Emergence of structures from parasitic species in a spatially distributed molecular system

Nathanael Aubert-Kato¹, Guillaume Gines², Teruo Fujii³, Yannick Rondelez²

¹ Department of Information Sciences, Ochanomizu University, Japan
² CNRS, ESPCI Paris, PSL Research University, Laboratoire Gulliver, France
³ Institute of Industrial Science, The University of Tokyo, Japan
naubertkato@is.ocha.ac.jp, {guillaume.gines, yannick.rondelez}@espci.fr, tfujii@iis.u-tokyo.ac.jp

Abstract

In this article, we implement a localized autocatalytic molecular system inside a closed microfluidic chamber. Due to the enzymatic nature of the catalytic process, a variety of parasitic species eventually emerge and compete with the legitimate molecular process for fuel. The behaviors observed range from the creation of large stable structures to that of small diffusing particles. Those results, along with the modularity of the molecular system, show that the proposed experimental setup can be used safely for further study of the evolution of parasitic behaviors at the molecular scale.

Introduction

Parasitism is a common process in nature, at scales ranging from viral infections to full ecosystems. At the molecular scale, as soon as a replication process is present, another, simpler, mechanism will appear to hijack that process (Schmid-Hempel, 2011). That phenomenon has been observed in several in-vitro repeated amplification setups. It is the case, for instance, of Spiegelman’s Monster, where ever shorter sequences of RNA took advantage of the replication provided by other chemical species present in the system (Kacian et al., 1972). Marshall and Ellington demonstrated a setup in which multiple classes of parasites evolved to increase in robustness, with the most robust class capable of withstanding extensive washing and decontamination of the experimental setup (Marshall and Ellington, 1999). Banshoet al. have proposed an in-vitro molecular ecosystem of self-replicating RNA molecules, allowing the evolution of both the hosts (replicators) and parasites (species only replicated without contribution) over time, while avoiding the complete collapse of the system (Bansho et al., 2016).

The legitimate system (host), if it is to survive, will need to provide defenses against parasitism, which will in turn lead to more advanced parasites (Van Valen, 1973; Schmid-Hempel, 2011). While the system may get stuck in a rock-paper-scissors oscillations, that mechanism, called Red Queen dynamics, can lead to open-ended evolution (Hickinbotham et al., 2021) or at least complex population-level behaviors (Rabajante et al., 2015).

As such, host-parasite or predator-prey systems are good objects of study for the emergence of complex behaviors. In that context, using in-vitro experiments is a promising approach, as unexpected behaviors are, by definition, difficult to predict with a model and prone to the reality gap. This advantage comes at the cost that performing the number of experiments required for an evolutionary algorithm can be prohibitive and analyzing results can be difficult as we may lack a clear understanding of the content of the system. The problem of the number of experiments can be mitigated in the case of modular systems or systems with few molecular components by using microfluidic devices to automate the experimental process and sample over the parameter space (Parrilla-Gutierrez et al., 2017; Genot et al., 2016). Yet, tracking the content of the system remains difficult.

On the other hand, using in-silico experiments allows us to set up a variety of experimental conditions that might be difficult to produce in a lab and gives us much faster timescales than in-vitro experiments. Such approach is particularly relevant to test for the probability of emergence of a given phenomenon (Rabajante et al., 2015; Hickinbotham et al., 2021) or scenarios based on hypothesis of the conditions of the early Earth (Scharf et al., 2015). For instance, chemical systems that are hard to reproduce in-vitro, such as the RNA world hypothesis (Gilbert, 1986), can be thoroughly examined under a large range of parameters, showing that, given the right conditions, RNA replicator system may not be overrun by parasites (Synak et al., 2020).

In this work, we aim to combine both approaches by providing a modular reaction framework that can implement complex reaction networks in a rational way, yet allows complex dynamics to appear. Recent advances in the field of molecular programming has opened the door to creating complex reaction networks in a rational way (Padirac et al., 2013), which gives us great flexibility to implement the design. Among such design strategies, enzyme-based systems such as EXPAR (EXPonential Amplification Reaction; Van Ness et al., 2003) or the PEN DNA toolbox (Montagne et al., 2011) are convenient for the task at hand. Small errors in the replica-
Figure 1: Autocatalytic system designed with the PEN DNA toolbox. In the presence of A DNA primer (blue), enzymes present in the environment will produce B (green) through \(A \rightarrow B\) strands. Conversely, B primers will trigger the production of A through \(B \rightarrow A\).

Figure 2: Beads colonies. Top: reaction chamber. The chamber is cut into a Parafilm layer positioned between glass slides and sealed with Araldite®. Bottom: reaction-diffusion system. Sepharose beads are functionalized with DNA strands of two possible types: \(A \rightarrow B\) and \(B \rightarrow A\). When both types of beads are in close proximity, the autocatalytic loop is completed, leading to high local enzymatic activity and primer concentration. Primers are prevented from diffusing too far by an exonuclease enzyme, which degrades free single-stranded DNA.

The experimental setup is based on the colony system of Gines et al. (2017), where two types of functionalized beads are spread in microfluidic chamber. When both types are in close vicinity, enzymatic activity is triggered, leading to the production of specific DNA molecules (signal) as well as, eventually, parasitic species. The beads are introduced in a 1-to-100 ratio, to favor the separation of area where both types are present. Based on previous work, having both type of beads would generate steady-state diffusion ‘cone’ of signal with a length scale of 500 \(\mu m\) (Gines et al., 2017). Our experimental results were consistent with such scale, thus confirming the creation of a rough island model. The microfluidic chamber is sealed by glue as evaporation could easily lead to contamination of the laboratory. Marshal and Ellington wrote “both parasites and their progeny remain in our lab to this day” (Marshall and Ellington, 1999), a properly terrifying cautionary tale.

Methods

The molecular setup is implemented with DNA-functionalized microbeads spread into a microfluidic chamber according to the protocol of Gines et al. (2017).

Molecular system

The “host” chemical reaction network is a two-step autocatalytic cycle implemented with the PEN (Polymerase-Exonuclease-Nickase) DNA toolbox (Montagne et al., 2011). The reaction network is summarized in Figure 1 and works as follows: a short DNA strand (signal) hybridizes with its complementary sequence at the 3’ end of a longer DNA strand (template). The signal strand is then extended by a first enzyme to match of complementary of the template strand. That production relies on dNTPs (deoxyribonucleotides) which act as the fuel of the system. The fully double-stranded structure is designed to contain the recognition site of a second enzyme, called nickase, which will
cut the extended signal strand. The cut restores the initial signal strand while releasing a new signal strand (output). Following the same mechanism, the output strand of the first template can hybridize to the 3’ site of a second template, producing in turn a new copy of the initial signal strand. As those steps repeats, the number of both species of signal strands increase over time. The accumulation of signal strands is limited by a third enzyme, called exonuclease, which degrades single-stranded DNA species present in the environment. DNA templates are chemically protected against that activity and remain stable over time.

**Functionalized beads**

We used streptavidin-conjugated Sepharose microbeads, with a media diameter of 34 ± 10 μm. That size range was selected to provide enough binding capacity for the molecular system while greatly reducing Brownian motion, allowing us to keep track of the activated area over long periods of time.

We followed Gines et al.’s protocol to bind to the beads to biotin-conjugated DNA templates. We first wash the beads by successive centrifugation and supernatant removal in the presence of Binding and Washing buffer. After repeating that process three times, the beads are resuspended into a mix of Binding and Washing buffer and DNA templates in TE buffer. Note that the functionalization is done below the maximum binding capacity of the beads to reduce leaks of DNA templates into the environment. See Gines et al. (2017) for additional details on the protocol.

We prepare two types of beads (Figure 2, bottom): beads functionalized with A→B DNA templates and beads functionalized with B→A DNA templates. When both types of beads are in close proximity, their respective outputs triggers the production of signal, thus completing the autocatalytic cycle. When only one type of bead is present, only spurious (unprimed) production from the polymerase can happen, the product of which should be quickly degraded by the exonuclease.

To keep the full reaction limited to certain area and create hot-spots of enzymatic activity, A→B beads (majority beads) are mixed in a 100x excess compared to the B→A beads (minority beads). Further, minority beads are also marked with an Atto633 barcode to easily differentiate both types.

**Chamber**

The reaction environment is made of a Parafilm sheet cut using a Rolan cutter plotter placed between two glass slides (Figure 2). Prior to assembly, the bottom glass was spincoated with Cytop CTL809-M to prevent the inhibition of the nicking enzyme (Nt.bsmNBI). The top glass slide is diamond cut to leave space for an inlet, allowing the introduction of the molecular system in the chamber. The assembled chamber is then backed at 60°C for three minutes to ensure the cohesion of the Parafilm layer to the glass slides.

The molecular mix (functionalized beads, buffer, and enzymes) is introduced by capillarity through the inlet, followed by oil (silicon) for waterproofing. Araldite®glue is then applied to seal the inlet as well as the sides of the top glass slide.

We prepared three identical chambers with different concentrations (0.5%, 1% and 2% of volume) of Bst full length polymerase (New England Biolabs, 5000 U ml⁻¹), as errors in the transcription is one of the major potential source of parasitic species. All other chemicals were kept in the same concentration across and chambers and came from the same master mix to avoid inconsistencies in mixing.

**Microscopy experiment**

Reaction chambers were kept at 42°C and monitored by an Olympus IX71 inverted microscope. We used two fluorescence sources: EvaGreen®, a non specific DNA-binding dye attaching to double-stranded species and Atto633 (red fluorophore, attached to minority beads).

We used an XY stage to take pictures of 30 regions of interest (10 per experimental settings) at 5 minute intervals. The regions of interest were manually selected by using the Atto633 fluorescence. We picked regions with at least one minority bead and 2 or more majority beads in close proximity to trigger a reasonably sized zone of enzymatic activity.

Raw fluorescence images were enhanced and analyzed with the Fiji software.

**Results**

Until the 800 min mark, our implementation displayed a similar behavior to that of Gines et al. In all conditions, colonies were able to start the autocalytic process without primer, in accordance with previous results (Figure 3).

Colonies show a growth in fluorescence intensity and radius until they reach a steady-state between 100 min and 400 min after the begin of the experiment. A typical colony is shown in Figure 4, showing the scale of the active area. Beyond that zone, DNA signal is only subject to the activity of the exonuclease and is quickly degraded.

Around the 1000 min mark, we could observe the appearance of parasitic species from an untemplated increase of DNA species in the environment in colonies from the system with the highest concentration of polymerase. In particular, we could observe the emergence of large structures of up to 100 μm (Figure 5) and a characteristic exponential increase in background fluorescence (Figure 6). We could also note a similar trend starting from 1500 min for lower enzymatic concentrations. In all case, the fluorescence increase preceded the appearance of structures, making it a good predictor for the presence of parasites.

We could observe three types of parasites. Type I parasites form large (in the 100μm range), elongated structures. Those structures, like all other parasitic structures, grow...
Figure 3: Normalized fluorescence over the region of interest. Due to the variable number of beads per area, the fluorescence is normalized by using the minimum and maximum intensity before parasitic emergence. Top: Fluorescence over the full experiment, showing the impact of parasites. Bottom: Data limited to the legitimate behavior range.

Figure 4: Active colony with a minority bead (center) surrounded by activated majority beads. A noticeable increase of fluorescence is seen over a radius of approximately 650 µm.

Figure 5: Emergence of a structure from untemplated polymerization. Starting at $t_0 = 1015$ min since the start of the experiment. Scale bar: 200 µm.

Figure 6: Background fluorescence over time.

<table>
<thead>
<tr>
<th>Bst polymerase fraction</th>
<th>Type I</th>
<th>Type II</th>
<th>Type III</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5%</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>1.0%</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2.0%</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 1: Out of 10 repeats for each condition, number of times a given type of behavior could be observed. Note that multiple behaviors could be observed in some spots at 2.0% Bst.

Thus, a short span of time before reaching their maximum size, possibly due to the local exhaustion of the dNTPs or an equilibrium between production and degradation. Type II parasites form spherical structure at the 20 µm. Such structure could be easily confused for legitimate beads if it was not for their spontaneous appearance in the environment. Type III parasites form a large number of very small structures (<10 µm) together with a much larger increase of background fluorescence than other types, hinting at the presence of single-stranded or at least unstructured parasites. Contrary to other types, type III parasites show a decrease in the number of structures and background fluorescence some time after their appearance. This phenomenon can be explained by the degradation of parasitic species by the exonuclease once production stops due to the exhaustion of dNTPs in the environment. The distribution of parasites for each polymerase conditions is shown in Table 1.

Examples of all parasitic behaviors are shown in Table 2 along with the legitimate behavior of the system.

Finally, we could observe reaction-diffusion front of parasite invading from other areas. In some cases, the front seemed to slow down past a certain point (e.g. position 23 in the raw data), while in other cases multiple fronts could be seen colliding (e.g. position 22). That later phenomenon, along with the timing of appearance of the different species, is a strong hint that multiple independent parasitic species did emerge in a single environment.

Discussion and Conclusion

In this work, we implemented a sealed reaction-diffusion system that provides independent spots of autocatalytic activity, applying the design of Gines et al. (2017). Those spots...
Table 2: Parasite types. The first column shows the initial configuration of the region of interest. The middle column shows the steady state configuration, before the emergence of parasite. The last column shows the region of interest at the maximum fluorescence intensity, corresponding the highest parasitic activity. Scale bar: 200 µm.
can in turn promote the emergence of a variety of parasitic structures, in a way that prevents contamination in the experimental laboratory. The implemented reaction network can be arbitrarily edited thanks to the modularity of the PEN DNA toolbox, the molecular framework we rely upon. Furthermore, events in the system can be monitored in real-time through fluorescence. The system thus provides a trade-off between a direct in-vitro implementation and an in-silico simulation of complex reaction networks.

We showed that the background fluorescence was a good indicator for the appearance of parasitic structures. That measure can be used either to track points of interest at a larger scale before zooming in on those points, or as a marker to stop an experiment and limit the risks of contamination. Note that total fluorescence can also be an indicator of parasitic activity, as structures made by the parasitic species have a very high fluorescence due to their high concentration in DNA. However, the overall plot is harder to read, as the local disposition of beads will yield different total fluorescence value, and may have delay in the autocatalytic start of different regions. A late exponential increase may simply be a new cluster rather than a parasite. Position 20 in the raw data is a good example. We suspected the appearance of parasites near the beginning of the experiment as a region without the minority type bead started producing a fluorescent signal. However, no leak was found in the background fluorescence until the very end of the experiment when parasitic species from other area started invading, suggesting that the early behavior is legitimate.

The various species in the system are in competition for fuel (dNTPs). As the system is closed, the total consumption of dNTPs will stop the evolution, but assuming that the proper precautions are taken (e.g. [Bansho et al. 2016]), it is possible to unseal the system to add more fuel or extract species for further analysis or implementation into a different environment. Nevertheless, the risk of contamination is a very real deterrent, as contamination would interfere with further, non-parasitic oriented research in the lab [Marshall and Ellington 1999].

Another limitation of the contamination risk is that it prevents the direct analysis of the content of the environment. In particular, we are unable to confirm that the different parasitic behaviors are indeed produced by unrelated molecular species. Depending on its local concentration, a given species (or family of species) could possibly create different shapes [He et al. 2008].

The current experimental setup opens the door to multiple venues of investigation. For instance, it is unclear why type I parasites could form asymmetrical structures. One possibility would be that those structures grow on dust or other particles present in the environment. If that was the case, there might be a direct connection between type I and type II parasites, the difference residing on the area that triggered the appearance of the structure. However, in that case, it is unclear why such structures could not be observed at lower polymerase concentration. One possibility would be that a minimum level of local production is required to overcome the degradation and start the growth. Similarly, type III parasites tend to emerge along type II parasites, which hints at a relation between the two types. However, due their independent emergence, it is reasonable to assume that at least some sequences are unrelated. Another hint comes from the observation that different parasitic species seem to have different impact on the legitimate colonies. The parasite selected as an example of type 2 behavior (Table 2; position 22 in the raw data) greatly dims the activity of the colony, while the other two examples (types 1 and 3; positions 23 and 29 in the raw data) have a much less dramatic impact. One potential explanation for that difference is that the sequence of the parasite present at position 22 is similar to that of legitimate signal species, thus interfering with the local activity, while other parasites are mostly independent, indirectly interfering by consuming dNTPs.

Despite the apparent connection between different types of parasites, it is remarkable to note that only type II parasite could form at lower polymerase concentration. While it is possible that other structures might form later on, a more likely explanation is that enough production is required to overcome the degradation from the exonuclease. Additional types of parasite may be observed at higher concentrations.

We can also note the fast propagation of parasitic species in the environment from the background fluorescence. Measurement of the propagation fronts at a larger scale would allow the estimation of the reactivity of the parasitic species [Zadorin et al. 2015]. In particular, comparing fronts across multiple repeats would allow us to confirm if independent parasitic sources have similar reactivity with respect to the enzymes present in the system.

Another application of focusing on multiple scales is the monitoring of the origin of the different parasitic species. As the current experiment focused on limited regions, we did not confirm the lack of appearance far from autocatalytic area. As at least some parasitic species appear to be resistant to exonuclease, most likely through secondary structures, they may have appeared for the spurious activity of the polymerase enzyme before diffusing beyond their original area.

Moreover, the present experimental setup is a useful tool for exploring emergent behaviors in an in-vitro setting, with application as a “breadboard” molecular system for research on both ALIFE [Bedau 2003] and the origins of life [Scharf et al. 2015]. In the current design, the DNA sequence of species A and B are completely independent. The system can thus be extended to provide longer cycles (A → B → C → ... → A), independent cycles (A₁ → B₁ → A₂, ..., Aₙ → Bₙ → Aₙ) or even hypercycles by having species of one cycle catalyzing another cycle [Fujii and Rondelez 2013]. While the autocatalytic network (host) is fixed, it is possible to design a system capable of switch-
ing between different reaction networks, using for instance one that would be more resistant to parasites (Urtel et al., 2019) but tuned to be less efficient than the default design (Montagne et al., 2016). Such design would provide a direct in-vitro implementation of two steps of the arms race. It is also possible to go beyond the PEN DNA toolbox and produce not only signal strands, but template strands as well. In that case, much more flexibility would be available to the system, at the cost that the exact reaction network would be unknown. That arms race introduces a soft selection pressure that is similar to the Red Queen dynamic (Van Valen, 1973; Schmid-Hempel, 2011). Such approach can be used, in particular, as a trigger for open-ended evolution (Raba-jante et al., 2015).

The system can be also extended by using molecular delays (Aubert et al., 2014; Kishi et al., 2018) and making molecules (including PEN DNA toolbox template) available at different timing, thus forcing a programmed evolutionary for the host side. One can go even further and have a more complex in vitro system such as a full RNA transcription mechanism (e.g. Vaidya et al., 2012).

A different set of nicking enzyme has been shown to trigger the \textit{ab initio} creation of DNA species in the presence of dNTP and Bst Polymerase (Antipova et al., 2014). The same mechanism could be exploited to increase the diversity of molecular species present in the environment, as they may rely on different enzymes to emerge or further evolve to take advantage of additional enzymes.

Additionally, methods have been proposed to prevent the apparition of parasites and PEN DNA toolbox and related chemical reaction frameworks (Gines et al., 2017; Urtel et al., 2019). Those methods rely careful sequence design for the templates and the addition of specific molecules in the reaction buffer. As such, those methods could be directly integrated into the current system and tune the difficulty of the environment for parasitic emergence. Moreover, by using both beads with DNA sequences prone to producing parasitic species and beads with DNA sequences preventing such mechanism, the difficulty of emergence can be made non uniform.

Finally, despite the closed nature of the system, inputs from the outside are still possible, for instance by changing the temperature and thus the activity of the various enzymes in the system in a non-linear way (Lobato-Dauzier et al., 2020), or by using azobenzene modified molecules (Asanuma et al., 2007), which can reversibly change conformation based on exposition to UV or visible light.

**Appendix**

<table>
<thead>
<tr>
<th>DNA sequences</th>
<th>Template</th>
<th>Sequence (5’ →3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A→B</td>
<td>bioteg-<em>C</em>A*CTGACUCCCTCAAGACTCAG-p</td>
<td></td>
</tr>
<tr>
<td>B→A</td>
<td>bioteg-<em>C</em>A*AAAAACACGACTGACTCTCT-p</td>
<td></td>
</tr>
</tbody>
</table>

“bioteg” (triethylene glycol linker) denotes a biotinylated synthon.

“s*” denotes a phosphitionate backbone, used to prevent exonucl ease activity. “p” denotes a 3’ phosphate modification, used to prevent polymerase extension.

**Data availability**

All microscopy images are available on Dryad [https://doi.org/10.5061/dryad.34tmpg4k7](https://doi.org/10.5061/dryad.34tmpg4k7). Use a software such as Fiji (Schindelin et al., 2012) to increase the contrast in the files.

**Acknowledgments**

This work was supported by JSPS KAKENHI Grants Number JP17K00399 and JP19KK0261.

**References**


