Mutations in artificial self-replicating tiles: A step toward Darwinian evolution

Feng Zhoua,1, Ruojie Shab, Heng Ni1, Nadrian Seemanb,1, and Paul Chaikina,1

aDepartment of Physics, New York University, New York, NY 10003; and bDepartment of Chemistry, New York University, New York, NY 10003

Published December 6, 2021.

Published by the National Academy of Sciences. This article is a PNAS Direct Submission. O.G. is a guest editor invited by the Editorial Board. The authors declare no competing interest.

Artificial self-replication and exponential growth holds the promise of gaining a better understanding of fundamental processes in nature but also of evolving new materials and devices with useful properties. A system of DNA origami dimers has been shown to exhibit exponential growth and selection. Here we introduce mutation and growth advantages to study the possibility of Darwinian-like evolution. We seed and grow one dimer species, AB, from A and B monomers that doubles in each cycle. A similar species from C and D monomers can replicate at a controlled growth rate of two or four per cycle but is unseeded. Introducing a small mutation rate so that AB parents infrequently template CD offspring we show experimentally that the CD species can take over the system in approximately six generations in an advantageous environment. This demonstration opens the door to the use of evolution in materials design.

Significance

In nature, mutation is the first step of evolution, where it provides the genetic variation for the natural selection to act. Here we take a system of artificial self-replicating tiles, DNA origami, that exhibit templated reproduction. We can generate a small fraction of mutations by introducing a mismatch in hybridization between parent and daughter. We can modify the origami functionality to affect the growth rate of the mutated species, giving it less or more evolutionary advantage, and to become dominant in several generations. The introduction of mutations into an artificial self-replicating system provides new directions for research into self-assembly processes.
in living systems. In the present case a low mutation rate is particularly important in that the forward and reverse mutations are equally limiting the final ratio of the species with high and low growth advantage.

**Results**

The present mutation and evolution experiments are based on the established self-replication procedure illustrated in Fig. 1. We use cross-tile DNA origami with perpendicularly organized DNA helices along the two axes and a characteristic thicker "equal sign" easily observable in an atomic force microscope (AFM) where the two perpendicular halves are joined Fig. 1A (7). Reversible "recognition" sticky ends, single-stranded DNA, containing seven nucleotides in each strand, were placed on the vertical arms for temperature-dependent reversible hybridization. Four photo-cross-linkable sticky ends containing 3-cyanovinylcarbazole (CNVK) were placed on one of the horizontal arms for permanently linking complementary A and B cross-tile origamis at a lower temperature (23). The replication process, Fig. 1B, proceeds by introduction of parent dimer AB seeds into a suspension of single A and B tiles, lowering the temperature so that A and B tiles assemble above or below AB tiles. The temperature is set such that A and B tiles will not hybridize in suspension. However, when they are attached to the dimer template the y-y' or z-z' sticky ends are held in close proximity and their local concentration is increased by $\sim 10^6$ and they hybridize. While held cold the sample is illuminated with ultraviolet (UV) light (365 nm) and the horizontal sticky ends are covalently bound by

![Figure 1](https://doi.org/10.1073/pnas.2111193118)

**Fig. 1.** (A) Schematics of DNA origami tiles for ladders/self-replication system. The seed A-B dimer is held together by six 11-bp complementary horizontal sticky ends (S and S'). Both A-B and C-D monomers have a unique set of horizontal sticky ends (Y and its complement Y' for A-B, and Z and its complement Z' for C-D), where there is a CNVK (denoted X in the Y' and Z' sequences) unit on each of the four strands in Y' and Z' (red lines), so no crossover reaction can happen between A-D or B-C. The seed and monomer A-B contains two sets of six vertical sticky ends ($\alpha$, $\beta$) that extend from the top of the tiles and another two sets of complementary sticky ends ($\alpha'$, $\beta'$) from the bottom of the tiles. The monomer C-D contains two sets of complementary vertical sticky ends ($\gamma$, $\delta$). $\alpha$, $\beta$ differ from $\gamma$, $\delta$ by a few nucleotides in the sequence design, so there is a small chance for self-assembly of A and C or B and D. Schematics of self-replication of (B) seed A-B using A and B monomer and (C) seed C-D using C and D monomer. The general replication cycle contains three steps: 1) annealing at 22°C for the seed-templated self-assembly; 2) exposure to a 365-nm UV LED at 22°C; and 3) heating to 48°C to break the vertical sticky end hybridization. The monomers hybridized on one side or two sides of the seed dimer (template) can be photochemically cross-linked to form new dimers. (Inset) The AFM images of the dimers show the orientation of "equal sign" to distinguish between the two species. (D) Mutation from A-B dimer to C-D dimer during the self-replication. The system starts with an A-B dimer and monomers A, B, C, and D. Mutation in the self-assembly can lead to change of the species during the self-replication.
photo-cross-linking of CNVK to form the new dimer structure. On heating the daughter dimers are released and form the parent templates for the next generations. On repeated cycles of cooling, UV illumination, and heating the number of offspring dimers grows exponentially. The growth rate is determined both by the thermodynamics of the recognition strands and by the relative concentrations of the monomers and parent dimers. If only half of the monomer vertical arms are functionalized with sticky ends, then the replication rate is less than a factor of 2 per cycle (SI Appendix, Fig. S1A). However, if both vertical arms are functionalized the parent dimer can nucleate a one-dimensional ladder and in one cooldown, UV cross-linking, and release we can have a litter of offspring and an exponential growth of up to a factor of ~10 per cycle (20) (SI Appendix, Fig. S1C).

For the mutation experiments we introduced a new set of self-replicating tiles C and D, with 90° rotation of the tiles and their AFM identifying equal signs, Fig. 1C. C and D vertical and horizontal arms are functionalized with different sequences than A and B tiles (Fig. 1C). A and C have similar complementary vertical strands, but with a small difference at the end of each strand. Thus, A has a small chance to bind with C, instead of itself, and vice versa. Similarly, B and D have a small probability to bind as compared to B/B (B/B implies B over B) or D/D. In order to simplify the mutation results, the two dimer species share the same CNVK horizontal sticky ends design but distinct DNA sequences, preventing the photo-cross-linking of AD or BC dimers. As illustrated in Fig. 1D, If the system starts with only the AB parent dimer, the AB nucleated ladder would mostly have AB dimer rungs but may contain a small portion of CD rungs. After UV exposure and heating, this produces many new AB dimers and a few “mutated” CD dimers. Note that AD and CB dimers are prohibited by the very noncomplementary cross-linking strands. We can follow the evolution of the dimer species by monitoring the change of ratio between the AB and CD dimers from AFM images, as we repeat replication cycles. We refer to CD dimers as mutations as the sample is seeded only with AB. CD results from unfavorable binding of mismatched DNA sticky ends on the vertical arms.

In order to control the mutation rate between AB and CD, we introduced mismatches onto one of the recognition sticky ends that bind monomers to the parent template. Normally β-strands on the top of B bind to complementary β-strands on the bottom of another B. However, if the δ-strands on the bottom of D are almost complementary to β then D will bind to B but with a lower probability than B with B. As shown in Fig. 2A, we introduced single mismatches in several recognition vertical sticky ends on D which is mutated from B. We then quantified the effect of changing the number of mismatched recognition strands on the B/D binding using fluorescence resonance energy transfer (FRET). As shown in Fig. 2B, the B monomer consisting of the regular DNA sticky ends was hybridizing with the D monomer consisting of the mutated complementary DNA sticky ends with the mismatches. The ε3 on B and ε5 on D form a FRET pair after hybridization and cause the decrease of the emitted light intensity of the ε3 compared to the free-diffusing ε3 in the control sample (SI Appendix, Fig. S2). By analyzing the change of the emission light intensity, we can plot the change of the FRET pair composition as temperature is changed, obtaining a melting curve (Fig. 2B). The melting curve shifts to lower temperature as we increased the number of mismatched strands in the D monomer. When the fifth mismatch is introduced, the hybridization starts at 23 °C. Since the photo-cross-linking is conducted at 22 °C in our replication process, we study the fraction of B/D dimers for one to four mismatch recognition sticky ends.

We mixed B monomers with two different monomers: B’ monomers with 0 mismatch and the D monomers with zero to four mismatches. During the annealing procedure from 48 °C to 20 °C, B monomers have a higher probability to hybridize with B’ and to form B/B dimers because of their higher melting temperature, compared to the D that forms the mutated dimer B/D (Fig. 2C). With the benefit of the equal-sign design, the regular B/B’ binding could be recognized with the parallel equal sign “||,” while the mutated B/D binding had the perpendicular equal signs “=||.” The equal-sign orientation was observed on the AFM, and we measured the hybridization selection by counting the two species after annealing (SI Appendix, Figs. S3 and S4). We found that the fraction of mutated BD binding, “=||,” would decrease with more mismatches introduced. We observed a continuous decrease in the fraction of the mutated B/D binding when more mismatches were introduced in the D monomer. With mutated monomers with three mismatches, we obtain 16% mutated binding, B/D ÷ (B/D + B/B) = 0.16. Similarly, for A/C the mutation rate is 16%. Considering that the specific sequence of the CNVK sticky ends only permits the photo-cross-linking between two regular monomers or two mutated monomers, the mutation rate leading to CD dimers is ~0.16×(0.162 + 0.842)~ 3.5% for three mismatches in the vertical sticky ends. In the following mutation self-replication experiments, the design for the mutated recognition strands contains three mismatches.

In Fig. 3 we show the results of a mutation experiment in which the growth rates of the two species AB and CD are approximately equal at 3.7 and 3.9 per cycle. In our litter/ladder self-replication systems the growth rate is controlled by the ratio of the number of AB dimers to the number of A and B monomers in each cycle and the number of CD dimers to C and D monomers per cycle. After each cycle the system is divided into two batches; one-third are used to seed the next generation and two-thirds are used for AFM analysis of the fractions of each species and the number of monomers. The mutation experiments were conducted with starting templates of AB dimers and a pool of the AB:A:B:C:CD monomers at a concentration of 0.1:1.6:1.6:1.6:1.6 nM (Fig. 3B). The refreshed monomers (AB:BC:CD = 1.6:1.6:1.6:1.6 nM) was repleted into the system after each self-replication cycle. We conducted 15 cycles of self-replication and tracked the evolution of the AB equal sign “||” composition using the AFM. Mutated the parent at the beginning of self-replication experiments, as mutated CD dimers were observed after the first cycle (Fig. 3E and SI Appendix, Fig. S5). The fraction of CD dimers was small at the beginning, but we observed more and more CD dimers through 15 cycles of self-replication, and the difference in the ratio of the two species got smaller (Fig. 3D). We further estimated the dimer composition after self-replication using the first-order master equations (SI Appendix, Eqs. S1 and S2) with the initial composition, the mutation rate, and the same growth rate for both species. The calculated fraction of the two species fits well with the experimental data (Fig. 3D). Analytically, if two species have the same mutation rate, one to the other, then the two species would eventually equalize if they have the same competitiveness, the growth rate.

We conducted a similar self-replication using two species with different growth rates. The mechanism of mutation and photo-cross-linking was the same between AB and CD dimers but the competitiveness in the self-replication growth rate was largely determined by the recognition strands design. As illustrated in Fig. 4A, we changed the recognition strands design of AB dimer to the one-sided design with the maximum replication rate of 2. Thus, species AB has a lower growth rate than species CD, which still forms ladders and produces litters. We conducted the mutation experiments with the AB dimer as the only seed template and A, B, C, and D monomers at a concentration of 0.1:1.6:1.6:1.6 nM (Fig. 4B). Fresh monomers (A:B:C:D = 1.6:1.6:1.6:1.6 nM) were replenished after every
self-replication cycle. The agarose gel showed that the dimers increased every cycle and the growth rate increased from 1.6 to 2.5 after 13 replication cycles, indicating that the mutated CD dimers with the higher growth rate dominated the system (SI Appendix, Table S3). On the AFM images the mutant CD dimer fraction increased after each cycle (Fig. 4E and F and SI Appendix, Fig. S7). In Fig. 4C and D we plot the fraction of CD dimers and the amplification of AB and CD dimers vs. the number of cycles. Starting from zero the mutated CD dimers reached the same level as the AB dimers within only five cycles of self-replication. The mutants took over the system afterward and kept increasing their proportion until reaching a plateau after nine cycles (Fig. 4C).

The experimental data fit well with simulation results. Based on the self-replication rate from the gel analysis (SI Appendix, Table S4) and the experimental composition of two species (Fig. 4C), we calculated the amount of the AB and CD dimers after each cycle of self-replication. The logarithm of the AB and CD dimers can be well-fitted to a linear curve within the first eight cycles of self-replication, indicating the exponential growth of the two species even though their continued a small mutation rate from AB to CD and CD to AB. Since the growth rate is the only competitiveness in this artificial self-replication, the mutated CD dimers with a higher growth rate eventually dominated the system independent of the initial seed. As our system approaches saturation, the experiment deviates from the simple prediction. We...
attribute this difference to accumulation of parasite, sterile, non-templated dimers which do not replicate according to our previous experiments (19). We allow for self-replication by adding monomers each cycle proportionate to the parent dimers and their growth rate. As shown in SI Appendix, Fig. S8, a nontemplated binding rate of 0.0007 dimers/monomer cycle well accounts for the experimental data in Fig. 4C.

**Conclusion**

We have developed an artificial system of DNA origami tiles of two species in which we can control the growth rates separately. Adding the ability of one species to mutate into the other, we have studied the evolution of the system where only one species is seeded. When growth rates are equal the system evolves to a steady state of equal populations. When one has the competitive advantage of faster growth it quickly becomes the dominant species, even when it only results from a mutation from the originally seeded and exponentially growing species. This is the expected result and a most elementary example of Darwinian evolution but here in an artificial self-replication system. Our system can be readily generalized to have many different types of mutations by adding E, F, G, H ... monomer tiles with different sticky end sequences and with different growth rates depending on their rigidity, structure, and sensitivity to pH, ionic concentration, and other physical and chemical properties of their environment. It opens the door to the use of human-made systems, devices, and materials that evolve to have desired properties. In a given environment mutations allow the creation of a set of species and evolution picks the species which grows fastest in that environment, mimicking nature but with artificial constructs.

**Materials and Methods**

**DNA Synthesis and Purification.** Scaffold DNA M13mp18 was purchased from Bayou Biolabs. CNVK phosphoramidite was purchased from Glen Research. The DNA strands containing the CNVK were synthesized on an Applied Biosystems
394 DNA synthesizer. The remaining DNA strands were purchased from Integrated DNA Technology, Inc. The sticky-end strands were purified using denaturing polyacrylamide gel electrophoresis. The staple strands with dyes (cy3 and cy5) were purified by IDT using high-performance liquid chromatography.

Formation of Individual DNA Origami Tiles. Staple strands and M13mp18 scaffold DNA were mixed in 1× TAE/Mg2+ buffer (40 mM Tris/C1HCl, 20 mM acetic acid, 2.5 mM ethylenediaminetetraacetic acid, and 12.5 mM magnesium acetate, pH 8.0). The molar ratio of scaffold:staple strand was 1:8, and the final concentration of the scaffold DNA was 10 nM. The mixture was heated to 70°C for 20 min and cooled to 20°C at a rate of 0.7°C/h in a thermocycling incubator. The resulting origami tiles were purified in 100K Millipore Amicon Ultra 0.5-mL centrifugal filters at 40°C five times to remove the excess staple DNA strands. The weight concentration of the DNA origami was measured by IMPLEN NanoPhotometer Pearl using the double strands mode at 260 nm. The molar concentration was calculated by dividing the weight concentration over the molar mass of DNA origami (around 7,500 base pairs [bp] × 650 Da/bp).

Formation of Seed Dimer. For the three types of seed dimer origami tiles, the monomers of the same species were preheated to 53°C for 30 min and then mixed at a 1:1 ratio. The mixture was cooled to 20°C at a rate of −0.7°C/h.

Self-Replication Cycling. Thirty microliters of solution containing seed dimers and monomer tiles was prepared at the following concentrations (in nanomolar):

1) Fig. 3C: seed AB:A:B = 0.1:3.2:3.2; seed CD:C:D = 0.1:3.2:3.2
2) Fig. 3D: seed AB:A:B:CD:C:D = 0.1:1.6:1.6:1.6:1.6
3) Fig. 4: seed AB:A1:B1:A2:B2:CD:C:D = 0.1:1.6:1.6:1.6:1.6:1.6:1.6
4) SI Appendix, Fig. S1A: seed AB:A1B1A2B2 = 0.4:1.6:1.6:1.6:1.2:1.2
5) SI Appendix, Fig. S1C: seed AB:A1B1A2B2 = 0.2:3.2:3.2:0.1:3.2:3.2, and 0.025:3.2:3.2, respectively.

A 10-μL sample was taken out to quantify the initial dimer amount and the rest of the solution was covered with 50 μL silicone oil (HR3-415: Hampton Research) to prevent the evaporation-induced concentration change. The solution was heated to 48°C for 30 min to break the vertical sticky-end hybridization and then annealed at 22°C for 1 h for the seed-templated self-assembly. The sample was exposed to a 365-nm UV light-emitting diode (LED) (M365LP1c, power intensity 24 mW/mm2; Thorlabs Inc.) at 22°C for 20 min. A 10-μL sample was then extracted for AFM imaging and nondenaturing 0.8% agarose gel electrophoresis to determine the resulting amplified dimer amount. For the multicycle self-replication, an additional 20 μL of fresh monomers of the same concentration was transferred into the original tube before the next cycle of annealing and photo-cross-linking.
Gel Electrophoresis. A 0.8% agarose gel was prepared and immersed in the 1× TAE/MgCl₂ buffer inside the gel electrophoresis box. Five-microliter samples were mixed with 1 μL nondenaturing tracking dye (1× TAE/MgCl₂ buffer, 50% glycerol, and trace amount of bromophenol blue and Xylene Cyanol FF). Gel was run at 4 to 5 V/cm at 48°C for 3 h. The gel was immersed in the ethidium bromide solution for staining and rinsed with water to remove excess stain. The fluorescent gel image was processed using ImageJ to quantify the amount of the dimers and monomers to calculate the dimer percentage (dimer%); the growth rate of each self-replication cycle was calculated by dividing the dimer% after UV cross-linking over the dimer% before UV cross-linking. When there is no monomer replenishment, the dimer% before and after UV would be marked as cycle 0 and cycle 1, and the growth rate was calculated as dimer% (cycle1)/dimer% (cycle0).

AFM Imaging. AFM imaging was performed in peakforce mode in air for the self-replication cycling. Two to five microliters of a diluted sample was deposited on a clean mica surface (Ted Pella, Inc.) for 1 min. Unless specified, the sample and mica were incubated at 48°C for 30 min and the deposition was conducted at 48°C. The mica was then washed with 50 to 70 μL of double-distilled H₂O three times and subsequently dried using N₂. PeakForce in air mode was performed on the Nanoscope V Multimode 8 scanning probe microscope (PeakForce QNM Software, ScanAsyst-HR accessory). Silicon nitride tips (ScanAsyst-Air; Bruker Nano, Inc.) were used for the Nanoscope V Multimode 8 SPM.

FRET Experiment. The FRET experiment was conducted using a Horiba PTI QuantaMaster 400 Fluorescent Spectrometer. Cy3 (donor) and Cy5 (acceptor) dye was modified on the vertical and horizontal sticky ends to make B-cy3 (donor), D-cy5 (acceptor), and cy3-FG-cy5 (donor acceptor). The concentration of B and D was 3.2 nM for the mutated vertical sticky end hybridization (Fig. 28); the concentration of seed was 1.6 nM and the B-cy3/D-cy5 tile was 3.2 nM for the 5'NH₂-horizontal sticky-end hybridization (SI Appendix, Fig. 59). The sample solution was covered with silicone oil to prevent evaporation. For the thermodynamic study of vertical sticky-end hybridization (Fig. 28 and SI Appendix, Fig. 52), the sample was annealed from 20°C to 48°C, then back to 20°C for three cycles. The ramping rate was 3 min°C, and the fluorescent signal was measured at each degree. For the study of horizontal sticky end hybridization (SI Appendix, Fig. 59), the annealing temperature range was from 10 to 40°C. The calculation method of the hybridization percentage can be found in our previous report (20).

Data Availability. All study data are included in the article and/or SI Appendix.

ACKNOWLEDGMENTS. This research has been primarily supported by Department of Energy (DOE) DE-SC0007991 to P.C., N.S., R.S., F.Z., and H.N. for initiation, design, analysis, and imaging. P.C. acknowledges support from the Center for Bio-Inspired Energy Sciences, an Energy Frontier Research Center funded by the DOE, Office of Sciences, Basic Energy Sciences, under Award DE-SC0000989, for initiation, DNA sequence design, preparation, and characterization confocal microscopy. R.S. and N.S. acknowledge partial support from NSF EFRI-1332411 and CCF-1526650 for laboratory supplies, under Award DMR-1420073 for synthesis and characterization of the DNA origami. R.S. and N.S acknowledge MURI W911NF-11-1-0024 from the Army Research Office and MURI N00014091118 from the Office of Naval Research for partial salary support. R.S. and N.S. acknowledge partial support from RGP0010/2017 from Human Frontiers Science Program. R.S. and N.S. acknowledge partial support from DOE DE-SC0007991 for DNA synthesis and partial salary support. The authors are grateful for shared facilities provided through the MRSEC program of the NSF under Award DMR-1420073.