nature chemistry | SUPPLEMENTARY INFORMATION
DOI: 10.1038/NCHEM.1869

Diversity in the dynamical behaviour of a compartmentalized programmable biochemical oscillator

Maximilian Weitz[1], Jongmin Kim[2], Korbinian Kapsner[1], Erik Winfree[2,3,4], Elisa Franco[5], and Friedrich C. Simmel[*,1]

[1] *Systems Biophysics and Bionanotechnology, Physik Department, Technische Universität München, Am Coulombwall 4a, 85748 Garching, Germany*
[2] *Bioengineering, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91125, USA*
[3] *Computation and Neural Systems, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91125, USA*
[4] *Computer Science, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91125, USA*
[5] *Mechanical Engineering, University of California at Riverside 900 University Avenue, Riverside, CA 92521, USA*

Contents

Supplementary Methods 3
- DNA Sequences 3
- Sample preparation 4
- Droplet generation 5
  - Droplet produced with 'shaken-not-stirred' method 5
  - Droplets produced in microfluidics 5
- Fluorescence data acquisition 7
- Time-lapse microscopy 7
- Analysis of time-lapse movies 7

Supplementary Note - Data Analysis 8
- Data sets 8
- Algorithms 8
  - Data filtering 8
  - Detection of steps in the time traces 9
  - Detection of extrema 10
  - Calculation of periods and amplitudes 10
- Data preprocessing 12
- Identification of valid data 12

Supplementary Note - Oscillator Data 15
- Sustained oscillator 15
- Damped oscillator 16
- Strongly damped oscillator 18

Supplementary Note - Compartmentalization 21

Distribution and quantification of biomolecules in droplets 21
Variability in microdroplet-encapsulated single-enzyme subsystems 24

Supplementary Note - Modelling **35**
Model Equations 35
Deterministic Simulation and Model Fits 38
    Sensitivity to enzyme concentrations 39
Deterministic Simulation with Partitioning Error 42
    Deterministic Simulation with Enzyme Loss and Partitioning Error 46
Stochastic Simulation 47

Supplementary Methods

DNA Sequences

The following DNA sequences were either ordered from IDT DNA (Belgium) (A1, A2), or biomers.net (Ulm, Germany) (T12 nt/t, T21 nt/t, dI1). These sequences are taken from [Kim, 2007] and [Kim and Winfree, 2011]. Oligonucleotides were HPLC purified, except for Texas Red labeled T21nt which was purified with polyacrylamide gel electrophoresis.

Table S1: Oscillator sequences

Name	Sequence
T12-t	5'-TTT CTG ACT TTG TCA GTA TTA GTG TGT AGT AGT AGT TCA TTA GTG TCG TTC GTT CTT TGT TTC TCC CTA TAG TGA GTC G
T12-nt	5'-AAG CAA GGG TAA GAT GGA ATG ATA ATA CGA CTC ACT ATA GGG AGA AAC AAA GAA CGA ACG ACA CTA ATG AAC TAC TAC TAC ACA CTA ATA CTG ACA AAG TCA GAA A
T21-t	5'-TTT CTG ACT TTG TCA GTA TTA TCA TTC CAT CTT ACC CTT GCT TCA ATC CGT TTT ACT CTC CCT ATA GTG AGT CG
T21-nt	5'-TexasRed-CAT TAG TGT CGT TCG TTC ACA GTA ATA CGA CTC ACT ATA GGG AGA GTA AAA CGG ATT GAA GCA AGG GTA AGA TGG AAT GAT AAT ACT GAC AAA GTC AGA AA
dI1	5'-GTG TGT AGT AGT AGT TCA TTA GTG TCG TTC GTT CAC AG
A1	5'-TAT TAC TGT GAA CGA ACG ACA CTA ATG AAC TAC TAC-Iowa Black RQ
A2	5'-TAT TAT CAT TCC ATC TTA CCC TTG CTT CAA TCC GT-Iowa Black RQ
rA1 (RNA)	5'-GGG AGA AAC AAA GAA CGA ACG ACA CUA AUG AAC UAC UAC UAC ACA CUA AUA CUG ACA AAG UCA GAA A
rI2 (RNA)	5'-GGG AGA GUA AAA CGG AUU GAA GCA AGG GUA AGA UGG AAU GAU AAU ACU GAC AAA GUC AGA AA

Table S2: Subsystem DNA sequences

Name	Sequence
Genelet-t	5'-TAG TTG TGA GTC GTA TTA ATT GAA TGG TTA TGT GGA TCT ATA GTG AGT CGT ATT AAT TGA AT
Genelet-nt	5'-ATT CAA TTA ATA CGA CTC ACT ATA GAT CCA CAT AAC CAT TCA ATT AAT ACG ACT CAC AAC TA
Rep-Q	5'-GTG AGT CGT ATT AAT TGA ATG GTT ATG T-BHQ2
Rep-F	5'-TAMRA-ATT CAA TTA ATA CGA CTC AC

Sample preparation

The final concentrations of oscillator DNA strands for all data sets were: T12 (80 nM), T21 (150 nM), dI1 (500 nM), A1 (150 nM), A2 (250 nM). T12 and T21 genelets were annealed seperately in 1x NEB RNAPol reaction buffer (B9012S) to a final stock concentration of 5 µM. The strands were annealed by heating up to 95 ° C for 5 min followed by cooling down to 20 °C over 80 min using a Thermomixer (Eppendorf). The transcription reaction mix contained 1x NEB RNAPol reaction buffer (B9012S), additional 24 mM $MgCl_2$ (6 mM already part of TX buffer), 7.5 mM each dNTP (Epicentre, RN02825) and 1x inorganic Pyrophosphatase (Sigma Aldrich, I1891-100UN). For data normalization purposes, we added 500 nM Alexa488 to the reaction mix. For details on data normalization see Supplementary Note - Data Analysis (Identification of valid data).

The inorganic Pyrophosphatase was solved in 20 mM TrisHCl (pH 7.2 at 25 °C), 1 mM $MgCl_2$ and 50 % glycerol yielding in a 100 µg/mL stock solution. According to the vendor, this concentration corresponds to 100x of the optimal PPase concentration.

For sustained oscillations we used E. coli RNase H (Ambion, AM2292, 10 units/µL) and T7 RNA polymerase (Epicentre, TM910K) in a volume relation of 1:10. T7 RNA polymerase (Epicentre, TM910K) was the last ingredient added to the reaction mix to a final concentration of 12.15 units/µL to start the reaction. Enzyme concentrations for all data sets shown are listed in Supplementary Table S3. For RNAP, the nominal concentration of enzyme stock for our specific batch was 3.6 µM according to the manufacturer. For RNase H, the nominal concentration of enzyme stock for our specific batch was not available from the manufacturer. By analyzing information on several other batches from the same vendor, which reported specific activity for RNase H ranging from 500,000–1,000,000 units/mg, we concluded that a reasonable nominal concentration of RNase H stock is 0.5–1 µM. We used 0.646 µM as a reasonable estimate for RNase H stock concentration. Following the suggested method to extend the lifetime of the transcription reaction [Milburn et al., U. S. Patent 5,256,555, 1993], we also included $\approx 0.001\,\mathrm{units}/\mu\mathrm{L}$ of inorganic pyrophosphatase (Sigma Aldrich, I1891-100UN) in the reaction mixture; based on the vendor reported activity $\geq$ 1000 units/mg and MW of 71 kDA, the bulk pyrophosphatase concentration was ≈ 12 nM. As pyrophosphatase is involved in enhancing RNA transcription and is not directly involved in the dynamics, we do not call it an 'essential enzyme' for the circuit dynamics and neglect this enzyme in our models.

After preparation of the reaction mix, it was added to pre-heated Hellma Quartz cuvettes (105.252-QS) without enzymes to record the maximum fluorescence intensity value, which corresponds to the maximum concentration of genelet T21 in the off-state. After 10 min of equilibration time, enzymes were added to the cuvettes and mixed by vigorously pipetting, yielding a total sample volume of 120 µL. For droplet generation, 4 aliquots of 10 µL were extracted before the sample in the cuvette was finally sealed with 35 µL hexadecane (Merk, 8.20633.1000) to prevent evaporation.

Table S3: Enzyme concentrations for sustained (s), damped (d) and strongly damped (sd) oscillator tuning. DNA concentrations are the same for all conditions, while enzyme concentrations were tuned.

Species	damped	sustained	strongly damped
T21 (nM)	150	150	150
A1 (nM)	150	150	150
dI1 (nM)	500	500	500
T12 (nM)	80	80	80
A2 (nM)	250	250	250
RNAP (nM)	218	218	109
RNase H (nM)	5	4	3.3

Droplet generation

Droplet produced with ‘shaken-not-stirred’ method

Droplets discussed in the paper were generated with the ‘shaken-not-stirred’ method. Therefore, we mixed 10 µL oscillator reaction mix with 45 µL FC-40 Fluorinert oil (Sigma-Aldrich, F9755) containing 1.8 % (w/w) E2K0660 non-ionic, biocompatible surfactant (RainDance Technologies, Lexington, MA 02421, USA) in Protein LoBind tubes (Eppendorf) using a benchtop Vortex mixer for 60 s at maximum speed. A volume of 45 µL emulsion was transferred with a pipette from the reaction tube to ibidi µ-slides VI$^{0.4}$ for microscope measurements. The flow chambers were sealed with PCR tape to protect the sample against evaporation. This technique offers the possibility to generate a big range of different sized compartments in a very short time. Example time traces of the different data sets are shown in the main text Figures 2, 3 and 4 and in Supplementary Note - Oscillator Data, where also size distributions of the droplets can be found.

Droplets produced in microfluidics

Microfluidic encapsulation methods offer an elegant way to generate populations of compartmentalized (bio-)chemical systems. Typically, the size distribution of the resulting reaction volumes is much narrower than those of droplets created with the ‘shaken-not-stirred’ method. Figure S1 shows an array of droplets produced in a microfluidic chip with flow-focus channel geometry. A typical time trace of the oscillator encapsulated utilizing the microfluidic technique is shown in Figure S1d. For droplet production, we used a hydrodynamic flow-focusing polydimethylsiloxane device as reported in [Holtze et al, 2008]. Typical pumping rates were 300 µL/h for oil and 100 µL/h for the reaction mix. To maintain hydrophobic channel walls which prevents from wetting, we heated our channels at 200 °C for 3 h according to [Kaneda et al, 2012]. Since we were interested on size dependencies we preferred to use droplets generated with the ‘shaken-not-stirred’ method for the analysis presented in the main text.

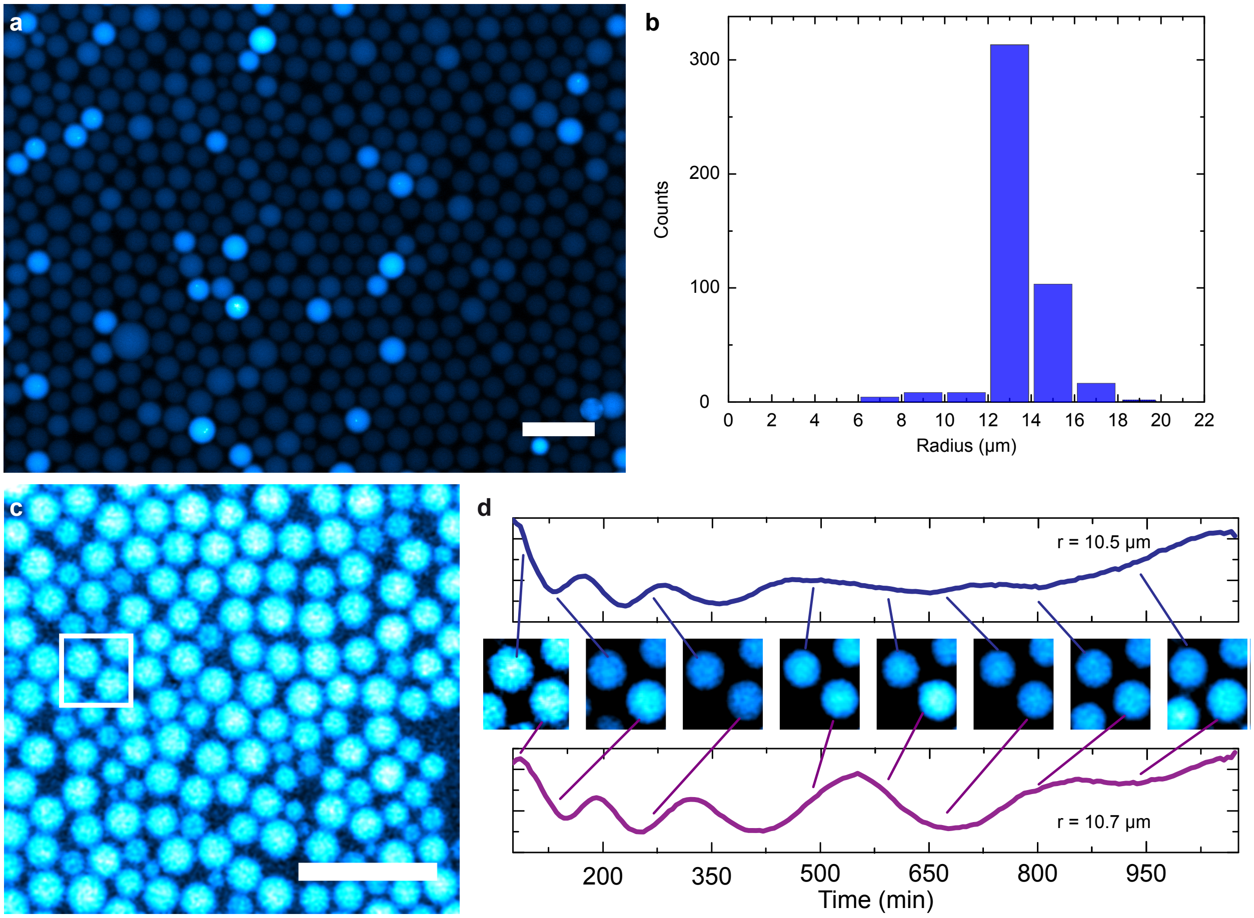

Figure S1: **(a)** Epifluorescence microscopy image of an array of droplets, which were produced in a microfluidic chip with flow-focus channel geometry. Droplets contain *in vitro* transcriptional oscillator reaction mix. Scale bar is 100 μm. **(b)** Corresponding size distribution of droplets shown in (a). **(c)** Epifluorescence microscopy image of microfluidic droplets containing damped oscillator. Scale bar is 100 μm. **(d)** Example fluorescence time traces of droplets in (c) which are surrounded by a white box. Droplets are associated with corresponding moments in the fluorescence time trace by solid lines. Note: due to the more elaborate production method, a time lag between encapsulation and start of observation of approximately 75 min exists because of the longer droplet generation time compared to the 'shaken-not-stirred' method. The y-axis corresponds to relative fluorescence values which are normalized for each trace to its maximum.

Fluorescence data acquisition

Constant wavelength analysis measurements were conducted on a Horiba Fluorolog 3 in Hellma Quartz cuvettes (105.252-QS). The sample temperature of 37 °C was maintained using a 4-sample changer with a heat bath temperature control. For fluorescence spectrometer measurements, final sample volumes were 80 µL. Data points were collected every minute with 0.1 s integration time. Excitation and emission wavelength for the dyes were chosen as suggested by IDT (Texas Red: $\lambda_{ex} = 598$ nm, $\lambda_{em} = 617$ nm; Alexa488: $\lambda_{ex} = 492$ nm, $\lambda_{em} = 517$ nm).

Time-lapse microscopy

Time-lapse microscopy measurements were conducted in an Olympus IX81 fluorescence microscope controlled with MicroManager 1.4. The microscope setup was equipped with an automated x-y-stage (Prior Scientific, Cambridge, UK) and an incubator box to maintain operation temperature of 37 °C. All observations were performed with 10x magnification objectives. Every 5 min, an image in brightfield mode, in TexasRed as well as GFP/ Alexa488 fluorescence mode (in combination with the appropriate filter set) were recorded for a total run time of at least 1000 min. The exposure times had to be adjusted according to the life-time dependent performance of the mercury fluorescence excitation light source. Typically we used 3 ms for brightfield, 1000 ms for TexasRed and GFP, and 500 ms for Alexa488. For later subtraction of background due to fluorescence of oil, a snapshot is recorded during each cycle of a position in a channel on the ibidi µ-slides VI$^{0.4}$ that contains oil-surfactant-mix only. Example microscopy movies accompany the SI.

Analysis of time-lapse movies

To analyze the time-lapse movies, we used the ImageJ plugin Cell Evaluator [Youssef, 2011]. The droplets were segmented and tracked using images recorded in brightfield mode. For the segmentation of objects (droplets), complexity of gray value brightfield images is reduced by converting it into a binary image. For this purpose, a threshold is set which was automatically determined with the method 'Otsu' implemented in the ImageJ plugin MultiThresholder written by Kevin Baler, Gabrial Landini and Wayne Rasband. Droplet radii were calculated assuming a circular shape of the segmented objects. The scaling from pixel to µm was 1 pixel = 0.641 µm. The radius then is given by $r = \sqrt{(A/\pi)} \cdot 0.641\,\mu\mathrm{m}$ with droplet's radius r in µm and area A in pixel.

Supplementary Note - Data Analysis

Data sets

In our experiments, the compartmentalized oscillators were observed using bright field and fluorescence microscopy. At each data acquisition time point, a bright field image (to locate droplets) and two fluorescence images (one for the oscillator and one for the additional normalization fluorophore, Alexa488, also present in the droplets) were collected.

The bright field images were used to identify and track the microdroplets. From the tracking process, a value for the area of each droplet is obtained as well as two fluorescence values. Thus each *time trace* corresponds to a set of three values (fluorescence signal $x_{fl,i}$ (a.u.), normalization signal $x_{norm,i}$ (a.u.), and droplet area $x_{area,i}$ (µm^2)) as a function of time.

From the fluorescence and normalization values $x_{fl,i}, x_{norm,i}$, a normalized *data set* is generated according to

$$x_{data,i} = \frac{x_{fl,i}}{x_{norm,i}}$$

This normalization procedure removes artifacts from spatial inhomogeneities in illumination, different illumination volumes within the droplets, and also lamp intensity fluctuations.

From each dataset, a *baseline* is generated by filtering with a Gaussian filter (see below) with a large standard deviation with a heuristically chosen value of $\sigma = 250$ min.

Baseline corrected data is then obtained from the dataset by subtracting the baseline. This removes long term fluctuations or drift in the signal. To remove noise from the baseline corrected data, once more a Gaussian filter is applied, this time with a low $\sigma = 25$ min, which takes care of fluctuations that are considerably faster than the oscillations expected from the oscillator. This filter produces a period cutoff of 80 min. The difference between the baseline corrected and the filtered data is the *noise*.

The *radius* of each droplet is calculated as the temporal mean value of the instantaneous radii calculated from the droplet areas at each time point. We assume that all segmented areas are circles, i.e.,

$$r = \langle r_i \rangle = \left\langle (x_{area,i}/\pi)^{1/2} \right\rangle$$

Algorithms

Data filtering

The data is filtered by a symmetric weighted moving average.

$$y_i = \frac{\sum_{k=1}^{i-1} x_k \cdot w_{i-k} + x_i \cdot w_0 + \sum_{k=i+1}^{n} x_k \cdot w_{k-i}}{\sum_{k=1}^{i-1} w_{i-k} + w_0 + \sum_{k=i+1}^{n} w_{k-i}}$$

where n is the number of data points, $x_k, k \in [1,n] \cap \mathbb{N}_0$ the raw data, $y_i, i \in [1,n] \cap \mathbb{N}_0$ the filtered data and $w_j, j \in [0, n-1] \cap \mathbb{N}_0$ the weights.

Gaussian filter The Gaussian filter uses a Gaussian as a weight function for the moving average:

$$w_\sigma(\tau) = e^{-\frac{\tau^2}{\sigma^2}}$$

The points are equally distributed in time so the weight w_j is calculated beforehand for better performance:

$$w_{\sigma,j} = e^{-\frac{(j \cdot \delta t)^2}{\sigma^2}}$$

Detection of steps in the time traces

During an experimental time course, the radius of each droplet and the fluorescence of the reference dye in each droplet should be approximately constant. If these quantities change abruptly (i.e. exhibit a "step" change in the time course), we take this as an indication that a droplet tracking issue occurred. In addition to software tracking errors, drifting droplets can suddenly overlap or leave the measurement area. To detect undesired steps in a time course x_i, the following algorithm is followed (example in Fig. S2):

1. A median filter is applied with filter-window size $ws = 2{\cdot}\mathrm{n}+1$ (we used $n = 5$) on the time course:

$$x_{filtered,i} = \text{median}(x_{i-n},\ \ldots,\ x_{i+n})$$

2. The discrete derivative of the filtered time course is calculated and n data points are cropped from the start and the end as median filtering produces artifacts at the boundaries.

3. Steps were identified as time points with significantly larger derivatives than expected from a 'normal' trace. For a heuristic criterion, first the 90% quantile of the absolute values of the derivatives was obtained. An outlier was defined as a data point with a derivative five times as large as this 90% quantile value.

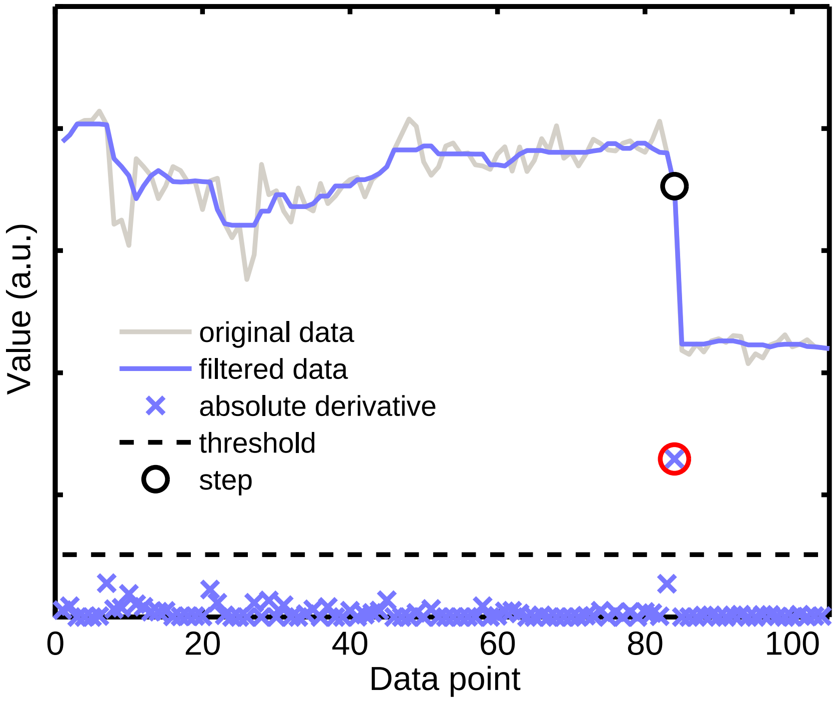

Figure S2: The data (gray line) is filtered (blue line) and the absolute value of the discrete derivative (blue crosses) is calculated. Each derivative point that is bigger than five times the 90% quantile of the whole population (dashed black line) is detected as a step (circles).

Detection of extrema

Due to the noisiness of the data, conventional extrema detection was not possible. We therefore applied an extremum detection procedure to baseline corrected data, according to which each extremum had to fulfill the following criteria:

1. A sign change has to occur between two extrema as baseline correction centers the data around zero.
2. Each extremum has to be the global extremum in its 'equal-sign' block.
3. The value difference between two consecutive extrema is larger than two times the noise level.
4. If an extremum is removed according to the previous criteria the neighboring 'equal-sign' blocks are merged together and handled as one.

Calculation of periods and amplitudes

The period and amplitude of each dataset is calculated from the extrema (cf. previous paragraph). The following criteria were applied:

1. If the first extremum happens to be at the very first time position, it is discarded (the data point might then sit next to the 'real' extremum occurring earlier in time).
2. Accordingly, if the last extremum is at the very last time position, it is also discarded.
3. If less than five extrema remain, the dataset is not identifiably oscillating and not considered for period or amplitude analysis. Five extrema define two full periods of an oscillation. Hence, if in our observation time window of 1000 minutes less than five extrema occur, the trace either does not oscillate, or the period T is larger than 500 minutes.
4. (a) The period of an oscillating trace is then obtained from the time difference of the nth and the mth extremum according to:

   $$T_{n,m} = \frac{t_m - t_n}{\frac{1}{2}(m-n)}$$

   We used $n = 1$ and $m = 3, 4, 5$ (corresponding to 3, 4 or 5 extrema for period determination). Empirically, we found $m = 3$ (3 extrema) to provide the most reliable data.

   (b) The amplitude of an oscillating trace is then obtained from the maximal absolute value difference between two consecutive extrema. But not all extrema were encountered. If the first extremum is a minimum it is discarded because the oscillator shows a very deep trough at the beginning. Maximal five extrema are considered.

The traces that were identifiably oscillating (had 5 or more extrema) are called 'oscillating' and the others are called 'non oscillating'.

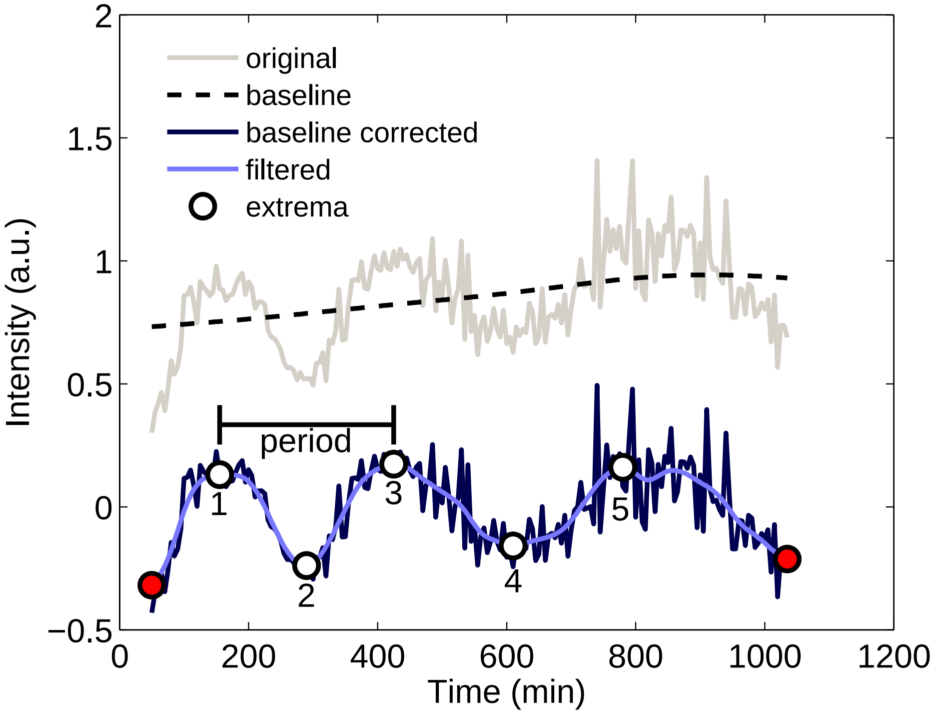

Figure S3: Example trace for data processing. The original data normalized by the reference dye (gray line) is shifted by the baseline (dashed black line) to remove drift and to get the data around zero (dark blue line). This baseline corrected trace is filtered (light blue line) and the extrema are calculated (circles). The first and the last extrema are discarded (filled red circles) because it is not sure that the real extrema are not earlier or later. Since five extrema remain the trace is marked as oscillating and the time difference between the first and the third extrema is taken as period. The value difference between the second and the third extremum is taken as amplitude (not indicated).

Table S4: Percentage of traces removed from filter in each experiment.

Filter #	sustained	damped	strongly damped
1. (too few data points)	44.24%	87.15%	28.39%
2. (low normalization)	22.65%	0.06%	2.58%
3. (unstable radius trace)	5.08%	2.21%	2.23%
4. (high noise)	9.61%	0.78%	50.46%
5. (acceleration)	0.14%	0.04%	0.30%
6. (bad FFT)	1.40%	0.29%	0.78%
7. (lying under a big one)	4.46%	2.78%	0.89%
passed filtering criteria	12.43%	6.69%	14.39%
5 or more extrema:‘oscillating’	4.99%	3.71%	2.58%

Data preprocessing

Traces were trimmed at the start and at the end, in order to ensure that the first and the last data points had a non-zero droplet area and normalization value. If the normalization or the area time courses had steps (cf. Note on Detection of steps in the time traces) the data was cut before the first occurring step.

Identification of valid data

In order to avoid artifacts from traces wrongly identified as 'oscillating' (false positives), a rigorous data reduction procedure was applied.

1. Traces with not enough data points were discarded (we chose traces with less than 40 points - corresponding to a time window of 200 min - as the cut-off).

2. Traces with zero normalization value at any time point were discarded - this means that the tracked region had to little reference dye and therefor the droplet is not correctly tracked. Example trace in Fig. S4a.

3. We discarded traces for which the droplet radius was not constant, i.e., if the standard deviation of the radius over time was larger than 10% of the mean radius. A change in radius signifies that the area tracked by the tracking algorithm was no droplet but a gap between several droplets. The same criterion also rules out growing or collapsing droplets, which in fact were not observed in our experiments. Example trace in Fig. S4b.

4. Traces with signal to noise ratio less than one were discarded - as the signal level we define the standard deviation of the filtered data, the noise level is the standard deviation of the noise. This criterion potentially also discards oscillating traces with very small amplitudes and genuine droplets with non oscillating behavior. Example trace in Fig. S4c.

5. Traces for which the oscillator apparently speeds up over time were discarded. Empirically and from simulations it is known that the transcriptional oscillator in bulk slows down over time. While speeding-up oscillations are suspicious but could be real, closer investigation revealed that they were artifacts, and the real cause was that they had a very noisy region were normally the second peak lies. They were very different and some could have a peak there but some not and all kind of intermediates. To be sure we only catch good tracked and analyzed traces we removed all of them. This filter was only applied on traces that give a distinct period value (cf. Calculation of periods and amplitudes). Example trace in Fig. S4d.

6. We discarded traces for which the calculated period had no dominant mode in Fourier space. To apply this criterion, the discrete Fourier transform (DFT) of each trace was calculated, and then the next local maximum in the vicinity of the frequency corresponding to the period obtained from the extrema was determined. If the local maximum in the DFT was too small (smaller than 1/3 of the global maximum) or too far from the expected frequency, the trace was discarded.

   This filter is used because if the trace has a relative high noise the extrema detection algorithms sometimes fails. To be sure the algorithm worked fine we double checked with this filter. This filter was only applied on traces that give a distinct period value (cf. Calculation of periods and amplitudes). Example trace in Fig. S4e.

7. Finally, small droplets were excluded that were lying below larger droplets. This selection was accomplished also using heuristic criteria - in particular, such small droplets formed a distinct

population with considerably higher fluorescence values of reference dye than comparably sized ‘normal’ droplets. Example trace in Fig. S4f.

In summary, this rigorous filtering procedure emphasizes the identification of clearly valid droplets and is not expected to distort the statistical statements made in the paper. Notice that after filtering there are still droplets that are not noticeable oscillating. Due to our filter restrictions we are not able to get an information of the real fraction of oscillating vs. non oscillating droplets.

It is possible, though, that many more droplets are oscillating than those used for the analysis, but these droplets either display very slow oscillations (too slow for our observation time window ($T \geq 500$ min)), or very small amplitudes (below our fluorescence detection limit ($SNR < 1$)). Also the smoothing filter produces a 80 min cutoff in period.

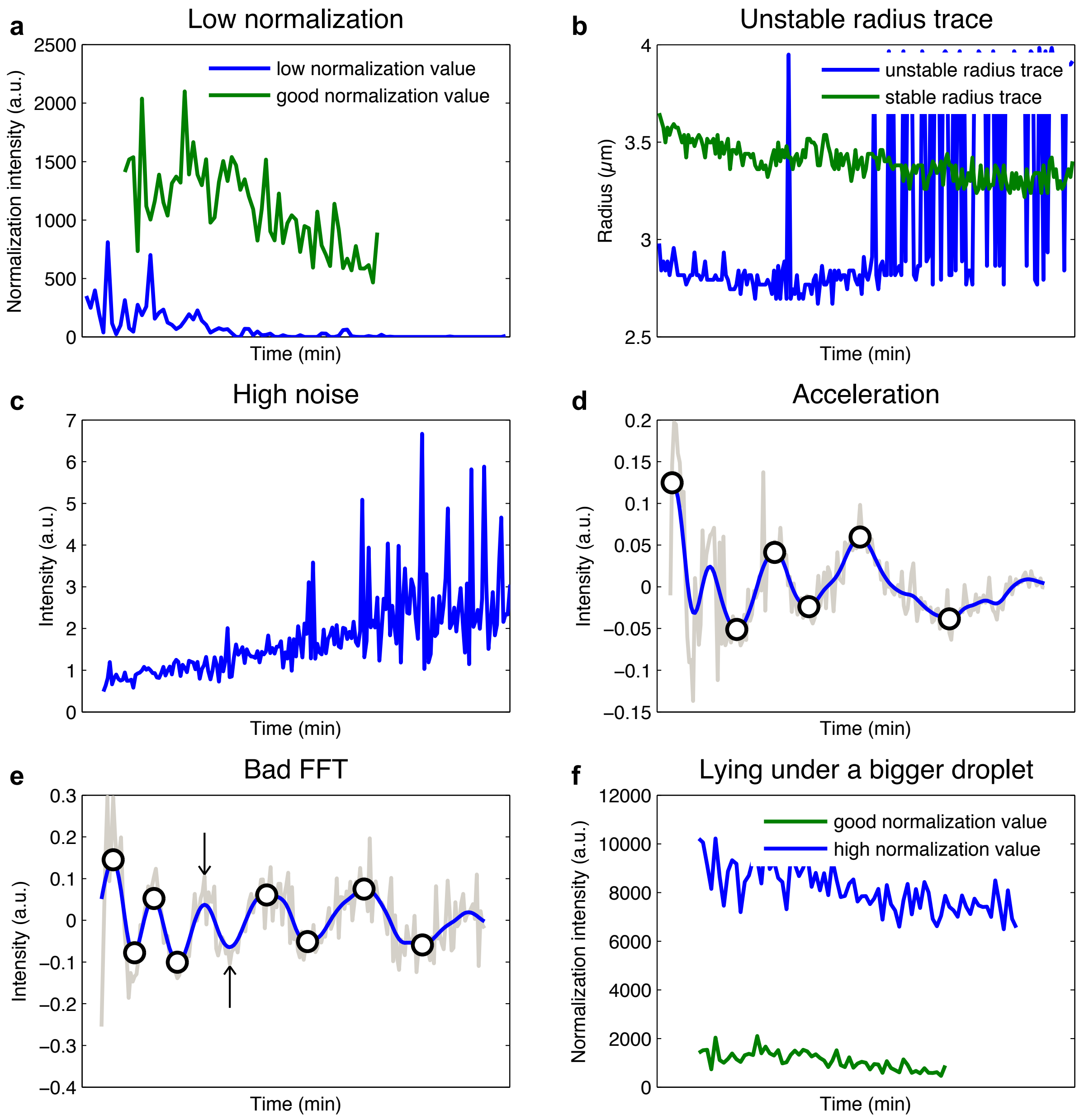

Figure S4: Example traces for each applied filter. Green lines are traces that pass the filter and blue lines are filtered out. **(a)** The filtered trace has a low level that goes down to zero. Although the good trace shows some noise it is good because we want to get rid of this noise in the signal time course with the normalization. **(b)** Radius trace of a tracked are showing strong fluctuations. Since the droplets are quite stable this can not be a well tracked droplet. **(c)** Example trace with very high noise. **(d)** Trace with the first period longer than the second one. **(e)** Failure trace of the extrema detection (missed one period due to relative high noise, arrows). Therefore it has a wrong mode in Fourier space. **(f)** Two traces with similar radii. Since they should show similar normalization value levels the one with the higher mean normalization value must lie under a bigger droplet.

Supplementary Note - Oscillator Data

Experiments for different oscillatory behaviors were performed using the same enzyme batches to maintain constant enzymatic performance in all experiments. Detailed analysis was performed for different oscillator tunings in the following called *sustained*, *damped* and *strongly damped oscillator* exhibiting sustained, damped and strongly damped oscillations. In the following sections, additional analysis results for the three oscillator tunings are shown. Figures S5, S8 and S10 show images of droplet arrays in brightfield mode as well as the corresponding size distributions for droplets which passed filter criteria #1 – #7 as described in Supplementary Note - Data Analysis (Identification of valid data). Figures S6, S9 and S11 show **(a)** example fluorescence traces (normalized to their maximum) for droplets that passed the filtering criteria, including identifiably oscillating traces (blue) and not identifiably oscillating traces (orange), **(b)** population averages of normalized amplitudes and corresponding standard deviations for all traces that passed the filtering criteria, **(c)** example fluorescence traces (normalized to their maximum) for droplets that exhibited identifiable oscillations for at least 900 min, **(d)** population averages of amplitudes and corresponding standard deviations for identifiably oscillating traces, as well as scatter plots of **(e)** amplitudes and **(f)** periods.

The concentrations of DNA strands and enzymes comprising the transcriptional oscillator for the discussed data sets can be found in Table S3.

Sustained oscillator

We tuned the oscillator into its sustained regime by adjusting the enzyme concentrations as stated in Table S3. Example traces for the sustained oscillator compartmentalized with the 'shaken-not-stirred' method are shown in the main text Figure 2. Figure S5 shows an brightfield image of droplets containing the damped oscillator and the size distribution of droplets which passed the filter criteria #1 – #7 described in Supplementary Note - Data Analysis. Figure S6 shows example traces for droplets which passed the filter criteria, the mean amplitude and the corresponding standard deviations as well as additional results on amplitude and period analysis which are not shown in the main text.

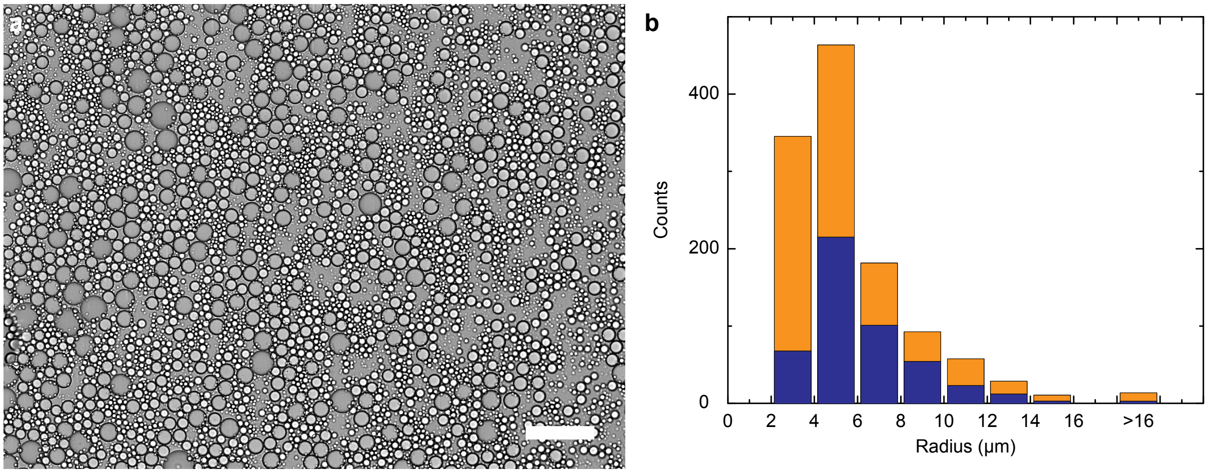

Figure S5: (a) Brightfield image of an array of droplets containing *in vitro* transcriptional oscillator with sustained tuning and **(b)** the corresponding size distribution for droplets which passed filter criteria #1 – #7 as described in Supplementary Note - Data Analysis. The scale bar is 100 µm. The two-colored bars represent the overall number of droplets which passed the filtering criteria. Blue regions indicate the fraction of droplets with identifiably oscillating behavior while orange regions indicate the fraction which was considered to be 'not oscillating' by our filtering procedure.

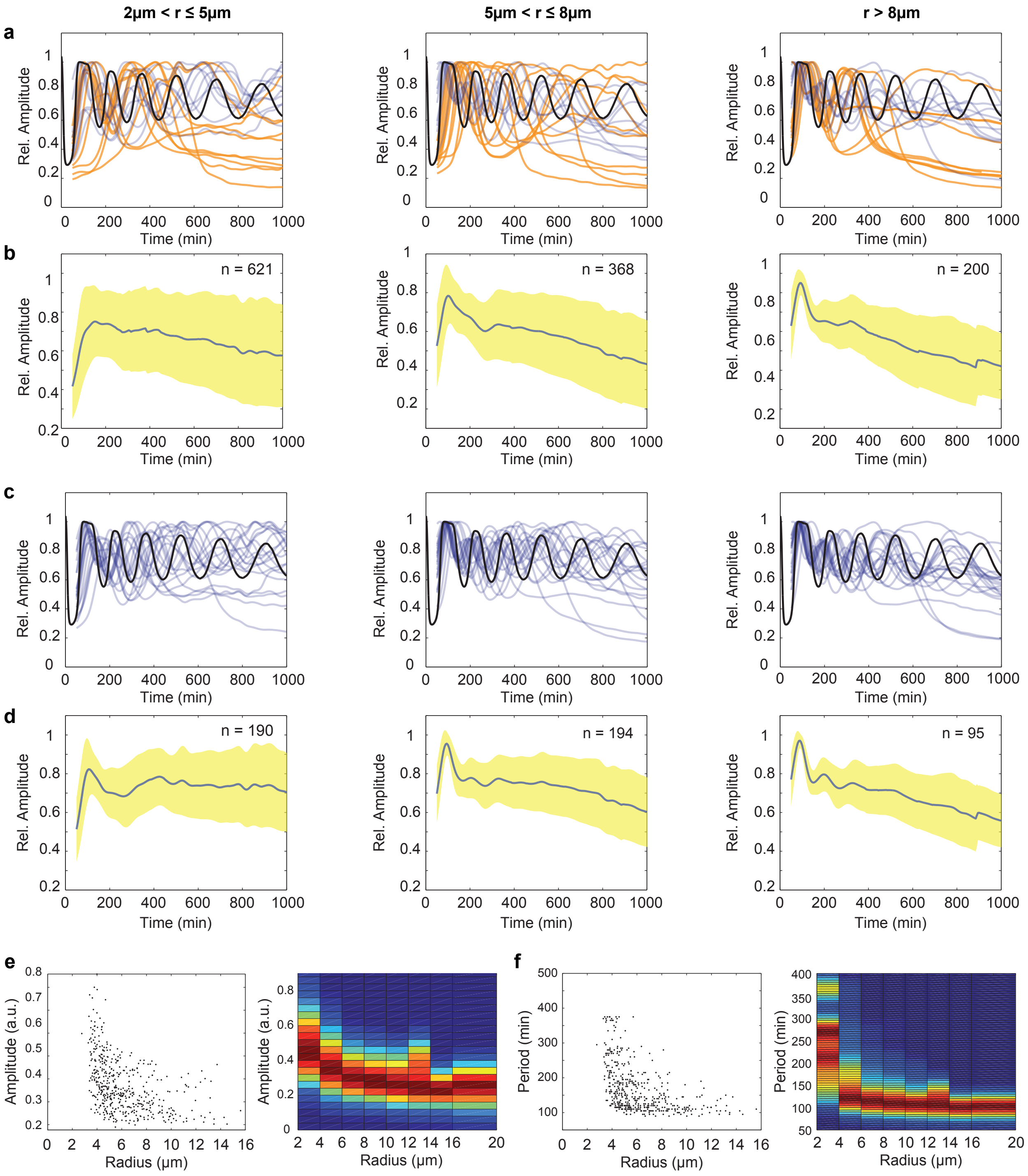

Figure S6: Analysis of the sustained oscillator. See text for description.

Damped oscillator

Further adjustment of the RNase H concentration (see Table S3) directed the oscillators dynamics into the damped regime. Droplets containing transcriptional oscillator with damped tuning are shown in Figure S7. Example traces for droplets which passed the filter criteria #1 - #7 described in Supplementary

Note - Data Analysis are shown in Figure S9a with their corresponding mean amplitude and standard deviation in Figure S9b. To this end, Figures S9c and S9d show example traces and standard deviations only for droplets which were considered to be 'oscillating' by our filtering procedure. Scatter plots for amplitudes and periods are shown in Figures S9e and S9f. A brightfield image of droplets containing the damped transcriptional oscillator and the size distribution of all droplets which passed filter criteria are shown in Figure S8.

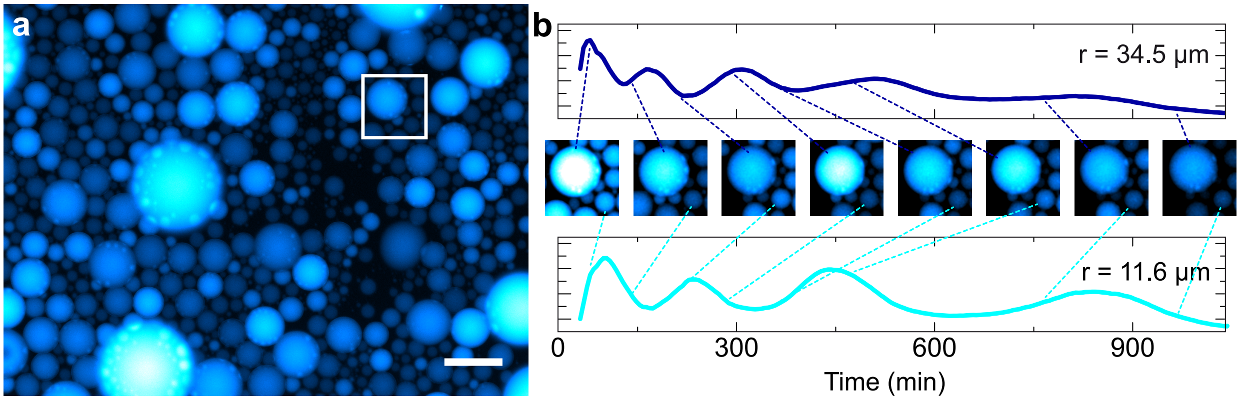

Figure S7: Compartmentalized transcriptional oscillator showing damped behavior. **(a)** Epifluorescence microscopy image of droplets containing damped oscillator. Scale bar is 100 µm. **(b)** Time series of oscillations in droplets for damped oscillator. Droplets are associated with corresponding moments in the fluorescence time trace by dashed lines.

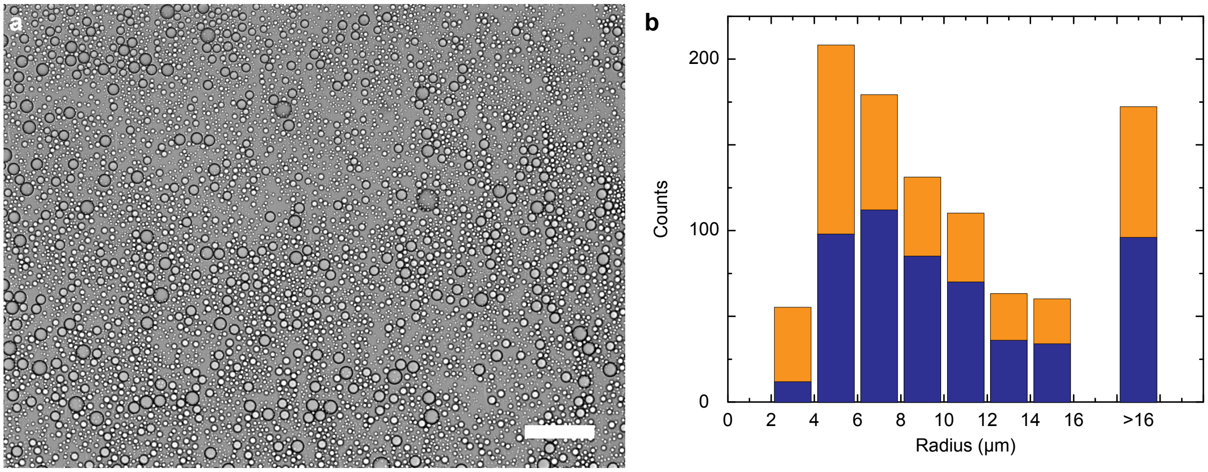

Figure S8: **(a)** Brightfield image of an array of droplets containing *in vitro* transcriptional oscillator with damped tuning and **(b)** the corresponding size distribution for droplets which passed filter criteria #1 – #7 as described in Supplementary Note - Data Analysis. The scale bar is 100 µm. The two-colored bars represent the overall number of droplets which passed the filtering criteria. Blue regions indicate the fraction of droplets with identifiably oscillating behavior while orange regions indicate the fraction which was considered to be 'not oscillating' by our filtering procedure.

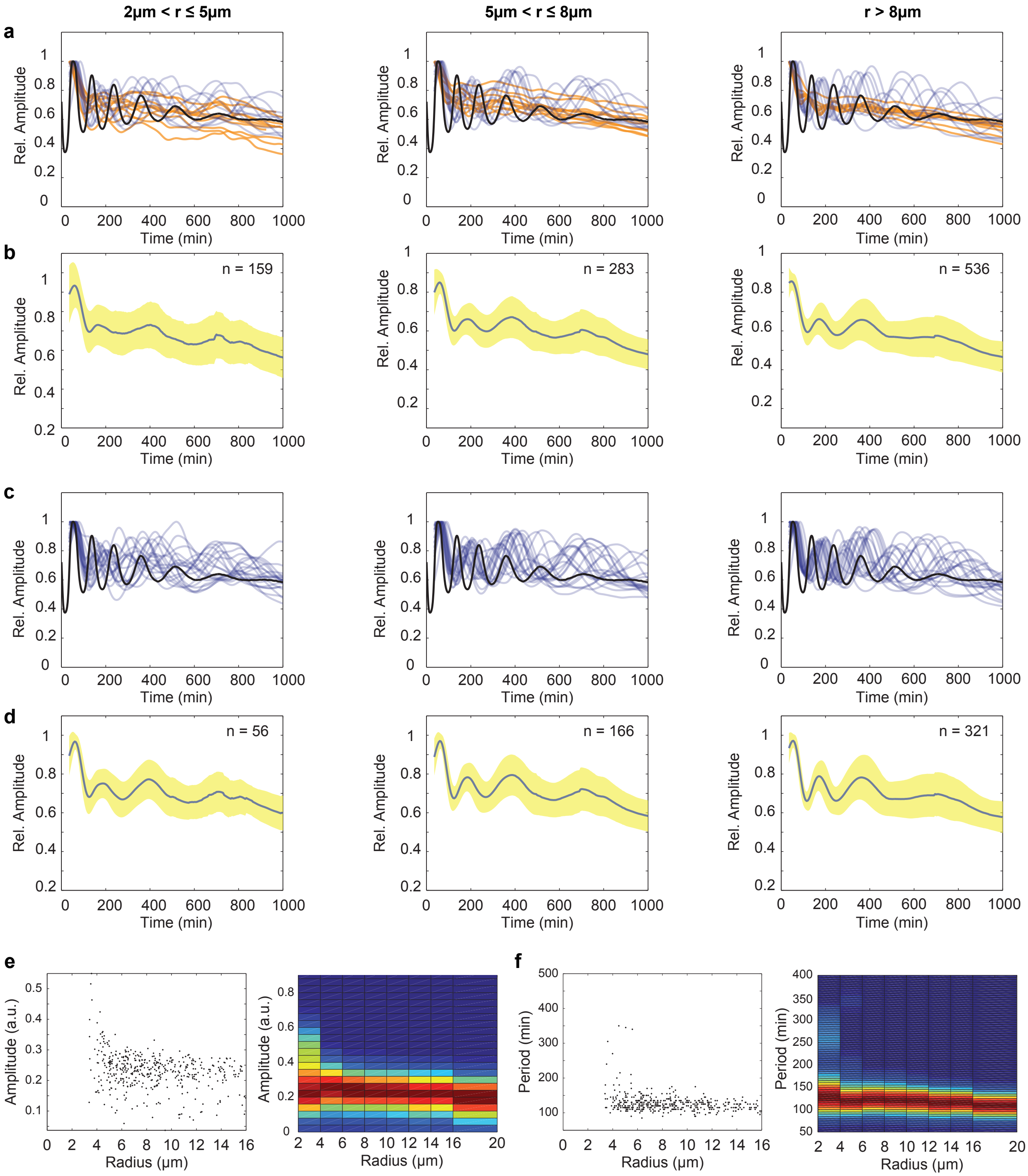

Figure S9: Analysis of the damped oscillator. See text for description.

Strongly damped oscillator

Bulk oscillator showing strongly damped behavior was observed after further variation of the ratio between T7 RNA polymerase and RNase H as well as decrease of total enzyme concentrations. Example traces for a large and a small droplet containing the transcriptional oscillator tuned for strongly damped behavior are shown in the main text in Figure 4. While large droplets follow qualitatively the bulk behavior, the molecular composition comprising the oscillator in the small droplets deviates from the bulk due to 'partitioning error' such that sustained oscillations can be observed for more than 1000 min. This size dependence is nicely illustrated in the size distribution of droplets containing the transcriptional oscillator with strongly damped behavior in Figure S10b. Here, droplets are shown which passed the filter criteria #1 – #7 described in the Supplementary Note on Data Analysis.. Only a small fraction of oscillating droplets (blue) is identified for droplets with radii $< 10\,\mu m$. Example trajectories in Figure S11 highlight that the number of identifiably oscillating traces dramatically decreases with larger radii. No identifiably oscillating traces are detected for radii $> 10\,\mu m$. Scatter plots for amplitudes and periods are shown in Figures S11e and S11f.

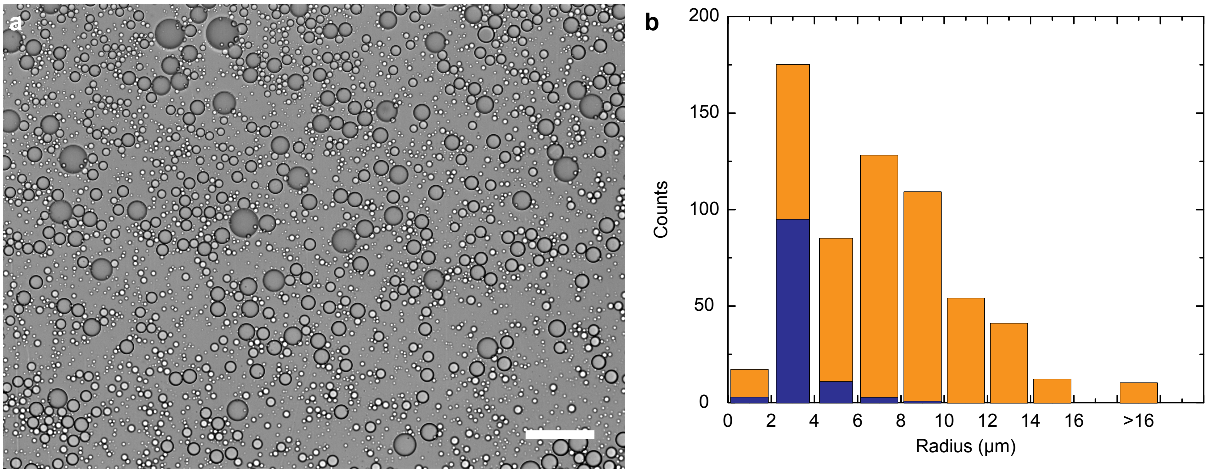

Figure S10: (a) Brightfield image of an array of droplets containing *in vitro* transcriptional oscillator with strongly damped tuning and **(b)** the corresponding size distribution for droplets which passed filter criteria #1 – #7 as described in Supplementary Note - Data Analysis. The scale bar is $100\,\mu m$. The two-colored bars represent the overall number of droplets which passed the filtering criteria. Blue regions indicate the fraction of droplets with identifiably oscillating behavior while orange regions indicate the fraction which was considered to be 'not oscillating' by our filtering procedure.

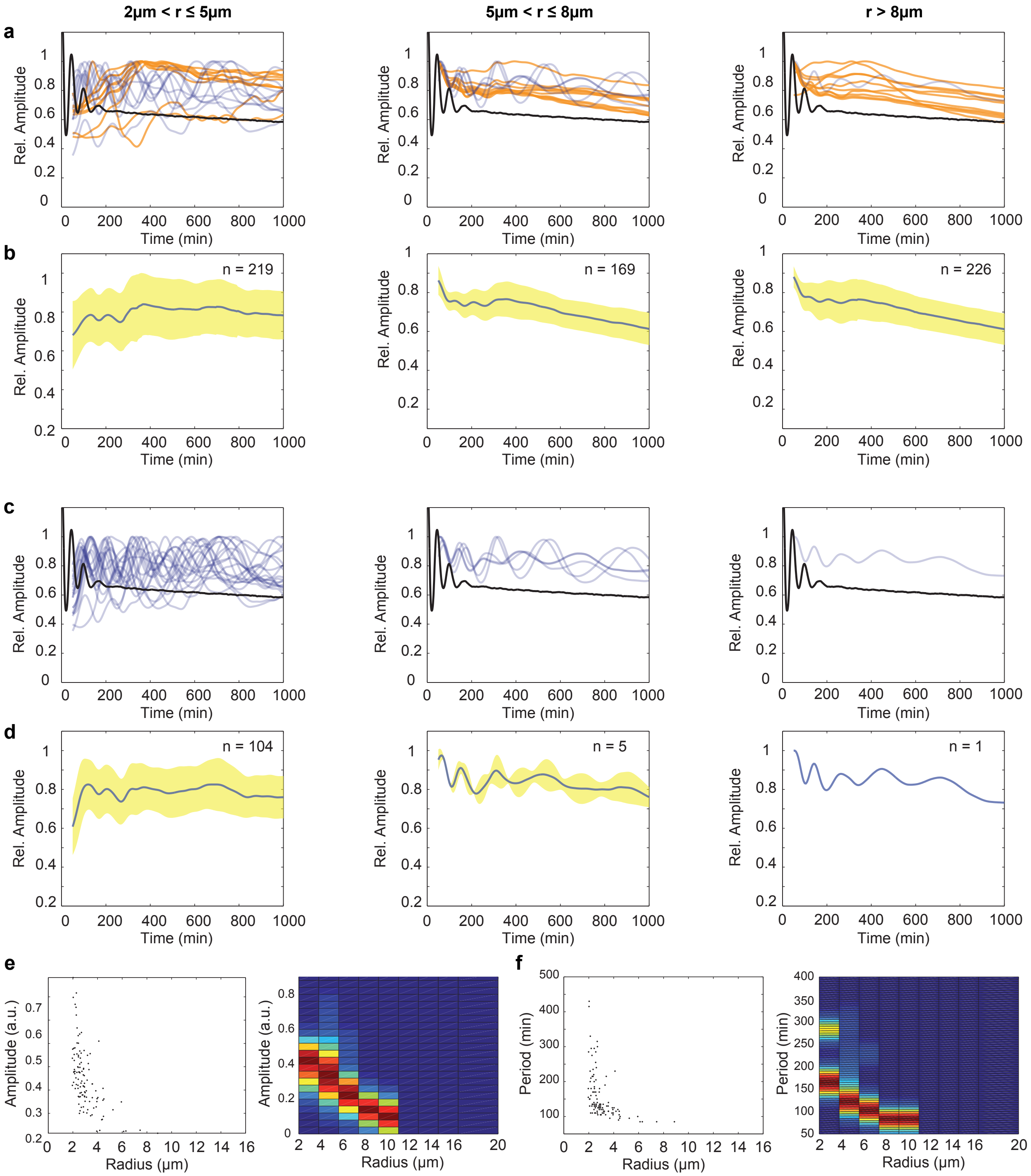

Figure S11: Analysis of the strongly damped oscillator. See text for description. For larger droplet sizes, fewer than 10 droplets showed identifiable oscillations. A vertical offset between droplet traces and the bulk trace exists due to the normalization of each trace to its maximum between $t = 45$ and $t = 1000$ min, i.e. the range in which droplet and bulk data are both present. (Fig. S6 and Fig. S9 are similarly normalized, but in those cases the first peak of the oscillation is strong enough and similar enough that the droplet and bulk traces superimpose more neatly.)

Supplementary Note - Compartmentalization

Distribution and quantification of biomolecules in droplets

To validate measurements of droplet-encapsulated molecules such as labeled DNA and proteins, we performed 'shaken-not-stirred' droplet control experiments using different fluorescent markers.

First, we ran simple experiments to assess the distribution of dyes inside a single droplet. We examined the distributions of 150 nM (bulk) Texas Red-labeled T21 genelet, and 2 µM (bulk) green fluorescent protein (GFP), simultaneously encapsulated in the same emulsion following the procedure previously described in the Supplementary Materials; the two species were suspended in 1X transcription buffer; for this specific experiment, we used an Olympus Apochromat 10x objective. The profile plot of a droplet containing Texas Red labeled DNA switches (T21 off) and GFP are shown in Figure S12a. The overlaying fluorescence cross sections indicate that no molecular species is predominantly attached to the oil-buffer interface and the molecular distribution of both in the compartments is similar. Since the utilized E2K0660 surfactant was invented for biological purposes, and, due to its passive PEG headgroups oriented towards the compartment, we assumed that degradation and adsorption at the inner droplet surface is not a major cause of enzyme loss. We further encapsulated RNAP modified with two different fluorescence labels. Fig. S12 d shows a fluorescence micrograph in 10x magnification of Alexa488-RNAP (a generous gift from Craig T. Martin and Luis E. Ramírez-Tapia of the University of Massachusetts) in microdroplets. Fig. S12 e shows bright field and fluorescence microscopy images in 40x magnification of Atto647-RNAP (generously provided by Evi Stahl from the Dietz lab at TUM) in microemulsion droplets using the same buffer conditions and reaction temperature as in oscillator time-lapse microscopy experiments (see Supplementary Materials). As for the oscillator, the sample was prepared in a reaction tube and mixed on a benchtop vortex before enzymes were added - in this case only RNAP. Thereafter, RNAP was added and the solution was mixed by vigorous pipetting. 10 µL were transferred to an Eppendorf Protein LoBind reaction tube containing 45 µL oil-surfactant-mix and encapsulated into microemulsions. Time-lapse movies reveal the presence of fluorescent aggregates in the droplets in the fluorescence channel corresponding to the RNAP dye label but not in the channel for the emission wavelength of a reference dye which was also present in solution. Microscopy images of bulk solution (without encapsulation) on passivated microscope slides did not show aggregation of RNAP.

Second, we compared the distribution of fluorescence for the different molecules across an ensemble of droplets. We examined the distributions of 200 nM (bulk solution) Alexa488-labeled T7 RNA polymerase, 150 nM (bulk) Texas Red-labeled T21 genelet, and 1 µM (bulk) green fluorescent protein (GFP). These molecules were suspended in 1X transcription mix (1x NEB transcription buffer, 24 mM $MgCl_2$, 7.5 mM each NTP, in nuclease free water), and were individually encapsulated in separate emulsions, following the procedure described in the Supplementary Materials. Brightfield and fluorescence images were recorded over a period of two hours using a NIKON Eclipse TI-E inverted microscope, NIS Elements software, CFI Plan Fluor 10X objective. Droplet fluorescence, area, and perimeter data were collected and processed using the same procedures described for the oscillator traces as outlined in Supplementary Materials (Analysis of time-lapse movies) and Supplementary Note - Data Analysis (however, no reference dye was utilized in these experiments). In addition, fluorescence traces were eliminated if, over the two hour period, 1) the standard deviation of radius measurements exceeded 0.25 µm; 2) the fluorescence standard deviation exceeded 4% the mean, and 3) measurements of perimeter and area of the droplet were inconsistent: specifically, say A is the mean measured area, and P is the mean measured perimeter; then, the radius computed from the perimeter is $R_P = P/2\pi$, with which we can compute an estimated area $A_P = \pi(R_P)^2 = P^2/4\pi$. We discard traces for which the difference $|A - A_P|$ is larger than $0.3A$. (In our experience, the area measurements had lower coefficient of variation than the perimeter, and therefore were used for most calculations; however, particularly inconsistent perimeter measurements suggested mis-identified droplets or corrupted images.)

Figure S13 shows the dependence of detected mean fluorescence intensity on mean droplet radius; fluorescence and radius data for individual droplets were averaged over 2 hours for Texas Red-labeled

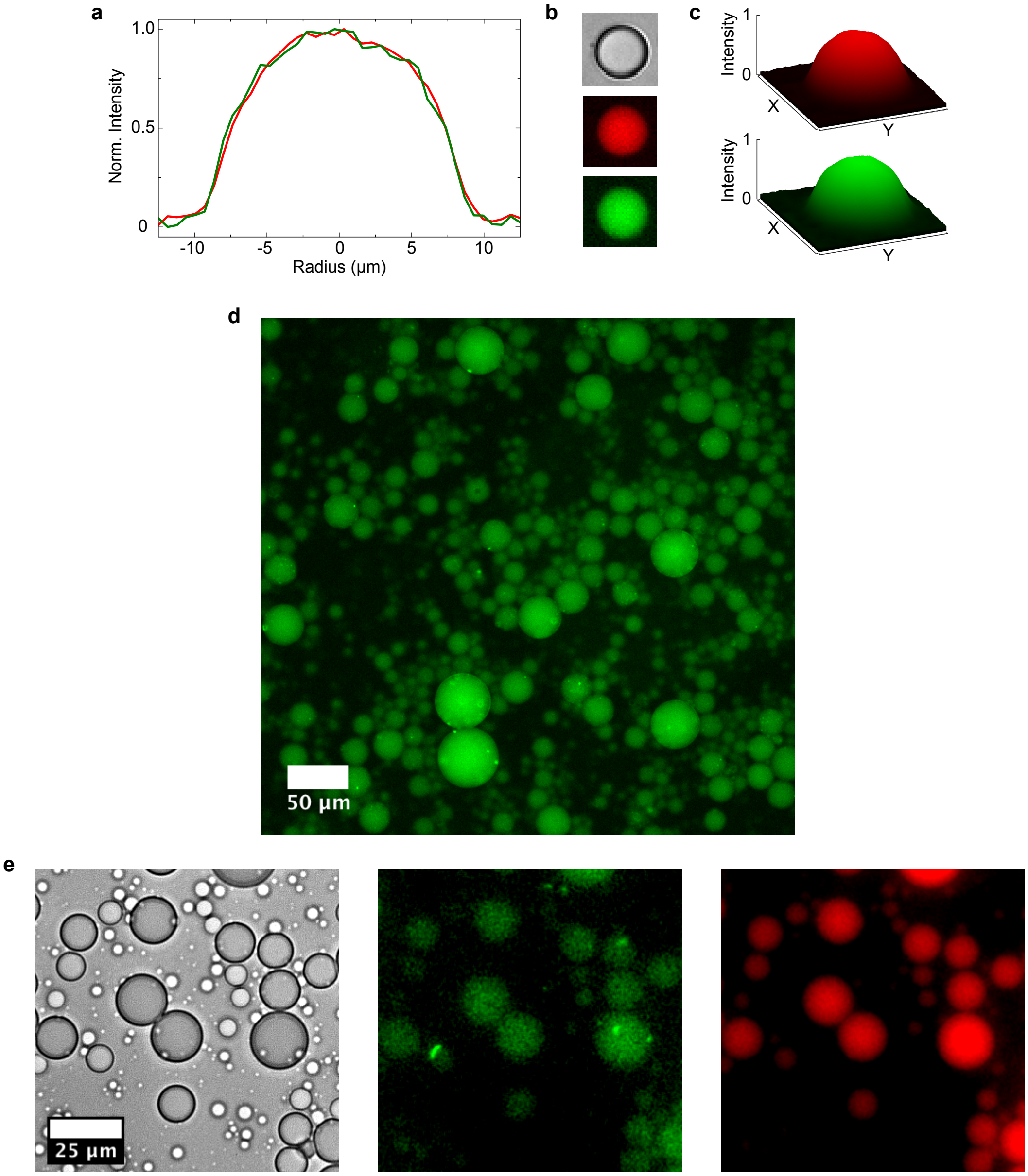

Figure S12: **(a)** Profile plots of normalized droplet fluorescence recorded in the Texas Red (red) and GFP (green) channel. **(b)** Brightfield (gray) and epifluorescence images in Texas Red (red) and GFP (green) channel of an emulsion droplet containing Texas Red labeled genelet T21 and green fluorescent protein. **(c)** Surface plots of droplets shown in (b). **(d)** Fluorescence microscopy image of Alexa488-RNAP in microemulsion droplets in 10x magnification. **(e)** Microscopy images in 40x magnification of Atto647-RNAP and Alexa488 in droplets in brightfield (gray), Atto647 (green) and Alexa488 (red) channel (false colors are used for better comparison with (d)). Thus, aggregation is observed in the two labeled RNAP samples (Alexa488-RNAP and Atto647-RNAP) but not in the control samples containing GFP, DNA template, or pure dye.

T21, and for 10 minutes for Alexa488-labeled RNAP and GFP which were affected by 5% bleaching on average. Using MATLAB, we fitted our mean fluorescence data to the mean droplet radius R with a power law $f(R) = a \cdot R^n$; the numerical fits resulted in n only slightly below a cubic fit. This suggests that the measured fluorescence intensity scales with the droplet volume, and thus with the total number of molecules in a droplet.

Do the deviations from the fit curve reflect partitioning noise? That is, do the higher and lower fluorescence signals for a given radius indicate droplets with higher- or lower-than-average concentrations? This depends on the signal-to-noise of the measurement, which might include artifacts such as background fluorescence, dark noise, and photon shot noise in the microscope and droplet image segmentation noise in the data analysis. To assess the quality of the fluorescence measurements, we used MATLAB to numerically simulate partitioning of our molecules using a Poisson model where we included measurement noise as a source of potential variability. We generated a population of droplets of radii ranging between 0.6 and 20 µm. For each droplet, we generated a random number of molecules according to a Poisson distribution with mean $\lambda = c_0 \cdot N_A \cdot V$, where c_0 is the bulk concentration of the species being partitioned, N_A is Avogadro's number, and V is the volume of the droplet. This random number of molecules was then converted to fluorescence counts using an averaged conversion factor computed as the ratio of the mean expected number of molecules per droplet (computed as $\lambda = c_0 \cdot N_A \cdot V$) and of the mean fluorescence signal over the droplet population. The simulated fluorescence counts were perturbed using zero-mean Gaussian noise using a standard deviation equal to the measured standard error of the mean (SEM) of the fluorescence signals over time. In addition, we also simulated the presence of Gaussian measurement noise of the radius (with standard deviation equal to the measured radius SEM) accompanied by possible binning errors.

Figure S14 shows logarithmic scatter plots of experimental data overlapped with the 10–90% confidence intervals for the simulated "noisy" Poisson partitioned fluorescent molecules (black lines). GFP and Alexa488-labeled RNAP fluorescence time traces exhibit larger measurement noise than Texas Red-labeled T21, and an average 5% bleaching (not observed in Texas Red-labeled T21), which contribute to a larger numerically estimated noise. While addition of simulated measurement noise does result in a distribution that is broader than the ideal Poisson partitioning, it does not entirely explain the fluorescence variability observed in the emulsions. Indeed, to match the observed variability, we would need to add not only the above measurement noise sources, but also Gamma-distributed partitioning noise with β several orders of magnitude larger than that inferred in the main text to explain the oscillator period variability. Consequently, we infer that our fluorescence measurements include uncharacterized noise sources that preclude direct observation of concentration variability due to molecular partitioning processes.

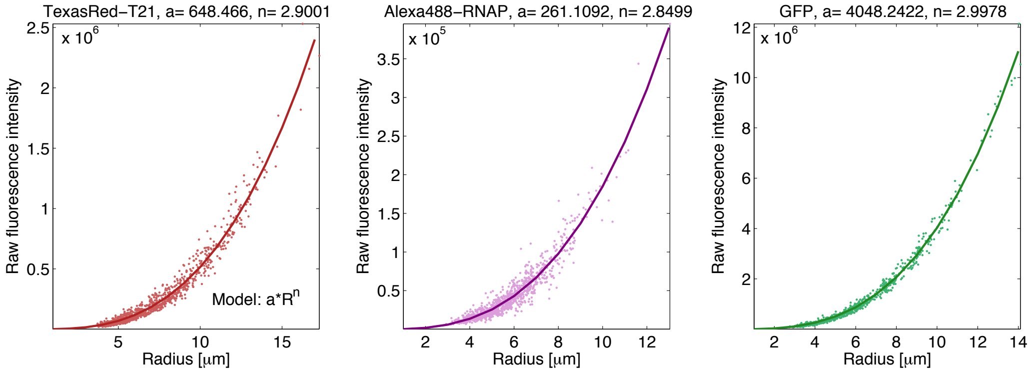

Figure S13: Distribution of TexasRed-labeled switch T21, Alexa488-labeled RNAP, and green fluorescent protein (GFP) in droplets produced by vortexing. Left: Intensity vs. radius for Alexa488-labeled RNAP. Center: Intensity vs. radius for TexasRed. Right: Intensity vs. radius for GFP.

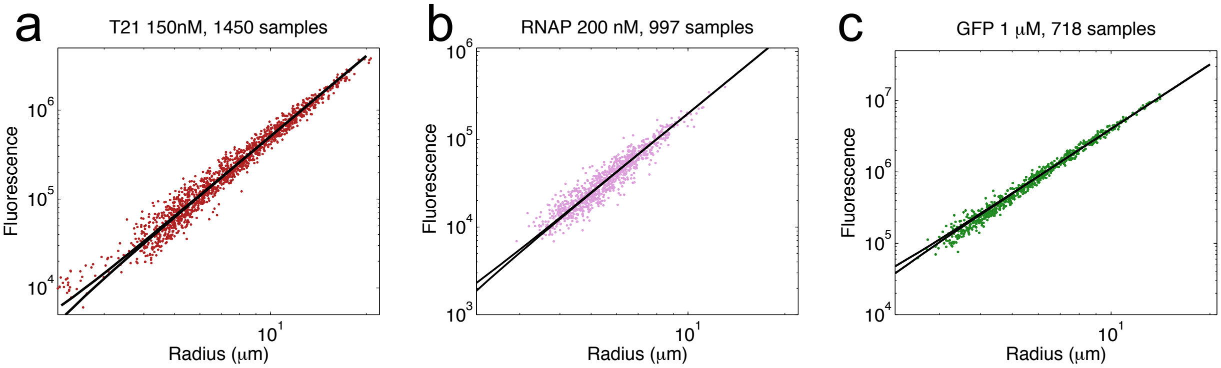

Figure S14: Log-log scatter plots of TexasRed-labeled switch T21, Alexa488-labeled RNAP, and green fluorescent protein (GFP) in droplets produced by vortexing. Dark lines are numerically simulated 10-90% confidence intervals for Poisson-distributed fluorescent molecules affected by zero-mean Gaussian measurement noise and radius binning errors.

Variability in microdroplet-encapsulated single-enzyme subsystems

In order to examine our hypothesis that partitioning in shaken-not-stirred droplets yielded broader-than-Poisson molecular count distributions, and to assess the accuracy and reliability of our experimental and analytical measurement techniques, we performed control experiments with single-enzyme subsystems. Specifically, we compartmentalized simple systems for only transcription or only degradation of RNA signals. Before getting into the details, let us summarize the questions and conclusions of these control experiments. (1) Do we find evidence for broader-than-Poisson molecular partitioning in these single-enzyme subsystems? Both systems do exhibit remarkable variability, which in the case of RNase H cannot be explained by Poisson partitioning and measurement noise. The RNAP subsystem has too many unknowns to allow conclusive attribution. Interestingly, while there is no evidence for enzyme loss in the RNase H subsystem, there is a hint that RNAP activity is reduced in smaller droplets. (2) Given that we can't directly measure partitioning variability from fluorescence signals, does measurement noise also interfere with characterization of the single-enzyme subsystems? Of course it does, but thankfully the temporal dynamics of the single-enzyme systems – such as rise times for RNAP transcription and exponential decay rate constants for RNAse H degradation – allows us to infer parameters that are relatively insensitive to large time-independent additive and multiplicative errors in the fluorescence signal as well as to time-dependent effects such as photon shot noise. (3) Given that the coefficient of variation for measurement noise is also larger for smaller droplets, which is the same signature that we expect for partitioning noise, can we distinguish the two? The simplicity of the RNase H experiments allowed us to quantitatively assess the magnitude of error introduced into estimated degradation rate constants by measurement noise, and while the measurement noise would obscure variability from Poisson partitioning, the experimentally observed variability is well beyond the measurement noise – comparable to Gamma partitioning with β between 10 and 100. (4) How do the noise sources identified in these control experiments affect the central measurements of the main text, that is, the radius-dependent variability of the oscillation period? We identified several noise factors: time-independent additive and multiplicative noise that varies from droplet to droplet (e.g. due to background fluorescence, photodetector dark current, spatial inhomogeneities in lighting or imaging) and which can be corrected for if appropriate baselines can be established; time-varying noise that is correlated between droplets (e.g. due to lamp intensity fluctuations even after normalization to the control dye) and which can be corrected for thanks to the large number of droplets analyzed; time-varying noise that is independent from droplet to droplet (e.g. due to photon shot noise, instrument readout noise, noise in image analysis for radius determination) that can be reduced to some extent by temporal averaging; and of course the biochemical noise due to partitioning, enzyme loss, and stochastic reaction dynamics that we are interested in. While direct measurement of partitioning in static fluorescent samples is precluded by measurement noise, and analysis of single-enzyme experiments is significantly complicated by measurement noise, the inference of oscillation period is robust to constant additive or multiplicative noise and surprisingly insensitive to the level of time-varying noise present in our experiments. Thus, we conclude that the radius-dependent period variability reported in Figure 5 of the main text is minimally affected by measurement noise, although conceivably the amplitude variability is more significantly affected. (5) While it might be considered 'interpretation noise' rather than 'measurement noise' or 'biochemical noise', it should be acknowledged that errors in our estimate of enzyme concentrations (or the active fraction thereof) would yield corresponding errors in our estimates for the degree of variability to be expected from Poisson partitioning. That is to say, 'broader-than-Poisson' distributions could alternatively be explained by errors in the concentrations of active enzymes. However, matching our lower bound, $\beta = 10$, would require a 10-fold error in the concentration, which we consider unlikely.

Schemes describing the two enzyme characterization subsystems are shown in Figure S15, along with sample measurement traces. The RNA transcription system consists of a fully duplex, hence transcriptionally active, template and a preformed reporter consisting of a top strand (labeled with the TAMRA fluorophore at its 5′ end) and a bottom strand (labeled with Black Hole Quencher-2 at its 3′ end). Sequences are given in Table S2. Once the transcription is initiated by RNAP, the RNA transcript displaces the TAMRA-labeled top strand from the reporter complex so that the fluorescence signal increases. The RNA degradation system contains the aforementioned reporter strands mixed with a separately produced RNA strand such that the TAMRA-labeled top strand is initially free in solution. Once the degradation reaction is initiated by RNase H, the RNA strand within RNA-DNA hybrid complex is degraded which allows the fluorophore-labeled top strand to bind to the quencher-labeled bottom strand, resulting in decreased fluorescence. All the other components except for DNA/RNA strands, RNAP and RNase H (i.e., transcriptional buffer components and pyrophosphatase) are chosen as in the oscillator reactions. The experiments were performed in the same manner as reported in the Supplementary Materials. For characterization of the RNA transcription subsystem, the fluorescence traces were analyzed by determining rise time, Δt_{rise}, the time when a fluorescence trace spans 20% to 80% of its maximum intensity value – we used this as a surrogate for the speed of transcription and hence the concentration of enzyme molecules. Using the timepoints from the normalized bulk trajectory, the rise time for bulk trajectory was about 12 min (Fig. S15c).

Since approximately 15 min have passed between the initiation of the reaction in bulk solution and

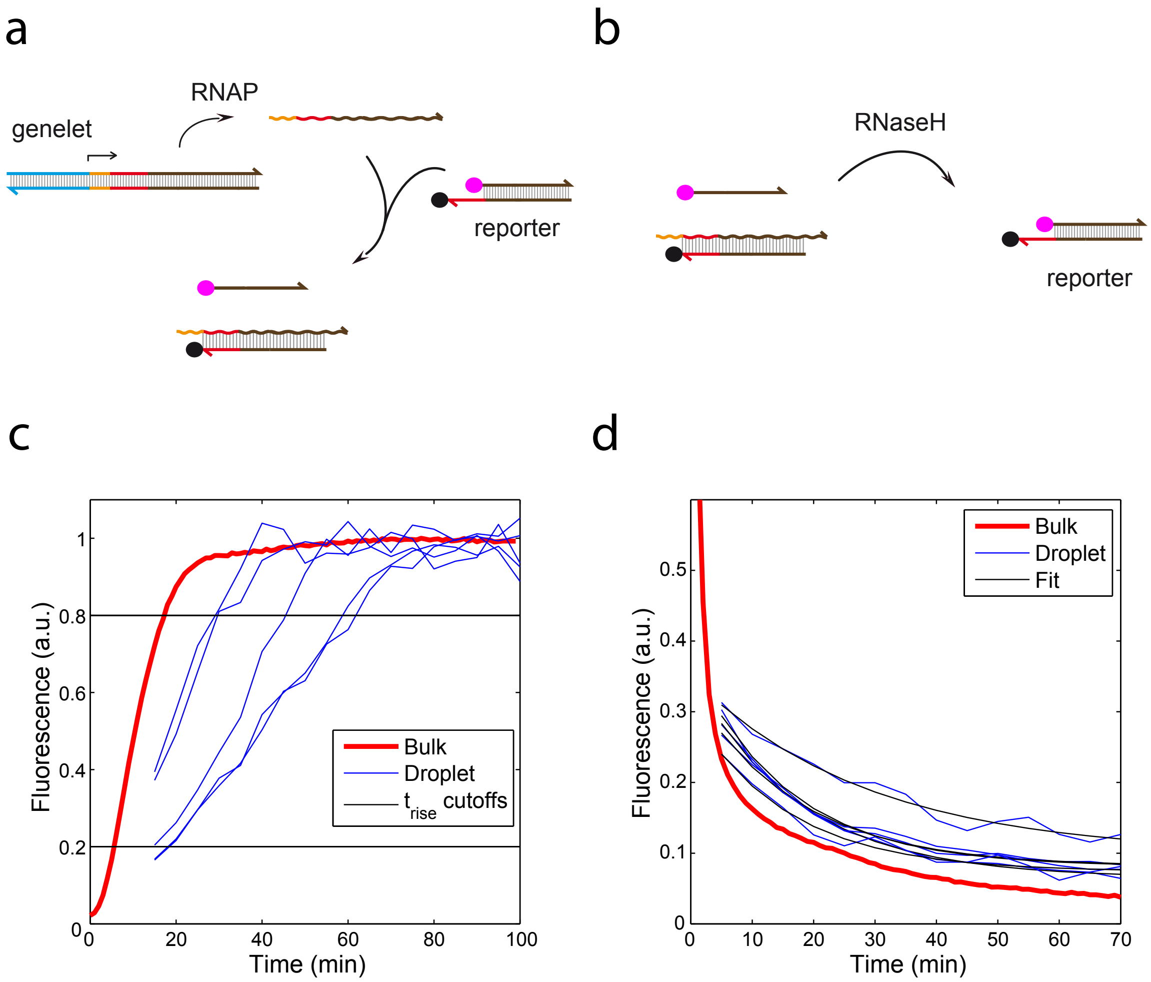

Figure S15: Scheme of fluorescence monitoring of RNA transcription and degradation subsystems. DNA strands are represented by colored lines where colors indicate sequence identity or complementarity; RNA signals are shown as wavelike lines; Black Hole Quencher-2 molecules are black circles; TAMRA fluorophore moieties are pink circles. (a) The RNA signal is transcribed constitutively from a genelet by RNAP; it then displaces a fluorophore-labeled DNA strand from a DNA reporter complex, increasing the fluorescence signal. Nominal concentrations for the experiment are as follows: [Template] = 50 nM, [Reporter] = 400 nM, [RNAP] = 200 nM. (b) Upon degradation of the RNA within the RNA-DNA hybrid complex by RNase H, subsequent binding of the fluorophore-labeled DNA strand and the quencher-labeled strand reduces the fluorescence signal. Nominal concentrations for the experiment are as follows: [RNA] $\approx$ 400 nM, [Reporter] = 400 nM, [RNase H] = 4 nM. (c, d) Bulk trajectories and five example traces from the RNA transcription subsystem (c) and RNA degradation subsystem (d); red lines show bulk trajectories; blue lines show rescaled (but not smoothed) experimental data; black lines show the corresponding rise-time cutoffs and exponential fits.

the beginning of observation under the microscope, the initial 15 min of data for microdroplets are not observed (Fig. S15c). Despite normalization by Alexa dye to control for variations in experimental measurements (as described in the Supplementary Materials), the experimental traces for microdroplets show a large variation in maximum values and presumably in the unobservable minimum values at the beginning. The accurate estimation of maximum and minimum values are critical to determine rise times, and therefore, we decided to perform rescaling and smoothing steps. First, all the raw traces $F_{raw}(t) = I_{TAMRA}(t)/I_{Alexa}(t)$ with more than 60 timepoints spanning 300 min (timepoints are separated by 5 min) were examined: for the 699 traces analyzed, the raw fluorescence did not show any further increase after 200 min, indicating that the reporter was presumably exhausted. Each jth raw trace was normalized by the average value of a 100-min window after 200 min to obtain the rescaled fluorescence $F_{1st,j}(t) = F_{raw,j}(t)/(\sum_{t=200}^{t=300} F_{raw,j}(t)/21)$. After the first normalization step, the average fluorescence of each trace $F_{1st}(t)$ reaches 1.0 at the end of transcription reaction. At this point, a plot of the partially-normalized traces reveals highly variable initial values $F_{1st}(20)$, which could be attributed to droplet-dependent baseline on the order of up to 20%, as well as highly correlated (from droplet-to-droplet) temporal fluctuations in the on-average-constant period from 200 to 300 min, also on the order of up to 20%, which we interpret as lamp intensity fluctuations. Therefore, as a second normalization step, the lamp fluctuations were corrected by dividing the fluorescence trace timepoint-by-timepoint by the median value of the population of droplets that are close to completion: $F_{2nd,j}(t) = F_{1st,j}(t)/(\mathrm{median}_i\,(F_{1st,i}(t) > 0.85))$. Since there were not enough traces that were close to completion for the initial parts of traces ($t = 15, 20...40$ min), we used the median value of the droplet population for $t = 45$ min for those initial parts. As closer examination revealed two distinct subpopulations of “lamp noise” patterns, which presumably correspond to droplets at different locations on the microscope slide, we performed this correction separately for the two subpopulations. Finally, the initial slope was estimated using the first four timepoints, i.e. $F_{2nd}(t = 15, 20, 25, 30)$ – in order to process traces that completely used up reporter quickly, first three timepoints $F_{2nd}(t = 15, 20, 25)$ were used if the initial value $F_{2nd}(t = 15)$ was higher than 0.4 or the initial discrete slope $F_{2nd}(t = 20) - F_{2nd}(t = 15)$ was higher than 0.15. The linear fit was considered acceptable if the relative slope error (calculated as the confidence interval normalized by the fit value, i.e., $E = (S_{1\sigma,max} - S_{1\sigma,min})/S$) was less than 0.35. (During this step, 47 out of 699 traces were excluded; other than especially high noise in the first few timepoints, these traces were indistinguishable from a random sample of the full population.) With reasonable estimates of slope and intercepts thus obtained, the initial timepoints for $F_{2nd}(t = 0, 5, 10)$ were calculated. Using the facts that the signal must be close to 0 prior to addition of RNAP and that the linear fit to the bulk trace crosses 0 around $t = 2$ min, all traces were normalized such that they cross 0 at $t = 2$ min: $F_{final}(t) = (F_{2nd}(t) - F_{2nd}(2))/(1 - F_{2nd}(2))$ with $F_{2nd}(2) = F_{2nd}(0) + 0.4 \times (F_{2nd}(5) - F_{2nd}(0))$.

These rescaled fluorescent traces were smoothed by 5 timepoint averages and plotted in S16a, where grey portions indicate the extrapolated timepoints. (Examples of rescaled but not smoothed traces are shown in S15c.) The rise times Δt_{rise} were calculated using the same procedure as for the bulk trace with the cutoffs of 20% and 80% of the final mean intensity value as indicated in Figure S16a. The scatter plot in Figure S16b shows that the rise times show larger variability for smaller droplets and that the rise times on average decrease for larger droplets. This latter trend, combined with the observation that the bulk rise time is less than all observed droplets, possibly indicates some loss of enzyme activity that is more significant for smaller droplets that presumably endured more splitting events.

The calculated Δt_{rise} values are large for slow production processes and small for fast production processes. Despite several caveats listed below, we attempt to obtain a rough estimate of RNAP activity using the $1/\Delta t_{rise}$ values. For analysis of $1/\Delta t_{rise}$ values, we chose to explore the fit function using the median of these values for radius bins with the largest number of data points, i.e. for those including 3 μm to 7 μm radii. While not a small number of very slow traces were observed (which would be ignored as outliers by the fit to median values), they may be the result of large loss in other

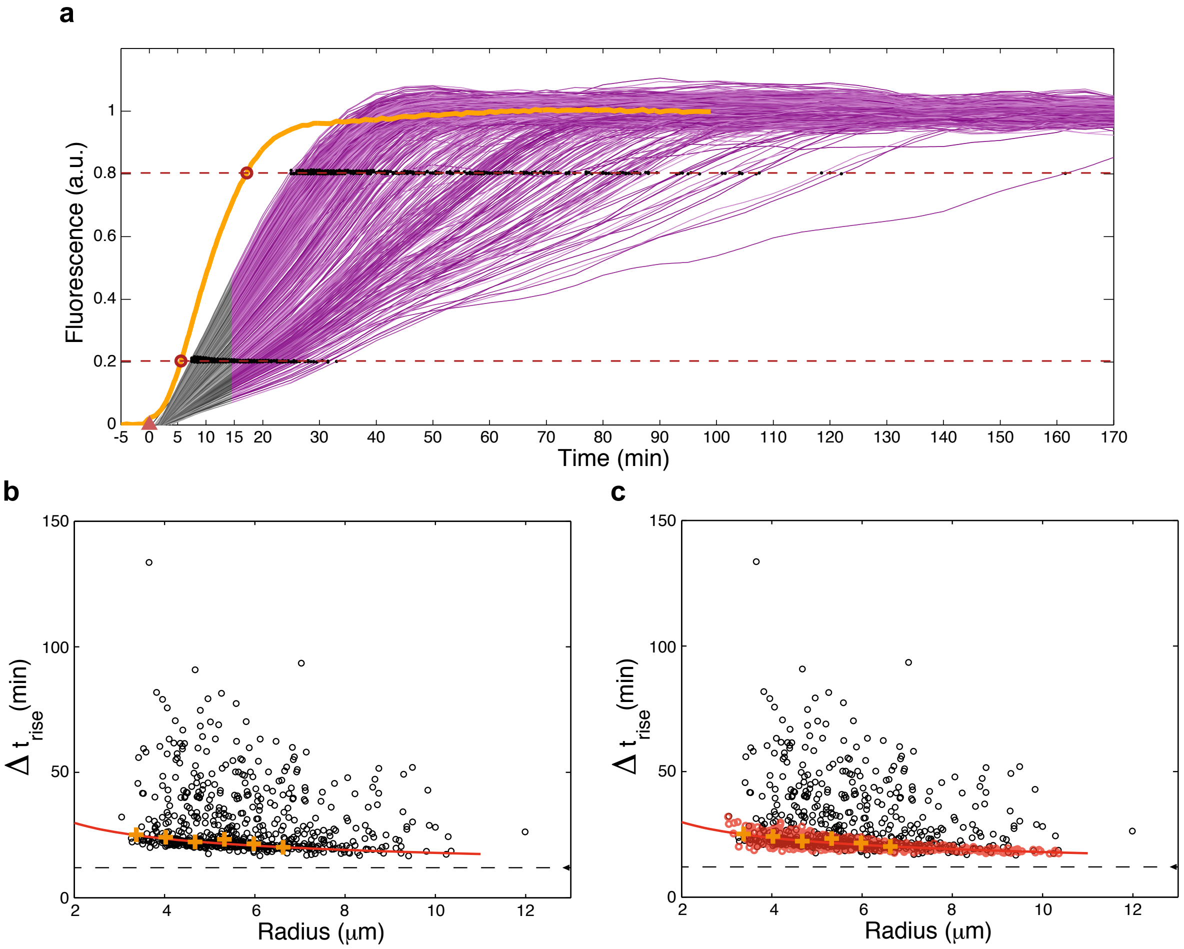

Figure S16: Transcription sub-system in microdroplets. (a) Rise time calculation for bulk and droplet trajectories. Orange line shows the bulk trajectory with pink triangle indicating the timepoint for RNAP addition; purple lines show rescaled and smoothed experimental traces with grey portions indicating extrapolated data points; red dashed lines show cutoffs for rise time calculation. (b) The scatter plot shows rise times Δt_{rise} for individual droplet traces with corresponding radii. The bulk value is indicated by a black dashed line. The median values of the Δt_{rise} data for 0.5 µm radius bins and the fitting curve to these values are shown in orange and red respectively. (c) Same as (b), but containing simulated data (red) sampled according to a Gamma distribution with $\beta = 100$. Note that both distributions are peaked around the median value – only few experimental outliers have very high Δt_{rise} values. Most of the experimental data points as well as the simulated data are close to the median fit.

components of transcription reactions, e.g. the DNA template and pyrophosphatase which are at lower concentration than RNAP. Further, those *very* slow traces would presumably not result in identifiably oscillating traces, in effect being excluded from our analysis of oscillation trajectories. We hypothesize that a loss of enzyme activity proportional to the number of droplet division events is a plausible scenario. Under that assumption, the droplet volume is dependent on the number of division events:

$$V_{droplet} = V_{bulk}/2^N, N = \log_2(V_{bulk}/V_{droplet}) = a - b \cdot \ln(R/\mu\text{m}).$$

If we further assume that the loss of activity (active concentration) is proportional to number of division events , which may be reasonably if disruption is predominantly localized to the point of droplet rupture:

$$c_{RNAP,droplet} = c_{RNAP,bulk} - L \cdot N = (c_{RNAP,bulk} - a \cdot L) + b \cdot L \cdot \ln(R/\mu\text{m})$$

Using the $1/\Delta t_{rise}$ as a proxy for RNAP activity, the resulting fit was as follows:

$$1/\Delta t_{rise} = 0.0278\,\text{min}^{-1} + 0.0107\,\text{min}^{-1} \times \ln(R/\mu\text{m}).$$

For comparison, the bulk Δt_{rise} value was 12 min: $1/\Delta t_{rise,bulk} = 0.083\,\text{min}^{-1}$. The reciprocal of the above fit function is shown in Fig. S16b. In order to test whether this fit adequately captures the variations observed in Δt_{rise}, a distribution sampled according to the Gamma distribution with β at 100 is shown as well (Fig. S16c). We note that alternative scenarios for enzyme loss may well also explain the observed trends in Δt_{rise} values. For instance, the data can also be fit by an RNAP loss proportional to the droplet area (corresponding to a surface effect) with an appropriate global (radius-independent) offset.

Although the trends present in Figure S16b are suggestive, we did not attempt to use the rise times to infer molecule distributions of RNAP quantitatively for several reasons. First, the results are sensitive to the details of multistep data processing that depended on several reasonable but not proven assumptions. It would be difficult to convincingly and quantitatively discriminate between droplet-size-dependent measurement noise and the true biochemical noise that we are interest in. Second, partitioning noise in this subsystem can effect rise times via molecules other than RNAP: partitioning noise in the concentration of reporter (400 nM) is likely to be small, but would introduce error in the first normalization step where a constant reporter maximum is assumed; if there is an excess (say 10 to 20%) of the reporter's bottom strand in the unpurified reporter stock, that would introduce a corresponding delay in the onset of fluorescence increases that would additionally be subject to partitioning noise; the fact that the template concentration was kept low (50 nM) to slow the time scale of the reaction, means that partitioning noise for the template is expected to exceed that of the RNAP itself (at 200 nM); and finally, as noted in Supplementary Note - Modelling (Deterministic Simulation with Partitioning Error), the 'non-essential' enzyme pyrophosphatase is present at a particularly low concentration ($\approx$ 12 nM) and thus particularly subject to partitioning noise – and a low pyrophosphatase concentration could lead to build up of pyrophosphate that inhibits RNAP activity. Consequently, variability in Δt_{rise} cannot be interpreted simply as a measure of partitioning noise in the number of RNAP molecules. However, it is reasonable to interpret the data shown in Figure S16 as evidence that this reaction subsystem is considerably effected by partitioning noise. A clearer case can be made for the RNase H subsystem, below.

For characterization of the RNA degradation subsystem, as shown in Figure S15b, we measured the speed of RNA degradation: we fitted an exponential decay function to the bulk and individual droplet fluorescence time traces:

$$f(t) = B + A \cdot \left(e^{-(t-t_0)/\tau} - 1\right),$$

where B is the baseline, A is the range of exponential curve, $t_0 = 5$ min is the delay between the initiation of degradation and observed traces, and τ is the degradation time constant. While an exponential decay is not the analytical solution for the RNA degradation reaction, it is a reasonable approximation when the RNA-DNA hybrid substrate level falls below the Michaelis constant of RNase H. Starting from the time t_0, which was 5 min after the mixing of enzyme and reporter complex substrates in the bulk, the droplet observation in the microscope was initiated. We chose a 70 min time window, in which the traces could be fit well by an exponential decay function, to determine the measure τ for the enzyme reaction velocity (e.g. Fig. S15d). Figure S17 shows the size distribution of droplets that were analyzed to obtain τ values, as well as the resulting data points for all individual droplets in a scatter plot.

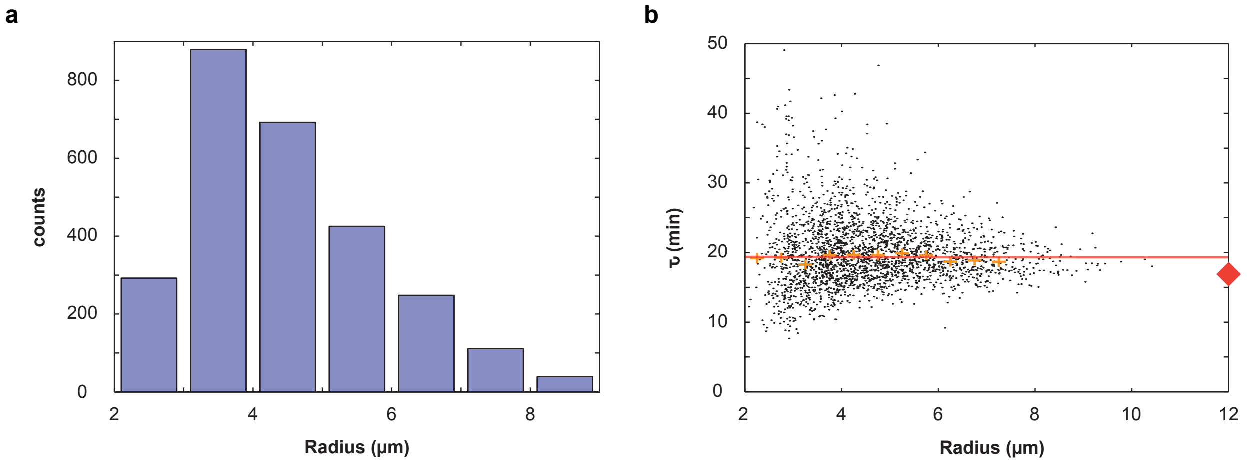

Figure S17: RNA degradation sub-system in microdroplets. (a) Size distribution of analyzed droplets. (b) The scatter plot shows characteristic degradation times τ for individual droplet traces in dependence of the corresponding radii. The median for τ values for 0.5 μm radius bins and the fitting curve to median values are shown in orange and red respectively. The bulk value is indicated by a red diamond. Due to the slower mean τ for droplets compared to the bulk it was assumed that there is a $10\,\%$ global loss of RNase H during encapsulation.

The calculated τ values are large for slow degradation processes and small for fast degradation processes. Therefore, we can obtain a rough estimate of RNase H activity using $1/\tau$ values. Analogous to the analysis for RNAP data, we chose to explore the fit function using median $1/\tau$ values for radius bins. The resulting fit was as follows:

$$1/\tau = 0.0516\,\mathrm{min}^{-1} + 0.0003\,\mathrm{min}^{-1} \times \ln(R/\mu\mathrm{m}).$$

For comparison, the bulk τ value was 16.9 min: $1/\tau_{bulk} = 0.0592\,\mathrm{min}^{-1}$. Unlike the RNAP data, the dependence of median $1/\tau$ values on droplet radius was negligible. This indicates a stronger sensitivity of RNAP with respect to our droplet production process, which seems to be reasonable in light of the fact that T7 RNAP is a much larger (99 kDa) and more complex enzyme than RNase H (17.6 kDa), and is expected to be more prone to denaturation and loss of function. Consequently, we infer that a global loss of RNase H acitivity during the transfer of reaction mixture is a plausible scenario. The reciprocal of the above fit function is shown in Fig. S17b.

As with the experimental results using the RNA transcription subsystem, the observed fluorescence traces in the RNase H subsystem show significant variation in terms of projected baseline and maximum/minimum ranges observed (Fig. S15d). Since measurement noise is also likely to increase for smaller droplets, it is important to establish that the observed variability cannot be explained as predominantly measurement noise, and thus can be safely interpreted as reflecting the biochemical noise, e.g. due to partitioning effects. As a first step, simulation results indicated that the fit value of τ is robust to changes in the baseline and scale of the data: artificially adjusting the baseline and/or scale of experimental traces by several fold and refitting the exponential curve resulted in changes in the baseline parameter (B) and the scale parameter (A) without affecting time constant (τ). Therefore, we conclude that we can obtain meaningful information by comparing τ from experimental trajectories (which may have baseline offsets and variable scale) and simulated trajectories (in which we don't include such effects).

The remaining question is whether timepoint-to-timepoint noise in the experimental data – which indeed is larger (relative to the mean) for smaller droplets – can account for the radius-dependent variability in Figure S17b. We assessed this by performing simulations of the degradation subsystem including biochemical noise from partitioning as well as measurement noise that is comparable to or greater than that which was observed experimentally for the simulated droplet size. The following set of reactions were implemented for simulating the RNA degradation subsystem:

$$\begin{aligned} \mathrm{Rep_F} + \mathrm{Rep_Q} &\xrightarrow{k_{bind}} \mathrm{Rep_F}\cdot\mathrm{Rep_Q} \\ \mathrm{RNaseH} + \mathrm{RNA}\cdot\mathrm{Rep_Q} &\underset{k_{-,H}}{\overset{k_{+,H}}{\rightleftharpoons}} \mathrm{RNaseH}\cdot\mathrm{RNA}\cdot\mathrm{Rep_Q} \xrightarrow{k_{cat,H}} \mathrm{RNaseH} + \mathrm{Rep_Q}, \end{aligned}$$

where k_{bind} was chosen to be $k_{TA,12}$, and parameters for RNase H were chosen as those for substrate A2·rI2, i.e., $k_{+,H}, k_{-,H,2}, k_{cat,H,2}$. Note that we did not fit the simulation results to these experimental traces but use the parameters obtained for oscillator traces. (See Supplementary Note - Modelling (Deterministic Simulation and Model Fits) for further details on parameter determination.) For simplicity, we assumed that the chemical kinetics can be well-approximated by mass action kinetics within the microdroplet-encapsulated RNA degradation subsystem. (Stochastic reactions are considered in Supplementary Note - Modelling (Stochastic Simulation) for oscillator traces, where the deviation from mass action kinetics is shown to be negligible.)

Assuming an independent distribution of molecular species upon formation of microemulsion, the molecule count N is expected to follow a Poisson distribution according to $p(N) = \frac{\lambda^N}{N!}e^{-\lambda}$, where the parameter $\lambda = c \cdot V \cdot N_A$ is the expected number of molecules for a bulk concentration c, a droplet volume V, and Avogadro's constant N_A. After Poisson sampling, the number of molecules within droplets of a given volume can be converted back to concentrations $c_p = N_p/(V \cdot N_A)$. Alternatively, the number of molecules in emulsion droplets N_γ can be sampled according to a gamma distribution – further details are provided in Supplementary Note - Modelling (Deterministic Simulation with Partitioning Error).

To assess the contribution of measurement noise (including but not limited to photon shot noise and droplet area mis-estimation), we included a volume-dependent measurement noise term for each time point in the simulation trajectories. First, we generated a nominal simulation trajectory without measurement noise $F_{sim}(t) = [\mathrm{Rep_F}](t)/[\mathrm{Rep_F^{tot}}]$ for the given droplet volume, the concentration of the molecular species, and the mode of partitioning investigated, i.e., no partitioning noise, partitioning with Poisson distribution, and partitioning with gamma distribution. Without partitioning noise, the nominal simulation trajectory matches the bulk simulation result. To determine the extent of measurement noise to be added, the number of fluorophore molecules is chosen as a measure for fluorescence counts observed – the smaller the number of fluorophore molecules, the smaller the fluorescence counts observed, increasing measurement errors. Thus, the trajectory with measurement noise at each timepoint is generated by sampling with the same mean number of $\mathrm{Rep_F}$ molecules as that of nominal trajectory but with variance 30 times higher than the number of $\mathrm{Rep_F}$ molecules of nominal trajectory. Specifically, the number of $\mathrm{Rep_F}$ molecules at time t equals $\mathrm{Rep_{F,count}}(t) = [\mathrm{Rep_F}](t) \cdot V \cdot N_A$, and thus, F_{noise} can be obtained by sampling accordingly, converting back to concentration of $\mathrm{Rep_F}$, and normalizing against the total concentration of $\mathrm{Rep_F}$:

$$F_{noise}(t) = \frac{\mathcal{N}(\mathrm{Rep_{F,count}}(t), 30 \cdot \mathrm{Rep_{F,count}}(t))}{V \cdot N_A \cdot [\mathrm{Rep_F^{tot}}]}.$$

Example simulation traces with only measurement noise are shown in Figure S18d; example traces with both partitioning and measurement noise are shown in Figure S18g. The extent of measurement noise was quantified by the least-squared-error between the exponential fit and the experimental traces (Fig. S18c) or the simulated traces without and with partitioning noise (Fig. S18f,i). The observed least-squared-error for simulated traces exceed that of experimental traces in many instances (the proportionality constant of variance for sampling F_{noise} was heuristically chosen for this), but are not significantly affected by partitioning noise. The distribution of τ values are much narrower without partitioning noise (Fig. S18e,h), suggesting that the measurement noise alone is not enough to explain the observed variability in τ.

The comparison of the τ distributions for experimental and simulated traces are shown as scatter plots in Figure S19, indicating that a wider-than-Poisson distribution of constituent molecules upon microdroplet formation is a plausible inference from the analysis of RNA degradation experiments. More specifically, simulated traces using partitioning noise following a phenomenological Gamma distribution with parameter β chosen between 10 and 100 appear to match the width of experimental distribution. Also note that there is no strong evidence for droplet-formation-dependent loss of RNase H enzyme activity during the 'shaken-not-stirred' emulsification procedure: the mean of the τ distribution has insignificant dependence on the radius and agrees reasonably well (to within about 10%) with the bulk value, consistent with a uniform loss prior to droplet formation.

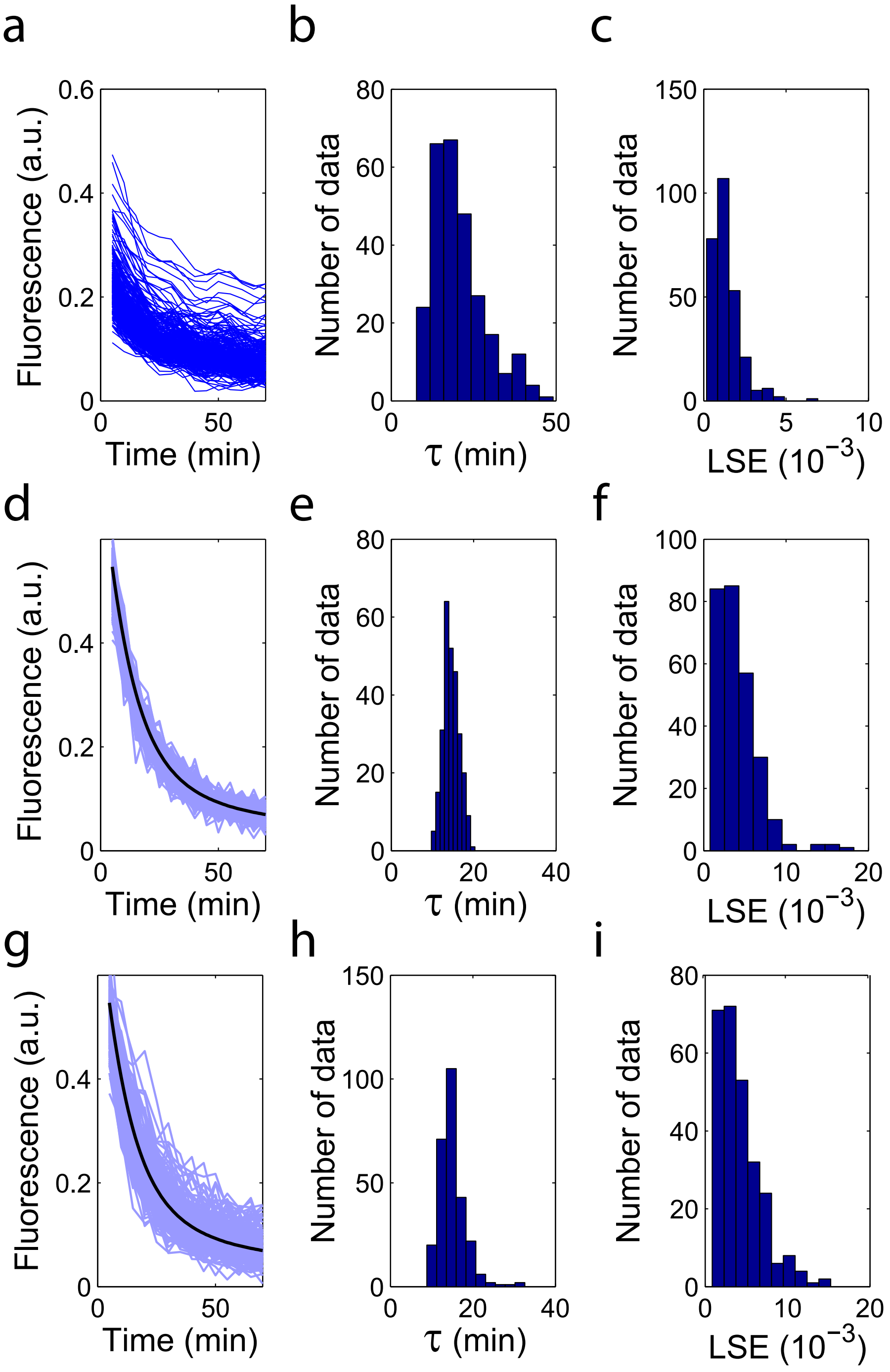

Figure S18: Comparison of RNase H subsystem experimental and simulation results for the 2 μm radius bin. (a-c) Experimental results: (a) normalized trajectories, (b) distribution of τ, (c) distribution of the least-squared-error between the fitted exponential trajectories and the measured trajectories. (d-f) Simulation results with only the 'measurement noise' term added, and no partitioning noise: (d) normalized trajectories, (e) distribution of τ, (f) distribution of the least-squared-error. (g-i) Simulation results using gamma partitioning ($\beta = 10$) and the measurement noise term: (g) normalized trajectories, (h) distribution of τ, (i) distribution of least-squared-error. Black lines in (d) and (g) are traces for the simulated bulk experiment.

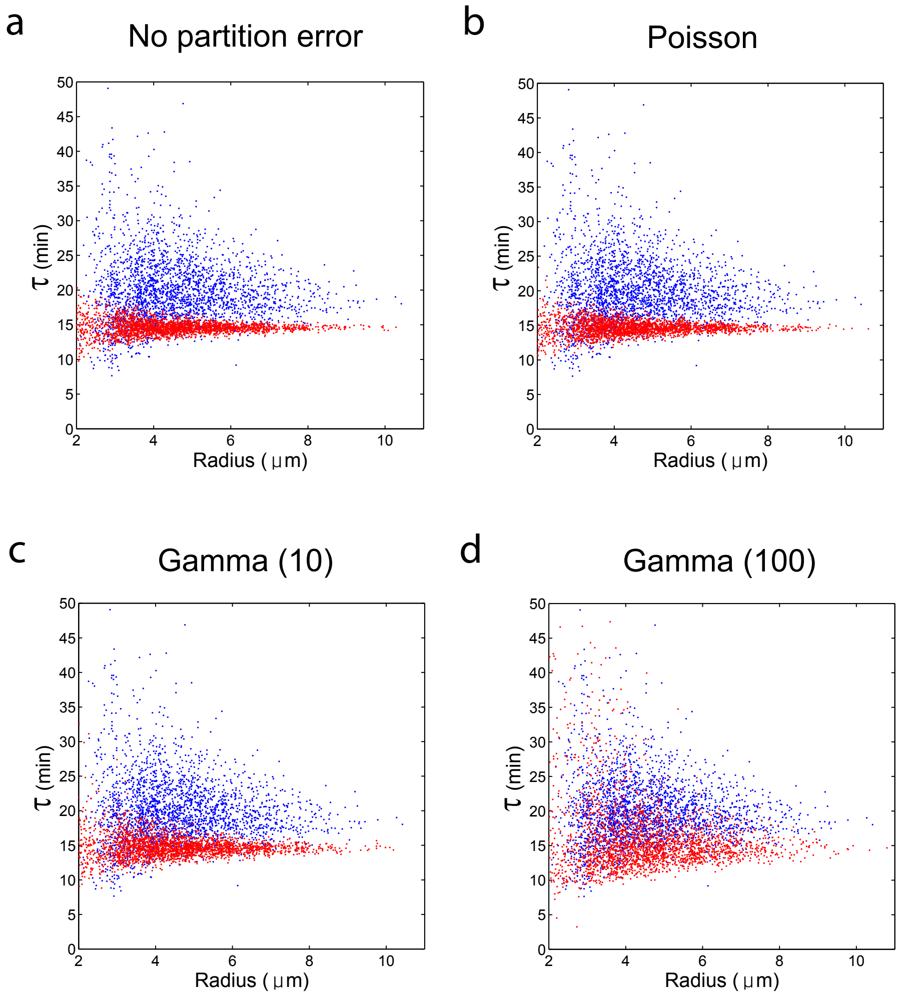

Figure S19: Distribution of τ for RNase H subsystem experiments compared to simulations with different partitioning variation of molecules during the droplet formation, without loss: (a) no partitioning error, (b) Poisson partitioning, (c) gamma partitioning with $\beta = 10$, (d) gamma partitioning with $\beta = 100$. The τ values for experiment and simulation are indicated by blue and red dots, respectively.

Supplementary Note - Modelling

Model Equations

A set of DNA and RNA hybridization reactions, branch migration reactions, and Michaelis-Menten enzyme reactions for the oscillator is as shown below. This set of reaction equations is identical to those used in the extended model of the Design I oscillator in [Kim and Winfree, 2011]. However, the enzyme species and enzyme complexes were treated as separate states in the system rather than relying on the Michaelis-Menten approximation as in prior work [Kim and Winfree, 2011, Franco et al, 2011]; this made stochastic simulation straightforward. For simulations of bulk reactions and non-stochastic simulations of droplets, standard mass action kinetics was used to convert these reaction equations to a set of ordinary differential equations as outlined in [Kim and Winfree, 2011]. Figure S20 shows the schematic diagram for reaction equations, excluding the reactions that involve incomplete degradation products (figure adapted from [Kim and Winfree, 2011]).

DNA/RNA hybridization and branch migration reactions

$$\begin{array}{rcll}
\mathrm{T12} + \mathrm{A2} & \xrightarrow{k_{TA,12}} & \mathrm{T12}\cdot\mathrm{A2} & \text{(Activation)}\\
\mathrm{T21} + \mathrm{A1} & \xrightarrow{k_{TA,21}} & \mathrm{T21}\cdot\mathrm{A1} & \text{(Activation)}\\
\mathrm{A1} + \mathrm{dI1} & \xrightarrow{k_{AI,1}} & \mathrm{A1}\cdot\mathrm{dI1} & \text{(Annihilation)}\\
\mathrm{rA1} + \mathrm{dI1} & \xrightarrow{k_{rAI,1}} & \mathrm{rA1}\cdot\mathrm{dI1} & \text{(Annihilation)}\\
\mathrm{A2} + \mathrm{rI2} & \xrightarrow{k_{AI,2}} & \mathrm{A2}\cdot\mathrm{rI2} & \text{(Annihilation)}\\
\mathrm{T12}\cdot\mathrm{A2} + \mathrm{rI2} & \xrightarrow{k_{TAI,12}} & \mathrm{T12} + \mathrm{A2}\cdot\mathrm{rI2} & \text{(Inhibition)}\\
\mathrm{T21}\cdot\mathrm{A1} + \mathrm{dI1} & \xrightarrow{k_{TAI,21}} & \mathrm{T21} + \mathrm{A1}\cdot\mathrm{dI1} & \text{(Inhibition)}\\
\mathrm{rA1} + \mathrm{A1}\cdot\mathrm{dI1} & \xrightarrow{k_{AIrA,1}} & \mathrm{rA1}\cdot\mathrm{dI1} + \mathrm{A1} & \text{(Release)}
\end{array}$$

Michaelis-Menten enzyme reactions

$$\begin{array}{rcccl}
\mathrm{RNAP} + \mathrm{T12}\cdot\mathrm{A2} & \underset{k_{-,ON,12}}{\overset{k_{+}}{\rightleftharpoons}} & \mathrm{RNAP}\cdot\mathrm{T12}\cdot\mathrm{A2} & \xrightarrow{k_{cat,ON,12}} & \mathrm{RNAP} + \mathrm{T12}\cdot\mathrm{A2} + \mathrm{rA1}\\
\mathrm{RNAP} + \mathrm{T21}\cdot\mathrm{A1} & \underset{k_{-,ON,21}}{\overset{k_{+}}{\rightleftharpoons}} & \mathrm{RNAP}\cdot\mathrm{T21}\cdot\mathrm{A1} & \xrightarrow{k_{cat,ON,21}} & \mathrm{RNAP} + \mathrm{T21}\cdot\mathrm{A1} + \mathrm{rI2}\\
\mathrm{RNAP} + \mathrm{T12} & \underset{k_{-,OFF,12}}{\overset{k_{+}}{\rightleftharpoons}} & \mathrm{RNAP}\cdot\mathrm{T12} & \xrightarrow{k_{cat,OFF,12}} & \mathrm{RNAP} + \mathrm{T12} + \mathrm{rA1}\\
\mathrm{RNAP} + \mathrm{T21} & \underset{k_{-,OFF,21}}{\overset{k_{+}}{\rightleftharpoons}} & \mathrm{RNAP}\cdot\mathrm{T21} & \xrightarrow{k_{cat,OFF,21}} & \mathrm{RNAP} + \mathrm{T21} + \mathrm{rI2}\\
\mathrm{RNaseH} + \mathrm{rA1}\cdot\mathrm{dI1} & \underset{k_{-,H,1}}{\overset{k_{+,H}}{\rightleftharpoons}} & \mathrm{RNaseH}\cdot\mathrm{rA1}\cdot\mathrm{dI1} & \xrightarrow{k_{cat,H,1}} & \mathrm{RNaseH} + \mathrm{dI1}
\end{array}$$

Reactions involving incomplete degradation product

$$\begin{array}{rcccl}
\mathrm{RNaseH} + \mathrm{A2}\cdot\mathrm{rI2} & \underset{k_{-,H,2}}{\overset{k_{+,H}}{\rightleftharpoons}} & \mathrm{RNaseH}\cdot\mathrm{A2}\cdot\mathrm{rI2} & \xrightarrow{k_{cat,H,2}} & \mathrm{RNaseH} + \mathrm{A2}\cdot\mathrm{sI2}\\
\mathrm{RNAP} + \mathrm{T12}\cdot\mathrm{A2}\cdot\mathrm{sI2} & \underset{k_{-,ON,12}}{\overset{k_{+}}{\rightleftharpoons}} & \mathrm{RNAP}\cdot\mathrm{T12}\cdot\mathrm{A2}\cdot\mathrm{sI2} & \xrightarrow{k_{cat,ON,12}} & \mathrm{RNAP} + \mathrm{T12}\cdot\mathrm{A2}\cdot\mathrm{sI2} + \mathrm{rA1}\\
\mathrm{A2} + \mathrm{sI2} & \underset{k_{s,-}}{\overset{k_{s,+}}{\rightleftharpoons}} & \mathrm{A2}\cdot\mathrm{sI2} & & \\
\mathrm{T12} + \mathrm{A2}\cdot\mathrm{sI2} & \xrightarrow{k_{TA,12}} & \mathrm{T12}\cdot\mathrm{A2}\cdot\mathrm{sI2} & & \\
\mathrm{T12}\cdot\mathrm{A2} + \mathrm{sI2} & \underset{k_{s,-}}{\overset{k_{s,+}}{\rightleftharpoons}} & \mathrm{T12}\cdot\mathrm{A2}\cdot\mathrm{sI2} & & \\
\mathrm{rI2} + \mathrm{A2}\cdot\mathrm{sI2} & \xrightarrow{k_{AI,2}} & \mathrm{A2}\cdot\mathrm{rI2} + \mathrm{sI2} & &
\end{array}$$

Thus, the dynamics of the synthetic oscillator is described by the following seventeen ordinary differential equations:

$$
\begin{aligned}
\frac{d[\mathrm{T21}]}{dt} &= -k_{TA,21}[\mathrm{T21}][\mathrm{A1}] + k_{TAI,21}[\mathrm{T21A1}][\mathrm{dI1}] - k_{+}[\mathrm{RNAP}][\mathrm{T21}] + (k_{-,OFF,21} + k_{cat,OFF,21})[\mathrm{RNAPT21}],\\
\frac{d[\mathrm{A1}]}{dt} &= -k_{AI,1}[\mathrm{A1}][\mathrm{dI1}] - k_{TA,21}[\mathrm{T21}][\mathrm{A1}] + k_{AIrA,1}[\mathrm{A1dI1}][\mathrm{rA1}],\\
\frac{d[\mathrm{dI1}]}{dt} &= -k_{AI,1}[\mathrm{A1}][\mathrm{dI1}] - k_{rAI,1}[\mathrm{rA1}][\mathrm{dI1}] - k_{TAI,21}[\mathrm{T21A1}][\mathrm{dI1}] + k_{cat,H,1}[\mathrm{RNaseHrA1dI1}],\\
\frac{d[\mathrm{rA1}]}{dt} &= -k_{rAI,1}[\mathrm{rA1}][\mathrm{dI1}] - k_{AIrA,1}[\mathrm{A1dI1}][\mathrm{rA1}] + k_{cat,ON,12}([\mathrm{RNAPT12A2}] + [\mathrm{RNAPT12A2sI2}])\\
&\quad + k_{cat,OFF,12}[\mathrm{RNAPT12}],\\
\frac{d[\mathrm{RNAPT21A1}]}{dt} &= k_{+}[\mathrm{RNAP}][\mathrm{T21A1}] - (k_{-,ON,21} + k_{cat,ON,21})[\mathrm{RNAPT21A1}],\\
\frac{d[\mathrm{RNAPT21}]}{dt} &= k_{+}[\mathrm{RNAP}][\mathrm{T21}] - (k_{-,OFF,21} + k_{cat,OFF,21})[\mathrm{RNAPT21}],\\
\frac{d[\mathrm{RNaseHrA1dI1}]}{dt} &= k_{+,H}[\mathrm{RNaseH}][\mathrm{rA1dI1}] - (k_{-,H,1} + k_{cat,H,1})[\mathrm{RNaseHrA1dI1}],\\
\frac{d[\mathrm{T12}]}{dt} &= -k_{TA,12}[\mathrm{T12}][\mathrm{A2}] + k_{TAI,12}[\mathrm{T12A2}][\mathrm{rI2}] - k_{TA,12}[\mathrm{T12}][\mathrm{A2sI2}] - k_{+}[\mathrm{RNAP}][\mathrm{T12}]\\
&\quad + (k_{-,OFF,12} + k_{cat,OFF,12})[\mathrm{RNAPT12}],\\
\frac{d[\mathrm{A2}]}{dt} &= -k_{AI,2}[\mathrm{A2}][\mathrm{rI2}] - k_{TA,12}[\mathrm{T12}][\mathrm{A2}] - k_{s,+}[\mathrm{A2}][\mathrm{sI2}] + k_{s,-}[\mathrm{A2sI2}],\\
\frac{d[\mathrm{rI2}]}{dt} &= -k_{AI,2}[\mathrm{A2}][\mathrm{rI2}] - k_{AI,2}[\mathrm{A2sI2}][\mathrm{rI2}] - k_{TAI,12}[\mathrm{T12A2}][\mathrm{rI2}] + k_{cat,ON,21}[\mathrm{RNAPT21A1}]\\
&\quad + k_{cat,OFF,21}[\mathrm{RNAPT21}],\\
\frac{d[\mathrm{sI2}]}{dt} &= -k_{s,+}[\mathrm{A2}][\mathrm{sI2}] + k_{s,-}[\mathrm{A2sI2}] + k_{AI,2}[\mathrm{A2sI2}][\mathrm{rI2}] - k_{s,+}[\mathrm{T12A2}][\mathrm{sI2}] + k_{s,-}[\mathrm{T12A2sI2}],\\
\frac{d[\mathrm{A2sI2}]}{dt} &= k_{s,+}[\mathrm{A2}][\mathrm{sI2}] - k_{s,-}[\mathrm{A2sI2}] - k_{AI,2}[\mathrm{A2sI2}][\mathrm{rI2}] - k_{TA,12}[\mathrm{T12}][\mathrm{A2sI2}] + k_{cat,H,2}[\mathrm{RNaseHA2rI2}],\\
\frac{d[\mathrm{T12A2sI2}]}{dt} &= k_{s,+}[\mathrm{T12A2}][\mathrm{sI2}] - k_{s,-}[\mathrm{T12A2sI2}] + k_{TA,12}[\mathrm{T12}][\mathrm{A2sI2}] - k_{+}[\mathrm{RNAP}][\mathrm{T12A2sI2}]\\
&\quad + (k_{-,ON,12} + k_{cat,ON,12})[\mathrm{RNAPT12A2sI2}],\\
\frac{d[\mathrm{RNAPT12A2sI2}]}{dt} &= k_{+}[\mathrm{RNAP}][\mathrm{T12A2sI2}] - (k_{-,ON,12} + k_{cat,ON,12})[\mathrm{RNAPT12A2sI2}],\\
\frac{d[\mathrm{RNAPT12A2}]}{dt} &= k_{+}[\mathrm{RNAP}][\mathrm{T12A2}] - (k_{-,ON,12} + k_{cat,ON,12})[\mathrm{RNAPT12A2}],\\
\frac{d[\mathrm{RNAPT12}]}{dt} &= k_{+}[\mathrm{RNAP}][\mathrm{T12}] - (k_{-,OFF,12} + k_{cat,OFF,12})[\mathrm{RNAPT12}],\\
\frac{d[\mathrm{RNaseHA2rI2}]}{dt} &= k_{+,H}[\mathrm{RNaseH}][\mathrm{A2rI2}] - (k_{-,H,2} + k_{cat,H,2})[\mathrm{RNaseHA2rI2}].
\end{aligned}
$$

The system preserves the conservation relations, $[\mathrm{T21}^{\mathrm{tot}}] = [\mathrm{T21}] + [\mathrm{T21A1}] + [\mathrm{RNAPT21}] + [\mathrm{RNAPT21A1}]$, and similarly for $[\mathrm{T12}^{\mathrm{tot}}]$, $[\mathrm{A1}^{\mathrm{tot}}]$, $[\mathrm{A2}^{\mathrm{tot}}]$, $[\mathrm{dI1}^{\mathrm{tot}}]$, $[\mathrm{RNAP}^{\mathrm{tot}}]$, and $[\mathrm{RNaseH}^{\mathrm{tot}}]$, where the superscript tot indicates that all complexes involving the given species are being counted. Using these conserved quantities, the remaining seven variables, [T21A1], [A1dI1], [rA1dI1], [T12A2], [A2rI2], [RNAP], and [RNaseH], are directly calculated from the concentrations of other species.

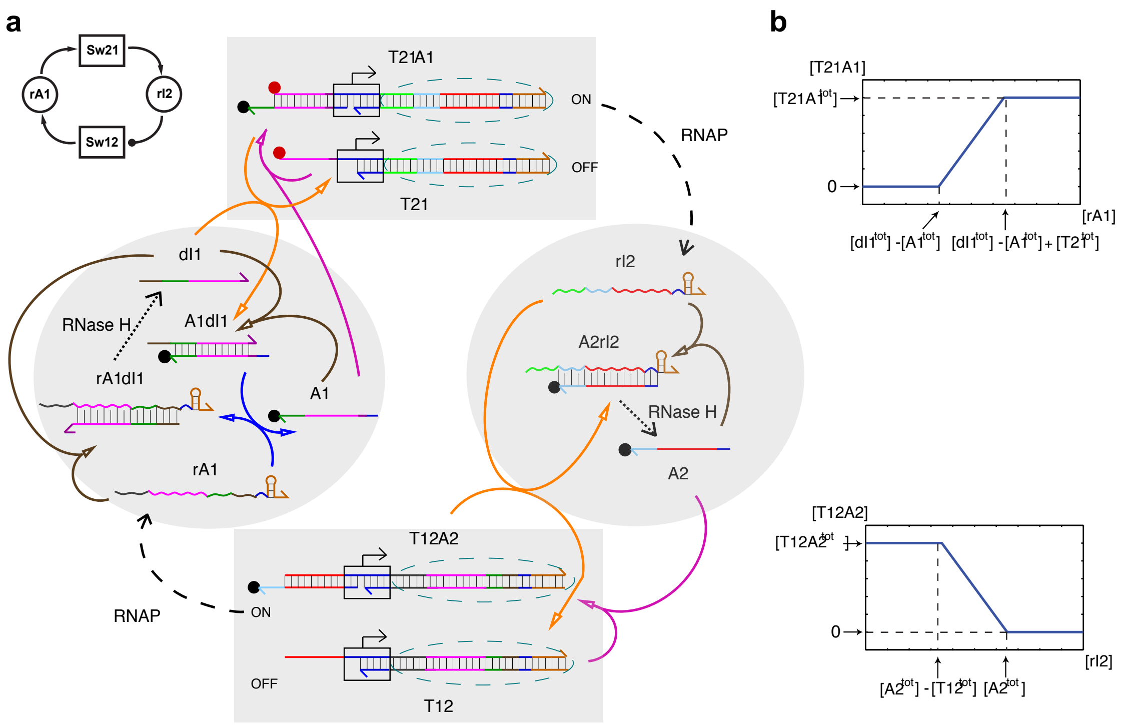

Figure S20: Enzyme and hybridization reactions for the oscillator, colored to facilitate describing the correspondence with the reaction equations. **(a)** Reaction diagram. On the top left is a block diagram for the oscillator, where arrowheads indicate activation or production and circular ends indicate inhibition. The block diagram corresponds to the detailed diagram highlighted in gray shaded areas: T21·A1 (ON-state switch SW21) and T21 (OFF-state switch SW21) are summarized by the SW21 block; RNA inhibitor rI2 together with its threshold, DNA activator A2, and their complex, A2·rI2, are summarized by the rI2 block; the SW12 and rA1 blocks are defined similarly. The sequence domains are color-coded to indicate identical or complementary sequences; for the switch templates, the dark blue sequence domain inside the rectangle indicates the T7 RNAP promoter sequence with arrows pointing in the direction of transcription, with transcription domains indicated by light blue dashed circles. For fluorescence monitoring, the OFF-state switch T21 is labeled with Texas Red fluorophore (red circle) and the activator A1 is labeled with Iowa Black RQ quencher (black circle). Four types of hybridization reactions are indicated by arrows: activation (magenta), inhibition (orange), annihilation (brown), and release (blue). **(b)** Theoretical end-states of hybridization reactions in the absence of enzymes. As the input RNA inhibitor rI2 concentration increases, initially the free DNA activator A2 is consumed without affecting switch state. When all free A2 is consumed (i.e., $[\mathrm{rI2}] = [\mathrm{A2}^{\mathrm{tot}}] - [\mathrm{T12}^{\mathrm{tot}}]$), rI2 displaces A2 from the T12·A2 complex in stoichiometric amounts until all A2 is consumed (i.e., $[\mathrm{rI2}] = [\mathrm{A2}^{\mathrm{tot}}]$), resulting in a piece-wise linear graph. Similarly, the response of switch SW21 to rA1 input is a piece-wise linear graph.

Deterministic Simulation and Model Fits

The kinetic simulations and parameter fittings were implemented in MATLAB. Differential equations were solved using the *ode23s* routine, while the cost function of model fits on experimental data was minimized using the *fmincon* routine. We settled on a cost function using least-squared errors of fluorescence trajectories and characteristics of oscillation (oscillation amplitude, frequency, and damping coefficient) between simulation results and experiments. The error for each experiment is calculated as follows:

$$E = \frac{1}{\mathrm{N_f}}\left(\sum_t (\Delta \mathrm{TR_t})^2\right) + (\Delta\mathrm{Amp})^2 + (\Delta\mathrm{Fre})^2 + (\Delta\mathrm{Damp})^2$$

where Δ indicates the difference between experimental and simulation values, $\mathrm{N_f}$ is the number of fluorescence measurements, TR is the normalized Texas Red fluorescence signal (calculated as $[\mathrm{T21}]/[\mathrm{T21^{tot}}]$ for simulation results), Amp is the amplitude of oscillation, Fre is the frequency of oscillation, and Damp is the damping coefficient of oscillation. The amplitude, frequency, and damping coefficient of oscillation were calculated from the normalized Texas Red fluorescence signals as follows: 1. The peaks and troughs of the oscillation were marked, excluding the initial transient, by calculating derivatives of smoothed trajectories. It is possible to find more peaks and troughs for bulk experiments and simulations due to the negligible noise level as compared to droplet experiments. Thus, we required that the difference between the neighboring peaks and troughs ($|TR(t_i) - TR(t_{i-1})|$) must be greater than 5% of the normalized signal. 2. The frequency of the oscillation was calculated as follows: $\mathrm{Fre} = 0.5 \times (n-1)/(t_n - t_1)$, where n is the total number of marks, t_n is the time of the last marked peak or trough, and t_1 is the time of the first marked peak or trough. 3. The amplitude is calculated as the maximum value of normalized fluorescence difference between successive marks as follows: $\mathrm{Amp} = max_i|TR(t_i) - TR(t_{i-1})|$. 4. The damping coefficient is calculated by fitting the normalized fluorescence differences over time as an exponential decay process ($A = A_0 \times e^{-\lambda t}$), where we fit $log(|TR(t_i) - TR(t_{i-1})|)$ as a linear function of time t_i to extract its slope $-\lambda$. Then, the damping coefficient is simply: $\mathrm{Damp} = \lambda$. (See [Kim and Winfree, 2011] for more detailed examples.)Overall, the contributions for cost functions were roughly equal for the least-squared-error and all three characteristics of oscillation: the normalized fluorescence was used to calculate least-squared-error and amplitude ranging from 0 to 1; the maximum frequency was on the order of 1/hr; the maximum damping coefficient was set to 1 when damping coefficients were not calculated because too few extrema were detected or when the damping was calculated to be more than 1.

Experimental conditions for bulk and microemulsion experiments are as shown in Table S3, where a sustained oscillation (s) was achieved by reducing RNase H concentration as compared to a damped oscillation (d), while a strongly damped oscillation (sd) was achieved by using a higher RNase H concentration with respect to the RNAP concentration (cf. Fig. S21). (See Supplementary Materials (Sample preparation) for details on experiments.) After the fit, the amplitude, frequency, and mean signal levels of the oscillator are generally captured well by the model (Fig. S21), although the phase of the damped tuning oscillations is somewhat off, as is the mean value and frequency for the strongly damped tuning. Poor fits can be expected for the initial part of fluorescence measurements due to 'bursting' enzyme kinetics [Jia and Patel, 1997]. At the same time, poor fits can be expected for the late part of fluorescence measurements due to buffer exhaustion, NTP depletion, and build-up of waste products. Finally, our models neglect the effects of pyrophosphatase, an important enzyme enhancing RNA transcription [Milburn et al., U. S. Patent 5,256,555, 1993]. A complex modeling approach may improve quantitative accuracy for *in vitro* transcription systems by explicitly including pyrophosphatase, NTPs, enzyme life-time, and product inhibition [Arnold et al, 2001].

During the fit, each parameter is constrained within a plausible range spanning about two orders of magnitude as shown in Table S5. For comparison, the maximum hybridization rates and the range of

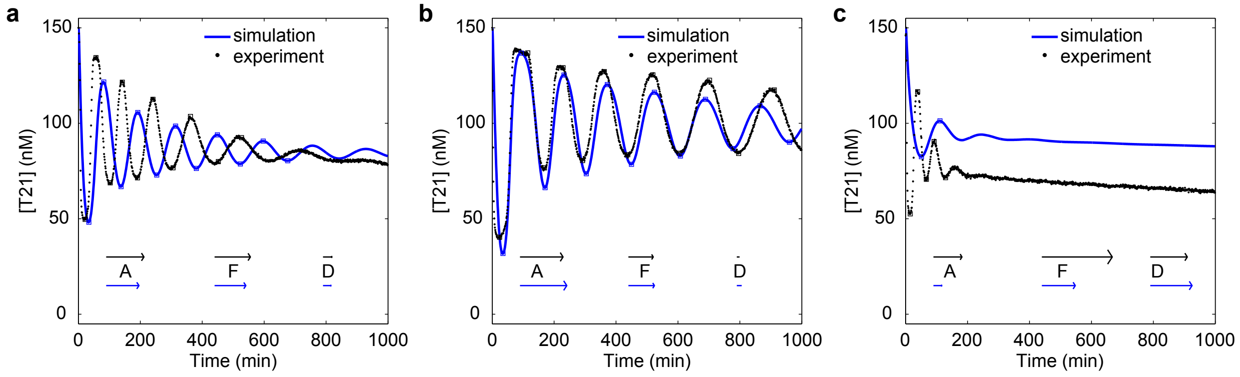

Figure S21: Experimental results and model fits for the synthetic oscillator measured in bulk. The experimental time-courses are plotted as dots (black) and simulation time-courses are plotted as lines (blue). The amplitude A, frequency F, and damping coefficients D of oscillation for experimental and simulation trajectories are shown as black arrows and blue arrows, respectively. The simulation results reproduced a damped oscillation **(a)**, a sustained oscillation **(b)**, and a strongly damped behavior **(c)**.

enzyme constants from other biochemical studies are listed [Rizzo et al., 2002, Martin and Coleman, 1987, Yurke and Mills, 2003]. The hybridization parameters in this study are close to those obtained in our previous study [Kim and Winfree, 2011]: only $k_{TA,21}$, $k_{AI,2}$, and $k_{TAI,12}$ showed differences by more than a few fold. The calculated Michaelis constants from our fit parameters ($K_{M,ON} = 332$–581 nM, $K_{M,OFF} = 1.86$–2.11 µM, and $K_{M,H} = 746$–952 nM) are larger than those provided in the references ($K_{M,ON} = 15$–37 nM, $K_{M,OFF} = 0.1$–1.1 µM, and $K_{M,H} = 16$–130 nM) and those from our earlier studies ($K_{M,ON} = 88$–209 nM, $K_{M,OFF} = 0.88$–1.37 µM, and $K_{M,H} = 82$–154 nM) [Kim and Winfree, 2011]. This could be due to the small size of bulk data set used for fitting with limited range of enzyme concentrations explored. However, the $k_{cat,OFF}$ and $k_{cat,H}$ values are also higher than our previous study [Kim and Winfree, 2011], indicating that the *effective* enzyme speeds (k_{cat}/K_M) are comparable. Like many kinetic models of biological regulation, our synthetic oscillator models have sloppy model features, where many parameters cannot effectively be constrained [Brown and Sethna, 2003]. Thus, diverse sets of reaction rate parameters are potentially compatible with experimental observation, in which similar fits can be achieved when appropriate trade-offs are made. However, our choice of parameters shows that the model we present here is quantitatively plausible compared to other biochemical studies. Thus, we expect that these parameters are more than suitable for qualitative analysis of the oscillator phase space and response to partitioning noise.

Sensitivity to enzyme concentrations

The amplitude, periods, and damping of oscillations are determined by the choice of operating points as shown in previous studies [Kim and Winfree, 2011, Franco et al, 2011]. Particularly of interest is the RNase H concentration: only a small amount of RNase H is required for oscillation, yet the characteristics of oscillation depends sensitively on the RNase H concentration as shown in Figure S22.

From simulation results with the given RNAP concentration of 218 nM, (possibly damped) oscillations with at least 2 complete cycles can be achieved for RNase H concentrations ranging from 3 nM to 6 nM. As the RNase H concentration increases, the periods of oscillation decrease. On the other hand, the amplitude is maximum while the damping coefficient is minimum (the oscillation is more ‘sustained’) for intermediate RNase H concentrations, close to the sustained tuning (RNase H = 4 nM).

Table S5: Parameter ranges and results for model fits

	i=2, j=1	i=1, j=2	Lower limit	Upper limit	Other studies
k_+ (/M/s)	1.90×10^5	–	10^5	10^7	–
$k_{-,ON,ij}$ (/s)	0.0446	0.0634	0.01	0.3	–
$k_{cat,ON,ij}$ (/s)	0.0186	0.0471	0.01	0.1	0.73–1.12
$k_{-,OFF,ij}$ (/s)	0.35	0.388	0.01	1	–
$k_{cat,OFF,ij}$ (/s)	0.0033	0.0131	0.001	0.03	0.11–0.18
$k_{+,H}$ (/M/s)	7.13×10^5	–	10^5	10^7	–
$k_{-,H,j}$ (/s)	0.125	0.0692	0.01	1	–
$k_{cat,H,j}$ (/s)	0.554	0.463	0.01	1	0.02–0.6
$k_{TA,ij}$ (/M/s)	1.59×10^5	9.47×10^3	3×10^3	3×10^5	$0–3\times10^6$
$k_{AI,j}$ (/M/s)	4.69×10^3	5.66×10^4	3×10^3	3×10^5	$0–3\times10^6$
$k_{TAI,ij}$ (/M/s)	1.92×10^4	5.05×10^3	3×10^3	3×10^5	$0–3\times10^6$
$k_{rAI,j}$ (/M/s)	1.50×10^4	–	3×10^3	3×10^5	$0–3\times10^6$
$k_{AIrA,j}$ (/M/s)	1.65×10^4	–	3×10^3	3×10^5	$0–3\times10^6$
$k_{s,+}$ (/M/s)	1.65×10^4	–	3×10^3	3×10^5	–
$k_{s,-}$ (/s)	0.0525	–	0.01	1	–

For instance, the effect of decreasing RNase H concentration by 25 % from the previous operating points are shown in Figure S23.

Not surprisingly, the previously damped oscillation (d) showed a sustained oscillation because the RNase H reduction brings its operating point close to the previously sustained oscillation. On the other hand, the previously sustained oscillation (s) became a damped oscillation upon RNase H reduction. Interestingly, the previously strongly damped behavior (sd) showed a sustained oscillation upon RNase H reduction: this transition to a sustained oscillation can be expected from the phase diagram shown in Figure 6c in the main text. The impact of adjusting the operating points will be discussed further in the next section (Deterministic Simulation with Partitioning Error).

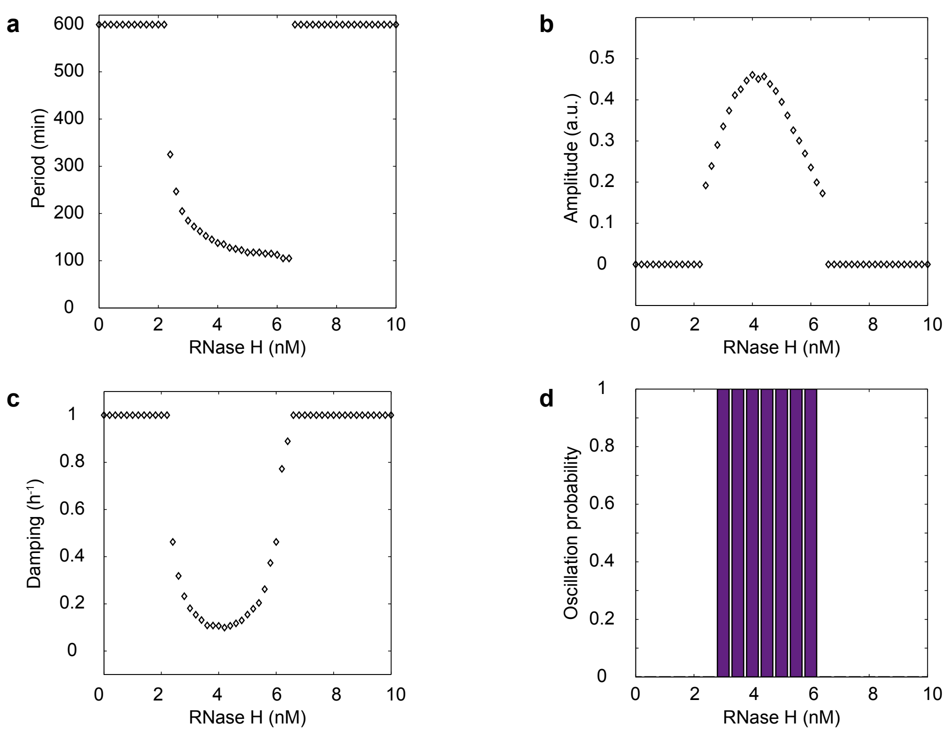

Figure S22: Simulation results of RNase H variation on the synthetic oscillator. The periods **(a)**, amplitudes **(b)** and damping **(c)** of oscillations were calculated as stated in Section Model Equations as long as one complete oscillation cycle was detected. When less than one cycle was detected, a maximum period of 600 min, a minimum amplitude of 0, and a maximum damping of 1 were assigned. The oscillation probability plot **(d)** indicated whether at least two complete cycles were detected. The experimental conditions for damped (d) and sustained (s) correspond to 5 nM and 4 nM RNase H, respectively.

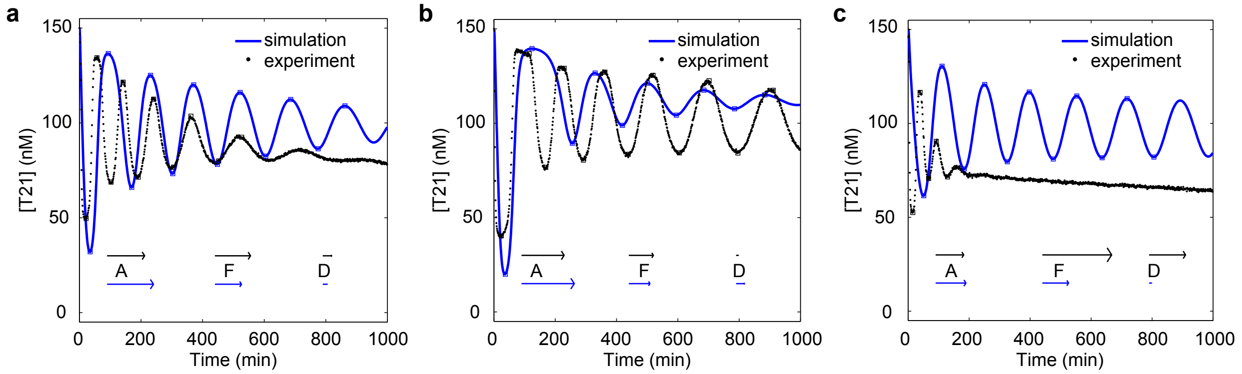

Figure S23: Experimental results and simulated RNase H loss for the synthetic oscillator. The experiments as well as simulation results without RNase H loss showed a damped oscillation **(a)**, a sustained oscillation **(b)**, and a strongly damped oscillation **(c)**, while the simulation results with RNase H loss showed a sustained oscillation **(a)**, a damped oscillation **(b)**, and a sustained oscillation **(c)**.

Deterministic Simulation with Partitioning Error

As we have seen in the previous section, a small change in enzyme concentrations leads to a noticeable change in oscillation characteristics. If we assume an independent distribution of molecular species upon formation of the microemulsion, the molecule count N is expected to follow a Poisson distribution according to $p(N) = \frac{\lambda^N}{N!}e^{-\lambda}$, where the parameter $\lambda = c \cdot V \cdot N_A$ is the expected number of molecules for a bulk concentration c, a droplet volume V, and the Avogadro's constant N_A. For simplicity, in this section, we assume that the chemical kinetics can still be well-approximated by mass action as described in Section Model Equations. (Stochastic reactions are considered in the following section, Stochastic Simulation.) Thus, the number of molecules within droplets were converted to concentration $c_p = N_p/(V \cdot N_A)$ where N_p is drawn according to Poisson distribution as described above. The deterministic simulation of ODEs were performed using the *effective* droplet concentration c_p instead of the bulk concentration c.

Using the initial conditions for sustained oscillations in bulk (s), we explore the effect of Poisson partitioning. As shown in Figure S24, the periods and amplitudes of oscillations are highly variable for small microemulsion droplets due to the initial variability in constituent concentrations, especially that of RNase H, present at very low concentration in the bulk solution. As the reaction volume increases, the Poisson partitioning effect becomes smaller.

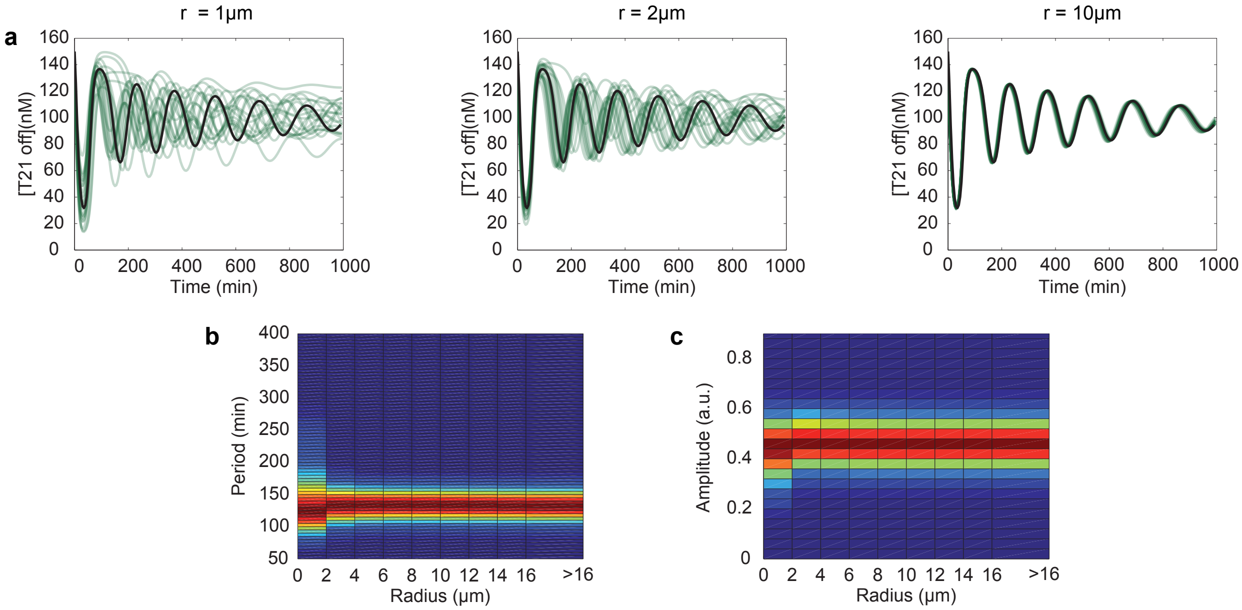

Figure S24: Deterministic simulation of oscillator for the sustained tuning assuming Poisson partitioning during droplet formation. **(a)** The deterministic simulation based on concentrations sampled according to Poisson partitioning are shown as blue lines, while the deterministic simulation based on bulk concentrations are shown as black lines. The simulation results are for microemulsion droplets of radius 1 μm (left), 2 μm (middle), and 10 μm (right). Colorplots represent the relative probabilities of **(b)** periods and **(c)** amplitudes for given droplet radii.

In Figure S25, the periods, amplitudes, damping, and fraction of oscillations are summarized for microemulsion droplets of 1 to 10 μm radii. For small reaction volumes, the periods are longer, the amplitudes are smaller, and the damping of oscillations are stronger on average with larger standard deviations. For droplets with radius 2 μm or less, a small fraction of trajectories failed to complete two oscillation cycles.

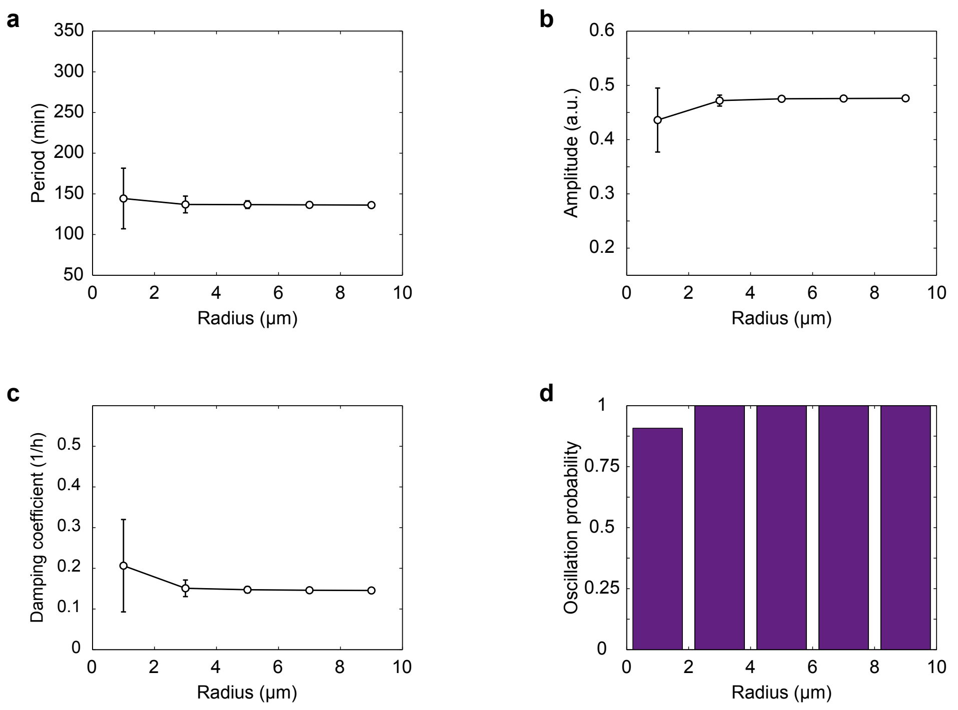

Figure S25: Summary of the oscillation characteristics for simulations using the sustained tuning with partitioning of molecules according to Poisson distribution during the droplet formation. Damping coefficients of oscillations were calculated as described above (Model Equations) as long as two complete oscillation cycles were detected. The amplitudes and periods are calculated as stated in the Supplementary Materials (Calculation of periods and amplitudes). The oscillation probability plot indicated whether at least two complete cycles were detected.

While the trends of oscillation characteristics agree with the experimental results, we observe that the variability of simulated oscillation trajectories are much smaller in magnitude as compared to those of the experimental traces (Fig. 5a). In fact, several potential sources of variability were ignored for the Poisson partitioning assumption. The adsorption of enzymes and DNA molecules may lead to lower than expected concentrations for a given volume, and the potential denaturation of enzymes can reduce the effective concentrations even further. Finally, our model does not account for the effects of pyrophosphatase, a catalyst which improves transcription efficiency; given its low concentration of ≈ 12 nM, which corresponds to ≈ 250 enzyme molecules in a 2 µm radius droplet, it is plausible that pyrophosphatase partitioning contributes to the observed variability. However, our calculations suggest that there is more than enough pyrophosphatase present to process all the pyrophosphate released by transcription, even at much lower enzyme concentrations, and therefore, the overall oscillator dynamics are not likely to be sensitive to pyrophosphatase concentration variability. Also, the molecules may cluster together during partitioning: some possibilities include enzyme-substrate complexes, clumping of proteins, and spermidine-mediated DNA clustering. Thus, we chose to explore the initial concentration variability using an overdispersion model, i.e., a probability distribution where the variance exceeds the mean. For instance, a gamma distribution

$$p(x;\alpha,\beta) = \frac{1}{\beta^\alpha}\frac{1}{\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}$$

with a shape parameter α and a scale parameter β. If the scale parameter β is chosen to be greater than 1, the overdispersion criteria is satisfied because the mean is $\alpha \cdot \beta$ and the variance is $\alpha \cdot \beta^2$.

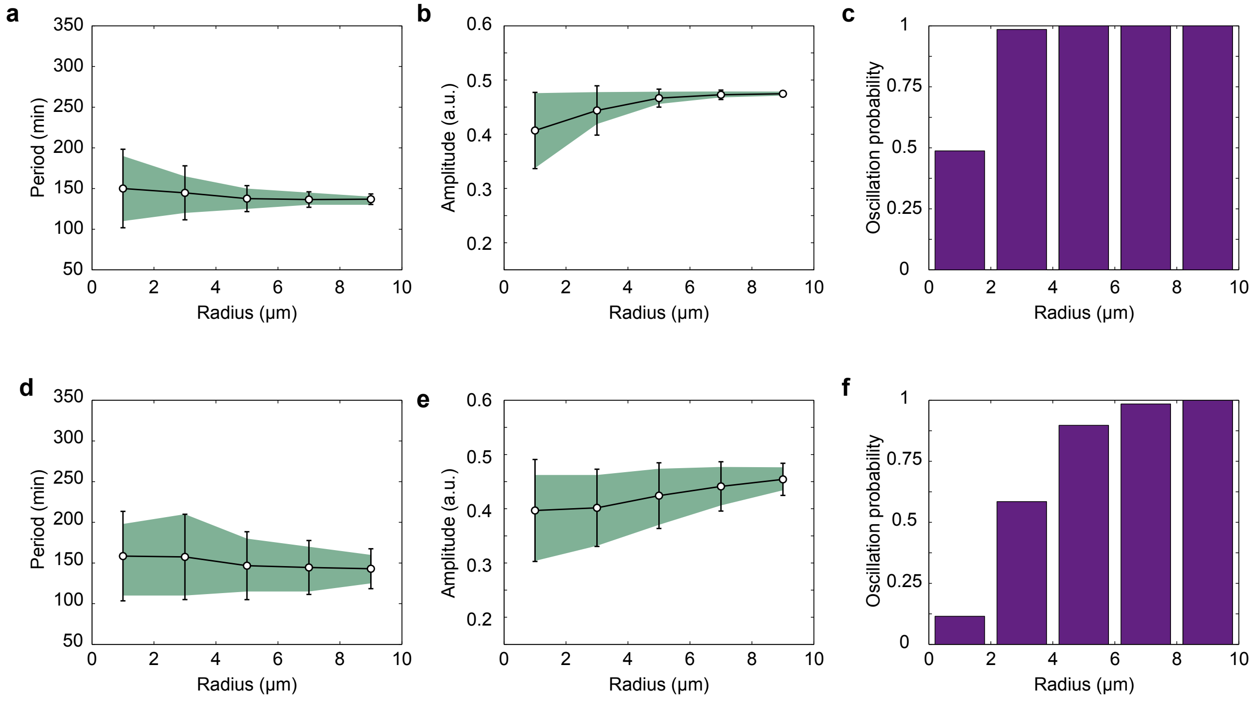

Figure S26: Summary of the oscillation characteristics for simulations using the sustained oscillator tuning upon partitioning of molecules according to gamma distribution. The scale factors for gamma distribution (β) were 10 for the top row and 100 for the bottom row.

The number of molecules in emulsion droplets N_γ are sampled according to a gamma distribution using $\alpha = c \cdot V \cdot N_A/\beta$ and $\beta = 10$, and converted to *effective* concentrations $c_\gamma = N_\gamma/(V \cdot N_A)$ to be plugged into ODEs similar to the Poisson partitioning case above. The sampling procedure according to a gamma distribution showed increased mean and standard deviation of oscillation periods with decreased oscillatory fractions for given droplet radii as compared to a Poisson partitioning (Fig. S26). Similarly, the oscillation amplitude showed increased variability for gamma sampling. When β is increased to 100, the variability for oscillation periods increased even further, analogous to experimentally observed variability (Fig. 5a). On the other hand, for small reaction volumes, the amplitudes are smaller on average with relatively large standard deviations (Fig. S26) unlike the experimental observation (Fig. 5b). This discrepancy may be due to imperfect parameterization of the model or could be related to the individual droplet-based amplitude normalization used for experimental data processing.

Next, using the initial conditions for a damped oscillation in bulk (d), we explore the effect of Poisson and Gamma partitioning. As shown in Figure S27, the oscillation periods increase and the fraction of oscillatory traces decrease as the droplet radii increase in agreement with the previous analysis for the sustained oscillation (s) shown above (Fig. S25 and Fig. S26). Further, the standard deviations of oscillation periods are smaller for a given partitioning model (whether Poisson or Gamma) as compared to the previous analysis. This is in agreement with a smaller experimental variability observed for this initial condition (Fig. 5c): the larger RNase H concentration for damped oscillation (d) as compared to the previous sustained oscillation (s) may contribute to this decreased variability. For small droplets, the amplitudes are larger on average with relatively large standard deviations (Fig. S27) in agreement with the experimental observations (Fig. 5d).

Finally, using the initial conditions which showed a strongly damped behavior in bulk (sd), we explore

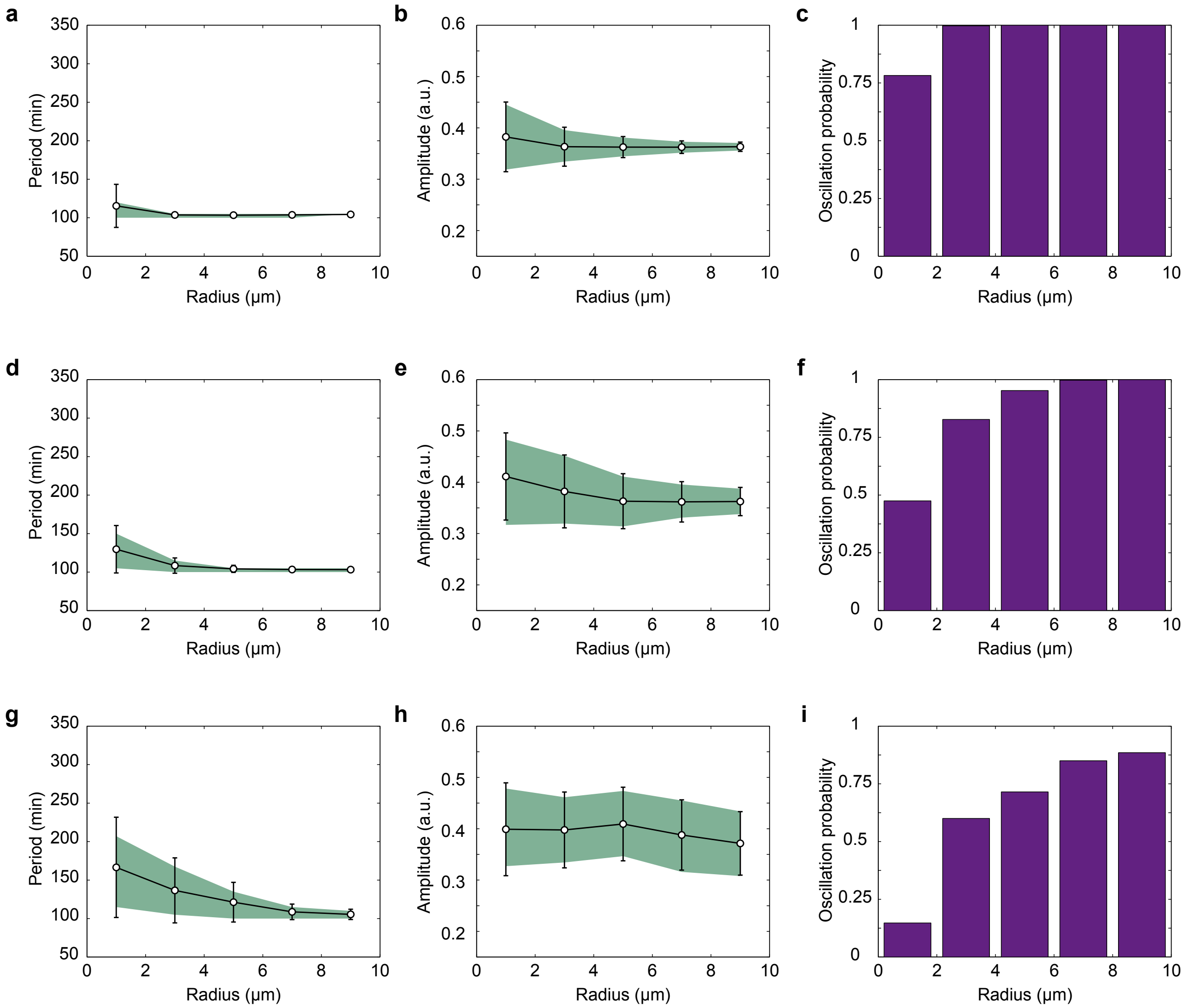

Figure S27: Summary of the oscillation characteristics for simulations using the damped tuning with partitioning variation of molecules during the droplet formation. The top row was based on Poisson partitioning. The scale factors for gamma distribution (β) were 10 for the middle row and 100 for the bottom row.

the effect of Poisson and Gamma partitioning. An interesting trend is observed in Figure S28: the big droplets are not able to support sustained oscillations much as in the bulk reaction, while some fraction of smaller droplets showed sustained oscillations following a Poisson partitioning. For overdispersion cases, the fraction of oscillatory traces increased for larger droplets: the oscillatory fraction peaked for droplets with radius 2 μm or less for Poisson distribution and gamma distribution with $\beta = 10$, while the oscillatory fraction peaked for droplets with 2–4 μm radius for $\beta = 100$. This can be explained by the increased variability of RNase H concentrations for droplets with smaller radii: as shown in Figure S23, a decrease of RNase H with respect to the initial condition in bulk reaction was enough to support sustained oscillations (cf. Fig. 6c). However, too much decrease in RNase H concentration will again result in the breakdown of sustained oscillations. Satisfyingly, this expected feature of sustained oscillation arising in small droplets due to the random sampling of initial conditions was indeed observed experimentally (Fig. 5ef).

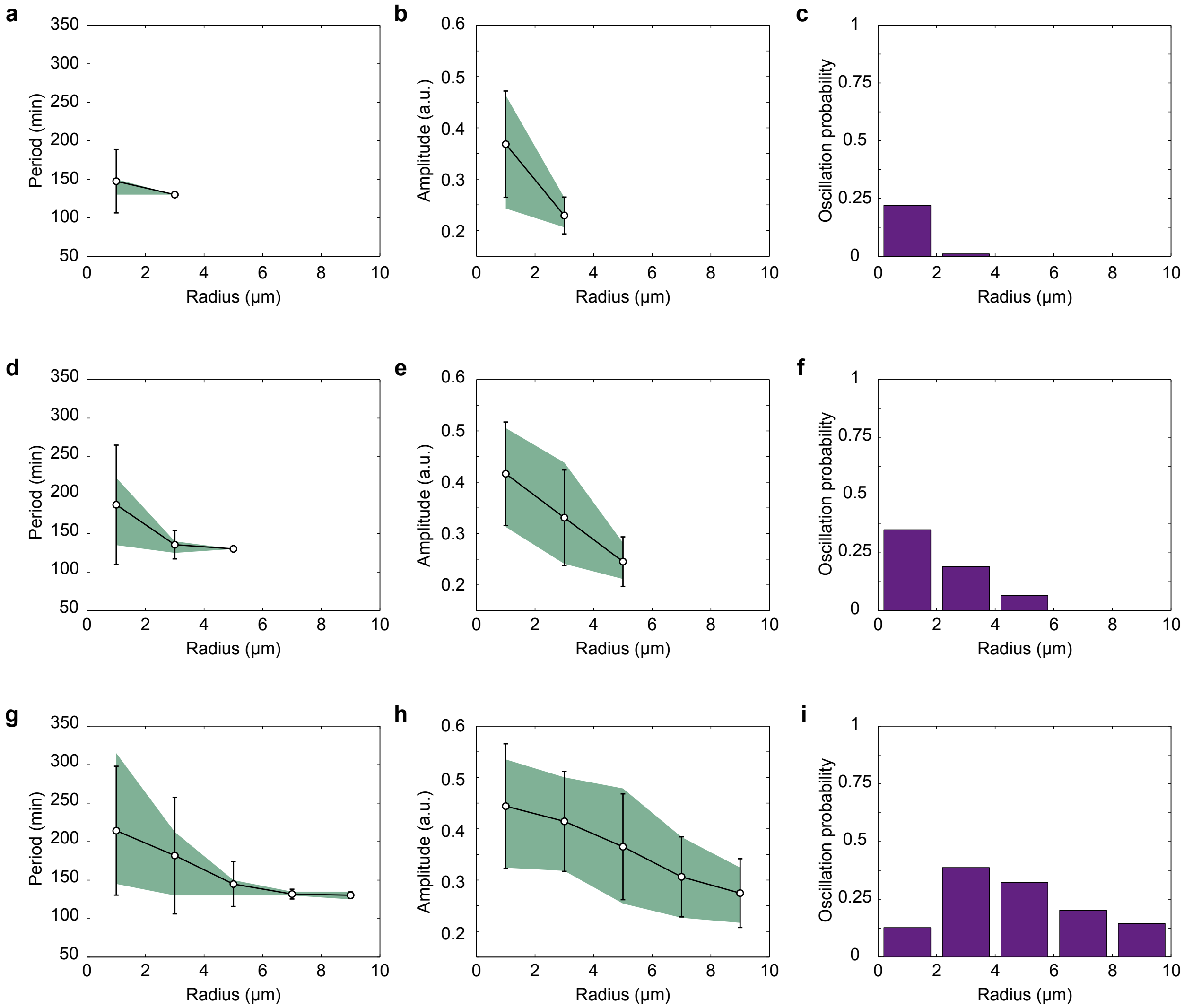

Figure S28: Summary of the oscillation characteristics for simulations using the strongly damped tuning with partitioning variation of molecules during the droplet formation. The top row was based on Poisson partitioning. The scale factors for gamma distribution (β) were 10 for the middle row and 100 for the bottom row.

Deterministic Simulation with Enzyme Loss and Partitioning Error

Based on the estimated enzyme loss during droplet formation as observed in Supplementary Note - Compartmentalization (Variability in microdroplet-encapsulated single-enzyme subsystems), we explored the simulation results when the enzyme loss was included as well as partitioning error upon droplet formation. For this, we use the following fit results obtained earlier for all three enzyme tunings (s, d, and sd). (For droplets with radii ranging from 2 μm to 10 μm, 12% to 11% of activity loss was obtained from the fit for RNase H data. However, since the evidence for radius-dependent loss was not clear in comparison to RNAP data, we chose to implement a conservative loss function that is independent of droplet radii.)

1) For RNAP, a log(R)-dependent loss function was implemented as follows (cf. Supplementary

Note - Compartmentalization):

$$c_{RNAP,droplet} = ((0.0278 + 0.0107 \times \ln(R/\mu\text{m}))/0.083) \times c_{RNAP,bulk}.$$

2) For RNase H, a 10% loss irrespective of the final droplet radii (R) was implemented:

$$c_{RNaseH,droplet} = 0.9 \times c_{RNaseH,bulk}.$$

Given droplet radii, the enzyme concentrations were chosen as above while all the other DNA and RNA species were assumed to retain their bulk concentrations as the mean values. The number of molecules in emulsion droplet N_γ are sampled according to a gamma distribution using $\alpha = c_{droplet} \cdot V \cdot N_A/\beta$ and $\beta = 10$ and 100, and converted to effective concentrations $c_\gamma = N_\gamma/(V \cdot N_A)$ to be plugged into ODEs as described in Section Deterministic Simulation with Partitioning Error. The simulation results are shown in Fig. 6b. We note again (cf. Supplementary Note - Compartmentalization) that these fits are phenomenological and that alternative enzyme loss models are consistent with the data of single-enzyme experiments.

Stochastic Simulation

The set of equations for the oscillator described in Section Model Equations can be used for stochastic simulation for small reaction volumes. Specifically, following the Gillespie algorithm for stochastic chemical reactions in a well-mixed solutions [Gillespie, 1977], the propensities for 35 elementary reactions are calculated using appropriate molecule counts and reaction rates: $N = \lceil c \cdot N_A \cdot V \rceil$, $k_{bi} = k_{bulk,bi}/(N_A \cdot V)$, $k_{uni} = k_{bulk,uni}$, where c is the bulk concentration, N_A is Avogadro's constant, $k_{bulk,bi}$ is bimolecular reaction rates in bulk, $k_{bulk,uni}$ is unimolecular reaction rates in bulk. Two random numbers were generated to choose the next time step and the next reaction. Using the experimental conditions of Reaction #2, stochastic simulation results are shown for reaction volumes ranging from 1 fL to 100 fL.

The stochastic simulation results showed that the stochastic fluctuations contribute significantly to variations in periods and amplitudes of oscillations for small reaction volumes (e.g., for 1 fL reaction volume, averaged across time points, $C.V. = 0.146$ for the number of T21 molecules among different time-courses, due to desynchronization). However, for volumes greater than 30 fL (corresponding to emulsion droplets with 1.87 μm radius), as long as the reactions are synchronized at the beginning with the same number of molecules for all molecular species, the stochastic fluctuations are small ($C.V. < 0.03$). Therefore, we conclude that the stochastic reaction dynamics cannot explain the diversity of oscillations observed in microemulsion droplets bigger than 100 fL (corresponding to emulsion droplets with 2.8 μm radius).

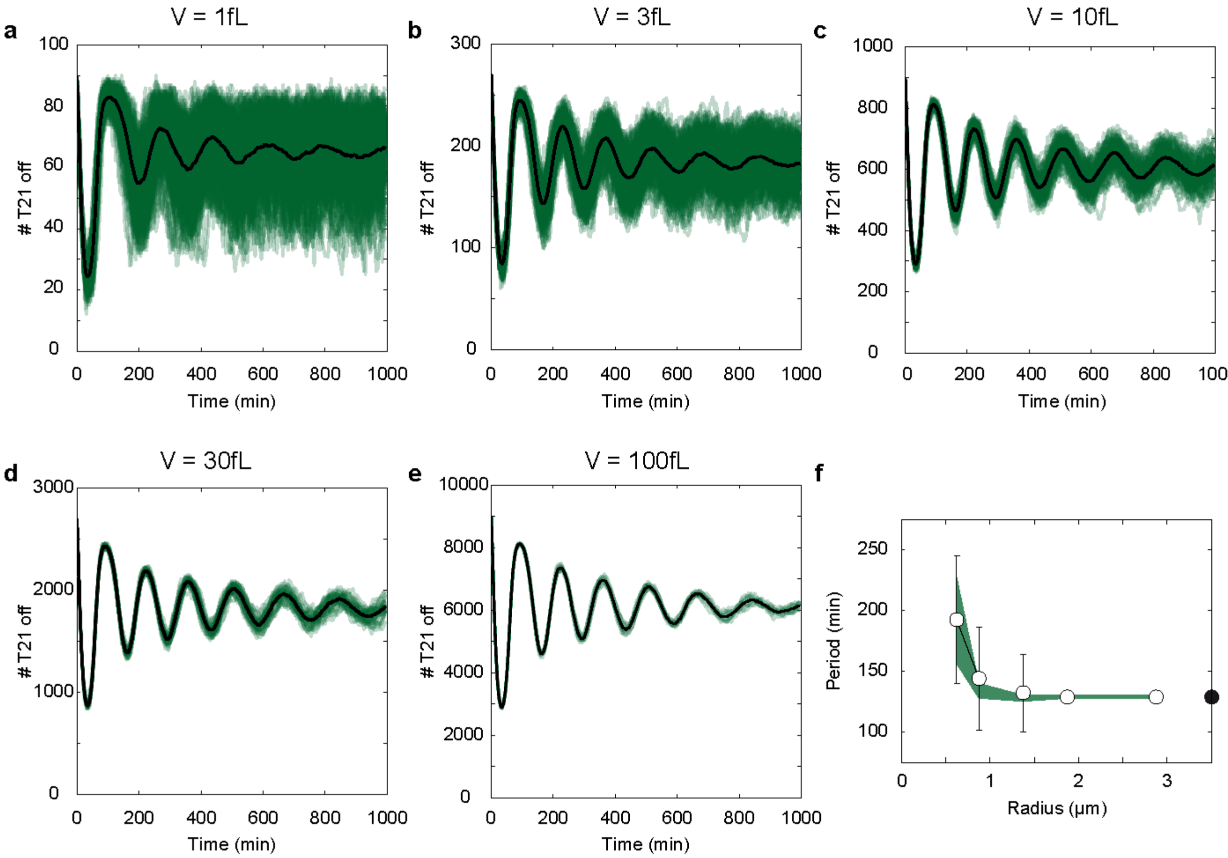

Figure S29: Stochastic simulations for the sustained oscillator tuning. Horizontal axes are time in minutes, and vertical axes are molecule counts for T21 in the transcriptional inactive state. The maximum number of T21 molecules for reaction volume V (fl) is $90 \times V$. Two hundred traces are generated for 1 fL **(a)** and 3 fL **(b)**, 100 traces for 10 fL **(c)**, 30 traces for 30 fL **(d)**, and 10 traces for 100 fL **(e)**. These volumes correspond to radii of 0.60, 0.87, 1.3, 1.87, and 2.8 μm respectively. The individual time-courses are plotted as blue lines and the averages of time-courses are plotted as black lines. The stochastic fluctuations decreased for large reaction volumes.

References

J. Kim. *In vitro synthetic transcriptional networks*. PhD thesis, California Institute of Technology, 2007.

J. Kim and E. Winfree. Synthetic *in vitro* transcriptional oscillators. *Molecular Systems Biology*, 7:465, 2011.

E. Franco, E. Friedrichs, J. Kim, R. Jungmann, R. Murray, E. Winfree, F. C. Simmel. Timing molecular motion and production with a synthetic transcriptional clock. *Proc. Natl. Acad. Sci. USA*, 108: E784–E793, 2011.

S. Milburn, M. Goldrick, and M. Winkler. Compositions and methods for increasing the yields of *in vitro* RNA transcription and other polynucleotide synthetic reactions, U. S. Patent 5,256,555, 1993.

C. T. Martin and J. E. Coleman. Kinetic analysis of T7 RNA polymerase-promoter interactions with small synthetic promoters. *Biochemistry*, 26:2690–2696, 1987.

J. Rizzo, L. K. Gifford, X. Zhang, A. M. Gewirtz, and P. Lu. Chimeric RNA-DNA molecular beacon assay for ribonuclease H activity. *Mol. Cell. Probes*, 16:277–283, 2002.

G. Russo, M. di Bernardo and E. D. Sontag. Global entrainment of transcriptional systems to periodic inputs. *PLoS Comput Biol*, 6(4):e1000739, 2010.

B. Yurke and A. P. Mills, Jr. Using DNA to power nanostructures. *Genet. Program. Evol. Mach.*, 4: 111–122, 2003.

S. Youssef, S. Gude and J. O. Rädler Automated tracking in live-cell time-lapse movies. *Integr. Biol*, 3: 1095–1101, 2011.

K. S. Brown and J. P. Sethna. Statistical mechanical approaches to models with many poorly known parameters. *Phys. Rev. E*, 68:021904, 2003.

Y. Jia and S. S. Patel. Kinetic mechanism of transcription initiation by bacteriophage T7 RNA polymerase. *Biochemistry*, 36:4223–4232, 1997.

S. Arnold, M. Siemann, K. Scharnweber, M. Werner, S. Baumann and M. Reuss. Kinetic modeling and simulation of *in vitro* transcription by phage T7 RNA polymerase. *Biotech. Bioeng.*, 72:548–561, 2001.

C. Holtze, A. C. Rowat, J. J. Agresti, J. B. Hutchinson, F. E. Angile, C. H. J. Schmitz, S. Koster, H. Duan, K. J. Humphry, R. A. Scanga, J. . Johnson, D. Pisignano and D. A. Weitz. Biocompatible surfactants for water-in-fluorocarbon emulsions. *Lab Chip*, 8 (10):1632–1639, 2008.

S. Kaneda, K. Ono, T. Fukuba, T. Nojiama, T. Yamamoto and T. Fujii. Modification of the glass surface property in PDMS-glass hybrid microfluidic devices. *Anal Sci.*, 28(1):39–44, 2012.

D. T. Gillespie. Stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81:2340–2361, 1977.