led to a fast saturation of quenching and excess DNA molecules did not affect the signal. Secondly, the intercalating dye showed increased fluorescence when bound to double stranded DNA. In the presence of pseudotemplate, the species had more binding partners, which led to increased fluorescence.

13.3 Definition of Position and Width

When experiments were started, the autocatalysis proceeded very fast in the region where pT was zero or very low and the signal from the intercalating dye EvaGreen increased over a broad range. The system separated into an area with increased signal and an area with low signal (Figure 25a, top). The transition zone behaved like a wavefront that traveled from low suppressor concentrations on the left to the right where pT concentrations were higher. The increased pT concentration also led to an increased signal of the wavefront (compare Figure 24f). To show the dynamics of the system in a single image, kymographs were created by integrating the intensity of a capillary over the width of the capillary in every time step and concatenating these profiles below each other (Figure 25a, bottom).

Here, three identical experiments were performed with 12.5 nM exonuclease (Cap. 1 - Cap. 3) and kymographs were created in the EvaGreen- and the Cy3.5-channels. For further analysis, the EvaGreen channel was used due to higher intensity and less artifacts. Figure 25b shows smoothed and normalized wave profiles extracted from a kymograph (Cap. 1) after 200, 400, 600, 800 and 950 minutes. The position of the wavefront is defined as the x-value that gives the normalized intensity closest to 0.5 and has a negative slope when a line is fitted in the region around it. In this experiment, all suppressor gradients were created in the same way (Figure 25c) and showed a good overlap. To determine velocities, a 4th order polynomial was fitted to the positions of the wavefronts. This is done to avoid strong fluctuations in the velocity plots, which occurred in experiments with slow waves or lower signal. Following the position and the velocity of the wavefronts over time showed that the results from different capillaries were comparable and that the wavefront slowed down, when traveling to higher pT concentrations (Figure 25e). However, the waves never stopped in this experiment, which is probably caused by a mistake during gradient generation that shifted the gradient to the right. For the high pT reaction solution, 400 nM of the phosphothioate protected suppressor were added that usually led to a concentration of ~200 nM on the right side of the capillary, but was probably lower.

To determine the widths of the wavefronts, the slopes of the line fits to the wave profiles (Figure 25b)

were used. The width w is defined as $w = 1/|slope|$ and its time evolution is shown for two of the capillaries in Figure 25f. The third capillary (Cap. 3 in Figure 25a) showed a decrease in fluorescence on the right side for x>40 mm, which made proper normalization of the wavefronts problematic. The results showed that this was not a problem for wave localization, but the slope of the line fit changed and thus the width changed. Averaging the widths for all indicated times gave $w = 2.93 \pm 0.15$ mm and $w = 2.62 \pm 0.15$ mm for capillaries 1 and 2, but $w = 3.80 \pm 0.51$ mm for capillary 3.

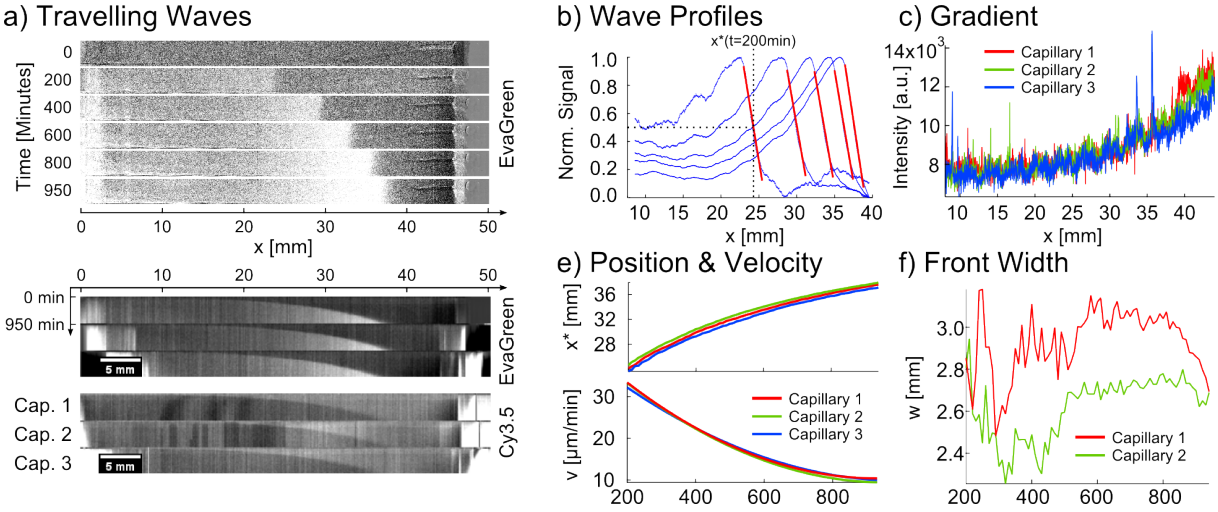


Figure 25: Travelling waves: a) Top: Images of a capillary in the EvaGreen channel at different times show a wavefront traveling from left to right. Bottom: By integrating the intensity over the width of the capillary, the spatiotemporal dynamics can be represented in kymographs. Here, three capillaries are shown in the EvaGreen- and Cy3.5 channels. b) Smoothed and normalized wave profiles (blue) after 200, 400, 600, 800 and 950 minutes with line fits around the region with the most negative slope (red). The fits allow to determine the wavefront position x*. c) The pT gradients were comparable between the three different capillaries. e) Time evolution of the position and velocity of the wavefronts show that they slowed down, but did not stop in the given time. f) Time evolution of the wavefront widths.

13.4 Wavefronts with Different Polymerase Concentrations

Changing the polymerase concentration changes the autocatalytic rate, thus the velocity of a traveling wave [34] and the threshold concentration of the suppressor (Figure 24). In an experiment that has been performed together with Adrian Zambrano, four different polymerase concentrations were used ranging from 8 U/ml to 64 U/ml and an exonuclease concentration of 12.5 nM. As for the previous experiment, the high pT reaction solution contained 400 nM of the phosphothioate protected suppressor.

Two problems with this experiments were the lower signal to noise ratio, when the amount of polymerase was reduced and the faster formation of parasites for higher polymerase concentrations (compare Section 7.1) which made it impossible to evaluate data after a while. Figure 26a shows the formation of parasites for a polymerase concentration of 64 U/ml after roughly 800 minutes. The parasites emerged independently in two different positions and led to an increase in fluorescence signal.

The increased fluorescence signal forming a horizontal line in the kymograph between minutes 970 and 1030 came from a temperature decrease to 40°C, which led to more stable double strands and thus more EvaGreen signal. Such temperature changes were done in different experiments to see if the system responses to external perturbations, which did not seem to be the case. The pT-gradients (Figure 26b) showed good overlap and the positions showed the expected behavior that wavefronts driven by faster autocatalysts with higher polymerase concentrations moved further into the suppressor gradients (Figure 26c). As shown in Figure 26d the velocities slowed down over time. However, in this plot the effects of the polymerase concentration on the velocity were not clear, especially the sample with 16 U/ml showed a large variation.

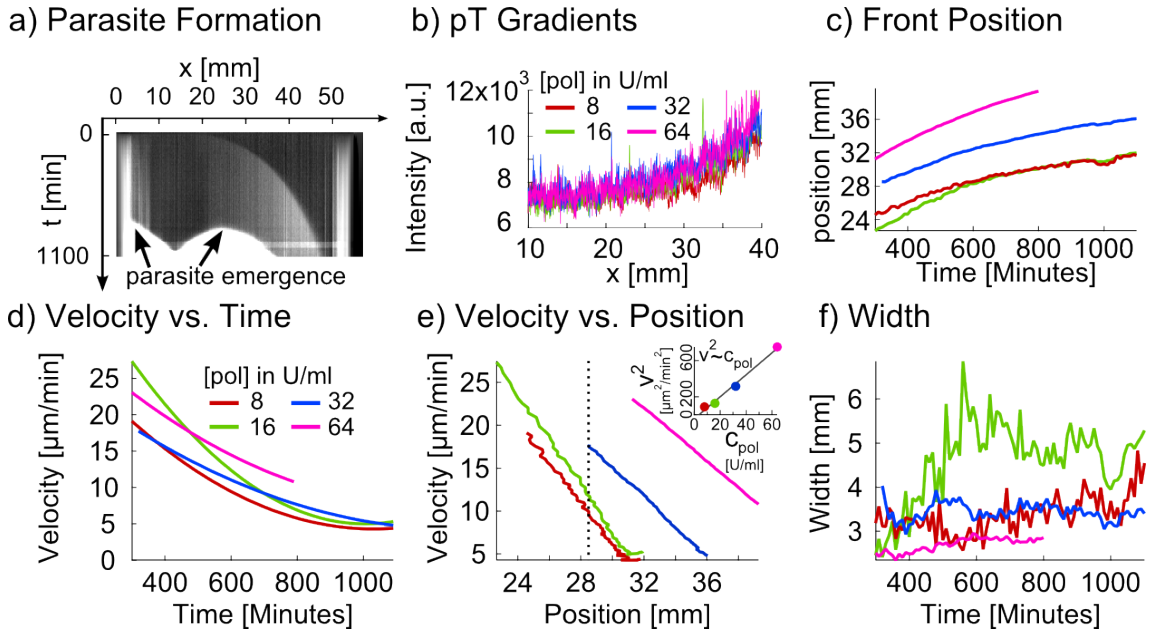


Figure 26: Wavefronts with different Polymerase Concentrations: a) The higher the polymerase concentration (here 64 U/ml), the earlier the parasites form. The parasitic species led to an increase in fluorescence and made it impossible to evaluate the kymopgraphs after their formation. b) The suppressor gradients in different capillaries show a good overlap. c) With increasing polymerase concentrations, the wave fronts move further into the suppressor gradients. d) The time evolution of the velocity shows a slowing down for all wavefronts. e) For a given position in the suppressor gradient, a higher polymerase concentration led to a higher velocity. The inset shows that the square of the velocity at the position marked with the dashed line scaled with the polymerase concentration in agreement with equation 7. For the 64 U/ml sample, a line was fitted to get the velocity at this position. f) The widths seemed to decrease with the polymerase concentration. However, the width for the 16 U/ml sample (green) was higher than the width of the 8 U/ml sample (red), but it is not clear whether the latter might have spread further when the experiment went on for a longer time.

Since the initial distribution of the autocatalyst was homogeneous at t=0 min, the initial replication took place at all the positions, where the suppressor concentration was low enough. This made it impossible to assign wavefront positions at early times. However, the polymerase concentration led to

different tolerated pT-concentrations and thus, when wavefronts formed, they were already at different positions in the gradient. Both contributions are assumed to have opposing effects: Higher polymerase concentrations is expected to increase the velocity under otherwise similar conditions, whereas higher suppressor concentration at the wavefront position should reduce the velocity. Thus, the velocites of the different wavefronts were compared at the same position in the capillary. Plotting the velocities against the position in the gradient showed the expected behavior (Figure 26e): Higher polymerase concentrations led to higher velocities at a given position in the gradient (assuming the gradients were identical). Comparing the velocities at the position 28.5 mm showed that the velocities were in agreement with eq. 7 (Figure 26e, inset). However, to obtain a value for the 64 U/ml sample at this position, a line had to be fitted in the velocity vs. position graph and the velocity was extrapolated to the intersection of the line with the dashed line in Figure 26e.

Determining the width worked well for the two samples with higher polymerase concentrations, but the signal was more noisy in the case of the lower concentrations (Figure 26f). The 8 U/ml sample showed a slight increase over time, which did not stop in the given time. The width of the 16 U/ml sample increased, but stabilized after roughly 600 minutes. From the supplementary information of *Rulands et al.* [59] the front width is expected to decrease with increasing autocatalytic rate and thus with the polymerase concentration. Averaging the values for the width over time starting from 600 minutes until the end of the measurement gave $w_8 = 3.53 \pm 0.42$ mm for the 8 U/ml sample. For the other polymerase concentrations, $w_{16} = 4.92 \pm 0.46$ mm, $w_{32} = 3.42 \pm 0.15$ mm and $w_{64} = 2.82 \pm 0.05$ mm. The error margins stem from standard deviations obtained from averaging the widths. For the higher polymerase concentrations, the expected trend was observed, whereas the lower polymerase concentrations led to lower widths. However, it is not clear whether the width has already reached its maximum value.

13.5 Wavefronts in Different Gradients

To investigate the influence of the shape of the suppressor gradients, experiments were made in which the minimal and maximal pT values of the gradients were changed. Here, the streptavidin protected suppressors were used.

Reference capillaries containing the reaction solutions with high and low suppressor concentrations were used to quantify the pT concentrations throughout the gradient. In the first step, the values for the high suppressor concentrations were 200 nM or 800 nM instead of the 400 nM used in the previous experiments. As usual, mixing the solutions resulted in a gradient, which was roughly a factor of 2 lower at the point with the highest concentration (Figure 27a). Two additional capillaries prepared with the pT solutions 400 nM and 600 nM were not evaluated because of artifacts, that led to a signal increase and caused problems during wave profile normalization.

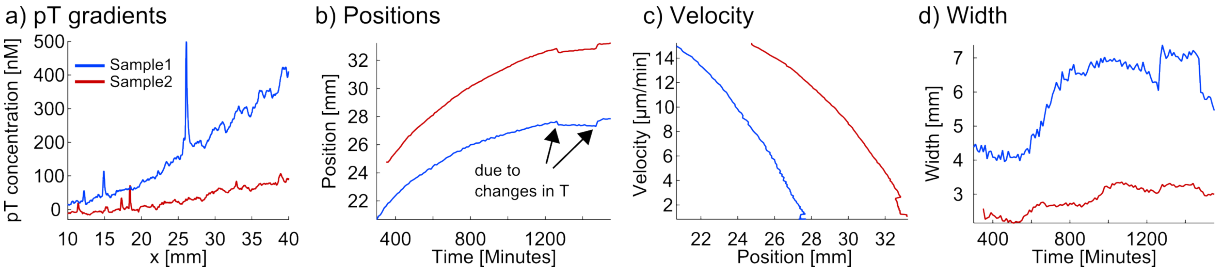


Figure 27: Change in the pT gradients on the high concentration side: a) Using reference capillaries allowed to quantify the pT concentrations in the gradients.In addition, the data had been smoothed. The peak in Sample 1 stems from a dirt particle. b) When the gradient was flatter, the wavefront moved further into it. However, the final positions corresponded to pT concentrations of 200.03 ± 13.4 nM in Sample 1 and 70.82 ± 14.43 nM in Sample 2. This difference was unexpected. Between minutes 1260 and 1470 the temperature was changed to 35°C, but this did not cause a long lasting change after the temperature was changed back to 45°C. c) Plotting the velocity vs. the front position shows that the dynamics in the gradients differed from each other. d) The wavefront was found to be wider in the steeper gradient.

Tracking the positions of the wavefront showed the expected result that the front moved further in the shallow gradient (Figure 27b). However, the pT concentrations in the final positions showed a large and unexpected difference. The final positions of the fronts corresponded to a pT concentration of 200.03 ± 13.4 nM for the steeper gradient and 70.82 ± 14.43 nM for the shallower gradient (error values from averaging over a 2 mm region). Thus, the threshold concentration in the 1-dimensional system seems to depend on the gradient shape as well. Plotting the velocities of the wavefront vs. the position showed that the dynamics were similar, but shifted in position (Figure 27c). The widths of the wavefronts were higher in the steeper gradient (Figure 27d).

In another experiment, the high pT concentration was varied, but the low pT concentration was set to 30 nM, which required an initially inhomogeneous distribution of the autocatalytic species (Figure 28a). Usually, a homogeneous distribution of 1 nM was used, but this would not allow replication because the lowest suppressor concentration was already too high (compare Figure 24). For this experiment, the gradients were prepared with two reaction solutions of different suppressor concentrations, but without the autocatalytic species. In a separate reaction solution which contains 30 nM of the suppressor the autocatalytic species was added at a concentration of 100 nM. To add this mix to the capillary that did not yet contain the autocatalytic species, a portion of the reaction solution in the low pT region was removed with a tissue. The mixture that contained the species was then used to refill the capillary from the side, where the sample was removed. This excess allowed the formation of the wavefront despite the relatively high pT concentration throughout the capillary (Figure 28b).

The kymograph in Figure 28c shows the initially inhomogeneous distribution with the high species concentration on the left side. The wavefront formed and moved further to the right, whereas the initial excess of the species in the region for x<5 mm vanished. Figure 28d shows that there were two different stages in the process of wave propagation, a faster one in the beginning and a slower one after around 800 minutes. The final positions of the fronts from Sample 1-4 were 30.16 ± 13.85 nM, 24.39 ± 5.70 nM,

45.74 ± 9.78 nM and 62.23 ± 5.36 nM. For the velocity plots, there were changes after 1500 minutes, which were an artifact of the polynomial fit to the positions (Figure 28e). The wavefronts did not only become slower, but also the width increased in a way that has not been observed in other experiments so far (Figure 28f). This increase did not stop in the experimental time of 2000 minutes.

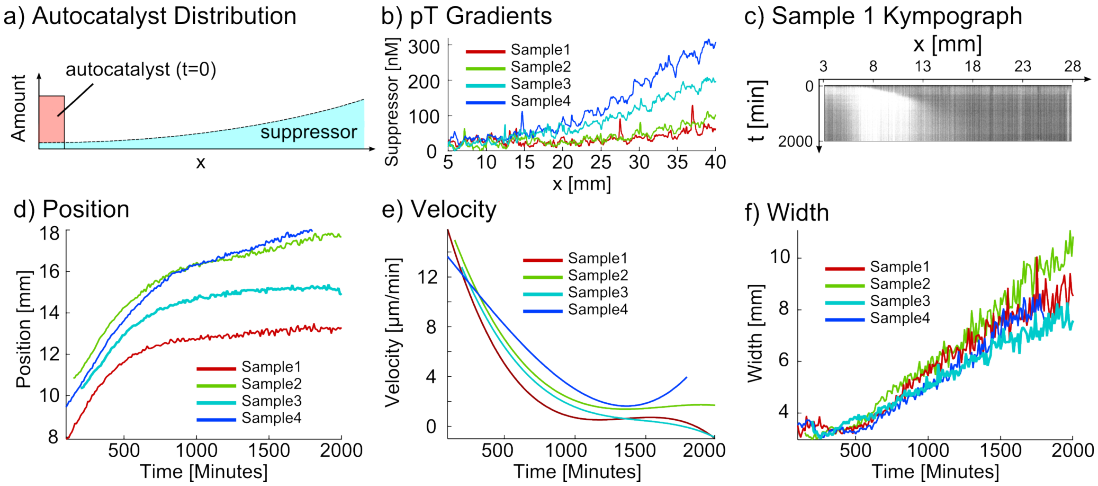


Figure 28: Inhomogeneous distribution of autocatalytic species in a gradient: a) To allow replication in a pT gradient, which had a minimal suppressor concentration of 30 nM, an inhomogeneous distribution of the autocatalytic species was chosen. b) The resulting gradients. c) The kymograph shows the initially high species concentration on the left side of the capillary, which was necessary to form the wavefront. After a while the species concentration became lower in this area, but the wavefront persisted. d) The wavefronts did not move far into the gradient and two regimes were distinguished. A fast phase until ~500 minutes and a slower phase after that. e) The velocity plot also shows the slowing down. The final velocity changes were artifacts from the fitting procedure. f) The widths of the wavefronts started to increase at around 500 min and did not reach a stable value in the experimental time.

14 Testing Shorter Pseudotemplates

The presented experiments were performed with pseudotemplates that bind the autocatalytic species and allow the polymerase to attach four additional bases (Figure 29a, left). This createds a double stranded 16-mer that was very stable at the used temperatures of 35°C-50°C (including the perturbations). Using shorter pseudotemplates reduced the length of the double stranded complex and increased dehybridization rates.

To test how the stability of this complex affects the degradation of pseudospecies, an experiment on pseudotemplates with different lengths was performed (Figure 29a,b). 100 nM of the pseudotemplates pT-A4, pT-A3 and pT-A2 were incubated separately in the enzyme-free reaction solution with 100 nM of their respective pseudospecies at 45°C. This led to complexes of 16 bp, 15 bp and 14 bp. In addition, pT-A4 was incubated with the autocatalytic species, which bound with 12 bp. After 20 minutes, exonuclease was injected to a final concentration of 30 nM to all samples and the degradation dynamics

were followed in the EvaGreen channel (Figure 29b). The signal decreased upon degradation of single stranded species and pseudospecies. The shorter the double stranded complex, the faster the degradation process. However, in the observed time of more than 800 minutes the degradation of the long pseudospecies was still not completed. In a control experiment, the samples also showed a signal decrease when water was added instead of exonuclease, but to a lower extent (Figure 29c). A control with the protected pseudotemplates in the presence of exonuclease led to a constant signal, indicating that only the pseudospecies was degraded as expected.

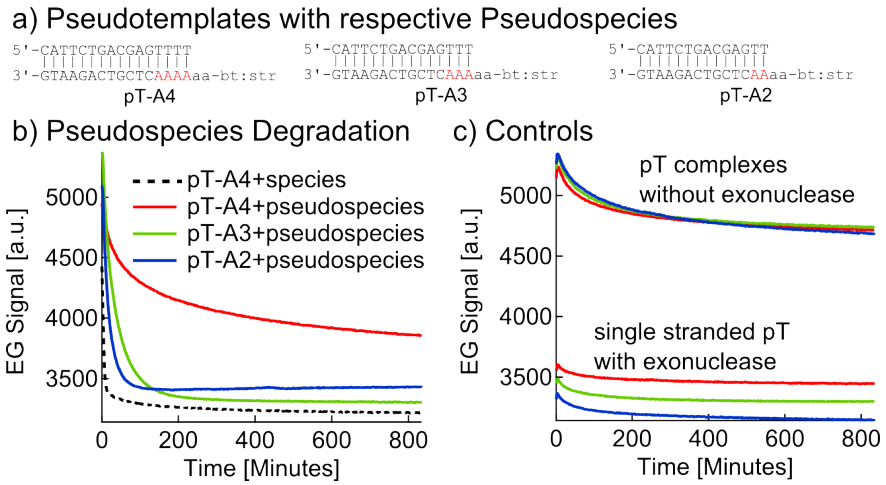


Figure 29: Different Pseuotemplates: a) Pseudotemplates of different lengths with their respective pseudospecies. b) The complexes shown in a) and pT-A4 with the autocatalytic species were incubated with exonuclease. The fluorescence decrease over time indicates the degradation of free (pseudo)species and thus the stability of the double strands.

The stability of the 16 bp pseudotemplate-pseudospecies complex is a probable explanation for the stability of the wavefront to temperature changes. Even if the free autocatalytic species were degraded upon temperature changes, the signal coming from the long double strand might have covered this effect. Using the shorter pseudotemplates might allow to explore the stability of the wavefronts by allowing a measurable response to perturbations and compare the results to theoretical predictions [59]. The stable binding of the used pseudotemplate pT-A4 also reduced the amount of unbound suppressor, which was assumed to be constant over time.

15 Conclusion

Despite the simplicity of the experimental system that is comprised of a degradable autocatalyst that reacts to a gradient of suppressor, a variety of phenomena can be studied. Some of them are in agreement with theoretical assumptions, whereas others are puzzling. The increasing suppressor concentration at the wavefront position led to a decrease in velocity. However, the suppressor concentrations that the wavefronts reached covered a surprisingly large range from 24 nM to 200 nM, depending on the shape

of the gradient.

Increasing the autocatalytic strength by changing the polymerase concentration also resulted in tolerating higher suppressor concentrations as well as higher velocities. The front widths showed an increasing trend when autocatalysis was weakened. The width dependence on the gradient shape remains unclear. The broadest wavefronts have been measured for very steep, but also for very shallow gradients. Especially the experiments with the shallow gradients showed a broadening of the wavefronts even after 2000 minutes. However, in this experiment a basal suppressor concentration of 30 nM was used which required an inhomogeneous initial distribution of the autocatalytic species. This might also have an influence on the behavior of the front, implying that the system has a memory.

To perturb the system, the temperatures were changed in several experiments when the wavefronts were already very slow. Temperature changes alter the binding kinetics of DNA and have an impact on the enzymes. Thus, changes in the position and width of the wavefront were expected. Surprisingly, neither increasing the temperatures to 50°C nor decreasing to 35°C had a long lasting effect on the position or the width of the front after the temperatures were changed back to the standard temperature of 45°C. A possible explanation is that the pseudotemplate-pseudospecies double strand was very stable. Thus, possible changes in the concentration of the autocatalytic species might have been unobserved, because of the strong fluorescent signal of the pseudotemplate-pseudospecies complex.

16 Acknowledgments

I want to thank a lot of people for a lot of support. I would have liked to do this with some delicious chocolate covered apples, but last time i tried this almost lead to poisoning of the entire Braun lab. So maybe it is better to do it in a written form.

- Thanks to **Dieter** for not only teaching a lot about science, but also about coffee, sledge riding and paragliding. Providing this kind of atmosphere in the lab allowed me to really enjoy the work. Also thanks for the trust and support to let me follow own ideas.
- Also thanks to the Paris folks **André, Jean-Christophe, Adrian and Jonathan**. It was a lot of fun working with you and, hopefully, we can continue for some more time.
- Thanking all the (former) postdocs, PhD-, Master- and Bachelor students is actually a tough task, since there are a lot of people with whom I did a lot of different activities in the lab and in my free time, but I will try my best: Thanks to **Manu, Zhenya, Noel and Victor** for a lot of fun bouldering sessions. Thanks to **Matzi and Lonzo** for the real-life soap-opera. Thanks to **Simon and Christof** for making their knowledge and skills accessible to less talented people. Thanks to **Ike** for making blueberry pancakes and mozzarella with tomato. Thanks to **Franzi** for teaching me how to use a pipette. Thanks to **Susi** for a lot of fun during building the microscope. Thanks to **Thomas, Sebastian, Tobi and Moritz** for conducting some experiments for me. Also thanks to **Mandy, Steffi, Mikey, Emil, Julian, Mara and Jonathan** for the great atmosphere in the lab. Thanks to all of you for making me enjoy being in the lab.
- Thanks for lots of helpful hints on the manuscript from **Zhenya, Steffi, Matzi, Ike and Manu**.
- Thanks to **Erwin and Marianne** for a lot of useful discussion on my projects.
- Also thanks to all the people from groups of **Gaub, Lipfert, Nash, Rädler, Lidl and Nickel** for being good neighbours. Especially to **Peter** for having tons of coffee and whoever joined in the weekly football sessions.
- Thanks a lot to my MGK/IRTG colleagues **Marilena, Katharina, Fabian and Mario** for doing the actual work in the graduate program, for which I get a lot of credit. Also thanks to the people from **NIM** and **CeNS** for a lot of help.
- Thanks to my Schwabing-Connection **André and Flo**, my friends from Salzburg and Munich and the snobs in Vienna.
- And of course thanks to **my family** for a lot of support.

References

[1] Alberts B, et al. (2002) *Molecular Biology of the Cell, Fourth Edition*.

[2] Kruger K, et al. (1982) Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* 31:147–157.

[3] Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35:849–857.

[4] Gilbert W (1986) The RNA World. *Nature* 319:618.

[5] Hud NV, Cafferty BJ, Krishnamurthy R, Williams LD (2013) The origin of RNA and "my grandfather's axe". *Chemistry & biology* 20:466–74.

[6] Joyce GF (2009) Evolution in an RNA world. *Cold Spring Harbor symposia on quantitative biology* 74:17–23.

[7] Powner MW, Gerland B, Sutherland JD (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* 459:239–242.

[8] Ferus M, et al. (2015) High-energy chemistry of formamide: A unified mechanism of nucleobase formation. *Proceedings of the National Academy of Sciences of the United States of America* 112:657–662.

[9] Kanavarioti A, Monnard PA, Deamer DW (2010) Eutectic Phases in Ice Facilitate Nonenzymatic Nucleic Acid Synthesis. *Astrobiology* 1:271–281.

[10] Mast CB, Schink S, Gerland U, Braun D (2013) Escalation of polymerization in a thermal gradient. *Proceedings of the National Academy of Sciences of the United States of America* 110:8030–8035.

[11] Mamajanov I, et al. (2014) Ester formation and hydrolysis during wet-dry cycles: Generation of far-from-equilibrium polymers in a model prebiotic reaction. *Macromolecules* 47:1334–1343.

[12] Morasch M, Mast CB, Langer JK, Schilcher P, Braun D (2014) Dry polymerization of 3',5'-cyclic GMP to long strands of RNA. *Chembiochem : a European journal of chemical biology* 15:879–83.

[13] Ferris JP (2005) Mineral Catalysis and Prebiotic Synthesis: Montmorillonite-Catalyzed Formation of RNA. *ELEMENTS* 1:145–150.

[14] Mutschler H, Wochner A, Holliger P (2015) Freeze-thaw cycles as drivers of complex ribozyme assembly. *Nature chemistry* 7:502–508.

[15] von Kiedrowski G, Wlotzka B, Helbing J, Matzen M, Jordan S (1991) Parabolic Growth of a Self-Replicating Hexadeoxynucleotide Bearing a 3'-5'-Phosphoamidate Linkage. *Angewandte Chemie* 30:423–426.

[16] Obermayer B, Krammer H, Braun D, Gerland U (2011) Emergence of information transmission in a prebiotic RNA reactor. *Physical Review Letters* 107:018101.

[17] Leu K, et al. (2013) Cascade of reduced speed and accuracy after errors in enzyme-free copying of nucleic acid sequences. *Journal of the American Chemical Society* 135:354–66.

[18] Lincoln Ta, Joyce GF (2009) Self-sustained replication of an RNA enzyme. *Science (New York, N.Y.)* 323:1229–1232.

[19] Aniela Wochner, James Attwater, Alan Coulson PH (2011) Ribozyme-Catalyzed Transcription. *Science* 061911:2009–2012.

[20] Vaidya N, et al. (2012) Spontaneous network formation among cooperative RNA replicators. *Nature* 491:72–77.

[21] Attwater J, Wochner A, Holliger P (2013) In-ice evolution of RNA polymerase ribozyme activity. *Nature Chemistry* 5:1011–1018.

[22] Ivica Na, et al. (2013) The paradox of dual roles in the RNA world: Resolving the conflict between stable folding and templating ability. *Journal of Molecular Evolution* 77:55–63.

[23] Baaske P, et al. (2007) Extreme accumulation of nucleotides in simulated hydrothermal pore systems. *Proceedings of the National Academy of Sciences of the United States of America* 104:9346–9351.

[24] Mast CB, Braun D (2010) Thermal Trap for DNA Replication. *Physical Review Letters* 104:188102.

[25] Krammer H, Möller FM, Braun D (2012) Thermal, autonomous replicator made from transfer RNA. *Physical Review Letters* 108:1–5.

[26] Mülhardt C (2009) *Der Experimentator: Molekularbiologie/ Genomics* Vol. 6. Auflage, pp 90–92.

[27] Mills, D. R., Peterson, R. L., & Spiegelman S (1967) An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proceedings of the National Academy of Sciences of the United States of America* 58:217–224.

[28] Kreysing M, Keil L, Lanzmich S, Braun D (2015) Heat flux across an open pore enables the continuous replication and selection of oligonucleotides towards increasing length. *Nature Chemistry* 7:203–208.

[29] Veening JW, et al. (2008) Bet-hedging and epigenetic inheritance in bacterial cell development. *Proceedings of the National Academy of Sciences of the United States of America* 105:4393–4398.

[30] Acar M, Mettetal JT, van Oudenaarden A (2008) Stochastic switching as a survival strategy in fluctuating environments. *Nature genetics* 40:471–475.

[31] Rulands S, Jahn D, Frey E (2014) Specialization and Bet Hedging in Heterogeneous Populations. *Physical Review Letters* 113:108102.

[32] Tautz D, Schlötterer C (1994) Simple sequences. *Current Opinion in Genetics and Development* 4:832–837.

[33] Lorenz R, et al. (2011) ViennaRNA Package 2.0. *Algorithms for Molecular Biology* 6:26.

[34] Zadorin AS, Rondelez Y, Galas JC, Estevez-Torres A (2015) Synthesis of Programmable Reaction-Diffusion Fronts Using DNA Catalyzers. *Physical Review Letters* 114:068301.

[35] Seidel SAI, Markwardt NA, Lanzmich SA, Braun D (2014) Thermophoresis in nanoliter droplets to quantify aptamer binding. *Angewandte Chemie - International Edition* 53:7948–7951.

[36] Seidel S (2014) Ph.D. thesis.

[37] Braun D, Goddard NL, Libchaber A (2003) Exponential DNA replication by laminar convection. *Physical Review Letters* 91:158103.

[38] Duhr S, Braun D (2006) Why molecules move along a temperature gradient. *Proceedings of the National Academy of Sciences of the United States of America* 103:19678–82.

[39] Maynard Smith J (1982) Evolution and the Theory of Games. *Darwin* 13:224.

[40] Sinervo B, Lively CM (1996) The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* 380:240–243.

[41] Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335–338.

[42] Reichenbach T, Mobilia M, Frey E (2007) Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games. *Nature* 448:1046–1049.

[43] Reichenbach T, Mobilia M, Frey E (2008) Self-organization of mobile populations in cyclic competition. *Journal of Theoretical Biology* 254:368–383.

[44] Kerr B, Riley MA, Feldman MW, Bohannan BJM (2002) Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* 418:171–174.

[45] Vanag VK, Epstein IR (2001) Pattern formation in a tunable medium: the Belousov-Zhabotinsky reaction in an aerosol OT microemulsion. *Physical Review Letters* 87:228301.

[46] Baccouche A, Montagne K, Padirac A, Fujii T, Rondelez Y (2014) Dynamic DNA-toolbox reaction circuits: A walkthrough. *Methods* 67:234–249.

[47] Montagne K, Plasson R, Sakai Y, Fujii T, Rondelez Y (2011) Programming an in vitro DNA oscillator using a molecular networking strategy. *Molecular systems biology* 7:466.

[48] Kim J, White KS, Winfree E (2006) Construction of an in vitro bistable circuit from synthetic transcriptional switches. *Molecular systems biology* 2:68.

[49] Franco E, et al. (2011) Timing molecular motion and production with a synthetic transcriptional clock. *Proceedings of the National Academy of Sciences of the United States of America* 108:E784–E793.

[50] Padirac A, Fujii T, Rondelez Y (2012) Bottom-up construction of in vitro switchable memories. *Proceedings of the National Academy of Sciences of the United States of America* 109:4–6.

[51] Fujii T, Rondelez Y (2013) Predator-prey molecular ecosystems. *ACS nano* 7:27–34.

[52] Padirac A, Fujii T, Estévez-Torres A, Rondelez Y (2013) Spatial waves in synthetic biochemical networks. *Journal of the American Chemical Society* 135:14586–92.

[53] Tan E, et al. (2008) Specific versus nonspecific isothermal DNA amplification through thermophilic polymerase and nicking enzyme activities. *Biochemistry* 47:9987–9999.

[54] Roekel HWHV, et al. (2015) Automated Design of Programmable Enzyme-Driven DNA Circuits. *ACS Synthetic Biolog7* 4:735–745.

[55] Padirac A, Fujii T, Rondelez Y (2012) Quencher-free multiplexed monitoring of DNA reaction circuits. *Nucleic Acids Research* 40:1–7.

[56] Padirac A (2012) Ph.D. thesis.

[57] Aubert N, Mosca C, Fujii T, Hagiya M, Rondelez Y (2014) Computer-assisted design for scaling up systems based on DNA reaction networks. *Journal of the Royal Society, Interface / the Royal Society* 11:20131167.

[58] Johnston DS, Nüsslein-Volhard C (1992) The origin of pattern and polarity in the Drosophila embryo.

[59] Rulands S, Klünder B, Frey E (2013) Stability of Localized Wave Fronts in Bistable Systems. 038102:1–5.

[60] Zadorin A, Rondelez Y, Gines G, Galas JC, Estevez-Torres A (2016) Synthesis of a reaction-diffusion French flag pattern. *submitted*.

[61] Montagne K, Gines G, Rondelez Y (2016) Degradation-controlled nonlinearities in DNA networks. *submitted*.

[62] Murray JD (2002) *Mathematical Biology : I . An Introduction , Third Edition* Vol. 1, p 576.

Part IV
Appendix

A Sequences used in Part I

For the majority of the presented experiments, the following sequences and their respective complementary sequences have been used:

Name	Length	Sequence from 5' to 3'	Primer
Short Hairpin	84	**TTGACTTAGGTAGATGAC**ATGAATAAGATCAGAGCAATCGACACTTATATGCATAGTCGCCTTACA**GTCATCTACCTAAGTCAA**	P1
Long Hairpin	124	**TAGGCGTCATTTACTCACACATTTTATGTCGATACTCT**TGCTCTGCTTATAGCAATTTCAAGCGACTTAAATGCATATGGTGTTGT**AGAGTATCGACATAAAATGTGTGAGTAAATGACGCCTA**	P2
Short Cooperating Hairpin	84	**GATCTAGTACCATCCATT**TGCTCTGCTTATAGCAATTTCAAGCGACTTAAATGCATATGGTGTTGT**AATGGATGGTACTAGATC**	P3
Crossbreed	104	GATCTAGTACCATCCATTTGCTCTGCTTATAGCAATTTCAAGCGACTTAAATGCATATGGTGTTGTAGAGTATCGACATAAAATGTGTGAGTAAATGACGCCTA	P2
Crossbreed [Complement]	104	TAGGCGTCATTTACTCACACATTTTATGTCGATACTCTACAACACCATATGCATTTAAGTCGCTTGAAATTGCTATAAGCAGAGCAAATGGATGGTACTAGATC	P3

Table 3: Sequence species used in Part I. Complementary strands for the hairpins are not shown here, except for the crossbreed, where the primer binding sites differ. Primer sequences can be found in Table 4. For linear growth experiments, the crossbreed was used with the Primer P2.

Name	Length	Sequence from 5' to 3'
P1	18	TTGACTTAGGTAGATGAC
P2	20	TAGGCGTCATTTACTCACAC
P3	18	GATCTAGTACCATCCATT

Table 4: Primers used in Part I. Sequences of the species can be found in Table 3

B Parameters for the Crossbreed Finite-Element Model from Part I

Parameter	Value
$D_{Long\,Hairpin}$	$8.24 \cdot 10^{-11}\ \frac{m^2}{s}$
$D_{Short\,Hairpin}$	$1.16 \cdot 10^{-10}\ \frac{m^2}{s}$
D_{CB}	$9.60 \cdot 10^{-11}\ \frac{m^2}{s}$
γ	$3.73 \cdot 10^{-3}\ \frac{m^2}{s}$
$r_{Long\,Hairpin}$	$4.22 \cdot 10^{-4}\ \frac{m^2}{s}$
$r_{Short\,Hairpin}$	$6.95 \cdot 10^{-4}\ \frac{m^2}{s}$
r_{CB}	$6.25 \cdot 10^{-4}\ \frac{m^2}{s}$
K	$41.4\,nM$
δ	$11.3\ \frac{m^3}{mol \cdot s}$

Table 5: Parameters used in the finite-element model described in Section 3.3

C Sequences used for Sequence Screen in Part II

Name	Sequence from 5' to 3'	Set
inp1	GTAGAGTCAACG	1
inp2	AATGAGTCTGTG	1
inp3	GAAGAGTCATGC	1
inp4	CAAGAGTCTAGC	2
inp5	GAAGAGTCATGG	2
inp6	GTGGAGTCTAGT	2
inp7	AATGAGTCTTCG	2
inp8	TAGGAGTCAGTT	2
inp9	GTTGAGTCTGTA	2
inh1	GTCAACGGTAGAGaa	1
inh2	GTCTGTGAATGAGaa	1
inh3	GTCATGCGAAGAGaa	1
inh4	GTCTAGCCAAGAGaa	2
inh5	GTCATGGGAAGAGaa	2
inh6	GTCTAGTGTGGAGaa	2
inh7	GTCTTCGAATGAGaa	2
inh8	GTCAGTTTAGGAGaa	2
inh9	GTCTGTAGTTGAGaa	2
ac1	CGTTGACTCTACCGTTGACTCTAC	1
ac2	CACAGACTCATTCACAGACTCATT	1
ac3	GCATGACTCTTCGCATGACTCTTC	1
ac4	GCTAGACTCTTGGCTAGACTCTTG	2
ac5	CCATGACTCTTCCCATGACTCTTC	2
ac6	ACTAGACTCCACACTAGACTCCAC	2
ac7	CGAAGACTCATTCGAAGACTCATT	2
ac8	AACTGACTCCTAAACTGACTCCTA	2
ac9	TACAGACTCAACTACAGACTCAAC	2

Table 6: Oligonucleotides used in Section 9.2, Figure 18: All oligonucleotides are without modifications. Autocatalytic templates are termed 'ac1'-'ac9' (24-mers), the inputs are 'inp1'-'inp9' (12-mers) and the inhibitors 'inh1'-'inh9' (15-mers). The underlined sequences are recognition sites of the nickase and small letters denote spacer bases that should not bind to prevent elongation of the inhibitors.

Name	Sequence from 5' to 3'	Comment
Input Sequences		
inp6	GTTGAGTCTGTA	
inp7'	GAATGAGTCTTCG	sequence: 5'-G+inp7-3' (Fig. 20)
inp9	GTTGAGTCTGTA	
Inhibitors		
+inh9+	AGTCTGTAGTTGAGTaa	inhibitor for species 9 (Fig. 19)
inh9++	GTCTGTAGTTGAGTCaa	inhibitor for species 9 (Fig. 19)
+inh9++	AGTCTGTAGTTGAGTCaa	inhibitor for species 9 (Figs. 19, 21)
Autocatalytic Templates		
btac6DY	B-aaACTAGACUCCACACTAGACTCCAC-DY530	strong inp6 template (Fig. 19)
btac7DY	B-aaCGAAGACUCATTCGAAGACTCATT-DY530	strong inp7 template (Fig. 19)
btac8DY	B-aaAACTGACUCCTAAACTGACTCCTA-DY530	strong inp8 template (Fig. 19)
btac9DY	B-aaTACAGACUCAACTACAGACTCAAC-DY530	strong inp9 template (Figs. 13, 19, 21)
bt.ac6-1.J	B-aaACTAGACUCCACACTAGACTCCA-JOE	medium inp6 template (Fig. 20)
ac7'-1	B-aaCGAAGACUCATTCCGAAGACTCAT-p	medium inp7' template (Fig. 20)
bt.ac9-1.J	B-aaTACAGACUCAACTACAGACTCAA-JOE	medium inp9 template (Figs. 20, 21)
bt.ac9-2.J	B-aaTACAGACUCAACTACAGACTCA-JOE	weak inp9 template (Fig. 20, 21)
Reporters		
btRep6	B-ACTAGACUCCAC-BMN5	reporter for inp6 (Fig. 20)
btRep7'	B-GAAGACUCATTC-BMN3	reporter for inp7' (Fig. 20)
btRep9	B-TACAGACUCAAC-FAM	reporter for inp9 (Figs. 13, 20, 21)

Table 7: Oligos for used for further experiments after extracting a set from Table 6: From the sequence screen (Section 9.2) the inputs inp6 and inp9 were chosen as two species for rock-paper-scissors. The third species is a variant from inp7 with a G on the 5'-end for better reporting.Nickase recognition sites are underlined (uracil modification 'U' prevents recognition by nickase). 'B' stands for biotin. Small letters indicate spacer bases to prevent inhibitor elongation by adding mismatches or to take the sterical hindrance into account that the streptavidin (when binding to 'B') poses to the polymerase. Elongation of templates and reporters is prevented by a phosphate 'p' or by one of the dyes DY530, JOE, BMN3 and BMN5. The dyes can of course also be used for reporting when they are quenched by the respective input.

Manuscript in preparation

Stepwise cooperation of molecular replicators

Georg C. Urtel and Dieter Braun*

Systems Biophysics, Physics Department, NanoSystems Iniative Munich and Center for Nanoscience
Ludwig-Maximilians-Universität München, Amalienstrasse 54, 80799 München, Germany

* corresponding author

Abstract. How can molecules with short lifetimes preserve their information over millions of years? For evolution to occur, information-carrying molecules need to replicate their information into new molecules before they themselves degrade. Our experiments reveal the advantages of cooperation between replicating molecules. The natural formation of crossbreeds from unfinished replicates allows two replicators to combine their replication initiator sites, boost their fitness and overcome exponential degradation. The mechanism shows a robust, stepwise evolution towards faster and more complex replicators. The gain of replicator fitness through cooperation sheds light onto the transition from prebiotic chemistry to the ever increasing complexity at the onset of life.

Introduction. The exploration of evolutionary dynamics under controlled conditions has led to fascinating insights into the mechanisms of Darwinian evolution [1–5]. Compared to the passive observation of evolution in nature, a synthetic, experimental approach offers a number of advantages: the studied ecosystems can be reduced and manipulated to focus on key aspects under defined boundary conditions, e.g. by controlling the supply of resources. Most importantly, evolution can be readily studied on the molecular level, minimizing the complexity involved with living organisms and focusing on central mechanisms [6–13].

The genetic substrate of molecular evolution experiments are oligonucleotides such as RNA or DNA. On a technical level, nucleic acids are a good target to study evolution with for several reasons. Since the secondary structure and dimer formation of DNA or RNA is predictable to a larger extent, the phenotypical features can be directly related to the genotype. Also, their chemical and thermodynamic properties are well known and can be detected over a wide range of concentrations, allowing to access both the deterministic regime and the single molecule regime where fluctuations become important [14]. Since common consensus states that life began via the evolution of nucleic acids [15,16] such experiments allow one to analyze the early stages of evolution.

One of the major problems to make the transition from inanimate matter to life, is the steady fight against the increase of entropy and the struggle to avoid thermodynamic equilibrium [17]. The genetic information content of replicators has to be preserved against processes such as diffusion, dilution, or molecular degradation. For example, in the case of oligonucleotides, base pairing can reduce cleavage and enhance sequence memory due to mutual interaction [18]. Intermolecular interactions are especially important for the functionality of ribozymes, which can be bred to replicate oligonucleotides [8,9,19]. In order to do so, ribozymes must hybridize to a replication template. Generally, these recognition mechanisms are crucial to initiate replication and to prevent unwanted side reactions such as parasite replication[19] in cooperating ecosystems [9,20].

All known RNA or DNA replicating mechanisms rely on a replication initiator site. At this location, the base-by-base replication machinery docks initially and proceeds to create double stranded oligonucleotides from single stranded ones. In our experiments, a short oligonucleotide - often termed a primer - allows the Taq replication enzyme to start the replication. For RNA-based replicases. very similar initiator mechanisms are used.

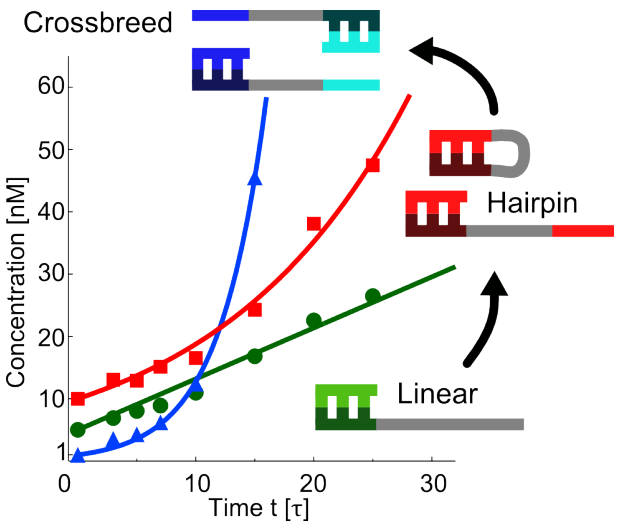


Figure 1. Three modes of replication. *(●, Linear): An oligonucleotide with one replication initiator site produces copies which cannot replicate further due to the lack of a binding site in the replicate. Replicates increase only linear in time. (■, Hairpin): Hairpins with one initiator site replicate exponentially, but the hairpin structure blocks of the initiator site. (▲, Crossbreed): Linear DNA replicates exponentially if two initiator sites are present at both ends. With the absence of a secondary structure, they easily outcompete hairpins with a single initiator site.*

In this paper, we demonstrate the unexpected finding that two hairpin replicators can readily cooperate by a simple mutation mechanism. They spontaneously make the important transition from one to two initiator sites. For this cooperation step to happen, the loop sequences of the hairpins must exhibit sequence similarities. A fast exponential replicator then emerges that combines the genetic information of both hairpins.

If an RNA ribozyme wants to copy itself, the product is its base pair complement and will most likely lack catalytic activity. Initially, its only function would be to act as a template for subsequent replication [21]. However, this again requires a replication initiator site, or else the system would merely accumulate non-functional oligonucleotides that cannot be further processed by the ribozymes. In this paper, we distinguish between three types of replication that depend on the initiator sites and secondary structures of the oligonucleotides (Fig. 1) and have an effect on the ratio of the growth rate to degradation rate r/γ:

1. If there is only one initiator site, linear strands without secondary structure produce a template that cannot replicate further due to the lack of a new initiator site.
2. If a replicator has a hairpin structure, its replicated complement has the same initiator site and exponential replication becomes possible. However, the secondary structure limits the access to the initiator site, causing slow replication ($r/\gamma < 1$).

The replication behavior of these two replicator types is shown experimentally in Figure 1. Both have only one initiator site and offer only slow replication. The third replicator type is faster and more complex:

3. If a replicator has two initiator sites, one at the initial strand and one at its complement, fast exponential growth is possible ($r/\gamma > 1$). However, the spontaneous emergence of two different, functional initiator sites is an unlikely event.

Material and Methods. We chose a polymerase chain reaction (PCR) as an experimental proxy for exponential replication under oscillatory dehybridizing conditions. Temperature oscillation was implemented by microscale thermal convection [22]. Phusion High Fidelity DNA Polymerase was used in a final concentration of 8 U/ml in 1x HF Buffer with 0.2 mM NTPs (all from New England Biolabs) and 1x EvaGreen (Biotum). The reaction volume was either 5 µl or 10 µl. Each primer was added with a concentration of 250 nM, and template concentration varied from 1 nM to 100 nM. The samples were cycled in a C1000 Touch™ Thermal Cycler with a CFX96 detection system (Biorad). For a PCR cycle, samples were denaturated for 10 s at 95 °C (initial cycles were 3 min longer) and annealed for 36 s at 67°C. Including heating, cooling, and readout, the total time for one cycle was 72s. The lid temperature was set to 105°C to prevent condensation and the samples were cooled down to 1°C or quenched with ice water after the last cycle. All used oligonucleotides were purchased from biomers.net and HPLC purified. Three hairpin species and one linear species together with their complements were used (stem sequence is bold, initiator site sequence is underlined):

The short hairpin 5'-**GATCTAGTACCATCCATT**TGCTCTGCTTATAGCAATTTCAAGCGACTTAAA TGCATATGGTGTTGT**AATGGATGGTACTAGATC**-3', the short control hairpin 5'-**TTGACTTAGGTAG-ATGAC**ATGAATAAGATCAGAGCAATCGACACTTATATGCATAGTCGCCTTACA**GTCATCTACCTAAGT-CAA**-3', the long hairpin 5'-**TAGGCGTCATTTACTCACACATTTTATGTCGATACTCT** TGCTCTGCT-TATAGCAATTTCAAGCGACTTAAATGCATATGGTGTTGT**AGAGTATCGACATAAAATGTGT-TAGTAAATGACGCCTA**-3' and the crossbreed 5'-GATCTAGTACCATCCATTTGCTCTGCTTATAG

CAATTTCAAGCGACTTAAATGCATATGGTGTTGTAGAGTATCGACATAAAATGTGTGAGTAAAT-GACGCCTA-3'.

The three primers 5'-TTGACTTAGGTAGATGAC-3', 5'-GATCTAGTACCATCCATT-3' and 5'-TAGGCGTCATTTACTCACAC-3' were labeled with Cy5 at their 5'-end in order to measure species concentrations using PAGE-gels (12.5% PAA gels, 19:1, 50% urea, Roth). More information on sequences can be found in the supplemental material S1. After PCR, samples were mixed with a 2x loading buffer (95% formamide, 5% glycerol and 50 μM EDTA) as preparation for denaturing PAGE-analysis. The errors of the given fitness and efficiency values are obtained from assuming a pipetting error of 10%. The crossbreed experiment in Figure 3 was performed three times; the error bars are standard deviations.

Results. We use DNA species, which replicate with PCR under temperature oscillations, to prevent product inhibition and allow for exponential growth [13,23]. This approach allows one to implement all three replication modes (Fig. 1). We observe linear growth if an oligonucleotide has only one initiator site and no secondary structure (Fig. 1, linear). The template concentration remains constant, whereas in every PCR cycle new products are created that cannot be replicated due to the lack of an initiator site. In contrast, A simple exponential replicator has a template with a hairpin structure (Fig. 1, hairpin). This allows the product strand to form an initiator site and therefore template itself. A more complex exponential replicator possesses a template without secondary structures - as in the case of linear growth - and is supplied with two primers for replication initiation (Fig. 1, crossbreed).

For the linear replicating species, a 5 nM template produces 0.82 ± 0.14 nM of copies in every PCR cycle. Thus, the concentration c increases linearly with time t: $c(t) = 5nM + 0.82nM \cdot t / \tau$, where $\tau = 72s$ is the length of a cycle. For exponential replicators we assume that $c(t) = c_{t=0} \cdot e^{\varepsilon \cdot t / \tau}$ with the efficiency $\varepsilon = r \cdot \tau$. The fitted efficiencies for the slow and the fast exponential replicator in Figure 1 are $\varepsilon = 0.06 \pm 0.01$ and $\varepsilon = 0.25^{+0.02}_{-0.01}$ respectively. This difference can be explained with the hairpin structure of the slow replicator, which causes inhibition of the initiator site. We find that the efficiencies also depend on the details of the initiator sites and the templates (supplemental material S1).

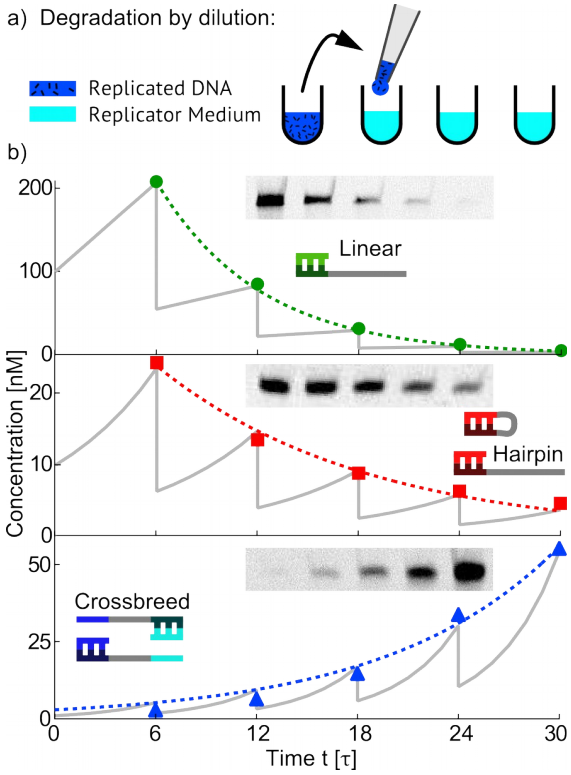


Figure 2. Modes of replication to survive degradation. *a) Molecular degradation is implemented in a controlled way by serial dilutions. Parts of the replicated DNA are transferred to fresh replication resources (polymerase, primers). b) Degradation leads to selection for fast replicators. For linear replicators (●), extinction is only a matter of time (supplemental material S8). Exponential replicators (■, ▲) survive if the replication rates exceed the dilution rate. This is only achieved for the replicator with two initiator sites. Data points are from PAGE-analysis, dashed curves are exponential fits, and gray theoretical curves show the fitted dynamics between serial dilutions (supplemental material S6).*

However, molecules do not only form in nature, but they also degrade. Typically, this happens in an exponential process [24]. Degradation is experimentally implemented by making serial dilution transfers of PCR products to fresh PCR medium (Fig. 2a), which leads to selection for fast replicators. Here, after n=6 PCR cycles, we use a dilution factor of $d = e^{\gamma \cdot n \cdot \tau} = 5$ i.e. one part of PCR product is added to four parts of fresh medium. In order to prevent extinction, the replication rate has to exceed the dilution rate. For exponential replication with the time $t' = n \cdot \tau$ between two dilution steps m and m+1, species go extinct if the fitness $F = \frac{c(t_{m+1})}{c(t_m)} = \frac{e^{\varepsilon \cdot t'/\tau}}{d} = e^{\varepsilon \cdot n - \ln(d)} < 1$.

Figure 2 shows how the dynamics behave under the selection pressure of degradation: Linear replicators with one primer will always go extinct (supplemental material S8). Templates supplied with two initiator sites reach a fitness $F = 1.81 \pm 0.08 > 1$, whereas a hairpin with one initiator site (red, $F = 0.62^{+0.05}_{-0.07}$) goes extinct due to its self-inhibition, despite being able to replicate exponentially. Interestingly, if two independently replicating hairpins (long hairpin 124 bp, short hairpin 84 bp with different initiator site sequences) at less challenging dilutions have evolved to use the same loop sequence information, it is possible for them to create crossbreeds - fast exponential replicators with two initiator sites (Fig. 3a).

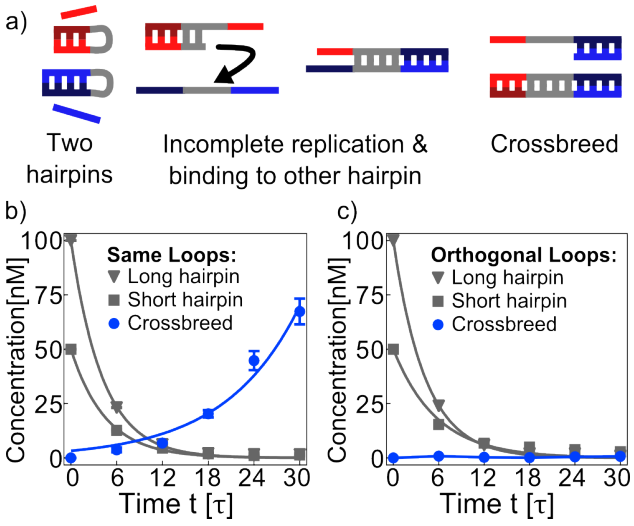


Figure 3. Crossbreeding under serial dilution. *a) The incomplete replication of one hairpin leads to the emergence of crossbreeds if two replicating hairpins have a matching loop sequence. An incompletely elongated primer can bind to the other template for another elongation step. b) Two hairpins with same loop sequences go extinct under dilution conditions. However, since the loop sequences are identical, the hairpins can form crossbreeds before they die out. The crossbreeds slowly emerge and takes over the entire population. c) Hairpins with orthogonal loop sequences vanish after few cycles and cannot form crossbreeds due to the lack of complementary loop sequences.*

This cooperative crossbreeding process is a consequence of non-processive replication, triggered for example by a premature temperature increase before the replication process is finished. An incomplete strand emerges, containing only a part of the complement of the loop sequence. At lower temperatures, the incomplete strand can bind to the loop of another hairpin (Fig. 3a). Subsequent elongation creates the first crossbreed strand with both initiator sites. The crossbreed sequences are the loop sequences, flanked by the stem sequences of the long and the small hairpins (supplemental material S2*)*. As a result, the crossbreeds have an intermediate length of 104 bases. This process is generic and can also be observed with other replicases such as Deep Vent (exo-) DNA Polymerase (supplemental material S3).

Upon mixing hairpins with the same loop sequence, the crossbreeds slowly emerge. They grow continuously and with their superior fitness take over the entire population (Fig. 3b). An exponential fit finds a fitness value of F=1.85 (with an extrapolated starting concentration $c_{t=0} = 3.3^{+0.7}_{-0.6}$ nM). The hairpin fitnesses are smaller than one ($F = 0.27 \pm 0.05$ for the small hairpin and $F = 0.24 \pm 0.04$ for the long hairpin). Only the crossbreeds are able to survive, but they preserve the genetic information of the extincted hairpins. The loop sequence is now flanked by the stem parts of the two different hairpins (supplemental material S2). As a control, a pair of hairpins was chosen that do not form stable complexes with each other. Mixing such orthogonal hairpins in replication-dilution experiments should not change their respective fitness.

Our results show that the fitness of the longer hairpins is not affected by the presence of a smaller species ($F = 0.25^{+0.05}_{-0.04}$ when the hairpin grows alone, $F = 0.24^{+0.05}_{-0.04}$ in a mix). However, the smaller hairpin has a reduced fitness of $F = 0.35^{+0.07}_{-0.06}$ in the mixture with the long hairpin (compared to $F = 0.62^{+0.05}_{-0.07}$, supplemental material S9). Without crossbreed formation, no species with a fitness F>1 emerges and all the sequence information is lost due to degradation (Fig. 3c).

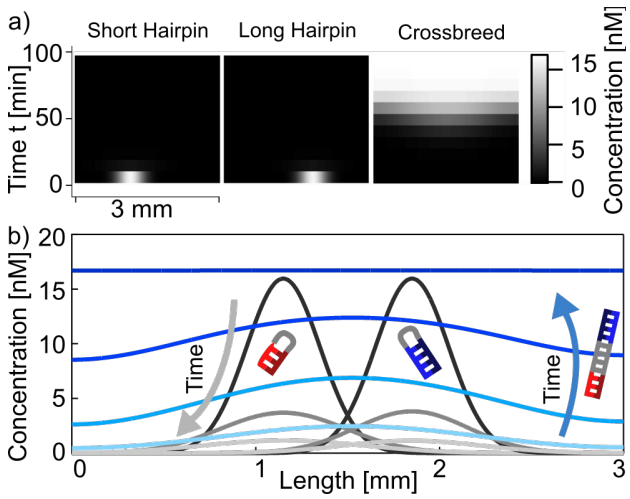


Figure 4. Spatial emergence of crossbreeds under diffusion. a) *Time evolution of hairpins and crossbreeds in a one-dimensional simulation. Although the hairpins vanish fast, the concentration overlap is sufficient to create the crossbreeds. Because the latter replicate faster than they are degraded, they take over the entire population until they reach their saturation limit. b) A different representation of the same simulation with hairpin concentrations after 0, 5 and 10 minutes (gray) and crossbreeds after 40, 50, 60 and 100 minutes (blue).*

To explore the influences of diffusion and the spatial distribution of our molecular model system, we perform a finite-element simulation of the measured system (Fig. 4 and supplemental material S7). In the beginning, two Gaussian distributed, related hairpin populations are present, which are challenged by diffusion and degradation. In a small overlapping region, the hairpin concentration is large enough to create the crossbreed. The crossbreed further replicates until a homogeneous distribution is reached in the end, whereas the hairpin species vanish exponentially.

Notably, if crossbreeds mutate the sequence of their initiator site, they can go back in evolution and form hairpins. Such a reverse transformation could be useful in environments where single stranded nucleic acids and thus hairpin structures are favored by chemical degradation. Thus, the observed process can switch between two states and form a risk spreading mechanism [18,25,26]. The overall genetic information is preserved, but the rearrangement allows for different phenotypes and the fast adaption to changing environments.

Outlook. Evolution is based on passing on the information of individual molecules. To prevent extinction, Darwinian evolution creates non-trivial cooperative strategies, such as altruistic behavior [27]. Individuals even sacrifice themselves in order to prevent the loss of their species' genetic information [28,29]. Our results show that the mechanism of kin selection can already be found for molecular species. Degradation enforces an evolution of replicators from initially linear replicators to slow, but exponentially growing hairpin structures. The latter can cooperate by binding incomplete replicates, forming a fast, exponentially growing crossbreed that maintains the original sequence information in a rearranged manner. This suggests that even in early evolution, cooperation and kin selection likely played a central role.

References.

1. Gore, J., Youk, H. & van Oudenaarden, A. Snowdrift game dynamics and facultative cheating in yeast. *Nature* **459,** 253–256 (2009).
2. Elena, S. F. & Lenski, R. E. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* **4,** 457–469 (2003).
3. Kerr, B., Riley, M. A., Feldman, M. W. & Bohannan, B. J. M. Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* **418,** 171–174 (2002).
4. Weber, M. F., Poxleitner, G., Hebisch, E., Frey, E. & Opitz, M. Chemical warfare and survival strategies in bacterial range expansions. *J. R. Soc. Interface* **11,** 20140172 (2014).
5. Chuang, J. S., Rivoire, O., Leibler, S., Chuang John S., O. R. & Leibler, S. Simpson's paradox in a synthetic microbial system. *Science* **323,** 272–5 (2009).
6. Fujii, T. & Rondelez, Y. Predator-prey molecular ecosystems. *ACS Nano* **7,** 27–34 (2013).
7. Padirac, A., Fujii, T., Estévez-Torres, A. & Rondelez, Y. Spatial waves in synthetic biochemical networks. *J. Am. Chem. Soc.* **135,** 14586–92 (2013).
8. Lincoln, T. a & Joyce, G. F. Self-sustained replication of an RNA enzyme. *Science* **323,** 1229–1232 (2009).
9. Vaidya, N. *et al.* Spontaneous network formation among cooperative RNA replicators. *Nature* **491,** 72–77 (2012).
10. Chen, I. a, Roberts, R. W. & Szostak, J. W. The emergence of competition between model protocells. *Science* **305,** 1474–1476 (2004).
11. Krammer, H., Möller, F. M. & Braun, D. Thermal, autonomous replicator made from transfer RNA. *Phys. Rev. Lett.* **108,** 1–5 (2012).
12. Kreysing, M., Keil, L., Lanzmich, S. & Braun, D. Heat flux across an open pore enables the continuous replication and selection of oligonucleotides towards increasing length. *Nat. Chem.* **7,** 203–208 (2015).
13. von Kiedrowski, G., Wlotzka, B., Helbing, J., Matzen, M. & Jordan, S. Parabolic Growth of a Self-Replicating Hexadeoxynucleotide Bearing a 3'-5'-Phosphoamidate Linkage. *Angew. Chemie* **30,** 423–426 (1991).

14. Traulsen, A., Claussen, J. C. & Hauert, C. Coevolutionary dynamics: From finite to infinite populations. *Phys. Rev. Lett.* **95,** 1–4 (2005).
15. Orgel, L. E. The origin of life-a review of facts and speculations. *Trends Biochem. Sci.* **23,** 491–495 (1998).
16. Robertson, M. P. & Joyce, G. F. The origins of the RNA world. *Cold Spring Harb. Perspect. Biol.* **4,** cshperspect.a003608– (2012).
17. England, J. L. Statistical physics of self-replication. *J. Chem. Phys.* **139,** 121923 (2013).
18. Obermayer, B., Krammer, H., Braun, D. & Gerland, U. Emergence of information transmission in a prebiotic RNA reactor. *Phys. Rev. Lett.* **107,** 018101 (2011).
19. Wochner, A., Attwater, J., Coulson, A. & Holliger, P. Ribozyme-catalyzed transcription of an active ribozyme. *Science* **332,** 209–212 (2011).
20. Gómez-Gardeñes, J., Campillo, M., Floría, L. M. & Moreno, Y. Dynamical organization of cooperation in complex topologies. *Phys. Rev. Lett.* **98,** 1–4 (2007).
21. Ivica, N. a. *et al.* The paradox of dual roles in the rna world: Resolving the conflict between stable folding and templating ability. *J. Mol. Evol.* **77,** 55–63 (2013).
22. Mast, C. B. & Braun, D. Thermal Trap for DNA Replication. *Phys. Rev. Lett.* **104,** 188102 (2010).
23. Szathmáry, E. & Maynard Smith, J. From replicators to reproducers: the first major transitions leading to life. *J. Theor. Biol.* **187,** 555–71 (1997).
24. Scheuring, I. & Szathmáry, E. Survival of replicators with parabolic growth tendency and exponential decay. *J. Theor. Biol.* **212,** 99–105 (2001).
25. Reiter, M., Rulands, S. & Frey, E. Range expansion of heterogeneous populations. *Phys. Rev. Lett.* **112,** 148103 (2014).
26. Rulands, S., Jahn, D. & Frey, E. Specialization and Bet Hedging in Heterogeneous Populations. *Phys. Rev. Lett.* **113,** 108102 (2014).
27. Nowak, M. a. Five rules for the evolution of cooperation. *Science* **314,** 1560–1563 (2006).
28. Hamilton, W. D. The genetical evolution of social behaviour. I. *J. Theor. Biol.* **7,** 1–16 (1964).
29. Hudson, R. E., Aukema, J. E., Rispe, C. & Roze, D. Altruism, cheating, and anticheater adaptations in cellular slime molds. *Am. Nat.* **160,** 31–43 (2002).

Acknowledgments.

We thank Jonathan Liu, Matthias Morasch and Friederike Möller for corrections on the manuscript. Financial support from the SFB1032 and the Simons Collaboration on the Origins of Life is gratefully acknowledged.

Author contributions. All authors contributed extensively to the work presented in this paper.

Supporting Online Information

Stepwise cooperation of molecular replicators

Georg C. Urtel and Dieter Braun

S1 Sequences

Four strand species were ordered: A short and a long hairpin, which can create two crossbreed species. One of them was ordered as well. As control a short control hairpin was used, which does not react with the long hairpin. Also the complementary sequences were ordered to start the experiments with double stranded species. The long hairpin has a 38-bp-stem, whereas the stem length of the short hairpins is 18 bp. All hairpins have the same loop length of 48 bases.

Name	Length	Sequence from 5' to 3'	Primer	ε
short hairpin	84	**GATCTAGTACCATCCATT**TGCTCTGCTTATAG CAATTTCAAGCGACTTAAATGCATATGGTGTT GT**AATGGATGGTACTAGATC**	P2	$0.064^{+0.003}_{-0.006}$
short hairpin / complement	84	**GATCTAGTACCATCCATT**ACAACACCATATGC ATTTAAGTCGCTTGAAATTGCTATAAGCAGAG CA**AATGGATGGTACTAGATC**	P2	$0.064^{+0.003}_{-0.006}$
crossbreed	104	GATCTAGTACCATCCATTTGCTCTGCTTATAG CAATTTCAAGCGACTTAAATGCATATGGTGTT GTAGAGTATCGACATAAAATGTGTGAGTAAAT GACGCCTA	P3	$0.254^{+0.001}_{-0.006}$
crossbreed / complement	104	TAGGCGTCATTTACTCACACATTTTATGTCGA TACTCTACAACACCATATGCATTTAAGTCGCT TGAAATTGCTATAAGCAGAGCAAATGGATGGT ACTAGATC	P2	$0.254^{+0.001}_{-0.006}$
long hairpin	124	**TAGGCGTCATTTACTCACACATTTTATGTCGA TACTCT**TGCTCTGCTTATAGCAATTTCAAGCG ACTTAAATGCATATGGTGTTGT**AGAGTATCGA CATAAAAT**GTGTGAGTAAATGACGCCTA	P3	0.047 ± 0.010
long hairpin / complement	124	**TAGGCGTCATTTACTCACACATTTTATGTCGA TACTCT**ACAACACCATATGCATTTAAGTCGCT TGAAATTGCTATAAGCAGAGCA**AGAGTATCGA CATAAAAT**GTGTGAGTAAATGACGCCTA	P3	0.047 ± 0.010
short control hairpin	84	**TTGACTTAGGTAGATGAC**ATGAATAAGATCAG AGCAATCGACACTTATATGCATAGTCGCCTTA CA**GTCATCTACCTAAGTCAA**	P1	0.103 ± 0.024
short control hairpin / complement	84	**TTGACTTAGGTAGATGAC**TGTAAGGCGACTAT GCATATAAGTGTCGATTGCTCTGATCTTATTC AT**GTCATCTACCTAAGTCAA**	P1	0.103 ± 0.024

***Table S1**: All used DNA templates, their lengths, sequences, exponential efficiencies and primers. Stem parts of hairpins are bold, replication initiator sites are underlined.*

Primers have a length of 18 or 20 bases and are labeled with Cy5 at their 5'-end, which allows to measure concentrations of PCR products in gels. The efficiencies in Table S1 suggest, that the primers and the stem length play an important role in replication.

Name	**Length**	**Sequence from 5' to 3'**	**Target**
P1	18	TTGACTTAGGTAGATGAC	short control hairpin
P2	18	GATCTAGTACCATCCATT	short hairpin, crossbreed
P3	20	TAGGCGTCATTTACTCACAC	long hairpin, crossbreed

Table S2: *The sequences of the three primers in use and their respective target templates.*

For the figures in the main text, the following sequences have been used:

Figure	**Curve**	**Sequences**
1	Linear	Crossbreed, P3
1	Hairpin	short hairpin (+complement), P2
1	Crossbreed	Crossbreed (+complement), P2, P3
2b	Linear	Crossbreed, P3
2b	Hairpin	Short control hairpin (+complement), P1
2b	Crossbreed	Crossbreed (+complement), P2, P3
3b		Short control hairpin (+complement), long hairpin (+complement), P1, P3
3c		Short hairpin (+complement), long hairpin (+complement), P2, P3

Table S3: *Sequence compositions for the experiments in Figure 1-3 of main article.*

S2 Conservation of information in the crossbreeds

The DNA sequences in Table S1 can be split up in several parts:

Part	Length	Sequence from 5' to 3'
A	18	GATCTAGTACCATCCATT
B	38	TAGGCGTCATTTACTCACACATTTTATGTCGATACTCT
C	18	TTGACTTAGGTAGATGAC
Loop1	48	TGCTCTGCTTATAGCAATTTCAAGCGACTTAAATGCAT ATGGTGTTGT
Loop2	48	ATGAATAAGATCAGAGCAATCGACACTTATATGCATAG TCGCCTTACA

Table S4: *All the sequences which were used in this paper can be described by a combination of the sequence parts listed in this table and their respective complementary sequences.*

With the sequence parts in Table S4 and the notation that complementary sequences of one of the parts are written with an asterisk on top, the conservation / rearrangement of the sequences during crossbreed creation can be described as follows (Fig S1): The two hairpin species needed to create the crossbreeds have the sequence pattern 5'-A-Loop1-A*-3' / 5'-A-Loop1*-A*-3' (short hairpin and its complement) and 5'-B-Loop1-B*-3' / 5'-B-Loop1*-B*-3' (long hairpin and its complement). To conserve the information, the crossbreeds have to contain the parts A, B and Loop1 as well as all the respective complements. This is the case, since the crossbreed sequences are 5'-A-Loop1-B*-3' / 5'-B-Loop1*-A*-3' as well as 5'-A-Loop1*-B*-3' / 5'-B-Loop1-A*-3'. The control hairpin has the sequence 5'-C-Loop2-C*-3' / 5'-C-Loop2*-C*-3'.

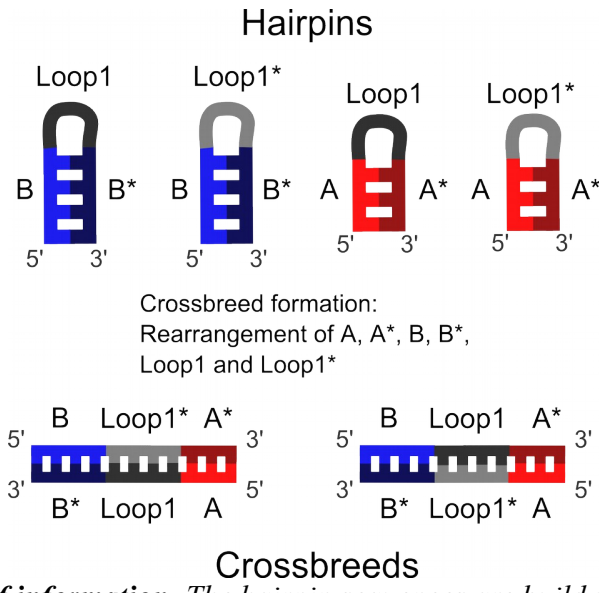


Figure S1. Conservation of information. *The hairpin sequences are build up of several parts and their complements (marked with *). In order to preserve the genetic information, the stem parts (A,A*,B,B*) as well as the loops (Loop1, Loop1*) have to be transferred to the crossbreeds. The double stranded crossbreeds indeed contain all parts, but in a different order.*

S3 A different polymerase shows that the crossbreed reaction is a generic processes

To check if the reaction that creates the crossbreeds is a generic process and not caused by a polymerase specific property, we used a different polymerase. We chose Deep Vent (exo-) DNA Polymerase in 1x ThermoPol Reaction Buffer (both purchased from NEB). The PCR was performed for 25 cycles with the short and the long hairpins and the same settings as described in the material and methods section. The only difference was a shorter initial denaturating phase of only one minute and lower primer concentrations for linear replicators to obtain more defined bands. Several positive and negative controls were performed to ensure that there are no primer dimers for any relevant primer combination and the right primer-template combinations give the correct bands (Fig. S2). The crossbreeds are only visible, if added in the beginning (lanes H and I) or the short and long hairpin are mixed (lane L). This shows that the described crossbreeding mechanism is polymerase-independent.

Negative Control	**Sequences**	**Positive Control**	**Sequences**
A	P2, P3	G	P3, long hairpin
B	P1, P3	H	P3, crossbreed
C	P3, short hairpin	I	P2, crossbreed
D	P3, short control hairpin	J	P2, short hairpin
E	P1, long hairpin	K	P1, short control hairpin
F	P2, long hairpin	L	P2, P3 short and long hairpin
		M	P1, P3 short control and long

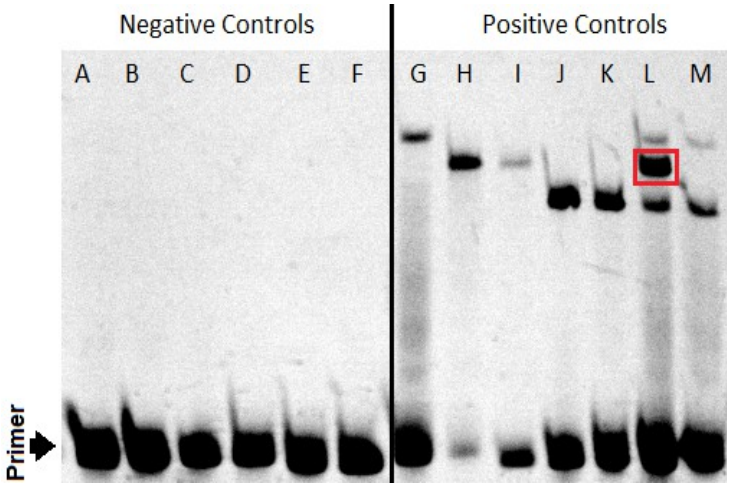


Figure S2. Crossbreeds emergence with other polymerase. *The table on the top lists the sequences which were mixed for every pocket of the gel. The crossbreeds do not only emerge when Phusion High Fidelity DNA Polymerase is used, but also with Deep-Vent (exo-)-polymerase. This becomes clear in lane L, where two hairpin species are mixed for a polymerase chain reaction (PCR), but another emerges (red box).*

S4 Gel Analysis

To differentiate the various species by their lengths and measure their concentrations, PAGE-gels (12.5% PAA, 19:1, 50% urea, Roth) where used. They were run in a TBE buffer for typically 15 minutes (in some cases longer, see below) with 400V.

To analyze the PAA gels (typical gels in Figure S2 or S3), a LabView software was used, which allows to integrate the signal from the Cy5-labeled oligonucleotides. To correct for inhomogeneous illumination, a white paper was imaged and the picture of it smoothed. The gel images where then divided by this reference image. As long as the primers stay within the gel, the total concentration of Cy5-labeled oligonucleotides is known in every lane. The concentration of single bands (i.e. single species) can be calculated with the following equation:

$$c_{band} = \frac{I_{band}}{I_{total}} \cdot c_{total} + c_{unlabeled}$$

With the oligonucleotide concentration of a single band c_{band}, the integrated intensities of the band I_{band} and the integrated intensity of the total lane I_{total}. c_{total} is the total concentration of labeled oligonucleotides, which corresponds to the initial primer concentration and $c_{unlabeled}$ corresponds to the starting concentration of templates, which have no Cy5-label.

For reactions with long hairpins, small hairpins and crossbreeds in one lane, a slightly different approach was chosen, since the separation of the strands after 15 minutes gel electrophoresis is not sufficient. After this time a first image was taken and the total concentration of labeled species (concentration of short hairpin + concentration of long hairpin + concentration of crossbreeds) was measured with a similar formula as above:

$$c_{3bands} = \frac{I_{3bands}}{I_{total}} \cdot c_{total}$$

With the total Cy5-labeled oligonucleotide concentration of the three bands c_{3bands}, the integrated intensities of the bands I_{3bands} and the integrated intensity of the total lane I_{total}. c_{total} is the total concentration of labeled oligonucleotides, which is again the initial primer concentration. Electrophoresis is further conducted for 20 minutes. This leads to a better separation of the the hairpins and the crossbreeds (Fig. S3), but primers move out of the gel. Since the total concentration of the three bands is known, the following formula can be applied:

$$c_j = \frac{I_j}{\Sigma_{j=1}^{3} I_j} \cdot c_{3bands} + c_{unlabeled,j}$$

Where c_j and I_j are the concentrations of labeled oligonucleotides and the intensities of the bands j=1,2,3. For the short and the long hairpin the unlabeled DNA $c_{unlabeled,j}$ has to be taken into account, which get reduced with every dilution step.

The only values which need to be measured are the spatially integrated intensities. The LabView program searches for local minima to the left and the right of the lanes and sets this points to zero intensity. This can lead to a slight overestimation of c_{band} which leads to a relatively large error for small concentrations.

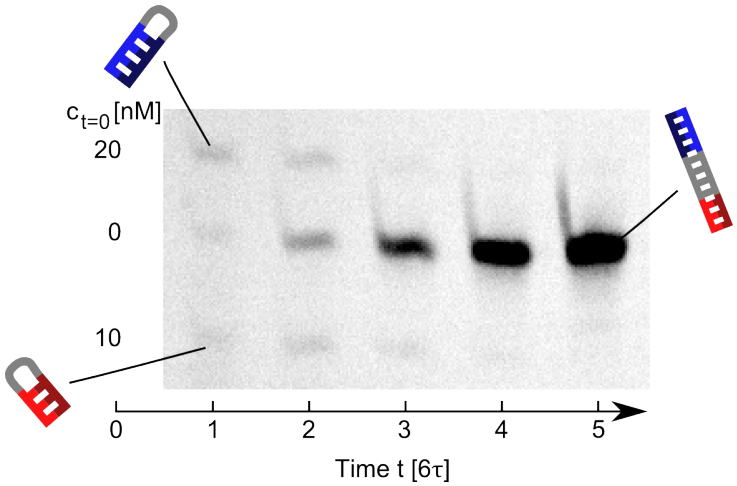


Figure S3. Emerging crossbreeds. *PAA gel showing the slow emergence of the crossbreeds and the following takeover of the entire population. Oligonucleotides are repeatedly replicated (for six PCR cycles) and then diluted (with a factor of d=5). The lanes represent the state of the system before a dilution step. In the initial mix, only the long and the short hairpins are supplied. However, before the first dilution, the crossbreeds are already slightly visible. This gel was run for 35 minutes to increase the separation width of the species within a lane.*

S5 Correction for pipetting error

We needed to change the dilution factor from d=5 to d<5 for the simple model in Figure 2b. This suggests that for the experiments with serial dilution transfers the small transferred volumes (1 µl or 2 µl) might lead to imprecise dilutions. To measure this, a 5 µl PCR with 1 nM crossbreed and 250 nM of each of its primers was prepared. The mix was cycled 30 times with the conditions used for all other experiments. 1 µl of the PCR product was transferred to a mix containing the same primer concentration as the initial mix. This mixture was stirred carefully, then 1 µl was transferred to a primer mix again. Three dilution steps were made and then everything was run on a PAA gel. The results show that the predicted concentrations deviate from the measured concentrations. Averaging over the first three dilution leads to a dilution factor of 3.85 instead of 5 (Fig. S4). A possible explanation is that for each dilution transfer, an additional portion (~0.4 µl) is transferred on the outside of the pipette tip. As a consequence, the dilution factor was not used as a fixed parameter in Figure 2b.

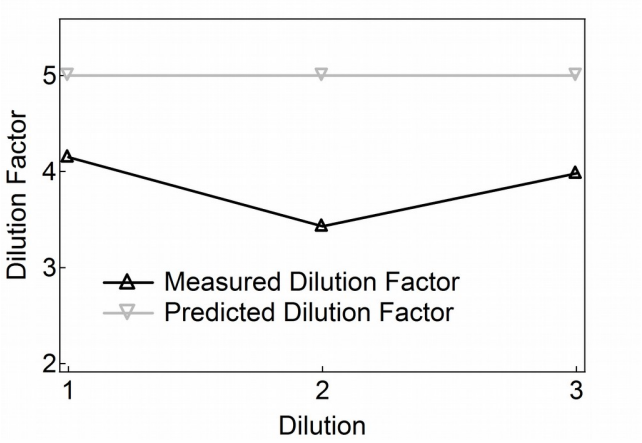


Figure S4. Measurement of dilution factor. *The measured dilution factor (black symbols) deviated from the desired dilution factor (gray symbols).*

S6 The simple replication model

For the different types of growth a simple Matlab code was written, which also takes the dilutions into account. For every species an array is created which has the concentrations for every cycle as entries. For every growth type (in the beginning without dilution), there is a different update rule for the concentration c :

Linear replication: $c_{j+1} = c_j + c_{template} \cdot \varepsilon_{linear}$

With the concentration after cycle j given by c_j , the template concentration $c_{template}$ and the growth rate ε_{linear} which is the fraction of template copied per cycle.

Exponential replication: $c_{j+1} = c_j \cdot (1+\varepsilon)$

With the efficiency $0 < \varepsilon < 1$, where $\varepsilon = 1$ corresponds to doubling of the concentration in every cycle. After several replication cycles, the concentration is divided by the effective dilution factor d^{eff} .

The parameters used for the curves in Figure 2 are shown in the following table:

Figure	**Replicator type**	**Templates and Primers**	$\varepsilon_{(linear)}$	d^{eff}	**Result**
2b	Linear	Crossbreed (+Complement), Primer P2	0.180	4.0	Extinction
2c	Exponential (one initiator site)	Short Control Hairpin (+Complement), Primer P1	0.140	3.8	Extinction $F = \frac{e^{\varepsilon \cdot t'/\tau}}{d} = 0.61$
2d	Exponential (two initiator sites)	Crossbreed (+Complement), Primer P2 & P3	0.275	2.9	Survival $F = \frac{e^{\varepsilon \cdot t'/\tau}}{d} = 1.80$

Table S5: *Different types of replicators and the used parameters for the simple model.*

In the case of exponential replication with two initiator sites (crossbreed), the dilution factor deviates strongly from the experiments with the other types of growth. In a control experiment, 1 nM of the crossbreed was placed with primers and PCR-Mastermix in the thermocycler, which was preheated to 95°C. After 10 seconds, the sample was cooled down to 1°C and placed on ice. The thermocycler was preheated again to 95°C and the sample inserted. After 10 of these cycles, an efficiency of $\varepsilon = 0.3$ was measured, suggesting a replication during initial heating and final cooling. If the crossbreed concentration is increased by 20% during heating, leading to a slightly higher initial concentration c^{eff} , the difference between d=4 and d=3 seems reasonable: $\frac{c^{eff}}{d} = \frac{c}{d^{eff}}$, where d^{eff} is an effective dilution factor. Assuming that c^{eff} differs from c by 30%: $c^{eff} = c \cdot 1.3$. Thus, $\frac{c \cdot 1.3}{d} = \frac{c}{d^{eff}}$. Assuming that the actual dilution factor is $d = 3.85$ (see S5), we get $d^{eff} = 3.0$, which is much closer to the value in Table S5. However, it is not clear if the heating or cooling has a higher influence on the measured efficiency. Also, for experiments with an annealing step at 67°C replication dynamics might differ.

S7 Finite element simulation for the modeling with a spatial dimension

To explore the replicator dynamics in one spatial dimensions, a COMSOL model was created which allows the species to diffuse. In contrast to the conducted experiments and the simple model, replication and molecular degradation were assumed to be permanent processes instead of needing heating/cooling cycles or stepwise dilutions. Testing the model with homogenous starting concentrations of long and short hairpins should show comparable results to the experiments in Figure 3c of the main article. Since in our experiments concentrations are measured before the dilution steps, but degradation is a permanent process in the model, we corrected for the dilution factor in the comparison (Fig. S5).

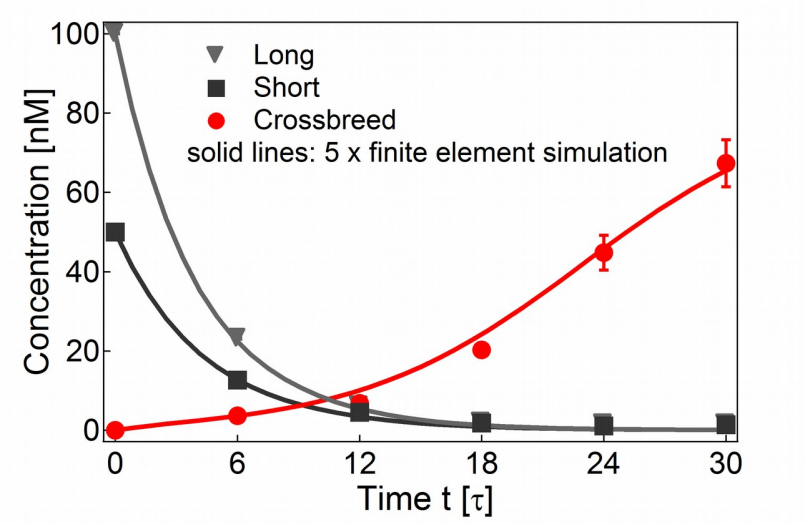


Figure S5: *Comparison of the COMSOL model with homogenous initial distribution of the long and the short hairpin (solid line) to the data (symbols). Since the measurements are made before the dilution, but degradation is a continuous process in the model, the COMSOL values are multiplied with the dilution factor of 5. The time scale is the same as in the experiments with* $\tau = 72\,s$.

For the model, growth was assumed to fulfill a logistic equation [S1–S3]. For the short and long hairpins (SH and LH), the equation is $\frac{dc_i}{dt} = D_i \cdot \frac{d^2 c_i}{dx^2} \cdot -\gamma \cdot c_i + r_i \cdot c_i \cdot \left(1 - \frac{c_i}{K}\right)$ with i={SH, LH}, the concentration c_i , the diffusion constant D_i , the carrying capacity K, the growth rate r_i and the decay rate γ . The crossbreed (CB) has an additional term to allow its de novo creation from the short and the long hairpin: $\frac{dc_{CB}}{dt} = D_{CB} \cdot \frac{d^2 c_{CB}}{dx^2} - \gamma \cdot c_{CB} + r_{CB} \cdot c_{CB} \cdot \left(1 - \frac{c_{CB}}{K}\right) + \delta \cdot c_{LH} \cdot c_{SH}$. The parameters were chosen as follows:

- For the diffusion constant the values were calculated with a formula for single stranded DNA at 75°C according to [S4]: $D = \frac{6.69 \cdot 10^{-19}}{8 \cdot 10^{-10} + (\#bases) \cdot 5.9 \cdot 10^{-11}} \left[\frac{m^2}{s}\right]$.
- The degradation rate was adjusted, such that after $t = 6 \cdot \tau$ a population without growth shrinks to a fifth of its initial concentration $\gamma = \frac{-\ln(1/5)}{6 \cdot \tau} = 3.73\text{E-}3 s^{-1}$ (this corresponds to experiments with dilutions d=5 after every sixth cycle).
- To measure the carrying capacity of the crossbreed, we made an experiment similar to the one in Figure 1, but with a higher initial concentration and fitted a logistic curve. It is important to note, that this carrying capacity is for growth without dilution/degradation. The obtained carrying capacity is 207.1 nM. The ultimate limitation for PCR is the primer concentration (here 500 nM), but around 200 nM the template and primer concentrations are already comparable and product inhibition will slow down replication. The measured carrying capacity is from an experiment without dilution and was divided by the dilution factor 5 for the COMSOL model.

- Growth rates are chosen closely relate to measured efficiencies for the hairpins. For the crossbreed, we fitted a logistic function to the data in Figure 3c. The fit gives a rather large efficiency $r_{CB}^{fit} = 6.53\text{E-}3s^{-1}$ (compared to values from the simple model, which when normalized to $\tau = 72s$ is $r_{CB} = 3.82\text{E-}3s^{-1}$), which is counterbalanced by a carrying capacity of 85 nM for higher concentrations.
- δ was varied to fit the data.

The steady state concentration of the crossbreed in the COMSOL model is 16.7 nM. Multiplied by the dilution factor of d=5, the result of 83.5 nM is close to the fitted value. All used parameters can be found in the following table:

Parameter	**Value**
D_{LH}	8.24E-11 m²/s
D_{SH}	1.16E-10 m²/s
D_{CB}	9.60E-11 m²/s
K	41.4 nM
γ	3.73E-3 1/s
r_{LH}	4.22E-4 1/s
r_{SH}	6.95E-4 1/s
r_{CB}	6.25E-3 1/s
δ	11.3 m³/(mol*s)

Table S6: *Used parameters for the finite element model.*

S8 Linear growth

For linear replication, where a template is used to create copies which are not used as template (Fig. 1 & Fig. 2b, green curves), the species will always go extinct if oligonucleotides are degraded. Linear replication can be described with the following differential equation:

$$\frac{dc}{dt} = \frac{dc_P}{dt} + \frac{dc_T}{dt}$$

where $c(t)$ is the total species concentration, including the template concentration and the product of the replication c_P. With the degradation rate $\gamma > 0$ and the growth rate $r > 0$ we can write

$$\frac{dc_P}{dt} = -\gamma \cdot c_P + r \cdot c_T \text{ and } \frac{dc_T}{dt} = -\gamma \cdot c_T .$$

Thus, for $\gamma = 0$ the increase of product concentration c_P should be proportional to the template concentration (Fig. S6). With $c_T = c_{T,t=0} \cdot e^{-\gamma \cdot t}$ the initial differential equation becomes:

$$\frac{dc}{dt} = -\gamma \cdot c_P + (r - \gamma) \cdot c_{T,t=0} \cdot e^{-\gamma \cdot t} \cdot r$$

Even if r >> γ the population will go extinct, since the degradation term dominates for large t:

$$\lim_{t\to\infty} \frac{dc}{dt} = -\gamma \cdot c_P$$

This means, that at some points all template molecules are vanished and there is no more replication, just degradation (this justifies the exponential fit for the linear replicator in Fig. 2b) and thus

$$\lim_{t\to\infty} c = 0 \,.$$

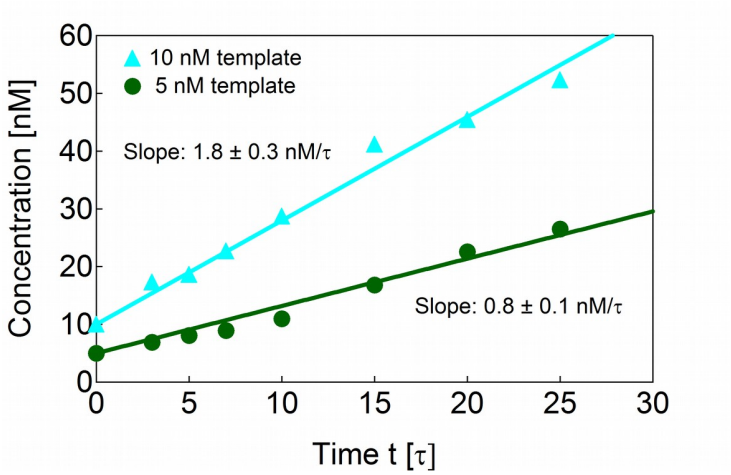


Figure S6. *Linear replication is observed if a linear DNA template is supplied with one primer. The slope of the growth is proportional to the template concentration.*

S9 Influence of parallel replication of hairpins

Mixing hairpins has an effect on the fitness of the short control hairpin, but not on the long hairpin. (Fig. S7). The fitness values of the long hairpin do not change in the presence of the short control hairpin ($F= 0.25^{+0.05}_{-0.04}$ when the long hairpin grows alone, $F=0.24^{+0.05}_{-0.04}$ in a mix). However, the short control hairpin is slowed down in the presence of the long hairpin. It has a fitness of $F= 0.62^{+0.05}_{-0.07}$ when it replicates alone, but only $F=0.35^{+0.07}_{-0.06}$ when replicating in the presence of the long hairpin. The overall result is not changed. The hairpins go extinct alone as well as in the mix.

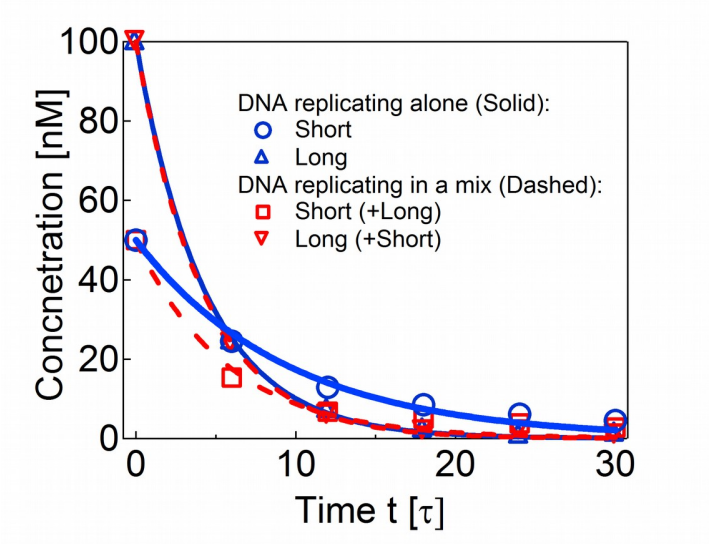


Figure S7. *The mixing of two hairpins can have an impact on their fitness.*

References for Supplementary Material

S1. Tsoularis, A & Wallace, J. Analysis of logistic growth models. *Math. Biosci.* **179,** 21–55 (2002).

S2. Swillens, S., Dessars, B. & Housni, H. El. Revisiting the sigmoidal curve fitting applied to quantitative real-time PCR data. *Anal. Biochem.* **373,** 370–376 (2008).

S3. Mast, C. B. & Braun, D. Thermal Trap for DNA Replication. *Phys. Rev. Lett.* **104,** 188102 (2010).

S4. Kreysing, M., Keil, L., Lanzmich, S. & Braun, D. Heat flux across an open pore enables the continuous replication and selection of oligonucleotides towards increasing length. *Nat. Chem.* **7,** 203–208 (2015).