Spontaneous Emergence of Self-Replicating Molecules Containing Nucleobases and Amino Acids

Liu, Bin; Pappas, Charalampos G; Ottelé, Jim; Schaeffer, Gael; Jurissek, Christoph; Pieters, Priscilla F; Altay, Meniz; Marić, Ivana; Stuart, Marc C A; Otto, Sijbren

Published in:
Journal of the American Chemical Society

DOI:
10.1021/jacs.9b10796

Publication date:
2020

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JACS

Article

Spontaneous Emergence of Self-Replicating Molecules Containing Nucleobases and Amino Acids

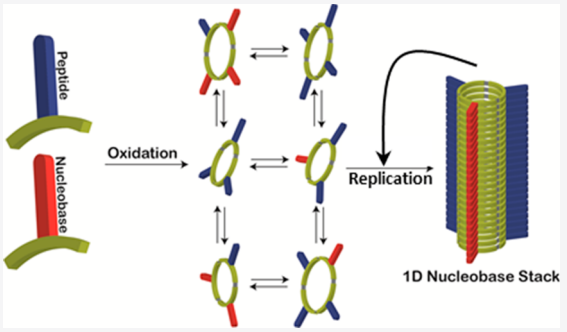
Bin Liu,[†] Charalampos G. Pappas,[†] Jim Ottelé, Gaël Schaeffer, Christoph Jurissek, Priscilla F. Pieters, Meniz Altay, Ivana Marić, Marc C. A. Stuart, and Sijbren Otto*

Cite This: J. Am. Chem. Soc. 2020, 142, 4184–4192

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The conditions that led to the formation of the first organisms and the ways that life originates from a lifeless chemical soup are poorly understood. The recent hypothesis of "RNA-peptide coevolution" suggests that the current close relationship between amino acids and nucleobases may well have extended to the origin of life. We now show how the interplay between these compound classes can give rise to new self-replicating molecules using a dynamic combinatorial approach. We report two strategies for the fabrication of chimeric amino acid/nucleobase self-replicating macrocycles capable of exponential growth. The first one relies on mixing nucleobase- and peptide-based building blocks, where the ligation of these two gives rise to highly specific chimeric ring structures. The second one starts from peptide nucleic acid (PNA) building blocks in which nucleobases are already linked to amino acids from the start. While previously reported nucleic acid-based self-replicating systems rely on presynthesis of (short) oligonucleotide sequences, self-replication in the present systems start from units containing only a single nucleobase. Self-replication is accompanied by self-assembly, spontaneously giving rise to an ordered one-dimensional arrangement of nucleobase nanostructures.



INTRODUCTION

Establishing possible pathways through which life can emerge from inanimate matter is one of the grand challenges in today's science.[1−3] In addressing this challenge the functional and structural characteristics of present-day life provide important guidance. At the same time, the overwhelming complexity of evolved life makes it challenging to extract its essence and identify pathways for its emergence.

Recently, a systems chemistry[4,5] view toward the challenges of the origins and synthesis of life is gaining popularity. The facts that many different types of molecules coexisted at the time of life's origin and that the same applies to present-day life, warrants an exploration of what may emerge upon allowing different compound classes to interact. For example, the notion of peptides and nucleic acids cooperating during the early stages of the emergence of life is becoming increasingly popular[6−9] and nucleobase-peptide chimera show unique self-assembly behavior[10−12] and can give rise to remarkably complex foldamers.[13]

Key in the transition of chemistry into biology is the acquisition of function. The core functional characteristics of life are its ability to replicate, to metabolize, and to be spatially segregated from its environment. Where life requires the functional integration of all of these characteristics, most research efforts still focus on one of these aspects in isolation.

Autocatalysis, the ability of systems (molecules, metabolic networks or compartments) to make copies of themselves, is central to all evolutionary scenarios.[14−18] Systems where autocatalysis is accompanied by information transfer and heredity are said to be self-replicating. Synthetic systems of self-replicators have been pioneered by von Kiedrowski using short DNA strands.[19] Subsequently, self-replicating molecules have been developed that feature most of the other important current biopolymers (i.e., RNA[20−24] and peptides[25−28]) as well as completely synthetic molecules.[29−33]

A major issue in replicator chemistry is the tendency for self-inhibition through replicator duplex formation. This causes many replicators to exhibit only parabolic growth (i.e., showing a kinetic order in replicator of 1/2) whereas exponential growth (first order in replicator) would be necessary for most scenarios of Darwinian evolution.[34−37] Another problem in replicator chemistry is the complexity of the structures associated with most self-replicators, which are unlikely to emerge spontaneously from simple starting materials.

Received: October 8, 2019
Published: February 5, 2020

https://dx.doi.org/10.1021/jacs.9b10796
J. Am. Chem. Soc. 2020, 142, 4184–4192

We recently introduced a new approach to self-replication that addresses both of these problems simultaneously.[38−40] This approach relies on (i) the creation of a mixture of molecules that continuously interconvert (a dynamic combinatorial library or DCL) and (ii) a self-assembly process that leads to the sequestration of molecules from this mixture, which subsequently get replenished. The combination of these two features is sufficient for the spontaneous and autocatalytic formation of self-replicating molecules. Given that networks of interconverting molecules and self-assembly processes are likely to have been widespread in prebiotic environments, this mechanism provides a likely path for the spontaneous emergence of replicators. Note that the building blocks that give rise to the network of interconverting molecules can be relatively simple,[41] while the structure of the emerging replicators can be relatively complex, consisting of many different building blocks connected in a way that is not a priori specified.[42] Furthermore, exponential replication is possible upon entering a growth-breakage cycle, in which mechanical energy is utilized to break replicator assemblies exposing more edges from which the assemblies grow.[43]

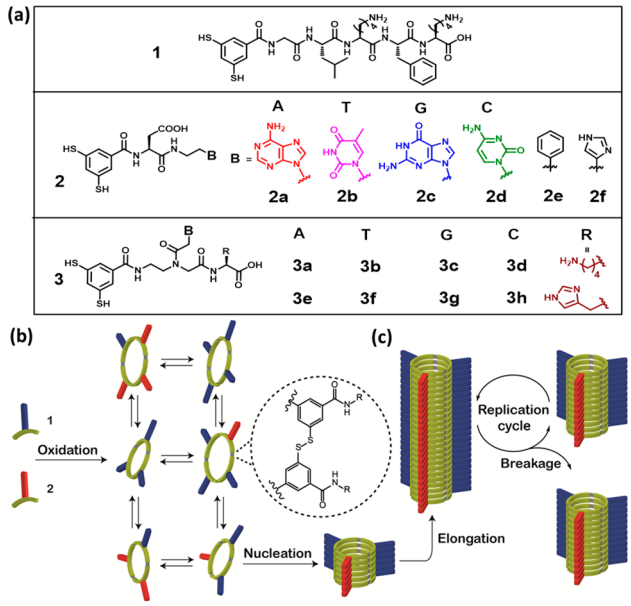
A systems approach to the emergence of self-replicating molecules, where different compound classes (i.e., amino acids, peptides, and nucleobases) coexist has thus far received only little attention. Efforts directed at PNA-based replicators, in which an amino acid replaces the phosphate-sugar backbone of DNA/RNA, come closest.[44−46] However, PNA remains very similar to DNA/RNA in architecture and behavior.

We now report the spontaneous emergence of new self-replicating molecules from molecular networks in which nucleobases and amino acids are both present. We show that this leads to chimeric replicators, which rely on the assembly of peptides and nucleobases into fibrous aggregates (but do not rely on base-pairing) resulting in the autocatalytic formation of a one-dimensional arrangement of nucleobases. The two different systems constructed herein allow for a direct comparison between replicator mutations. The peptide-nucleobase system shows that mutations are easily accommodated during replication, while in the PNA system, replicator mutation is impeded as it requires a change in ring size. While the building blocks used were not selected for prebiotic relevance, they do illustrate the potential of the assembly driven replication mechanism that might well extend to other types of molecules.

RESULTS AND DISCUSSION

We have constructed molecular networks based on thiol−disulfide chemistry.[47] Three families of relatively simple dithiol building blocks were prepared featuring peptides (**1**; Scheme 1a) or amino-acid nucleobase conjugates (**2** and **3**). Oxidizing (mixtures of) these dithiols generates DCLs of macrocyclic disulfides with different (compositions and) ring sizes, which interconvert through reaction of the disulfides with residual thiolate anion (Scheme 1b). We previously reported that DCLs made from adenine-containing building block **2a** were dominated by a foldamer, consisting of 15 subunits of **2a**, accompanied by small amounts of trimers and tetramers (Figure 1a).[13] Analysis of DCLs prepared from analogous building blocks **2b**, **2c**, and **2d**, containing thymine, guanine, and cytosine, respectively, led to similar conclusions. Specifically, in the libraries consisting of **2c** and **2d**, most of the building blocks are converted into the folded 15mer (see the Supporting Information, Figures S34a and S44a).

Scheme 1. (a) Building Block Structures; (b) Oxidation of These Dithiol Building Blocks Gives Rise to Dynamic Combinatorial Libraries of Macrocyclic Disulfides; (c) Assembly of a Specific Ring Size into Stacks Leads to the Autocatalytic Formation of More of These Rings through a Fiber Growth-Breakage Cycle



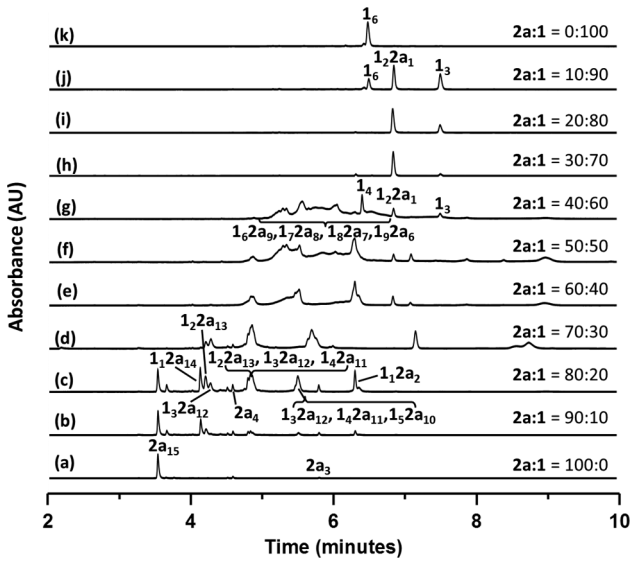


Figure 1. UPLC-MS analysis of DCLs made from different ratios of building blocks **1** and **2a** at a total building block concentration of 2.0 mM in 50 mM borate buffer pH 8.2 stirred at 1200 rpm for 25 days. From bottom (a) 100 mol % of **2a** to top (k) 100 mol % of **1**.

The library made from **2b** contains, besides 15mer, a significant amount of trimer and tetramer macrocycles (see the Supporting Information, Figure S21a). In contrast, DCLs made from peptide-functionalized building block **1** readily produces self-replicating hexamers, partially driven by the assembly of the peptide side groups into β-sheets.[48] The mechanisms by which **1** replicates is analogous to that shown in Scheme 1c and bears similarities to amyloid formation.[43,49]

We hypothesized that mixing building blocks **1** and **2** might lead to the formation of mixed macrocycles that retain the

https://dx.doi.org/10.1021/jacs.9b10796
J. Am. Chem. Soc. 2020, 142, 4184−4192

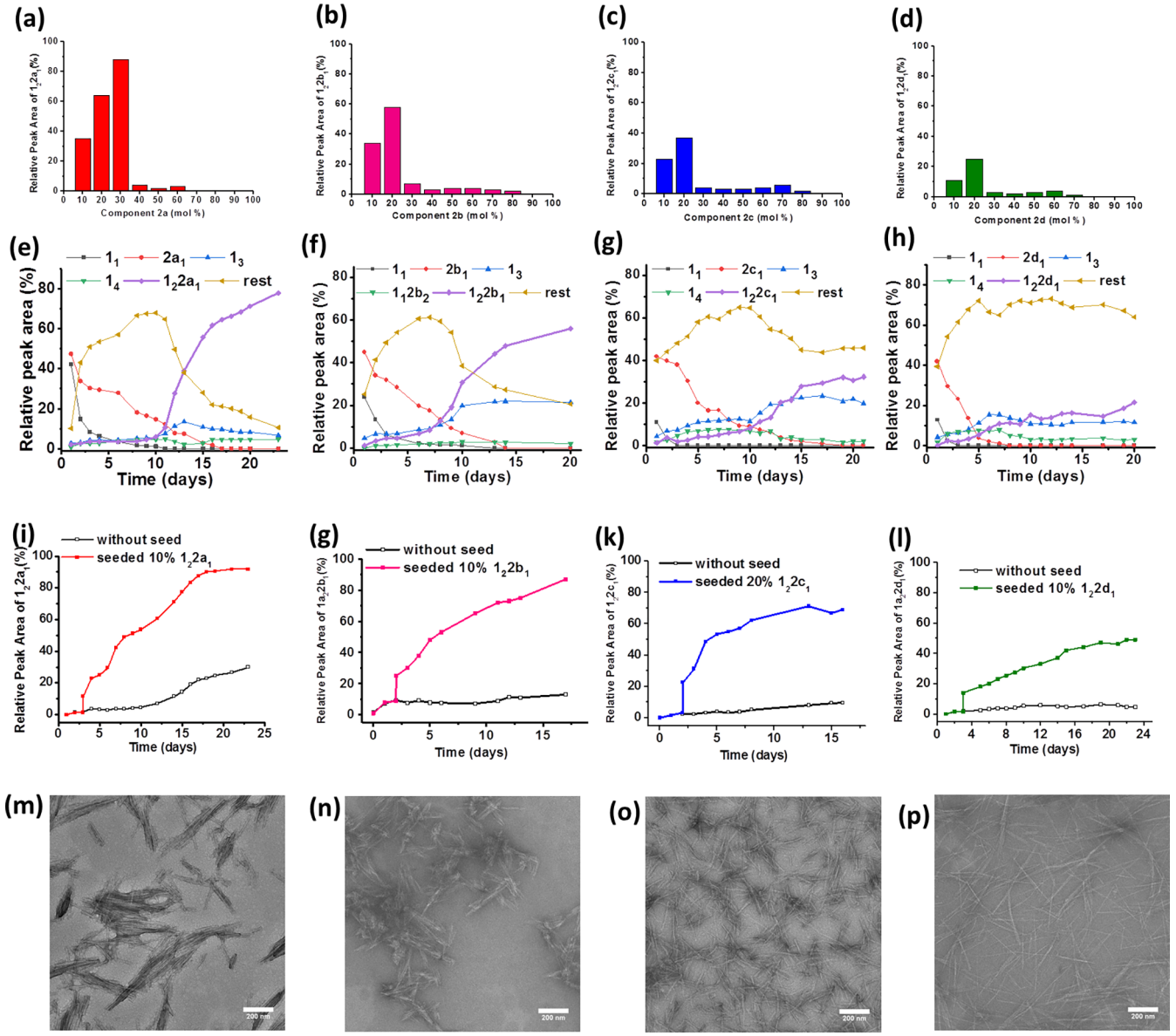


Figure 2. Fraction of trimers (a) $\mathbf{1}_2(\mathbf{2a})_1$, (b) $\mathbf{1}_2(\mathbf{2b})_1$, (c) $\mathbf{1}_2(\mathbf{2c})_1$, and (d) $\mathbf{1}_2(\mathbf{2d})_1$ obtained in DCLs made from a mixture of building blocks **1** and **2** (2.0 mM total in 50 mM borate buffer at pH 8.2 stirred at 1200 rpm) at different ratios. Change in product distribution with time in DCLs made from (e) **1** (1.4 mM) and **2a** (0.6 mM), (f) **1** (1.6 mM) and **2b** (0.4 mM), (g) **2c** (0.4 mM) and (h) **2d** (0.4 mM). Kinetics of formation of mixed trimers (i) $\mathbf{1}_2(\mathbf{2a})_1$, (j) $\mathbf{1}_2(\mathbf{2b})_1$, (k) $\mathbf{1}_2(\mathbf{2c})_1$, and (l) $\mathbf{1}_2(\mathbf{2d})_1$ in DCLs made from **1** (1.0 mM) and **2** (0.5 mM) in the absence (open symbols) and presence (closed symbols) of 10 or 20 mol % of the corresponding preformed trimer (added at day 3). Seeding experiments were conducted by preparing one mother solution, splitting it in two and adding seed to one of these. Negative staining TEM images of the assemblies formed in DCLs dominated by (m) $\mathbf{1}_2(\mathbf{2a})_1$, (n) $\mathbf{1}_2(\mathbf{2b})_1$, (o) $\mathbf{1}_2(\mathbf{2c})_1$, and (p) $\mathbf{1}_2(\mathbf{2d})_1$. For additional TEM images, see Figures S146–S149.

ability of the peptides to form β-sheets, while also featuring nucleobases. To test this hypothesis, we dissolved peptide building block **1** and adenine containing building block **2a** (1.0 mM each) in 50 mM borate buffer (pH 8.2) in a capped vial and stirred the solution in the presence of air. The DCL reached a stationary distribution after stirring the solution for 3 weeks. Ultra-performance liquid chromatography mass spectrometry (UPLC-MS) analysis revealed that a rather complex and poorly resolved mixture was produced, composed of many mixed macrocycles from trimers up to cyclic 15mers (Figure 1f).

We then prepared similar DCLs at different building block ratios (**1**:**2a** = 90:10, 80:20, 70:30, 60:40, 50:50, 40:60, 30:70, 20:80, and 10:90; see Figure 1). Interestingly, the DCLs made using ratios of **1**:**2a** of 90:10, 80:20, and 70:30 were dominated by trimer $\mathbf{1}_2(\mathbf{2a})_1$. The highest yield of this trimer was observed at a ratio of 70:30 (Figure 2a). The almost complete absence of other macrocycles in the latter sample suggests that $\mathbf{1}_2(\mathbf{2a})_1$ selectively benefits from a special stabilizing effect (vide infra).

To further investigate the nature of the unusual stability of $\mathbf{1}_2(\mathbf{2a})_1$ we monitored the kinetics of its growth. Upon mixing **1** and **2a** a variety of different macrocycles were formed in the first 10 days (Figure 2e). However, after 10 days the rate of growth of $\mathbf{1}_2(\mathbf{2a})_1$ increased and it became the main product of the library, at the expense of the other macrocycles. The formation of $\mathbf{1}_2(\mathbf{2a})_1$ was highly depended on mechanical agitation. Nonagitated libraries gave only small amounts of

https://dx.doi.org/10.1021/jacs.9b10796
J. Am. Chem. Soc. 2020, 142, 4184−4192

$\mathbf{1}_2(\mathbf{2a})_1$ even after 20 days (Supporting Information, Figure S1). The sigmoidal growth of $\mathbf{1}_2(\mathbf{2a})_1$, together with the strong dependence on agitation suggests that this macrocycle is a self-replicator that grows through self-assembly.[13] The autocatalytic nature of the formation of $\mathbf{1}_2(\mathbf{2a})_1$ was confirmed in a seeding experiment. Upon adding 10 mol % of preformed $\mathbf{1}_2(\mathbf{2a})_1$ to a sample made from **1** and **2a** in a 2 to 1 ratio its growth dramatically accelerated (Figure 2i). The rate of emergence of $\mathbf{1}_2(\mathbf{2a})_1$ is dependent on temperature (see Supporting Information, Figure S159). Analysis by transmission electron microscopy (TEM) of the sample dominated by $\mathbf{1}_2(\mathbf{2a})_1$ confirmed the presence of fibrous aggregates of lengths ranging from 50 nm to 1 μm, corresponding to 100–2000 nucleobases (Figure 2m). In order to probe whether replication was exponential, rather than parabolic,[43] we determined the kinetic order in self-replicator $\mathbf{1}_2(\mathbf{2a})_1$ by seeding with different amounts of preformed replicator. The initial rate of replication was determined by UPLC analysis (see the Supporting Information, Figure S158 for the kinetic data). The slope of the plot gave a value of the order in replicator $\mathbf{1}_2(\mathbf{2a})_1$ of 1.24 ± 0.13, consistent with exponential growth (Figure 3). Taken together, these data suggest that

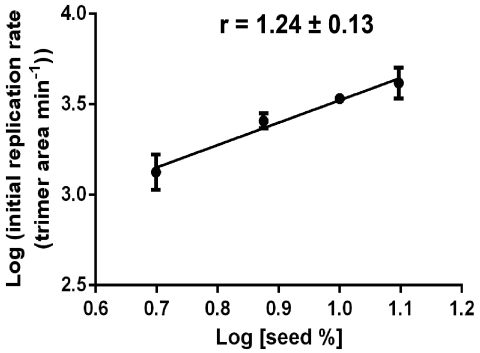


Figure 3. Determination of the order in replicator $\mathbf{1}_2(\mathbf{2a})_1$. The initial replication rate is plotted versus the concentration of replicator. The data points correspond to seeding concentrations of 5.0, 7.5, 10, and 12.5% relative to the stock solution (500 μM in building blocks). The error bars denote the standard deviation based on three individual measurements.

$\mathbf{1}_2(\mathbf{2a})_1$ self-replicates through the fiber growth-breakage mechanisms shown in Scheme 1c, similar to the mechanism we observed previously for replicators formed from only building block **1**.

Subsequently, analogous mixed DCLs made from peptide **1** and the different nucleobase building blocks **2b** – **2d** (containing thymine, guanine and cytosine, respectively) were also studied. The results show that in these systems $\mathbf{1}_2(\mathbf{2b})_1$, $\mathbf{1}_2(\mathbf{2c})_1$, and $\mathbf{1}_2(\mathbf{2d})_1$ can also be obtained (Figure 2b–d and f–h, Supporting Information, Figures S21, S34, and S44), albeit with somewhat reduced yields. Seeding experiments demonstrated that all of these mixed trimers are self-replicators (Figure 2g–l). As with $\mathbf{1}_2(\mathbf{2a})_1$, samples enriched in the mixed trimers all revealed fibrous aggregates when analyzed by TEM (Figure 2n–p). Circular dichroism (CD) spectra on samples dominated by $\mathbf{1}_2(\mathbf{2a})_1$ showed the bands that are indicative of β-sheets, suggesting that the peptide assembles in this type of secondary structure (see the Supporting Information, Figure S2). A less ordered chiral environment was observed for the other mixed trimer replicators.

The fidelity of replication of these systems is remarkable, particularly for the adenine replicator $\mathbf{1}_2(\mathbf{2a})_1$, which showed very little tendency to mutate into $\mathbf{1}(\mathbf{2a})_2$ or $\mathbf{1}_3$. Thus, mutating a nucleobase unit into a peptide unit or vice versa does not happen frequently. This spurred us to investigate the extent to which the nucleobase could be mutated to another nucleobase. We investigated this question through a series of cross-seeding experiments on libraries made from peptide building block **1** and nucleobase building block **2** mixed in the ratio optimal for producing the corresponding $\mathbf{1}_2(\mathbf{2a})_1$ replicator. After 3 days of stirring, these samples were seeded with 10 mol % of trimer replicators based on one of the other three nucleobases. As shown in Figure 4, all nucleobase

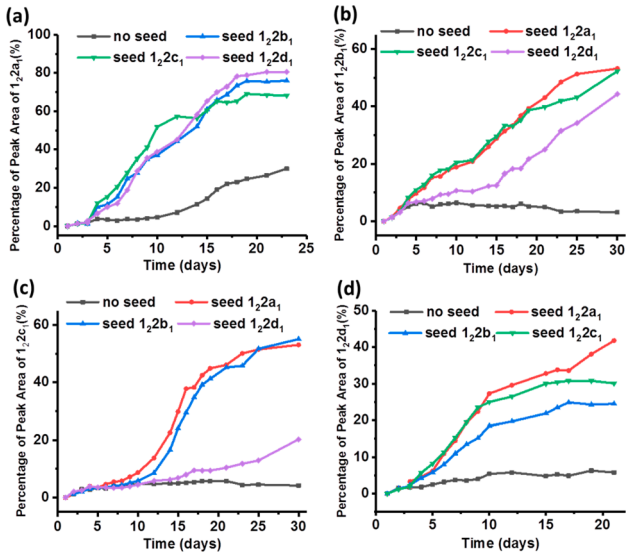


Figure 4. Effect of seeding on the growth of $\mathbf{1}_2\mathbf{2}_1$ replicators in DCLs prepared from **1** and (a) **2a**, (b) **2b**, (c) **2c**, and (d) **2d** in 2 to 1 ratio (1.5 mM total) in the presence of 10 mol % of seed containing $\mathbf{1}_2(\mathbf{2a})_1$ (red circles), $\mathbf{1}_2(\mathbf{2b})_1$ (blue triangles), $\mathbf{1}_2(\mathbf{2c})_1$ (green triangles), or $\mathbf{1}_2(\mathbf{2d})_1$ (magenta diamonds), which was added at day 3. Seeding experiments were conducted by preparing one mother solution, splitting it in four and adding seed to one of these.

replicators are able to cross-catalyze the formation of any other nucleobase replicator. In most cases the efficiency with which they do so is similar and also similar to the efficiency of autocatalysis upon self-seeding. Only the cytosine replicator $\mathbf{1}_2(\mathbf{2d})_1$ is less efficient at seeding other replicators and, in turn, also benefits less from being seeded by the other nucleobase replicators. Given the small energy differences involved in base stacking propensity of nucleobases in DNA,[50,51] it is hard to fully rationalize the cross-seeding effects.

We also performed analogous experiments on DCLs which contained all four nucleobases, which were seeded with one of the four $\mathbf{1}_2\mathbf{2}_1$ replicators. These experiments probe the replication fidelity of the individual nucleobase replicators. Unfortunately, we were unable to separate $\mathbf{1}_2(\mathbf{2a})_1$ and $\mathbf{1}_2(\mathbf{2d})_1$ in our UPLC analysis. Figure 5 shows that all seeds enhance the formation of the $\mathbf{1}_2\mathbf{2}_1$ replicators relative to their spontaneous emergence. However, the extent to which a specific nucleobase replicator is able to selectively enhance the formation of copies of itself seems limited and the incorporation of other nucleobases into trimers occurs readily.

Taken together, the seeding experiments of Figures 4 and 5 suggest that mutation of the nucleobase residue during

https://dx.doi.org/10.1021/jacs.9b10796
J. Am. Chem. Soc. 2020, 142, 4184–4192

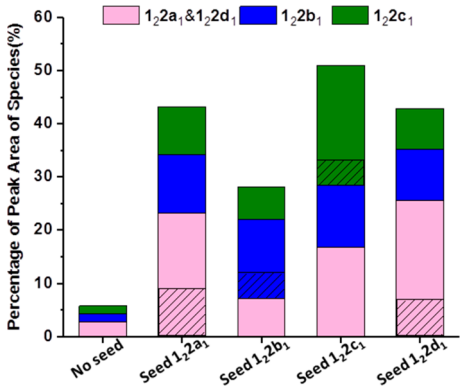


Figure 5. Effect of seeding on the distribution of $\mathbf{1}_2\mathbf{2}_1$ replicators in DCLs prepared from **1** (1.0 mM) and **2a**, **2b**, **2c**, and **2d** (0.125 mM each) in the presence of 10 mol % of seed containing $\mathbf{1}_2(\mathbf{2a})_1$, $\mathbf{1}_2(\mathbf{2b})_1$, $\mathbf{1}_2(\mathbf{2c})_1$, or $\mathbf{1}_2(\mathbf{2d})_1$. Due to co-elution, the amounts of $\mathbf{1}_2(\mathbf{2a})_1$ and $\mathbf{1}_2(\mathbf{2d})_1$ could not be separately quantified. The striped areas correspond to the amount of seed.

replication is facile. Hence, when different nucleobase building blocks are present during replicator growth, the resulting fibers will feature essentially random sequences of nucleobases along the fiber axis.

In order to probe the importance of the nucleobase motif for replication, the nucleobases were replaced by a simple phenyl ring. Thus, building block **2e** was synthesized, following a protocol analogous to that of the nucleobase building blocks. The DCLs that were prepared by mixing this building block in different ratios with peptide building block **1** did not provide any indications for the emergence of a self-replicator (see the Supporting Information, Figure S55). Additionally, cross-seeding a DCL prepared from **1** and **2e** in a 2:1 ratio with 10 mol % of any of the four nucleobase replicators also failed to promote the growth of $\mathbf{1}_2(\mathbf{2e})_1$ (see the Supporting Information, Figure S4). These results indicate that, in the context of the present system, the tendency of the canonical nucleobases to form replicators is superior to that of a simple phenyl analog.

In order to further investigate the importance of the amphiphilicty in the structures of the building blocks, nucleobases were replaced by a simple imidazole moiety to obtain building block **2f**. A series of DCLs were set up by mixing building blocks **2f** and **1** in different ratios. DCLs made using ratios of **1**:**2f** of 80:20, 70:30, 60:40 were dominated by replicator $\mathbf{1}_2(\mathbf{2f})_1$ (Figure S139), similar to what was previously observed for the mixture containing **1** and **2a**. Moreover, cross-seeding experiments demonstrated that the addition of replicator $\mathbf{1}_2(\mathbf{2a})_1$ can accelerate the growth of replicator $\mathbf{1}_2(\mathbf{2f})_1$ (Figure S145). These results suggest that the emergence of the mixed trimer replicator from the binary systems benefits from the presence of a hydrophilic aromatic ring in the monomeric units.

Encouraged by the results described above, we investigated whether we can also obtain self-replicating molecules that feature amino acid and nucleobase subunits without the need for peptide building block **1**. We changed our building block design and utilized the PNA motif, linked to an amino acid, while maintaining the aromatic dithiol unit, resulting in building blocks **3a**–**h**. The PNA unit carries one of the four canonical nucleobases, while the amino acid residue is L-lysine or L-histidine. These amino acids were chosen to have different charges: where the amine side chain of lysine is fully protonated at neutral pH, the imidazole ring of histidine is only partially protonated. These building blocks were synthesized using conventional Fmoc/tBu solid-phase peptide synthesis.

We prepared a set of DCLs by dissolving building blocks **3a**–**h** (3.8 mM) separately in 50 mM sodium borate buffer at pH 8.2 under continuous stirring (1200 rpm). Rapid partial oxidation (80% conversion of thiols to disulfides) of the resulting solutions was performed by adding sodium perborate solution (80 mM), followed by slower further oxidation mediated by oxygen present in the air.

We first explored the behavior of DCLs made from the lysine-containing building blocks **3a**–**d**. The corresponding DCL made from adenine building block **3a** rapidly became dominated by the cyclic tetramer, which constituted 85% of the library material after 30 min. Over time, the tetramer formed in quantitative yield (see Figure 6a), and the composition remained unchanged for the duration the sample was monitored (6 days). The DCL made from thymine building block **3b** behaved somewhat differently. Initially a mixture of cyclic trimer, tetramer, hexamer, and other oligomers was formed, which, after 2 days, gave way to the pentamer macrocycle (Figure 6b). Analogous experiments using cytosine building block **3d** gave rise to cyclic trimer (Figure 6c). Rather different behavior was observed for DCLs made from guanine building block **3c**, which contained a broad range of macrocycles up to 14mers (Figure S77). The composition remained unchanged for up to 30 days.

In order to investigate whether macrocycles $(\mathbf{3a})_4$, $(\mathbf{3b})_5$, and $(\mathbf{3d})_3$ are self-replicators, seeding experiments were conducted. UPLC analysis revealed that, upon addition of 10 mol % seed, the formation of these macrocycles is significantly faster compared to the nonseeded control samples (Figure 6d–f). TEM analysis of the samples dominated by these macrocycles revealed the presence of fibrillar assemblies (Figure 6g–i). CD spectra of the samples showed strong induced CD signals for the aromatic dithiol core (peaks observed at 250 and 275 nm) as well as for the nucleobases (300 to 350 nm), suggesting that these experience a chiral microenvironment in the assemblies (see the Supporting Information, Figure S5). Taken together, these data suggest that $(\mathbf{3a})_4$, $(\mathbf{3b})_5$, and $(\mathbf{3d})_3$ are self-replicators. Interestingly, the different nucleobases impose different ring sizes on the replicators. We previously noted that, for peptide-based replicators, the size of the ring is inversely correlated with the strength of the (hydrophobic) interactions between the building blocks in the assemblies.[39] Surprisingly, in the present nucleobase system the trend in size of the selected macrocycle (thymine > adenine > cytosine) is exactly the opposite. Judging from water/chloroform partitioning coefficient, the hydrophilicity of the nucleobases follows the trend thymine (0.45) < adenine (0.78) < cytosine (3.00), where the number in parentheses corresponds to $\log(C_{water}/C_{CHCl3})$.[52] Thus, the mode of assembly of the rings into the stacks appears to involve more than just hydrophobic binding. What these additional interaction patterns are remains obscure as attempts at elucidating the assembly structure have so far been unsuccessful.

Similar to $\mathbf{1}_2(\mathbf{2a})_1$, the kinetic order in self-replicator $(\mathbf{3b})_5$ (1.21 ± 0.14; Figure 7 and S102) indicates that also $(\mathbf{3b})_5$ is capable of exponential growth.

https://dx.doi.org/10.1021/jacs.9b10796
J. Am. Chem. Soc. 2020, 142, 4184–4192

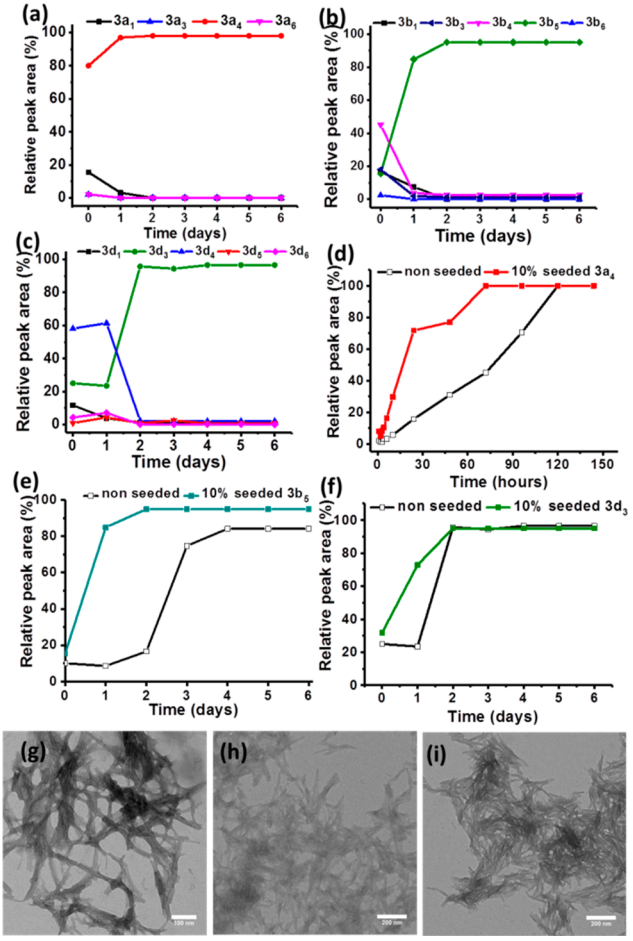


Figure 6. Change in product distribution with time in stirred (1200 rpm) DCLs made from (a) **3a**, (b) **3b**, and (c) **3d** (3.8 mM) in 50 mM borate buffer pH 8.2. Kinetics of formation of: (d) $(\mathbf{3a})_4$, (e) $(\mathbf{3b})_5$, and (f) $(\mathbf{3d})_3$ in DCLs made from the corresponding building block (3.8 mM) in the absence (open symbols) and presence (closed symbols) of 10 mol % of the corresponding preformed seed (added at day 0). All of the DCLs were oxidized to 80% using sodium perborate, with the exception for the seeding experiment of building block **3a** (graph 6b), which was performed starting from a 100% reduced library. Negative staining TEM images of the assemblies formed in DCLs: dominated by (g) $(\mathbf{3a})_4$, (h) $(\mathbf{3b})_5$, and (i) $(\mathbf{3d})_3$. Additional TEM images are shown in Figures S150, S151, and S153.

We then extended our exploration to the series of histidine containing building blocks **3e**–**3h**. For adenine building block **3e** a tetramer macrocycle was obtained while for thymine and cytosine analogs **3f** and **3g**, respectively, trimer macrocycles were observed (see the Supporting Information, Figure S101a-c). All of these macrocycles were proven to be self-replicators (Figure S101d–f), adopting ordered supramolecular assemblies (Figures S101g–i and S157), as observed for the lysine containing building blocks.

Of the four nucleobases explored only guanine failed to produce self-replicators. As shown below, DCLs made from building block **3c** produced a range of macrocycles, while building block **3g** gave rise to tetramer. Guanines are known to assemble into quadruplexes that are stabilized by cations, in particular potassium.[53] Indeed, upon addition of 50 mM potassium bromide to the DCL made from **3c** a dramatic change in product distribution was observed: the macrocyclic trimer was formed in essentially quantitative yield (Figure 8a).

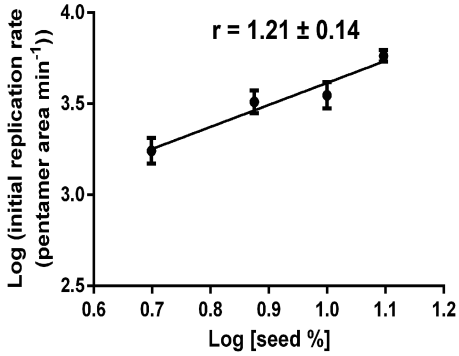


Figure 7. Determination of the order in replicator $(\mathbf{3b})_5$. The initial replication rate is plotted versus the concentration of replicator. The data points correspond to seeding concentrations of 5.0, 7.5, 10, and 12.5% relative to the stock solution (500 μM in building block **3b**). The error bars denote the standard deviation based on three individual measurements.

This transformation at the molecular level was accompanied by a change of the macroscopic appearance from a clear solution to a viscous suspension.

The CD spectrum changed upon addition of KBr from a low intensity featureless trace to a reveal a strong signal at 268 nm

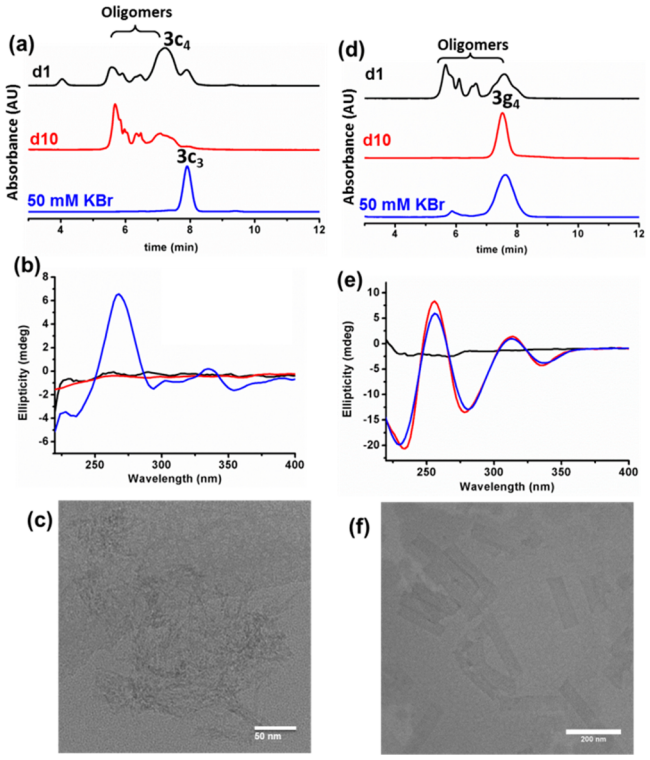


Figure 8. Cation-induced guanine based supramolecular assembly. UPLC analysis of DCLs prepared from (a) **3c** and (d) **3g** (3.8 mM in 12.5 mM borate buffer pH 8.2) after oxidation (80%) using sodium perborate at day 1 (upper line), day 10 (middle line), and after the addition of 50 mM potassium bromide (KBr). CD spectra of DCLs made from (b) **3c** and (e) **3g** before (black and red) and after (blue) addition of KBr. (c) Cryo-TEM image of DCL prepared from **3c** upon the addition of KBr, corresponding to the trimer and (f) negative stain TEM image of **3g** after 10 days of the reaction corresponding to the tetramer. Additional TEM images are shown in Figures S152 and S156.

https://dx.doi.org/10.1021/jacs.9b10796
J. Am. Chem. Soc. 2020, 142, 4184–4192

(Figure 8b) typical for CD spectra of G-quadruplexes.[54] Bundles of thin fibers were revealed using cryo-TEM (Figure 8c). In the DCL made from **3g**, also a family of large oligomers was observed initially. However, over time, these products gave way to the tetramer macrocycle (Figure 8d), which self-assembled into twisted tape-like structures (Figure 8e,f). Addition of KBr did not significantly alter composition of this DCL, nor its nanostructure.

Finally, in order to investigate the possibility to mutate the nucleobases during the replication of the macrocycles based on the PNA building blocks, a series of cross-seeding experiments were performed. Analysis of DCLs made by mixing all four nucleobases revealed broad and overlapping peaks containing mixtures of trimers and tetramers which could not be adequately separated with the available chromatographic techniques (see the Supporting Information, Figure S103). Analysis was drastically improved in the three-component system (A, T, and C). However, seeding experiments failed to induce the amplification of a specific macrocycle (see the Supporting Information, Figure S105). Notably, in the binary system containing A and T, seeding experiments with $(\mathbf{3b})_5$, triggered the autocatalytic formation of a family of mixed pentamers (Figures 9 and S127), whereas a mixture of trimers

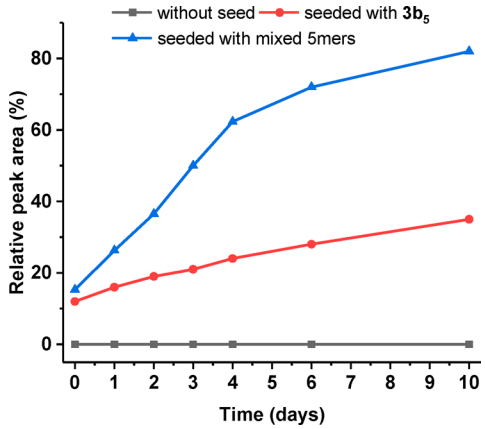


Figure 9. Seeding induced the growth of replicators $(\mathbf{3a})_n(\mathbf{3b})_{5-n}$ in DCLs prepared from **3a** and **3b** ([**3a**] = [**3b**] = 1.0 mM) in the presence of 10 mol % of seed containing $(\mathbf{3b})_5$ (red circles) or $(\mathbf{3a})_n(\mathbf{3b})_{5-n}$ (blue triangles). All of the DCLs were oxidized to 80% using sodium perborate, followed by addition of the seed.

and tetramers dominated in the system in the absence of seed. Binary mixtures of other nucleobases showed limited response to seeding (see Supporting Information, Figures S116, S120, S123, S127, and S137). Overall, these results suggest that in the PNA system replication with mutation is less efficient compared to the one composed of nucleobase- and peptide-based building blocks, most likely as a consequence of the fact that the individual nucleobases in the PNA system have a strong effect on the nature of the self-replicating macrocycles, each preferring a different ring size.

CONCLUSIONS

We have shown that exponential replicators featuring nucleobases and amino acids can emerge spontaneously from mixtures of relatively simple building blocks. The autocatalytic supramolecular polymerization of specific nucleobase-containing rings into stacks leads to a linear arrangement of nucleobases. The present system has the advantage over previously reported assemblies of nucleobase analogues[55] in that they form autocatalytically. The ease with which these long noncovalent oligomers are produced is in stark contrast to the difficulty in obtaining long oligomers through more conventional nonautocatalytic condensation reactions of (activated) nucleotides. Interactions between nucleobases through hydrogen bonding (i.e., base-pairing) do not appear to play a role in the assembly and replication steps. However, we speculate that the fibrous assemblies might be a stepping stone toward systems in which information transfer occurs through base-pairing interactions. Thus, structures like our replicator fibers might be important for closing the gap between the building blocks of life and the formation of long functional information polymers. Investigations are currently underway aimed at obtaining replicator assemblies featuring nucleobases that are accessible for base-pairing.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.9b10796.

Synthetic procedures, NMR spectra, UPLC and LC-MS methods, methods of DCL and sample preparation, and mass spectra (PDF)

AUTHOR INFORMATION

Corresponding Author

Sijbren Otto − *Centre for Systems Chemistry, Stratingh Institute, University of Groningen 9747 AG Groningen, The Netherlands*; orcid.org/0000-0003-0259-5637; Email: s.otto@rug.nl

Authors

Bin Liu − *Centre for Systems Chemistry, Stratingh Institute, University of Groningen 9747 AG Groningen, The Netherlands*

Charalampos G. Pappas − *Centre for Systems Chemistry, Stratingh Institute, University of Groningen 9747 AG Groningen, The Netherlands*

Jim Ottelé − *Centre for Systems Chemistry, Stratingh Institute, University of Groningen 9747 AG Groningen, The Netherlands*

Gaël Schaeffer − *Centre for Systems Chemistry, Stratingh Institute, University of Groningen 9747 AG Groningen, The Netherlands*

Christoph Jurissek − *Centre for Systems Chemistry, Stratingh Institute, University of Groningen 9747 AG Groningen, The Netherlands*

Priscilla F. Pieters − *Centre for Systems Chemistry, Stratingh Institute, University of Groningen 9747 AG Groningen, The Netherlands*

Meniz Altay − *Centre for Systems Chemistry, Stratingh Institute, University of Groningen 9747 AG Groningen, The Netherlands*

Ivana Marić − *Centre for Systems Chemistry, Stratingh Institute, University of Groningen 9747 AG Groningen, The Netherlands*

Marc C. A. Stuart − *Centre for Systems Chemistry, Stratingh Institute, University of Groningen 9747 AG Groningen, The Netherlands*; orcid.org/0000-0003-0667-6338

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.9b10796

Author Contributions

†B.L. and C.G.P. contributed equally.

https://dx.doi.org/10.1021/jacs.9b10796
J. Am. Chem. Soc. 2020, 142, 4184−4192

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful for support from the ERC (AdG 741774), the EU (MCIF 745805 - DSR), NWO (VICI grant), Zernike Dieptestrategie and the Dutch Ministry of Education, Culture and Science (Gravitation program 024.001.035).

■ REFERENCES

(1) Kauffman, S. A. At home in the universe. *Math. Soc. Sci.* **1997**, *1*, 94−95.
(2) Schwille, P.; Spatz, J.; Landfester, K.; Bodenschatz, E.; Herminghaus, S.; Sourjik, V.; Erb, T. J.; Bastiaens, P.; Lipowsky, R.; Hyman, A.; Dabrock, P.; Baret, J.-C.; Vidakovic-Koch, T.; Bieling, P.; Dimova, R.; Mutschler, H.; Robinson, T.; Tang, T.-Y. D.; Wegner, S.; Sundmacher, K. MaxSynBio: avenues towards creating cells from the bottom up. *Angew. Chem., Int. Ed.* **2018**, *57*, 13382−13392.
(3) Ruiz-Mirazo, K.; Briones, C.; de la Escosura, A. Prebiotic systems chemistry: new perspectives for the origins of life. *Chem. Rev.* **2014**, *114*, 285−366.
(4) Ashkenasy, G.; Hermans, T. M.; Otto, S.; Taylor, A. F. Systems chemistry. *Chem. Soc. Rev.* **2017**, *46*, 2543−2554.
(5) Mattia, E.; Otto, S. Supramolecular systems chemistry. *Nat. Nanotechnol.* **2015**, *10*, 111−119.
(6) Williams, L. D. Yin and yang: polypeptide and polynucleotide. *Origins of Life Community* **2012**, *2012*.
(7) Kunnev, D.; Gospodinov, A. Possible emergence of sequence specific RNA aminoacylation via peptide intermediary to initiate Darwinian evolution and code through origin of life. *Life* **2018**, *8*, 44.
(8) Carter, C. What RNA world? Why a peptide/RNA partnership merits renewed experimental attention. *Life* **2015**, *5*, 294−320.
(9) Bai, Y. S.; Chotera, A.; Taran, O.; Liang, C.; Ashkenasy, G.; Lynn, D. G. Achieving biopolymer synergy in systems chemistry. *Chem. Soc. Rev.* **2018**, *47*, 5530−5530.
(10) Chotera, A.; Sadihov, H.; Cohen-Luria, R.; Monnard, P. A.; Ashkenasy, G. Functional assemblies emerging in complex mixtures of peptides and nucleic acid-peptide chimeras. *Chem. - Eur. J.* **2018**, *24*, 10128−10135.
(11) Liu, P.; Ni, R.; Mehta, A. K.; Childers, W. S.; Lakdawala, A.; Pingali, S. V.; Thiyagarajan, P.; Lynn, D. G. Nucleobase-Directed Amyloid Nanotube Assembly. *J. Am. Chem. Soc.* **2008**, *130*, 16867−16869.
(12) Stephanopoulos, N. Peptide-oligonucleotide hybrid molecules for bioactive nanomaterials. *Bioconjugate Chem.* **2019**, *30*, 1915−1922.
(13) Liu, B.; Pappas, C. G.; Zangrando, E.; Demitri, N.; Chmielewski, P. J.; Otto, S. Complex molecules that fold like proteins can emerge spontaneously. *J. Am. Chem. Soc.* **2019**, *141*, 1685−1689.
(14) Pross, A. Causation and the origin of life. Metabolism or replication first? *Origins Life Evol. Biospheres* **2004**, *34*, 307−321.
(15) Kosikova, T.; Philp, D. Exploring the emergence of complexity using synthetic replicators. *Chem. Soc. Rev.* **2017**, *46*, 7274−7305.
(16) Bissette, A. J.; Fletcher, S. P. Mechanisms of autocatalysis. *Angew. Chem., Int. Ed.* **2013**, *52*, 12800−12826.
(17) Patzke, V.; von Kiedrowski, G. Self-replicating systems. *Arkivoc* **2007**, 293−310.
(18) Duim, H.; Otto, S. Towards open-ended evolution in self-replicating molecular systems. *Beilstein J. Org. Chem.* **2017**, *13*, 1189−1203.
(19) Vonkiedrowski, G. A self-replicating hexadeoxynucleotide. *Angew. Chem., Int. Ed. Engl.* **1986**, *25*, 932−935.
(20) Paul, N.; Joyce, G. F. A self-replicating ligase ribozyme. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12733−12740.
(21) Lincoln, T. A.; Joyce, G. F. Self-sustained replication of an RNA enzyme. *Science* **2009**, *323*, 1229−1232.
(22) Hayden, E. J.; von Kiedrowski, G.; Lehman, N. Systems chemistry on ribozyme self-construction: evidence for anabolic autocatalysis in a recombination network. *Angew. Chem., Int. Ed.* **2008**, *47*, 8424−8428.
(23) Vaidya, N.; Manapat, M. L.; Chen, I. A.; Xulvi-Brunet, R.; Hayden, E. J.; Lehman, N. Spontaneous network formation among cooperative RNA replicators. *Nature* **2012**, *491*, 72−77.
(24) Jayathilaka, T. S.; Lehman, N. Spontaneous covalent self-assembly of the azoarcus ribozyme from five fragments. *ChemBioChem* **2018**, *19*, 217−220.
(25) Lee, D. H.; Granja, J. R.; Martinez, J. A.; Severin, K.; Ghadiri, M. R. A self-replicating peptide. *Nature* **1996**, *382*, 525−528.
(26) Yao, S.; Ghosh, I.; Zutshi, R.; Chmielewski, J. A pH-modulated, self-replicating peptide. *J. Am. Chem. Soc.* **1997**, *119*, 10559−10560.
(27) Rubinov, B.; Wagner, N.; Rapaport, H.; Ashkenasy, G. Self-replicating amphiphilic beta-sheet peptides. *Angew. Chem., Int. Ed.* **2009**, *48*, 6683−6686.
(28) Nanda, J.; Rubinov, B.; Ivnitski, D.; Mukherjee, R.; Shtelman, E.; Motro, Y.; Miller, Y.; Wagner, N.; Cohen-Luria, R.; Ashkenasy, G. Emergence of native peptide sequences in prebiotic replication networks. *Nat. Commun.* **2017**, *8*, 434.
(29) Tjivikua, T.; Ballester, P.; Rebek, J. A self-replicating system. *J. Am. Chem. Soc.* **1990**, *112*, 1249−1250.
(30) Dieckmann, A.; Beniken, S.; Lorenz, C. D.; Doltsinis, N. L.; von Kiedrowski, G. Elucidating the origin of diastereoselectivity in a self-replicating system: selfishness versus altruism. *Chem. - Eur. J.* **2011**, *17*, 468−480.
(31) Bottero, I.; Huck, J.; Kosikova, T.; Philp, D. A synthetic replicator drives a propagating reaction-diffusion front. *J. Am. Chem. Soc.* **2016**, *138*, 6723−6726.
(32) Sadownik, J. W.; Kosikova, T.; Philp, D. Generating system-level responses from a network of simple synthetic replicators. *J. Am. Chem. Soc.* **2017**, *139*, 17565−17573.
(33) Kosikova, T.; Philp, D. Two synthetic replicators compete to process a dynamic reagent pool. *J. Am. Chem. Soc.* **2019**, *141*, 3059−3072.
(34) von Kiedrowski, G. In *Bioorganic chemistry frontiers*; Springer: Berlin, 1993; pp 113−146.
(35) Szathmáry, E.; Gladkih, I. Sub-exponential growth and coexistence of non-enzymatically replicating templates. *J. Theor. Biol.* **1989**, *138*, 55−58.
(36) Szathmáry, E. Simple growth laws and selection consequences. *Trends. Ecol. Evol.* **1991**, *6*, 366−370.
(37) Lifson, S.; Lifson, H. Coexistence and Darwinian selection among replicators: response to the preceding paper by Scheuring and Szathmáry. *J. Theor. Biol.* **2001**, *212*, 107−109.
(38) Carnall, J. M. A.; Waudby, C. A.; Belenguer, A. M.; Stuart, M. C. A.; Peyralans, J. J. P.; Otto, S. Mechanosensitive self-replication driven by self-organization. *Science* **2010**, *327*, 1502−1506.
(39) Malakoutikhah, M.; Peyralans, J. J. P.; Colomb-Delsuc, M.; Fanlo-Virgos, H.; Stuart, M. C. A.; Otto, S. Uncovering the selection criteria for the emergence of multi-building-block replicators from dynamic combinatorial libraries. *J. Am. Chem. Soc.* **2013**, *135*, 18406−18417.
(40) Leonetti, G.; Otto, S. Solvent composition dictates emergence in dynamic molecular networks containing competing replicators. *J. Am. Chem. Soc.* **2015**, *137*, 2067−2072.
(41) Bartolec, B.; Altay, M.; Otto, S. Template-promoted self-replication in dynamic combinatorial libraries made from a simple building block. *Chem. Commun.* **2018**, *54*, 13096−13098.
(42) Sadownik, J. W.; Mattia, E.; Nowak, P.; Otto, S. Diversification of self-replicating molecules. *Nat. Chem.* **2016**, *8*, 264−269.
(43) Colomb-Delsuc, M.; Mattia, E.; Sadownik, J. W.; Otto, S. Exponential self-replication enabled through a fibre elongation/breakage mechanism. *Nat. Commun.* **2015**, *6*, 8427.
(44) Ura, Y.; Beierle, J. M.; Leman, L. J.; Orgel, L. E.; Ghadiri, M. R. Self-assembling sequence-adaptive peptide nucleic acids. *Science* **2009**, *325*, 73−77.
(45) Singhal, A.; Nielsen, P. E. Cross-catalytic peptide nucleic acid (PNA) replication based on templated ligation. *Org. Biomol. Chem.* **2014**, *12*, 6901−6907.

https://dx.doi.org/10.1021/jacs.9b10796
J. Am. Chem. Soc. 2020, 142, 4184−4192

(46) Ploger, T. A.; von Kiedrowski, G. A self-replicating peptide nucleic acid. *Org. Biomol. Chem.* **2014**, *12*, 6908−6914.
(47) Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. Dynamic combinatorial libraries of macrocyclic disulfides in water. *J. Am. Chem. Soc.* **2000**, *122*, 12063−12064.
(48) Frederix, P.; Ide, J.; Altay, Y.; Schaeffer, G.; Surin, M.; Beljonne, D.; Bondarenko, A. S.; Jansen, T. L. C.; Otto, S.; Marrink, S. J. Structural and spectroscopic properties of assemblies of self-replicating peptide macrocycles. *ACS Nano* **2017**, *11*, 7858−7868.
(49) Lutter, L.; Serpell, C. J.; Tuite, M. F.; Xue, W. F. The molecular lifecycle of amyloid - Mechanism of assembly, mesoscopic organisation, polymorphism, suprastructures, and biological consequences. *Biochim. Biophys. Acta, Proteins Proteomics* **2019**, *1867*, 140257.
(50) Jafilan, S.; Klein, L.; Hyun, C.; Florian, J. Intramolecular base stacking of dinucleoside monophosphate anions in aqueous solution. *J. Phys. Chem. B* **2012**, *116*, 3613−3618.
(51) Bommarito, S.; Peyret, N.; SantaLucia, J. Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Res.* **2000**, *28*, 1929−1934.
(52) Cullis, P. M.; Wolfenden, R. Affinities of nucleic-acid bases for solvent water. *Biochemistry* **1981**, *20*, 3024−3028.
(53) Peters, G. M.; Davis, J. T. Supramolecular gels made from nucleobase, nucleoside and nucleotide analogs. *Chem. Soc. Rev.* **2016**, *45*, 3188−3206.
(54) Davis, J. T. G-quartets 40 years later: From 5 ′-GMP to molecular biology and supramolecular chemistry. *Angew. Chem., Int. Ed.* **2004**, *43*, 668−698.
(55) Cafferty, B. J.; Gallego, I.; Chen, M. C.; Farley, K. I.; Eritja, R.; Hud, N. V. Efficient self-assembly in water of long noncovalent polymers by nucleobase analogues. *J. Am. Chem. Soc.* **2013**, *135*, 2447−2450.

https://dx.doi.org/10.1021/jacs.9b10796
J. Am. Chem. Soc. 2020, 142, 4184−4192