SUPPLEMENTARY INFORMATION

Continuous nonenzymatic cross-replication of DNA strands with *in situ* activated DNA oligonucleotides

Evgeniia Edeleva, Annalena Salditt, Julian Stamp, Philipp Schwintek, Job Boekhoven, and Dieter Braun

Abstract: Continuous enzyme-free replication of oligonucleotides is central for open-ended evolution experiments that mimic the origin of life. Here, we studied a reaction system, whereby two 24mer DNA templates cross-catalyzed each other's synthesis from four 12mer DNA fragments, two of which were in situ activated with the condensing agent 1-ethyl-3-(3-dimethylamino-propyl)carbodiimide (EDC). We circumvented the problem of product inhibition by melting the stable product duplexes for their reuse as templates in the following ligation step. The system reproduced itself through ligation/melting cycles and survived exponential dilution. We quantified EDC-induced side reactions in a detailed kinetic model. The model allowed to analyze effects of various reaction rates on the system's kinetics and confirmed maximal replication under the chosen conditions. The presented system enables to study nonenzymatic open-ended evolution experiments starting from diverse sequence pools.

Table of Contents

Supplementary Figures

Figure S1. ssDNA analysis with Bioanalyzer gel electrophoresis.
(a) Full-timescale fluorescence traces of the reaction samples at 0 and 1 hours (solid lines) and calibration samples (dashed lines) from the microfluidic gel electrophoresis from Bioanalyzer (referred to as gel electrophoresis from now on). Calibration traces contained each strand at a fixed concentration of 50 nM and were required for quantification of the reactions. The normalization strand was included in each sample to correct for the fluorescence differences between different wells of the Bioanalyzer chip. A marker is used in standard Bioanalyzer applications instead of the additional normalization strand. In this work, we did not include the marker, because the marker peak overlaid with some of the fragment peaks.
(b) Zoom into the calibration traces (dashed lines) and Gaussian function fits (solid red lines). The replicates clearly separated from the templates, allowing quantification of the concentration of each replicate in the reaction. Since the peaks of the two templates overlaid, two calibration samples were included in every Bioanalyzer run. The area of each peak was quantified by fitting a Gaussian function (solid red lines).
(c) The increase in the peak area with the strand concentration showed a linear dependence for each template and product strand. Beyond 400 nM, we observed a nonlinear dependence and, therefore, kept the concentrations used in the analysis with Bioanalyzer below this value.

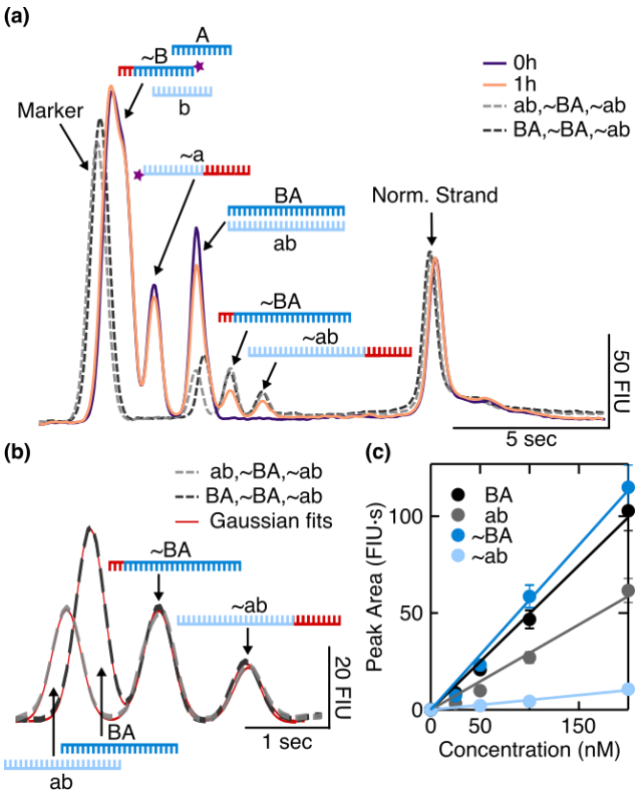


Figure S2. Control reactions in the absence of 3'-phosphate groups on the corresponding fragments and in the absence of templates.
(a) The gel electrophoresis traces in the presence (dashed lines) and absence (solid lines) of the 3'-phosphate groups. As expected, no replicates were formed in the absence of the 3'-phosphate group, since the condensation reaction between the 3'- and the 5'-ends of the corresponding fragments could not occur. Here, the area of the first peak of the 5-hour control reaction trace (solid orange line) is large because the peak of the ~a fragment without a phosphate group overlaid with the template peaks ab and BA. The area increase resulted from the hourly feeding of the ~a fragment.
(b) The gel electrophoresis traces of the reactions started in the presence (dashed lines) and absence (solid lines) of the templates. Clear product peaks were observed in the reaction with templates. In the absence of the templates, three new peaks (black stars with arrows) appeared that showed a very different peak signature from the peaks in the reaction with templates. Since the new peaks were at the different positions from the reaction products, we concluded that the untemplated reaction did not occur to detectable levels. Based on the mass spectrometry data, the new peaks could be the 5'-3' circularization products of the phosphorylated fragments or other side reactions that occurred in the absence of templates. All the replication reactions in this manuscript included template strands and did not show the peaks marked with the stars. This raises the interesting prospect that templated ligation can provide additional selection pressures under *in situ* activation conditions. The direct, untemplated ligation like polymerization is not expected at the low micromolar fragment concentrations.

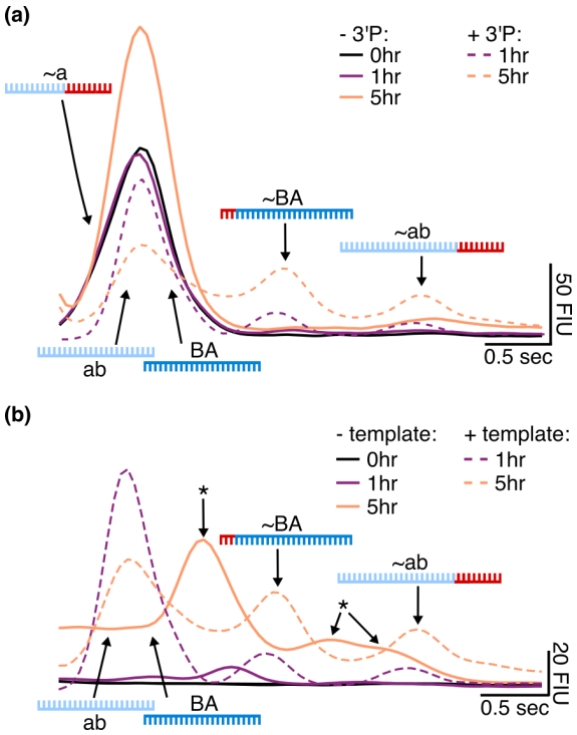


Figure S3. Fragment peaks on the gel electrophoresis traces, and the 3'-phosphate group dependent fragment side reaction.
(a) The gel electrophoresis traces of the samples where the four fragments ~B, b, A, and ~a were mixed together and incubated in the reaction buffer for 0 or 4 hours (solid lines). Running each strand separately (dashed lines) allowed the clear identification of the species peaks.
(b) The gel electrophoresis traces of the samples where the fragments A and ~a were mixed together and incubated in the reaction buffer for 0, 1, or 3 hours (solid lines). Fragment ~a was used either with (left graph) or without (right graph) a 3'-phosphate group. Over time, a shoulder on the right of the ~a fragment peak emerged only in the case when the fragment ~a had a 3'-phosphate group. Running each strand separately (dashed lines) allowed the clear identification of the species peaks.
(c) A likely pathway for the fragment side reaction is shown. The EDC-activated phosphate intermediate O-phosphorylisourea (left) could isomerize to produce N-phosphorylurea[1,2] (right) that would not be able to support the condensation reaction. Since fragments with 3'-phosphate groups modified in such a way would still be able to bind to the corresponding templates, we would expect them to exhibit an inhibiting effect on the ligation reaction. R_1=C_2H_6, R_2=CH_3–CH_2–CH_2–N–$(CH_3)_2$

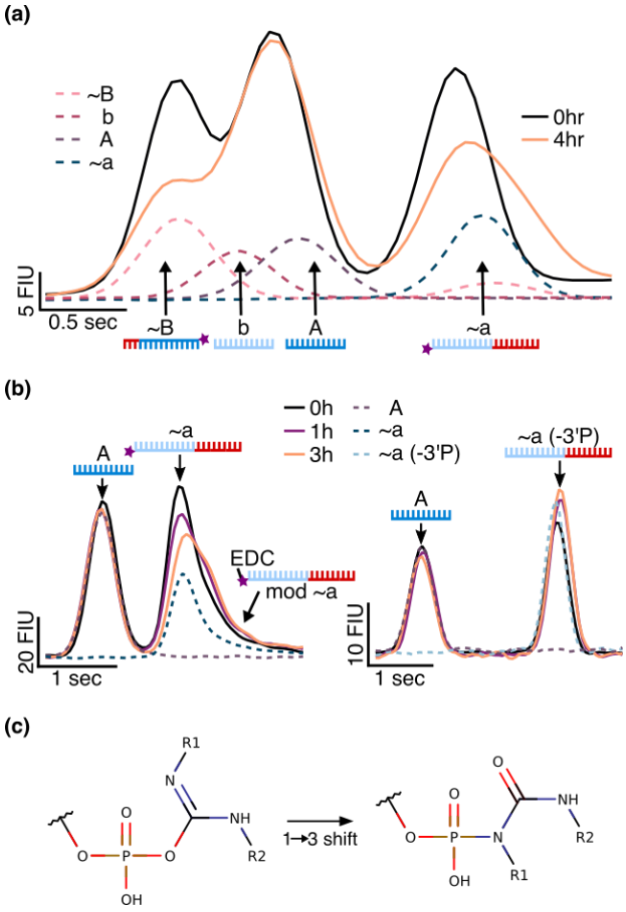


Figure S4. HPLC-MS analysis of side product formation.
Fragments A and ~a were incubated together in the reaction buffer under reaction conditions for 0 or 180 min, precipitated, re-dissolved in water, and analyzed on an HPLC-MS setup (SI 2).
(a) HPLC UV absorption data. Three new peaks marked with red arrows appeared after the 180 minutes of incubation in EDC. The ranges for mass spectrum extraction are highlighted in grey.
(b) Mass spectrum data extracted from different regions of the UV chromatograms. Exact masses and corresponding species are listed in Table S2. No additional masses appeared around the peaks corresponding to the isotope and ion distributions of fragment A after the incubation in EDC. An expected 5'-3' circularization by-product of ~a (~a mass minus water mass) appeared in one of the new peaks of the UV chromatogram after 180 minutes of the incubation in EDC. A by-product with a mass corresponding to a structure shown in Figure S3c appeared in three different ranges of the UV chromatogram after 180 minutes of incubation in EDC. It could represent ~a modified with EDC on the phosphate group as in Figure S3c or a modification of one of the nucleobases. Since the oligomer A did not show a corresponding by-product, it is likely that ~a was modified either on the phosphate group or on the unprotected 5' C-tag. However, the 5' C-tag did not seem to influence the ligation reaction (Figure 4a) while the side reaction by-product clearly influenced the reaction (Figure 3c). This suggests that the observed by-product was indeed a modification on the phosphate group as in Figure S3c. It should also be noted that the bases of oligomers A and ~a were hybridized to each other, possibly providing an additional protection.

In summary, the data revealed for ~a but not for a non-phosphorylated fragment A the by-products of circularization and EDC-modification. This suggests that the bases, all predominantly hybridized to complementary strands in the exponential reaction conditions of the reaction, were not significantly affected by EDC. Very likely this was the result of using only short temperature spikes in the reaction.

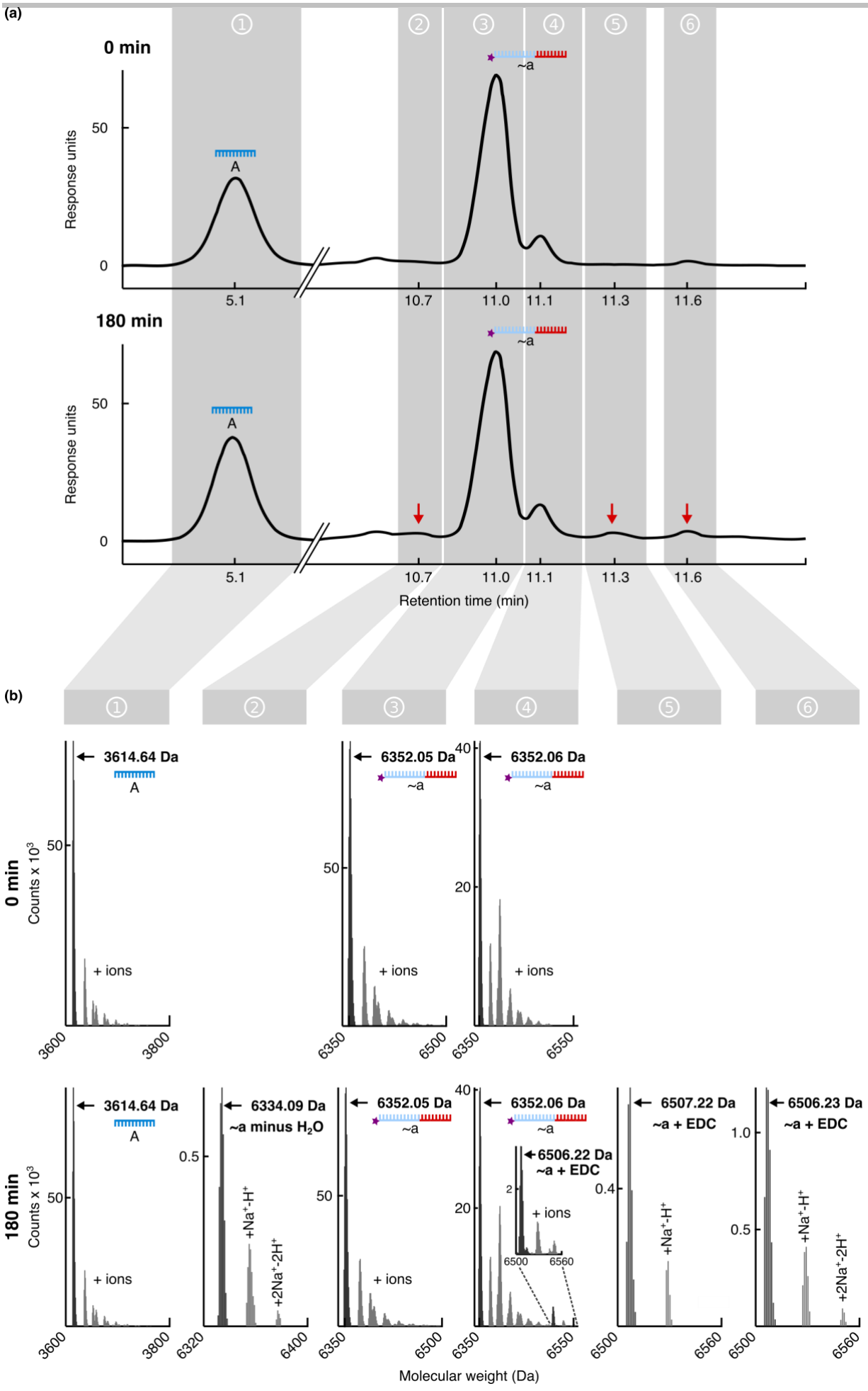

(a)
0 min
180 min
Response units
A
~a
Retention time (min)
(b)
0 min
180 min
Counts x 10^3
3614.64 Da
6352.05 Da
6352.06 Da
6334.09 Da
~a minus H2O
6506.22 Da
~a + EDC
6507.22 Da
6506.23 Da
+ ions
+Na+-H+
+2Na+-2H+
Molecular weight (Da)

Figure S5. Effect of EDC on the melting behavior of the reaction species.
(a) Baseline-corrected melting curves (0.083 °C/s heating rate) of the complementary templates and fragments in the control buffer (squares) and in the reaction buffer (circles). Presence of EDC strongly increased the thermal stability of the duplexes.
(b) Baseline-corrected melting (heating) and annealing (cooling) curves of the complementary templates and complementary fragments in the presence of EDC (in the reaction buffer) used to quantify the dissociation rates.

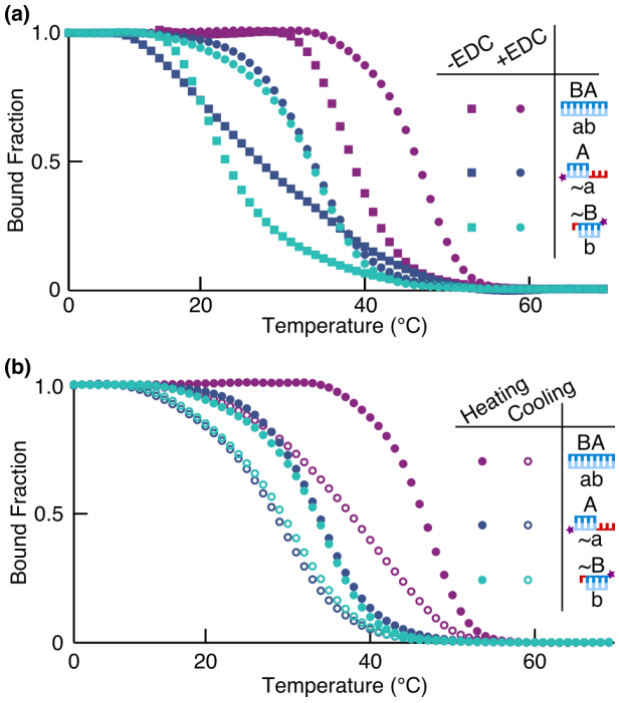
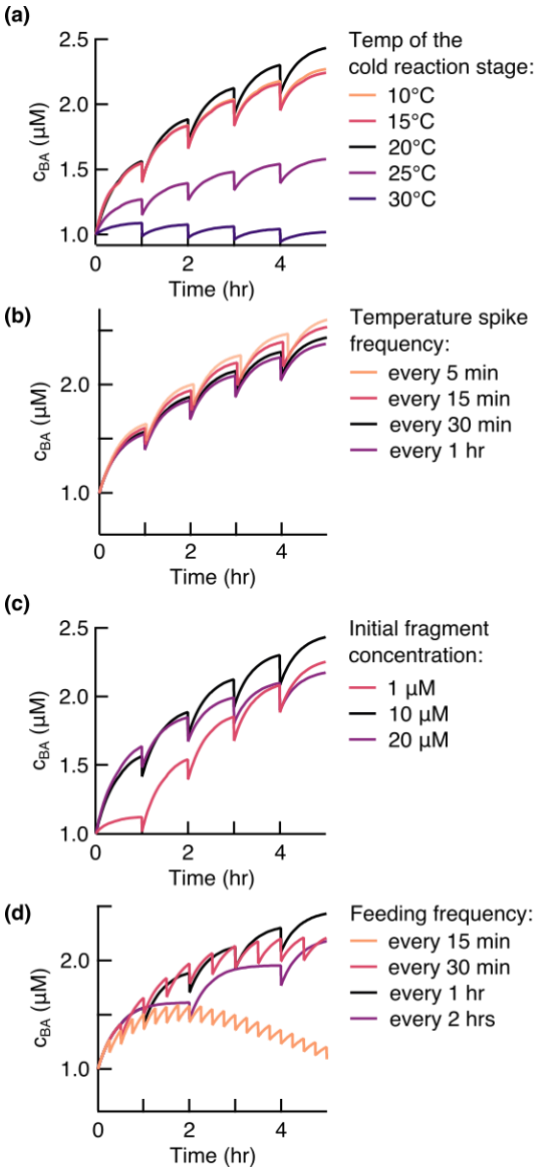


Figure S6. The influence of the reaction parameters on the replication reaction.
Black lines show the fit results to the experimental data.
(a) The result of the modulation of the temperature of the cold stage of the reaction on the replication. At 20°C, the fraction of the strands in the bound configuration approaches 1.0 (Figure S5 b). This allows the two fragment strands to be brought together by the complementary template resulting in the efficient ligation reaction. Lower temperature does not significantly affect the replication (yellow and pink curves), while 5°C and 10 °C higher temperatures are predicted to significantly suppress the replication (light and dark purple curves). The effects originate from the higher dissociation rate of duplexes at the higher temperature. However, the side reaction and ligation rates were assumed temperature independent.
(b) One more parameter that could be varied in the temperature cycling scheme is the frequency of the temperature spike. We chose a frequency of 30 min in our cycling scheme, as according to the inset in Figure 1a, that is when the product formation starts to saturate. Simulation of faster or slower cycling schemes showed no significant difference. Temperature spike every 1 hour just prolongs the time in saturation. More frequent temperature spikes may enhance the reaction, however, the more deleterious effects of side reactions at high temperatures are not considered in the simulation. We thus considered cycling every 30 min as the optimal scheme that provides the turnover of templates but keeps the side reactions to the minimum.
(c) The result of the modulation of the initial fragment concentration on the replication. Initial fragment concentration should be balanced between providing excess over template strands but not saturating the system with too many fragments that participate in the side reactions.
(d) The simulation (see the graph below) shows that introducing a longer interval between the feeding steps (purple curve, 2-hours interval) would just prolong the saturated state. On the other hand, a more often feeding (pink and yellow curves, 30- and 15-minute intervals) would result in diluting the reaction faster than the product is produced, because each feeding step also dilutes the reaction.



SUPPLEMENTARY INFORMATION

SI 1: Data quantification with the Bioanalyzer instrument

The Bioanalyzer instrument provides an automated gel electrophoresis system running on a microfluidic chip. For the analysis shown in this study, we used the Agilent Small RNA Kit. Per chip, a maximum of 12 samples could be analyzed. Although the Agilent Small RNA Kit was designed for the detection of short RNA strands (6-150nt), we established the procedure for the detection and quantification of short ssDNA strands (~ 10-100nt).

The full-timescale fluorescence traces of the example reaction and calibration samples are shown in Figure S1a. The area and position of the peak that belongs to a specific strand depends on its length, sequence, and present modifications (*e.g.* the 3'-phosphate group in our system). Thus, the identification of peaks required studying traces of samples where different combinations of species were mixed together (Figure S1a and S3a).

Calibration experiments showed that the area of the peak was proportional to the ssDNA strand concentration (Figure S1c). Using this, the concentrations of the emerging products were calculated from the two calibration traces ran on the same Bioanalyzer chip. The two calibration traces included a template or its complement and both product strands with the known concentration of 50 nM each (Figure S1b). The FIU/nM (fluorescence intensity signal per concentration) value was calculated for each chip from the average peak area for each product from the two calibration curves. The peak area was calculated by fitting Gaussian functions (Figure S1b). Due to a varying gel/dye composition and subsequently slightly varying strength of the fluorescence signal between Bioanalyzer chips, the two calibration samples were included in every run.

In addition, we observed a slight fluorescence signal variation between different wells of the same Bioanalyzer chip. Including an internal control should exclude this fluorescence variation. According to the company protocol, a commercially available marker solution with a normalization strand has to be included. However, the marker strand overlaid with some of the fragment peaks and could not be used as a fluorescence control in our system (Figure S1a). Instead, we included an 84 nt-long DNA strand as a normalization strand (sequence in the Experimental section in the main text). All the experimental curves of the same chip were normalized to the average peak area of the normalization strand from the two calibration curves of the same chip.

Figures 1c and 1d in the main text present curves normalized to the peak area of the normalization strand. The slight baseline shift observed for hours 4 and 5, we attribute to the effect of fragment feeding, which increased the material amount 4- and 5-fold in the fragment peaks after 4 and 5 hours, respectively. A slight bump to the left of the templates peak in Figure 1c observed for a 2-hour curve had a stochastic origin and occurs for some Bioanalyzer curves. The effect of templates dilution is strongly pronounced in Figure 1c due to the nonlinear relationship between the fluorescence signal and the concentration of the double stranded DNA. Only single stranded DNA has a linear relationship to the fluorescence signal in the Agilent Small RNA Kit assay (Figure S1c). The effect of template dilution in Figure 1d is not pronounced because a single strand of the template was used in this experiment and the effect of the 10% dilution was hidden by the 10% experimental error.

To quantify the error that analysis on the Bioanalyzer instrument yielded, we performed a control experiment, where in addition to the marker solution, the same sample consisting of one template, two products, and the normalization strand was pipetted into every well. It showed that the area normalization between the curves of different wells could be performed by normalizing either to the peak area of the marker strand or to the peak area of the normalization strand introduced by us. The well-to-well error left after the normalization was 10%, which corresponds to the pipetting error. The 10% pipetting error was thus included as the predominant error source in the results.

In each experiment, the amplification factor was quantified separately for BA and ab motifs as $A_{BA} = c_{\sim BA+BA}^{t=\Delta t}/c_{BA}^{t=0}$ and $A_{ab} = c_{\sim ab+ab}^{t=\Delta t}/c_{ab}^{t=0}$, where A_{BA} and A_{ab} are the amplification factors for BA and ab, $c_{\sim BA+BA}^{t=\Delta t}$ is the BA concentration originating from template BA and the replicate ~BA at the end of the reaction ($c_{\sim ab+ab}^{t=\Delta t}$ represents the same for the ab concentration), and $c_{BA}^{t=0}$ is the BA concentration originating only from the template at the start of the reaction ($c_{ab}^{t=0}$ represents the same for the ab concentration). The average amplification factor and its standard deviation in the main text were obtained from four different experiments where individual amplification factors for BA and ab were calculated (Table S1 for details). Variation in the precipitation yields and pipetting errors are the likely sources of discrepancies between the amplification factors in different experiments.

Table S1. Reaction conditions and amplification factors for BA/ab motifs in four reactions.

Exp. Nr.	$c_{BA}^{t=0}$ (µM)	$c_{ab}^{t=0}$ (µM)	Δt (hrs)	A_{BA}	A_{ab}
1	1	1	6	2.12	2.08
2	0.7	0.7	6	1.52	1.66
3	0.5	0.5	6	2.36	3.03
4	1	1	5	2.21	2.44

SI 2: Details to the HPLC-MS data acquisition and analysis

For the analysis of the reaction byproducts, a bioinert HPLC-MS system by Agilent Technologies was used. The HPLC consists of a quaternary pump with solvent degasser, a temperature-controlled sample holder cooled to 16°C with multisampler, a column of the type AdvanceBio Oligonucleotides (4.6 x 150 mm, 2.7 µm) held at 30°C, and a UV absorption diode array detector (DAD), recording chromatograms at 259 nm. The outflow of the HPLC is injected into the ESI-QTOF with two injection needles providing direct comparison with a reference and, therefore, a high degree of accuracy.

4µL of sample was eluted with a mixture of two solvents. Solvent A: DI water with 8mM TEA and 200mM HFIP, and solvent B: LC grade Methanol. At a flow rate of 0.6mL/min the following gradient was run: 0 min 20% B, 15 min 28% B, 15.1 min 95% B, 19 min 95% B, 19.1 min 20% B, Stoptime 29 min.

The MS was run in a negative mode with 8 L/min drying gas flow at 300°C, 45 psig nebulizer, acquisition time of 200 ms per spectrum, scanning 5 spectra per second in a mass range of 90 to 3200 m/z, fragmentor voltage 175V, capillary voltage 4 kV, and nozzle voltage 2 kV.

The UV absorption peaks were first assigned using reference samples (reference UV data not shown because the peaks were also assigned based on the extracted MS spectra, for which the data is presented in Figure S4). Next, the LCMS data was processed with the Agilent MassHunter Qualitative Analysis Navigator Software B.08.00 to deconvolute the mass spectra corresponding to different ranges of the UV absorption peaks of the analyte. The ranges are specified in the Table S2. The most important peaks in the mass spectra are noted with the arrows in Figure S4. The rest of the assigned masses are presented in the Table S2.

Table S2. Assigned masses corresponding to different ranges for MS spectra extraction from Figure S4.

	0 min	**180 min**
Range 1 (4.98-5.31 min)	3614.64 Da: A 3636.62 Da: A + Na^+ - H^+ 3652.59 Da: A + K^+ - H^+ 3658.60 Da: A + 2Na^+ - 2H^+ 3674.57 Da: A + Na^+ + K^+ - 2H^+ 3680.58 Da: A + 3Na^+ - 3H^+	3614.64 Da: A 3636.62 Da: A + Na^+ - H^+ 3652.59 Da: A + K^+ - H^+ 3658.60 Da: A + 2Na^+ - 2H^+ 3674.57 Da: A + Na^+ + K^+ - 2H^+ 3680.58 Da: A + 3Na^+ - 3H^+
Range 2 (10.63-10.78 min)		**6334.09 Da: ~a - H_2O** **6355.05 Da: ~a - H_2O + Na^+ - H^+** **6377.03 Da: ~a - H_2O + 2Na^+ - 2H^+**
Range 3 (10.79-11.05 min)	6352.05 Da: ~a 6374.05 Da: ~a + Na^+ - H^+ 6389.02 Da: ~a + K^+ - H^+ 6396.03 Da: ~a + 2Na^+ - 2H^+	6352.05 Da: ~a 6374.05 Da: ~a + Na^+ - H^+ 6389.02 Da: ~a + K^+ - H^+ 6396.03 Da: ~a + 2Na^+ - 2H^+
Range 4 (11.06-11.24 min)	6352.06 Da: ~a 6374.05 Da: ~a + Na^+ - H^+ 6389.02 Da: ~a + K^+ - H^+	6352.06 Da: ~a 6374.05 Da: ~a + Na^+ - H^+ 6389.02 Da: ~a + K^+ - H^+ **6506.22 Da: ~a + EDC** **6528.20 Da: ~a + EDC + Na^+ - H^+** **6544.12 Da: ~a + EDC + K^+ - H^+** **6550.17 Da: ~a + EDC + 2Na^+ - 2H^+**
Range 5 (11.26-11.46 min)		**6507.22 Da: ~a + EDC** **6529.19 Da: ~a + EDC + Na^+ - H^+**
Range 6 (11.47-11.88 min)		**6506.23 Da: ~a + EDC** **6529.20 Da: ~a + EDC + Na^+ - H^+** **6550.20 Da: ~a + EDC + 2Na^+ - 2H^+**

Table Legend: "A" denotes fragment A; "~a" denotes fragment ~a; "~a + EDC" corresponds to the product of ~a modification with EDC as in Figure S3c. The masses presented in the table correspond to the isotopes with the highest counts for each identified group of isotope distributions. In case of ~a, two isotopes are expected to be almost equally abundant: isotopes with masses 6351.09 Da and 6352.09 Da with 100% and 96.07% normalized abundances. Since both masses are highly abundant, we identify sometimes 6351.09 Da and sometimes 6352.09 Da mass as appearing with the highest counts in different isotope distributions. The same is true of the molecule ~a + EDC where the two isotopes of the highest abundance have masses of 6506.23 Da and 6507.23 Da.

SI 3: Details of the kinetic model and the fitting method

The kinetic modeling was realized in MATLAB. The model is based on 23 rate equations (Schemes 1 and 2 in the main text) and 23 ordinary differential equations (ODEs) describing the rate equations (Table S4). To simulate temperature cycling, the ODEs were alternately solved with two sets of rate constants (Table S3).

We obtained the k_{on} value of 1440 μM^{-1} hr^{-1} for a short DNA oligonucleotide composed only of adenine and thymine nucleobases from literature[3,4]. The rates k_{on}^{12nt} and k_{on}^{24nt} were assumed length independent and were both set equal to this k_{on} value.

Since we wanted to explicitly model the temperature cycling, we needed to determine $k_{off}^{12nt,20℃}$, $k_{off}^{12nt,65℃}$, $k_{off}^{24nt,20℃}$, and $k_{off}^{24nt,65℃}$.

The values for $k_{off}^{12nt,20℃}$ and $k_{off}^{24nt,20℃}$ were quantified experimentally by analyzing melting curves as described in the Experimental section of this paper and in the review[5]. First, we showed that the duplexes in the reaction buffer were significantly more stable than in the control buffer (reaction buffer without EDC) (Figure S5a). Therefore, measuring the dissociation rates in the control buffer would have yielded overestimated values. Then, we measured melting and annealing curves in the reaction buffer and observed a hysteresis between them for all the tested species combinations (Figure S5b). This means that a slower temperature ramp was required to reach the dynamic equilibrium at each temperature point of data acquisition. However, we expected that the slower temperature ramp would increase the extent of EDC-induced unwanted modifications of the DNA bases in the single-stranded form[6]. Instead of slowing down the temperature ramp, we decided to estimate the minimum and maximum $k_{off}^{nt,20℃}$ values from the melting and annealing curves, respectively. We quantified $k_{off}^{24nt,20℃}$ and $k_{off}^{12nt,20℃}$ to be within the bounds of (2 to 64100) 10^{-4} hr^{-1} and (17 to 339) hr^{-1}, respectively. Although the range for the $k_{off}^{nt,20℃}$ values is large, these dissociation rates are slower than the melting kinetics of the implemented thermal cycling. The reaction model was not expected to be sensitive to variations of this parameter as indeed was the case.

The values $k_{off}^{24nt,65℃}$ and $k_{off}^{12nt,65℃}$ estimated in the same way yielded respective bounds of (0.4 to 9.6) 10^{6} and (0.4 to 1.1) 10^{8} hr^{-1}. The result of the large dissociation rates is the absence of duplex strands at equilibrium at 65 °C in our system. Without significantly altering the simulation results, we approximated the rates $k_{off}^{24nt,65℃}$ and $k_{off}^{12nt,65℃}$ to be 100 times higher than the chosen k_{on} value. This allowed the dissociation of the majority of the DNA complexes during the simulated melting steps and reduced the computational load for the simulation.

The rate of EDC hydrolysis at the pH value 6.5 of the reaction was taken from literature[7–9] to be 0.047 hr^{-1}, leading to 75% of the initial EDC concentration left after 6 hours of the reaction, maintaining the >2500-fold excess of EDC over the DNA oligonucleotides. Therefore, we did not explicitly model the hydrolysis of EDC in the simulation.

As mentioned in the main text, the rates k_{feed} and k_{dil} were fixed to model the experimental feeding and dilution concentrations. The quantification of $k_{s.r.}$ and k_{lig} was provided in the main text.

The ODEs were solved using an *ode15s* solver to obtain concentrations of all the reaction species at time points chosen by the solver. The absolute and relative tolerances of the solver were set to 10^{-3} and 10^{-1}, respectively.

Table S3. Two sets of rate constants that model temperature cycling in the system.

Rate Constant	20ºC for 30 min	65ºC for 16 sec
Dissociation for 12nt long strands	$\boldsymbol{k_{off}^{12nt,20℃}}$	$\boldsymbol{k_{off}^{12nt,65℃}}$
Dissociation for 24nt long strands	$\boldsymbol{k_{off}^{24nt,20℃}}$	$\boldsymbol{k_{off}^{24nt,65℃}}$
Association for all strands	$\boldsymbol{k_{on}}$	$\boldsymbol{k_{on}}$
Ligation	$\boldsymbol{k_{lig}}$	$\boldsymbol{k_{lig}}$
Feeding	0	$\boldsymbol{k_{feed}}$
Dilution	0	$\boldsymbol{k_{dil}}$
Side reaction	$\boldsymbol{k_{s.r.}}$	$\boldsymbol{k_{s.r.}}$

Since the solution of the kinetic model reproduced the experimental behavior of the system sufficiently well, we used this kinetic model to fit the experimental data for the side reaction rate. The fit was realized with the *lsqcurvefit* tool of MATLAB with $k_{s.r.}$ as the only fit parameter.

The MATLAB code that allows to run the simulation, fit the data, and plot the results is attached to this paper.

Table S4. ODEs describing the experimental system (terms describing side reaction pathways are highlighted in bold).

Left side	Right side
dB/dt	$-k_{dil}*B+k_{feed}-\boldsymbol{k_{s.r.}}*\mathbf{B}+k_{off}^{12nt}*([B\cdot b]+[B\cdot ab]+[B\cdot A\cdot ab])-k_{on}^{12nt}*(B*b+B*ab+B*[A\cdot ab])$
dA/dt	$-k_{dil}*A+k_{feed}+k_{off}^{12nt}*([A\cdot a]+[A\cdot ab]+[B\cdot A\cdot ab]+[\mathbf{A}\cdot\mathbf{a}^*]+[\mathbf{B}^*\mathbf{\cdot A\cdot ab}])-k_{on}^{12nt}*(A*a+A*ab+A*[B\cdot ab]+\mathbf{A}*\boldsymbol{a}^*+\mathbf{A}*\mathbf{[B^*\cdot ab]})$
db/dt	$-k_{dil}*b+k_{feed}+k_{off}^{12nt}*([B\cdot b]+[BA\cdot b]+[BA\cdot a\cdot b]+[\mathbf{B}^*\cdot\mathbf{b}]+[\mathbf{BA\cdot a^*\cdot b}])-k_{on}^{12nt}*(b*B+b*BA+b*[BA\cdot a]+\mathbf{b}*\mathbf{B}^*+\mathbf{b}*\mathbf{[BA\cdot a^*]})$
da/dt	$-k_{dil}*a+k_{feed}-\boldsymbol{k_{s.r.}}*\mathbf{a}+k_{off}^{12nt}*([A\cdot a]+[BA\cdot a]+[BA\cdot a\cdot b])-k_{on}^{12nt}*(a*A+a*BA+a*[BA\cdot b])$
dBA/dt	$-k_{dil}*BA+k_{off}^{12nt}*([BA\cdot a]+[BA\cdot b]+[\mathbf{BA\cdot a^*}])-k_{on}^{12nt}*(BA*a+BA*b+\mathbf{BA}*\mathbf{a}^*)+k_{off}^{24nt}*[BA\cdot ab]-k_{on}^{24nt}*BA*ab$
dab/dt	$-k_{dil}*ab+k_{off}^{12nt}*([A\cdot ab]+[B\cdot ab]+[\mathbf{B}^*\mathbf{\cdot ab}])-k_{on}^{12nt}*(ab*A+ab*B+\mathbf{ab}*\mathbf{B}^*)+k_{off}^{24nt}*[BA\cdot ab]-k_{on}^{24nt}*BA*ab$
d[BA·ab]/dt	$-k_{dil}*[BA\cdot ab]+k_{on}^{24nt}*BA*ab-k_{off}^{24nt}*[BA\cdot ab]+k_{lig}*([B\cdot A\cdot ab]+[BA\cdot a\cdot b])$
d[BA·b]/dt	$-k_{dil}*[BA\cdot b]+k_{off}^{12nt}*([BA\cdot a\cdot b]+[\mathbf{BA\cdot a^*\cdot b}]-[BA\cdot b])+k_{on}^{12nt}*(BA*b-[BA\cdot b]*a-[\mathbf{BA}\cdot\mathbf{b}]*\mathbf{a}^*)$
d[BA·a]/dt	$-k_{dil}*[BA\cdot a]-\boldsymbol{k_{s.r.}}*\mathbf{[BA\cdot a]}+k_{off}^{12nt}*([BA\cdot a\cdot b]-[BA\cdot a])+k_{on}^{12nt}*(BA*a-[BA\cdot a]*b)$
d[B·ab]/dt	$-k_{dil}*[B\cdot ab]-\boldsymbol{k_{s.r.}}*\mathbf{[B\cdot ab]}+k_{off}^{12nt}*([B\cdot A\cdot ab]-[B\cdot ab])+k_{on}^{12nt}*(B*ab-[B\cdot ab]*A)$
d[A·ab]/dt	$-k_{dil}*[A\cdot ab]+k_{off}^{12nt}*([B\cdot A\cdot ab]+\mathbf{[B^*\cdot A\cdot ab]}-[A\cdot ab])+k_{on}^{12nt}*(A*ab-[A\cdot ab]*B-[\mathbf{A}\cdot\mathbf{ab}]*\mathbf{B}^*)$
d[B·A·ab]/dt	$-k_{dil}*[B\cdot A\cdot ab]-\boldsymbol{k_{s.r.}}*\mathbf{[B\cdot A\cdot ab]}+k_{on}^{12nt}*(B*[A\cdot ab]+A*[B\cdot ab])-2*k_{off}^{12nt}*[B\cdot A\cdot ab]-k_{lig}*[B\cdot A\cdot ab]$
d[BA·a·b]/dt	$-k_{dil}*[BA\cdot a\cdot b]-\boldsymbol{k_{s.r.}}*\mathbf{[BA\cdot a\cdot b]}+k_{on}^{12nt}*(b*[BA\cdot a]+a*[BA\cdot b])-2*k_{off}^{12nt}*[BA\cdot a\cdot b]-k_{lig}*[BA\cdot a\cdot b]$
d[B·b]/dt	$-k_{dil}*[B\cdot b]-\boldsymbol{k_{s.r.}}*\mathbf{[B\cdot b]}+k_{on}^{12nt}*B*b-k_{off}^{12nt}*[B\cdot b]$
d[A·a]/dt	$-k_{dil}*[A\cdot a]-\boldsymbol{k_{s.r.}}*\mathbf{[A\cdot a]}+k_{on}^{12nt}*A*a-k_{off}^{12nt}*[A\cdot a]$
dB*/dt	$-\mathbf{k_{dil}}*\mathbf{B}^*+\boldsymbol{k_{s.r.}}*\mathbf{B}+\mathbf{k_{off}^{12nt}}*(\mathbf{[B^*\cdot b]}+\mathbf{[B^*\cdot ab]}+\mathbf{[B^*\cdot A\cdot ab]})-\mathbf{k_{on}^{12nt}}*(\mathbf{B}^**\mathbf{b}+\mathbf{B}^**\mathbf{ab}+\mathbf{B}^**\mathbf{[A\cdot ab]})$
da*/dt	$-\mathbf{k_{dil}}*\mathbf{a}^*+\boldsymbol{k_{s.r.}}*\mathbf{a}+\mathbf{k_{off}^{12nt}}*(\mathbf{[A\cdot a^*]}+\mathbf{[BA\cdot a^*]}+\mathbf{[BA\cdot a^*\cdot b]})-\mathbf{k_{on}^{12nt}}*(\mathbf{a}^**\mathbf{A}+\mathbf{a}^**\mathbf{BA}+\mathbf{a}^**\mathbf{[BA\cdot b]})$
d[B*·ab]/dt	$-\mathbf{k_{dil}}*\mathbf{[B^*\cdot ab]}+\boldsymbol{k_{s.r.}}*\mathbf{[B\cdot ab]}+\mathbf{k_{off}^{12nt}}*(\mathbf{[B^*\cdot A\cdot ab]}-\mathbf{[B^*\cdot ab]})+\mathbf{k_{on}^{12nt}}*(\mathbf{B}^**\mathbf{ab}-\mathbf{[B^*\cdot ab]}*\mathbf{A})$
d[BA· a*]/dt	$-\mathbf{k_{dil}}*\mathbf{[BA\cdot a^*]}+\boldsymbol{k_{s.r.}}*\mathbf{[BA\cdot a]}+\mathbf{k_{off}^{12nt}}*(\mathbf{[BA\cdot a^*\cdot b]}-\mathbf{[BA\cdot a^*]})+\mathbf{k_{on}^{12nt}}*(\mathbf{BA}*\mathbf{a}^*-\mathbf{[BA\cdot a^*]}*\mathbf{b})$
d[B*·A·ab]/dt	$-\mathbf{k_{dil}}*\mathbf{[B^*\cdot A\cdot ab]}+\boldsymbol{k_{s.r.}}*\mathbf{[B\cdot A\cdot ab]}+\mathbf{k_{on}^{12nt}}*(\mathbf{B}^**\mathbf{[A\cdot ab]}+\mathbf{A}*\mathbf{[B^*\cdot ab]})-\mathbf{2}*\mathbf{k_{off}^{12nt}}*\mathbf{[B^*\cdot A\cdot ab]}$
d[BA· a*·b]/dt	$-\mathbf{k_{dil}}*\mathbf{[BA\cdot a^*\cdot b]}+\boldsymbol{k_{s.r.}}*\mathbf{[BA\cdot a\cdot b]}+\mathbf{k_{on}^{12nt}}*(\mathbf{b}*\mathbf{[BA\cdot a^*]}+\mathbf{a}^**\mathbf{[BA\cdot b]})-\mathbf{2}*\mathbf{k_{off}^{12nt}}*\mathbf{[BA\cdot a^*\cdot b]}$
d[B*·b]/dt	$-\mathbf{k_{dil}}*\mathbf{[B^*\cdot b]}+\boldsymbol{k_{s.r.}}*\mathbf{[B\cdot b]}+\mathbf{k_{on}^{12nt}}*\mathbf{B}^**\mathbf{b}-\mathbf{k_{off}^{12nt}}*\mathbf{[B^*\cdot b]}$
d[A· a*]/dt	$-\mathbf{k_{dil}}*\mathbf{[A\cdot a^*]}+\boldsymbol{k_{s.r.}}*\mathbf{[A\cdot a]}+\mathbf{k_{on}^{12nt}}*\mathbf{A}*\mathbf{a}^*-\mathbf{k_{off}^{12nt}}*\mathbf{[A\cdot a^*]}$

References

1 H. G. Khorana and A. R. Todd, *J. Chem. Soc.*, 1953, **0**, 2257–2260.

2 M. Mikolajczyk, P. Kielbasinski and W. Basinski, *J. Org. Chem.*, 1984, **49**, 899–908.

3 T. E. Ouldridge, P. Šulc, F. Romano, J. P. K. Doye and A. A. Louis, *Nucleic Acids Res.*, 2013, **41**, 8886–8895.

4 D. Y. Zhang and E. Winfree, *J. Am. Chem. Soc.*, 2009, **131**, 17303–17314.

5 J.-L. Mergny and L. Lacroix, *Oligonucleotides*, 2003, **13**, 515–537.

6 V. F. Drevitch, R. I. Salganik, D. G. Knorre and E. G. Mal'Ygin, *Biochim. Biophys. Acta - Nucleic Acids Protein Synth.*, 1966, **123**, 207–209.

7 A. Williams and I. T. Ibrahim, *Chem. Rev.*, 1981, **81**, 589–636.

8 N. Wrobel, M. Schinkinger and V. M. Mirsky, *Anal. Biochem.*, 2002, **305**, 135–138.

9 M. Tena-Solsona, B. Rieß, R. K. Grötsch, F. C. Löhrer, C. Wanzke, B. Käsdorf, A. R. Bausch, P. Müller-Buschbaum, O. Lieleg and J. Boekhoven, *Nat. Commun.*, 2017, **8**, 15895.

Author Contributions

E.E. designed and performed the experiments. A.S., J.S., and P.S. performed the experiments. D.B. conceived and supervised the project. E.E., J.B., and D.B. wrote the paper.