Unbounded growth patterns of reproducing, competing polymers – similarities to biological evolution

– Supplementary Material –

Emanuel Gregor Worst\textsuperscript{1}, Philipp Zimmer\textsuperscript{2}, Eva Wollrab\textsuperscript{1}, Karsten Kruse\textsuperscript{2}, Albrecht Ott\textsuperscript{1}

\textsuperscript{1}Experimental Physics, Saarland University, D-66041 Saarbruecken,
\textsuperscript{2}Theoretical Physics, Saarland University, D-66041 Saarbruecken

1 Experiment

1.1 Rationale behind the DNA sequences used in the experiments

In this section we present and discuss the DNA sequences chosen for our experiments.

For the first experiment (Fig. 3A) we used 20 b DNA single strands of sequence $\Sigma = \text{GGC ATC TTA GTA CAT ATC AG}$ and its complement $\Sigma^c = \text{CTG ATA TGT ACT AAG ATG CC}$. These sequences were chosen randomly, given that the melting temperature of the resulting double strand is in the range of highest ligase activity, and single strands form no secondary structures.

The experiment presented in Fig. 3C is based on sixteen pair-wise complementary 10 b long single DNA strands. Their sequences were chosen to minimize cross-hybridization among non-complementary sequences. At the same time, ligated sequences that act as templates have to hybridize very specifically to avoid overhangs. The melting temperatures of the eight corresponding 10 bp double strands, need to be as close as possible.

To find appropriate sequences we use the following strategy: We initially generated random 10 bp single-stranded sequences and their complements. To narrow the expected melting temperatures, in the following we consider only those sequences with the same number of (C or G)s as (A or T)s. This way, we drastically reduce the computational time for the following steps. From the remaining pool of 290,000 sequence pairs, we create all possible sets, consisting of sixteen 10 b as well as eight 20 b and four 40 b pairwise complementary single-stranded sequences that result from ligating pairs of the 10 b sequences.

A suitable set of sequences for the sixteen pair-wise complementary DNA single strands following the conditions outlined above is composed of (and that we used in the experiment presented in Fig. 3C): $S_1 = \text{TGT GCC TTT G}$ and $S_1^c = \text{CAA AGG CAC A}$, $S_2 = \text{ACA CCT GAC C}$ and $S_2^c = \text{GAT CAG GTG T}$, $S_3 = \text{GGA GTT GGT T}$ and $S_3^c = \text{AAC CAA CTC C}$, $S_4 = \text{ATT TGG ACC G}$ and $S_4^c = \text{CGG TCC AAA T}$, $S_5 = \text{GAA GAC TAG C}$ and $S_5^c = \text{GCT AGT CTT C}$. $S_6 = \text{CTA CAT GAC G}$ and $S_6^c = \text{CGT CAT GTA G}$, $S_7 = \text{GGG AAG TTG T}$ and $S_7^c = \text{ACT TCC C}$, $S_8 = \text{GAA GAC TAG C}$ and $S_8^c = \text{GCT AGT CTT C}$. 

1
We chose the following pair-wise complementary 20 b sequences to result from ligation: 

\[ \begin{align*} 
S_1 & = \text{TGT GCC TTT GAC ACC TGA TC} \quad \text{and} \quad S_1^c = \text{GAT CAG GTG TCA AAG GCA CA}, \\
S_2 & = \text{GGA GTT GGT TAT TTG GAC CG} \quad \text{and} \quad S_2^c = \text{CGG TCC AAA TAA CCA ACT CC}, \\
S_3 & = \text{CTT ACA CGA GCT ACA TGA CG} \quad \text{and} \quad S_3^c = \text{CGT CAT GTA GCT CGT GTA AG}, \\
S_4 & = \text{GGG AAG TTG TGA AGA CTA GC} \quad \text{and} \quad S_4^c = \text{GCT AGT CTT CAC AAC TTC CC}. 
\end{align*} \]

The 40 b molecules have the sequences: 

\[ \begin{align*} 
S_1 & = \text{TGT GCC TTT GAC ACC TGA TCG} \\
S_2 & = \text{GAG TTG GTT ATT TGG ACC G} \\
S_3 & = \text{CTT ACA CGA GCT ACA TGA CGG GGA} \\
S_4 & = \text{AGT TGT GAA GAC TAG C} \\
S_5 & = \text{GCT AGT CTT CAC AAC TTC CCC GTC} \\
S_6 & = \text{ATG TAG CTC GTG TAA GCG GTC CAA ATA ACC AAC TCC GAT CAG GTG TCA AAG GCA CA}. 
\end{align*} \]

1.2 Quantification of DNA abundance

To quantify the amounts of DNA in a sample, we perform polyacrylamide gel electrophoresis (PAGE, Fig. S1A) with fluorescent staining (see Methods). Gel pictures are analyzed using ImageJ [54, 55]. After subtracting background fluorescence, the total area beneath a peak is proportional to the amount of double-stranded DNA in that band (Fig. S1B,C). To quantify the amounts of DNA double strands, in Figure 3, we directly use the corresponding calibration data for DNA lengths matching the calibration DNA (Fig. S1B). For other lengths, we use interpolated values.

For the experiments presented in Figure 4, we normalized all fluorescence intensities to the fluorescence intensity of the DNA band of the smallest DNA molecules of the reference sample extracted at the time point zero. From these ratios we derive the relative amount of DNA molecules in each band.

1.3 Determination of the spontaneous ligation and autocatalytic proliferation rates

In this section we show how we estimate the template-directed proliferation rate constant \( \alpha \) and the free ligation rate constant \( \beta \) from the experimental data. To determine the value of \( \beta \), we consider situations where new molecules are mainly formed by spontaneous ligation. In that case, the proliferation of molecules \( A_n \) is described by

\[ \frac{d}{dt}[A_{2n}] = \beta [A_n]^2. \]

Here, the \([A_n]\) denotes the number of double-stranded DNA molecules \( A_n \), and \( n \) corresponds to the number of base pairs of the initially present DNA double strands. Spontaneous ligation dominates in the early phases of an experiment, when predominantly the initial DNA strands are present. We obtain the value of \( \beta \) from

\[ \frac{[A_{2n}(t_1)]}{t_1} = \beta [A_n](0)^2, \]

where \([A_{2n}(t_1)]\) is the number of double-stranded DNA molecules at time \( t_1 \) and \([A_n](0)\) is the number of DNA molecules at the beginning of the experiment.
Figure S1: Quantification of DNA abundance. A) Example of a polyacrylamide gel. Lanes (from left to right) 1-4 correspond to dilutions of a home made DNA-standard, which consists of the 80 bp, 40 bp, 20 bp, and 10 bp DNA double strands (sequences given in Sec. 1.1). Numbers above the bands indicate the (known) amount of double-stranded DNA in fmol. Lane 5: DNA Ladder with bands of double-stranded DNA of lengths 10 bp, 20 bp, ... 150 bp (O’RangeRuler 10 bp DNA Ladder, ready-to-use, Thermo Fisher Scientific Inc, Waltham, MA, USA). Lanes 6-10: samples extracted from the experiment presented in Fig. 3C. The cycle numbers are indicated. B) Calibration curve obtained from the gel in (A). The line is a linear fit with a slope of $(1.017 \pm 0.045) \text{fmol}^{-1}$. C) Left to right: Profile plots of the fluorescence intensity of the 10 bp bands of lanes 1-4 (A). Numbers indicate the absolute fluorescence intensity of each band. D) Profile plots of the fluorescence intensity of the 10 bp bands of lanes 6-10 (A).
Figure S2: Abundance of the 10 bp (filled symbols) and 20 bp double strands (open symbols) as a function of the number of cycles in two experiments where free ligation occurs. This is to determine the spontaneous ligation rate constant $\beta$ for the 10 bp strands. These strands of sequences S$_1$-S$_8$ have an initial concentration of 3021 fmol (squares) or 4123 fmol (triangles). The lower 10 bp concentration leads to an abundance of 20 bp double strands of 126 fmol that can be detected after 18 cycles. At higher 10 bp concentrations the first 20 bp strands are measured after 20 cycles (270 fmol). As a result of both experiments $\beta = 8 \cdot 10^{-7} \text{cycle}^{-1} \text{fmol}^{-1}$.

where $t_1$ is the time frame until the ligated DNA double strands $A_{2n}$ can be detected for the first time.

Starting with the 20-bp DNA double strands of sequence $\Sigma$ (Sect. 1.1) at 60°C, we find $[A_{20}](0) = 4961$ fmol and $[A_{40}](t_1) = 266$ fmol with $t_1 = 8$ cycles and $\beta = 13.5 \cdot 10^{-7} \text{cycle}^{-1} \text{fmol}^{-1}$ (Fig. 3A). Experiments starting with the 10-bp double strands of sequences S$_1$-S$_8$ (Sect. 1.1) at 37°C, $[A_{10}](0) = 3021$ fmol and $[A_{20}](t_1) = 126$ fmol with $t_1 = 18$ cycles, we find $\beta = 8 \cdot 10^{-7} \text{cycle}^{-1} \text{fmol}^{-1}$. To test, whether the value of $\beta$ depends on the initial amount of DNA, we consider the experiment (Fig. S2) where $[A_{10}](0) = 4123$ fmol. $[A_{20}](t_1) = 270$ fmol with $t_1 = 20$ cycles, and we find again $\beta = 8 \cdot 10^{-7} \text{cycle}^{-1} \text{fmol}^{-1}$). Apparently, the spontaneous ligation rate constant $\beta$ is insensitive to the concentration, the temperature, and the molecular identity. In the following we use the average value of $\beta = 9.8 \cdot 10^{-7} \text{cycle}^{-1} \text{fmol}^{-1}$.

We determine the autocatalytic reproduction rate through

$$\frac{d}{dt}[A_{2n}] = \alpha[A_{2n}] - \beta \sum_{k=1}^{n}[A_{2n-k}] [A_k] \frac{\sum_{k=1}^{n}[A_{2n-k}][A_k]}{\sum_{k=1}^{n}[A_{2n-k}] [A_k]}.$$  \hspace{1cm} (3)

$\frac{d}{dt}[A_{2n}]$ is approximated by the difference of the amount $[A_{2n}]$ between two successive time points. Note that in general the value of $\alpha$ obtained in this way presents an upper estimate, because other processes than free and autocatalytic ligation could contribute to the generation of $A_{2n}$.

1.4 Further enzymatic approaches to suppress overhang-ligation

In the experiment presented in Fig. 4C, Mungbean nuclease removes overhangs and thus suppresses overhang ligation. At the same time, this enzyme also digests single-strands in a se-
Table 1: Autocatalytic proliferation rate $\alpha [A_n]$ and its ratio $\omega_n$ to the free ligation rate for the experiment presented in Fig. 3C. $[A_n]$ denotes the number of double-stranded DNA molecules $A_n$, and $n$ corresponds to the number their base pairs. $t_1$ and $t_2$ define the time interval (cycles) in which the autocatalytic redproliferation of a $[A_n]$ is analyzed. The average value for $\omega \approx 510$.

| $t_1 - t_2$ | $\alpha [A_{20}]$ $[f mol^{-1} \cdot cycle^{-1}]$ | $\omega_{20}$ |
|-------------|---------------------------------|-------------|
| 0-6         | $10.14 \cdot 10^{-6}$             | 10.35       |
| 6-12        | $8.75 \cdot 10^{-6}$              | 8.93        |
| $t_1 - t_2$ | $\alpha [A_{40}]$ $[f mol^{-1} \cdot cycle^{-1}]$ | $\omega_{40}$ |
| 25          | $201.00 \cdot 10^{-6}$            | 205.11      |
| 26          | $96.05 \cdot 10^{-6}$             | 98.01       |
| $t_1 - t_2$ | $\alpha [A_{80}]$ $[f mol^{-1} \cdot cycle^{-1}]$ | $\omega_{80}$ |
| 28          | $596.00 \cdot 10^{-6}$            | 608.16      |
| 29          | $437.15 \cdot 10^{-6}$            | 446.07      |
| 30          | $1578.65 \cdot 10^{-6}$           | 1610.87     |
| 31          | $1080.69 \cdot 10^{-6}$           | 1102.74     |

Digestion of freely ligated templates effectively inhibits the template-directed reproduction of DNA strands. As a consequence, only a small population of 40 bp double strands emerges, and longer strands are no more detectable (Fig. 4C).

We check whether a reduced activity of Mungbean nuclease leads to the emergence of populations of longer than 40 bp strands. To this end, Mungbean nuclease is added for the first time at a later time point, namely, after six cycles. Before the initial addition, the samples are incubated as in the experiment presented in Fig. 3A followed by a single incubation step at 30°C for 5 min.

Again, from the seventh cycle on, we use modified cycles: 15 min at 30°C, 10 min at 60°C, 1 min at 66°C, 10 min at 60°C, 1 min at 84°C followed by 5 min at 30°C. Furthermore, now only half of the concentration (0.5 units) of Mungbean nuclease is added to all samples at the beginning of each cycle. The concentration of Taq DNA ligase is 8.0 units/µl and the concentration of each of the two complementary 20 b single-stranded DNA of sequences $\Sigma$ and $\Sigma^c$ is 500 nM.

Fig. S3 shows that the reduction of the enzymatic activity of the nuclease leads to more intense 40 bp bands, but also to the appearance of weak 60 bp bands that result from overhangligation. Beyond the 60 bp double strands no longer DNA bands are detectable. This result supports our idea that unbounded elongation results from overhang ligation.

The two restriction enzymes AfeI and BstUI in the experiment presented in Fig. 4D reduce overhang formation since they digest the inhomogeneous DNA double strands of sequences $\sigma\sigma^c$ and $\sigma^c\sigma$ that primarily form overhangs. However, the homogeneous DNA double strands of sequences $\sigma\sigma$ and $\sigma^c\sigma^c$ are not digested. These strands also have a low probability for
Table 2: Autocatalytic proliferation rate $\alpha[A_n]$ and its ratio $\omega_n$ to the free ligation rate for the experiment presented in Fig. 3A. $[A_n]$ denotes the number of double-stranded DNA molecules $A_n$, and $n$ corresponds to the number their base pairs. $t_1$ and $t_2$ define the time interval (cycles) in which the autocatalytic redproliferation of a $[A_n]$ is analyzed. The average value for $\omega \approx 40$.

| $t_1 - t_2$ | $\alpha \cdot [A_{40}]$ [fmol$^{-1}$·cycle$^{-1}$] | $\omega_{40}$ | $t_1 - t_2$ | $\alpha \cdot [A_{60}]$ [fmol$^{-1}$·cycle$^{-1}$] | $\omega_{60}$ |
|------------|-------------------------|---------|------------|-------------------------|---------|
| 6-8        | $4.97 \cdot 10^{-6}$     | 5.08    | 14         | $20.01 \cdot 10^{-6}$   | 20.42   |
| 8-10       | $2.35 \cdot 10^{-6}$     | 2.4     | 15         | $10.16 \cdot 10^{-6}$   | 10.37   |
| 10-12      | $0.5 \cdot 10^{-6}$      | 0.51    | 16         | $41.21 \cdot 10^{-6}$   | 42.05   |
| 13         | $10.47 \cdot 10^{-6}$    | 10.68   | 17         | $37.18 \cdot 10^{-6}$   | 37.93   |
| 14         | $3.74 \cdot 10^{-6}$     | 3.82    | 18         | $58.42 \cdot 10^{-6}$   | 59.62   |
| 15         | $-0.35 \cdot 10^{-6}$    | -0.35   | 19         | $35.36 \cdot 10^{-6}$   | 36.08   |
| 16         | $7.31 \cdot 10^{-6}$     | 7.46    |            |                         |         |
| 17         | $5.53 \cdot 10^{-6}$     | 5.65    |            |                         |         |
| 18         | $26.98 \cdot 10^{-6}$    | 27.53   |            |                         |         |
| 19         | $1.04 \cdot 10^{-6}$     | 1.06    |            |                         |         |
| $t_1 - t_2$ | $\alpha \cdot [A_{80}]$ [fmol$^{-1}$·cycle$^{-1}$] | $\omega_{80}$ | $t_1 - t_2$ | $\alpha \cdot [A_{100}]$ [fmol$^{-1}$·cycle$^{-1}$] | $\omega_{100}$ |
| 16         | $31.12 \cdot 10^{-6}$    | 31.76   | 18         | $279.57 \cdot 10^{-6}$  | 285.27  |
| 17         | $17.73 \cdot 10^{-6}$    | 18.09   | 19         | $36.82 \cdot 10^{-6}$   | 37.57   |
| 18         | $57.06 \cdot 10^{-6}$    | 58.22   | 20         | $46.74 \cdot 10^{-6}$   | 47.69   |
| 19         | $35.27 \cdot 10^{-6}$    | 35.99   | 21         | $74.68 \cdot 10^{-6}$   | -76.20  |
| 20         | $0.58 \cdot 10^{-6}$     | 0.60    | 22         | $17.57 \cdot 10^{-6}$   | 17.92   |
| $t_1 - t_2$ | $\alpha \cdot [A_{120}]$ [fmol$^{-1}$·cycle$^{-1}$] | $\omega_{120}$ | $t_1 - t_2$ | $\alpha \cdot [A_{140}]$ [fmol$^{-1}$·cycle$^{-1}$] | $\omega_{140}$ |
| 18         | $363.22 \cdot 10^{-6}$   | 370.63  | 20         | $73.65 \cdot 10^{-6}$   | 75.15   |
| 19         | $0.06 \cdot 10^{-6}$     | 0.06    | 21         | $18.66 \cdot 10^{-6}$   | 19.07   |
| 20         | $54.63 \cdot 10^{-6}$    | 55.74   | 22         | $44.06 \cdot 10^{-6}$   | 44.96   |
| $t_1 - t_2$ | $\alpha \cdot [A_{160}]$ [fmol$^{-1}$·cycle$^{-1}$] | $\omega_{160}$ | $t_1 - t_2$ | $\alpha \cdot [A_{180}]$ [fmol$^{-1}$·cycle$^{-1}$] | $\omega_{180}$ |
| 21         | $53.37 \cdot 10^{-6}$    | 54.45   | 21         | $42.01 \cdot 10^{-6}$   | 42.87   |
| 22         | $58.11 \cdot 10^{-6}$    | 59.23   | 22         | $80.77 \cdot 10^{-6}$   | 82.42   |
| 23         | $4.60 \cdot 10^{-6}$     | 4.70    | 23         | $30.51 \cdot 10^{-6}$   | 31.13   |
| $t_1 - t_2$ | $\alpha \cdot n_{200}$ [fmol$^{-1}$·cycle$^{-1}$] | $\omega_{200}$ |
| 22         | $153.83 \cdot 10^{-6}$   | 156.97  |
| 23         | $29.73 \cdot 10^{-6}$    | 30.34   |
overhang formation that leads again to unbounded strand elongation in the course of the experiment. In the following, we combine Mungbean nuclease, AfeI and BstUI, to completely suppress overhang-ligation and to test if overhang ligation ceases.

The incubation scheme and the concentration of the Taq DNA ligase agrees with the experiment presented in Fig. 4D. However, temperatures are increased a later time point, namely, after 22 cycles. Also, the concentration of the 10 bp strands is reduced to 200 nM. As in the experiment presented in Fig. 4D, 1 unit of AfeI and of BstUI are added to the samples after eight and twelve cycles. 0.5 units of both enzymes as well as 0.25 units of Mungbean nuclease are added after 15 and 18 cycles. A final addition of 0.5 units of the two restriction enzymes is performed after 22 cycles.

Fig. S4 shows a regime where 10 bp, 20 bp and 40 bp double strands can be detected, whereas the DNA abundance of the 30 bp double strands is below the detection limit. However, the resulting abundance of the 40 bp strands does not suffice to form detectable amounts of longer DNA strands.

2 Theory: Refined analysis for $\beta \ll \alpha [A_n]$}
Figure S4: Combination of enzymatic activity of Mungbean nuclease, AfeI and BstUI to suppress overhang ligation. Time evolution of the 10 bp strands of sequence $\sigma$ in presence of the three nucleases. The restriction enzymes are added at the end of cycle 8, Mungbean nuclease is present from the beginning of cycle 15 on. Inset: Electrophoresis gel. Cycle numbers of sample extraction are indicated. DNA Ladder consists of bands with DNA double strands of lengths 10 bp, 20 bp, ...150 bp (O’RangeRuler 10 bp DNA Ladder, ready-to use, Thermo Fisher Scientific Inc, Waltham, MA, USA).

2.1 Duration of the plateau-phase

In the supercritical case $\alpha > \alpha_c$, there are periods during which one molecule length dominates the system (Fig. 3D, S5). For further quantitative analysis, we define the molecule of length $A_n$ to dominate if $n[A_n]$ exceeds a threshold-value that we fix to 70% of $[A_1]$ at $t = 0$. The qualitative results do not depend on the exact threshold-value.

In periods with a dominating molecule length, the corresponding number of molecules changes only weakly with time (Fig. S5). There are two possible mechanisms leading to this plateau-phase: first, the production and loss rates of the dominating molecule length balance each other (Fig. S5A). Second, both rates are essentially zero (Fig. S5B). In the latter case the average duration of the plateau-phase is

$$\tau = \frac{1}{B} \cdot [A_n]^{-2} \quad (4)$$

and the distribution of lifetimes is exponential, $1 - e^{-t/\tau}$ (Fig. S6) indicating that the appearance of a new length-doubled molecule is a Poisson process. This occurs if autocatalytic reproduction is very large so that all shorter molecules are quickly consumed in formation of the dominating molecules. Further changes of the length distribution of the molecules is inhibited until a new length-doubled molecules forms spontaneously. In our simulations with $[A_1] = 2^{17}$ this occurs for $\alpha \geq 10^3$. The values of $\tau$ obtained from the simulation agree with the theoretical values (4): $\tau_{A_4} = 2.35 \cdot 10^{-10}$, $\tau_{A_8} = 9.37 \cdot 10^{-10}$, $\tau_{A_{16}} = 3.79 \cdot 10^{-9}$, and $\tau_{A_{32}} = 1.50 \cdot 10^{-8}$ (Fig. S6). Furthermore, with increasing $\alpha$ the periods of coexistence shorten (Fig. S5).

The dependence of $\tau$ on $[A_n]$ remains the same as in the previous case if the formation and disappearance rates of the dominating molecule length are finite (Fig. S7). Deviations can be detected only if $\alpha$ is close to $\alpha_c$. 


Figure S5: Plateau phases in the stochastic simulations. A) Plateau phases due to balanced production and loss (same data as in Fig. 3D). The preferentially occurring molecules initially grow exponentially (dashed lines). The characteristic times are $6.67(5) \cdot 10^{-10}$ for $A_2$, $2.63(7) \cdot 10^{-9}$ for $A_4$, $1.1(6) \cdot 10^{-8}$ for $A_8$, $1.0(5) \cdot 10^{-8}$ for $A_{16}$, $2.1(4) \cdot 10^{-7}$ for $A_{32}$, and $1.1(4) \cdot 10^{-6}$ for $A_{64}$. Fluctuations of the plateau duration are small compared to the duration. B) Plateau phases for $\alpha = 10^3 \gg \alpha_c$.

Figure S6: Cumulative distribution of the plateau duration for plateaus of different preferentially occurring molecules (bold lines). Dashed lines: CDF for exponential distribution, $1 - e^{-\tau}$. In the simulations, $[A_1] = 2^{18}$ and $\alpha = 10^6$. Distributions were obtained from 10,000 runs. The characteristic times are given in the text.

Figure S7: Dependence of plateau duration on molecule size $n$ (A) and on the initial concentration $[A_1]$ (B). In (A), the black line represents a power law fitted to the data for $\alpha = 10^6$ with exponent 2.06(2). Simulations were performed for $[A_1] = 2^{15}$. In (B), the value of the autocatalytic rate constant is $\alpha = 10^5$. Averages were taken over 50 runs and error bars indicate the standard deviation.
Figure S8: Appearance of other dominating molecule lengths. It may be the case that a substantial fraction of $A_n$ remains in the system during plateau phases of $A_{2n}$. Furthermore, in addition to the family of molecules with $n = 2^k$ (dashed lines) another family founded by $A_{33}$ emerges (full lines). Later, molecule $A_{97} = A_{33} + A_{64}$ is more abundant than $A_{256}$. Both molecules continue to double their lengths. Parameters are $\alpha = 100$, $[A_1] = 2^{18}$.

### 2.2 Exponential growth during plateau phase

Let us now consider the case when the production and loss rates equal each other. Let $\alpha > \alpha_c$, $\beta$ be negligible and $A_{2n}$ be the dominating molecule length. In this case $A_{4n}$ grows exponentially (Fig. S5A). To understand this exponential growth let us neglect reactions of all molecules other than $A_n$, $A_{2n}$, and $A_{4n}$. We are thus left with

$$
A_n + A_n \rightarrow A_{2n}
$$

$$
A_{2n} + A_{2n} \rightarrow A_{4n}.
$$

When $[A_{2n}]$ is stationary ($[\dot{A}_{2n}] = 0$), $[A_n]$ is decaying and $[A_{4n}]$ is growing according to

$$
\frac{d}{dt}[A_{4n}] = \alpha [A_{4n}] [A_{2n}]^2
$$

implying exponential growth of $[A_{4n}]$. Furthermore mass conservation implies $4[\dot{A}_n] + 2[\dot{A}_{2n}] + [\dot{A}_{4n}] = 0$ and consequently the decay of $A_n$ is also exponential. The characteristic timescale is $\alpha^{-1}[A_{2n}]^{-2}$. Simulations confirm these relations (Fig. S5A).

### 2.3 Effects for $\alpha > \alpha_c$

We observe that during the plateau phase of $[A_{2n}]$ not all molecules $[A_n]$ are necessarily concatenated. In this case, following an initial decay after the plateau, $A_{2n}$ can rise again because $A_{4n}$ is consumed to produce $A_{8n}$ (Fig. S8 for $n = 1$).

If $A_k$ is not completely consumed during the plateau phase of $A_l$, autocatalytic production of $A_{k+l}$ may be initiated. This can lead to co-development of other molecule families.

### 2.4 Distribution of appearance times

In Figure S9 we present the average times of appearance of molecule $A_n$ as a function of $n$. This dependence is quadratic for $\alpha > \alpha_c$. 

10
Figure S9: Appearance time $t_{\text{app}, n}$ of molecules $A_n$ for $\alpha = 0.1$ (blue) and $\alpha = 0.00001$ (red). For $\alpha = 0.1$, $t_{\text{app}, n} \sim n^{2.10(2)}$, the red line is a guide to the eye.

Figure S10: Consequences of backreactions, disintegration into monomers. Parameters as in Fig. 3D but $\delta = 1,000$. Length-doubling is virtually unaffected by the back reactions and the molecules with $n = 3 \cdot 2^k$ remain readily detectable.

2.5 Back-reactions

As we did not observe DNA breakage in the experiments, we neglect back reactions in our theoretical analysis. To give a measure on how strong the impact of back reactions is on the stability of the dynamics of the preferentially occurring molecules, we consider disintegration into monomers $A_n \rightarrow \delta nA$. For rates $\delta < 1000$, the time course is essentially the same as for the case without back reactions, $\delta = 0$ (Fig. S10).