Departure from Randomness: Evolution of Self-Replicators that can Self-Sort through Steric Zipper Formation

Marcel J. Eleveld[1], Juntian Wu[1], Kai Liu[1], Jim Ottelé[1], Omer Markovitch[1,2,3], Armin Kiani[1], Sijbren Otto[1]*

Abstract

Darwinian evolution of synthetic self-replicating entities is likely to have an important role in the emergence of life from inanimate chemical matter. Darwinian evolution of self-replicators requires that these (i) have structural space accessible to them; (ii) occupy only part of this space at any one time, and (iii) navigate this space through a process of mutation and selection. We now report a system of self-replicating hexameric macrocycles that emerges upon mixing two building blocks. It occupies a subset of possible building block sequences. Specific interactions between the building blocks, most likely through steric zipper formation involving the interdigitation of a phenylalanine residue of one building block between two leucine residues of the other building block, results in the preferential formation of a hexamer with a sequence in which the two building blocks alternate. When this system was exposed to two different replication-destruction regimes, different replicator mutant distributions were selected for. When the destruction process was non-selective (mediated by outflow in an open system) the fastest replicating sequences dominated, overriding the preference for zipper formation observed in a closed vial. However, when destruction was mediated chemically (and therefore potentially selective) the replicator mutant that combined adequate resistance to reduction with adequate replication speed and was capable of steric zipper formation, became dominant. These results constitute a rudimentary form of Darwinian evolution where replicators adapt to a changing selection pressure through mutation and selection.

Introduction

How chemistry can be turned into biology is one of the grand challenges of contemporary science.[1-6] Darwinian evolution is likely to play an important role in this transition. Self-replicating molecules[7-11] are suitable candidates on which evolution can act. Self-replicators make copies of themselves by bringing building blocks in close proximity, thereby catalyzing their ligation to form the new copy and transmitting the information contained in the original replicator to the next generation (inheritance). While many different types of synthetic self-replicating molecules have been developed[12-21] since the

---

[1] Center for Systems Chemistry, Stratingh Institute, University of Groningen, Groningen, the Netherlands.
[2] Groningen Institute for Evolutionary Life Sciences, University of Groningen, the Netherlands.
[3] Blue Marble Space Institute of Science, Seattle, WA, 98154, USA.
*Correspondence: s.otto@rug.nl

1

first report in 1986,[22] Darwinian evolution of these has not yet been reported. In order to develop evolvable systems of self-replicators several conditions need to be met:

Firstly, the system should be designed such that it can access a space (i.e. self-replicating molecules with different structures) that potentially harbors multiple heritable states (i.e. different replicator quasi species[23-25]). In most published studies, self-replicators are only supplied with building blocks that are their exact fragments and only a single product can be formed. In order to allow for evolution the system needs to be supplied with more building blocks, so that they can be incorporated into the replicators in different combinations, creating opportunities for evolution.

Secondly, once a larger replicator space is made available to the system, this space needs to partially (ideally sparsely) populated. It is desirable that, at any given time, the system only occupies a very small subset of the available structure space.[1] The incorporation of building blocks into replicators should not be random, as this would result in the simultaneous population of all possible structures under most conditions.

Thirdly, starting from a state where a subset of all possible replicator structures is populated, incursions into the remainder of the replicator space may then occur through a process of mutation and selection. Only few examples have been reported of self-replicating systems with the ability to mutate[26-31] and even fewer instances where not all possible mutants were (immediately) populated.[28-29] Selection is most easily implemented if the process of replication is supplemented by a process of replicator "death", which can occur through physical removal or chemical destruction of a fraction of the replicator population. Replication needs to be faster than removal/destruction in order for a replicator to sustain its population. Operating systems in such out-of-equilibrium regime[32-35] allows for selection of replicators based on their dynamic kinetically stability[36-37] (as opposed to selective formation of the thermodynamically most stable, or fastest replicator, as would happen in closed systems). The first examples of self-replication operated under such conditions have been reported.[38-41]

We decided to probe whether we can: (i) enlarge replicator space; (ii) selectively populate part of it; and (iii) navigate it through a process of mutation and selection, using the systems of self-replicating molecules that we developed previously.[1, 42-43] These replicators emerge spontaneously from a dynamic combinatorial library (DCL) made from a peptide-functionalized dithiol building block (such as **1**), when it is stirred and slowly oxidized by atmospheric oxygen. Under these conditions initially a mixture of macrocycles of various sizes is formed that constantly exchange building blocks. After a nucleation event a self-replicating macrocycle of a distinct macrocycle size can emerge that catalyzes its own formation. Self-replication is driven by stacking of the rings into fibers, reinforced by the formation of β-sheets between the peptide chains. Binding of precursors to fiber sides guides this material to the growing fiber ends.[44] Mechanical agitation pushes this process into a fiber elongation-fragmentation regime, which enables exponential growth of the self-replicating macrocycle (Figure 1).[45]
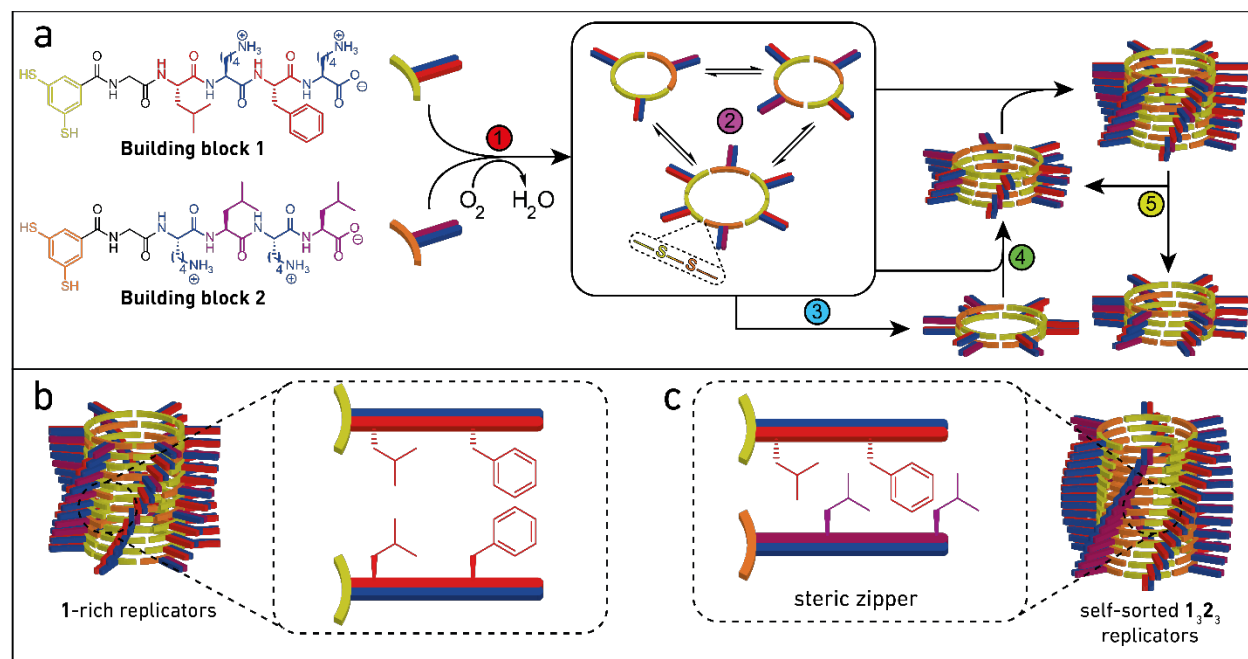
Figure 1. *(a) Mechanism of self-replication. Monomeric building blocks (**1** and **2**) are oxidized by oxygen from the air (1) to produce a mixture of small macrocycles that constantly exchange building blocks (2). A nucleation event (3) can cause hexameric replicators to catalyze the formation of more hexamers, resulting in fiber formation (4). Mechanical agitation causes fragmentation of fibers with a sufficient length (5), increasing the number of growing fiber ends. This fragmentation/elongation mechanism allows exponential replicator growth. Building block **2** also makes a self-replicator that replicates using this mechanism. In the graphical representation aromatic dithiol cores of **1** and **2** are indicated as yellow and orange, respectively. The hydrophobic amino acid residues point to the same side of the peptide backbone, depicted by the red or blue bars in the graphic representation. The hydrophilic residues point to the opposing side, depicted by the blue bar. (b) In a dynamic combinatorial library (DCL) containing **1** and **2**, initially **1**-rich replicators will be dominant, because they are produced fastest. (c) After some time self-sorted **1**$_3$**2**$_3$-replicator will dominate the DCL, because it has an increased stability when a steric zipper is formed when **1** and **2** are alternating within the macrocycle.*

In most of our previous work we observed the emergence of a self-replicator of one specific macrocycle size, even though several ring sizes are in principle accessible.[43] In this regard replicator emergence is already a (sometimes stochastic[46]) departure from randomness. However, when using mixtures of building blocks we mostly obtain, for a given ring size, a statistical distribution of different sequences (although during the growth process the distribution can be temporarily non-statistical).[29, 47] In these cases only the macrocycle size is transferred to the next generations, but no information regarding the building-block composition of this macrocycle is inherited. Building blocks tend to be incorporated almost randomly. This tendency of multi-building-block systems to produce many different possible products, is a well-known phenomenon that lies at the basis of one of the fundamental paradoxes of the origin of life: the asphalt paradox.[48] Simply put: chemistry tends to diversify, yet biology uses only a select subset of chemical reactions and structures. Identifying mechanisms that allow for an escape from randomness and impart selectivity in the incorporation of building blocks into replicators is important to

3

address this paradox. The challenge is now to discover what can make self-replicating systems selective in producing a specific building block sequence. We reasoned that tuning the interactions between building blocks may lead to their self-sorting[49-52] which can be *narcissistic* when there is a favorable interaction between the same building blocks or *social* when there is a favorable interaction between different building blocks.

We now report a two-building-blocks system that departs from randomness by producing hexameric self-replicators with specific building-block sequence preferences. In a closed vial, initially a subset of possible hexameric self-replicators is produced that is rich in one of the building blocks. Upon equilibration the composition changes and becomes dominated by a socially self-sorted hexameric self-replicator with the building blocks (three of each) arranged in an alternating fashion, driven by the formation of a steric zipper structure. Exposing the system to a replication-destruction regime, and applying different selection pressures then led to the formation of different subsets of mutants through a process of mutation and selection, amounting to a rudimentary instance of Darwinian evolution.

Results and discussion

Selective occupation of a subset of replicator space

Our first aim was to develop a system of self-replicating molecules where only subset of potential replicators are populated, aiming specifically for selective formation of specific sequences from among many different possibilities. We built on our previous work where self-replicators were made from dithiol building blocks appended with a peptide containing alternating hydrophobic and hydrophilic amino acids, starting with a hydrophobic one.[43] Mixing two of such building blocks should open up a combinatorial space of different sequences with specific replicator ring sizes. A study of 9 binary mixtures of such building blocks, differing by one or two amino acid residues, invariably produced a set of replicators with a specific macrocycle size, while the distribution of the two building blocks within this ring was dictated by statistics (see SI Section S18). Apparently changes in the amino-acid sequence, while maintaining the registry of hydrophobic and hydrophilic amino acids, did not allow a departure from randomness. We hypothesized that making the building blocks structurally more different should lead to better discrimination between them. This consideration led us to focus on the building block pair **1** and **2**. Building block **1** is one of our most reliable candidates for producing self-replicators.[40, 43-44, 53-54] Building block **2** was designed to have the same dithiol aromatic core as **1**, but with a pentapeptide chain of alternating hydrophilic and hydrophobic residues, where the first residue (after the Gly linker) is hydrophilic rather than hydrophobic. Because of the hydrophobic-hydrophilic inversion of the sequence **2** is out-of-register with **1** and it should be unfavorable for different building blocks to stack on top of each other to form a β-sheet.

Similar to **1**, a DCL of **2** in a closed vial, upon oxidation in a buffered aqueous solution at 45 °C produces initially a mixture of small macrocycles (trimers and tetramers) and, after a lag phase, a hexameric macrocycle 2_6 emerges and grows to dominate the DCL (Figure S17). To confirm that 1_6 and 2_6 are self-replicators and to probe the extent to which the two building blocks can cross-catalyze each other's

4

formation, we performed (cross-)seeding experiments. First a DCL of **1** was prepared and divided over three vials, to which 10% of pre-formed 1_6, 2_6 or no seed were added. A similar experiment was performed using a DCL of **2** at 45 °C, (2_6 is only formed efficiently at elevated temperatures). The DCLs were not stirred to prevent spontaneous nucleation of any self-replicators as much as possible and solely investigate the growth of already present replicator fibers. The DCLs were analyzed daily using UPLC-MS (Figure 2).
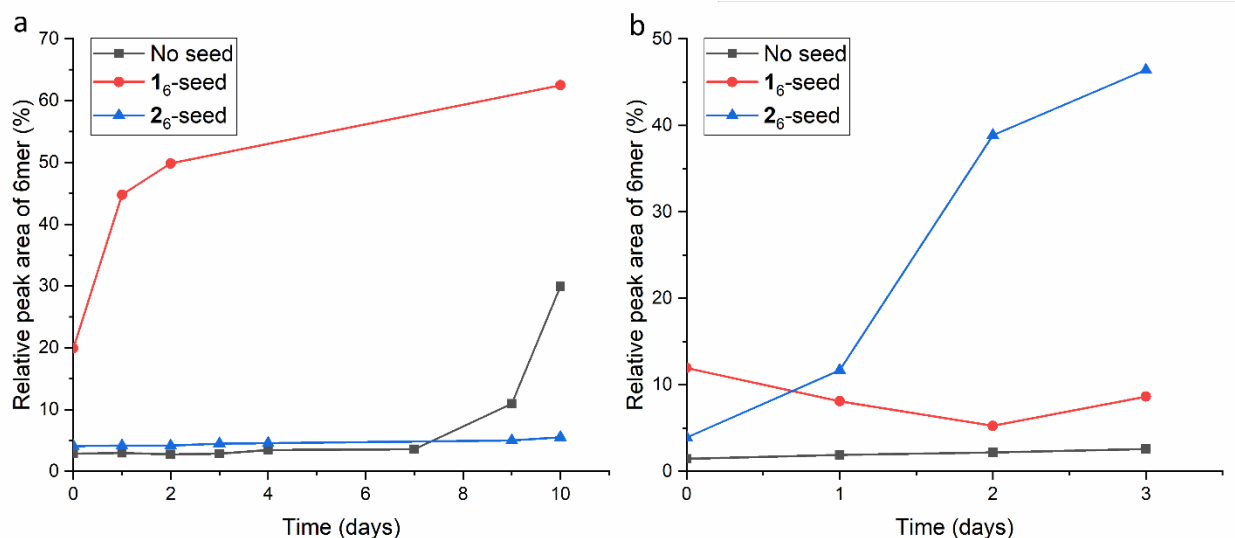

Figure 2. *Kinetic traces of 6mer formation in buffered aqueous solution (100 mM B_2O_3, pH 8.2) based on UPLC-MS analysis of (cross-)seeding experiments on (a) DCLs made from **1** (2.0 mM) seeded with 10% preformed 1_6, 2_6 or no seed at room temperature without stirring and (b) DCLs made from **2** (2.0 mM) seeded with 10% preformed 1_6, 2_6 or no seed at 45 °C without stirring.*

A seed of 1_6 only grows substantially in a DCL made from **1** and does not show growth in a DCL made from **2**. Similarly, 2_6 only grows when it is added to a DCL made from **2** and does not seed the formation of 1_6 when it was added to a DCL made from **1**. These observations lend support to the hypothesis that the different recognition motifs make it unfavorable for the different building blocks to stack on top of each other within a β-sheet. Additionally, the observation that growth of 1_6 and 2_6 from their respective building blocks is accelerated with the addition of seeds proves that 1_6 and 2_6 are self-replicators. We then proceeded to probe if mixing these building blocks together in a single sample could lead to the preferential formation of a sub-set of all possible self-replicators.
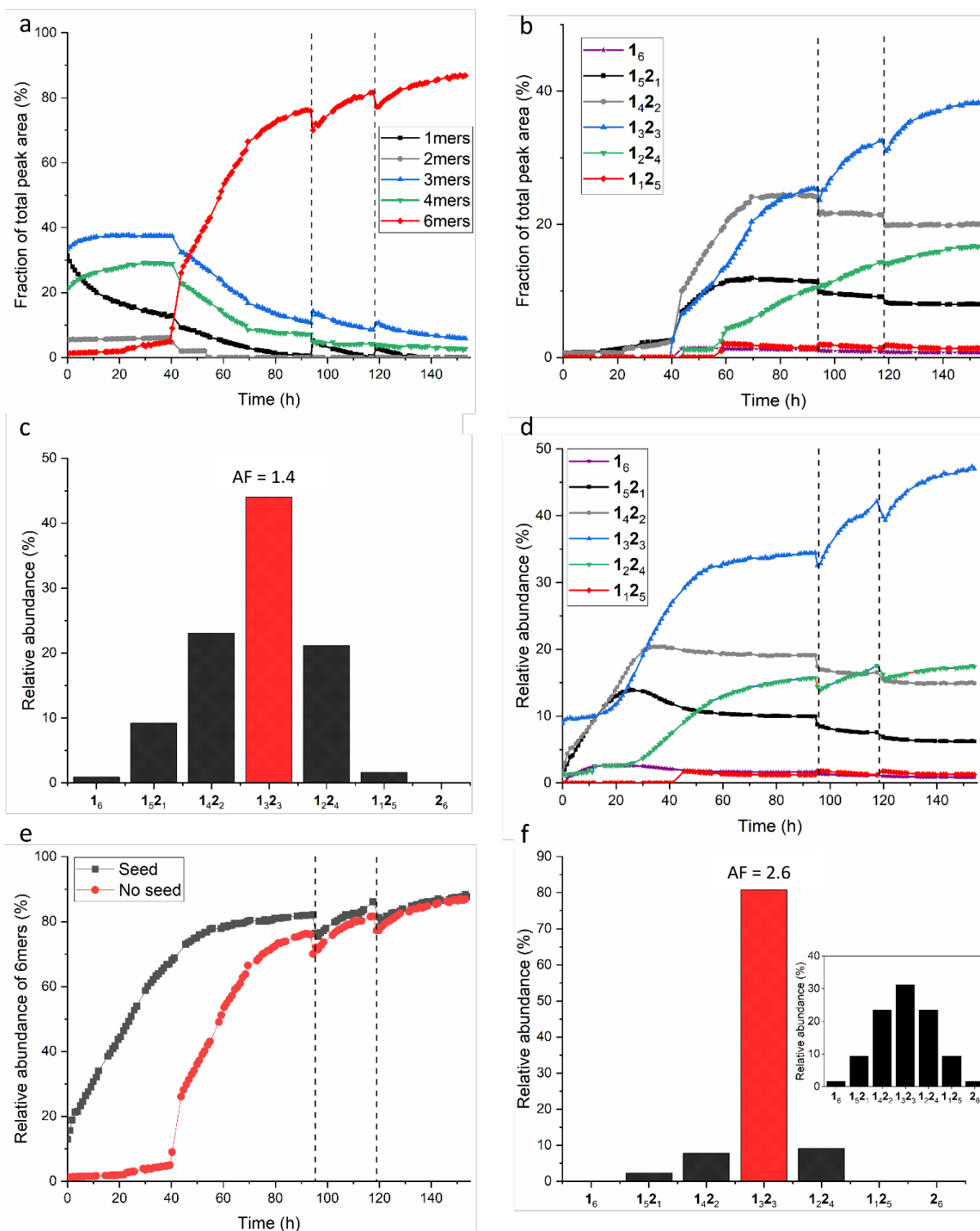
5

Figure 3. *(a) Changes in the distribution of ring sizes in a DCL (agitated by stirring at 1200 rpm at 40 °C)[55] made from equimolar amounts of **1** (1.0 mM) and **2** (1.0 mM). (b) Changes in the amounts of the different hexamer rings in the DCL in panel a showing the emergence of the self-sorted **1**$_3$**2**$_3$ mixed replicator at the expense of the **1**-rich replicators. The dotted vertical lines indicate the addition of a*

6

*mixture of 4.2% **1** and 4.2% **2** to ensure continued disulfide exchange. (c) Relative abundance of hexameric replicators at the end of the experiment in panels a and b. The distribution is enriched in 1_32_3 with an AF of 1.4; (d) Distribution of hexameric replicators in a DCL where 10 mol% seed was added at the start of the experiment. The experimental conditions are, otherwise, identical to those of the experiment in panels a-c. (e) Comparison of growth of the total amount of hexameric replicators between the seeded (panel d) and the non-seeded (panel a-c) samples. (f) Relative abundance of hexameric replicators in a DCL at 45 ^{o}C under conditions that are, otherwise, identical to those of the experiment in panel a-c. The distribution is enriched in 1_32_3 with an AF of 2.6. For reference, a statistical distribution of hexamers is shown as an inset.*

A DCL was prepared containing equimolar amounts of **1** and **2** in aqueous borate buffer. This library was stirred at elevated temperature (1200 rpm, 40 ^{o}C) and analyzed continuously by RP-UPLC(-MS).[55] Figure 3a shows the change distribution over the different rings sizes and Figure 3b shows the changes in the distribution of different hexameric rings. After ~40 hours a nucleation event took place and the hexamers started to grow, initially dominated by the **1**-rich hexamers (1_42_2, 1_52_1, together with a small amount of 1_6). Shortly after the nucleation event the growth of 1_6 halted. At approximately 70 hours the growth of 1_52_1 also stopped and at approximately 80 hours also 1_42_2 reaches a plateau. The growth of hexamer 1_32_3 only slows down when almost all free thiols (monomers) are depleted, at which point it already started to dominate the library. At two different time points small amounts of **1** and **2** were added to the DCL to reinvigorate thiol-mediated disulfide exchange, leading to a further increase in the relative amount of 1_32_3. The final distribution of hexamers is enriched in socially self-sorted 1_32_3 (Figure 3c), which even consumed part of the initially formed **1**-rich replicators. This behavior is consistent with this specific hexamer being thermodynamically more stable than those formed initially. The preferential formation of 1_32_3 was also observed in DCLs made using different ratios of **1** and **2** (see SI section S10 and S12). In order to probe if the system exhibits self-replication a seeding experiment was performed (Figure 3d), where at the start of the experiment 10 mol% of a pre-formed 1_32_3 was added to the DCL. The growth of the hexamers started immediately, following a similar trend as observed in the unseeded DCL. Since addition of the seed negates the lag phase that is observed in the non-seeded DCL (Figure 3e), we conclude that the mixed hexamers in this system are self-replicating.

The degree of self-sorting was quantified by estimating the amplification factor of the self-sorted species.[56] Where in a statistical distribution of two-building blocks (A+B) 31% of the formed hexamers have a composition of A_3B_3, in the mixed DCL of **1** and **2** (as shown in Figure 3c) the 1_32_3 hexamer accounts for 44% of all formed hexamers, resulting in an amplification factor (AF) of 1.4. When the same experiment is performed at 45 ^{o}C, instead of 40 ^{o}C, a similar behavior is observed, but the final distribution of hexamers is more enriched in 1_32_3, with an AF of 2.6 (Figure 3f) without having to add any additional monomers. Note that the maximum possible AF is 3.2.

Surprisingly, the two-building blocks that were designed to prefer self-interactions along the replicator stack (i.e. between macrocycles) show a pronounced preference for mixing within a macrocycle. We hypothesized that there is an unforeseen favorable interaction between **1** and **2** within the hexamer ring. It is tempting to speculate that the 1_32_3 hexamer is formed specifically, because only with this

7

composition it is possible to have an isomer where **1** and **2** are alternating within the macrocycle, thereby maximizing the interactions between **1** and **2**.

In order to probe whether the **1**$_3$**2**$_3$ hexamer indeed had an alternating sequence, we isolated **1**$_3$**2**$_3$ from the DCL by collecting fractions by UPLC, followed by fragmentation of the disulfide bonds during MALDI-TOF analysis (efforts to fragment the disulfides directly by UPLC-MS were not successful). We verified that the fragments do not scramble in the gas phase (see SI Section 8). The obtained fragmentation pattern confirmed that the **1**$_3$**2**$_3$ hexamer distribution is indeed strongly dominated by the isomer where **1** and **2** are alternating within the macrocycle. From the ratio between the different dimer fragments that were observed, the ratio between the different **1**$_3$**2**$_3$ isomers was calculated and an amplification factor for the alternating isomer (AF$_{121212}$) of 11.2 was found (see SI Section 9 for detailed calculations).

Additionally the expected distribution of hexamers was determined from a kinetic model. The kinetic model simulates a two-building-block system that allows exchange of building blocks between all possible compositions of hexamers (see SI Section S11), which produces the expected statistical distribution when both building blocks are treated indiscriminately. However, when it becomes more favorable for two different building blocks to be next to each other (**1**-**2**) compared to being next to themselves (**1**-**1** and **2**-**2**) the distribution of hexamers becomes narrower. This pairing preference was simulated by assuming that the disulfide bonds connecting **1** with **2** exchange more slowly to form other hexamers than the disulfide bonds connecting two building blocks of the same structure. An approximate rate constant reduction of a factor 10 for exchange of **1**-**2** compared to **1**-**1** and **2**-**2** resulted in a distribution that was closest to the experimentally observed hexamer distribution.

Intrigued by the preference of **1** and **2** to alternate in the self-sorted replicator we investigated what could be the driving force behind this. Thioflavin T (ThT) assays on **1**$_6$, **2**$_6$ and self-sorted **1**$_3$**2**$_3$ (SI Section 6) showed comparable fluorescence for **1**$_6$ and **2**$_6$, indicating that both of these replicators form β-sheets. The self-sorted **1**$_3$**2**$_3$ replicator gave a much larger ThT fluorescence intensity compared to **1**$_6$ and **2**$_6$, suggesting that **1**$_3$**2**$_3$ forms a more well-defined structure that restricts the conformational degrees of freedom of the bound ThT more strongly[57-58] than fibers formed in DCLs made from either **1** or **2**. We found that there is only a preference for building blocks **1** and **2** to alternate when the macrocycles they reside in are assembled into fibers (see SI Section 16), which provides another indication that β-sheet formation plays a role in this phenomenon. In a β-sheet the hydrophobic residues of each building block are likely to be located next to each other on the same side of the β-sheet: the hydrophobic face (indicated in red in Figure 1). Similarly, all charged residues are located on the opposite side of the β-sheet: the hydrophilic face. We envisage that in an assembly the hydrophobic residues of **1** can interdigitate in between the hydrophobic residues of **2** when their respective β-sheets are adjacent within a fiber, forming zipper-like structures, known as steric zippers.[59-60] Steric zipper formation is the driving force behind various amyloid assemblies[59] and provides an extra stability to their supramolecular structure.[61]

A structural investigation of a range of peptide sequences (SI section 4 for details) revealed that the interdigitation of a phenylalanine residue in between two leucine residues is particularly effective at promoting steric zipper formation. Replacing phenylalanine with smaller residues (leucine) or larger residues (1-naphthyl or tryptophan) resulted in a lower degree of self-sorting. A potential role of π

8

stacking could be ruled out, because substituting phenylalanine with cyclohexyl alanine only gave a minor change in the degree of self-sorting. The relative orientation of the peptides is crucial to observe self-sorting, as no self-sorting was observed when the hydrophobic residues where in-register or when peptides with opposite stereochemistries were mixed in a single DCL. Finally, substituting the lysine residues with arginine or ornithine had minor effects on zipper formation, confirming that it is mainly the interactions between the hydrophobic amino-acid side chains that govern zipper formation.

Navigating replicator structure space by mutation and selection

Next we wanted to probe the extent of heredity in replicators obtained from the mixture of building blocks **1** and **2.** Before applying any selection pressure we investigated the extent to which replicators pass on information about ring size and composition during replication through seeding experiments. Equimolar mixtures of **1** and **2** (75-80% oxidized) were seeded with 5 mol% of $\mathbf{1}_6$ or $\mathbf{1}_3\mathbf{2}_3$. The macrocycle composition of the seed had a substantial influence on the replication process: Using $\mathbf{1}_3\mathbf{2}_3$ as seed resulted in the conversion of 32% of the available precursors into hexamer replicator after 12 h, while upon seeding with $\mathbf{1}_6$ only 12% of precursor was converted during the same time (see SI Section 13). This difference likely reflects the fact that the precursor composition (in terms of **1**:**2** ratio) is more suited for growth of $\mathbf{1}_3\mathbf{2}_3$ than for growth of $\mathbf{1}_6$. Analysis of the composition of the newly formed replicator revealed that, at an early stage of replication (after 1 h) $\mathbf{1}_3\mathbf{2}_3$ accounted for 21% of the newly formed replicator after seeding with $\mathbf{1}_3\mathbf{2}_3$, while seeding with $\mathbf{1}_6$ only 13% of newly formed replicator had this composition. Growth of $\mathbf{1}_6$ was minimal upon seeding with $\mathbf{1}_6$ and not detectable upon seeding with $\mathbf{1}_3\mathbf{2}_3$. Its closest mutant ($\mathbf{1}_5\mathbf{2}_1$), however, accounted for 36% of the newly formed hexamers (after 1 h) when $\mathbf{1}_6$ was used as a seed, compared to 25% when seeding with $\mathbf{1}_3\mathbf{2}_3$. These number show that some degree of heredity exists, but also that the rate of mutation is relatively high.

We then proceeded to impose a selection pressure on the replicator system by implementing two different replication-destruction regimes; one where "destruction" was based on physical removal (outflow) and one where destruction was mediated chemically through disulfide bond reduction.

In the first implementation (Figure 4a) the replicator system is constantly supplied with solution of building blocks and a photoredox catalyst to enhance the rate of oxidation,[54] allowing the formation of new replicators. At the same time, part of the replicator solution is constantly removed by outflow. This physical removal of material is non-selective, so every replicator has an equal probability of being removed. Hence this replication/destruction regime selects for the fastest replication.

A self-sorted distribution of self-replicators (Figure 4c) was used at the start of the flow-experiment (Figure 4b,c). At a constant flow rate equimolar amounts of building blocks **1** and **2** and photoredox catalyst Ru(bpy)$_3^{2+}$ (in a separate syringe) were infused into the reaction mixture using a syringe pump, while at the same time and with the same flow rate part of the DCL solution was removed, maintaining a constant sample volume. The sample was continuously irradiated to maintain a sufficient degree of oxidation. During the experiment the total abundance of the hexameric self-replicators decreases from >90% to ~20% at the steady state obtained after 120h, corresponding to about 4 turnovers (Figure 4b). Especially the hexamer with composition $\mathbf{1}_3\mathbf{2}_3$ showed a drastic drop in abundance. The distribution obtained at the steady state is centered around the $\mathbf{1}_5\mathbf{2}_1$ and $\mathbf{1}_4\mathbf{2}_2$ self-replicators (Figure 4d). This

9

composition is similar to that observed transiently the early phases of growth in the closed-vial experiments shown in Figure 3, except that now also replicator 1_6 is significantly populated.
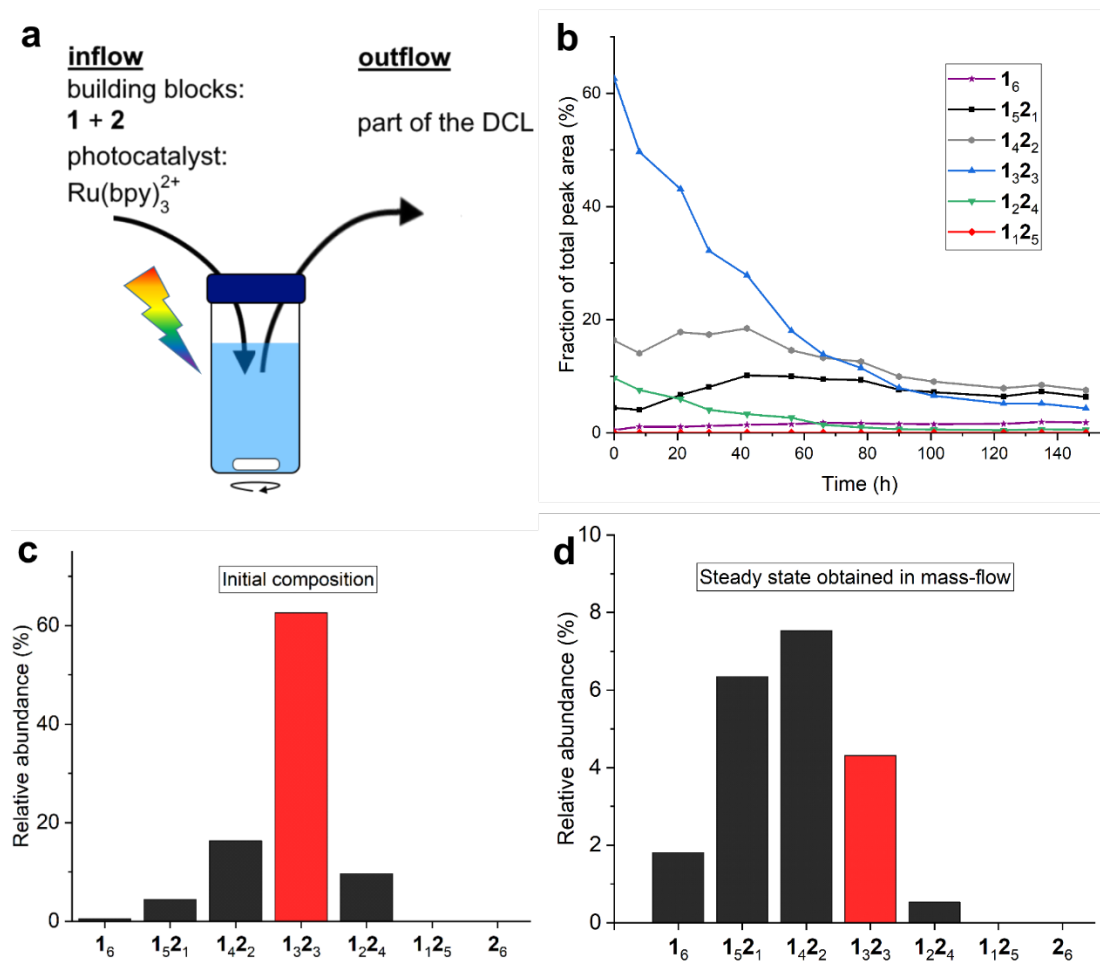

Figure 4. Replicator selection in an out-of-equilibrium regime implemented by mass-flow. (a) Schematic depiction of the flow set-up. (b) Change in relative abundances of hexameric self-replicators with different building block compositions over the course of the flow-experiment. The proportion of self-replicators (especially $1_3 2_3$) decreased from >90% at the start of the experiment to ~20% at the steady state. (c) Distribution of hexameric self-replicators at the start of the experiment and (d) the corresponding distribution at the steady state.

For the second implementation of a replication-destruction regime we continuously supplied the replicator system with oxidant (to convert dithiol building blocks to small-ring disulfides, which are the precursors to the replicators) and reductant (converting replicators and precursors back to building blocks) using previously developed methodology.[40] We again started from a pre-formed self-sorted distribution of hexamers (Figure 5b) and constantly added sodium perborate ($NaBO_3$ as oxidant) and tricarboxyethylphosphine (TCEP as reductant) (see Figure 5a). Under these conditions selection is not only based on the rate of replication but also on the resilience to reduction.
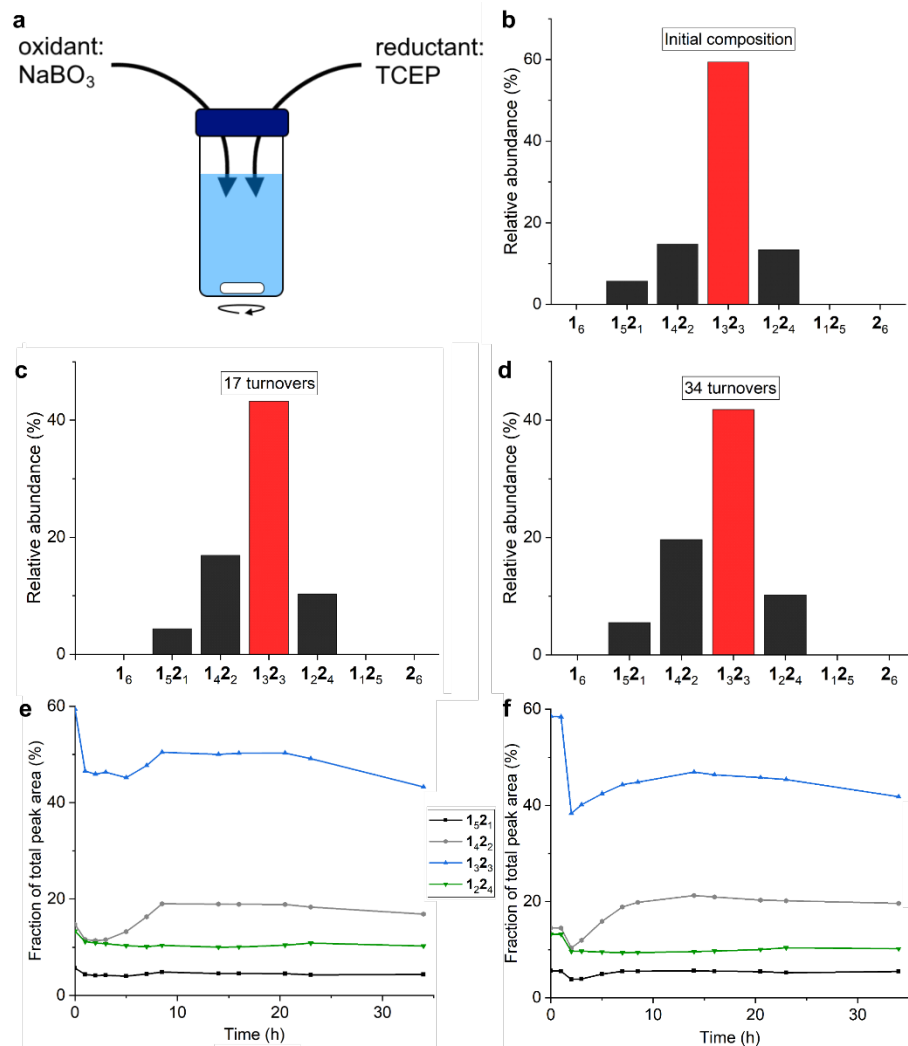
10

Figure 5. *Replicator selection in an out-of-equilibrium regime implemented by redox-infusion. (a) Schematic depiction of the redox-infusion set-up. (b) Hexamer composition before redox infusion started. The total amount of hexamer is 93% of the library material and the AF for 1_32_3 = 2.04. (c) Hexamer composition after 17 redox turnovers (turnover time = 2 h) when the hexamer accounts for 75% of library material and the AF for 1_32_3 = 1.85. (d) Hexamer composition after 34 redox turnovers (turnover time = 1h) when the total amount of hexamer is 77% of library material and the AF for 1_32_3 = 1.73. (e) Relative abundances of hexameric self-replicators with different compositions over the course of the redox-infusion experiment with a turnover time of 1h, and (f) corresponding experiment with a turnover time of 2h.*

After 17 turnovers the hexamer distribution is still rich in the self-sorted 1_32_3 hexamer with an AF of 1.85 (Figure 5c,e). Even when the turnover rate is doubled the hexamer distribution remains enriched in 1_32_3 after 34 turnovers (Figure 5d,f). Since the efficiency of replication 1_32_3 is low compared to many of its competing replicators (vide infra), these data suggest that 1_32_3 has an increased resistance to reduction compared to these faster competitors. Support for this hypothesis comes from separate experiments where we investigated the reduction of the self-replicators by TCEP in more detail. When a DCL

11

containing a self-sorted distribution of **1** and **2** is partially reduced with TCEP, the fast replicators $1_5 2_1$ and $1_4 2_2$ are relatively more reduced than $1_3 2_3$ (See SI Section 14). However, $1_3 2_3$ is also reduced faster than $1_2 2_4$. It seems that selection for $1_3 2_3$ in the redox infusion regime is a consequence of two opposing trends: the rate of replication is higher for the **1**-rich replicators (vide supra), while the resilience to destruction is more pronounced for the **2**-rich replicators. Apparently $1_3 2_3$ becomes dominant as it combines a reasonable replication speed with a reasonable resilience to reduction. On top of that also steric zipper formation may contribute to its dynamic kinetic stability.

The steady-state replicator mutant distributions obtained under the different selection pressures (Figure 4d vs Figure 5c,d) are clearly different, yet obtained from a similar starting distribution. These results amount to a change of the replicator distribution through a process that has the hallmarks of Darwinian evolution: navigating replicator space through replication, mutation and selection.

Conclusions

Achieving Darwinian evolution in synthetic self-replicating systems requires that the replicators occupy only part of the structural space available to them and that the occupancy of this space changes as the selection pressure changes. We have demonstrated that mixing two building blocks, that individually form self-replicators, enlarges the structure space available to the self-replicators and can lead to various types of behavior. In most cases where two building blocks with a relatively similar structure are mixed, they are both incorporated in the self-replicator yielding a statistical distribution of all possible combinations within a certain macrocycle size (all of sequence space is populated). However, when specific interactions occur between building blocks, a departure from randomness through self-sorting can occur, which causes the system to occupy a select part of sequence space where these interactions are maximized. Specifically, in systems of self-replicators made from building blocks **1** and **2** we found that social self-sorting results in a strong amplification (compared to a random distribution) of the specific sequence isomer of $1_3 2_3$, where the two building blocks alternate in the macrocycle, as confirmed by fragmentation of isolated macrocycles using MALDI-TOF. This specific isomer is most likely stabilized by the formation of a steric zipper structure in which a phenyl ring of the phenylalanine residue of **1** is sandwiched between the two leucine residues of building block **2**, in an arrangement in which these two building blocks are next to each other in a macrocycle, while part of adjacent parallel β-sheets (formed by the stacking of the macrocycles) at the same time. Such steric zipper interaction patterns have been observed previously in amyloid formation,[59-60] albeit, for as far as we are aware, not yet noted for the specific leucine– phenylalanine – leucine arrangement that we discovered. Seeding experiments confirmed that the building block sequence was moderately heritable, and that sequence replication is error-prone, while the fidelity for copying ring size is high. We also demonstrated that this system shows signs of evolving when subjected to two different replication-destruction regimes. Depending on the selection pressure that is imposed, different replicator mutant distributions are selected for. When replicator "destruction" is mediated by outflow and therefore is non-selective, the replicators that are produced fastest become dominant. When destruction is mediated by reduction and is therefore potentially selective, replicator $1_3 2_3$, that combines an adequate speed of replication with an adequate resilience to destruction, while possibly also benefitting from steric zipper formation, becomes dominant. This behavior represents a rudimentary form of Darwinian evolution; it involves

12

replication, mutation and selection. However, in the present system the structure space available to the replicators is still limited and the extent to which the mutant distributions change is not dramatic. The next step is to enlarge replicator space further which would allow for more significant evolutionary changes, ultimately aiming for Darwinian evolution to become open-ended,[11, 62] where the opportunities for evolutionary inventions are effectively limitless and perpetual. The challenge is to ensure that replication remains sufficiently accurate as the number of accessible mutants increases.

Acknowledgments

Author contributions

M.J.E. and S.O. conceived the concept; M.J.E. designed the experiments and conducted them with the help of J.W. (who performed the seeding experiments at low stirring rate and redox-infusion experiments), K.L. (who did the mass flow experiments with photocatalyst), J.O. (who performed the experiments on the replicators showing statistical building blocks distributions and part of the TEM measurements) and A.K. (who performed part of the TEM measurements). Together, these authors analyzed the data. O.M. performed and analyzed the kinetic simulations; M.J.E. and S.O. wrote the manuscript.

Declaration of interests

The authors declare no competing interests.

References

1.   Otto, S., An Approach to the De Novo Synthesis of Life. *Acc. Chem. Res.* **2022,** *55* (2), 145-155.
2.   Adamski, P.; Eleveld, M.; Sood, A.; Kun, A.; Szilagyi, A.; Czaran, T.; Szathmary, E.; Otto, S., From self-replication to replicator systems en route to de novo life. *Nat. Rev. Chem.* **2020,** *4* (8), 386-403.
3.   Krishnamurthy, R.; Hud, N. V., Introduction: Chemical Evolution and the Origins of Life. *Chem. Rev.* **2020,** *120* (11), 4613-4615.
4.   Szostak, J., How Did Life Begin? *Nature* **2018,** *557* (7704), S13-S14.
5.   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.
6.   Pascal, R.; Pross, A., A Roadmap toward Synthetic Protolife. *Synlett* **2017,** *28* (1), 30-35.
7.   Kosikova, T.; Philp, D., Exploring the emergence of complexity using synthetic replicators. *Chem. Soc. Rev.* **2017,** *46* (23), 7274-7305.
8.   Le Vay, K.; Weise, L. I.; Libicher, K.; Mascarenhas, J.; Mutschler, H., Templated Self-Replication in Biomimetic Systems. *Adv. Biosyst.* **2019,** *3* (6), 1800313.
9.   Clixby, G.; Twyman, L., Self-replicating systems. *Org. Biomol. Chem.* **2016,** *14*, 4170-4184.
10.   Patzke, V.; von Kiedrowski, G., Self replicating systems. *Arkivoc* **2007,** *v*, 293-310.
11.   Duim, H.; Otto, S., Towards open-ended evolution in self-replicating molecular systems. *Beilstein J. Org. Chem.* **2017,** *13*, 1189–1203.

12.   Tjivikua, T.; Ballester, P.; Rebek, J., A self-replicating system. *J. Am. Chem. Soc.* **1990,** *112*, 1249-1250.

13.   Yao, S.; Ghosh, I.; Zutshi, R.; Chmielewski, J., A pH-modulated, self-replicating peptide. *J. Am. Chem. Soc.* **1997,** *119*, 10559-10560.

14.   Lee, D. H.; Granja, J. R.; Martinez, J. A.; Severin, K.; Ghadiri, M. R., A self-replicating peptide. *Nature* **1996,** *382*, 525-528.

15.   Rout, S. K.; Rhyner, D.; Riek, R.; Greenwald, J., Prebiotically Plausible Autocatalytic Peptide Amyloids. *Chem.-Eur. J.* **2022,** *28* (3), e202103841.

16.   Wang, B.; Sutherland, I. O., Self-replication in a Diels-Alder reaction. *Chem. Commun.* **1997**, 1495-1496.

17.   Terfort, A.; von Kiedrowski, G., Self-replication by condensation of 3-amino-benzamidines and 2-formylphenoxyacetic acids. *Angew. Chem. Int. Ed. Engl.* **1992,** *31*, 654-656.

18.   Bag, B. G.; von Kiedrowski, G., Stepwise replication of a Troger's base analogue. *Angew. Chem. Int. Ed.* **1999,** *38*, 3713-3714.

19.   Rubinov, B.; Wagner, N.; Rapaport, H.; Ashkenasy, G., Self-Replicating Amphiphilic beta-Sheet Peptides. *Angew. Chem. Int. Ed.* **2009,** *48*, 6683-6686.

20.   Kassianidis, E.; Philp, D., Design and implementation of a highly selective minimal self-replicating system. *Angew. Chem. Int. Ed.* **2006,** *45*, 6344-6348.

21.   Sadownik, J. W.; Philp, D., A simple synthetic replicator amplifies itself from a dynamic reagent pool. *Angew. Chem. Int. Ed.* **2008,** *47*, 9965-9970.

22.   von Kiedrowski, G., A self-replicating hexadeoxynucleotide. *Angew. Chem. Int. Ed. Engl.* **1986,** *25*, 932-935.

23.   Nowak, M. A., What is a quasispecies. *Trends Ecol. Evol.* **1992,** *7*, 118-121.

24.   Eigen, M.; Winkler-Oswatitsch, R., *Steps towards life*. Oxford University Press: Oxford, 1992.

25.   Gross, R.; Fouxon, I.; Lancet, D.; Markovitch, O., Quasispecies in population of compositional assemblies. *BMC Evol. Biol.* **2014,** *14*, 265.

26.   Hong, J. I.; Feng, Q.; Rotello, V.; Rebek, J., Competition, cooperation, and mutation - improving a synthetic replicator by light irradiation. *Science* **1992,** *255*, 848-850.

27.   Feng, Q.; Park, T. K.; Rebek, J., Crossover reactions between synthetic replicators yield active and inactive recombinants. *Science* **1992,** *256* (5060), 1179-1180.

28.   Saghatelian, A.; Yokobayashi, Y.; Soltani, K.; Ghadiri, M. R., A chiroselective peptide replicator. *Nature* **2001,** *409*, 797-801.

29.   Sadownik, J. W.; Mattia, E.; Nowak, P.; Otto, S., Diversification of self-replicating molecules. *Nat. Chem.* **2016,** *8*, 264-269.

30.   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.

31.   Lincoln, T. A.; Joyce, G. F., Self-Sustained Replication of an RNA Enzyme. *Science* **2009,** *323*, 1229-1232.

32.   van Rossum, S. A. P.; Tena-Solsona, M.; van Esch, J. H.; Eelkema, R.; Boekhoven, J., Dissipative out-of-equilibrium assembly of man-made supramolecular materials. *Chem. Soc. Rev.* **2017,** *46* (18), 5519-5535.

33.   Ragazzon, G.; Prins, L. J., Energy consumption in chemical fuel-driven self-assembly. *Nat. Nanotech.* **2018,** *13*, 882-889.

34.   Sorrenti, A.; Leira-Iglesias, J.; Markvoort, A. J.; de Greef, T. F. A.; Hermans, T. M., Non-equilibrium supramolecular polymerization. *Chem. Soc. Rev.* **2017,** *46* (18), 5476-5490.

35.   Weissenfels, M.; Gemen, J.; Klajn, R., Dissipative Self-Assembly: Fueling with Chemicals versus Light. *Chem* **2021,** *7* (1), 23-37.

36.   Pross, A., *What is life? How chemistry becomes biology*. Oxford University Press: Oxford, 2012.

37.   Pross, A.; Pascal, R., How and why kinetics, thermodynamics, and chemistry induce the logic of biological evolution. *Beilstein J. Org. Chem.* **2017,** *13*, 665-674.

38.   Bandela, A. K.; Wagner, N.; Sadihov, H.; Morales-Reina, S.; Chotera-Ouda, A.; Basu, K.; Cohen-Luria, R.; de la Escosura, A.; Ashkenasy, G., Primitive selection of the fittest emerging through functional synergy in nucleopeptide networks. *Proc. Natl. Acad. Sci. U. S. A.* **2021,** *118* (9), e2015285118.

39.   Maity, I.; Wagner, N.; Mukherjee, R.; Dev, D.; Peacock-Lopez, E.; Cohen-Luria, R.; Ashkenasy, G., A chemically fueled non-enzymatic bistable network. *Nat. Commun.* **2019,** *10*, 4636.

40.   Yang, S.; Schaeffer, G.; Mattia, E.; Markovitch, O.; Liu, K.; Hussain, A. S.; Ottelé, J.; Sood, A.; Otto, S., Chemical fueling enables molecular complexification of self-replicators. *Angew. Chem. Int. Ed.* **2021,** *60*, 11344-11349.

41.   Liu, B.; Wu, J. T.; Geerts, M.; Markovitch, O.; Pappas, C. G.; Liu, K.; Otto, S., Out-of-Equilibrium Self-Replication Allows Selection for Dynamic Kinetic Stability in a System of Competing Replicators. *Angew. Chem.-Int. Edit.* **2022,** *61* (18), e202117605.

42.   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.

43.   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.

44.   Maity, S.; Ottele, J.; Santiago, G. M.; Frederix, P.; Kroon, P.; Markovitch, O.; Stuart, M. C. A.; Marrink, S. J.; Otto, S.; Roos, W. H., Caught in the Act: Mechanistic Insight into Supramolecular Polymerization-Driven Self-Replication from Real-Time Visualization. *J. Am. Chem. Soc.* **2020,** *142* (32), 13709-13717.

45.   Colomb-Delsuc, M.; Mattia, E.; Sadownik, J. W.; Otto, S., Exponential self-replication enabled through a fibre elongation/breakage mechanism. *Nat. Commun.* **2015,** *6*, 7427.

46.   Schaeffer, G.; Eleveld, M. J.; Ottelé, J.; Kroon, P. C.; Frederix, P. W. J. M.; Yang, S.; Otto, S., Stochastic Emergence of Two Distinct Self-Replicators from a Dynamic Combinatorial Library. *J. Am. Chem. Soc.* **2022,** *144* (14), 6291-6297.

47.   Altay, M.; Altay, Y.; Otto, S., Parasitic Behavior of Self-Replicating Molecules. *Angew. Chem. Int. Ed.* **2018,** *57* (33), 10564-10568.

48.   Benner, S. A., Paradoxes in the Origin of Life. *Orig. Life Evol. Biosph.* **2014,** *44* (4), 339-343.

49.   Wu, A. X.; Isaacs, L., Self-sorting: The exception or the rule? *J. Am. Chem. Soc.* **2003,** *125*, 4831-4835.

50.   He, Z. F.; Jiang, W.; Schalley, C. A., Integrative self-sorting: a versatile strategy for the construction of complex supramolecular architecture. *Chem. Soc. Rev.* **2015,** *44* (3), 779-789.

51.   Northrop, B. H.; Zheng, Y. R.; Chi, K. W.; Stang, P. J., Self-Organization in Coordination-Driven Self-Assembly. *Acc. Chem. Res.* **2009,** *42* (10), 1554-1563.

52.   Hsu, C. W.; Miljanic, O. S., Self-sorting through dynamic covalent chemistry. In *Dynamic covalent chemistry: principles, reactions and applications*, Zhang, W.; Jin , Y., Eds. Wiley: 2018; pp 253-286.

53.   Ottele, J.; Hussain, A. S.; Mayer, C.; Otto, S., Chance emergence of catalytic activity and promiscuity in a self-replicator. *Nat. Catal.* **2020,** *3* (7), 547-553.

54.   Santiago, G. M.; Liu, K.; Browne, W. R.; Otto, S., Emergence of light-driven protometabolism on recruitment of a photocatalytic cofactor by a self-replicator. *Nat. Chem.* **2020,** *12* (7), 603-607.

55.   Markovitch, O.; Ottele, J.; Veldman, O.; Otto, S., Automated device for continuous stirring while sampling in liquid chromatography systems. *Comm. Chem.* **2020,** *3* (1), 180.

56.   Saha, M. L.; Schmittel, M., Degree of molecular self-sorting in multicomponent systems. *Org. Biomol. Chem.* **2012,** *10* (24), 4651-4684.

57.   Biancalana, M.; Makabe, K.; Koide, A.; Koide, S., Molecular Mechanism of Thioflavin-T Binding to the Surface of beta-Rich Peptide Self-Assemblies. *J. Mol. Biol.* **2009,** *385* (4), 1052-1063.

15

58.   Wu, C.; Wang, Z. X.; Lei, H. X.; Duan, Y.; Bowers, M. T.; Shea, J. E., The Binding of Thioflavin T and Its Neutral Analog BTA-1 to Protofibrils of the Alzheimer's Disease A beta(16-22) Peptide Probed by Molecular Dynamics Simulations. *J. Mol. Biol.* **2008,** *384* (3), 718-729.

59.   Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A. O.; Riekel, C.; Eisenberg, D., Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* **2007,** *447* (7143), 453-457.

60.   Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D., Structure of the cross-beta spine of amyloid-like fibrils. *Nature* **2005,** *435* (7043), 773-778.

61.   Park, J.; Kahng, B.; Hwang, W., Thermodynamic Selection of Steric Zipper Patterns in the Amyloid Cross-beta Spine. *PLoS Comput. Biol.* **2009,** *5* (9), e1000492.

62.   Ruiz-Mirazo, K.; Umerez, J.; Moreno, A., Enabling conditions for 'open-ended evolution'. *Biol. Philos.* **2008,** *23* (1), 67-85.