

How to model DNA replication in stochastic models of synthetic gene circuits (and why)

Samuel E. Clamons and Richard M. Murray*

Caltech, Pasadena, CA, United States

E-mail: murray@cds.caltech.edu

Abstract

Biocircuit modeling sometimes requires explicit tracking of a self-replicating DNA species. The most obvious, straightforward way to model a replicating DNA is structurally unstable and leads to pathological model behavior. We describe a simple, stable replication mechanism with good model behavior and show how to derive it from a mechanistic model of ColE1 replication.

Introduction

Although most synthetic biocircuits use DNA, models of synthetic biocircuits typically do not explicitly describe the dynamics of those DNA species. For example, the original repressilator model tracked mRNA and protein species, but not DNA,¹ and the model for the first genetic toggle switch simply tracked “repressor 1” and “repressor 2”.² For many circuits under many modeling assumptions, it is sufficient to assume that all DNA species are held at a fixed concentration by the cell, by means of mysterious machinery whose details are irrelevant to understanding the circuit.

In some cases, however, it is useful or necessary to explicitly represent DNA as a dynamic species. Some circuits, for example, use DNA in different states as a dynamic component or

readout (for example, integrase-based state machines³). Explicit representation of DNA can also be useful when DNA binding is slow relative to the other circuit processes (for example, CRISPR-based transcription factor networks under some conditions⁴). Stochastic models, in particular, are often most naturally expressed using explicit DNA species.

Explicitly-modeled DNA often requires some mechanism of replication – again particularly in stochastic models. Unfortunately, the obvious replication implementation $DNA \rightarrow DNA + DNA$ is a trap that leads to pathological circuit behavior.

Trivially-replicating plasmids have no well-defined copy number

An obvious way to model replication of a DNA species is the straightforward



shown visually in Figure 1A. We refer to this mechanism as “trivial self-replication”. Although trivial replication is intuitively appealing, we strongly recommend against its use.

Adding a dilution reaction $DNA \xrightarrow{\gamma} \emptyset$, trivially self-replicating DNA has dynamics

$$\frac{dDNA}{dt} = (\alpha - \gamma)DNA$$

Notice that at steady state, we have $\alpha - \gamma = 0$, *independent of DNA concentration*. In other words, the only finite, non-zero steady state of the trivial replication mechanism occurs when production is precisely balanced by dilution. If production is slightly faster than dilution, then the DNA’s concentration will explode unphysically (and unbiologically) to infinity. Conversely, if dilution is slightly faster than production, *DNA* will always fall to zero concentration and die out. Such a steady state is structurally unstable

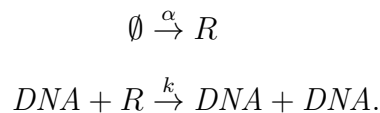
In a deterministic model of a trivially-replicating plasmid, production and dilution *can* be

balanced perfectly, giving a nominally constant concentration of DNA, but this mechanism of replication rejects no disturbances—any addition or removal of DNA will remain permanently uncorrected.

Stochastic simulations of trivially-replicating DNA cannot even achieve this level of marginal stability. Stochastic simulation is, by its nature, noisy; a stochastically simulated, trivially replicating DNA will random walk in copy number until, practically speaking, it either dies out by wandering to zero or explodes to a concentration too large to simulate (see Figure 1A).

Zero-order replication recovers good steady-state properties

The trivial replication mechanism is unstable because it has both production and degradation of DNA depend linearly on the concentration of DNA itself. A simple way to add stability to the replication model is to remove that dependence by conditioning replication on a “dummy replication trigger” produced at a constant rate (shown diagrammatically in Figure 1B):



This mechanism has ODE dynamics

$$\begin{aligned}\frac{dR}{dt} &= \alpha - kR * DNA \\ \frac{dDNA}{dt} &= kR * DNA - \gamma DNA\end{aligned}$$

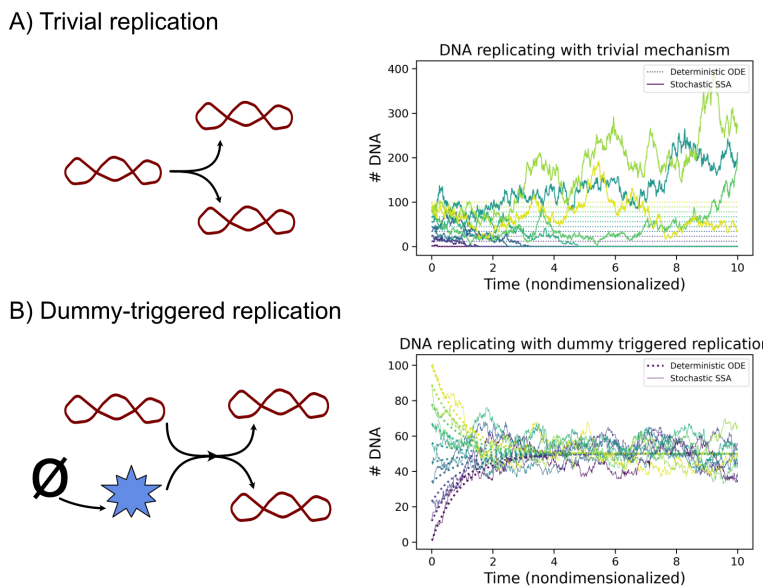


Figure 1: **A)** The trivial replication mechanism, in which DNA spontaneously self-replicates. This mechanism is unstable and produces random-walking DNA concentrations. **B)** The dummy-triggered replication mechanism, in which replication is triggered by a dummy molecule produced at a constant rate. This mechanism, coupled to dilution, is stable and rejects disturbances.

At steady state, $R = \frac{\alpha}{k * DNA}$, which cancels out DNA in DNA 's production term and gives $DNA = \frac{\alpha}{\gamma}$. This steady state is stable. As long as k is fast relative to other replication dynamics (α and γ), the dummy-triggered replication mechanism emulates zero-order replication of DNA.

Because dummy-triggered replication leads to a stable steady state concentration of DNA, as shown in Figure 1B, we recommend it over the trivial replication mechanism.

Biological models of plasmid replication approximately reduce to zero-order replication

Dummy-triggered replication is not meant to accurately describe a real biological processes. Real cells do not control replication using consumable molecules like R . The critical property of the dummy-triggered replication mechanism is not its realism, but its ability to achieve

the (biologically important!) property of having a defined, stable steady state.

Nevertheless, under a few relatively mild assumptions, at least one real-world DNA replication mechanism can be reduced to a zero-order replication mechanism equivalent to the dummy-triggered replication mechanism, .

Consider the ColE1 plasmid replication system, first crystallized mathematically by Brendel & Perelson in 1993, shown on the left in Figure 2.⁵ We use ColE1 as an example because it has a particularly simple and well-understood replication mechanism.

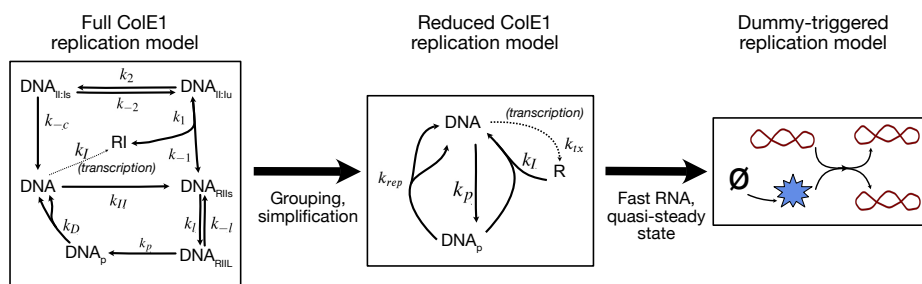


Figure 2: A mechanistic model of ColE1 (left), which can be conceptually approximated by a simplified ColE1 model (middle), which under fast RNA dynamics reduces zero-order replication equivalent to that of the dummy-triggered replication model (right).

Although the details of the Brendel & Perelson model are a bit messy, they implement a straightforward logic:

- *DNA* produces an RNA species *RI*. More *DNA* leads to more *RI*.
- Occasionally, *DNA* will spontaneously enter a primed state from which it can replicate. (Biologically, the plasmid produces an RNA primer *RNAII* which can initiate replication.)
- *RI* can react with a replication-primed *DNA*, un-priming it and destroying the *RI*. (*RI* binds to *RIII* to form an inert complex that is removed by RNases.)
- If not stopped by an *RI*, a primed *DNA* can spontaneously replicate into two non-primed *DNA*. (DNA polymerase initiates replication using *RIII* as a primer.)

We can represent this logic more clearly with a three-species reduction of the Brendel & Perelson model, shown in the middle in Figure 2, consisting of “unprimed” DNA , “primed” DNA_p , and feedback RNA R .

We can further simplify the three-species model by assuming that RNA transcription and degradation are fast compared to other replication dynamics. This should be a reasonable assumption, as both $RNAI$ and $RNAII$ have estimated half-lives of ≈ 2 minutes,⁵ while the plasmid’s doubling time is roughly one *E. coli* cell generation time, or at least ≈ 20 minutes. Under this assumption, the three-species reduced ColE1 model can be approximated with one of two different single-species models, depending on the relative magnitudes of rate parameters k_p and k_{rep} . With $k_{rep} \gg k_p$, total DNA replication rate becomes constant with sufficiently large copy number; when $k_p \gg k_{rep}$, replication rate is always constant. With the right parameterization, we have arrived back at a simple zero-order replication mechanism.

See supplemental notebook “simple_bp_model_reduction.ipynb” for derivations.

Acknowledgments

The authors thank John Marken and Andrew Halleran for insightful discussions and thoughtful editing.

Supporting Information

For more information and worked examples using the models described here, see our supplementary IPython notebooks, included as supplemental files or at https://github.com/sclamons/plasmid_replication_modeling (see the “code” folder). All notebooks use the BioSCRAPE modeling package to simulate systems with replicating DNA using both standard Gillespie stochastic simulation algorithm (SSA)⁶ and a lineage-based simulation algorithm incorporating cell growth and division.

See supplementary notebook “model_comparison_and_fitting.ipynb” for:

- Full descriptions of the trivial, dummy-triggered replication, full ColE1, and three-species ColE1 models.
- Examples of replicating DNA using each of the models with the BioSCRAPE simulation package,
- Replication of Figure 1.
- Comparison of model output to empirically-measured plasmid copy numbers taken from Shao et al., 2020,⁷ including parameter inference.

See supplementary notebook “simple_bp_model_reduction.ipynb” for:

- Derivations of the reductions made in section .
- Random-parameter simulations evaluating the accuracy of reductions to the three-species reduced ColE1 replication model.
- Parameter conditions for stability of the three-species reduced ColE1 model.

See supplementary notebook “CRISPRlator.ipynb” for:

- Worked examples of a 5-node CRISPRi-based repressilator^{8,9} implemented on a plasmid using each of the trivial, dummy-triggered replication, full ColE1, and three-species reduced ColE1 models.
- Simulations showing synchronization of the CRISPRlator across many generations of growing and dividing cells, and dependence of that synchronization on circuit copy number, using the dummy-triggered replication model.

See supplementary notebook “temporal_logic_gate.ipynb” for:

- Worked examples of a population-based DNA integrase circuit capable of sensing the order of and time between arrival of two signals.¹⁰

- Demonstration of the same circuit moved to a single cell bearing the circuit on a 30 or 100 copy number plasmid (instead of in many cells bearing a single genomic copy) using dummy-triggered replication.

References

- (1) Elowitz, M. B., and Leibler, S. (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403, 335–338.
- (2) Gardner, T. S., Cantor, C. R., and Collins, J. J. (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403, 339–342.
- (3) Roquet, N., Soleimany, A. P., Ferris, A. C., Aaronson, S., and Lu, T. K. (2016) Synthetic recombinase-based state machines in living cells. *Science* 353.
- (4) Jones, D. L., Leroy, P., Unoson, C., Fange, D., Ćurić, V., Lawson, M. J., and Elf, J. (2017) Kinetics of dCas9 target search in *Escherichia coli*. *Science* 357, 1420–1424.
- (5) Brendel, V., and Perelson, A. S. (1993) Quantitative Model of ColE1 Plasmid Copy Number Control. *Journal of Molecular Biology* 229.
- (6) Gillespie, D. T. (1977) Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry* 81, 2340–2361.
- (7) Shao, B., Rammohan, J., Anderson, D. A., Alperovich, N., Ross, D., and Voigt, C. A. (2021) Single-cell measurement of plasmid copy number and promoter activity. *Nature Communications* 1475.
- (8) Clamons, S. E., and Murray, R. M. (2017) Modeling Dynamic Transcriptional Circuits with CRISPRi. *bioRxiv*
- (9) Santos-Moreno, J., Tasiudi, E., Stelling, J., and Schaerli, Y. (2020) Multistable and dynamic CRISPRi-based synthetic circuits. *Nature Communications* 11.

- (10) Hsiao, V., Hori, Y., Rothmund, P. W., and Murray, R. M. (2016) A population-based temporal logic gate for timing and recording chemical events. *Molecular Systems Biology* 869.