Check for updates

From self-replication to replicator systems en route to de novo life

Paul Adamski[1], *Marcel Eleveld*[1], *Ankush Sood*[1], *Ádám Kun*[2,3,4], *András Szilágyi*[2,3,4], *Tamás Czárán*[2,4], *Eörs Szathmáry*[2,3,5]✉ *and Sijbren Otto*[1]✉

Abstract | The process by which chemistry can give rise to biology remains one of the biggest mysteries in contemporary science. The de novo synthesis and origin of life both require the functional integration of three key characteristics — replication, metabolism and compartmentalization — into a system that is maintained out of equilibrium and is capable of open-ended Darwinian evolution. This Review takes systems of self-replicating molecules as starting points and describes the steps necessary to integrate additional characteristics of life. We analyse how far experimental self-replicators have come in terms of Darwinian evolution. We also cover models of replicator communities that attempt to solve Eigen's paradox, whereby accurate replication needs complex machinery yet obtaining such complex self-replicators through evolution requires accurate replication. Successful models rely on a collective metabolism and a way of (transient) compartmentalization, suggesting that the invention and integration of these two characteristics is driven by evolution. Despite our growing knowledge, there remain numerous key challenges that may be addressed by a combined theoretical and experimental approach.

Life, as we know it, is bewilderingly complex. It is not clear how life originated or how its complexity came about. Moreover, it is unclear whether it is possible to synthesize life de novo from simple chemical components. These questions are among the grand challenges of contemporary science and at the heart of systems chemistry[1–4]. Although topical, we will not aim to answer the question of the origin of life, nor will we address issues related to prebiotic plausibility. Instead, we will specifically cover research that targets the de novo synthesis of life. In addressing this challenge, we are partially guided by extant biochemistry but are certainly neither constrained by it nor will we necessarily converge on it. We identify the concepts and challenges in life's de novo synthesis and argue that many of these also extend to the origin of life.

Although life is remarkably difficult to define[5–7], every living system exhibits metabolism, is able to reproduce and is separated from its environment (FIG. 1a). Metabolism involves the harvesting of energy, which is required because living systems are dissipative — they require energy input for their maintenance. Along with energy conversions, metabolism also involves conversions of matter to afford building blocks, which enable self-maintenance and reproduction. During reproduction, the system makes copies of itself with sufficient accuracy that the integrity of the species is maintained across generations. Nevertheless, reproduction is sufficiently error-prone so as to allow for Darwinian evolution through mutation and selection. Finally, compartmentalization keeps the components of a living system together and separate from the environment.

On a coarse level, synthesizing life requires the functional integration of replication[8–14], compartmentalization[15–19] and metabolism[20–25] into a system that remains out of thermodynamic equilibrium (FIG. 1a). Preferably, a population of such synthetic systems may also have the capacity to undergo Darwinian evolution in an open-ended sense, because evolvability is the ultimate hallmark of the living world. Living systems are out of equilibrium in that they experience effectively irreversible processes of reproduction and degradation. This repeated process of formation and destruction is driven by continual material and/or energy input, giving rise to a non-equilibrium steady state, referred to as dynamic kinetic stability[26–30]. As we detail below, present efforts towards the development of de novo life focus on integrating replicative, compartmentalized and metabolic subsystems (initially targeting different binary combinations), and developing out-of-equilibrium systems, ultimately to enable Darwinian evolution.

We will begin our discussion by considering self-replicating molecules, a topic that is one of many that could be taken as a starting point for life (FIG. 1a). As the problem of the synthesis of life is underdetermined, it is not yet clear what the most appropriate approach is.

[1]Centre for Systems Chemistry, Stratingh Institute, University of Groningen, Groningen, Netherlands.

[2]Institute of Evolution, MTA Centre for Ecological Research, Tihany, Hungary.

[3]Parmenides Center for the Conceptual Foundations of Science, Parmenides Foundation, Pullach, Germany.

[4]MTA-ELTE Theoretical Biology and Evolutionary Ecology Research Group, Eötvös University, Budapest, Hungary.

[5]Department of Plant Systematics, Ecology and Theoretical Biology, Eötvös University, Budapest, Hungary.

✉e-mail: szathmary.eors@gmail.com; s.otto@rug.nl

https://doi.org/10.1038/s41570-020-0196-x

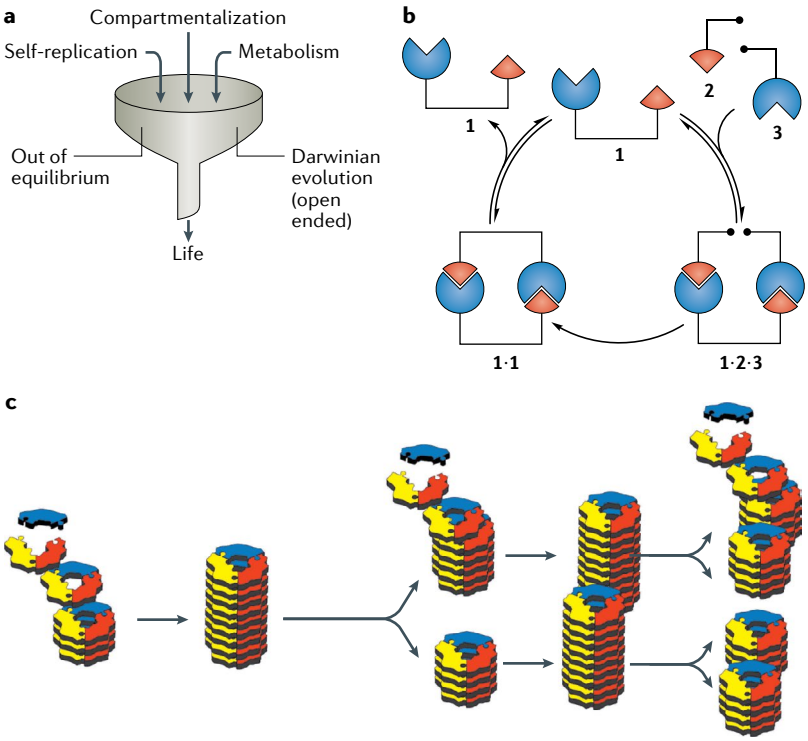


Fig. 1 | **Fundamental features of life and mechanisms of self-replication. a** | De novo life involves the functional integration of replication, metabolism and compartmentalization. The conditions must enable open-ended Darwinian evolution and the system must remain out of equilibrium, in a state of dynamic kinetic stability. **b** | The first mechanism of self-replication involves duplex formation, whereby replicator **1** binds building blocks **2** and **3** and templates their conversion into a single copy of **1**. Although only two recognition sites are shown here, the same mechanism that affords dimers can in principle yield oligomers. **c** | The second mechanism of self-replication features supramolecular polymerization. The replicator-catalysed ligation of building blocks gives rise to a stack of replicator copies, which can exhibit exponential replication on entering a stack growth–breakage regime. The example here is a cyclic oligomer but the same mechanism may also yield linear oligomers[54].

Systems chemistry
The study of properties that emerge from mixtures of interacting molecules. One of the key foci is the analysis and synthesis of diverse autocatalytic systems and their possible couplings.

Metabolism
Chemical processes that form the constituents of a living system from (often simple) raw materials (the food set) and connect the internal maintenance of the system to an external energy source.

Darwinian evolution
Evolution by natural selection that requires units that multiply and have heredity and variability. There should be hereditary traits that affect the chance of reproduction and/or survival of the units.

In the context of this Review, we define self-replication as the ability of a system to autonomously catalyse its copying, such that information in the system components is transferred to the next generation. This autocatalysis typically takes the form of replication templated by specific non-covalent interactions between information-containing molecules. Our treatment will focus on self-replicating systems comprising completely synthetic molecules as well as bio-inspired systems featuring peptides and nucleic acids. Darwinian evolution of self-replicators requires that there be different kinds of replicators that keep their identities during the process. This property, referred to in biology as heredity, is not exact, such that there is variability in the population. Entities that multiply, exhibit heredity and show variability are regarded as units of evolution, a population of which can undergo evolution by natural selection if hereditary traits influence the survival and/or multiplication of the units[31]. For systems to exhibit stable evolvability, there are specific quantifiable conditions that need to be met, including a minimum accuracy of replication.

Molecules that self-replicate do so by one of two distinct mechanisms. The first of these has the minimal requirement that self-replicator **1** reversibly binds its building blocks **2** and **3**, thereby bringing the reactive ends of these compounds together, such that they condense to give **1** as part of the replicator duplex **1·1** (FIG. 1b). On dissociation of **1·1**, two replicator molecules become available for the next cycle of replication. This mechanism is related to but simpler than the way DNA is replicated, which occurs through cross-catalysis (not shown). In the simplest form of cross-catalysis, **1** induces the formation of a complementary molecule **1′** by binding the precursors **2′** and **3′**, while, in turn, **1′** templates the formation of **1** from **2** and **3**. Where the mechanism shown in FIG. 1b has a tendency to halt at the stage of the duplex, another template-based replication mechanism allows additional rounds of replication without requiring replicators to dissociate from one another (FIG. 1c). This mechanism affords assemblies of multiple replicators, and when these assemblies are large enough they become susceptible to mechanically induced breakage. This growth–breakage mechanism enables exponential replication, a property that has important implications for Darwinian evolution.

Although autocatalysis is essential for self-replication, not all autocatalytic systems are self-replicating. For example, autopoiesis involves assemblies of molecules (typically micelles or vesicles) that reproduce autocatalytically[18,32–35] but are not self-replicating because the interactions between the molecules in the assemblies lack the specificity required for molecular-level information transfer. Likewise, some autocatalytic networks, including the formose reaction[36], are not self-replicating because they lack the ability to transfer information in the molecules, which is required to exhibit heredity[10]. The reader should note that the terminology used in this field can vary, and scientists may speak of replication in the broad sense and informational replication in a narrower sense[37]. The latter case is then referred to as hereditary replication[38].

Having defined the scope of this Review, let us now briefly explain its structure. The first section describes the features that self-replicating systems need to begin to resemble life and the challenges associated with incorporating these features. The path from individual self-replicators to de novo life most likely involves communities of different and interacting replicators, so the second section summarizes the insights obtained from modelling (theoretical) replicator communities. The stage is then set to survey experimental progress, and we, in particular, discuss advances towards Darwinian evolution, replicator community dynamics and the integration of replication with metabolism and compartmentalization.

Steps in the path from self-replication to life

The ability to self-replicate is a necessary but insufficient condition for life, which requires additional characteristics to be assimilated[39]. The most obvious characteristic is the ability to undergo Darwinian evolution, wherein replication proceeds with mutation and the resulting mutants undergo competitive selection. Incorporating Darwinian evolution into systems of self-replicators requires different building blocks to be present such that different (mutated) offspring can form. Moreover,

Compartmentalization
A system that enables spatial gradients (whereas chemists often consider bulk, well-stirred systems). Passive compartmentalization can be provided by absorptive surface and rock pores. Active compartmentalization rests on boundaries (such as membranes created through autopoiesis).

we need to introduce a mechanism of selection. In biology, selection occurs under conditions in which total resources are finite, such that population growth is eventually balanced by death. Both replication (growth) and death (decay) can depend on external (environmental) conditions, thus enabling the fittest molecular species — that with the highest replication rate and/or lowest death rate — to displace all others. In a purely competitive situation, this follows exponential exclusion kinetics and truly is 'survival of the fittest'. Systems of self-replicating molecules reported thus far have mostly failed to incorporate mechanisms of death. Death can be introduced by physically removing replicators or by degrading them chemically. The most popular way to incorporate death is through serial transfer[40–42], in which a small fraction of a replicator-containing sample is transferred to a fresh solution of building blocks several times. In the limit of many such transfers, the system starts to resemble a continuously stirred tank reactor in which a solution of building blocks flows in and the resulting replicators flow out at another location at the same flow rate[43].

Unlike in a closed system, the populations of replicators in a replication–destruction system are not necessarily governed by the thermodynamic stability of the individual replicators, so the predominant replicator need not be the most thermodynamically stable. This departure from thermodynamically controlled replicator distribution is possible owing to the coupling of the system to the energy source (in the form of appropriate reagents) that drives the replication–destruction process, making it an open system. Thus, a replication–destruction regime can have an out-of-equilibrium character, which has been described in terms of dynamic kinetic stability[26–30] determined by a balance between rates of replication and destruction of individual replicators. The resulting replicator populations, when they have the capacity to mutate, can transition between different (steady) states and lead to quasi-species, defined by Eigen as the winning subset of replicators in mutation–selection balance[44,45].

An important aspect of Darwinian evolution and ecology is the competitive exclusion principle, which states that a given niche can only be stably occupied by one species. If two species compete for the same resource, only the fittest survives. The same principle can also hold mathematically for replicators that compete for common building blocks but only when the kinetic order of the replicator in the replication reaction equals (or exceeds) its kinetic order in the destruction process[42,46]. Thus, because the destruction process is normally first order in the replicator being destroyed, the replication reaction also needs to be (at least) first order in the replicator; this implies a need for exponential replication. However, the vast majority of self-replicators **1** that operate by the complexation–dissociation mechanism (FIG. 1b) have a rate of replication that is only of order 0.5 in [**1**] (the square-root law of autocatalysis[47]), which is a consequence of self-inhibition in view of the duplex **1·1** needing to dissociate before its components can replicate. Replicators obeying the square-root law of autocatalysis exhibit parabolic growth dynamics. Such replicators, when competing for a common resource, continue to coexist indefinitely, limiting their potential for Darwinian evolution[46,48] (BOX 1).

Another highly counterintuitive, yet potentially common, mechanism leading to parabolic growth dynamics and replicator coexistence[49,50] is competition in an open chaotic flow (OCF)[51,52] regime. The OCF model has found prebiotic relevance through the phenomenon of thermophoresis, whereby molecules (such as nucleotides and DNA) can accumulate in a fluid owing to a temperature gradient that may be present in a thermal vent[53]. However, thermophoresis in general does not necessarily involve chaotic flows. The fractal nature of the differently populated fluid domains renders replicator growth

Box 1 | **Replicator growth and selection consequences**

Replicator populations with access to constant resources exhibit a rate of offspring production at a given time t that is proportional to the concentration of replicators [A] (Eq. 1).

$$\frac{d[A]}{dt} = r[A] \quad (1)$$

Here, r denotes the Malthusian parameter for the population, which is the intrinsic growth rate. Solving this differential equation yields the well-known exponential growth equation (Eq. 2).

$$[A] = [A]_0 e^{rt} \quad (2)$$

This potentially steep curve is the fundamental engine of evolutionary selection. The quotient $[A_1]/[A_2]$ of two exponentially growing populations with Malthusian parameters r_1 and r_2, respectively, also increases or declines exponentially. In this way, the inferior species becomes diluted and competitively excluded over a very short time, even without growth limitation and even if the difference between the growth rates is small. Growth behaviour is dramatically different if replication is template-directed and the newly formed copy reversibly inhibits the template (FIG. 1b). The copy remains associated with its template for a while in an inert complex that maintains an equilibrium with its dissociated and potentially replicating components (Eq. 3).

$$A + R_1 + R_2 \xrightarrow{k_1} A \cdot A \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} 2A \quad (3)$$

Here, R_1 and R_2 denote the resources for replicating species A, whereas k_1, k_2 and k_{-2} are the rate constants of replication, association and dissociation, respectively. If $k_1 < k_2 \ll k_{-2}$, as would be physico-chemically plausible assuming that dissociation is the slowest process, the system becomes self-regulated, such that the higher the concentration [A], the stronger its inactivation by dimerization (Eq. 4).

$$\frac{d[A]}{dt} = r[A]^p \quad (4)$$

where $r = k_1(k_2/k_{-2})^{1/2}$ (REF.[130]) and p denotes the kinetic order. When $p = 1$ there is exponential growth, and $0 < p < 1$ corresponds to the parabolic regime. In almost all experimentally investigated systems, $p \approx 1/2$, such that growth is quadratic in time[130] (Eq. 5).

$$[A] = ([A]_0^{1/2} + rt/2)^2 \quad (5)$$

Dilute conditions favour dissociation of inert dimers into active monomeric replicators[46], enabling the system to maintain an arbitrarily diverse set of replicators with different Malthusian parameters, thus preventing Darwinian selection. Lastly, if $p = 2$ then we predict hyperbolic growth (Eq. 6).

$$[A] = ([A]_0^{-1} - rt)^{-1} \quad (6)$$

In contrast to the previous cases, hyperbolic growth is so fast that the population size diverges in a finite time, $t_c = (r[A]_0)^{-1}$. If multiple species are initially present, the one with the smallest t_c will eventually predominate because not only fitness r_i but also the initial abundance $[A_i]_0$ of each species i determine which species wins.

Out of equilibrium
Any state that is not at equilibrium.

Dynamic kinetic stability
A persistent state of an open chemical system resulting from a cyclic process of formation (replicative or otherwise) and destruction, occurring effectively irreversibly (that is, formation and destruction reactions are kinetically directed and not each other's microscopic reverse), driven by continual material and/or energy input.

Self-replication
The ability of a system to autonomously catalyse the formation of copies of itself, such that information contained in the molecules that constitute the system is transferred to the next generation.

Exponential replication
An autocatalytic process with constant per-capita growth. Exponential replication leads to infinite concentration in infinite time. In competition, it entails survival of the fittest.

Autopoiesis
A complex process in which a system is able to produce more of itself and its constituent molecules.

Quasi-species
The weighted distribution of mutants centred around one or several master sequences in a mutation–selection balance. The quasi-species is the target of selection in a system of replicating individuals who replicate without cooperating with one another.

Competitive exclusion principle
The principle that, in a replication–destruction regime harbouring self-replicators capable of exponential growth that compete for the same precursors (from which they replicate), one replicator will drive all others to extinction.

Eigen's paradox
The paradox that accurate replication needs complex machinery, yet obtaining such complex self-replicators through evolution requires sufficiently accurate replication.

kinetics, in effect, parabolic in OCF, again ensuring the survival of all species. This absence of selection affords no evolutionary change, so replication in OCFs, similar to any other process yielding parabolic population growth, can be responsible only for temporary replicator diversification that has, sooner or later, to be followed by selection through a different mechanism. Only a few experimental examples of self-replicators capable of exponential growth exist, and they mostly do so by supramolecular polymerization[54,55] (FIG. 1c).

To date, the fidelities (replication accuracies) of synthetic self-replicating systems rely solely on molecular recognition and lack the sophisticated error-correction machinery that promotes fidelity in DNA replication. Thus, error-prone replication appears unavoidable and must be accommodated in any scenario that involves self-replicators becoming more complex, which requires a greater amount of information to be copied during replication. Increasing the amount of information in self-replicators leads to Eigen's paradox — self-replicators must contain a lot of information to replicate accurately, yet obtaining self-replicators containing a lot of information already requires accurate replication. Several solutions to this chicken-and-egg problem have been proposed, all of which involve communities of replicators that each store limited information but are able to cooperate such that, collectively, they contain and can replicate a large amount of information with sufficient accuracy. This notion provides a strong impetus to develop communities of replicators and study their collective dynamics.

Darwinian evolution is one of the most powerful engines of invention. Yet it is still largely unclear how this creative potential can be exploited in synthetic self-replicators. Making systems evolve by replication, mutation and selection is possible, but the discovery of new functions from these systems remains rare. The autonomous, continuous, never-ending invention of new functions is central to the idea of open-ended evolution[56,57], which represents another key feature of life. Among the most desirable functions that could be invented by self-replicating systems is the ability to catalyse other reactions. Specifically, when replicators acquire the ability to catalyse reactions that benefit their replication efficiency (for example, by converting materials in their environment into resources from which they replicate), they start to acquire metabolism. Another evolutionarily desirable invention would be genotype–phenotype separation. Such separation boosts the potential for further inventions because it allows genotype evolution to become less constrained by being partially decoupled from phenotype fitness. Whereas the genotype and phenotype in extant biochemistry are linked through the genetic code, there are perhaps other, simpler, mechanisms that allow for this division of labour.

Besides open-ended evolution and out-of-equilibrium conditions, the synthesis of life requires two additional features: metabolism and compartmentalization. All these features need to be incorporated into a chemical supersystem to create a minimal form of life[5,26]. The most obvious approach to minimal life is to proceed stepwise and first target the functional integration of two features, resulting in infrabiological systems[58], which are not living but exhibit exciting life-like properties[59]. To acquire metabolism, self-replicators need to catalyse not only their own formation but also the formation of the building blocks from which they replicate (the first experimental examples of such behaviour have recently been reported and are described below[60,61]). The second aspect of metabolism — sustaining a non-equilibrium state — may initially be satisfied by the replication–destruction regime imposed on the system by its environment. At a later stage, it would be desirable for systems of replicators to tap into energy sources to drive endergonic reactions associated with system maintenance. Also, the functional coupling of replication with compartmentalization needs to be achieved. Although it is relatively straightforward to house self-replicating systems in vesicular compartments, coupling the replication process with compartment growth and division is an unsolved problem that would benefit from more investigation[62].

Theory of replicator community dynamics

The important insights obtained from theoretical studies can guide experimental work on replicator community dynamics, a field that is presently still in its infancy. When replicators, such as biological organisms, are members of populations, they can constitute ecological communities. The interaction between populations can be positive or negative. Thus, if the presence of A decreases the density of population B, then A has a negative effect on B. If A increases the density of B, then this effect on B is positive. Interactions are not necessarily symmetrical. Elementary combinatorics covers the cases of competition (−,−), predation/parasitism (+,−) and mutualism (+,+). All of these dynamics have been observed in experimental systems of replicators.

The presence of two or more different replicators in a community does not violate the competitive exclusion principle if the different replicators each occupy different food niches by requiring different resources/precursors for their growth. The apparent contradiction between the ecological principle of competitive exclusion and the evolutionary requirement of sustaining a sufficiently diverse set of replicators in spite of their common food source has been the main concern with origin-of-life models. Almost all theoretical attempts to resolve the paradox invoke mutualistic (cooperative) interactions between replicator species, showing that the dynamic effect of cooperation can overrule the destructive power of both competition and parasitism to maintain replicator coexistence.

A historically important model of replicator community dynamics is the hypercycle, proposed by Eigen as a solution to the paradox that bears his name[44]. In hypercyclic coupling, molecular species A and B help the replication (rather than the formation) of each other (BOX 2). This resembles a mutualistic link in biology, exemplified by plant–pollinator systems. The name indicates that the replication cycle of each species is further catalysed by the other species, and this further aid also forms a topological cycle. Each member is an autocatalyst and a heterocatalyst at the same time, so replication kinetics

Box 2 | **Hypercyclic and metabolic replicators**

According to Eigen and Schuster[99] there is a strict upper limit on the amount of information sustainable in a simple mutation–selection dynamics context: $L < \ln\sigma/\mu$. Here L is the maximum sustainable length of a master sequence with per-base mutation rate μ and selection advantage σ over its mutants. Prebiotic replication without enzymatic catalysis must have been inaccurate ($\mu \approx 0.01$), such that L was limited to ~100, which is too short to code for even a primordial genome with a small number of different functions. Eigen and Schuster suggested to resolve this error threshold[127] problem by assuming that each replicator should specifically catalyse the replication of another replicator and receive similar catalytic aid from yet another, in a cyclic topology. The dynamics of an n-membered hypercycle are formalized (Eq. 7).

$$\frac{d[A_i]}{dt} = [A_i](k_i[A_{i-1}] - \Phi) \quad (7)$$

where $[A_i]$ is the concentration and k_i the replication rate constant of species i ($1 \le i \le n$), and Φ is the non-selective efflux of replicators that keeps the total replicator concentration constant ($[A] = [A]_0$, where $[A]$ is the total concentration of all replicators). Small hypercycles ($n \le 4$) converge to fixed points[138] but larger systems oscillate[139–141]. The hypercycle is vulnerable in an evolutionary sense[63,142] because selfish mutants that accept but do not give catalytic aid may proliferate and destroy the community. Shortcut parasites help a distant member of the hypercycle instead of their dedicated target, thus reducing the length of the hypercycle and decreasing its information content[64,75,143–145].

Another distinct model is the metabolic replicator, whose dynamics in a spatially homogeneous, well-mixed setting can also be formalized (Eq. 8).

$$\frac{d[A_i]}{dt} = [A_i](r_i M - \Phi) \quad (8)$$

Here, M represents the flux of a common metabolic process determined by values of $[A_i]$, with Φ denoting the excess production. In a well-mixed system, the fastest replicator will exclude all others, such that metabolism is defunct. Thus, because the metabolic help supplied by the community is the same for all replicators in the system, there is no coexistence because of competitive exclusion. With spatial inhomogeneity that might arise from transient compartmentalization or surface-bound dynamics, useful replicator species coexist and have sufficient resistance against parasites. This evolutionary stability comes from the self-thinning property of parasites. Indeed, a parasite cannot participate in the metabolism, so the lack of metabolic completeness kills all replicators in a small area around the parasite whereas metabolically complete local groups survive[51,67,68,146].

is overall second order — as the replication rate depends on the product of the concentrations of the template and its helper species. It has been shown that this system is ecologically stable, despite the fact that the rate constants can be arbitrarily different[57]. It took a while to realize that the system is, unfortunately, evolutionarily unstable[63] because a parasite A′ that does not help B but grows (accepting the help by B) faster than A would kill the entire system in a continuously stirred tank reactor.

The parasite problem of the hypercycle model has been addressed through different spatially explicit models of ecological communities. One solution to the problem would be to grow hypercycles on a surface, such as on a mineral, for example[64]. However, this model is insufficiently robust to perturbations, including the patchy distribution of rates of replicator desorption from the surface.

Open-ended evolution
A process whereby Darwinian replicator evolution proceeds indefinitely in a non-trivial manner. It may come in three forms: weak, strong and ultimate. In the weak form, novel phenotypes (not seen before, perhaps a new form of beak on a bird) arise indefinitely. The strong form requires evolutionary innovations, such as a novel catalytic or motor activity. The ultimate form allows for a major transition to occur, with the emergence of higher units of evolution from lower ones, such as reproducing protocells from replicating molecules.

Metabolically coupled replicator community models

We now consider three general mechanisms that can give rise to dynamic coexistence in replicator communities even in the presence of parasites. These models assume an explicit metabolic (mutualistic) coupling between the replicators and, essentially, implement the RNA world scenario in different spatially resolved ways, by means of either transient compartmentalization (FIG. 2a), spreading on a surface combined with metabolic coupling (FIG. 2b) or reproducing compartments known as protocells (FIG. 2c).

Transient compartmentalization model. In the transient compartmentalization model, a community of replicators is subject to local replication–global dispersion cycles[52,65,66] (FIG. 2a). For example, replication occurs in local groups that inhabit mineral patches or pores, and then the groups are washed away, get mixed and, subsequently, become re-localized for replication, where local groups are assumed to form as random samples of the global pool. Evolutionary survival requires the 'helper' molecule to 'feel' its own presence, even in the presence of parasites[52]. Thus, a single replicase molecule does not qualify: it cannot replicate itself because it needs a copy of itself as a template. By contrast, a molecule (such as a ribozyme) that catalyses a metabolic reaction to give monomers useful for its own replication does qualify. Indeed, even a single molecule can 'scratch its own back' in this way, as has recently been observed experimentally (see below). Importantly, parasite-only groups are infertile. The bottom line is that although altruistic replicators suffer from a relative replicative disadvantage within local groups, groups with more altruists ultimately contribute more to the global pool. Parasites are typically not completely displaced, but they are kept at bay, as has been observed in chemical experiments (see below). Replicators contributing to a common good can also be maintained by this mechanism[66].

Metabolic cooperation on mineral surfaces. Surface confinement alone cannot fully rescue hypercycles from parasitic invasion. Spatial confinement, together with metabolic coupling, has proven more successful. This is evident from simulations of a metabolically (rather than hypercyclically) coupled replicator system (MCRS)[67–69] (FIG. 2b). This model involves different replicators aiding the community without hypercyclic coupling, with each replicator being assumed to contribute to a common metabolism in a multiplicative synergistic manner. Replicators can only grow if there is a full set of complementary partners in a local region referred to as a metabolic neighbourhood. A locally balanced composition entails more efficient metabolism. Of course, spatial confinement is not at play in a well-mixed flow model, in which this system collapses (BOX 2). This is not the case if we model it as a reaction–diffusion system in 2D, as is implemented in cellular automata[64]. The reaction is the replication of molecules as a function of metabolic neighbourhood and the replication rate of the focal molecule. Diffusion only happens in the two dimensions of the surface and the system is inherently stochastic. Modelling shows that this environment allows for the stable coexistence of replicators, owing to the dual effects of a cost of commonness and an advantage of rarity. These dynamics emerge because a fast replicator faces the risk of a lack of metabolic complementation in a local neighbourhood (if it overgrows, it will be locally dead),

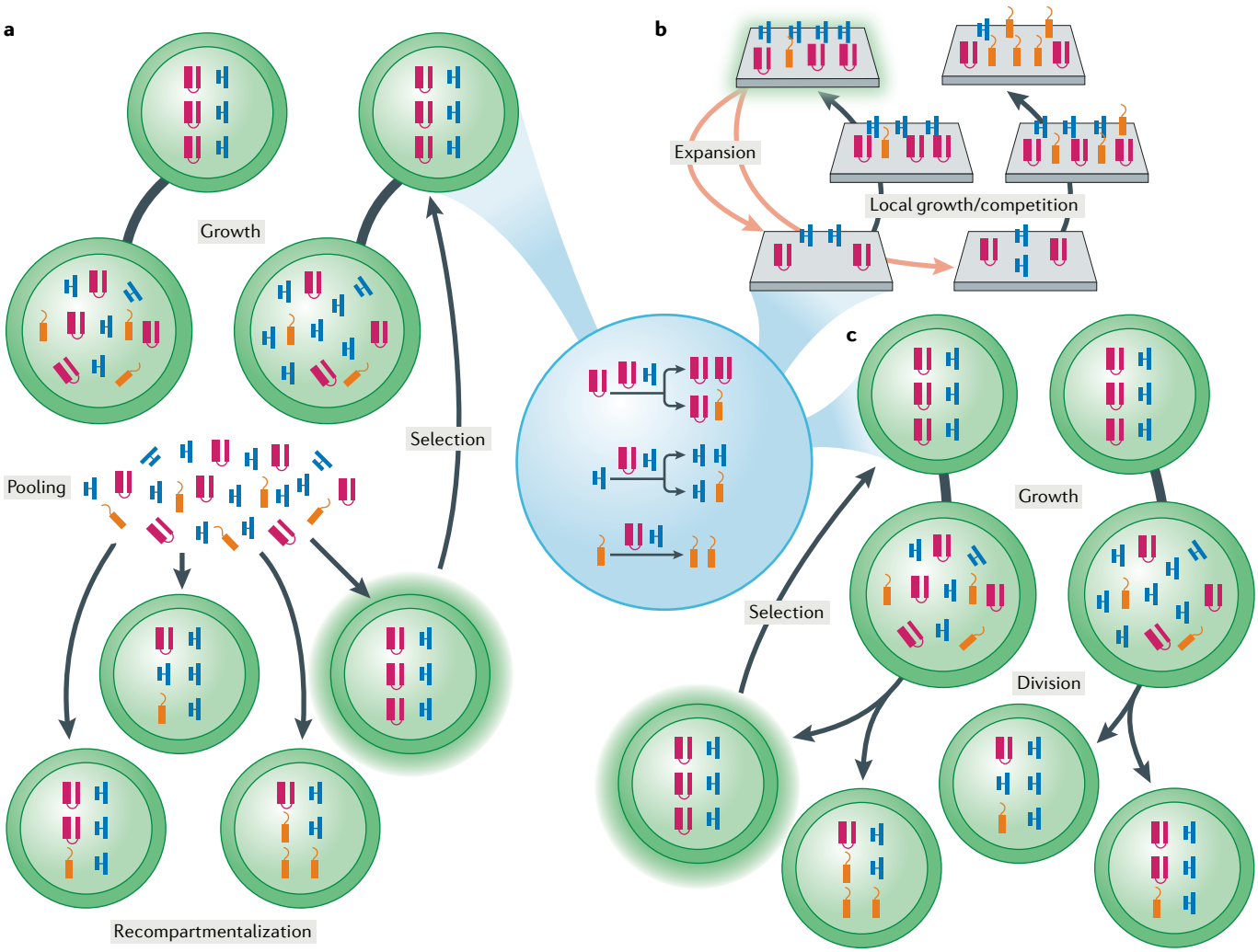


Fig. 2 | **Evolutionarily robust models of replicator communities. a** | In the transient compartmentalization model, selection acts on small temporary random samples of the metabolic replicator pool to sustain replicator diversity. **b** | The metabolically coupled replicator system model involves local group selection, which provides negative feedback on the densities of metabolic replicators to allow them to coexist. **c** | The reproducing compartment model features autopoietic compartment-level reproduction, such that selection maintains lineages of protocells harbouring constitutively mutualistic replicator sets. Each model resists destruction from parasitic replicators (orange symbols) produced by mutation. Blue and magenta symbols represent functional metabolic replicator species.

Predation
The process in which one replicator consumes another (the prey) for its own replication. The predator population benefits, and the prey population suffers.

Mutualism
An ecological coupling between two populations from which both benefit. Analogous to a two-membered hypercycle.

Hypercycle
A replicator set in which the autocatalytic replication of each member is heterocatalytically aided by another member conforming to cyclic topology.

Error threshold
The critical value of the mutation rate, above which errors accumulate and soon lead to the complete loss of information (error catastrophe) upon multiple rounds of replication. Stable selection requires that the error rate lies below the error threshold.

whereas a slow replicator is more likely to be locally complemented by the molecules with complementary functions. The rate of diffusion in an MCRS does not matter because the system converges to the transiently compartmentalized system.

The rate of diffusion matters a lot, however, if one is to maintain a hypercycle-like system. The minimalist and most realistic version of the hypercycle is a self-replicase that can copy another instance of itself. By virtue of replication, in this case, being error-prone, we face the Eigen problem in a more pronounced form — both the template and the replicase activity of the molecule can be adversely affected by mutations. But growth and spreading on a surface come to the rescue, because even if there is a three-way trade-off between template efficiency, replicase speed and accuracy (as a worst case), efficient replicases emerge, along with an enzymatically deficient mutant parasite cloud[70]. There is one constraint, in that the rate of diffusion must remain limited, as otherwise the system collapses owing to parasite load. This is a case of 'strong altruism', in which one replicase molecule cannot feel its own presence. Limited diffusion leads to what in evolution is called kin selection, whereby good molecules are likely to meet their own descendants, which are also likely to be good. In other words, random assortment into groups does not work for strong altruism, such that when diffusion is fast this model does not converge to a favourable transient compartmentalization model. Experimental manifestations of surface-confined replication are rare[71] and so far lack the metabolic component.

Reproducing compartments. Arguably, the most efficient model of compartmentalization favouring community coexistence is one that reproduces (autopoietic) protocells (FIG. 2c). In this model, replicators continuously

Parasites
Replicators that take help from another without paying back. The helper pays a cost in terms of fitness by maintaining its helping capacity. Saving this cost, the parasite has a replicative advantage.

'sit in the same boat', the rocking of which is not to the advantage of any species. This gives rise to what in evolution is called group selection. This was incorporated in the late 1980s into what is known as the stochastic corrector model (SCM)[72,73], which operates under the same assumptions as the (more recently developed) MCRS model except that the SCM describes a population of reproducing protocells. This is 'multilevel selection of the second type'[74], in which not only the template replicators but also the compartments reproduce, with the fission rate of the latter depending on the template composition and synergy among the templates. Realizing such a complex system experimentally remains an unmet challenge.

Replication is a stochastic process because it proceeds in terms of integer numbers of molecules rather than concentrations. Thus, during reproduction of compartments, the replicators assort themselves independently between the two daughter protocells. Higher-level natural selection acts on this stochastically generated variation between cells. This selection counteracts the malign intracellular competition among replicators by means of what biologists refer to as intragenomic conflict. If mutation rates are high, as is referred to in the Eigen paradox, metabolic coupling among replicators is better than a combination of metabolic and hypercyclic coupling. In the latter case, the mutation load is doubled and each replicator must have dual enzymatic functionalities: one for metabolic action and another for replicase function[75,76].

Catalysis features prominently in the majority of models of evolutionarily stable replicator communities. In these models, replicators are assumed to have direct catalytic capacity and can thus act without a genetic code for translation. The archetypical example is the ribozyme, but this is by no means unique in that any chemical realization of the same principle leads to the same dynamics and questions. Investigations in the context of the SCM on evolutionary dynamics of systems relying on collective catalysis for their metabolism show that there is a tendency for different types of catalytic activity to become linked. The reason for this is straightforward: unlinked replicators (not coupled to one another) exert an assortment load on protocells. In particular, there is a lowering of fitness due to chance loss of one replicator type because of internal competition and chance assortment upon fission. The assortment load can be reduced if catalytic activities are linked, as might occur through catalyst promiscuity, which is when the same catalyst enhances the rate of more than one reaction[76], as has recently been realized experimentally in a system of replicators (see below). Another way to reduce load would be to link replicators into 'chromosomes', which must somehow be handled during division. A handling mechanism might involve an accurate segregation mechanism (as exists in bacteria) or must rely on there being several copies of chromosomes, as would be the case for the original SCM mechanism for unlinked replicators[77]. The latter, more primitive scenario of chromosome formation works as part of an extended SCM model[78] in which replicators are allowed to join and break stochastically. As replicators are also catalysts, there is a 'dosage effect' — replicator numbers in the protocells do matter for metabolic efficiency. The emerging pattern is a family of chromosomes with a balanced replicator composition, where each replicator is typically present in multiple copies (a kind of 'multigene family') on each chromosome. This balanced representation is a direct consequence of the SCM favouring a balanced replicator composition.

Box 3 | **Cooperation in molecular networks**

The concepts of replication lie at the intersection of chemistry and biology, and the diversity of researchers in these fields have to tackle diverse and often misleading terminology. A notorious example is the confusion of collectively autocatalytic networks with hypercycles[121]. These two systems are indeed different, as becomes apparent on considering the stoichiometry of the following examples, which each comprise two replicator species. A collectively autocatalytic set is presented (Eqs. 9 and 10).

$$R_2 + A_1 \rightarrow A_1 + A_2 \quad (9)$$

$$R_1 + A_2 \rightarrow A_2 + A_1 \quad (10)$$

Here, A_i are informational molecules and R_i are the corresponding resources. Both A_1 and A_2 mutually catalyse each other's formation rather than replication, such that the whole set grows autocatalytically. By contrast, a two-membered hypercycle is described (Eqs. 11 and 12).

$$R_1 + A_1 + A_2 \rightarrow 2A_1 + A_2 \quad (11)$$

$$R_2 + A_2 + A_1 \rightarrow 2A_2 + A_1 \quad (12)$$

Here, each replicator catalyses the replication cycle of the other. Keeping the concentrations of R_i constant, the collective autocatalysis results in first-order growth, whereas hypercyclic organization results in second-order growth kinetics. The dynamic consequences are qualitatively different. In the biological literature, the theory of cooperation, partly resting on evolutionary game theory, is built on quadratic interaction dynamics, of which the hypercycle is but one example[147]. In this sense, a 'selfish' replicator would just catalyse its own replication (Eq. 13).

$$R_1 + 2A_1 \rightarrow 3A_1 \quad (13)$$

An example of this system would be a replicase that helps only its own replication cycle. Another ambiguity in the field is related to the analysis of collective autocatalysis. A more general version of the first system can be presented (Eqs. 14, 15 and 16).

$$R_2 + A_1 \rightarrow A_1 + A_2 \quad (14)$$

$$R_1 + A_2 \rightarrow A_2 + A_1 \quad (15)$$

$$R_3 + A_3 \rightarrow 2A_3 \quad (16)$$

In this case, A_3 denotes a direct autocatalyst that grows individually, whereas the other members can grow only collectively by means of collective autocatalysis. Here, we suggest the terms 'individualist' for A_3 and 'collectivist' for A_1 and A_2. If they were instead referred to as 'selfish' and 'cooperative'[113,148], this would mistakenly suggest that the essentially quadratic replicator equation and the theory of biological cooperation are applicable to these linearly growing systems, which is not the case. For example, a key concept in game theory is the Nash equilibrium[147], in which the whole system is in dynamic equilibrium so that no member gains by deviating from its present strategy in the context of the others' strategy being fixed. This equilibrium (in our case, as stationary replicator concentrations in a flow reactor) can be calculated and, unsurprisingly, does not coincide with the equilibria for the collectively autocatalytic system. Non-autocatalytic members of the same collectively autocatalytic system are not 'agents' because only the system as a whole is. We recommend against describing collectively autocatalytic systems in terms of 'chemical game theory'[148], which sounds rather confusing.

Ribozyme
An RNA molecule that can act as a catalyst.

Collectively autocatalytic set
A reaction network in which no member is itself autocatalytic but the members catalyse the production or formation (but not the replication) of other members of the set. The set is collectively autocatalytic if the formation of every member is catalysed by at least one other member of the set.

The final aspect of replicator community dynamics that we discuss here is sexual reproduction. This can take the form of genetic recombination between individual replicators or between sets of replicators. Here, we consider only the latter case, and only in the context of protocells undergoing fusion–fission cycles (primitive sex). Simulating an extended SCM reveals that there are conflicting forces for and against sex[79]. Namely, assortment load is alleviated because lost replicators can be regained by fusion, but the parasite load increases because such cheaters can hop from boat to boat. Overall, it is predicted that moderate sex among protocells is advantageous, especially when 'sick' protocells are more likely to fuse.

A recent population biology model[39] affirms the possibility that the replication efficiency and population structure (of which an evolving replicator catalysing the synthesis of a membrane-forming molecule would be a good example) can evolve. Unfortunately, this particular model includes some unrealistic chemical assumptions (for example, that the products of successful replication resulting from cooperative molecular associations can preferentially reassociate in bulk solution without membranes or surfaces to limit free diffusion), which prompt us to warn that real progress can only be expected if chemistry and evolution are both taken seriously at the same time.

Collectively autocatalytic sets

A collectively autocatalytic set is a system of compounds that, although not individually being directly autocatalytic, confer autocatalysis on the system as a whole. For example, in a reflexively autocatalytic and food-generated set (RAF), the formation of every member from available (simple) building blocks is catalysed by some other members of the whole set[80,81]. A hypercycle would not satisfy these criteria because its autocatalytic member cycles are linked by a cyclic loop of heterocatalytic aid responsible for second-order autocatalysis[46] (BOX 3). Small-molecule autocatalytic cycles and networks (such as the formose reaction) also do not qualify because their component chemical transformations are stoichiometric. The different steps in such cycles are analogous to different stages of the life cycle of a reproducing organism[82]. Predictions[5] and experiments[83] show that the metabolic networks of all existing cells contain at least an obligate autocatalytic core, on which present-day enzymatic catalysis is superimposed. A key open question in origin-of-life research is whether some such complete cycles (including the reverse citric acid cycle[84]) can run without enzymatic aid. A didactic form of collective autocatalysis is RNA replication, in which the plus and minus strands catalyse each other's formation, resulting in autocatalysis of the pair. If we take mutations into account, we can make a simple generalization in which building blocks A_i ($i = 1, 2, \ldots$) may catalyse, to different degrees, the incorporation of complementary B_j blocks in the other strand, and vice versa. Most nucleic acid[85] and peptide[86] versions of collective autocatalysts (known also as cross-catalytic autocatalysis) are generalizations of this template mechanism. The network based on the *Azoarcus* intron RNA[41] also requires templating by base pairing, as we describe in the section on Replicator community dynamics.

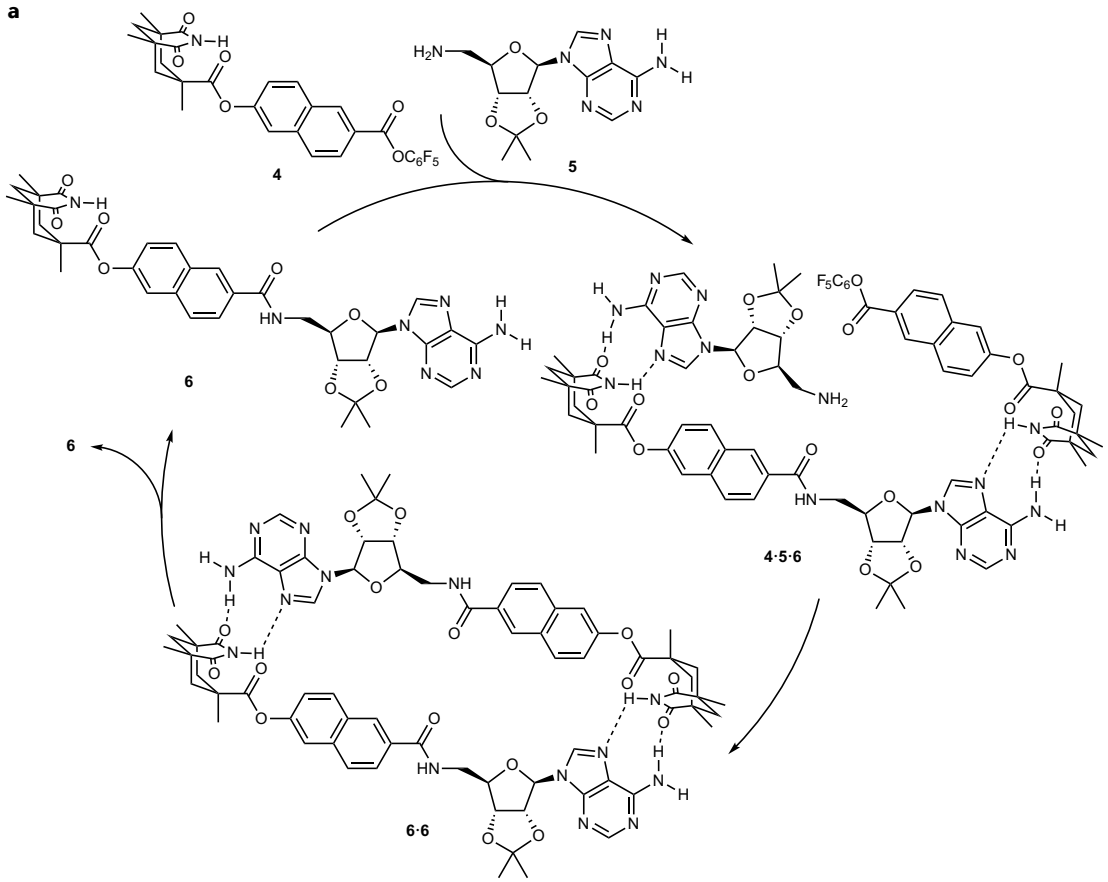
Early theoretical studies of collective autocatalytic sets[87,88] included the case in which direct templating does not play a major role in an RAF. This generalization raises theoretical and empirical questions about the plausibility of the spontaneous formation of such RAFs. Despite initial scepticism, there is good news on the theoretical front in that constraints on the probabilities of catalysis are more relaxed than were previously thought[80,89]. Moreover, compartmentalized versions of such RAFs are even expected to show some evolvability (although on a low level relative to nucleic acid template replication)[22]. Naturally, the merit of this and other proposals will be judged by the success or failure of targeted experiments. Nobody has yet observed the spontaneous formation of a generalized RAF set from a simple set of resources in a chemical experiment. Such an observation would be significant, in that it would support the 'start complex' idea of early evolution[90]. We have so far described many theoretical models of self-replication and now look to summarize experimental progress towards observing some of these phenomena.

Experimental systems of self-replicators

Molecules that can self-replicate autonomously remain relatively rare and their properties have recently been reviewed[8–14]. We consider here a selection of the most important self-replicators, the variety of which is encouraging as this shows that the replication mechanisms we have described (FIG. 1b,c) can be implemented with very different chemistries and with molecular recognition motifs that need not resemble those commonly found in biochemistry. FIGURE 3 shows two examples of completely synthetic replicators that operate by the duplex formation mechanism of FIG. 1b, featuring different hydrogen-bonding recognition motifs and different coupling chemistries, based on amide bond formation (FIG. 3a) and 1,3-dipolar cycloaddition (FIG. 3b). FIGURE 4 shows that self-replication through the duplex mechanism can also be implemented with DNA (FIG. 4a), RNA (FIG. 4b,c) and α-helical peptides (FIG. 4d). Peptide-based molecules can also self-replicate through assembly into β-sheets, giving rise to supramolecular polymers as shown in FIG. 5a,b. A similar mechanism of assembly-driven self-replication can also occur with fully synthetic molecules that form 2D sheets through π-stacking interactions (FIG. 5c). The remainder of this Review focuses on systems that are relevant to the hurdles that need to be overcome to proceed from self-replication to de novo life. We first summarize progress towards achieving Darwinian evolution, then continue with work on replicator community dynamics and finish with efforts directed at integrating self-replication with metabolism and compartmentalization.

Towards Darwinian evolution in systems of self-replicators

The prospect of having synthetic self-replicators undergo Darwinian evolution is becoming increasingly realistic. Numerous challenges associated with this goal have


a
4
5
6
6
4·5·6
6·6

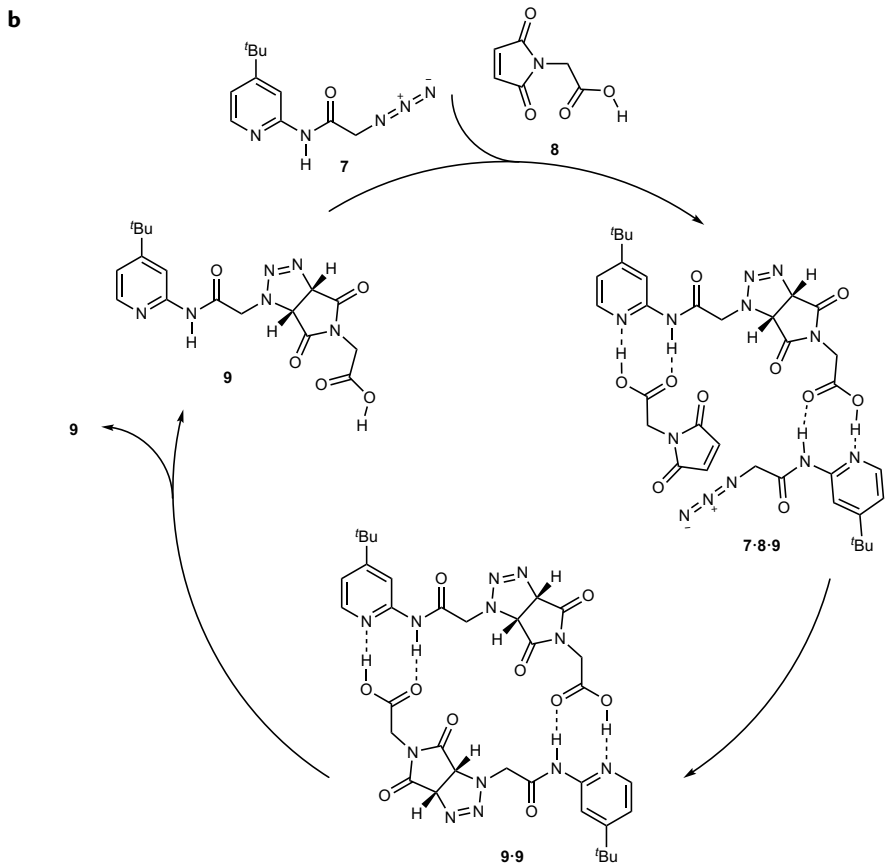

b
7
8
9
9
7·8·9
9·9

Fig. 3 | **Completely synthetic self-replicating molecules. a** | A self-replicating system in which ester **4** and amine **5** react to form self-complementary template **6**. Through building-block pre-organization in the complex [**4**·**5**·**6**], the template enables self-replication[128]. **b** | A similar mechanism is at play in a cycloaddition-based system, where azide **7** reacts with maleimide **8** to form self-complementary template **9** (REF.[129]).

already been met, the first of which is the propensity for self-replicators to self-inhibit and thereby grow only parabolically. Darwinian evolution is most readily performed with exponential replicators (BOX 1), and even though most self-replicators are parabolic, several examples of and/or protocols for exponential replication now exist. Transitioning from parabolic to exponential replication is possible if the replication reaction has a transition state whose geometry is sufficiently different to that of the product, such that the product duplex (but not the transition state) is strained and can readily dissociate. This intuitive strategy has led to close-to-exponential replication of α-helical peptides[91–93] and synthetic replicators based on cycloaddition reactions[94]. Exponential growth in a system of RNA replicators has been realized through an approach resembling directed evolution[95]. However, how this system avoids stalling in the replicator duplex state remains unclear. There also exists a protocol through which replicators that are parabolic in solution can become exponential when anchored to a surface[71]. In such a case, replication is followed by thermally induced dissociation of the replicator duplex, whereafter the monomeric replicators bind vacant sites on the surface, thus causing replicators to spread exponentially over successive heating–replication cycles. This study represents a rare example of surface-confined replicators, which are relevant to the MCRS model described above. Another promising strategy for achieving exponential replication involves supramolecular polymerization (FIG. 1c) of replicator stacks that can be broken mechanically, thereby liberating new growth sites. This strategy is likely to be generally applicable because self-assembly is a general phenomenon. Indeed, although the polymerization was first implemented with β-sheet-based replicators[54,55], it has since been realized with other building blocks[96–98].

For replication to be exponential it must be at least first order in the replicator. If replication is of a higher order, then it can lead to phenomena such as bistability. For example, hyperbolic growth (second-order autocatalysis) results in survival of the common rather than the fittest in competition[99]. Kinetic modelling in the right parameter window has been used to show that second-order autocatalysis allows for the onset of bistability and bifurcation[100]. Such bistability has been realized in a synthetic system based on α-helical peptides that replicate by forming three-helix bundles[101].

After overcoming self-inhibition, the second challenge with achieving Darwinian evolution involves incorporating mutation into the replication process. Despite their importance in biology, mutation and replication fidelity have received surprisingly little attention in synthetic systems. Notably, there exists a system of parabolic α-helix-based replicators featuring mutants that are infertile but are able to cross-catalyse the formation of the parent replicator through a type of error correction[102]. A similar error correction mechanism has been observed in a network of peptide isomers[103], which differs from the former system because the mutants here form spontaneously during the reaction. In both systems, the mutants are more efficient as cross-catalysts of native peptide sequences than as autocatalysts.

The spontaneous diversification of replicator sets has been observed for dithiols decorated with peptide chains[104] (FIG. 5b). Starting from a mixture of building blocks **27b** and **27c** led to the emergence of a set of replicators rich in **27b** (technically dehydro-**27b**). After several days, this first set then promoted, through cross-catalysis, the formation of a second set, rich in the remaining building block **27c**. This diversification of replicators, each utilizing different resources ('food' sources **27b** and **27c**), bears a crude resemblance to the formation of bacterial species as observed in biology.

Once exponential replication and mutation have been incorporated into synthetic self-replicators, the next step towards Darwinian evolution involves implementing a replication–destruction regime. Such a regime has, to a limited extent, been implemented through serial dilution, whereby a small fraction of replicator solution is transferred to a fresh solution of resources (food). Transfer of the replicator to a solution with excess food is repeated after the replicator consumes most of the original food[40,41,105]. In this protocol, the replicators that are not transferred are effectively dead because they no longer generate offspring. One limitation of this protocol is that replicator destruction is typically not selective, such that evolution in such systems primarily selects for replication speed rather than resistance to destruction. A disconcerting consequence of selection for replication speed is apparent from experiments on the enzyme-mediated replication of a RNA sequence conducted by Spiegelman and colleagues[106]. Starting from a long RNA oligomer and conducting several rounds of serial transfer affords an RNA sequence that is dramatically shortened, a result of shorter sequences tending to replicate more quickly than longer ones. This tendency to spontaneously decrease replicator complexity has become known as the 'Spiegelman monster'. This monster can be overcome by conducting a variation of this experiment in a flow system with a thermal gradient, such that thermophoresis leads to the selective retention of long RNA sequences, preventing their destruction by outflow[107,108].

With many challenges having been met, the prospect of evolving systems of replicators by Darwinian evolution is imminent. Biology gives us innumerable examples of how Darwinian evolution is a great engine of invention. Inventions can also manifest themselves in synthetic systems even before evolution. The selection of self-replicators at the stage of their emergence from molecular networks can inadvertently be accompanied by the emergence of functions (for example, catalysis) beyond mere replication. Furthermore, replicating systems comprising multiple fragments have been described, increasing opportunities for evolution[41,104,109], but the fidelity of replication in such systems is to some extent problematic. Here, we would need to solve Eigen's paradox, most likely by developing specific replicator community dynamics.

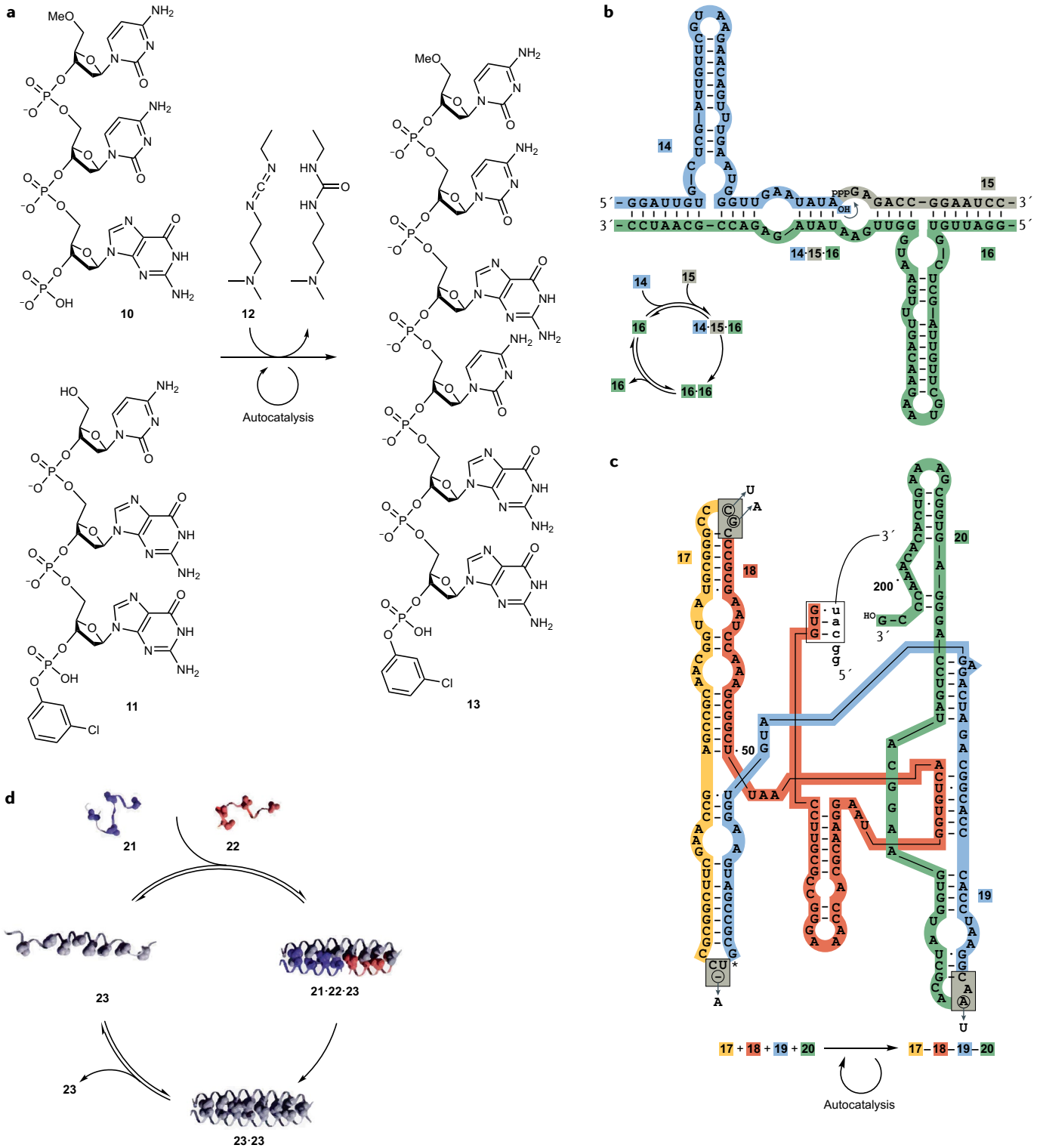


Fig. 4 | **Self-replicating molecules featuring nucleic acids or peptides.** **a** | The first example of a non-enzymatic template-directed self-replicator was hexanucleotide **13**. Trinucleotides **10** and **11** bind **13**, and activation of the phosphate of **10** with coupling agent **12** makes it susceptible to nucleophilic attack from the OH group of **11**. This affords a phosphodiester group and thus another copy of **13** (REF.[130]). **b** | Secondary structure of the modified R3C self-replicating ligase ribozyme. The self-complementary template molecule **16** can bind two fragments **14** and **15** to form the complex **14·15·16**. In this complex, a phosphodiester bond can be formed between **14** and **15**, liberating a pyrophosphate, resulting in another copy of **16** (REF.[131]). **c** | Secondary structure of the modified *Azoarcus* ribozyme, partitioned into four RNA fragments: **17**, **18**, **19** and **20**. The four fragments can non-covalently self-assemble to form the complex **17·18·19·20** that can catalyse recombination reactions to form the covalently linked ribozyme **17–18–19–20**. The covalently linked ribozyme is also catalytically active in the recombination reactions[132,133]. **d** | The thiobenzyl ester of 17-residue peptide fragment **21** and 15-residue peptide fragment **22** form a 32-residue α-helical peptide **23** through amide formation by native chemical ligation. Peptide **23** templates its own formation through interhelical hydrophobic interactions, resulting in parabolic replication[134]. Part **b** adapted with permission from REF.[131], Copyright (2002) National Academy of Sciences, USA. Part **c** adapted with permission from REF.[132], Elsevier. Part **d** adapted from REF.[134], Springer Nature Limited.

Parabolic replication
Replication that is slower than exponential because the per-capita growth rate decreases with increasing replicator concentration.

Collectivism
The propensity of an informational molecule to join a collectively autocatalytic set rather than replicating itself directly.

Dynamic combinatorial library
A set of continuously interconverting oligomeric molecules made by linking building blocks together through a reversible reaction.

Replicator community dynamics

As experimental self-replicators become more sophisticated one must address the Eigen paradox. Although there is at present no experimental work on this, theoretical models suggest that the answers will almost invariably rely on communities of coexisting replicators, as we described above. Recently, experimental work on the dynamics exhibited by replicator communities has started and, as in ecology, different interactions have been observed[110]. We now discuss examples of collectivism, competition and parasitism/predation. The terminology used here comes from replicator and game theory and in some instances will differ from the terminology used in the original publications (BOX 3). In this way, we attempt to unify the language such that experimentalists and theoreticians from these different fields can understand each other.

Collectivism has been observed in a system of RNA replicators based on group I self-splicing introns from the ciliate *Azoarcus* species. This system can autocatalytically assemble from its four[41] or even five[109] pieces. This core self-replication reaction can be incorporated into a large collectively autocatalytic set[111] by combinatorially altering certain complementary triplets (the internal guide sequence of the catalyst and the 3′ end-tag sequence of one of the substrate molecules; FIG. 4c). Remarkably, the *Azoarcus* intron system shows anabolic and catabolic capabilities because it can generate intermediates that become progressively larger and it can transform (by recombination) some reactants that cannot directly be incorporated into the growing complex into resources that can directly sustain the autocatalytic system[112]. Theoretical analysis of experimental data suggests that the combinatorics of internal guide sequence-tag triplets and their binding strengths together tip the interaction topology balance from individualism to collectivism[113]. Notably, a uniformly random distribution of binding strengths (and the resulting rate constants) would markedly lower the degree of collectivism. A series of serial dilution experiments have been conducted to assess how adding a new node to an existing three-membered network affects the replication rate[114]. When purely individualistic and fully collectivist networks are chemically balanced, such that the set and quantity of resources remain constant, they grow at equal rates. Increasing the number of Watson–Crick pairings by adding a fourth member to the core tends to increase both the total and core replication rates, except when the fourth member is merely a receiver that drains resources from the core. This effect is strongest when the newly added member is bidirectionally linked to the core and is also directly autocatalytic. These experimental findings resonate with the theoretical prediction that the most successful extensions to autocatalytic cores should be molecular 'vitalists' that are autocatalysts that also heterocatalytically aid the core itself[115].

The competition between replicators for common building blocks has been investigated in a fully synthetic set of small molecules that replicate through dimer formation[116] (FIG. 1b). In particular, the reaction network of two competing replicators was coupled to a dynamic combinatorial library[116] containing the resources for replication. In batch experiments, where the systems were allowed to approach equilibrium, the ability of replicators to process the dynamic combinatorial library to their own advantage is limited. This results in continuous coexistence of both replicators, as can be assessed by seeding the system with a specific replicator at the start of the experiments. Instead, a reaction–diffusion protocol can potentially lead to a more selective outcome of competition between replicators[117]. Further work has probed replicator community dynamics by analysing system-level responses[118] using four autocatalytic and partially cross-catalytic replicators. Here, the output of the replicator network is determined by the nature of the instructing template, which we call the replicator seed. The network topology is such that the input of a single template resulted in the system-level upregulation of two interlinked replicators. Being able to deliberately tune replicator community dynamics is likely to be important in overcoming the Eigen paradox.

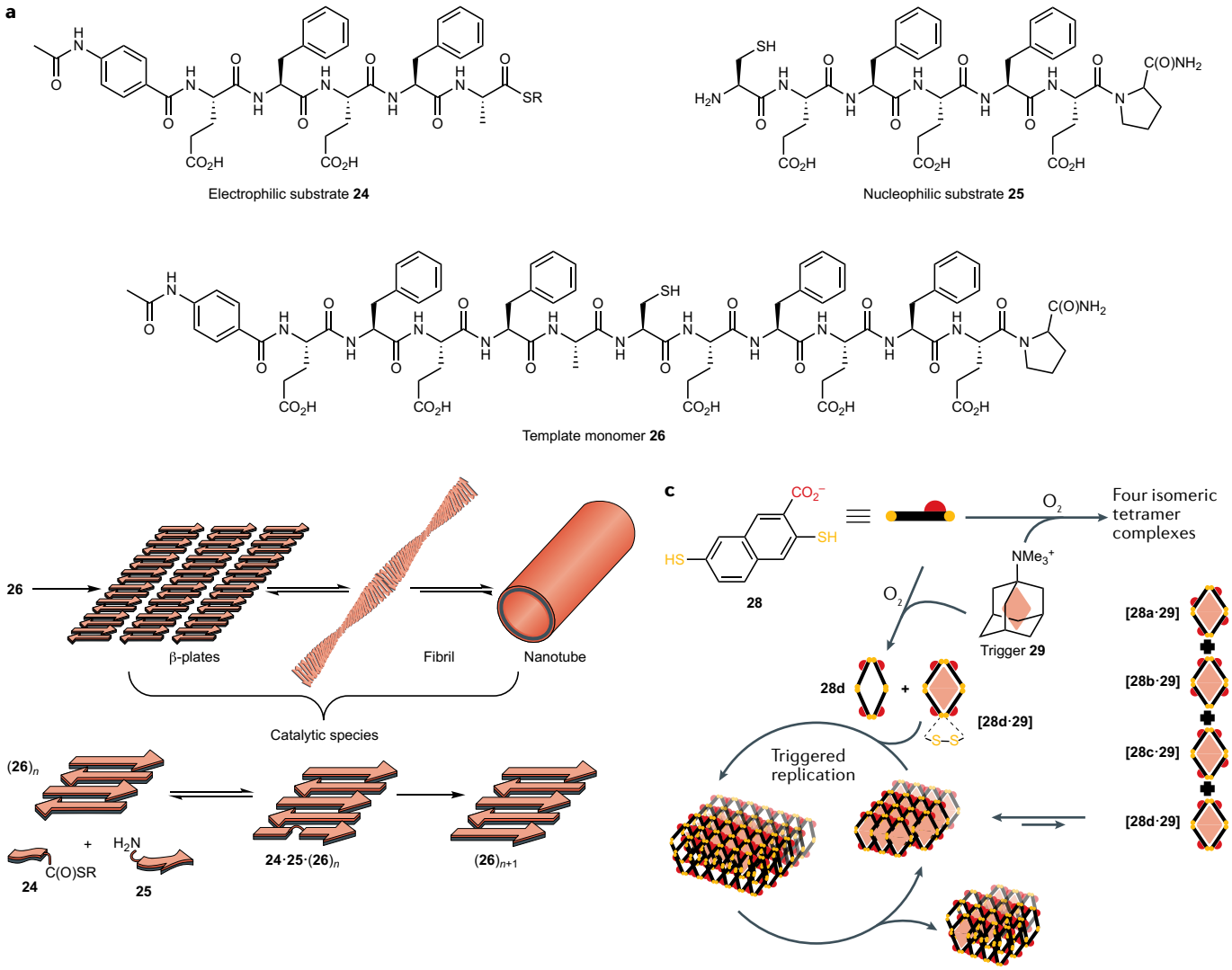
Parasitic and predatory behaviour, which are familiar in biology, have recently been observed in autonomously self-replicating molecules. Parasitic replicators were found to form exclusively through cross-catalysis by a parent replicator and, under certain conditions, subsequently consumed their parent[119]. Experimental efforts have also been directed towards constructing hypercycles of self-replicators[41,120]. However, in both cases these efforts have tripped up because of a misunderstanding of the hypercycle[121], which relies on catalysis of self-replication (BOX 2), not of replicator formation as was reported in the two experimental systems. Experimental hypercycles of self-replicators have yet to be reported, which is not discouraging because, in any case, they have a limited ability to solve Eigen's paradox.

Integrating self-replication with metabolism

For systems of self-replicators to acquire metabolism they first need to catalyse reactions. The first example of a self-replicator that catalyses a reaction other that its own formation features an imidazolidinone moiety to enable organocatalytic hydride reduction and Friedel–Crafts alkylation[122] (FIG. 6a). However, performing this exogenous catalysis does not aid the self-replicator because it does not afford it any additional resources. Furthermore, the solvent conditions for catalysis were incompatible with those for replication.

Several examples of self-replicators have recently been reported to catalyse reactions under conditions compatible with replication, including reactions that directly benefit replication. For example, it is possible to modify the resources for a replicator so they can only be used after being liberated by a reaction catalysed by the replicator itself. This has been demonstrated using self-replicating RNAs based on the *Azoarcus* ribozyme[112] (FIG. 6b). In this system, both the liberation and the self-replication reactions involve phosphodiester chemistry.

We have recently found that our peptide-appended dithiol replicators not only catalyse their own replication but also retro-aldol and carbamate hydrolysis (Fmoc-deprotection) reactions[61] (FIG. 6c). Catalysis is an emergent property of the replicators, the resources/precursors for


a
Electrophilic substrate 24
Nucleophilic substrate 25
Template monomer 26
26
β-plates
Fibril
Nanotube
Catalytic species
(26)n
24
25
24·25·(26)n
(26)n+1
c
28
Four isomeric tetramer complexes
Trigger 29
28d
[28d·29]
Triggered replication
[28a·29]
[28b·29]
[28c·29]
[28d·29]

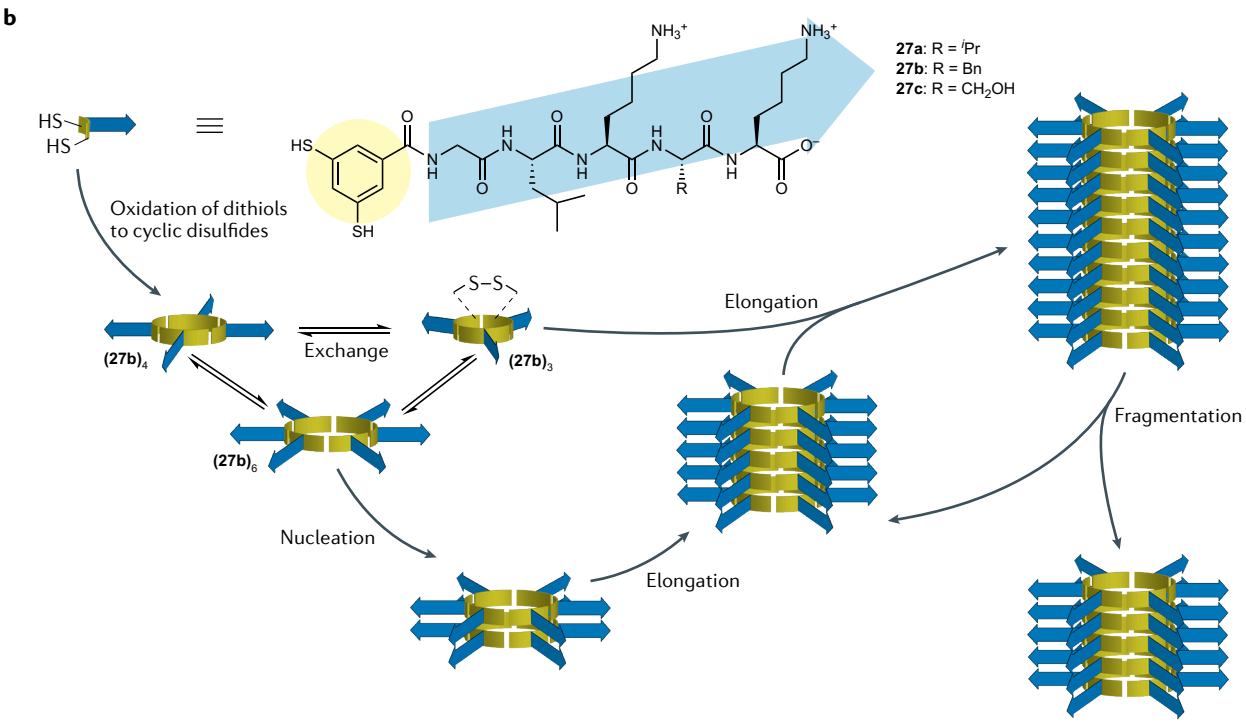

b
27a: R = iPr
27b: R = Bn
27c: R = CH2OH
Oxidation of dithiols to cyclic disulfides
(27b)4
Exchange
(27b)3
(27b)6
Elongation
Nucleation
Fragmentation

Fig. 5 | **Self-replication driven by supramolecular polymerization. a** | Peptide fragments **24** and **25**, with alternating hydrophobic and hydrophilic amino acids, react through native chemical ligation to produce peptide **26**. Transient β-sheet nanostructures $(\mathbf{26})_n$, including β-plates and fibrils, catalyse the ligation, generating the new catalytic end $(\mathbf{26})_{n+1}$. Replication progresses exponentially as more aggregates and catalytic ends are produced[135]. **b** | Dithiol building block **27**, upon oxidation, produces a mixture of disulfide macrocycles that interconvert through disulfide exchange. Nucleation of macrocycle $(\mathbf{27})_6$ followed by elongation at the fibre ends shifts the composition of the dynamic mixture towards formation of more $(\mathbf{27})_6$. Fragmentation of fibres by mechanical forces generates more growing fibre ends, enabling exponential replication[136,137]. **c** | Dithiol **28** oxidizes in the presence of the template **29** to form four isomeric tetramer complexes [**28a–d**·**29**]. Complex [**28d**·**29**] replicates owing to formation of a 2D sheet once its concentration surpasses the critical aggregation concentration. Complex [**28d**·**29**] then acts as a seed for replication of uncomplexed **28d**[96]. Part **a** adapted with permission from REF.[135], American Chemical Society. Part **c** adapted with permission from REF.[96], American Chemical Society.

which are virtually inactive for the exogenous reactions. Importantly, this is a chance invention, with selection for replication inadvertently orienting residues into catalytically active geometries. Importantly, Fmoc deprotection liberates dibenzofulvene (**42**), which speeds up the formation of precursors of the replicator by enhancing the oxidation of the dithiol starting material **27b** into the small macrocycles $(\mathbf{27b})_3$ and $(\mathbf{27b})_4$ from which the replicators grow (FIG. 6d). The same replicators can also bind and activate different photocatalytic cofactors[60] (FIG. 6e). Thus, photo-irradiation accelerates the oxidation of dithiol building blocks into disulfide replicator precursors, thereby aiding self-replication. Interestingly, the same self-replicator is able to catalyse all of these very different reactions. Although this catalytic promiscuity emerged naturally, even before any evolutionary processes, as we noted above, theoretical studies suggest that catalytic promiscuity is also evolutionarily advantageous.

The systems described here (FIG. 6b,d,e) are the first examples of replicators exhibiting a proto-metabolism, in the sense that they catalyse the formation of their own precursors from molecules in their environment. Even though some of them harvest light energy, these systems do not (yet) use energy to drive endergonic reactions, which, besides building precursors, is the second important aspect of metabolism.

Integrating self-replication with compartmentalization

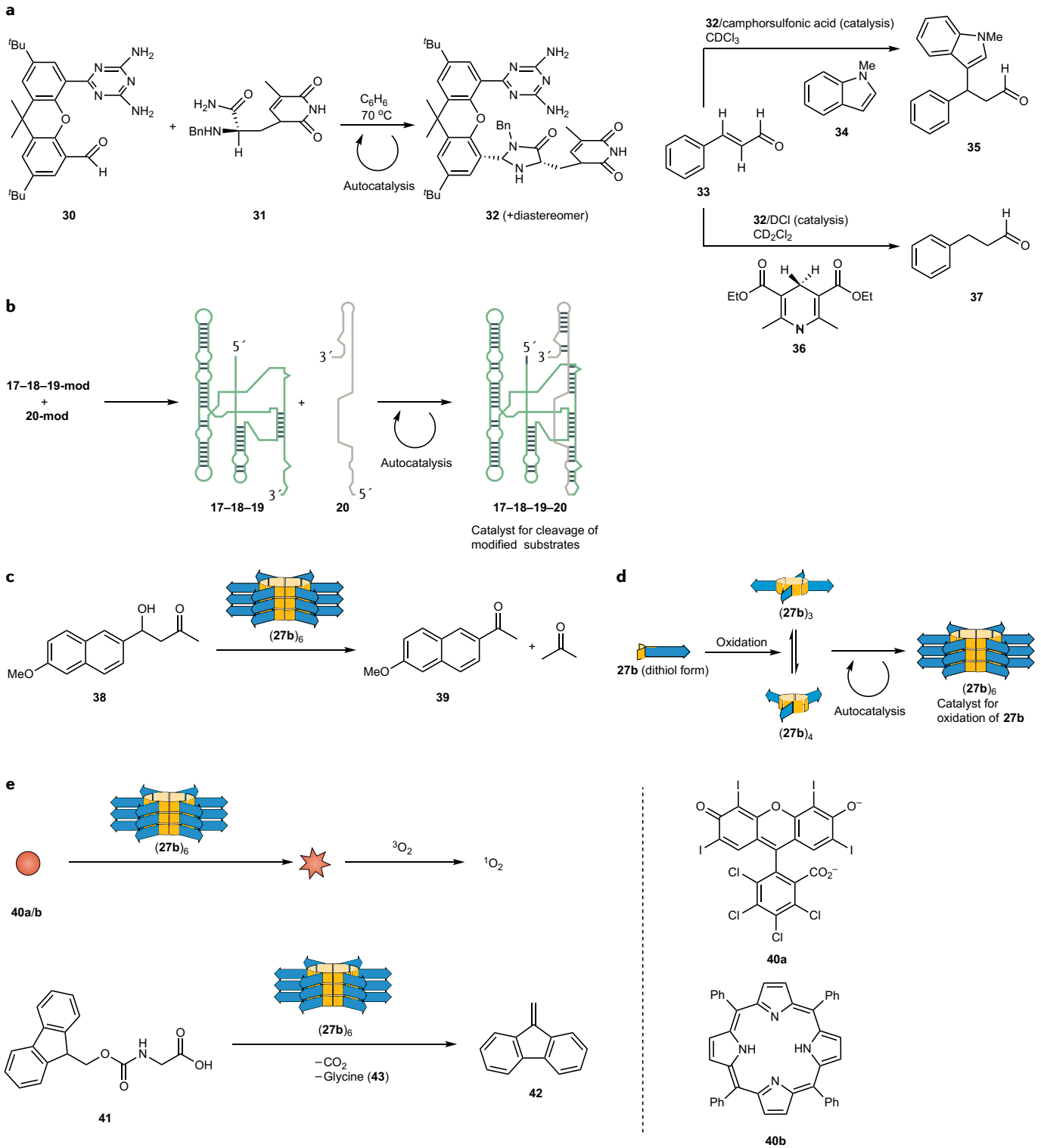
Despite the promising evolutionary features revealed by theoretical work on the reproducing compartment model (see above), experimental systems in which replication is coupled to compartment growth and division have yet to be realized. However, some interesting efforts in this direction have been described, mostly relying on non-autonomous replicator systems that use extant enzyme machinery to replicate oligonucleotides. A system has been developed in which enzyme-mediated replication is coupled to membrane growth and division through electrostatic interactions between anionic DNA that is being replicated at the cationic vesicle membrane[62]. Similarly, template-directed RNA polymerization has been performed inside coascervate compartments[19]. RNA replication can also be mediated by the Qβ replicase, which tends to lead to the emergence of short RNAs that act as parasites — the Spiegelman monster. Performing replication in a water–oil emulsion represses the takeover by parasites while maintaining the Qβ replicase activity[123]. Parasites can, in principle, still form, but this affects only a limited number of droplets from which they cannot escape. When performing serial transfer experiments in microdroplet format, replicating RNAs can survive and exhibit host–parasite oscillation dynamics, even when challenged with overall parasite concentrations that would cause replicators to die out in bulk solution[124]. This study also showed that a host and parasite can co-evolve during these experiments.

A notable recent collaborative study has afforded a transient compartmentalization system featuring a ribozyme that can cleave a polynucleotide substrate[125]. Different forms of a parasite repeatedly arose in the experiments. Both the ribozyme and parasite were replicated by the Qβ replicase, which was added together with activated nucleotides. Experiments were performed in the bulk and in microdroplets, with the latter repeatedly being broken and their contents mixed within a common pool by external manipulation (each iteration of this step can informally be called a 'generation'). In the bulk, the parasite took over the system, analogous to the classic Spiegelman experiments[106]. When introduced into microdroplets in a regime in which the fate of droplets did not depend on the activity of the ribozyme, the extinction of the ribozyme was merely slowed down. By contrast, the case of droplet sorting according to metabolite concentration produced by the ribozyme allowed it to coexist with the parasite. In a different experiment, the Qβ replicase was allowed to evolve using an externally provided in vitro translation system and the necessary resources. The experiments were performed by small-volume serial transfer[124] and in a droplet-containing automated flow reactor[126]. The bulk experiment naturally resulted in the extinction of the replicator. Serial transfer involved repeated droplet formation and mixing and gave rise to oscillations in the replicator and parasite populations. The basic growth rate of the ribozyme was much lower than that of the parasite, which was an order of magnitude shorter (we have noted above the inherently faster replication of shorter sequences). In this system, the replicase evolved such that it selectively replicated the ribozyme rather than the parasite.

Conclusions and outlook

Three decades of research on self-replicating systems have provided us with an increasingly clear path towards de novo life. Although there are many unsolved problems, these are starting to take the form of well-defined and addressable research questions. Answering these questions will likely be easiest by integrating theory and experiment, an approach that is, unfortunately, still relatively rare. We now outline the most pressing current challenges and how these might be overcome.

Implementing Darwinian evolution of self-replicating systems under conditions where dynamic kinetic stability governs replicator distributions is becoming within reach. Still lacking are ways to subject self-replicators



to selective death. Indeed, most present experimental models featuring death rely on serial dilution or physical removal, processes that are indiscriminate. Another challenge to experimentally realizing Darwinian evolution is ensuring a sufficiently large state space for a system to explore and evolve into, while also having a sufficiently high replication fidelity to allow the system to maintain its identity in the face of the many possibilities for mutation. Probably the greatest challenge is to manage state space and experimental conditions such that evolution becomes open-ended and the system repeatedly invents new functions. The search for open-ended evolution in a synthetic system is one of the few problems for which theory is unlikely to provide much guidance. Indeed, how does one allow a simulation to make inventions? Nevertheless, we know that the laws of chemistry and physics facilitate open-ended evolution and it is encouraging that the first observations of

Fig. 6 | **Emergent catalysis in self-replicating systems. a** | Self-replicator **32** can catalyse its own formation by templating the condensation of **30** with **31**. On replication, an imidazolidinone functionality is formed, which acts as an organocatalyst for hydride reduction (**33** + **34** → **35**) and Friedel–Crafts alkylation reactions (**33** + **36** → **37**)[122]. **b** | The *Azoarcus* ribozyme (**17–18–19–20**) can catalyse its own formation from **17–18–19** and **20** through phosphodiester transesterification. The replicator can also liberate more **17–18–19** and **20** from modified substrates **17–18–19-mod** and **20-mod**, respectively, by cleaving off a short RNA fragment[112]. **c** | The autocatalytic hexameric macrocycle $(\mathbf{27b})_6$ can catalyse the retro-aldol reaction that converts **38** into aldehyde **39** and Me_2CO. The catalysis is only observed when $(\mathbf{27b})_6$ is stacked into fibres and not for the smaller macrocycles. The reaction is performed at the same time as the replicator forms[61]. **d** | Double-positive feedback system in which self-replicator $(\mathbf{27b})_6$ not only acts as an autocatalyst but also promotes the formation of the precursor molecules that it needs to replicate. **e** | This mechanism can be realized with photocatalytic cofactor **40**, which bind to fibres of $(\mathbf{27b})_6$ and becomes activated, converting $^3O_2$ into $^1O_2$ to accelerate thiol oxidation. This principle has been demonstrated using two different photoactive cofactors: Rose Bengal **40a** and tetraphenylporphyrin **40b**[60]. Catalysed oxidation also occurs upon cleavage of **41** by the fibres of $(\mathbf{27b})_6$. After H^+ abstraction, **41** is converted into **42**, **43** and CO_2. Product **42** speeds up the oxidation of **27b** into $(\mathbf{27b})_3/(\mathbf{27b})_4$ (REF.[61]). Part **b** adapted with permission from REF.[112], Oxford University Press.

chance inventions made by self-replicators have recently been made[61].

Theory recommends that experimentalists consider communities of coexisting replicators. Although research is now moving in this direction, developing an experimental system that solves Eigen's paradox remains a huge challenge. We first need to learn how to engineer specific dynamics into replicator communities that might well rely on different types of molecular structure. For example, the structures of present-day ribosomes suggest a role for co-evolution of peptides and nucleic acids. In this vein, studies on systems of replicators that contain the structural elements of these compound classes would likely be useful[127]. It is intriguing to note that different theoretical models converge on solutions to Eigen's paradox that feature replicators with catalytic capabilities in combination with spatial confinement. Thus, from a bottom-up analysis of pushing Darwinian evolution of self-replicators in the direction of increasing complexity, features emerge that are central to the other two key characteristics of life: metabolism (heavily dependent on catalysis) and compartmentalization. Hence, the confluence of replication, metabolism and compartmentalization seems an innate process in Darwinian evolution. Life, based on these three pillars, appears to be the logical outcome of a process of evolution that starts from mere self-replication.

Published online 1 July 2020

1. Ruiz-Mirazo, K., Briones, C. & de la Escosura, A. Prebiotic systems chemistry: new perspectives for the origins of life. *Chem. Rev.* **114**, 285–366 (2014).
2. Ludlow, R. F. & Otto, S. Systems chemistry. *Chem. Soc. Rev.* **37**, 101–108 (2008).
3. Whitesides, G. M. & Ismagilov, R. F. Complexity in chemistry. *Science* **284**, 89–92 (1999).
4. Ashkenasy, G., Hermans, T. M., Otto, S. & Taylor, A. F. Systems chemistry. *Chem. Soc. Rev.* **46**, 2543–2554 (2017).
5. Gànti, T. *The Principles of Life* (Oxford Univ. Press, 2003).
6. Popa, R. *Between Necessity and Probability: Searching for the Definition and Origin of Life* (Springer, 2004).
7. Chodasewicz, K. Evolution, reproduction and definition of life. *Theory Biosci.* **133**, 39–45 (2014).
8. Kosikova, T. & Philp, D. Exploring the emergence of complexity using synthetic replicators. *Chem. Soc. Rev.* **46**, 7274–7305 (2017).
9. Duim, H. & Otto, S. Towards open-ended evolution in self-replicating molecular systems. *Beilstein J. Org. Chem.* **13**, 1189–1203 (2017).
10. Bissette, A. J. & Fletcher, S. P. Mechanisms of autocatalysis. *Angew. Chem. Int. Ed.* **52**, 12800–12826 (2013).
11. Patzke, V. & von Kiedrowski, G. Self replicating systems. *Arkivoc* **5**, 293–310 (2007).
12. Le Vay, K., Weise, L. I., Libicher, K., Mascarenhas, J. & Mutschler, H. Templated self-replication in biomimetic systems. *Adv. Biosyst.* **3**, 1800313 (2019).
13. Clixby, G. & Twyman, L. Self-replicating systems. *Org. Biomol. Chem.* **14**, 4170–4184 (2016).
14. Robertson, A., Sinclair, A. J. & Philp, D. Minimal self-replicating systems. *Chem. Soc. Rev.* **29**, 141–152 (2000).
15. Monnard, P.-A. & Walde, P. Current ideas about prebiological compartmentalization. *Life* **5**, 1239–1263 (2015).
16. Beales, P. A., Ciani, B. & Mann, S. The artificial cell: biology-inspired compartmentalization of chemical function. *Interface Focus.* **8**, 20180046 (2018).
17. Hanczyc, M. M., Fujikawa, S. M. & Szostak, J. W. Experimental models of primitive cellular compartments: encapsulation, growth, and division. *Science* **302**, 618–622 (2003).
18. Bissette, A. J. & Fletcher, S. P. Novel applications of physical autocatalysis. *Orig. Life Evol. Biosph.* **45**, 21–30 (2015).
19. Poudyal, R. R. et al. Template-directed RNA polymerization and enhanced ribozyme catalysis inside membraneless compartments formed by coacervates. *Nat. Commun.* **10**, 490 (2019).
20. Stewart, J. E. The origins of life: the managed-metabolism hypothesis. *Found. Sci.* **24**, 171–195 (2018).
21. Nghe, P. et al. Prebiotic network evolution: six key parameters. *Mol. Biosyst.* **11**, 3206–3217 (2015).
22. Vasas, V., Fernando, C., Santos, M., Kauffman, S. & Szathmáry, E. Evolution before genes. *Biol. Direct.* **7**, 1 (2012).
23. Ralser, M. An appeal to magic? The discovery of a non-enzymatic metabolism and its role in the origins of life. *Biochem. J.* **475**, 2577–2592 (2018).
24. Muchowska, K. B., Varma, S. J. & Moran, J. Synthesis and breakdown of universal metabolic precursors promoted by iron. *Nature* **569**, 104–107 (2019).
25. Engelhart, A. E., Adamala, K. P. & Szostak, J. W. A simple physical mechanism enables homeostasis in primitive cells. *Nat. Chem.* **8**, 448–453 (2016).
26. Pross, A. *What is Life?* (Oxford Univ. Press, 2016).
27. Pascal, R., Pross, A. & Sutherland, J. D. Towards an evolutionary theory of the origin of life based on kinetics and thermodynamics. *Open. Biol.* **3**, 130156 (2013).
28. Pross, A. Stability in chemistry and biology: life as a kinetic state of matter. *Pure Appl. Chem.* **77**, 1905–1921 (2005).
29. Pross, A. & Pascal, R. A roadmap toward synthetic protolife. *Synlett* **28**, 30–35 (2016).
30. Pross, A. & Pascal, R. How and why kinetics, thermodynamics, and chemistry induce the logic of biological evolution. *Beilstein J. Org. Chem.* **13**, 665–674 (2017).
31. Maynard Smith, J. *The Problems of Biology* (Oxford Univ. Press, 1986).
32. Bissette, A. J., Odell, B. & Fletcher, S. P. Physical autocatalysis driven by a bond-forming thiol–ene reaction. *Nat. Commun.* **5**, 4607 (2014).
33. Colomer, I., Morrow, S. M. & Fletcher, S. P. A transient self-assembling self-replicator. *Nat. Commun.* **9**, 2239 (2018).
34. Walde, P., Wick, R., Fresta, M., Mangone, A. & Luisi, P. L. Autopoietic self-reproduction of fatty acid vesicles. *J. Am. Chem. Soc.* **116**, 11649–11654 (1994).
35. Hardy, M. D. et al. Self-reproducing catalyst drives repeated phospholipid synthesis and membrane growth. *Proc. Natl Acad. Sci. USA* **112**, 8187–8192 (2015).
36. Breslow, R. On the mechanism of the formose reaction. *Tetrahedron Lett.* **1**, 22–26 (1959).
37. Orgel, L. E. Molecular replication. *Nature* **358**, 203–209 (1992).
38. Maynard Smith, J. & Szathmáry, E. *The Major Transitions in Evolution* (Oxford Univ. Press, 1995).
39. Levin, S. R., Gandon, S. & West, S. A. The social coevolution hypothesis for the origin of enzymatic cooperation. *Nat. Ecol. Evol.* **4**, 132–137 (2020).
40. Hordijk, W., Vaidya, N. & Lehman, N. Serial transfer can aid the evolution of autocatalytic sets. *J. Syst. Chem.* **5**, 4 (2014).
41. Vaidya, N. et al. Spontaneous network formation among cooperative RNA replicators. *Nature* **491**, 72–77 (2012).
42. Lifson, S. & Lifson, H. Coexistence and Darwinian selection among replicators: response to the preceding paper by Scheuring and Szathmary. *J. Theor. Biol.* **212**, 107–109 (2001).
43. Blokhuis, A., Lacoste, D. & Gaspard, P. Reaction kinetics in open reactors and serial transfers between closed reactors. *J. Chem. Phys.* **148**, 144902 (2018).
44. Eigen, M. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465–523 (1971).
45. Eigen, M. & Winkler-Oswatitsch, R. *Steps Towards Life* (Oxford Univ. Press, 1992).
46. Szathmáry, E. & Gladkih, I. Sub-exponential growth and coexistence of non-enzymatically replicating templates. *J. Theor. Biol.* **138**, 55–58 (1989).
47. von Kiedrowski, G. Minimal replicator theory I.: parabolic versus exponential growth. *Bioorg. Chem. Front.* **3**, 113–146 (1993).
48. Luisi, P. L. *The Emergence of Life: From Chemical Origins to Synthetic Biology* (Cambridge Univ. Press, 2006).
49. Károlyi, G., Scheuring, I. & Czárán, T. Metabolic network dynamics in open chaotic flow. *Chaos* **12**, 460–469 (2002).
50. Károlyi, G., Péntek, Á., Scheuring, I., Tél, T. & Toroczkai, Z. Chaotic flow: the physics of species coexistence. *Proc. Natl Acad. Sci. USA* **97**, 13661–13665 (2000).
51. Könnyű, B. & Czárán, T. Spatial aspects of prebiotic replicator coexistence and community stability in a surface-bound RNA world model. *BMC Evol. Biol.* **13**, 204 (2013).
52. Wilson, D. S. A theory of group selection. *Proc. Natl Acad. Sci. USA* **72**, 143–146 (1975).
53. Agerschou, E. D., Mast, C. B. & Braun, D. Emergence of life from trapped nucleotides? Non-equilibrium behavior of oligonucleotides in thermal gradients. *Synlett* **28**, 56–63 (2017).
54. Rubinov, B., Wagner, N., Rapaport, H. & Ashkenasy, G. Self-replicating amphiphilic β-sheet peptides. *Angew. Chem. Int. Ed.* **48**, 6683–6686 (2009).

55. Colomb-Delsuc, M., Mattia, E., Sadownik, J. W. & Otto, S. Exponential self-replication enabled through a fibre elongation/breakage mechanism. *Nat. Commun.* **6**, 7427 (2015).
56. Ruiz-Mirazo, K., Peretó, J. & Moreno, A. A universal definition of life: autonomy and open-ended evolution. *Orig. Life Evol. Biosphere* **34**, 323–346 (2004).
57. de Vladar, H. P., Santos, M. & Szathmáry, E. Grand views of evolution. *Trends Ecol. Evol.* **32**, 324–334 (2017).
58. Szathmáry, E. Life: in search of the simplest cell. *Nature* **433**, 469–470 (2005).
59. Walde, P. *Prebiotic Chemistry* (Springer, 2005).
60. Santiago, G. M., Liu, K., Browne, W. R. & Otto, S. Emergence of light-driven protometabolism upon recruitment of a photocatalytic cofactor by a self-replicator. *Nat. Chem.* https://doi.org/10.1038/s41557-020-0494-4 (2020).
61. Ottelé, J. O., Hussain, A. S., Mayer, C. & Otto, S. Chance emergence of catalytic activity and promiscuity in a self-replicator. *Nat. Catal.* https://doi.org/10.1038/s41929-020-0463-8 (2020).
62. Kurihara, K. et al. Self-reproduction of supramolecular giant vesicles combined with the amplification of encapsulated DNA. *Nat. Chem.* **3**, 775–781 (2011).
63. Maynard Smith, J. Hypercycles and the origin of life. *Nature* **280**, 445–446 (1979).
64. Boerlijst, M. C. & Hogeweg, P. Spiral wave structure in pre-biotic evolution: hypercycles stable against parasites. *Phys. D.* **48**, 17–28 (1991).
65. Michod, R. Population biology of the first replicators: the origin of genotype, phenotype and organism. *Am. Zool.* **23**, 5–14 (1983).
66. Szathmáry, E. Natural selection and dynamical coexistence of defective and complementing virus segments. *J. Theor. Biol.* **157**, 383–406 (1992).
67. Czárán, T. & Szathmáry, E. *Coexistence of Replicators in Prebiotic Evolution* (Cambridge Univ. Press, 2000).
68. Czárán, T., Könnyű, B. & Szathmáry, E. Metabolically coupled replicator systems: overview of an RNA-world model concept of prebiotic evolution on mineral surfaces. *J. Theor. Biol.* **381**, 39–54 (2015).
69. Szilágyi, A., Könnyű, B. & Czárán, T. Dynamics and stability in prebiotic information integration: an RNA world model from first principles. *Sci. Rep.* **10**, 51 (2020).
70. Szabó, P., Scheuring, I., Czárán, T. & Szathmáry, E. In silico simulations reveal that replicators with limited dispersal evolve towards higher efficiency and fidelity. *Nature* **420**, 340–343 (2002).
71. Luther, A., Brandsch, R. & von Kiedrowski, G. Surface-promoted replication and exponential amplification of DNA analogues. *Nature* **396**, 245–248 (1998).
72. Szathmáry, E. & Demeter, L. Group selection of early replicators and the origin of life. *J. Theor. Biol.* **128**, 463–486 (1987).
73. Grey, D., Hutson, V. & Szathmáry, E. A re-examination of the stochastic corrector model. *Proc. R. Soc. Lond.* **262**, 29–35 (1995).
74. Damuth, J. & Heisler, I. L. Alternative formulations of multilevel selection. *Biol. Philos.* **3**, 407–430 (1988).
75. Zintzaras, E., Santos, M. & Szathmáry, E. "Living" under the challenge of information decay: the stochastic corrector model versus hypercycles. *J. Theor. Biol.* **217**, 167–181 (2002).
76. Szilágyi, A., Kun, Á. & Szathmáry, E. Early evolution of efficient enzymes and genome organization. *Biol. Direct.* **7**, 38 (2012).
77. Maynard Smith, J. & Szathmáry, E. The origin of the chromosome I. Selection for linkage. *J. Theor. Biol.* **164**, 437–446 (1993).
78. Szilágyi, A., Kovács, P. V., Szathmáry, E. & Santos, M. Evolution of linkage and genome expansion in protocells. Preprint at *bioRxiv* https://doi.org/10.1101/746495 (2019).
79. Vig-Milkovics, Z., Zachar, I., Kun, Á., Szilágyi, A. & Szathmáry, E. Moderate sex between protocells can balance between a decrease in assortment load and an increase in parasite spread. *J. Theor. Biol.* **462**, 304–310 (2019).
80. Hordijk, W., Steel, M. & Kauffman, S. A. Molecular diversity required for the formation of autocatalytic sets. *Life* **9**, 23 (2019).
81. Ashkenasy, G., Jagasia, R., Yadav, M. & Ghadiri, M. R. Design of a directed molecular network. *Proc. Natl Acad. Sci. USA* **101**, 10872–10877 (2004).
82. Szathmáry, E. A classification of replicators and λ-calculus models of biological organization. *Proc. Biol. Sci.* **260**, 279–286 (1995).
83. Kun, Á., Papp, B. & Szathmary, E. Computational identification of obligatorily autocatalytic replicators embedded in metabolic networks. *Genome Biol.* **9**, R51 (2008).
84. Muchowska, K. B. et al. Metals promote sequences of the reverse Krebs cycle. *Nat. Ecol. Evol.* **1**, 1716–1721 (2017).
85. Sievers, D. & von Kiedrowski, G. Self-replication of hexadeoxynucleotide analogues: autocatalysis versus cross-catalysis. *Chem. Eur. J.* **4**, 629–641 (1998).
86. Yao, S., Ghosh, I., Zutshi, R. & Chmielewski, J. Selective amplification by auto- and cross-catalysis in a replicating peptide system. *Nature* **396**, 447–450 (1998).
87. Kauffman, S. A. Cellular homeostasis, epigenesis and replication in randomly aggregated macromolecular systems. *J. Cybern.* **1**, 71–96 (1971).
88. Kauffman, S. A. Autocatalytic sets of proteins. *J. Theor. Biol.* **119**, 1–24 (1986).
89. Mossel, E. & Steel, M. Random biochemical networks: the probability of self-sustaining autocatalysis. *J. Theor. Biol.* **233**, 327–336 (2005).
90. Kauffman, S. A. *The Origins of Order* (Oxford Univ. Press, 1993).
91. Issac, R., Ham, Y.-W. & Chmielewski, J. The design of self-replicating helical peptides. *Curr. Opin. Struc. Biol.* **11**, 458–463 (2001).
92. Li, X. & Chmielewski, J. Peptide self-replication enhanced by a proline kink. *J. Am. Chem. Soc.* **125**, 11820–11821 (2003).
93. Issac, R. & Chmielewski, J. Approaching exponential growth with a self-replicating peptide. *J. Am. Chem. Soc.* **124**, 6808–6809 (2002).
94. Wang, B. & Sutherland, I. O. Self-replication in a Diels–Alder reaction. *Chem. Commun.* **16**, 1495–1496 (1997).
95. Robertson, M. P. & Joyce, G. F. Highly efficient self-replicating RNA enzymes. *Chem. Biol.* **21**, 238–245 (2014).
96. Nowak, P., Colomb-Delsuc, M., Otto, S. & Li, J. Template-triggered emergence of a self-replicator from a dynamic combinatorial library. *J. Am. Chem. Soc.* **137**, 10965–10969 (2015).
97. Komáromy, D. et al. Self-assembly can direct dynamic covalent bond formation toward diversity or specificity. *J. Am. Chem. Soc.* **139**, 6234–6241 (2017).
98. Simmel, F. C. Nanostructure evolution. *Nat. Mater.* **16**, 974–976 (2017).
99. Eigen, M. & Schuster, P. The hypercycle. A principle of natural self-organisation Part A: emergence of the hypercycle. *Naturwissenschaften* **64**, 541–565 (1977).
100. Wagner, N., Mukherjee, R., Maity, I., Peacock-Lopez, E. & Ashkenasy, G. Bistability and bifurcation in minimal self-replication and nonenzymatic catalytic networks. *ChemPhysChem* **18**, 1842–1850 (2017).
101. Mukherjee, R., Cohen-Luria, R., Wagner, N. & Ashkenasy, G. A bistable switch in dynamic thiodepsipeptide folding and template-directed ligation. *Angew. Chem. Int. Ed.* **54**, 12452–12456 (2015).
102. Severin, K., Lee, D. H., Martinez, J. A., Vieth, M. & Ghadiri, M. R. Dynamic error correction in autocatalytic peptide networks. *Angew. Chem. Int. Ed.* **37**, 126–128 (1998).
103. Nanda, J. et al. Emergence of native peptide sequences in prebiotic replication networks. *Nat. Commun.* **8**, 434 (2017).
104. Sadownik, J. W., Mattia, E., Nowak, P. & Otto, S. Diversification of self-replicating molecules. *Nat. Chem.* **8**, 264–269 (2016).
105. Lincoln, T. A. & Joyce, G. F. Self-sustained replication of an RNA enzyme. *Science* **323**, 1229–1232 (2009).
106. Mills, D. R., Peterson, R. E. & Spiegelman, S. An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proc. Natl Acad. Sci. USA* **58**, 217–224 (1967).
107. Mast, C. B., Schink, S., Gerland, U. & Braun, D. Escalation of polymerization in a thermal gradient. *Proc. Natl Acad. Sci. USA* **110**, 8030–8035 (2013).
108. Kreysing, M., Keil, L., Lanzmich, S. & Braun, D. Heat flux across an open pore enables the continuous replication and selection of oligonucleotides towards increasing length. *Nat. Chem.* **7**, 203–208 (2015).
109. Jayathilaka, T. S. & Lehman, N. Spontaneous covalent self-assembly of the *Azoarcus* ribozyme from five fragments. *ChemBioChem* **19**, 217–220 (2018).
110. Lee, D. H., Severin, K. & Ghadiri, M. R. Autocatalytic networks: the transition from molecular self-replication to molecular ecosystems. *Curr. Opin. Chem. Biol.* **1**, 491–496 (1997).
111. Mizuuchi, R. & Lehman, N. Limited sequence diversity within a population supports prebiotic RNA reproduction. *Life* **9**, 20 (2019).
112. Arsène, S., Ameta, S., Lehman, N., Griffiths, A. D. & Nghe, P. Coupled catabolism and anabolism in autocatalytic RNA sets. *Nucleic Acids Res.* **46**, 9660–9666 (2018).
113. Mathis, C., Ramprasad, S. N., Walker, S. I. & Lehman, N. Prebiotic RNA network formation: a taxonomy of molecular cooperation. *Life* **7**, 38 (2017).
114. Yeates, J. A. M., Nghe, P. & Lehman, N. Topological and thermodynamic factors that influence the evolution of small networks of catalytic RNA species. *RNA* **23**, 1088–1096 (2017).
115. Wächtershäuser, G. *Vitalysts and Virulysts: A Theory of Self-Expanding Reproduction* (Columbia Univ. Press, 1994).
116. Kosikova, T. & Philp, D. Two synthetic replicators compete to process a dynamic reagent pool. *J. Am. Chem. Soc.* **141**, 3059–3072 (2019).
117. Bottero, I., Huck, J., Kosikova, T. & Philp, D. A synthetic replicator drives a propagating reaction–diffusion front. *J. Am. Chem. Soc.* **138**, 6723–6726 (2016).
118. Sadownik, J. W., Kosikova, T. & Philp, D. Generating system-level responses from a network of simple synthetic replicators. *J. Am. Chem. Soc.* **139**, 17565–17573 (2017).
119. Altay, M., Altay, Y. & Otto, S. Parasitic behavior of self-replicating molecules. *Angew. Chem. Int. Ed.* **57**, 10564–10568 (2018).
120. Lee, D. H., Severin, K., Yokobayashi, Y. & Ghadiri, M. R. Emergence of symbiosis in peptide self-replication through a hypercyclic network. *Nature* **390**, 591–594 (1997).
121. Szathmáry, E. On the propagation of a conceptual error concerning hypercycles and cooperation. *J. Syst. Chem.* **4**, 1 (2013).
122. Kamioka, S., Ajami, D. & Rebek Jr, J. Autocatalysis and organocatalysis with synthetic structures. *Proc. Natl Acad. Sci. USA* **107**, 541–544 (2010).
123. Urabe, H. et al. Compartmentalization in a water-in-oil emulsion repressed the spontaneous amplification of RNA by Qβ replicase. *Biochemistry* **49**, 1809–1813 (2010).
124. Bansho, Y., Furubayashi, T., Ichihashi, N. & Yomo, T. Host–parasite oscillation dynamics and evolution in a compartmentalized RNA replication system. *Proc. Natl Acad. Sci. USA* **113**, 4045–4050 (2016).
125. Matsumura, S. et al. Transient compartmentalization of RNA replicators prevents extinction due to parasites. *Science* **354**, 1293–1296 (2016).
126. Yoshiyama, T., Ichii, T., Yomo, T. & Ichihashi, N. Automated in vitro evolution of a translation-coupled RNA replication system in a droplet flow reactor. *Sci. Rep.* **8**, 11867 (2018).
127. Liu, B. et al. Spontaneous emergence of self-replicating molecules containing nucleobases and amino acids. *J. Am. Chem. Soc.* **142**, 4148–4192 (2020).
128. Tjivikua, T., Ballester, P. & Rebek Jr, J. A self-replicating system. *J. Am. Chem. Soc.* **112**, 1249–1250 (1990).
129. Quayle, J. M., Slawin, A. M. Z. & Philp, D. A structurally simple minimal self-replicating system. *Tetrahedron Lett.* **43**, 7229–7233 (2002).
130. von Kiedrowski, G. A self-replicating hexadeoxynucleotide. *Angew. Chem. Int. Ed.* **25**, 932–935 (1986).
131. Paul, N. & Joyce, G. F. A self-replicating ligase ribozyme. *Proc. Natl Acad. Sci. USA* **99**, 12733–12740 (2002).
132. Hayden, E. J. & Lehman, N. Self-assembly of a group I intron from inactive oligonucleotide fragments. *Chem. Biol.* **13**, 909–918 (2006).
133. Draper, W. E., Hayden, E. J. & Lehman, N. Mechanisms of covalent self-assembly of the *Azoarcus* ribozyme from four fragment oligonucleotides. *Nucleic Acids Res.* **36**, 520–531 (2008).
134. Lee, D. H., Granja, J. R., Martinez, J. A., Severin, K. & Ghadiri, M. R. A self-replicating peptide. *Nature* **382**, 525–528 (1996).
135. Rubinov, B. et al. Transient fibril structures facilitating nonenzymatic self-replication. *ACS Nano* **6**, 7893–7901 (2012).
136. Otto, S. et al. Mechanosensitive self-replication driven by self-organization. *Science* **327**, 1502–1506 (2010).
137. Malakoutikhah, M. et al. Uncovering the selection criteria for the emergence of multi-building-block replicators from dynamic combinatorial libraries. *J. Am. Chem. Soc.* **135**, 18406–18417 (2013).
138. Hofbauer, J. A difference equation model for the hypercycle. *SIAM J. Appl. Math.* **44**, 762–772 (1984).
139. Guillamon, A., Fontich, E. & Sardanyés, J. Bifurcations analysis of oscillating hypercycles. *J. Theor. Biol.* **387**, 23–30 (2015).
140. Hofbauer, J., Mallet-Paret, J. & Smith, H. L. Stable periodic solutions for the hypercycle system. *J. Dynam. Differ. Equat.* **3**, 423–436 (1991).

141. Stadler, P. F. & Nuño, J. C. The influence of mutation on autocatalytic reaction networks. *Math. Biosci.* **122**, 127–160 (1994).
142. Bresch, C., Niesert, U. & Harnasch, D. Hypercycles, parasites and packages. *J. Theor. Biol.* **85**, 399–405 (1980).
143. Boerlijst, M. C. & Hogeweg, P. *Selfstructuring and Selection: Spiral Waves as a Substrate for Evolution* (Addison-Wesley, 1991).
144. Kim, P.-J. & Jeong, H. Spatio-temporal dynamics in the origin of genetic information. *Phys. D.* **203**, 88–99 (2005).
145. Szilágyi, A. et al. Ecology and evolution in the RNA world: dynamics and stability of prebiotic replicator systems. *Life* **7**, 48 (2017).
146. Könnyű, B. & Czárán, T. Specialization of early replicators in the metabolic replicator model system. *Orig. Life Evol. Biosphere* **39**, 339–340 (2009).
147. Hofbauer, J. & Sigmund, K. *Evolutionary Games and Population Dynamics* (Cambridge Univ. Press, 1998).
148. Yeates, J. A. M., Hilbe, C., Zwick, M., Nowak, M. A. & Lehman, N. Dynamics of prebiotic RNA reproduction illuminated by chemical game theory. *Proc. Natl Acad. Sci. USA* **113**, 5030–5035 (2016).

Acknowledgements
The authors thank A. Pross for discussions leading to the definition of dynamic kinetic stability. They are grateful for the EU Attract program (EmLife), the ERC (Advanced Grant 'Towards de-novo life'; ToDL 741774), the Dutch Ministry of Education, Culture and Science (Gravitation program 024.001.035) and the National Research, Development and Innovation Office (NKFIH, Hungary) under grant numbers GINOP-2.3.2-15-2016-00057 and K119347, and the Volkswagen Stiftung initiative 'Leben? — Ein neuer Blick der Naturwissenschaften auf die grundlegenden Prinzipien des Lebens' under the project 'A Unified Model of Recombination in Life' for financial support. A. Szilágyi was supported by the Bolyai János Research Fellowship of the Hungarian Academy of Sciences.

Author contributions
All authors contributed equally to the preparation of this manuscript.

Competing interests
The authors declare no competing interests.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Springer Nature Limited 2020