Chem Soc Rev

Review

Exploring the emergence of complexity using synthetic replicators

Tamara Kosikova[a] and Douglas Philp[a]

Received 00th January h,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

A significant number of synthetic systems capable of replicating themselves or entities that are complementary to themselves have appeared in the last 30 years. Building on an understanding of the operation of synthetic replicators in isolation, this field has progressed to examples where catalytic relationships between replicators within the same network and the extant reaction conditions play a role in driving phenomena at the level of the whole system. Systems chemistry has played a pivotal role in the attempts to understand the origin of biological complexity by exploiting the power of synthetic chemistry, in conjunction with the molecular recognition toolkit pioneered by the field of supramolecular chemistry, thereby permitting the bottom-up engineering of increasingly complex reaction networks from simple building blocks. This review describes the advances facilitated by the systems chemistry approach in relating the expression of complex and emergent behaviour in networks of replicators with the connectivity and catalytic relationships inherent within them. These systems, examined within well-stirred batch reactors, represent conceptual and practical frameworks that can then be translated to conditions that permit replicating systems to overcome the fundamental limits imposed on selection processes in networks operating under closed conditions. This shift away from traditional spatially homogeneous reactors towards dynamic and non-equilibrium conditions, such as those provided by reaction-diffusion reaction formats, constitutes a key change that mimics environments within cellular systems, which possess obvious compartmentalisation and inhomogeneity.

Introduction

Complex systems pervade the world around us: stock markets, metabolic pathways, ecosystems and the weather all exhibit significant complexity[1] in their behaviours. The ideas surrounding complexity are embedded within the core of most major scientific disciplines, including computer science, biology, mathematics and physics. In chemistry, these concepts have represented a rather niche interest for a long time. Historically, the avoidance of complexity has been a direct consequence of the analytical intractability of these systems both in a chemical and a mathematical sense. The blossoming of synthetic organic chemistry[2] during the 20th century was driven by a focus on the creation and manipulation of covalent bonds with exquisite control and selectivity. This control has been exploited in the elegant syntheses of a staggering array of chemical compounds using strategies that are designed to exploit the sequential, programmed application of chemical reactions with predictable outcomes, affording ultimately the target structures. By contrast, the deliberate creation of mixtures in synthetic chemistry has only become popular with the advent of combinatorial approaches[3] for the generation of compound libraries for screening purposes. The avoidance of mixtures in synthetic organic chemistry is somewhat surprising given the complexity of the chemical networks that operate in biological systems. Such networks exploit a wide range of reaction pathways that have significant numbers of interconnections. These interconnections are key in allowing the system to introduce checkpoint controls and feedback loops, created through interactions between distant network nodes, thereby permitting regulation of overall pathway. These complex networks, built using the interactions and reactions between molecules, allow biological systems to adapt and respond rapidly to external stimuli and process chemical feed stocks in defined ways as required.

Over the past 20 years, systems-based approaches have emerged,[4] first in the field of biology and, later, also in chemistry. These approaches are directed at understanding chemical and biological complexity using a holistic approach. In particular, systems chemistry[5] aims to explore the connections between the properties of individual components (molecules) and the emergence of complex, system-level behaviour arising as a result of the interactions of these components. In exploiting synthetic chemistry for the design and development of systems with complex and potentially life-like properties, systems chemistry strives to develop a better understanding of the governing principles of assembly and function in complex systems, thereby shedding light on the origins of biological complexity.

The presence of living organisms means that, indisputably, the transition[6] from simple chemical building blocks to a world awash with living organisms must have occurred at least once. A fundamental component of this transition is the emergence of self-replication[7]—*i.e.* a process by which a chemical entity templates its own synthesis. On the early Earth, self-replication must have served not only as the means of transferring

[a] School of Chemistry and EaStCHEM, University of St Andrews, North Haugh, St Andrews, Fife KY16 9ST, UK

information to molecular progeny, but also as the amplification mechanism by which certain molecular constitutions became dominant in chemical mixtures. The process of replication is ubiquitous in Nature, and the reliance on template-directed processes for the copying and transmission of information using DNA as the genetic blueprint is well understood[8] in modern biology. The structural features and uniformity of the current genetic material suggest that all modern living systems emerged from a single Last Universal Common Ancestor[9] (LUCA), marking the transition from an inanimate world to a recognisably living entity. Although our current understanding and agreement as to what constitutes life remains[10] controversial, it is clear that the transformation to a living world required that the simple, chemical components on the prebiotic Earth were elaborated into larger functional biopolymers. These functional materials must have been capable of both catalysis and replication, with the capacity to couple to primitive metabolic cycles and membrane-based compartments. There are numerous models and theories that strive to explain how this gradual transition could have taken place, and these models can generally be grouped into one of three main classes, differing in their approaches[4b,11] to the experimental study of emergence of life—the RNA world,[12] metabolism first,[13] and the compartmentalistic approach[14]—which are complemented by more integrative systems approaches.[15] These schools of thought tend to disagree on the identity of the molecular species believed to have emerged first during the process of chemical evolution. Moving beyond the restrictions and challenges imposed by studying molecules with specific relevance to current biology, systems chemistry[5] explores the importance of replication processes, driven by autocatalysis and crosscatalysis, as observed through the prism of synthetic model systems. Through such study of replication phenomena relevant to the origin of life—*i.e.* processes arising *long before* the emergence of the LUCA, it is possible to arrive at a set of minimal structural, recognition and reaction requirements needed for a system to be capable of replication, delineating boundary conditions for the emergence of the enzymatic machinery available to modern cells.

In this review, we summarise the principal modes of replication (minimal self-replication and reciprocal replication) and review a new self-replication model reported recently based on an informational leaving group. Using examples of synthetic replicators, we explore the increasingly complex networks constructed from replicating systems based on oligonucleotides, peptides and small organic molecules. Building on these systems, we examine replicating systems coupled to dynamic processes and their role in driving the outcome of replication. Finally, we discuss replicating systems that operate under reaction-diffusion conditions—an area that has received limited attention to date. We finish this review with a brief critical overview of the work achieved to date and outline of potential future prospects for the field of synthetic replicators that will facilitate our understanding of replication phenomena and the emergency of complexity in general.

Overview of replication models

A molecular self-replicating system is one that is capable of transmitting structural information through an autocatalytic process. Self-replication forms a subset of autocatalytic reactions, where the reaction product acts as a specific autocatalyst for its own formation from its constituent building blocks, and, thus, the rate of the autocatalytic reaction correlates directly with the amount of catalytically active template present within the reaction mixture.

Fig. 1a presents a minimal model of self-replication comprising three reaction channels. In this model, two molecules **A** and **B** are equipped with complementary reactive (orange and green) and recognition sites (yellow and blue). These two components can react through an uncatalysed, template-independent reaction to produce template $\mathbf{T^{AB}}$ (Channel 1, **Fig. 1a**)—a process associated with the rate constant k_{bi}. Alternatively, template formation can proceed through two recognition-mediated channels. Components **A** and **B** possess complementary recognition sites that enable them to associate in a binary complex [**A•B**] (Channel 2, **Fig. 1a**). This association preorganises the reactive sites such that the reaction between **A** and **B** proceeds within this complex in a pseudounimolecular manner to form a closed, catalytically inactive template $\mathbf{T^{AB*}}$. The second recognition-mediated pathway (Channel 3, **Fig. 1a**) exploits the open template $\mathbf{T^{AB}}$. This template possesses recognition sites complementary to those present in **A** and **B**. The recognition of **A** and **B** by $\mathbf{T^{AB}}$—governed by the association constant, K_a^{Ind}, and the level of cooperativity in the system—allows the formation of a ternary, catalytically active complex [$\mathbf{A•B•T^{AB}}$]. Within this complex, the template $\mathbf{T^{AB}}$ preorganises the reagents **A** and **B** resulting in the acceleration of the reaction between **A** and **B**. The level of rate acceleration in the corresponding rate constant k_{auto} relative to k_{bi} can be determined by calculating the kinetic effective molarity[16] ($EM_{kinetic}$, **Fig. 1c**). This template-directed reaction affords a template dimer [$\mathbf{T^{AB}•T^{AB}}$] whose stability is described the duplex association constant, K_a^{Duplex}. The dissociation of [$\mathbf{T^{AB}•T^{AB}}$] releases two template molecules that can participate in further autocatalytic reactions. The values of K_a^{Ind} and K_a^{Duplex} can be used to calculate the thermodynamic effective molarity[17] (EM_{thermo}, **Fig. 1c**) as a means of indirectly comparing the stability of the template duplex relative to the corresponding ternary complex.

It should be noted that in addition to the parameters $EM_{kinetic}$ and EM_{thermo}, which can be used to describe and quantify the behaviour of replicating systems, two other parameters—namely autocatalytic reaction order p and autocatalytic efficiency ε, introduced[7i] first by von Kiedrowski in 1993 in his detailed work on the minimal replicator theory—have been used extensively to describe self-replicating systems. Despite their widespread use, in particular in early studies of artificial replicators, these parameters, however, do not permit‡ direct comparison of replicating systems and their discussion in this review will therefore be limited.

The minimal model of self-replication presented thus far relies on self-complementary recognition. Molecular replication

can also proceed through an alternative, reciprocal mechanism, where two different template molecules are present, each equipped with recognition sites that are complementary those on the other template, thereby enabling formation of a template heteroduplex. The minimal model of reciprocal replication (**Fig. 1b**) constitutes two uncatalysed reaction channels and two template-directed crosscatalytic pathways. This simplest form of reciprocal system is constructed from two pairs of building blocks, each equipped with complementary reactive sites: **A** can react only with **C** and **B** can react only with **D**. Bimolecular reactions of these components afford two templates, $\mathbf{T}^{\mathbf{AC}}$ and $\mathbf{T}^{\mathbf{BD}}$ (Channels 4 and 5, **Fig. 1b**), at rates that can, again, be described by the appropriate bimolecular rate constants, k_{bi}. These templates do not have the necessary self-complementary recognition sites required for self-replication. However, the reciprocal nature of the recognition processes allows the $\mathbf{T}^{\mathbf{BD}}$ template to crosscatalyse the formation of $\mathbf{T}^{\mathbf{AC}}$ through the formation of a catalytically active ternary complex [**A•C•T**$^{\mathbf{BD}}$] and *vice versa*. These two crosscatalytic template-mediated channels (Channels 6 and 7, **Fig. 1b**) are each described by their respective pseudounimolecular rate constants, k_{cross}, and afford a template heteroduplex that dissociates to return one molecule of $\mathbf{T}^{\mathbf{AC}}$ and one of $\mathbf{T}^{\mathbf{BD}}$ back to the start of the cycles, allowing them to participate in further crosscatalytic cycles. As in the case of the self-replicating model, the availability of the free catalytically active templates in solution depends on the relevant association constant, in this case $K_a^{Heteroduplex}$, and the corresponding EM_{thermo}.

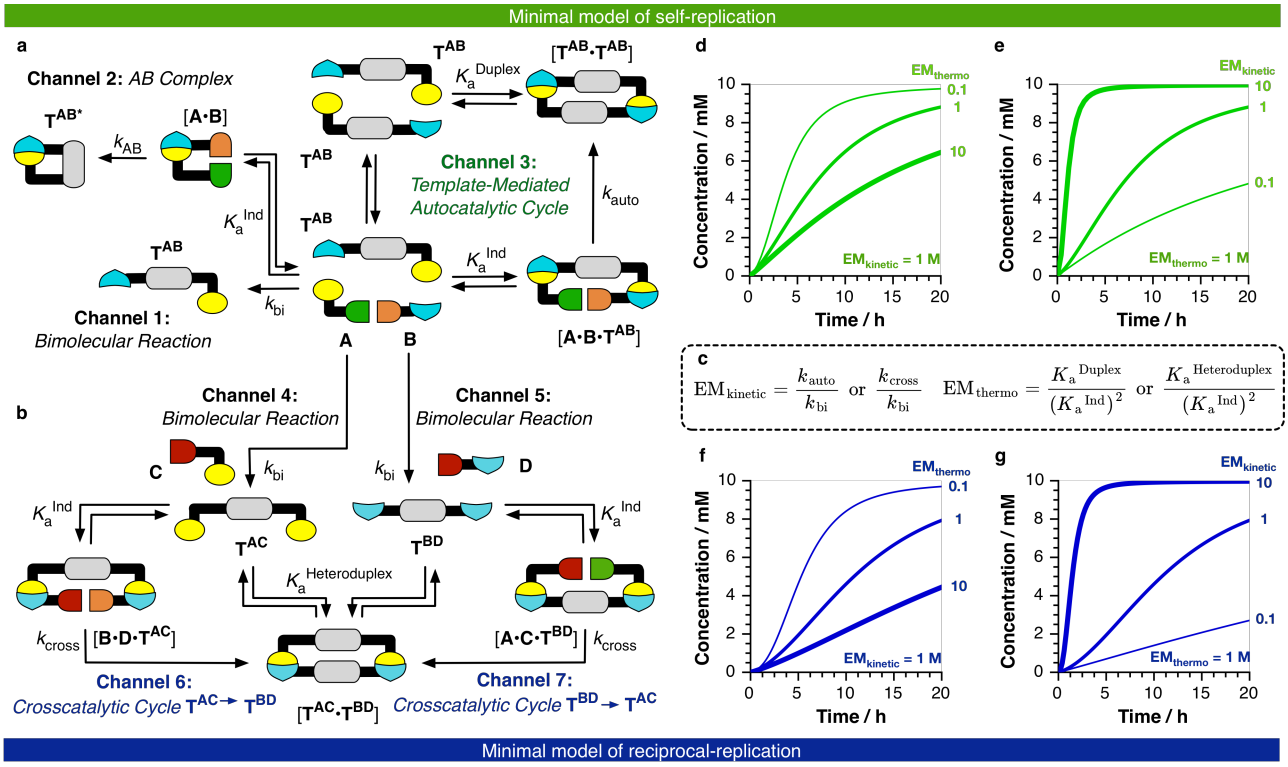


Fig. 1 Cartoon representations of **a** the minimal model of self-replication and **b** the minimal model of reciprocal replication. **a** Components **A** and **B** possess reactive (green/orange) and recognition (blue/yellow) sites. Formation of template $\mathbf{T}^{\mathbf{AB}}$ can proceed through a slow bimolecular reaction (Channel 1) and template-mediated self-replicating pathway (Channel 3). Channel 2 described the formation of catalytically inactive product $\mathbf{T}^{\mathbf{AB}}$* through a binary reactive complex [**A•B**]. **b** Two templates, $\mathbf{T}^{\mathbf{AC}}$ and $\mathbf{T}^{\mathbf{BD}}$, containing complementary recognition sites (yellow and blue) are formed initially through the bimolecular reactions of **A** with **C** (Channel 4) and **B** with **D** (Channel 5). Once formed, these templates can participate in two template-mediated reciprocal pathways (Channels 6 and 7), where $\mathbf{T}^{\mathbf{AC}}$ is formed ***via*** the catalytically active ternary complex [**A•C•T**$^{\mathbf{BD}}$], and $\mathbf{T}^{\mathbf{BD}}$ ***via*** the analogous complex [**B•D•T**$^{\mathbf{AC}}$]. **c** Equations for determination of effective kinetic molarity ($EM_{kinetic}$) and effective thermodynamic molarity (EM_{thermo}) of a replicating system. **d–g** Effect of $EM_{kinetic}$ and EM_{thermo} on the time course profile of a minimal self-replicating system ($\mathbf{T}^{\mathbf{AB}}$, green, **d**,**e**) and minimal reciprocal replicating system ($\mathbf{T}^{\mathbf{AC}}$ or $\mathbf{T}^{\mathbf{BD}}$, blue, **f**,**g**). Simulated conditions: 10 mM, $k_{bi} = 0.001\ M^{-1}\ s^{-1}$, $K_a^{Ind} = 1000$ M; **d**,**f**: parameters k_{auto} and k_{cross} were varied, while K_a^{Duplex} and $K_a^{Heteroduplex} = 10^6\ M^{-1}$; **e**,**g**: parameters K_a^{Duplex} and $K_a^{Heteroduplex}$ were varied, while k_{auto} and $k_{cross} = 0.001\ s^{-1}$.

The kinetic profiles of replicating systems can display, and are often associated with, sigmoidal time courses. The simulated kinetic profiles for replicators based on the minimal model of self-replication (**Figs. 1d** and **1e**) and reciprocal replication (**Figs. 1f** and **1g**), however, reveal that the appearance of the concentration *vs.* time profiles for these systems is strongly dependent on the kinetic and thermodynamic parameters. Although, the presence of a sigmoidal reaction profile is directly related to the processes and reactions inherent to the minimal models described above (**Figs. 1a** and **1b**), it is not diagnostic of their presence. The rate of self-replication is directly related to the amount of free, catalytically active template present in the reaction mixture. Therefore, a lag period in the formation of product is observed at the beginning of each reaction time course—at this time, only unreacted fragments are present and the reaction proceeds

through the template-independent bimolecular pathway. Once a sufficient concentration of the template has built up in solution to associate with the unreacted materials, which is dependent on the strength (K_a^{Ind}) of the relevant recognition process, the autocatalytic cycle is established. For this reason, the maximum reaction rate for a self-replicating system is not observed at the start of the reaction, but instead at a later time point. Undertaking the same reaction in the presence of preformed template, added at the start of the reaction, should result in a shortening or disappearance of the lag period and an enhanced initial rate of reaction, if the system is self-replicating. For this reason, template-instructed experiments are used to confirm the capacity of a molecular framework to replicate.

In a reciprocal replicating system (**Fig. 1b**), the formation of both templates will proceed at the rate of the respective uncatalysed bimolecular reactions until sufficient concentrations of templates, required for efficient crosscatalysis mediated by the assembly and reaction of building blocks in the catalytically active ternary complexes, have built up in the reaction mixture. In the presence of free template molecules, each active crosscatalytic reaction can proceed through the template-directed pathway, and, as a result, the corresponding reaction rate correlates directly with the amount of free template present in solution.

Template duplexes play a crucial role in both self-replication and reciprocal replication, resulting in the sensitivity of these systems to product inhibition. In both processes, product inhibition can dramatically reduce the efficiency of replication, as the template is sequestered within product duplexes, and is, therefore, not available as the monomeric template required for efficient auto- and crosscatalysis. In the ideal situation, these systems would be designed such that the catalytically active ternary complexes are more stable than the product duplexes formed, which would facilitate the dissociation of these duplexes.

Generally, there are two reaction parameters can be varied fairly easily in order to improve the catalytic efficiency of a replicating system operating through one of the minimal models: concentration and temperature. By altering the concentration at which the reaction is performed, the proportion of free and bound template in solution can be manipulated. For example, changing the reaction concentration can be useful for controlling the relative contribution of the bimolecular and the recognition-mediated pathways towards the production of template. If, for example, the reaction concentration is halved, the rate of the bimolecular reaction will decrease by a factor of four. By contrast, the rate of the pseudounimolecular reaction is directly proportional to the concentration of the ternary complex. Hence, there is a more complex relationship between this reaction process and concentration depending on whether the overall concentration is above or below the K_d for the ternary complex. Replicating systems generally contain a large number of components that can be present in bound or unbound states within the reaction mixture. Therefore, in order to deconvolute the overall effect of concentration on the reaction time course, it is often useful to employ kinetic simulations, utilising rate and association constants that are known or for which reasonable estimates can be made.

Alterations in the temperature at which the experiment is performed can also be exploited for increasing the efficiency of replication. For example, an increase in the reaction temperature will reduce the association constant for the product duplex, in turn decreasing product inhibition. Simultaneously, however, the increase in temperature also affects the assembly of the building blocks into the catalytically active complexes as K_a^{Ind} is also decreased. Changes in reaction temperature also affect the reaction processes and sometimes these changes to complex stability and reactivity will act in concert. For example, a decrease in the reaction temperature slows down both the bimolecular and unimolecular reaction rate, whilst simultaneously increasing the stabilities of all complexes relying on non-covalent interactions. The latter can perhaps lower the quantity of template required for its assembly with the reaction components in to catalytically active complexes. Ultimately, predicting the overall effect of changing temperature on the efficiency of replication can pose a significant challenge.

Taking inspiration from the capacity of non-coding tRNA to act both as a leaving group and carry information, as well as their previous work on nucleotide-based systems, Herdewijn and co-workers described[18] a model (**Fig. 2**) that exploits the concept of an informational leaving group (ILG) to create a replication process in which the template duplex is weakly bound and, thus, less susceptible to product inhibition. In this ILG-based model, a template molecule $\mathbf{T^{EF}}$ is formed by the reaction of two precursors, **E** and **F** (Channel 1, **Fig. 2a**). Each precursor contains two recognition sites and one reactive site. Interestingly, the increased number of recognition sites means that the components **E** and **F** can associate together to form two binary complexes, [**E**•**F**]-**1** and [**E**•**F**]-**2** (**Fig. 2b**) as well as into larger oligomeric assemblies. The template forming reaction in this model is associated with a concomitant displacement of the leaving group, **LG** (**Fig. 2a**), to afford template $\mathbf{T^{EF}}$ that is equipped with three recognition sites, and not template $\mathbf{T^{EF}}$* (**Fig. 2b**), the formation of which would be expected in the absence of an operative ILG mechanism. In a manner similar to the minimal model of self-replication, the produced template $\mathbf{T^{EF}}$ can associate with the precursors, **E** and **F**, to form a catalytically active ternary complex [**E**•**F**•$\mathbf{T^{EF}}$]. Template $\mathbf{T^{EF}}$ preorganises the reactive sites of **E** and **F**, thus accelerating their reaction within the ternary complex to produce [$\mathbf{T^{EF}}$•$\mathbf{T^{EF}}$•**LG**] (Channel 2, **Fig. 2a**). Since the formation of each template molecule is associated with the release of the leaving group **LG**, the original template $\mathbf{T^{EF}}$ in the resulting complex [$\mathbf{T^{EF}}$•$\mathbf{T^{EF}}$•**LG**] is participating in two individual recognition-mediated interactions with (i) another molecule of $\mathbf{T^{EF}}$ and (ii) **LG**, which makes this complex significantly less stable in theory than the analogous template dimer formed in the minimal model (see [$\mathbf{T^{AB}}$•$\mathbf{T^{AB}}$], **Fig. 1a**).

Chem Soc Rev

Review

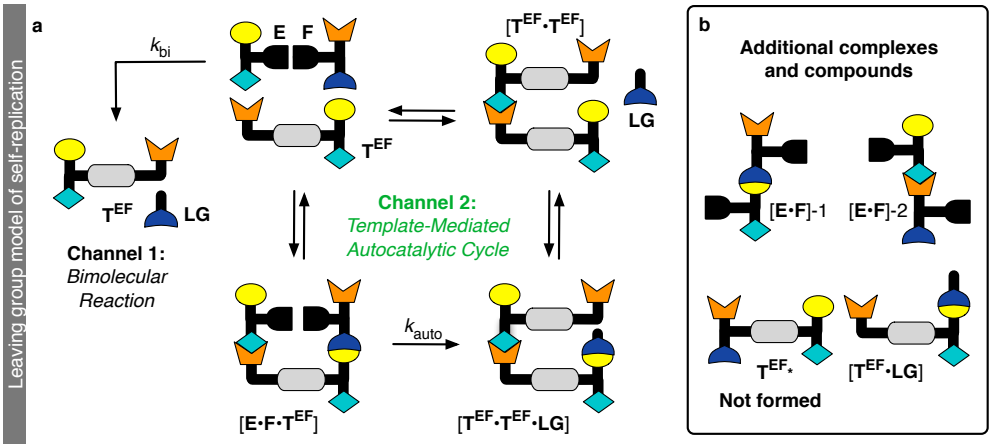


Fig. 2 **a** Cartoon representation of a minimal model of self-replication based on an informational leaving group (ILG) strategy reported by Herdewijn and co-workers in 2016. Precursors **E** and **F** are each equipped with two recognition sites and a reactive site (black), and react to form a molecule of template $\mathbf{T^{EF}}$—a process which is associated with a concomitant release of a leaving group **LG** (Channel 1). The template $\mathbf{T^{EF}}$ can catalyse the reaction between **E** and **F** by preorganising their reactive sites in a catalytically active ternary complex [$\mathbf{E\cdot F\cdot T^{EF}}$]. The pseudo-unimolecular ligation step (Channel 2) produces a ternary complex that contains two template molecules and **LG**, [$\mathbf{T^{EF}\cdot T^{EF}\cdot LG}$]. Dissociation of **LG** produces a template complex [$\mathbf{T^{EF}\cdot T^{EF}}$], which exists in equilibrium with two free catalytically active molecules of $\mathbf{T^{EF}}$. **b** The recognition sites on precursors **E** and **F** allow them to form binary complexes [**E·F**]-1 and [**E·F**]-2 (as well as longer oligomeric assemblies, not shown here). Formation of template $\mathbf{T^{EF}}$* in the ILG-based model is not possible. While the association of **LG** with template $\mathbf{T^{EF}}$ in complex [$\mathbf{T^{EF}\cdot LG}$] could affect the efficiency of replication, simulations by Herdewijn and co-workers show that ILG replicators can outperform systems based on the classical minimal model. Currently, there are no experimental examples of this replication model.

Dissociation of **LG** from the ternary complex [$\mathbf{T^{EF}\cdot T^{EF}\cdot LG}$] produces a template dimer [$\mathbf{T^{EF}\cdot T^{EF}}$], which is in equilibrium with the free form of the catalytically active template $\mathbf{T^{EF}}$. This new ILG-based model of self-replication provides an interesting mechanism for destabilising the product duplex, with the potential to improve replication efficiency. Using kinetic simulations, the authors demonstrated the performance of the ILG-based model under different temperature regimes and compared their model to the well-established minimal model of self-replication discussed previously, showing that it can outperform replicators based on the classical minimal model despite the potential for the formation of the unproductive complex [$\mathbf{T^{EF}\cdot LG}$]. Currently, despite the considerable potential of this model, its overall viability and, in particular, the possible challenges associated with product inhibition arising from the association of the free template $\mathbf{T^{EF}}$ with the leaving group **LG** in complex [$\mathbf{T^{EF}\cdot LG}$] are untested experimentally.

Synthetic minimal self-replicating systems

The establishment of theoretical requirements[7] for the creation of self-replicating systems spawned reports of experimental synthetic systems capable of templating their own synthesis or that of other molecules. The abiotic replicating systems reported to date range from simple replicators operating in isolation in a reaction medium to a collection of increasingly complex networks of replicators, with each example enabling us to advance our understanding of the molecular origins of life and the associated increase in complexity. In this section, examples of minimal synthetic replicating systems based on oligonucleotides, peptides and small molecules will be presented.

Following extensive work[19] on template-directed synthesis of oligonucleotides by Orgel and co-workers, the von Kiedrowski laboratory reported[20] in 1986 the first example of non-enzymatic self-replication in a model chemical system based on an oligonucleotide strand with a palindromic sequence (**Fig. 3a**). In the replication cycle, a trinucleotide **CCG** (protected at the 5' end), activated using a carbodiimide (**EDCI**) *in situ*, was coupled to trinucleotide **CGG** (protected at the 3' end), affording a hexanucleotide template **CCGCGG**. The complementarity of the template sequence, now protected at both the 3' and 5' ends, to the building blocks **CCG** and **CGG** enabled the association of the template with the trinucleotide precursors *via* hydrogen-bonding-mediated recognition between the Watson-Crick base pairs, thus facilitating the formation of phosphodiester bond, although with relatively low efficiency. The ability of the hexanucleotide to replicate was improved when the sequence was redesigned[21] to

take advantage of an amine as the nucleophile on the cytosine in the **CGG** trinucleotide instead of a hydroxyl group. This replicator based on a 3',5'-phosphoamidate linkage exhibited parabolic growth and sigmoidal reaction profile, with an autocatalytic efficiency‡ ε of 420. In subsequent work, von Kiedrowski and co-workers have designed[22] a self-replicating oligonucleotide system assembled from three building blocks and employed[23] DNA strands immobilised on solid supports to overcome the product inhibition that limited the replication efficiency of their earlier designs.

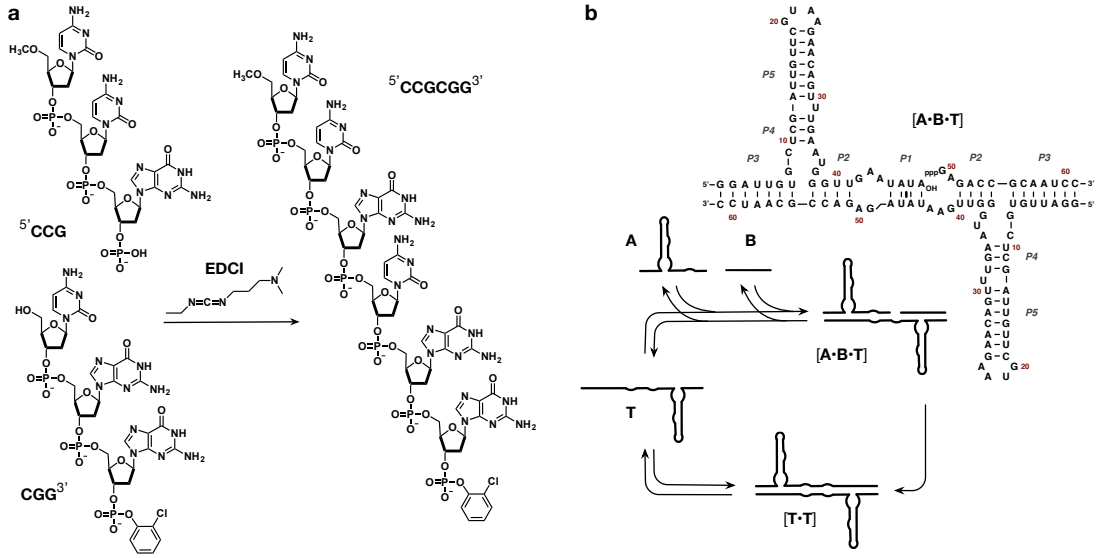


Fig. 3 **a** A self-replicating hexanucleotide with a palindromic **CCGCGG** sequence capable of catalysing its own formation by preorganising two smaller trinucleotides **CCG** and **CGG** in a catalytically active ternary complex, as reported by von Kiedrowski in 1986. **b** Schematic representation of the minimal replication cycle of a self-replicating RNA ribozyme **T**, formed from two smaller subunits, **A** and **B**, as reported by Paul and Joyce in 2002. The structural details of the self-replicating ligase ribozyme illustrate the two-fold symmetry of the template.

Soon after the first example reported[20] by von Kiedrowski, Zielinski and Orgel demonstrated[24] a self-replicating minimal system based on the chemistry of 3'-amino-3'-deoxy nucleotides. The replication of the resulting tetranucleotide analogue template (ε = 340), however, was also hampered by product inhibition. In 2002, Paul and Joyce reported[25] the first example of an RNA-based self-replicating system, exploiting[26] an R3C RNA ligase ribozyme. The study employed an adapted version of an R3C ligase ribozyme, capable of catalysing the formation of a 3',5'-phosphodiester bond between two individual RNA molecules. The RNA ribozyme template **T** was designed to be capable of ligating two RNA subunits **A** and **B** (**Fig. 3b**) *via* a ternary complex, producing an exact copy of itself. The reaction was mediated by a nucleophilic attack of the 3'-hydroxyl group of **A** on the α-phosphate of the 5'-pppG of subunit **B** to give a template duplex [**T•T**], with a two-fold centre of symmetry. When the formation of the RNA ribozyme was examined in the presence of preformed template, the system showed a clear increase in the initial rate of template formation. The enhanced reaction rate observed in the template-instructed experiment, however, was found to be only temporary, indicating that, while the ternary complex [**A•B•T**] contributes to the formation of the template, a competing process in the system prevents the replication cycle from operating efficiently. Kinetic fitting revealed that two processes contribute to the formation of template **T**. Specifically, the enhanced rate of replication observed early on in the reaction was attributed to ligation mediated by [**A•B•T**] complexes, which were produced by the association of preformed complexes [**B•T**] with **A**. By contrast, the second, slower phase was dominated by the bimolecular reaction of **A** and **B** in the absence of template. The authors suggested that the decrease in the efficiency of the designed RNA system over time stems from the similarities in the nucleotide sequences of components **A** and **B**, which results in inhibition mediated by the formation of an inactive binary complex [**A•B**], the binding of which could not be disrupted by the addition of preformed template. The authors found that the deleterious effect exerted by the strong [**A•B**] complex on replication could be circumvented by either premixing **T** with **B** prior to the addition of **A** or by the addition of an excess of **A** to the reaction mixture.

The transfer of information in the synthetic oligonucleotide-based synthetic replicators discussed thus far relies on specific, well-defined patterns of molecular recognition elements—*i.e.* the recognition is directed by hydrogen-bonding interactions between donor and acceptor elements, encoded within each oligonucleotide sequence that permit the assembly into catalytically active complexes. Formation of a self-replicating peptide, likewise, necessitates that the peptide template be able to associate with smaller peptide fragments in some form of a catalytically active complex. In comparison to oligonucleotides, however, peptides possess a considerably richer structural lexicon, arising from the increased number of building blocks used—*i.e.* 20 amino acids compared to the four nucleotides in replicating systems based on DNA. Inter-peptide recognition is determined not only by the primary sequence of the amino acids, but also by the secondary and tertiary structures created by the interactions between those amino acids.

The first experimental demonstration of peptide replication was reported[27] in 1996 by Ghadiri and co-workers. The design of the 32-residue self-replicating α-helical peptide (**Fig. 4a**) was inspired[28] by the leucine zipper domain of the yeast transcription factor GCN4. This peptide replicator design exploited a simple protein folding motif—an α-helical coiled-coil, distinguished by peptide sequences composed of heptad repeats $(abcdefg)_n$, which produce two coiled-coils wrapped around each other with a slightly left-handed, superhelical twist. The sequence of the reported peptide replicator (**Fig. 4a**) implements six substitutions relative to the wild type GCN4. Of particular interest is the substitution of a neutral, hydrophilic asparagine residue (position 16 in the sequence), located within the core hydrophobic region, with a hydrophobic valine residue, which enabled equilibration between a dimeric and trimeric coiled-coil structure (*i.e.* catalysis could be mediated by one- and/or two-stranded α-helical coil).

Monomeric coiled-coil peptides are generally present as random coils in aqueous solutions. However, these peptides can adopt a completely α-helical structure provided that a suitable template framework for directing their assembly is present. As with other minimal replicating systems, an autocatalytic peptide system built from two smaller complementary peptides, each equipped with a reactive group, needs to assemble on a peptide sequence that positions these fragments in an orientation that promotes their reaction. In a situation where these fragments are the constituent parts of a longer template sequence, the product formed by their reaction constitutes an identical copy formed through self-replication. The self-replicating peptide described by Ghadiri was capable of forming both dimeric and trimeric assemblies, with both **T** and [**T**•**T**] serving as potential autocatalytic templates. The recognition mediating the template-directed reaction in this peptide system was afforded by interactions between complementary hydrophobic and electrostatic peptide surfaces (**Figs. 4a** and **4b**). Specifically, residues at positions *a* and *d* (**Fig. 4a**, red) within the peptide sequence drive the inter-helical recognition through hydrophobic interactions, playing a pivotal role in determining the stability and orientation of coiled-coil peptides. Residues in positions *e* and *g* (**Fig. 4a**, blue) within the heptad repeat are responsible for driving the intra-component recognition through electrostatic interactions. Residues *b*, *c* and *f* (**Fig. 4a**, yellow), on the other hand, are located on the solvent exposed surface, and do not therefore contribute to the recognition. The ligation site (**Fig. 4a**, orange) is positioned on the solvent-exposed surface in order to avoid potential interference with the hydrophobic core responsible for recognition. The peptide coupling strategy exploited by Ghadiri and co-workers employed a thioester-promoted native peptide ligation (**Fig. 4c**), first described[29] by Kent and co-workers in 1994. Peptide template **T** is formed through the reaction of an N-terminal, 17-residue electrophilic fragment **E** (**Fig. 4d**), activated as a thiobenzyl ester, and a 15-residue C-terminal nucleophilic fragment **N** (**Fig. 4d**), bearing a free cysteine residue. The native ligation reaction proceeds through an intermediate thioester (**Fig. 4c**), which undergoes intramolecular rearrangement to produce the final, more thermodynamically stable amide bond at the ligation site.

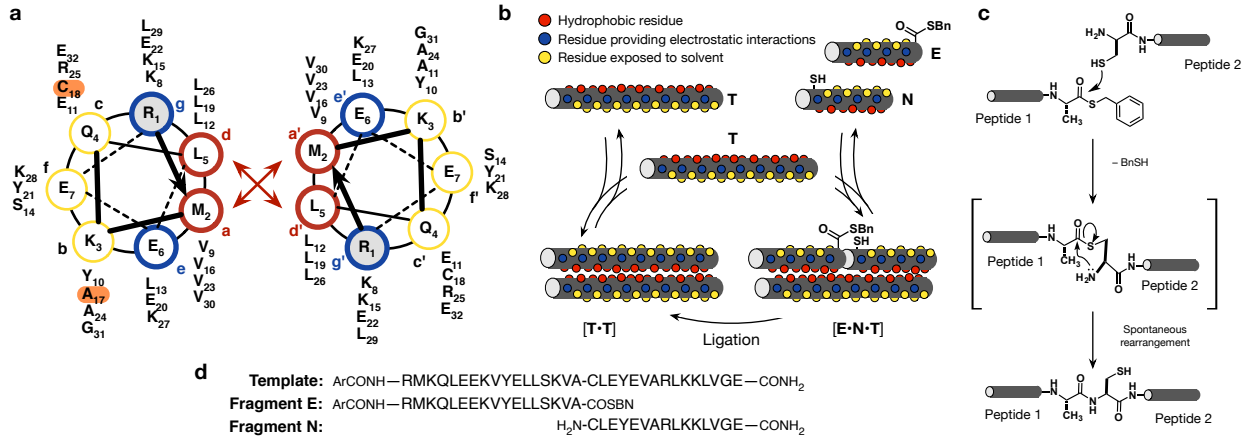
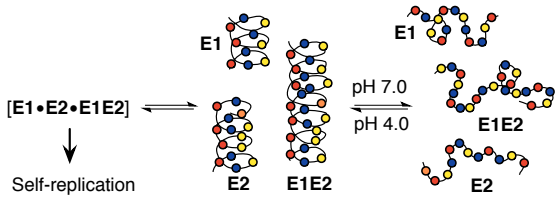


Fig. 4 **a** and **b** Design of an α-helical, coiled-coil peptide capable of self-replication, reported by Ghadiri and co-workers in 1996, featuring a heptad repeat $(\boldsymbol{abcdefg})_n$. Recognition between peptides and their assembly into complexes is mediated by the recognition between the hydrophobic residues at positions ***a*** and ***d*** (red) and electrostatic interactions between residues at positions ***e*** and ***g*** (blue). Residues ***b***, ***c*** and ***f*** (yellow) are exposed to the solvent surface and do not contribute to recognition. Position of the two residues, alanine (pre-activated as thiobenzyl ester) and cysteine, required for native chemical ligation (see **c**) is highlighted in orange. **c** Mechanism of native chemical ligation that leads to the formation of self-replicating peptide. The ligation to produce template **T** occurs between an N-terminal pre-activated thiobenzyl ester on **E** (peptide 1) and C-terminal cysteine located on nucleophilic fragment **N** (peptide 2). **d** Sequences of fragments **E** and **N** and template **T**.

The ability of the designed coiled-coil peptide to self-replicate was established unambiguously through template-instructed kinetic experiments, where the ligation reaction was examined in the presence of increasing quantities of preformed peptide template **T**, added at the reaction onset. The replication process displayed parabolic growth, where the initial rate of ligation correlated with the square root of the concentration of the initial template added, suggesting that replication is limited by product inhibition despite the relatively high autocatalytic efficiency (ε = 500). The authors established that the efficiency of replication is extremely sensitive to the identity of the residues within the peptide sequence by exploring conservative

substitutions of the residues at the key positions within the sequence of the peptide, *a* and *d* (*e.g.* a peptide containing an alanine residue instead of valine at position 9 displayed no significant template-assisted catalytic activity). Similarly, in the presence of guanidinium hydrochloride, a chaotropic reagent, revealed no observable enhancement in recognition-mediated formation of template. The authors also probed the potential contributions of the reactions between the binary complexes and the individual smaller fragments, *i.e.* [**T**•**E**] with **N** and [**T**•**N**] with **E**, towards the production of peptide **T**, by examining reactions with "crippled" peptide sequences containing a single mutation within the hydrophobic recognition-mediating core of the peptide fragments. Kinetic analyses of the data obtained with these fragments confirmed that the addition of the mutated templates, formed by the reaction of a "crippled" and a native fragment, capable of associating with **E** or **N** into binary complexes only, afforded no enhancement in the rate of formation of the native peptide. Taken together, the authors were able to demonstrate for the first time that a recognition-mediated, enzyme-free peptide replication is possible in systems exploiting the coiled-coil structural motif.

This first example[27] of a self-replicating peptide from Ghadiri and co-workers was followed rapidly by a number of other examples[30–32] exploiting similar design principles. Notable examples were described by Chmielewski and co-workers, who reported[31,32] the kinetic analyses of two peptide replicators, the functions of which could be modulated through environmental control of pH[31] (**E1E2** system, **Fig. 5**) and ionic strength[32] (referred to as the **K1K2** system), thereby allowing peptide self-replication to be switched on and off through the application of environmental stimuli. These two systems, together with the replicators reported by Ghadiri and co-workers, however, all suffered from varying degrees of product inhibition.

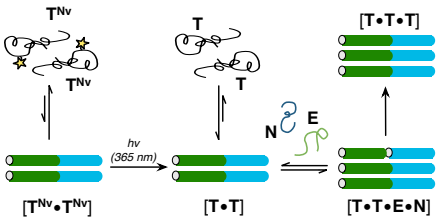


E1E2: Ac—ELYALEKELGALEKELA-CLEKELGALEKELYALEK—$CONH_2$
E1: Ac—ELYALEKELGALEKELA-COSR
E2: H_2N-CLEKELGALEKELYALEK—$CONH_2$

Fig. 5 A cartoon representation of a self-replicating peptide **E1E2** modulated by pH, as described by Chmielewski and co-authors. The recognition-mediated reaction processes in the system, and, thus also the formation of the catalytically active complex, are only effective at low pH (in this case, pH = 4), at which the two glutamate residues are protonated. R = $(CH_2)_2CONH_2$.

Since these initial reports, researchers in the area of peptide self-replication have made remarkable progress towards the design and implementation[33,34] of peptide replicators in which product inhibition is minimised or almost eliminated. For example, Issac and Chmielewski demonstrated[33] that efficiency of peptide replication increases considerably when the length of the peptide sequence is reduced by one heptad. This shortened peptide replicator (26 residues *vs.* 32 in the original design) exhibited autocatalytic efficiency (ε) of 100000, reaching nearly exponential replication. The decrease in product inhibition of this shortened system was further corroborated by a 20 °C decrease in the melting temperature of the re-designed 26-residue peptide replicator relative to the parent **E1E2** system. Chmielewski and Li reported[34] an alternative strategy for overcoming product inhibition, in which the destabilisation of the product duplex and the consequent enhancement in replication efficiency was driven by the incorporation of a proline residue at a strategic location within the replicator sequence introducing[35] a significant kink. In this case, the proline-containing template replicated with a 260-fold increase in catalytic efficiency (ε = 320000) relative to the system lacking the proline.

Ashkenasy and co-workers introduced[36] an attractive strategy for exploiting light as a control trigger for peptide replication. The experimental design was again based on a dimeric coiled-coil assembly and exhibited high levels of sequence specificity for replication. The peptide template contains a photocleavable moiety (**Fig. 6**, yellow star), 6-nitroveratryloxycarbonyl (**Nv**) bound to a lysine residue (*g*), that participates in electrostatic interactions important for replication. This caging element afforded a peptide template that has a significantly reduced propensity for dimerisation and association with smaller peptide fragments, **N** and **E**. Exposure to light as a stimulus led to efficient removal of the **Nv** group, thereby re-establishing the ability of the template to self-replicate through the normal association pathway (**Fig. 6**).



T^{Nv}: Ar—KKGGRVARLEKKVSALEKKVG-CLEYEVARLK^{Nv}KLVGE—$CONH_2$
T: Ar—KKGGRVARLEKKVSALEKKVG-CLEYEVARLKKLVGE—$CONH_2$
E: Ar—KKGGRVARLEKKVSALEKKVG-COSR'
N: H_2N-CLEYEVARLKKLVGE—$CONH_2$

Fig. 6 **a** Replicating peptide system controlled by light. In the absence of light (left), peptide template **T^{Nv}** exists as a random coil, incapable of dimerisation. Exposure to monochromatic UV light cleaves the caging moiety (yellow star, position ***g***), producing active template **T**. The activated template forms a dimeric coiled-coil [**T•T**] that is capable of associating with smaller fragments **E** and **N** in a quaternary complex [**T•T•E•N**]. K^{Nv} (highlighted in red) represents 6-nitroveratryloxycarbonyl protected lysine; R′ = ethanesulfonic acid.

The authors demonstrated that the amount of template within the mixture can be modulated by the length of exposure of the system to light, thereby providing a direct control over the rate of replication. By exploiting two other nucleophilic fragments, the authors also demonstrated that the concept of light-induced replication can be used to implement AND logic

operation, where both light and addition of preformed template were used as instructions.

Biology is dominated by nucleic acids and proteins, which play key informational and catalytic roles. As a consequence, theories[12,15] for the emergence of life on the prebiotic earth emphasise the role of these classes of molecules and their emergence in the transition from chemical matter to life. As a result of its capacity[37] to not only store hereditary information but to also act as a catalyst, RNA is often suggested to have played an essential part in the appearance of biological complexity. Despite the obvious relevance of RNA for living systems, there are certain difficulties associated[38] with the formation of strands of RNA of sufficient length and in large enough quantities from the corresponding building blocks. The Eschenmoser group has investigated[11a,39] the "*Etiology of Nucleic Acid Structure*", examining the chemical rules that underlie the origin and specific nature of RNA, as well as potential RNA alternatives and predecessors. Peptide nucleic acid[40] (PNA) based oligomers have been identified[41] as a potential genetic material that could have preceeded RNA on prebiotic earth. Unlike in DNA and RNA, the backbone in PNA is comprised of an uncharged, achiral, pseudopeptide framework. In addition, the components required for its formation, as well as that of other related peptide nucleic acids, have been detected[42] in experiments mimicking potentially plausible prebiotic conditions and have been detected[41b,41c] in the Murchison meteorite. Furthermore, the possibility of information transfer from PNA to RNA and DNA to PNA has been demonstrated[43] experimentally, further corroborating the potential role of PNAs as a prebiotically relevant class of molecules. Fairly recently, the first example of a synthetic self-replicating PNA (**Fig. 7**) has been reported[44] by Plöger and von Kiedrowski.

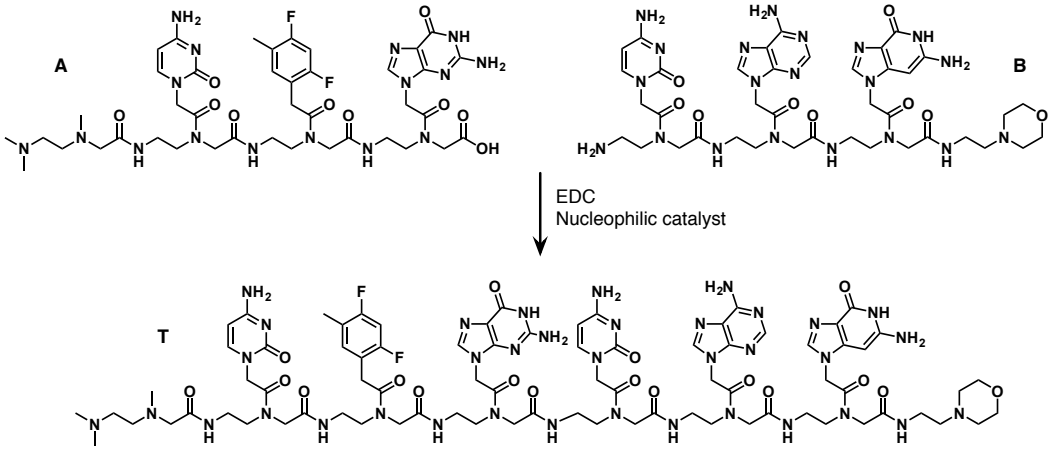


Fig. 7 PNA-based self-replicating template **T** reported by Plöger and von Kiedrowski, formed by the reaction of two trimeric building blocks, **A** and **B**.

The authors employed two building blocks (**Fig. 7**), namely, component **A** bearing two nucleobases and difluorotoluene residue and component **B** bearing three nucleobases, for the construction of the self-complementary hexa-PNA template **T**. The ability of the designed PNA sequence to template its own formation from **A** and **B** was demonstrated successfully using ^{19}F NMR spectroscopy kinetic experiments employing preformed template **T** as instruction. Similarly to the behaviour observed for the early oligonucleotide designs, the dissociation of the product template represented the rate-limiting step in this system. Through optimisation of the nucleophilic catalyst, pH and co-solvent, the authors were able to improve the autocatalytic efficiency (ε) of their system by two orders of magnitude. In parallel, Nielsen and Singhal have reported[45] an example of a PNA-based reciprocal replication system, where four pentameric precursor PNAs could react together *via* two primary crosscatalytic pathways and two pathways generating self-complementary products. Interestingly, while the two reciprocal templates acted as crosscatalysts, neither of the self-complementary decameric PNA products displayed any appreciable autocatalytic activity. The authors also investigated the influence of oligomer length on replication, demonstrating that a shorter system constructed from tetrameric building blocks (octameric templates) was less efficient than that utilising pentameric components.

In terms of the chemical evolution that took place on the early Earth, two features are of significant interest: the extant reaction processes and the local environment within which these processes take place. These considerations are tightly interconnected, influencing together which biomolecules are formed, where and how. Ideally, these areas of interest to chemical evolution could be probed with prebiotically plausible molecules—such as the examples of oligonucleotide and peptide replicators described thus far. Nevertheless, as we are only beginning to unravel the nature of the chemistry that played a role in the transition from a non-living world to a living one, attempting to investigate the key processes in the chemical evolution with molecules that are relevant prebiotically brings together two levels of difficulty, which could be perhaps understood more easily in isolation. It can be instructive, therefore, to move beyond the restrictions and challenges imposed by studying molecules with a specific relevance to current biology—focusing instead on the study of the complex phenomena and reaction processes that are

relevant to the behaviour of complex mixtures in the prebiotic context. In particular, the fundamental study of replication processes is a key component in unravelling this problem. While taking inspiration from the complexity of natural systems, synthetic systems that are based on small molecules strive to exploit structural and interactional simplicity in network components, achieving synthetic systems with well-defined chemistries and interactions—ones that can be analysed and characterised experimentally. The next section provides an overview of minimal replicators constructed from small organic molecules.

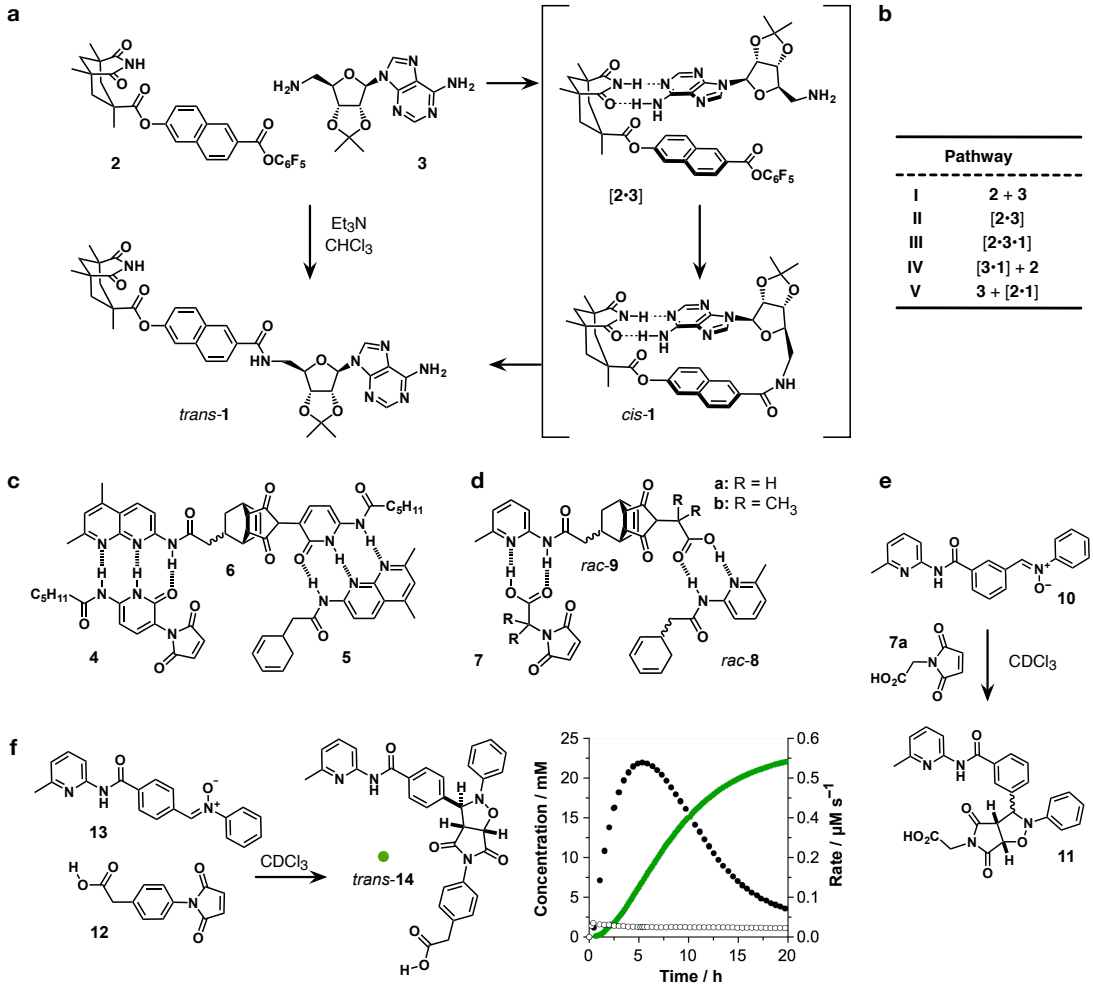


Fig. 8 **a** Rebek's self-replicating system, in which components **2** and **3** react to form template **1**. The reaction can also proceed through a [**2•3**] binary complex, affording product ***cis*-1**, which can isomerise to give the more stable template ***trans*-1**. **b** Five pathways identified by Reinhoudt and co-workers for Rebek's self-replicating system shown in **a**. Recognition-mediated complex formation is denoted by square brackets and reaction between components by +. **c** A Diels-Alder reaction based replicator reported by Wang and Sutherland. Reaction between maleimide **4** and diene **5** produces template **6**, capable of templating its own formation ***via*** the ternary catalytic complex [**4•5•6**], mediated by hydrogen-bonding recognition. **d** A Diels-Alder replicating system inspired by the Wang and Sutherland replicator reported by von Kiedrowski and co-workers. The reaction components, maleimide **7** and diene **8** assemble with the template **9** in a ternary catalytically active complex [**7•8•9**] ***via*** hydrogen-bonding mediated recognition**. e** Replicator design exploiting the reaction between maleimide **7a** and nitrone **10**. The reaction of these components produces two diastereoisomeric products ***trans*** and ***cis*** but the ***trans*** diastereoisomer of template **11** only is capable of templating its own formation ***via*** [**7a•10•11**] complex. **f** Reaction between nitrone **13** and maleimide **12** produces a highly diastereoselective self-replicator ***trans*-14** (green circles in the concentration ***vs.*** time profile) capable of amplifying itself at the expense of the ***cis*** cycloadduct (ratio of [***trans***]/[***cis***] > 125). The rate profiles for ***trans*-14** and ***cis*-14** are shown in black circles and white squares, respectively.

The Rebek laboratory described[46] the first example of a small molecule-based self-replicator in 1990. This system (**Fig. 8a**) exploited the Et_3N-catalysed formation of an amide bond as the strategy for the formation of template **1** from an adenine derivative **2** and an imide of Kemp's triacid[47] **3**. Rebek's design, however, did not exhibit a sigmoidal reaction profile—a finding that was attributed to the binary reactive complex pathway, mediated by [**2•3**], being more efficient than the pathway mediated by the ternary complex [**2•3•1**]. Reaction within [**2•3**] produced template *cis*-**1**, which could isomerise to give template *trans*-**1**. Rebek and co-workers demonstrated that the formation of template **1** is recognition-mediated, observing a drop in the reaction rate when **3** was reacted with a recognition-disabled, *N*-methylated version of **2**, or in the

presence of a competitive inhibitor (2,6-bis(acylamino) pyridine). In 1994, Menger and co-workers published[48] a study demonstrating that the formation of **1** is catalysed by the addition of simple amides, thus raising doubts as to the self-replicating nature of Rebek's system. Following a lengthy debate, finally the Reinhoudt laboratory provided[49] evidence in 1996 that resolved the argument between Rebek and Menger. Through a full kinetic analysis of the system, the authors were able to identify five different pathways (**Fig. 8b**) that contributed to the formation of **1**. The full kinetic analysis revealed that the performances of the various reaction pathways are, in fact, strongly dependent on the reaction concentration employed. Specifically, higher concentrations increase the contribution of the bimolecular pathway (I) at the expense of pathway II mediated by the binary complex [**2•3**]. At such concentrations, the replication facilitated by the ternary complex [**2•3•1**] (III) contributes 46% at most, and only if preformed template is added at the beginning of the reaction. The analysis also revealed that the contribution of pathway IV was more significant at higher concentrations (*i.e.* amide catalysis implicated by Menger and co-workers needs to be considered in higher concentration regimes), whereas pathway V was found to be relatively insignificant. Ultimately, the formation of **1** was found to proceed primarily through pathway II. Even before these results were reported, the Rebek laboratory re-engineered[50] their system by changing the naphthyl spacer to a biphenyl linker, thereby increasing the efficiency of replication considerably. The adenine-based replicating systems[46,50,51] were complemented by reports[52] of thymine-derivative incorporating replicators, as well as mixed systems by Rebek and co-workers.

In 1997, Wang and Sutherland reported[53] the design and experimental implementation (**Fig. 8c**) of a self-replicating system based on the Diels-Alder reaction between a maleimide **4**, acting as the 2π component, and cyclohexadiene **5**, acting as the 4π component. The reaction of these components in CD_2Cl_2 to form template **6** exhibited a sigmoidal reaction profile. Self-replication within this system was confirmed by examining the reaction in the absence and in the presence of preformed template **6**. While the authors undertook analysis of the kinetic and thermodynamic processes governing the system, no discussion of the stereochemical features of the self-replicator was provided. Specifically, both diene **5** and template **6** are chiral, and the reaction of **4** with **5** can result in four different diastereoisomers (two *endo* and two *exo*). The authors assigned the product observed experimentally as *endo*-**6**, despite providing no analytical evidence to support this assignment.

The possibility of homochiral and heterochiral self-replication presented[53] by the Wang and Sutherland replicator inspired von Kiedrowski and co-workers to undertake[54] a significantly more detailed mechanistic and stereochemical study on similar replicating system. The authors replaced the heterocyclic recognition sites on the original reaction components with an amidopyridine and a carboxylic acid (**Fig. 8d**), first reported[55] by Hamilton, to give maleimide **7** and diene **8**. Initially, the authors examined the reaction between *rac*-**8** and **7a**, as well as its methyl-substituted variant **7b**. The reaction profiles for the formation of *rac*-**9a** and *rac*-**9b** products exhibited a lag period, which shortened dramatically in the presence of the corresponding preformed racemic template. Through kinetic fitting of the experimental NMR data, the authors were able to establish that their replicator design retained its replication efficiency, which was similar to that reported by Wang and Sutherland. Through comprehensive computational analyses, the authors were able to rationalise the near exponential growth by the conformational constraints present in the product duplexes. Subsequently, the authors examined the homo- and heterochiral reaction pathways, namely the reactions of each enantiomer of the diene, *i.e.* *R*-**8** and *S*-**8**, with maleimide **7a** in the absence of instructional template, followed by analysis in the presence of enantiopure template *R*-**9a**. The kinetic results revealed that the enantiopure template *R*-**9a** exerts a similar catalytic effect on both pathways—reactions of **7a** with *R*-**8** and **7a** with *S*-**8**, thus confirming that both homo- and heterochiral catalytic pathways are effective in the system. In later work, von Kiedrowski and co-workers examined[56] a replicating system based on a fulvene framework. By investigating the catalytic relationships between the four products (two *endo* and two *exo*), the authors were able to demonstrate the strong sensitivity of the system to small changes in the disposition of recognition elements.

The Diels-Alder reaction has also been exploited[57] as the ligation step in recognition-mediated reactions and replicating systems by the Philp laboratory. In this work, Philp and co-workers utilised the Diels-Alder reaction in the design of two structurally similar furan- and maleimide-based families of replicators as platforms for investigating the effect of structural variation on the efficiency of replication and other recognition-mediated channels in each system. Comprehensive kinetic analyses[57b,57c] of the various reactions showed that a system with a high degree of conformational freedom is more likely to react through the binary complex pathway preferentially. As a consequence, the development of a highly efficient self-replicating system typically necessitates a certain degree of rigidity as well as a suitable disposition of recognition sites in space—a requirement for the formation of catalytically active ternary complexes.

In addition to replicating systems that exploit the Diels-Alder reaction, Philp and co-workers have pioneered the use of 1,3-dipolar cycloaddition reactions as the ligation step in the construction of a number of minimal replicating systems. Following their initial work, which exploited[58] the reaction between a maleimide as the dipolarophile and an azide as the 1,3-dipole, Philp and co-workers have developed[59,60] replicating systems employing a nitrone as the 1,3-dipole. The reaction of a nitrone with a maleimide results in the formation of two diastereoisomeric products—normally given the descriptors[§] *trans* and *cis*. These two cycloadducts possess significantly different geometries and the replicating systems employing this 1,3-dipolar cycloaddition reaction therefore provide an excellent platform for investigating the transfer of stereochemical information.

In 2001, Philp and co-workers reported[59] the 1,3-dipolar cycloaddition reaction of maleimide **7a** with nitrone **10** (**Fig.**

8e). Molecular recognition in the system was provided by the association of a carboxylic acid moiety with the 6-methyl amidopyridine group. In this system, the *trans* diastereoisomer of the template **11** possesses the conformation necessary for successful docking of the two building blocks **7a** and **10** to form the appropriate ternary catalytic complex [**7a**•**10**•*trans*-**11**], required to drive the self-replication cycle. Initially, the authors examined the reaction of nitrone **10** with the recognition-disabled methyl ester of maleimide **7a**. In the absence of recognition, this reaction afforded the corresponding recognition-disabled analogues of *trans*-**11** and *cis*-**11** cycloadducts in a ratio of 4:1. By contrast, the reaction between nitrone **10** and maleimide **7a**, where recognition is active, proceeded much more efficiently, exhibiting a sigmoidal reaction profile, and the ratio of the two diastereoisomeric products (trans and cis) was now ~6:1. When this reaction was repeated in the presence of preformed *trans*-**11**, the lag period observed previously disappeared and the diastereoselectivity for the *trans* product increased to 9:1. However, addition of *cis*-**11** had no effect on the reaction. Through these kinetic experiments, the authors confirmed that the self-replicating template, *trans*-**11**, is capable of transmitting structural information successfully *via* the ternary complex pathway. While this nitrone-based self-replicating design performed efficiently, only modest increases in both the reaction rate and the diastereoisomeric ratio was achieved.

Exploiting the information available from the structure–reactivity studies[57b,57c] and computational modelling, Philp and Kassianidis designed[60] a structurally optimised replicator, incorporating modifications in the design of both the nitrone and the maleimide elements, intended to disfavour the reactivity *via* the binary complex pathway. This re-engineered replicator (**Fig. 8f**) was formed by the reaction of a more extended phenylacetic acid maleimide **12** with nitrone **13**. These structural changes facilitated the formation of a significantly more stereoselective replicator *trans*-**14** ([*trans*]/[*cis*] = 115), even in the absence of preformed template. The system could be further biased towards the formation of the *trans* diastereoisomer by the addition of preformed template of *trans*-**14** or by reduction of the reaction concentration.

The numerous examples of minimal self-replicating templates based on oligonucleotides, peptides, and small molecules described in this section demonstrate unmistakably that self-replication is not limited to complex biological systems as we know them today. These examples, whilst minimal in their operation and design, represent a key step in the understanding and examination of replication processes under more complex scenarios and environments, which will be examined next.

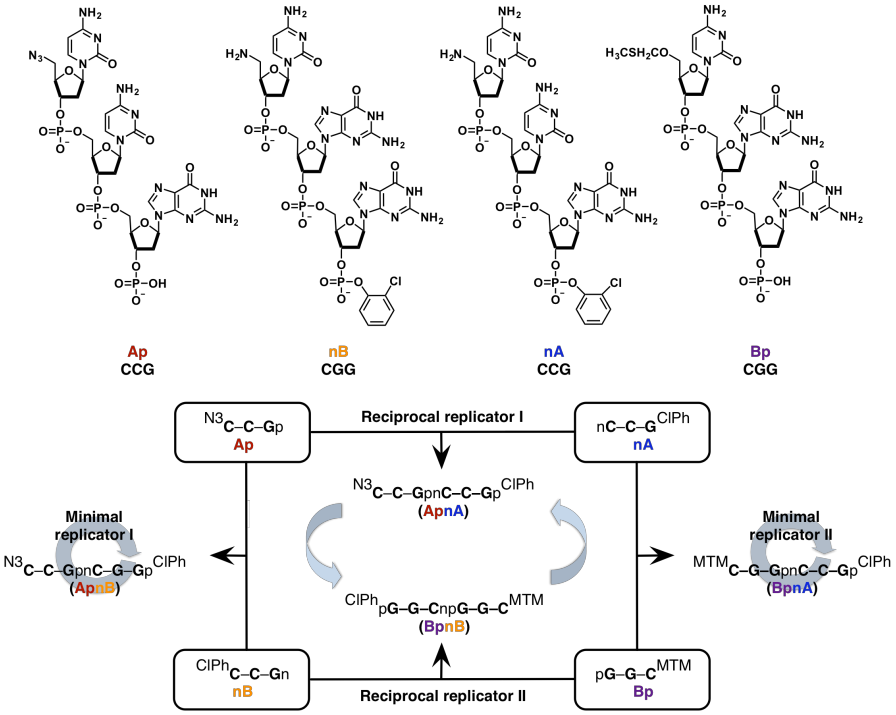


Fig. 9 Multicyclic replicating network composed of four trinucleotides: **Ap** (**CCG**), **nB** (**CGG**), **nA** (**CCG**) and **Bp** (**CGG**). The components can react to form two self-replicating templates (**ApnB** and **BpnA**) and two reciprocally replicating templates (**ApnA** and **BpnB**). The letters **n** and **p** correspond to the presence of a free amine group at the 5′ end and free phosphate group at the 3′ end, respectively. MTM = methylthiomethyl ester; ClPh = ***o***-chlorophenyl.

Networks of synthetic replicators

Building on the design of their minimal replicating system, the von Kiedrowski laboratory has extended[61] their minimal oligonucleotide-based replicators[20,62] into a multicyclic system (**Fig. 9**), capable of both autocatalysis and crosscatalysis. The

network was composed of four DNA-based trinucleotides: **Ap** (**CCG**), **nB** (**CGG**), **nA** (**CCG**) and **Bp** (**CGG**). Trinucleotides **Ap** and **Bp** were equipped with a reactive electrophilic phosphate group (denoted with **p**) at the 3' end, while the 5' prime ends of these two components were protected as an azide and methylthiomethyl ester, respectively, in order to prevent any self-condensation reactions from occurring. Trinucleotides **nB** and **nA** incorporated a free nucleophilic amine (marked as **n**) at the 5' end. In the cases of **Ap** and **Bp**, their 3' phosphates were protected using an *o*-chlorophenyl group in each case in order to prevent any unwanted ligation reactions. When these compounds were allowed to react together, they could form four templates. Specifically, the condensation of **Ap** and **nB** results in the formation of a self-complementary autocatalytic template **ApnB** (**CCGCGG**). Similarly, the reaction of **nA** with **Bp** affords another self-replicating template **BpnA** (**GGCCG**). The hexanucleotide templates, produced by the reaction of **Ap** with **nA** (**CCGCCG**) and **Bp** with **nB** (**GGCGGC**) are capable of reciprocal replication only.

In this multicyclic system, formation of template molecules from the individual components assembled in ternary complexes proceeded by an attack of the 5' amine on the 3' phosphate, at a comparable rate for all templates. Recognition in all four ternary complexes was mediated by Watson-Crick base pairing, and as a result of the similarities in the strength of recognition, the concentration reactions leading to the four templates proceeded with similar efficiencies in the absence of instruction. Subsequently, the formation of the native templates was examined in the presence of preformed templates, which incorporated a phosphodiester bond (**n**) instead of the phosphoramidate linkage (**np**) found in the templates formed through the reaction of building blocks **Ap**, **nB**, **nA** and **Bp**. These experiments revealed, that in each case, the addition of preformed template resulted in an increase in the concentration of the corresponding complementary strand possessing the native **np** linkage (*e.g.* addition of **AnB** would result in the amplification of self-replicator **ApnB**, whereas the addition of **ApA** would enhance the formation of reciprocal replicator **BpnB**). Interestingly, in each of these instructed experiments, small but noticeable up-regulation of the corresponding partner replicator (*e.g.* **BpnA** to **ApnB**) was also observed. The fact that instructing the network with one template resulted in increased concentration of another replicator (self- or reciprocal replicator) can be viewed as a system-level property, which stems from the interactions in the system.

Kim and Joyce have, likewise, expanded[63] the minimal R3C ligase self-replicating system into a network (**Fig. 10**) operating through reciprocal mechanism, where two ribozymes, **E** and **E'**, template the formation of each other from the corresponding subunits: **A**, **B**, **A'** and **B'**. While the ribozymes contain complementary recognition and catalytic elements, their sequences are not identical. The loss of self-complementarity prevents the association of the substrate subunits, which hindered replication in the original self-replicating system described earlier. In this system, therefore, the ribozymes catalyse their syntheses *via* the assembly of the appropriate reaction components in ternary catalytically active complexes [**A'•B'•E**] and [**A•B•E'**].

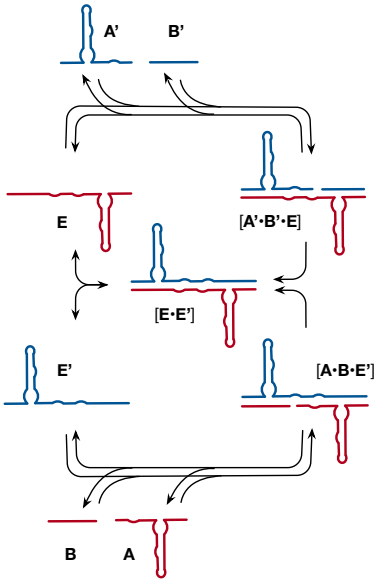


Fig. 10 Crosscatalytic replication network comprised of two R3C ligase ribozymes. Ribozyme **E'** (blue) is created from components **A'** and **B'**, and catalysed the reaction between **A** and **B** to give **E** (red). Similarly, product **E** mediates the formation of **E'** from **A'** and **B'**. These reciprocal pathways are mediated by the formation of the catalytically active ternary complexes [**A•B•E'**] and [**A'•B'•E**].

Kinetic analyses of the individual crosscatalytic pathways demonstrated that both templates catalyse the formation of the reciprocal replicator as expected. Interestingly, the results also confirmed the formation of two chimeric products, where reaction of **A'** and **B** was catalysed on **E**', and ligation of **A** and **B'** on **E**, which suggests a certain degree of tolerance of the ribozyme design to the presence of mismatches. The reactions leading to the two chimeric products, which could potentially self-replicate, were, however, found to be considerably slower than the two target reciprocal pathways.

Analysis of the reciprocal system as a whole (all four components) revealed that in the absence of added template **E** or **E'**, the chimeric product formed by the reaction of **A'** and **B** dominates the product pool. Nevertheless, in the presence of preformed template of **E** or **E'**, the formation of the corresponding target crosscatalytic product was confirmed. In each case, the addition of preformed template resulted also in increased formation of the ribozyme employed as instruction.

Subsequently, this work was extended further by Lincoln and Joyce, who utilised[64] *in vitro* evolution to create an optimised crosscatalytic pair **E** and **E'**, capable of exponential amplification. The authors demonstrated that efficient replication could be sustained indefinitely when a small portion (4%) of the reaction mixture was transferred after 5 h into a fresh batch of reactants. By introducing mutations into the structures of the ribozyme pair, the authors created 12 pairs of reciprocal ribozymes, **E1–E1'** to **E12–E12'** (each pair was constructed from four substrates, *e.g.* **E12–E12'** from **A12**, **A'12**, **B12** and **B'12**; 48 substrates overall). Out of these pairs,

pair **E1–E1'** exhibited the highest efficiency of replication—20-fold amplification in just 5 h. Lincoln and Joyce performed a serial transfer experiment, starting from with all 12 pairs of enzymes (each at 0.1 μM) and all 48 substrates (5 μM each). This mixture could produce 132 possible pairs of recombinant enzymes and 12 non-recombinant pairs (*e.g.* **E3** with **E3'**). Over 100 h, the authors performed 20 successive transfers (5% of reaction mixture in each case), which should exert a selection pressure on the system, allowing progressive elimination of reaction members that replicate inefficiently. Analysis of the 100 clones obtained from the final reaction mixture revealed that seven of them were non-recombinant pairs, and the most abundant products were the recombinant templates formed by the reaction of **A5** with **B2**, **A5** with **B3** and **A5** with **B4** (and the corresponding partner in each case). Subsequently, the Joyce laboratory has reported[65] a comprehensive analysis of the kinetic properties of the **E** and **E'** ribozymes and utilised[66] it to improve the design of the self-replicating ribozyme. Joyce and co-workers have also applied their expertise to the design and implementation[67] of ligand-dependent RNA enzyme replicators (aptazymes) and replicators based on L-RNA.

Both the Ghadiri and Chmielewski laboratories have been extremely successful in developing and incorporating their individual self-replicating peptides into more complex networks where multiple catalytic pathways operate simultaneously. Chmielewski and co-workers have combined[68] the two self-replicating peptides that can be modulated through environmental conditions, **E1E2**[31] (pH) and **K1K2**[32] (ionic strength), in to a single system where both auto- and crosscatalytic processes can operate. The expanded peptide network was assembled from four shorter peptide fragments, **E1**, **E2**, **K1** and **K2**, which permitted formation of the native[31,32] peptide templates, **E1E2** and **K1K2**, and two recombinant proteins, **E1K2** and **K1E2**. These mixed templates are capable of self-associating *via* anti-parallel coiled-coils and capable of associating with each other *via* formation of parallel coiled-coils. Addition of individual peptide templates to the reaction mixture consisting of the four fragments, at pH 7.5, allowed the authors to identify several unexpected, crosscatalytic pathways. Under neutral conditions, **E1K2** template is produced most rapidly. Despite the preference of the system for the production of **E1K2**, the authors were able to amplify **E1E2** selectively by lowering the pH of the reaction mixture to 4. Similarly, by undertaking the reaction under high salt conditions at neutral pH, the authors were able to direct the system towards enhanced production of **K1K2**. Using this framework, the authors demonstrated successfully that production of a particular peptide replicator can be amplified selectively from a mixture of reactive components by careful modulation of the reaction environment, such as the pH and salt concentration, providing support for the potential role of proteins in the origin of life.

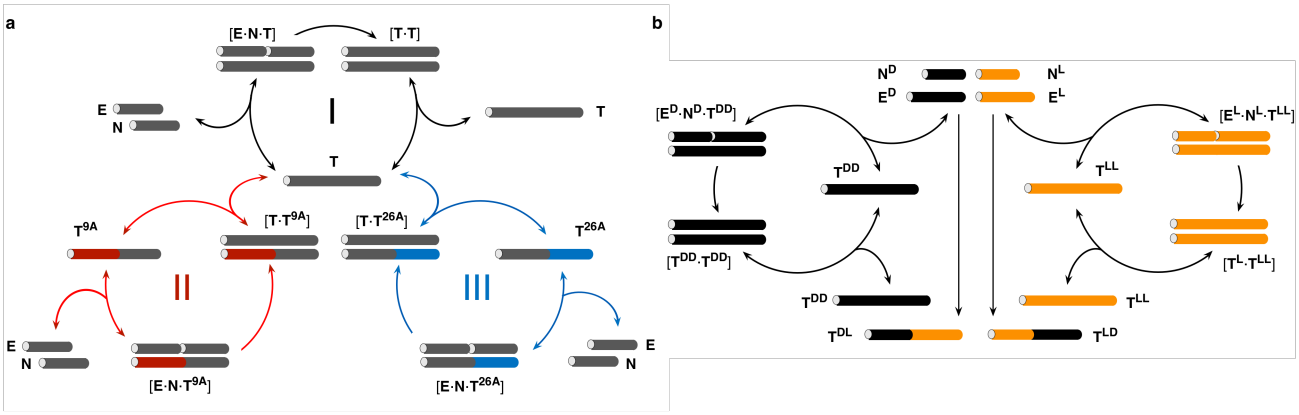


Fig. 11 **a** Schematic representation of an error-correcting, autocratic peptide network. A mixture of peptide fragments **E** and **N**, and their single-alanine mutants, $\mathbf{E^{9A}}$ and $\mathbf{N^{26A}}$, (simulating spontaneous generation of errors) results in a wild type template **T** (grey cylinder) and single mutation containing templates $\mathbf{T^{9A}}$ (red cylinder) and $\mathbf{T^{26A}}$ (blue cylinder). The self-organised network amplifies the template **T** selectively by subjugation of the mutant templates for the production of **T**. The double mutant $\mathbf{T^{9A/26A}}$ is not shown as it was determined to be catalytically inactive. **b** Schematic representation of stereospecific peptide replicators. The electrophilic fragments, $\mathbf{E^L}$ and $\mathbf{E^D}$, and nucleophilic fragments, $\mathbf{N^L}$ and $\mathbf{N^D}$, combine to form four templates. The homochiral templates $\mathbf{T^{LL}}$ and $\mathbf{T^{DD}}$ are capable of autocatalysis, while the heterochiral templates are formed through uncatalysed bimolecular reactions only. Black and orange cylinders represent peptide regions comprised of D- and L-amino acids, respectively.

Following the extension of their initial replicator design[27,30] into a symbiotic peptide network where two replicators co-existed[69] in a mutually symbiotic relationship, Ghadiri and co-workers designed[70] a dynamic peptide network capable of error-correction (**Fig. 11a**). Selective amplification of a single peptide sequence within this simultaneously auto- and crosscatalytic system was achieved by subjugation of the mutant peptides for the synthesis of the wild type peptide, **T** (**Fig. 11a**, grey). Slow spontaneous generation of errors/mutants, as observed in biological systems, was simulated by production of closely related mutant peptides through bimolecular reaction of smaller fragments incorporating mutations. In addition to the native electrophilic and nucleophilic fragments, **E** and **N** (**Fig. 11a**, grey), the network included their single alanine mutants, $\mathbf{E^{9A}}$ (**Fig. 11a**, red) and $\mathbf{N^{2A}}$ (**Fig. 11a**, blue). Reaction of these fragments affords four different peptide templates, the native **T**, $\mathbf{T^{9A}}$ and $\mathbf{T^{26A}}$ with a single mutated residue and double mutant $\mathbf{T^{9A/26A}}$.

At neutral pH, the reaction system exhibits a strong preference for the formation of mutation-free **T**. The double mutant $\mathbf{T}^{9A/26A}$ was shown to be completely inactive catalytically, whereas the two templates incorporating a single "error" displayed some crosscatalytic activity only and this activity was directed towards enhanced formation of native replicator **T**. The template **T**, however, was found to be a selfish autocatalyst, incapable of acting as a crosscatalytic platform for the formation of the mutated templates. Therefore, the autocatalytic cycle producing **T** works in concert with the two crosscatalytic pathways to achieve selective production of error-free **T**. Within this peptide network, the authors have demonstrated an example of a system exhibiting two complex phenomena simultaneously, namely error-correction and sequence-specific replication, with potential significance in stabilisation of the genotype of self-replicating molecules.

The biological world is overwhelmingly a homochiral reaction space, and yet, the origins[71] of this homochirality have yet to be established and are a source of significant on-going debate. In order to explore the possible role of peptide replicators in this process, Ghadiri and co-workers have designed[72] a network of stereospecific replicating peptides (**Fig. 11b**). The extension of their original peptide replicator **T** to a system composed of two enantiomeric electrophilic fragments, $\mathbf{E}^{L}$ and $\mathbf{E}^{D}$, and two enantiomeric nucleophilic components $\mathbf{N}^{L}$ and $\mathbf{N}^{D}$ provided a platform for these investigations. Reactions of fragments with the same stereochemistry afford homochiral templates $\mathbf{T}^{LL}$ (**Fig. 11b**, orange) and $\mathbf{T}^{DD}$ (**Fig. 11b**, black), whereas reaction of mixed fragments creates heterochiral templates $\mathbf{T}^{LD}$ and $\mathbf{T}^{DL}$. Analyses of the individual pathways leading to the two homochiral and two heterochiral products revealed that although $\mathbf{T}^{LL}$ and $\mathbf{T}^{DD}$ are formed from their constituent components efficiently, the heterochiral peptide templates, $\mathbf{T}^{LD}$ and $\mathbf{T}^{DL}$, are only formed through template-independent pathways. The authors propose that this lack of template-mediated activity in the heterochiral products stems from the diminished ability of these two templates to form coiled-coil helical assemblies. Detailed kinetic analyses and template-instructed experiments revealed that $\mathbf{T}^{LL}$ is capable of stereospecific self-replication, as addition of homochiral $\mathbf{T}^{DD}$ or the heterochiral templates exerted no influence on its formation. The ability of the homochiral template $\mathbf{T}^{DD}$, formed from unnatural D-amino acids only, to self-replicate was not discussed in the study. The authors also found that the formation of the heterochiral product $\mathbf{T}^{LD}$ from its fragments proceeds at the same rate in the presence of instructing template (all four templates were examined) as it does its absence. The stereoselective system exhibited strong sensitivity to changes in even a single amino acid residue, which permitted the amplification of a single homochiral template, once produced. In particular, while $\mathbf{T}^{LL}$ peptides containing a single mutation showed little autocatalytic aptitude, they were able to accelerate the formation of the homochiral $\mathbf{T}^{LL}$ strongly through crosscatalytic pathways. In this study, the authors demonstrated, for the first time, the feasibility of stereospecific replication. Taken into consideration the viability of stereospecific replication together with the diminished ability of chiral peptides incorporating mutations in their primary sequence to self-replicate, these results suggest that a replicating peptide biopolymer might have played[73] a role in the origin of biological homochirality.

The design of peptide replicators continued to evolve, exploring larger and more interconnected systems. In 2004, Ghadiri and co-workers described[74] a bottom-up approach to designing a peptide network composed of 81 structurally similar 32-residue coiled-coil peptides. Analysis of the numerous peptides relied initially on prediction methods in order to estimate the relative stability of all substrate-template complexes available to the system. The authors analysed the stability of each complex in order to predict the auto- (**Fig. 12**, red arrows) and crosscatalytic pathways (**Fig. 12**, black arrows), as well as the overall network topology (**Fig. 12**). The design principles employed were tested experimentally on a smaller, 9-node subsystem (**Fig. 12**, dark grey) within the network, which showed a good agreement with the predictions made using graph analysis. Ghadiri and co-workers also demonstrated that the efficiency of certain network pathways can be selectively modulated by employing various chemical triggers, *i.e.* different instructing templates. The smaller 9-peptide system was built from a single nucleophilic component **N** and nine different electrophilic peptide fragments, $\mathbf{E}^{1}$ to $\mathbf{E}^{9}$. The authors exploited substitutions at four key electrophilic residues, located at positions *e* and *g* within the peptide sequence (**Fig. 12**, represented by the four-letter code), in order to design peptides with varied ability to form aggregates, and, thus, also different catalytic efficiencies. Theoretical analysis predicted that twenty crosscatalytic pathways and three autocatalytic pathways are plausible in this network. Experimental analysis of a mixture containing all of the electrophilic fragments with a sub-stoichiometric amount of **N** revealed that all nine possible products are formed, with $\mathbf{T}^{1}$, $\mathbf{T}^{2}$, $\mathbf{T}^{4}$, $\mathbf{T}^{7}$ and $\mathbf{T}^{8}$ reaching the highest concentrations. Undertaking comprehensive kinetic analyses of the individual reaction pathways as well as the network as a whole demonstrated that the rate of formation of the templates examined was noticeably different in isolation relative to their rates when the full network was examined. These differences highlight the potential of system-level properties manifesting themselves in networks of interconnected components that are not observed when the components are examined in isolation.

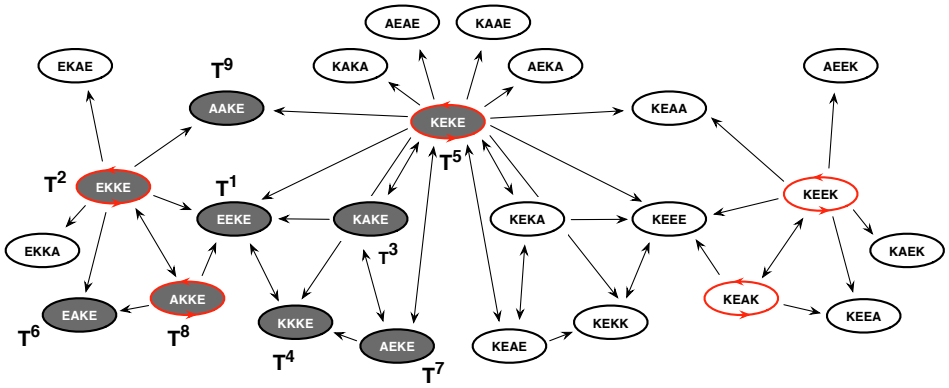


Fig. 12 Graph representation of a directed self-organised peptide network established from an array of 81 structurally similar coiled-coil sequences (32 residues each) ***in silico***. The network is comprised of 25 nodes (molecular species) and joined by 53 edges (arrows), representing auto- (red circular arrows) and crosscatalytic (black arrows, where the arrow direction indicates catalytic relationship) processes. Sub-network of nine nodes, **T**[1] to **T**[9], highlighted in dark grey, was investigated experimentally, showing a good agreement with the predicted results. The four letters within each node (template) represent the identity of the amino acids within the peptide sequence at residues 8, 13, 15 and 20 (positions ***e*** and ***g***).

Ghadiri and Ashkenasy examined[75] a sub-network of this smaller 9-node system (dark grey, **Fig. 12**), demonstrating its capacity to perform basic Boolean logic functions, such as OR and NOR (neither X nor Y), when instructed with chemical inputs (preformed templates)—control operations not unlike those observed[76] in complex biological systems. The system (**Fig. 13**) is comprised from five nodes (peptide templates **T**[1], **T**[3], **T**[4], **T**[5] and **T**[7]) that are interconnected through 15 edges, representing both auto- and crosscatalytic pathways. As an example, the NOR logic operation (highlighted in red, **Fig. 13**) is expressed by the autocatalytic formation of **T**[3]. In this Boolean function, template **T**[3] is formed efficiently only in the absence of both the **E**[5] and **E**[7] fragments. If either of these fragments is present, however, formation of **T**[3] is diminished as any quantity of this template that is produced is utilised primarily as a crosscatalytic template for the formation of **T**[5] and/or **T**[7].

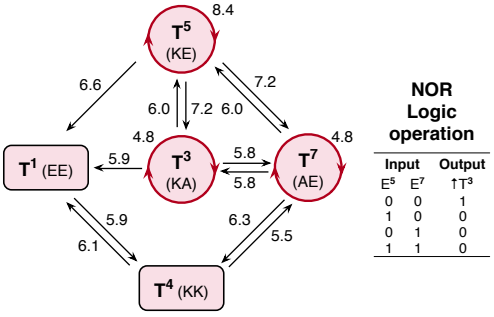


Fig. 13 A directed, weighted peptide network capable of performing logic operations. Five peptide templates are interconnected through 15 auto- and crosscatalytic pathways. The sub-network (red) of three templates, **T**[3], **T**[5] and **T**[7] expresses NOR logic operation, where self-replication of **T**[3] proceeds efficiently only in the absence of both electrophilic fragments **E**[5] and **E**[7].

Networks based on auto- and crosscatalytic templates constructed from small synthetic molecules have also continued to advance in complexity. Both the Rebek[51a,51c,52c] and Philp[77,78] laboratories have exploited their replicators for the fabrication of interconnected networks incorporating more than a single replicator. For example, Philp and co-workers have combined[57a,77] a 3-furan based replicator[57c] system with the efficient self-replicating system shown[60] in **Fig. 8f** to construct a network comprised of four building blocks: **12**, **13**, **15** and **16** (**Fig. 14**). Reaction of these components affords three established independent self-replicators *trans*-**14**, *endo*-**17** and *exo*-**17**. Alternatively, combination of the reaction components can also produce two mutually complementary templates, *trans*-**18** and *exo*-**19**, capable of taking part in reciprocal replication. Examination of the reaction of all four components revealed that the system makes *trans*-**14** preferentially (accounting for over 70% of the product pool).

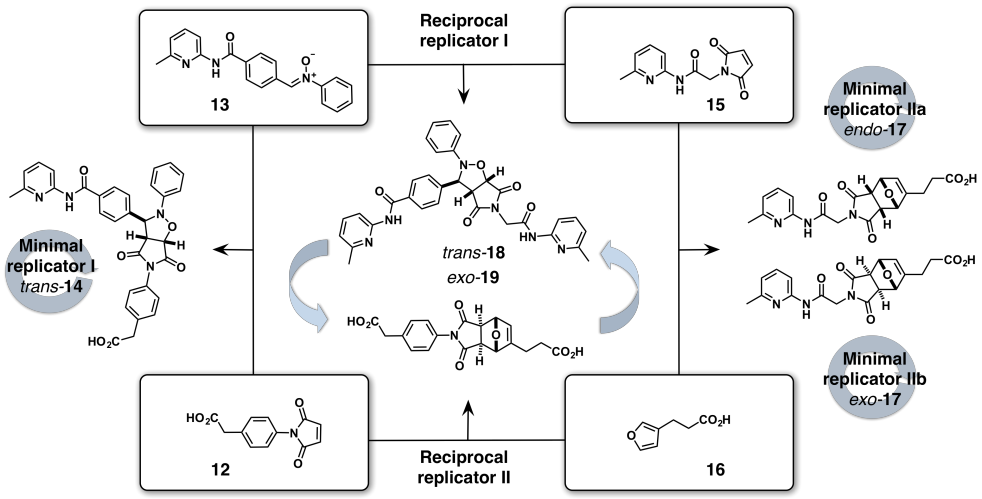


Fig. 14 **a** A multicyclic replication network built from two maleimides, **12** and **15**, nitrone **13** and a furan **16**. Reaction of these components produces three minimal self-replicating systems ***trans*-14**, ***endo*-17** and ***exo*-17** and two reciprocal products, ***trans*-18** and ***exo*-19**. Strong preference of the system for the formation of ***trans*-14** imposed a limit on the instructability of the system with preformed templates. Cycloadducts exhibiting no recognition-mediated activity are omitted for clarity.

Initially, the authors envisaged that the network could be directed towards increased production of a specific product by addition of preformed templates. However, the strong bias of the system for *trans*-**14** as a result of the uneven replication efficiencies resulted in a network that is, in fact, somewhat insensitive to the effects of instructional templates. Using kinetic simulations, the authors investigated[77] a number of conditions, varying the catalytic efficiencies for the auto- and crosscatalytic replicators, as well as the amount of preformed template used as input. Interestingly, the simulations demonstrated that the reciprocal replicators are more responsive to instruction, when compared to the autocatalytic templates—most likely as a result of the mutually reinforcing nature of the reciprocal replication processes.

Recently, we reported[78] a network where the outcome of competition between two small-molecule replicators is determined by a unidirectional crosscatalytic relationship. The two replicators described in this work (**Figure 16a**), referred to as $\mathbf{T^p}$ (a variation of replicator *trans*-**14** shown in **Fig. 8f**) and $\mathbf{T^m}$, are formed by 1,3-dipolar cycloaddition reactions of nitrone **N**, equipped with a 6-methylamidopyridne site, with two maleimides that differ in the location of their carboxylic acid site relative to the maleimide ring—*para* for maleimide $\mathbf{M^p}$ and *meta* for maleimide $\mathbf{M^m}$. The authors perform a comprehensive set of kinetic experiments, demonstrating that, in isolation, $\mathbf{T^p}$ is considerably more efficient at templating its own formation $\mathbf{T^m}$—and outcome that is attributed to the higher efficiency of the pseudounimolecular reaction within the corresponding ternary complex and lower homoduplex association constant of $\mathbf{T^p}$ in comparison to $\mathbf{T^m}$. Interestingly, when replicators $\mathbf{T^p}$ and $\mathbf{T^m}$ are examined in a scenario where they have to compete for a limited quantity of the shared building block **N** (**Figure 16b**), replicator $\mathbf{T^m}$ outcompetes $\mathbf{T^p}$. Using a combination of kinetic analysis, fitting, simulations, and modelling, the authors have traced this system-level outcome to the critical crosscatalytic pathway that operates in this network, which allows $\mathbf{T^m}$ to exploit $\mathbf{T^p}$ as a template for its formation but not *vice versa*. As a result of the small differences in the recognition sites present in the two templates, the template duplex formed by replicator $\mathbf{T^m}$ is stronger than that formed by $\mathbf{T^p}$, thus allowing $\mathbf{T^m}$ to sequester $\mathbf{T^p}$ in a heteroduplex $[\mathbf{T^p \bullet T^m}]$ that is more stable than the $[\mathbf{T^p \bullet T^p}]$ homoduplex. Attempts to instruct the network of two replicators to make one replicator with increased preference revealed that the initial imbalance between $\mathbf{T^p}$ and $\mathbf{T^m}$ achieved by the addition of preformed template is eroded over time (**Figure 16b**) as a result of the reaction environment employed—*i.e.* well-stirred batch reactor (WSBR) format. This decrease in replicator ratio over time is directly related to the progressive exhaustion of building blocks, which limits the efficiency of the replication processes operating in the system. Importantly, this work highlights that the outcome of competition in a network of replicators depends not only on the catalytic relationships inherent to the system but also on the reaction environment in which it is examined.

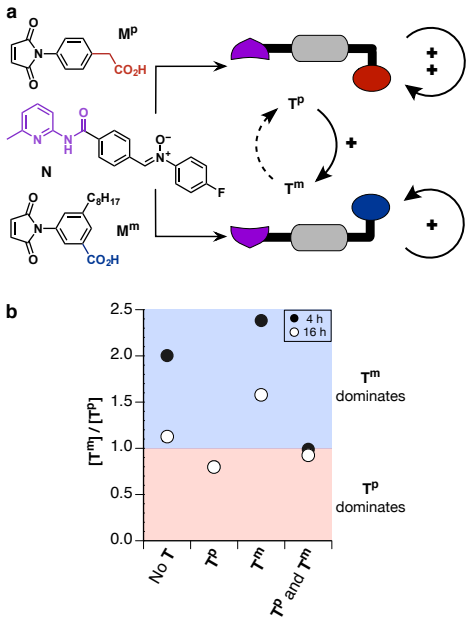


Fig. 15 **a** Cartoon and structural representation of a network of two replicators, $\mathbf{T^P}$ and $\mathbf{T^m}$, created by 1,3-dipolar cycloaddition reactions of nitrone **N** with maleimides $\mathbf{M^P}$ and $\mathbf{M^m}$ (red/blue recognition sites on the maleimides are complementary with the recognition site on nitrone, shown in purple). **b** Effect of adding preformed template (added at t = 0) on the ratio of $[\mathbf{T^m}]/[\mathbf{T^P}]$ determined after 4 h (black circles) and 16 h (white circles). Concentrations were determined using 470.3 MHz $^{19}F\{^1H\}$ NMR spectroscopy ($[\mathbf{M^P}] = [\mathbf{M^m}] = [\mathbf{N}]$ = 5 mM, if present, [template] = 1 mM, CDCl3, 5 °C). Figure adapted with permission from Ref. 78. Copyright, 2017 American Chemical Society.

Aided by computational methods and kinetic simulations, the investigations of self-replicating systems have progressed dramatically over the last 30 years, from the first examples of minimal, often inefficient, replicating systems to significantly more varied and interconnected networks that explore the interplay between various recognition and reaction processes taking place in parallel. The experimental implementations of self- and reciprocal replicating systems complement the theoretical models, leading to a better understanding of the principles governing the reactivity and information transfer in synthetic replicating systems. The interconnectedness of the components in these systems endows them with the capacity to respond to stimuli, allowing complex function to emerge. In particular, in addition to expressing the capacity for replication, the resulting chemical systems can also express other functionalities such as Boolean logic operations,[36,75,79] error-correction,[70] stereospecific[72] replication and the replication[80] of mechanically-interlocked molecules. Understanding the system-level behaviour in interconnected networks and, thus, the possibility of harnessing these complex replicating systems in the construction, selection and amplification of higher order assemblies, necessitates the understanding of the kinetic and thermodynamic forces driving the recognition and reaction processes between the individual components in each system.

Replication phenomena under dynamic conditions

Although Nature presents scientists with an abundance of complex systems, their study is often hampered by the sheer number of elements that comprise these systems and the high density of the connections between the system components. Additionally, there are significant challenges associated with their analysis. Yet, despite these challenges, the desire to understand and study the emergence and function of complex systems is strong. In the past 20 years, dynamic covalent chemistry[3,81] (DCC) has emerged as an efficient and direct protocol for the construction of chemical networks comprised of interconnected components from simple building blocks. This approach is driven by reversible reactions between a set of components that can create a virtual library of products resulting from all permutations of the structural and recognition features present within the building blocks used to form the library. This library, when exposed to a target, should result in the selection or amplification[82] of a specific product in possession of features, inherited from its constituent building blocks, that permit the most optimal interactions with the target.

DCC exploits the reversible covalent bond formation between various building blocks equipped with compatible reactive sites. The benefit afforded by the robustness of reversible bond formation coupled with the general combinatorial approach permits the formation of structurally diverse dynamic covalent libraries (DCLs) of exchanging components whose composition are generally under thermodynamic control. The equilibrium distribution of such libraries can be examined to reveal the most thermodynamically stable distribution of products in the absence of any binding partners (*i.e.* target). As a result of the reversible bond formation inherent in the DCC approach, the dynamic systems created possess inherent capacity for error checking and error correction.

Dynamic covalent libraries can be instructed by the addition of an external stimulus (target), capable of directing the library away from its original thermodynamic distribution towards a new composition. This new distribution of material within the library reflects the new, most thermodynamically stable state of the *entire* system and is a function of the specific interaction between the target and the library members. Ideally, the compounds that are detected in the library at higher concentrations after addition of a target are those capable of engaging in stronger recognition with the added template. In this manner, DCC offers an efficient approach to screening a large number of virtual components (all potential combinations that can be synthesised *in situ* from the library building blocks) and discovering those components that interact with the stimulus the most strongly. In practice, the strength of binding can be determined by comparing the concentrations of library members in the absence of instruction, relative to their concentrations observed when the library is permitted to evolve in the presence of an instructing stimulus.

When it comes to replicators and networks of replicators, the reaction environment is, in fact, a parameter that is only

beginning to be explored in systems chemistry, and the majority of replicating systems, as illustrated by the examples presented thus far, have been examined under the well-established, closed system conditions (*i.e.* WSBR conditions). In this respect, the DCC approach presents an extremely useful tool for the construction of complex networks with an added component of a dynamically-exchanging pool of components—a reaction environment for the study of chemical networks that is one step closer to the dynamic, often heterogeneous environment typical for biological networks that achieve their function from mixtures of precursors. Therefore, the examination of template-mediated processes under dynamic conditions has the potential to further our understanding of the requirements that could have allowed a replicator to exploit a mixture of components for its own synthesis during the processes of chemical evolution.

Building on the significant progress in the coupling[83] of kinetically driven irreversible reaction processes to dynamic covalent systems, attempts at integrating replication processes with the DCC approach have started[7a,81d,84] to appear in the last 10 years, and these will be summarised in this section.

Ashkenasy and co-workers have extended[85] the light-induced replication protocol described[36] in **Fig. 6** to a dynamic peptide system replicating reversibly under thermo-dynamic control (**Fig. 16**). The adapted system permits reversible formation of two peptide products, $\mathbf{R^1}$ and $\mathbf{R^2}$, from electrophilic fragments $\mathbf{E^1}$ (**Fig. 16a**, green), $\mathbf{E^2}$ (**Fig. 16a**, purple) and a shared nucleophilic peptide building block **N** (**Fig. 16a**, grey). The replication in this dynamic network was made reversible by the substitution of the reactive cysteine residue employed previously, with a thioglycolic acid at the N-terminus of the nucleophilic fragment **N**. Ligation reaction between the smaller peptide fragments produced a thioester bond at a central position in each peptide template, allowing for reversible *trans*-thioesterification. Peptide templates $\mathbf{R^1}$ (green) and $\mathbf{R^2}$ (purple) differ only in the nature of the electrophilic fragment, where $\mathbf{R^1}$ contains a glutamate residue and $\mathbf{R^2}$ incorporates a lysine residue at position *e* (13) within the heptad repeat, directly opposite a lysine residue at position *g*. As a consequence of repulsive electrostatic interactions between lysine residues, $\mathbf{R^2}$ is not capable of forming the stable catalytically active intermediates required for efficient self-replication. In the absence of external triggering, therefore, $\mathbf{R^1}$ only is capable of replication through the dimeric template-mediated thioesterification.

In the absence of template, reaction of $\mathbf{E^1}$ with **N** displayed a sigmoidal reaction profile. As expected for a self-replicating system, the rate of template formation was found to correlate with the quantity of preformed template added to the reaction mixture. The efficiency of replication for $\mathbf{R^1}$ was found to be similar to that observed in the original peptide system where replication was not reversible. Interestingly, doping with preformed template $\mathbf{T^1}$, containing the native non-reversible peptide linkage, revealed that this template is capable of crosscatalysing the formation of $\mathbf{R^1}$. Behaviour of the small DCL (**Fig. 16a**) comprised of three building blocks, $\mathbf{E^1}$, $\mathbf{E^2}$ and **N**, was examined in the absence and in the presence of an instructing chemical trigger, *e.g.* template $\mathbf{R^1}$, $\mathbf{T^1}$, $\mathbf{T^{1Nv}}$ or $\mathbf{T^{2C}}$.

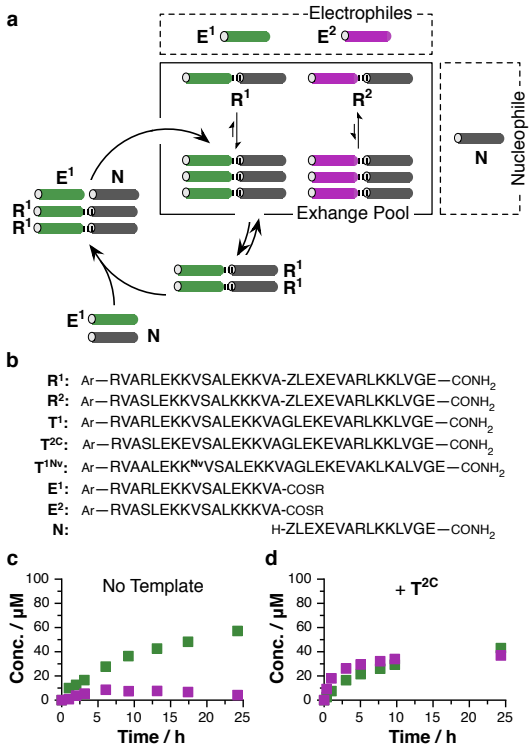


Fig. 16 **a** A replicating peptide network under thermodynamic control designed by Ashkenasy and co-workers. Two electrophilic components, $\mathbf{E^1}$ (green) and $\mathbf{E^2}$ (purple) can react with a nucleophilic component **N** (grey) to form reversibly peptide templates $\mathbf{R^1}$ and $\mathbf{R^2}$. $\mathbf{R^1}$ only is capable of forming catalytically active template assemblies, allowing it to replicate efficiently. **b** Labels and sequences of peptides employed. Ar = 4-acetamidobenzoate; Z = –$SHCH_2CO$; X = Lys-Ar; SR = 2-mercaptoethane sulfonate; Nv = 6-nitroveratryloxycarbonyl. **c** Formation of $\mathbf{R^1}$ (green) and $\mathbf{R^2}$ (purple) from [**N**] = [$\mathbf{E^1}$] = [$\mathbf{E^2}$] = 100 µM. **d** Formation of $\mathbf{R^1}$ (green) and $\mathbf{R^2}$ (purple) from [**N**] = [$\mathbf{E^1}$] = [$\mathbf{E^2}$] = 100 µM in the presence of $\mathbf{T^{2C}}$ (50 µM).

Kinetic analysis in the absence of template (**Fig. 16c**) revealed interesting behaviour, namely, while the concentration of $\mathbf{R^1}$ continued to increase throughout the reaction, as expected based on its replication efficiency, the initial increase in the concentration of $\mathbf{R^2}$ was followed by its hydrolysis to its constituent components, before the system reached equilibrium. Hydrolysis of $\mathbf{R^2}$ increased the transient concentration of the shared building block **N**, which, in turn, allowed $\mathbf{R^1}$ to form at the expense of $\mathbf{R^2}$. Instructing the DCL with $\mathbf{R^2}$ resulted in delayed equilibration to $\mathbf{R^1}$. Utilising an external, preformed template ($\mathbf{T^{2C}}$) that incorporates two glutamate residues (*g*13 and *e*13)—capable of forming stable heteromeric complexes with $\mathbf{R^2}$ (bearing lysine residues)—as input, the authors were able to alter the outcome of the competition between $\mathbf{R^1}$ and $\mathbf{R^2}$, resulting in nearly equal concentrations (**Fig. 16d**) of both templates at equilibrium.

The possibility of altering the behaviour of this network through the application of light as a stimulus was examined by introducing the photocleavable template $\mathbf{T^{1Nv}}$. Application of light to the reaction mixture resulted in the cleavage of the protecting photocleavable moiety, exposing the lysine residue (residue *e*, position 13) on this template, allowing it to act as a

crosscatalyst for the formation of $\mathbf{R^1}$ (in its protected form, $\mathbf{T^{1Nv}}$ acts as a weak crosscatalyst for $\mathbf{R^2}$). In this work, Ashkenasy and co-workers demonstrated experimentally for the first time that reversible peptide replication is possible, and that the product distribution within an interconnected peptide network under thermodynamic control depends on both the catalytic efficiency of each template as well as their thermodynamic stability. These results are of clear relevance for the understanding of the process of molecular evolution and formation of metabolic networks, and might possibly be extended to networks exhibiting chemical evolvability in the future.

Ashkenasy and co-workers exploited reversible and environmentally responsive replication for the design and implementation[86] of a larger peptide replicator network more functionally reminiscent of a prebiotic network. This more complex network (**Fig. 17**) was assembled from five peptide fragments, namely, three thioester incorporating electrophiles, $\mathbf{E^1}$ to $\mathbf{E^3}$, and two thiol-containing nucleophiles, **N** and $\mathbf{N^1}$. Reactions of these five building blocks furnish six different peptide replicators: $\mathbf{R^1}$, $\mathbf{R^2}$, $\mathbf{R^3}$, $\mathbf{R^{1A}}$, $\mathbf{R^{2A}}$ and $\mathbf{R^{3A}}$ ($\mathbf{R^1}$ and $\mathbf{R^2}$ were described in **Fig. 16**). The major distinguishing feature of these six replicators is their ability to form the coiled-coil assembly required for efficient replication. This ability is governed by the identity of the residue responsible for electrostatic interactions (**Fig. 17**, blue) at position *e*13 ($\mathbf{R^1}$/$\mathbf{R^{1a}}$ = glutamate, $\mathbf{R^2}$/$\mathbf{R^{2a}}$ = lysine and $\mathbf{R^3}$/$\mathbf{R^{3a}}$ = alanine), directly opposite the lysine residue at position *g*' (8) and the residue within the hydrophobic region responsible for inter-helical interactions (**Fig. 17**, red) at position *d*26 (**R** peptides = leucine, $\mathbf{R^a}$ peptides = alanine). The adaptive nature of this network and its capacity to adopt different network topology was demonstrated by changing the environmental conditions. Two template products, $\mathbf{R^1}$ and $\mathbf{R^3}$, were found to form particularly efficiently under all examined reaction conditions. Template $\mathbf{R^2}$ incorporates destabilising lysine–lysine interactions at the recognition interface and all alanine-containing templates displayed less efficient replication, or hydrolysed after a short time period as a result of their lower coiled-coil stability. In spite of the preference of the network for the formation of $\mathbf{R^1}$ and $\mathbf{R^3}$, the network could be directed towards enhanced formation of the $\mathbf{R^2}$ peptide by addition of the external template $\mathbf{T^{2C}}$, described previously, as instruction, or by exposure to high salt conditions that mask the unfavourable lysine–lysine interactions. As a consequence of the reversible nature of the peptide system, however, $\mathbf{R^2}$ hydrolysed over time into its constituent components, allowing $\mathbf{R^1}$ and $\mathbf{R^3}$ peptides to be formed in the highest yields. Ashkenasy and co-workers also probed the behaviour of this reaction network comprised of six replicators under thermodynamic control using kinetic simulations, and demonstrated successfully that the product distribution patterns observed experimentally can be reproduced.

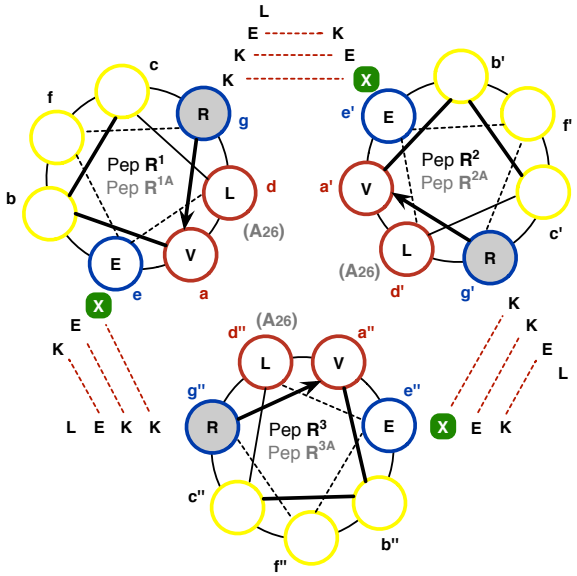


Fig. 17 Helical wheel representation of a trimeric coiled-coil formed by the association of six potential peptides: $\mathbf{R^1}$, $\mathbf{R^2}$ and $\mathbf{R^3}$ and $\mathbf{R^{1a}}$, $\mathbf{R^{2a}}$ and $\mathbf{R^{3a}}$ (subscript a denotes substitution of leucine at position ***26***d with alanine, highlighted in grey). Residue X (green) in position ***e***13 varies across the templates: glutamate = $\mathbf{R^1}$ / $\mathbf{R^{1a}}$, lysine = $\mathbf{R^2}$ / $\mathbf{R^{2a}}$ and alanine = $\mathbf{R^3}$ / $\mathbf{R^{3a}}$. Recognition between peptides and their assembly into complexes is mediated by the recognition between the hydrophobic residues at positions ***a*** and ***d*** (red) and electrostatic interactions between residues at positions ***e*** and ***g*** (blue). Residues ***b***, ***c*** and ***f*** (yellow) are exposed to the solvent surface and do not contribute to recognition.

An interesting and rather different strategy for self-replication under dynamic conditions has been reported[87] by Otto and co-workers, who utilised a dithiol building block **20** bearing a short oligopeptide sequence (**Fig. 18**), capable of associating reversibly through covalent disulfide exchange to give a dynamic library of macrocycles. These macrocycles can interact through non-covalent peptide–peptide interactions between hydrophobic (leucine) and hydrophilic (lysine) α-amino acids, arranged in an alternating pattern, allowing them to assemble into fibres, held together through β-sheet type interactions.

By employing a control building block **21**, lacking a peptide chain (**Fig. 18**), the authors demonstrated that in the absence of the recognition mediated by the peptide chains, the DCL, stirred in a borate buffer at pH 8, formed trimeric and tetrameric macrocycles preferentially. Although the behaviour of a DCL constructed from **20** displayed a product distribution that was initially similar to that of the control library employing **21**, preferential formation of hexamer and heptamer was detected over time. The appearance of these products was characterised by a sigmoidal concentration *vs.* time profile and was accompanied by a decrease in the concentration of the initially dominant trimer and tetramer. Interestingly, this behaviour was observed only when the DCL was agitated mechanically—stirring produced heptamer and shaking hexamer preferentially.

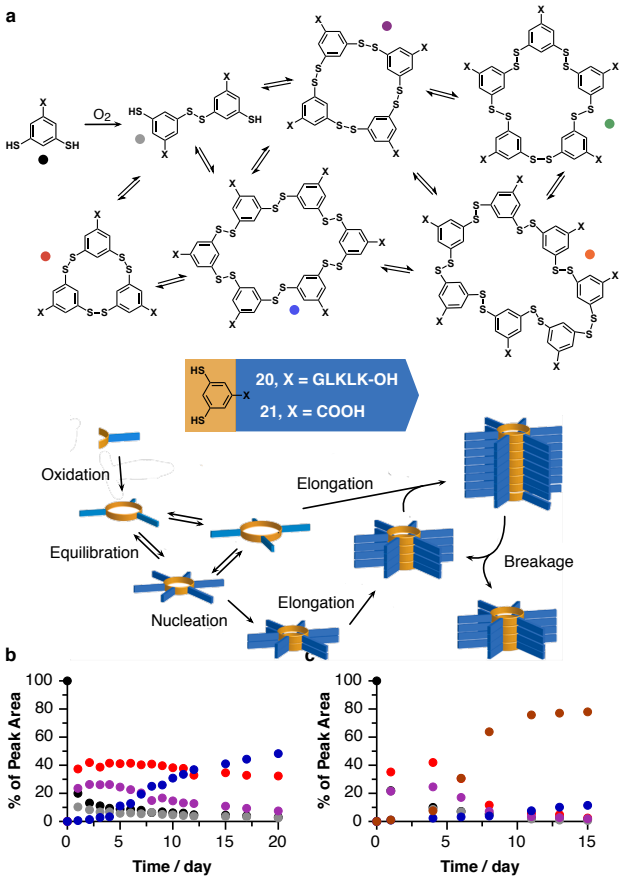


Fig. 18 **a** Dynamic library created from dithiol building blocks. Changes in product distribution observed over time upon agitation of a solution of **20** (3.8 mM) by **b** shaking (500 rpm) and **c** stirring (1200 rpm) as determined by high-performance liquid chromatography analyses. The colours of the circles shown in **b** and **c** match the library members shown in **a**. Figure adapted with permission from Ref. 88. Copyright 2015, American Chemical Society.

On day 4 the composition of the reaction mixture was dominated by the trimer and tetramer. Seeding the reaction mixture at this time with preformed hexamer and shaking, or preformed heptamer and stirring, resulted in the disappearance of lag period associated with the formation of the added template, either hexamer or heptamer. Exploiting a variety of control experiments and characterisation tools, Otto and co-workers were able to propose a model for the mechanosensitive fibre growth based on two processes, namely fibre elongation and fibre breakage—the latter of which is critical for the exponential transition from macrocycles to fibres in an autocatalytic process that can be compared to crystallisation. In subsequent work, Otto and co-workers were able to demonstrate[89] that fibre length, and thus also the rate of replication, are directly proportional to the stirring rate—*i.e.* the fibre length decreases as the rate of stirring increases.

Otto and co-workers have extended[90] their studies on mechanosensitive self-assembly driven fibre replicators to a dynamic system where two distinct sets of replicators emerge from a DCL constructed from two structurally similar building blocks **A** and **B** (**Fig. 19**), examined previosuly[91] in isolation and shown to form hexamers and octamers preferentially, respectively.

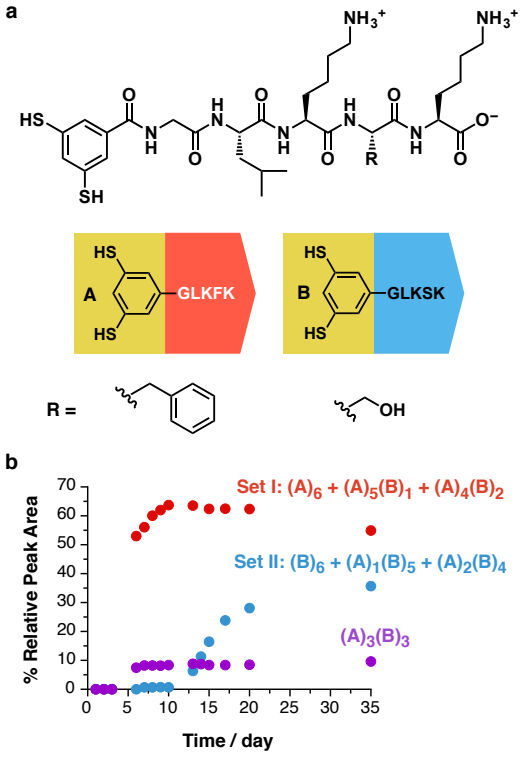


Fig. 19 **a** Dynamic library created from dithiol building blocks **A** (red) and **B** (blue), which differ in the nature of their R group. **b** Changes in product distribution observed over time in a library formed from building blocks **A** and **B** (3.8 mM each) in aq. borate buffer (50 mM, ph 8). The first set of replicators rich in dithiol block **A** is shown in red and the second set of replicators rich in **B** is shown in blue. The intermediate component $\mathbf{A_3B_3}$ that connects these two sets of replicators is depicted in purple.

The authors examined the behaviour of an equimolar ([**A**] = [**B**] = 3.8 mM) mixture of these components, monitoring all changes by ultraperformance liquid chromatography mass spectrometry (UPLC-MS) over 35 days. After three days, the mixture was dominated by a set of **A**-rich hexamers ($\mathbf{A_6}$, $\mathbf{A_5B}$, $\mathbf{A_4B_2}$, and $\mathbf{A_3B_3}$). The formation of this first set of replicators altered the pool of building blocks available in the reaction environment, and, as a consequence, gave rise to a second set of replicators, rich in **B** ($\mathbf{A_2B_4}$, $\mathbf{AB_5}$, and $\mathbf{B_6}$), after seven days. Interestingly, on repeating the experiments 12 times, although consistent behaviour was observed generally, the appearance of hexamer $\mathbf{A_3B_3}$ switched between the first and second set of replicators. The use of two different building blocks allows the replicator fibres to undergo "mutation" by incorporating different monomers. The authors examined the relationship between the two sets of replicators using a series of template-instructed experiments, showing that the first set can template the formation of the second set, transferring information about its macrocycle size if the added seed is rich in **B**. Exploiting a combination of theoretical and experimental methods, Frederix *et al.* have recently reproted[92] a highly detailed study of fibre-based replicators created from building block **B** (**Fig. 19**),

showing that that formation of organised architectures is mediated by cooperative β-sheet-driven self-assembly, and the aromatic core of the macrocyclic component from which the fibres are constructed is critical to the stability of the fibres. Concurrently, the Otto laboratory have also reproted[93] a study investigating the exchange of building blocks in stacks of self-replicating macrocycles.

The replication mechanism that underlies the fibre systems reported by Otto and co-workers is considerably different from the synthetic replicating systems discussed thus far. Even the larger-scale peptide replicators reported by the Ghadiri, Chmielewski and Ashkenasy laboratories exploit the well-established minimal models where ternary (or, in some peptide replicators, quaternary) complexes catalyse the replication processes. By contrast, the work reported by Otto and co-workers represents an important example of an alternative strategy for self-replication, driven by self-assembly and mechanical forces. This mode of self-replication is significantly less well defined than that described by the minimal model, and is beyond the scope of this review. Nevertheless, it should be noted that, to date, analogous self-replication (or self-reproduction/self-synthesis) has been demonstrated using self-assembled fibres, nanowires, micelles and vesicles. For more in-depth discussion of this body of work, the reader is directed to Refs. 94 and 95.

In one of the earliest examples examining self-replication in reversible systems, von Kiedrowski and Terfort explored[96] an amidinium-carboxylate salt bridge as an alternative to nucleotide base pairing recognition in order to drive self-replication of a small molecule-based synthetic system. The authors investigated the condensation reaction of several amines and aldehydes (**Fig. 20**).

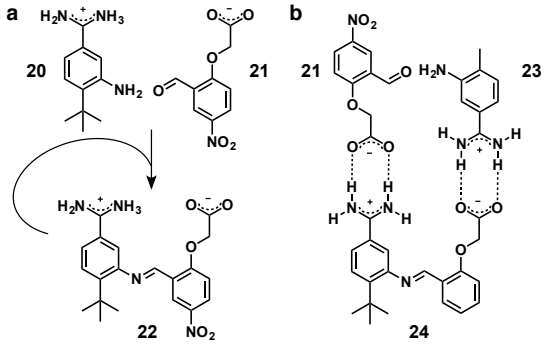


Fig. 20 **a** Small-molecule synthetic replicating system, exploiting carboxylate–amidinium salt bridge formation, investigated by Terfort and von Kiedrowski based on condensation of amines and aldehydes. **a** Amine **20** and aldehyde **21** form imine **22**, capable of self-replication (limited by product inhibition). **b** Crosscatalytic reaction of amine **23** with **21** mediated by imine template **24** showed exponential growth. Counter ions are omitted for clarity.

The imines formed were suitable templates for the association of the unreacted components in catalytically active ternary complexes. Analysis of the reaction between amine **20** and aldehyde **21** (DMSO-d_6), to give imine **22** (**Fig. 20a**), showed that the rate of product formation increased as the amount of preformed template **22** added to the reaction was progressively increased. Interestingly, examination of a structurally similar system where an amine **23** reacted with aldehyde **21** in the presence of imine **24** (**Fig. 20b**) showed exponential growth. While this system constitutes an example of a crosscatalytic replication, the authors showed that it is possible to overcome product inhibition in small molecule-based replicating systems.

Xu and Giuseppone reported[97] an example of a more complex imine library based on small organic molecules, where a single exchange pool component is stabilised by its ability to form a recognition-mediated duplex. Design of the self-complementary motif, necessary for the internal, template-mediated stabilisation was inspired by Rebek's replicator.[46] Condensation of three aldehydes (**Al^1** to **Al^3**) and two amines (**Am^1** and **Am^2**) afforded six different imine products (**Fig. 21**). Only the imine product **Al1–Am1** was capable of forming a recognition-mediated duplex [**Al^1–Am^1•Al^1–Am^1**]. Kinetic analysis of the full library revealed that the initial advantage in the rate of formation of **Al^1–Am^1**, eroded over time. **Fig. 21** (black circles) shows that this erosion stems from the thermodynamic preference of the library for products **Al^2–Am^2** and **Al^3–Am^2** (empty circles). The study showed clearly that the level of selectivity that can be achieved for a product formed *via* recognition-mediated reaction processes in the library is limited by the thermodynamic stabilities of the imine products, and the reversible nature of the system.

Contemporaneously, the Philp laboratory developed[98] an imine-based system (**Fig. 22**) formed from an aromatic aldehyde **25** and an amine **26**. These two components can react to form imine **27**, which is in dynamic equilibrium with its constituent building blocks. As a result of the complementarity between its carboxylic acid and 4,6-dimethyl amidopyridine sites, the formed imine has the capacity to assemble the unreacted aldehyde and amine in a catalytically active ternary complex [**25•26•27**] that can accelerate the reaction between **25** and **26**. Interestingly, the authors demonstrated that, while the addition of preformed imine **27** template removed the lag period observed in the absence of template, its addition also resulted in a decrease in the overall quantity of the newly formed imine **27** produced within the system. In fact, kinetic simulations revealed that the observed decrease is proportional to the amount of template added—a thermodynamic boundary imposes a limit on the formation of imine **27**, and addition of preformed template is not sufficient in itself to allow the system to break from this constraint. Instruction of this dynamic imine system with a reduced amine counterpart (**28**) of imine **27** confirmed that this crosscatalytic pathway operates efficiently, and without the decrease in imine formation observed in the presence of **27**. Nevertheless, even the addition of the reduced template **28** did not allow the system to shift away beyond its thermodynamic limit.

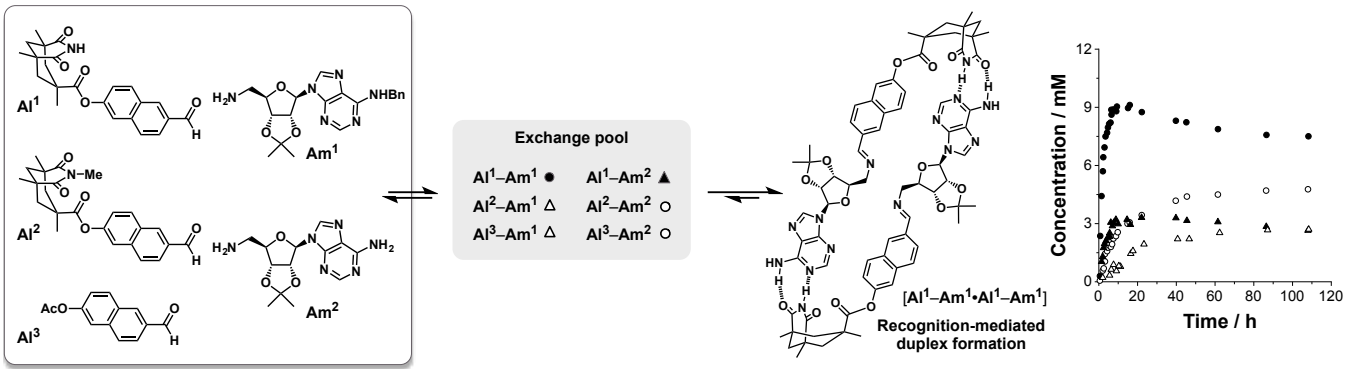


Fig. 21 **a** Dynamic library composed from three aldehydes, **Al**^{1} to **Al**^{3} and two amines, **Am**^{1} and **Am**^{2}, affording 6 different imine products. Only imine **Al**^{1}**–Am**^{1} (black circles) is capable of stabilisation through the formation a recognition-mediated duplex [**Al**^{1}**–Am**^{1}**•Al**^{1}**–Am**^{1}]. The initial advantage afforded in rate eroded over time as **Al**^{2}**–Am**^{2} and **Al**^{3}**–Am**^{2} (empty circles) products formed as the more thermodynamically stable products. The imine **Al**^{1}**–Am**^{2} (black triangle) product was similarly affected, decreasing in concentration over time. **Al**^{2}**–Am**^{1} and **Al**^{3}**–Am**^{1} (empty triangles) formed slowly and unselectively.

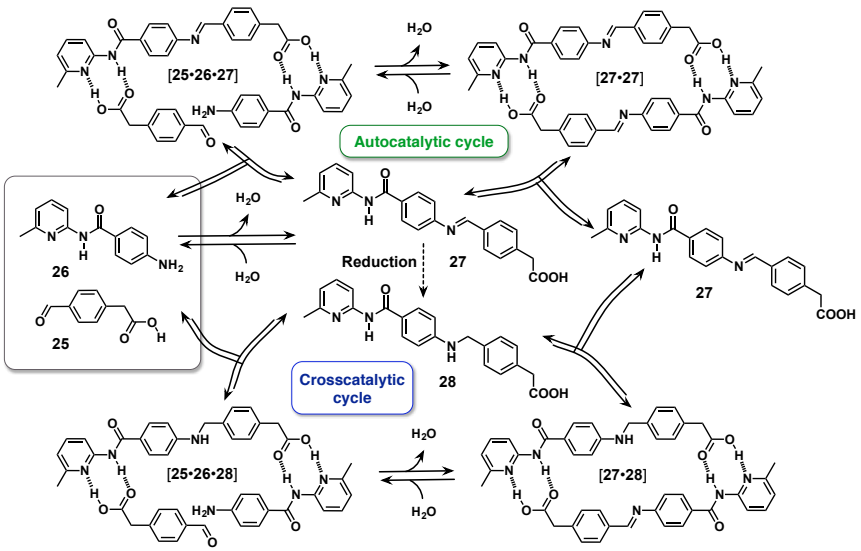


Fig. 22 A dynamic imine system incorporating formation of a self-replicator. Aldehyde **25** and amine **26** can react reversibly to form imine **27**, which is capable of templating its own formation ***via*** the catalytically active ternary complex [**25•26•27**]. The reversible nature of the covalent bond forming step employed, however, imposes a limit on the production of the self-replicating imine **2**7. Reduction of the imine to the corresponding amine **28** afforded a crosscatalytic template.

The examples of replicating systems examined under dynamic conditions reveal that the thermodynamic nature of DCLs imposes a limitation on the level of amplification that can be achieved upon addition of preformed template capable of binding one or several of the library components. As a consequence, the degree of amplification in the systems described so far tends to correlate with the amount of template added and the concentrations and binding affinities of the components present within a particular library. In addition to fine-tuning these reaction and thermodynamic parameters as a strategy for obtaining enhanced selectivity for the best affinity binder, a DCL can also be coupled to kinetically controlled processes in order to achieve increased selectivity. In this manner, the limits imposed on a system by the thermodynamic regime can be broken. By exploiting methods that allow the added target, *e.g.* template to not only interact but also to react with the library components, the best binders can be removed irreversibly from the thermodynamic pool *via* kinetic selection processes. An additional benefit arises from such utilisation of kinetically controlled processes, namely easier isolation of the amplified, now more stable library products. As well as potentially enhanced selectivity, the combination of thermodynamic and kinetic selection processes within a single system presents us with the possibility of exploring networks and systems with an additional layer of complexity that are closer in nature to the diverse interconnected networks present in nature.

Exploitation of template-mediated irreversible replication processes, in particular, presents an attractive strategy for amplifying the degree of perturbation within a dynamic system, and, thus, also the level of target amplification. In this case, an exchange pool of interconverting components (**Fig. 23**) contains a few selected members with the capacity to *either* interact or react unselectively with the added target. Only a

small subset of these components (*e.g.* **AZ** in **Fig. 23**), however, possess the ability to *both* react and interact with the target species, thus displaying amplification *via* irreversible template-mediated processes.

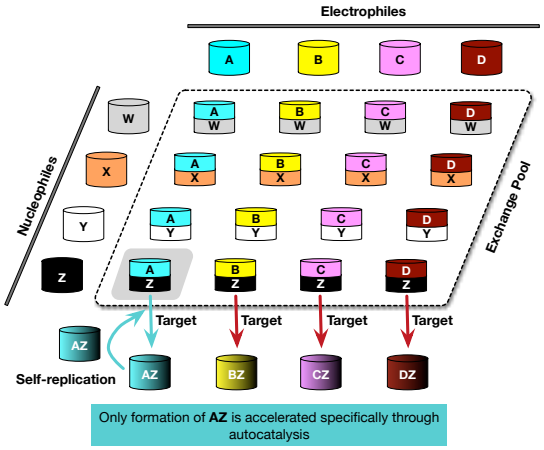


Fig. 23 Cartoon representation of a dynamic covalent library, where a replication process facilitates selective amplification (blue arrow) of a library member that is capable of both interacting and irreversibly reacting with the added target. Components capable of reacting but not interacting with the added target are transformed slowly and unselectively (red arrows).

Complementing their work[99] on DCLs driven by irreversible recognition-mediated reaction processes, Sadownik and Philp reported[100] a library of imines and nitrones that was coupled to a self-replication process. In this case, the position of the carboxylic acid on the maleimide component was selected to promote the template-mediated reaction *via* a ternary complex. Employing the *para*-substituted maleimide **12**, shown previously to form an efficient *trans* diastereoselective self-replicating system (**Fig. 8f**), the authors showed that instruction of the imine-nitrone dynamic library (**Fig. 24**) assembled from nitrone **29** and imine **30** (20 mM each) in CD_2Cl_2 saturated with *p*-toulene sulfonic acid resulted in rapid transformation of the recognition-enabled nitrone **31**. In fact, even in the absence of any template instruction, more than 48% of the library was transformed to products after 16 h, with *trans*-**32** (**Fig. 23**, green box) being by far the most dominant product at 80% ([*trans*-**32**]/[*cis*-**32**] = 21).

By contrast, nitrone **29**, which lacks the 6-amidopyridine recognition site, formed the *cis*-**33** and *trans*-**33** only very slowly and with low diastereoselectivity. The level of selectivity for *trans*-**32** was magnified further when the library was examined in the presence of 10 mol% of preformed *trans*-**32** as instruction. This addition of preformed template enabled *trans*-**32** to efficiently template its own formation *via* the template-mediated pathway from the onset of the reaction. In this instructed system, the overall conversion for all cycloadducts in the DCL was 64% after 16 h, of which 88% was constituted by *trans*-**32** ([*trans*-**32**]/[*cis*-**32**] = 38).

Examination of self-replicating systems within a dynamic environment such as that provided by dynamic covalent libraries present a unique opportunity to examine replication phenomena in a scenario where the overall outcome is affected by the interplay between the dynamic and replication processes. In this context, replicating species have to accomplish their own formation using the building blocks distributed in a pool of dynamically exchanging building blocks. As the reaction progresses, the replication processes perturb the dynamic state, operating under conditions that are temporarily driven away from equilibrium. In the end, however, the dynamic system reaches an equilibrium state once the building blocks have been exhausted.

Thus far, the examples of replicating systems coupled to DCLs have predominantly focused on examples involving a single replicator and studies incorporating two or more replicators are considerably under-explored. In the future, the established examples presented here will hopefully be extended to systems that examine networks of replicators competing for shared resources under dynamic conditions, thus helping to demonstrate how the outcomes of such systems vary from those observed under conditions where the building blocks are fully assembled at the reaction onset. Although DCLs can help us understand the origins of complexity in interconnected networks, the emergence of life on the prebiotic Earth most likely involved non-homogeneous environments. In order to explore replication processes under such conditions in the laboratory, it will be necessary to transition to reaction formats where replication processes do not become self-limiting such as those provided by reaction-diffusion and flow[101] environments

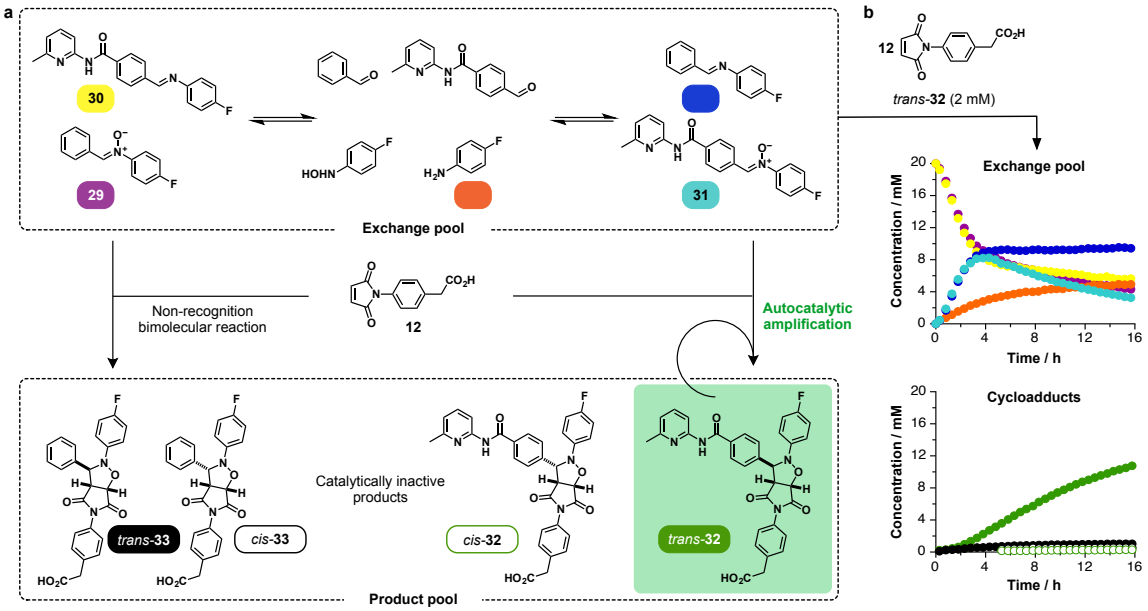


Fig. 24 **a** A dynamic covalent library assembled from nitrone **29** and imine **30**, which can equilibrate in CD_2Cl_2 saturated with ***p***-toluene sulfonic acid to give recognition-enabled nitrone **31** and a recognition-disabled imine. Within this library, two components are equipped with a reactive nitrone site and two with an amidopyridine recognition site. Only nitrone **31**, however, possesses both the recognition and reactive elements required for the reaction with maleimide **12** ***via*** a template-directed self-replicating pathway, forming the ***trans*-32** (green box) rapidly and with high diastereoselectivity. Nitrone **29** reacts with **12** to form ***cis*-33** and ***trans*-33** only ***via*** the slow, unselective bimolecular pathway. **b** Instruction of the library assembled from nitrone **29** and imine **30** (20 mM each) with **12** and 20 mol% (2 mM) of preformed template of ***trans*-32** results in dramatic amplification of ***trans*-32** from the DCL pool. The time course profiles show the changes in the exchange pool and cycloadduct components (corrected for template added) over time as determined by 470.4 MHz $^{19}F\{^1H\}$ NMR spectroscopy (0 °C).

Replication processes in non-standard reaction formats

The world around us is filled with molecular matter that is constantly interacting and reacting—processes that are assisted by the motion of molecules through space. On the early Earth, too, the emergence of an entity capable of replication probably occurred in a non-homogeneous, far-from-equilibrium environment, where the outcomes of all recognition and reaction processes were influenced by diffusion—a physical process with a key role in controlling[102] biochemical processes within cells. Such prebiotically relevant reaction-diffusion (R-D) environments differ markedly from the WSBR conditions typically employed for the analysis of self-replicating systems in the laboratory, as evidenced by the examples of replicating systems discussed in this review thus far. The absence of concentration gradients in a reaction performed under WSBR conditions means that diffusion plays little or no role in the outcome observed. The closed reaction environment imposes a limit on the selectivity that can be achieved in a network of two competing recognition-mediated reactive processes, and also limits the complexity of outcomes that can be achieved in interconnected chemical systems. In order to move away from the barriers imposed by kinetic selection, it is necessary to explore out-of-equilibrium, non-homogeneous reaction conditions.

Examples of systems exhibiting non-linear and spatiotemporal phenomena are considerably more common in the world around us than might seem at first sight—ranging from our unpredictable weather[1f,103] to patterns[104] on the skins of animal or precipitation patterns[104c,105] in rocks. For a system to sustain its dynamic, non-equilibrium state, a continuous supply of energy (light, heat, chemical energy, *etc.*) and matter is necessary, allowing spontaneous formation of, for example, self-organising patterns and R-D fronts. Oscillatory and autocatalytic processes are the core driving forces for the spontaneous generation[1d,102b,104c] of patterns and R-D fronts in nature. In the laboratory setting, the coupling of autocatalysis with diffusion in a non-homogeneous and unstirred environment can, likewise, gives rise to a propagating R-D front. The earliest examples of such complex phenomena demonstrated in chemical laboratories were based on inorganic chemistries, and despite their considerable initial controversy, an extensive body of work now exists[1d,1e,106] in the literature on inorganic systems that exhibit different spatiotemporal behaviour under non-equilibrium conditions—with the Belousov–Zhabotinsky oscillating reaction[107] representing perhaps the best-known example. Reports of organic systems exhibiting such behaviour, however, are significantly more scarce—in fact, the propagating R-D front detected[108] in the Co(II)-catalysed autoxidation of benzaldehyde was, for a long time, the only example of an R-D front involving a small organic molecule.

Such propagating fronts represent an opportunity to study replication processes, which are inherently autocatalytic, under

unfamiliar conditions. The concentration gradient created by the addition of a small quantity of preformed autocatalyst at a specific location to a solution of reactants will drive the propagation of an R-D front mediated by the replicating template, thereby allowing the replicator to operate at its optimum efficiency for a significant portion of the reaction.

In the last thirty years, several examples of R-D fronts based on RNA and DNA frameworks capable of replication have been reported. McCaskill and co-workers reported[109] the first example of an R-D front observed in a system based on replicating RNA. The propagating RNA fronts were initiated by the addition of preformed molecules of RNA at particular locations in an essentially two-dimensional capillary tube reactor (**Fig. 25**), containing a solution of RNA polymerase, nucleotide building blocks and buffer. Interestingly, McCaskill and co-workers were able to demonstrate that the fronts can emerge stochastically and are capable of evolving as they progress in space over time.

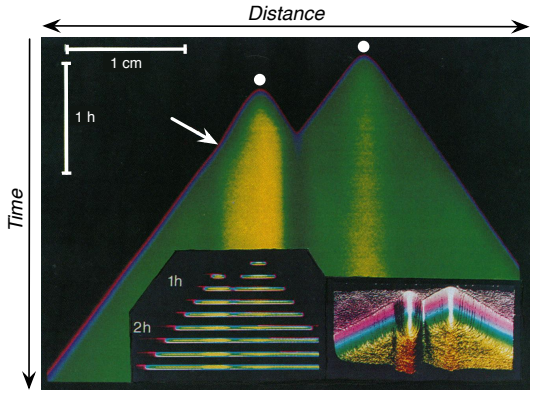


Fig. 25 A plot showing the distance travelled over time by propagating RNA fronts initiated by the addition of two single RNA molecules (white circles). The white arrow indicated a change in the velocity of the RNA wave. Concentration of RNA is represented using a colour scale, where black denotes the lowest and orange the highest concentration. Figure adapted with permission from Ref. 109a. Copyright 1993, National Academy of Sciences.

Until 2013, this study was the only example of its kind. Estévez-Torres, Rondelez and co-workers reported[110] an example of travelling concentration waves (**Fig. 26**) in a biochemical network exhibiting predator-prey[111] type of oscillations. The network presented in this study represents an extension of their previous work on DNA-based predator-prey systems, employing carefully designed DNA oligonucleotide-based molecules (**Fig. 26c**), connected by a shared encoding sequence. Namely, the predator-prey network was constructed from three components: prey **N**, predator **P**, and grass **G**—the template required for growth of the prey (**Figs. 26b** and **26c**). The reaction processes within this network are controlled by three purified enzymes, namely a polymerase, nicking enzyme and exonuclease. In the absence of these enzymes, the replication reactions in the network would stall. The prey utilises the grass for its formation, and in turn, the predator consumes the prey component in order to form another molecule of itself. Both predator and prey can decay through the action of the exonuclease enzyme. By examining the molecular network under R-D conditions—*i.e.* within the environment of an unstirred, 8 mm wide and 200 μm deep circular reactor, the authors were able to demonstrate DNA-based travelling prey-predator R-D fronts.

Recently, Estevez-Torres and co-workers have introduced[112] a more general method for achieving control over the reaction and diffusion parameters of the DNA components employed in programmable R-D networks based on an autocatalytic node of the DNA polymerase exonuclease nicking enzyme (PEN) toolbox.[113]

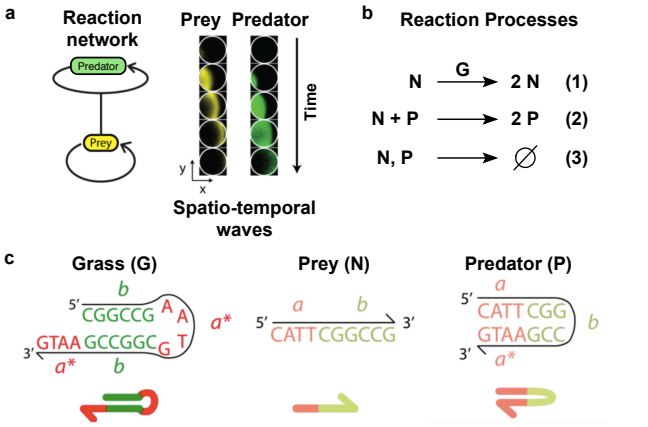


Fig. 26 **a** A biochemical network exhibiting predator-prey type of oscillations, connected by a common DNA oligonucleotide sequence capable of forming a propagating reaction-diffusion front. **b** Three reactions at the core of the predator-prey network: (1) autocatalytic growth of prey on the grass template, (2) autocatalytic growth of predator, with a consumption of prey, and (3) decay of predator/prey. **c** Structure of grass (**G**), prey (**N**) and predator (**P**). Complementary DNA sequences are highlighted in the same colour, whereas dark and light shade represent regions that can and cannot be destroyed through the action of an exonuclease. Figure adapted with permission from Ref. 110. Copyright 2013, American Chemical Society.

The notion of competitive autocatalysis in R-D systems was investigated[114] by Showalter and co-workers using simulations as early as 1998. The authors examined the possibility of achieving complete selectivity for a single product in a system composed of two autocatalytic processes competing for a shared building block **A** (**Eqs. 1** and **2**, **Fig. 27**). The autocatalytic products, **B** and **C**, each have a specific diffusion coefficient ((D_B and D_C) and their formation from their constituent components (reaction of **A** and **B**, or **A** and **C**) is governed by a specific rate constant (k_B and k_C).

Using computer simulations, Showalter and co-workers showed that the selectivity observed in a propagating R-D front depends on the relative ratios of the rate constants (k_C/k_B) and the diffusion coefficients (D_B/D_C) specific to the two competing replicators—**B** and **C**. **Fig. 27** shows an example simulation, where the two autocatalytic species have identical diffusion coefficients ($D_B = D_C$) and the rate constant k_B is twice k_C. The outcome of this simulation shows that the species with the higher rate constant, **B**, as the kinetically favoured product, forms a wave with constant velocity. In contrast, product **C**, travels from the initial seeding point almost exclusively through

diffusion, propagating over a significantly smaller distance. This observation is a direct consequence of coupling of the reaction processes where **A** → **B** is faster with diffusion of **B** from the starting point into areas containing fresh reagent **A**. In this situation where the diffusion coefficients are identical, the rate constant for the formation of the competing species plays a crucial role in determining the selectivity. This straightforward example of competition between **B** and **C** was elaborated into a simulation where the two species are more closely matched in their replicating abilities. More specifically, Showalter and co-workers showed that when $D_B/D_C > k_C/k_B$, the autocatalytic template **B** forms a propagating wave selectively (favoured by diffusion) while if $D_B/D_C < k_C/k_B$, the shared component **A** gets converted predominantly to product **C**, favoured kinetically, instead.

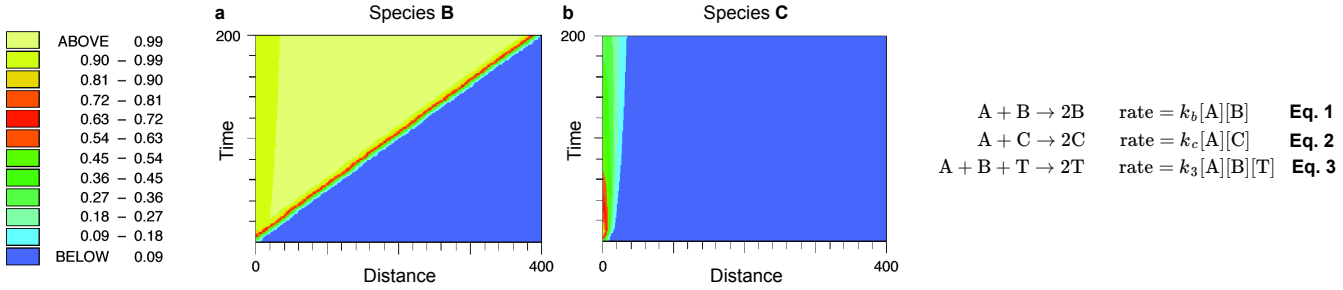


Fig. 27 Simulated time ***vs.*** distance travelled plots for the modelled evolution of autocatalytic species **B** and **C**, shown in **a** and **b**, respectively, when $D_B/D_C = 1$ and $k_C/k_B = 0.5$. Concentration levels of each product formed within the simulation are represented using the colour key shown on the left-hand side, where blue represents the lowest and yellow the highest concentration. Figure adapted with permission from Ref. 114. Copyright, 1998, Royal Society of Chemistry.

This theoretical study showed that R-D conditions can offer selectivity in a network of competing autocatalytic reactions far beyond that of a WSBR system, where the final composition always comprises a mixture of products. The degree of selectivity in a system of competing autocatalytic reactions coupled to diffusion processes can be fine-tuned by altering the reaction and diffusion parameters of the reacting species within a particular system. Despite the clear potential utility of this study as a general model, the simulations employed in this work examine the formation of autocatalytic products, **B** and **C**, as a simple second order reaction with respect to the reaction components, **A** and **B** or **C** (**Eqs. 1** and **2**, **Fig. 27**). This approach differs markedly from that of the minimal model of self-replication, introduced in the earlier sections, where a template molecule is required to establish an autocatalytic cycle (**Eq. 3**, **Fig. 27**), with an overall reaction order of up to 3. Additionally, the study by Showalter and co-workers fails to take into consideration any crosscatalytic interactions and thus the results of the simulations are not immediately applicable to networks of replicators where such pathways might operate.

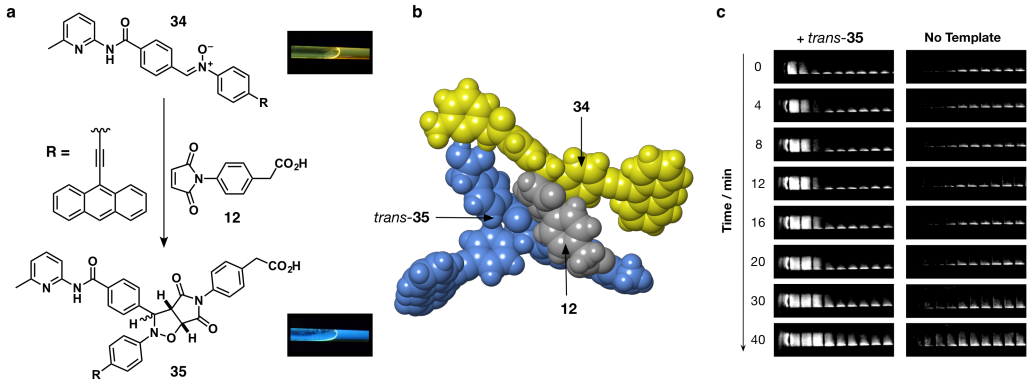


Fig. 28 **a** Design of an optically active self-replicating system formed from maleimide **12** and a 9-ethynylanthracene containing nitrone **34**. Formation of replicator ***trans*-35** from nitrone **34** and maleimide **12** is associated with a change in fluorescence from bright yellow (nitrone) to blue (cycloadduct). **b** Calculated (RM1) space-filling structure of the transition state [**12**•**34**•***trans*-35**] allowing the template-directed formation of ***trans*-35**. **c** Processed grey-scale images collected over time, using a 365 nm UV lam to illuminate a template instructed reaction-diffusion experiment (left-hand column, +***trans*-35**) and control condition, lacking template (right-hand column, no template). Figure adapted with permission from Ref. 115. Copyright 2016, American Chemical Society.

In 2016, building on these simulations and the repertoire of reaction-diffusion fronts based on oligonucleotides, Philp and co-workers reported[115] a replicating system with an optical signature that permits monitoring of the progress of the replication progress within the R-D environment in real time. This replicating system (**Fig. 28**), based on the design of a well-established efficient replicating system, first reported[60] by Kassianidis and Philp in 2006 (**Fig. 8f**), is capable of

functioning within an R-D environment, and initiating and sustaining a chemical propagating R-D front, even in the absence of enzymes, unlike the RNA and DNA examples reported to date. The replicating system was re-engineered and incorporated a 9-ethynylanthracene fluorescent tag (**Fig. 28a**) within the nitrone component **34**. This nitrone exhibits bright yellow fluorescence in $CDCl_3$ solution when irradiated with a long-wavelength UV light. The emission wavelength undergoes a dramatic change towards the blue region of the spectrum upon reaction of the nitrone with maleimide **12**. The ability of the redesigned system *trans*-**35** to self-replicate through the ternary catalytically active complex [**12**•**34**•*trans*-**35**] (**Fig. 28b**) was confirmed through a series of kinetic experiments, which revealed that production of **35** proceeds very efficiently, with extremely high diastereoselectivity for the *trans* diastereoisomer ([*trans*]/[*cis*] ratio of cycloadducts is *ca.* 100). Addition of 10 mol% of preformed *trans*-**35** template to the reaction components necessary for its formation shortened the lag period dramatically, confirming that the system maintains its ability to self-replicate despite the presence of the significantly more bulky anthracene tag, with an $EM_{kinetic}$ of 16.2 M.

The change in fluorescence, associated with the formation of replicator *trans*-**35** was confirmed as the signature of the underlying autocatalytic reaction by UV-Vis spectroscopy. The ability of the designed system to initiate and sustain a propagating R-D front in response to the addition of a small amount of a solution of preformed template to an unstirred solution of the unreacted building blocks was assessed within the environment of a 50 μL microsyringe (**Fig. 28c**), as opposed to a flat plate, as a result of the highly volatile character of the $CDCl_3$ reaction solvent employed in the formation of replicator *trans*-**35**. One 50 μL syringe was filled completely with a 5 mM solution of **12** and **34** (No template, **Fig. 28c**) whilst the second syringe, pre-filled with the same building blocks contained 2 μL of *trans*-**35** (+*trans*-**35**, **Fig. 28c**), added at one end of the syringe. Examination of the processed images acquired by illumination of both syringes at a sequence of time points with a 365 nm UV light (**Fig. 28c**) showed that addition of preformed replicator template solution allowed *trans*-**35** to establish a propagating R-D front. In contrast, no front was visible in the experiment lacking the preformed template, and, instead, the syringe underwent a uniform change in the colour of the emission, from yellow to blue. This experimental demonstration of an R-D front driven by a small-molecule organic synthetic replicator, described[115] by the Philp laboratory opens up the possibilities of exploring networks of interconnected replicators within R-D environment, *i.e.* under far-from-equilibrium conditions that could allow such systems to express the phenomenon of selective replication.

Future work will hopefully see a progressive shift towards the study of replicating systems under R-D and flow conditions, whilst retaining the benefits of analyses and characterisation of replicators under established WSBR conditions. One of the upcoming challenges of systems chemistry will be to work towards the delineation of the parameters that affect the behaviour of self-replicating systems under non-equilibrium conditions and how this behaviour differs from that observed in WSBR formats, be it purely kinetically driven or coupled to dynamic processes.

Concluding remarks and outlook

The fundamental role of auto- and crosscatalytic processes in the transition from simple chemical building blocks to LUCA and extant biology is evident, providing molecular entities or networks with the capacity to sustain and amplify themselves through repeated rounds of replication. Over the last three decades, a plethora of molecular systems capable of copying themselves or entities complementary to them in the reciprocal sense have been developed and experimentally demonstrated. This body of work is critical for our understanding of synthetic replicating systems and the requirements for their operation as well as interaction with each other. In an attempt to understand the progressive increase in complexity and emergence[4b,5a,7a] of function, *i.e.* beyond replication, which marks the transition from chemistry to life, systems chemistry has started to exploit the coupling of the broad palette afforded by synthetic chemistry and DCC with the molecular recognition toolkit pioneered[116] by the field of supramolecular chemistry to engineer increasingly complex reaction networks that are constructed from simple components and possess the capacity to react and interact in multiple ways. The coupling of the catalytic relationships inherent within replicating systems with the design principles of systems chemistry offers the opportunity to adopt a bottom-up approach to complexity through the creation of networks of replicators with well-defined connectivity. In these systems, complex behavioural outcomes—ranging from stereoselective replication, error-correction and Boolean logic operations to the fabrication of complex architectures and capacity for diversification—are facilitated by the interactions of the various components.

Building on these essential foundations, one must ask—what comes next for the field molecular replicators? Slowly, but surely, it is becoming apparent that there are several key changes that can facilitate the expression of increasingly complex and emergent behaviours, *i.e.* ones that are associated with living entities and include, for example, selective replication, switching[76a,76c] from one state to another and appearance of homochirality, in systems constructed from the bottom-up. The first step necessitates a transition away from constant and homogeneous reaction formats, such as the WSBRs that have been traditionally employed as the reactors of choice for the study of artificial replicators, towards more dynamic, nonhomogeneous and thus non-equilibrium environments. This change mimics the environment found in living cellular systems, where diffusion, concentration gradients and regulatory feedback loops play central roles in sustaining life under obviously nonhomogeneous conditions. Under such far-from-equilibrium reaction conditions, provided in the laboratory by propagating R-D and flow environments or the presence of droplets, vesicles and micelles, replicating systems can overcome the inherent limits that are placed on selection processes in networks operating under homogeneous closed

conditions. Through this change, synthetic replicators can progress towards the emergence of complex behaviour, such as selective formation of one species over another from a mixture of competing replicators. The second critical step in bridging the gap between synthetic replicators and living systems, which are characterised by a state of dynamic kinetic stability[15a,117] and integrate replication with metabolism and compartmentalisation, will require the field of artificial replicators to apply the bottom-up approach to the construction of systems that couple[4b,5a,7a,118] replication phenomena with metabolic and boundary subsystems. Subsequently, these two changes will serve as a framework for the design of replicating systems that are susceptible to and capable of spontaneous mutation. The coupling of such systems to destructive processes under non-equilibrium conditions should, ultimately, enable the creation of systems capable of moving beyond simple information transfer, towards environment-dependent artificial replicators that can undergo open-ended[4b,5a,7a] evolution. It is inevitable that these changes and increases in complexity, although from the bottom-up, will be accompanied by increasingly difficult challenges when it comes to the tractability of the experimental systems. For this reason, our experimental progress and capacity to translate between theory and practice will be tied closely not only to the advances in technology and analytical techniques, but will also rely significantly on computational tools[18,71a,71b,119] and scientific collaborations. Ultimately, whilst preserving the importance of characterising all individual components and their interactions and reactions, these avenues of research will be at the core of driving our understanding of the key interplay between replication phenomena, other reaction processes and the environment, which underlies the progress towards more complex outcomes in terms of behaviour in synthetic replication networks.

Acknowledgements

This work was supported by University of St Andrews and the award of a Postgraduate Studentship from Engineering and Physical Sciences Research Council (EP/K503162/1) to T.K.

Notes and references

‡ In 1993, von Kiedrowski introduced[7i] the concept of catalytic efficiency (ε) and reaction order (p) as parameters for describing and quantifying the behaviour of replicating systems. It should be noted that the concept of kinetic effective molarity ($EM_{kinetic}$) discussed in this review is not directly comparable to the values of ε reported in the literature for replicating systems (some of which are quoted in this review). Values of $EM_{kinetic}$ are derived from the ratio of a first order rate constant and a second order rate constant and this ratio has defined units of M, hence its description as "effective molarity". By contrast, the parameter ε is derived from the ratio of two rate constants, one of which has units that are dependent on the value of p. Strictly, this parameter cannot be described as "effective molarity" since it does not have units of M (unless p = 1) and values of ε are not strictly comparable as they most likely have different units.

§ The *cis* and *trans* notation reflects the relative configuration of the three protons located on the bicyclic ring structure formed in the cycloaddition reaction. In the *trans* cycloadduct, the proton derived from the nitrone is located on the opposite face of the bicyclic ring structure to the two protons that originate from the maleimide. In the *cis* cycloadduct, the protons derived from the nitrone and maleimide components are located on the same face of the fused ring system. In the absence of effects originating from molecular recognition, the *trans* and *cis* cycloadducts are normally formed in a 3:1 ratio.

1 (a) S. H. Strogatz, *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry, and Engineering*, Westview Press, USA, 2015; (b) C. Oestreicher, *Dialogues Clin. Neurosci.*, 2007, **9**, 279–289; (c) G. M. Whitesides and R. F. Ismagilov, *Science*, 1999, **284**, 89–92; (d) I. R. Epstein and J. A. Pojman, *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos*, Oxford University Press, New York, 1998; (e) I. R. Epstein and K. Showalter, *J. Phys. Chem.*, 1996, **100**, 13132–13147; (f) J. Gleick, *Chaos: Making a New Science*, Penguin Books USA Inc., Viking Books, New York, USA, 1987.

2 (a) E. Negishi, *Angew. Chem. Int. Ed.*, 2011, **50**, 6738–6764; (b) R. A. Shenvi, D. P. O'Malley and P. S. Baran, *Acc. Chem. Res.*, 2009, **42**, 530–541; (c) Y. Chauvin, *Angew. Chem. Int. Ed.*, 2006, **45**, 3741–3747; (d) K. C. Nicolau, D. Vourloumis, N. Winssinger and P. S. Baran, *Angew. Chem. Int. Ed.*, 2000, **39**, 44–122; (e) E. J. Corey, *Angew. Chem. int. Ed.*, 1991, **30**, 455–465.

3 (a) A. Herrmann, *Chem. Soc. Rev.*, 2014, **43**, 1899–1933; (b) F. B. L. Cougnon and J. M. L. Sanders, *Acc. Chem. Res.*, 2012, **45**, 2211–221; (c) *Dynamic Combinatorial Chemistry*, J. N. H. Reek and S. Otto, Eds.; Wiley-VCH, Weinheim, 2010; (d) *Dynamic Combinatorial Chemistry in Drug Discovery, Bioorganic Chemistry, and Materials Science*, B. L. Miller, Ed., Wiley, Hoboken, NJ, 2010; (e) J.-M. Lehn, *Chem. Soc. Rev.*, 2007, **36**, 151–160; (f) P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders and S. Otto, *Chem. Rev.*, 2006, **106**, 3652–3711; (g) S. Otto, R. L. E. Furlan and J. K. M. Sanders, *Drug Discov. Today*, 2002, **7**, 117–125; (h) J.-M. Lehn, *Chem. Eur. J.*, 1999, **5**, 2455–2463.

4 (a) A. de la Escosura, C. Briones, K. Ruiz-Mirazo, *J. Theor. Biol.*, 2015, **381**, 11–22; (b) K. Ruiz-Mirazo, C. Briones, A. de la Escosura, *Chem. Rev.* 2014, **114**, 285–366; (c) H.-Y. Chuang, M. Hofree and T. Ideker, *Annu. Rev. Cell Dev. Biol.*, 2010, **26**, 721–744; (d) A. Friboulet and D. Thomas, *Biosens. Bioelectron.*, 2005, **20**, 2404–2407; (e) H. Kitano, *Science*, 2002, **295**, 1662–1664.

5 (a) G. Ashkenasy, T. M. Hermans, S. Otto and A. F. Taylor, *Chem. Soc. Rev.*, 2017, **46**, 2543–2554; (b) S. Islam and M. W. Powner, *Chem*, 2017, **2**, 470–501; (c) E. Mattia and S. Otto, *Nat. Nanotechnol.*, 2015, **10**, 111–119; (d) G. Clixby and L. Twyman, *Org. Biomol. Chem.*, 2015, **14**, 4170–4184; (e) G. von Kiedrowski, S. Otto and P. Herdewijn, *J. Syst. Chem.*, 2010, **1**:1; (f) J. R. Nitschke, *Nature*, 2009, **462**, 736–738; (g) J. Stankiewicz and L. H. Eckardt, *Angew. Chem. Int. Ed.*, 2006, **45**, 342–344.

6 (a) J. D. Sutherland, *Nat. Rev. Chem.*, 2017, **1**: 0012; (b) K. Ruiz-Mirazo and A. Moreno, *Metode Sci. Stud. J.*, 2016, **6**, 151–159; (c) *NASA Astrobiology Strategy*, Lindsay Hays, Ed., 2015; (d) J. D. Sutherland, *Angew. Chem. Int. Ed.*, 2015, **55**, 104–121; (e) H. Rauchfuss, *Chemical Evolution and the Origin of Life*, translated by T. N. Mitchell, Springer, Berlin, 2008; (f) S. F. Mason, *Chemical Evolution*, Oxford University Press Inc., New York, 1991.

7 (a) H. Duim and S. Otto, *Beilstein J. Org. Chem.*, 2017, **13**, 1189–1203; (b) A. J. Bissette and S. P. Fletcher, *Angew. Chem., Int. Ed.*, 2013, **52**, 12800–12826; (c) J. Huck and D. Philp, *Supramolecular Chemistry: From Molecules to Nanomaterials*, John Wiley & Sons, Ltd.: New York, 2012; Vol. 4, pp 1415−1446; (d) R. Plasson, A. Brandenburg, L. Jullien and H. Bersini, *Artif. Life*, 2011, **17**, 219–236; (e) A. Vidonne and D. Philp, *Eur. J. Org. Chem.*, 2009, **2009**, 593–610; (f) Z. Dadon, N. Wagner and G. Ashkenasy, *Angew. Chem. Int. Ed.*, 2008, **47**, 6128–6136; (g) V. Patzke and G. von Kiedrowski, *ARKIVOC*, 2007, **46**, 293–310.; (h) N. Paul and G. F. Joyce, *Curr. Opin. Chem. Biol.*, 2004, **8**, 634–639; (i) G. von Kiedrowski, *Bioorg. Chem. Front.*, 1993, **3**, 113–146.

8 (a) *Nobel Lectures, Physiology or Medicine 1942–1962*, Elsevier Publishing Company, Amsterdam, 1964; (b) J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737–738; (c) J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 964–967.

9 (a) D. L. Theobald, *Nature*, 2010, **465**, 219–222; (b) S. A. Tsokolov, *Astrobiology*, 2009, **9**, 401–412; (c) N. Glansdorff, Y. Xu and B. Labedan, *Biol. Direct*, 2008, **3**:29; (d) N. R. Pace, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 805–808.

10 (a) S. Tirard, *Orig. Life Evol. Biosph.*, 2010, **40**, 215–220; (b) S. A. Benner, *Astrobiology*, 2010, **10**, 1021–1030; (c) A. Lazcano, *Chem. Biodivers.*, 2008, **5**, 1–15; (d) C. E. Cleland and C. F. Chyba, *Orig. Life Evol. Biosph.*, 2002, **32**, 387–393; (e) *Origins of Life: the Central Concepts*, G. F. Joyce, D. W. Deamer, G. Fleischaker, Eds., Jones and Bartlett, Boston, 1994, Foreword.

11 (a) S. A. Kauffman, *Life*, 2011, **1**, 34–48; (b) A. Lazcano, *Cold Spring Harb. Perspect. Biol. 2*, 2010, **11**:a002089; (c) A. Eschenmoser, *Tetrahedron*, 2007, **63**, 12821–12844.

12 (a) P. G. Higgs and N. Lehman, *Nature Rev. Genet.*, 2015, **16**, 7–17; (b) T. Czárán, B. Könnyu and E. Szathmáry, *J. Theor. Biol.*, 2015, **381**, 39–54; (c) M. Neveu, H.-J. Kim and S. A. Benner, *Astrobiology*, 2013, **13**, 391–403; (d) J. W. Szostak, *J. Syst. Chem.*, 2012, **3**:2; (e) N. Vaidya, M. L. Manapat, I. A. Chen, R. Xulvi-Brunet, E. J. Hayden and Niles Lehman, *Nature*, 2012, **491**, 72–77; (f) C. Anastasi, F. F. Buchet, M. A. Crowe, A. L. Parkes, M. W. Powner, J. M. Smith and J. D. Sutherland, *Chem. Biodivers.*, 2007, **4**, 721–739; (g) W. Gilbert, *Nature*, 1986, **319**, 618.

13 (a) K. B. Muchowska, S. J. Varma, E. Chevallot-Beroux, L. Lethuillier-Karl, G. Li and J. Moran, *Nat. Ecol. Evol.*, 2017, doi:10.1038/s41559-017-0311-7; (b) J. Peretó, *Chem. Soc. Rev.*, 2012, **41**, 5394–5403; (c) R. Shapiro, *Q. Rev. Biol.*, 2006, **81**, 105–125; (d) E. Smith and H. J. Morowitz, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 13168–13173; (e) L. E. Orgel, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 12503–12507.

14 (a) P. L. Luisi, *Orig. Life Evol. Biosph.*, 2014, **44**, 335–338; (b) P. L. Luisi, *The Emergence of Life*, Cambridge University Press, 2006; (c) D. Segré, D. Lancet, *EMBO Rep.*, 2000, **1**, 217–222; (d) V. Norris and D. J. Raine, *Orig. Life Evol. Biosph.*, 1998, **28**, 523–537.

15 (a) A. Pross and R. Pascal, *Beilstein J. Org. Chem.*, 2017, **13**, 665–674; (b) S. Matsumura, A. Kun, M. Ryckelynck, F. Coldren, A. Szilágyi, F. Jossinet, C. Rick, P. Nghe, E. Szathmáry and A. D. Griffiths, *Science*, 2016, **354**, 1293–1296; (c) B. H. Patel, C. Percivalle, D. J. Ritson, C. D. Duffy and J. D. Sutherland, *Nat. Chem.*, 2015, **7**, 301–307; (d) M. Powner and J. Sutherland, *Phil. Trans. R. Soc. B*, 2011, **366**, 2870–2877; (e) A. Lazcano, *Orig. Life Evol. Biosph.*, 2010, **40**, 161–167.

16 (a) A. J. Kirby, *Adv. Phys. Org. Chem.*, 1980, **17**, 183–278; (b) M. I. Page, *Chem. Soc. Rev.*, 1973, **2**, 295–323; (c) M. I. Page and W. P. Jencks, *Proc. Natl. Acad. Sci. U. S. A.*, 1971, **68**, 1678–1683.

17 (a) P. Motloch and C. A. Hunter, *Adv. Phys. Org. Chem.*, 2016, **50**, 77–118; (b) C. A Hunter and H. L. Anderson, *Angew. Chem. Int. Ed.*, 2009, **48**, 7488–7499; (c) C. A. Hunter, *Angew. Chem. Int. Ed.*, 2004, **43**, 5310–5324.

18 E. Bigan, H.-P. Mattelaer and P. Herdewijn, *J. Mol. Evol.*, 2016, **82**, 93–109.

19 T. Inoue, G. F. Joyce, K. Grzeskowiak, L. E. Orgel, J. M. Brown and C. Reese, *J. Mol. Biol.*, 1984, **178**, 669–676; (b) T. Inoue and L. E. Orgel, *Science*, 1983, **219**, 859–862; (c) R. Lohrmann and L. E. Orgel, *J. Mol. Biol.*, 1980, **142**, 555–567; (b) L. E. Orgel and R. Lohrman, *Acc. Chem. Res.*, 1974, **7**, 368–377

20 G. von Kiedrowski, *Angew. Chem. Int. Ed. Engl.*, 1986, **25**, 932–935.

21 G. von Kiedrowski, B. Wlotzka, J. Helbing, M. Matzen and S. Jordan, *Angew. Chem. Int. Ed. Engl.*, 1991, **30**, 423–426.

22 T. Achilles and G. von Kiedrowski, *Angew. Chem. Int. Ed. Engl.*, 1993, **32**, 1198–1201.

23 A. Luther, R. Brandsch and G. von Kiedrowski, *Nature*, 1998, **396**, 245–248.

24 W. S. Zielinski and L. E. Orgel, *Nature*, 1987, **327**, 346–347.

25 N. Paul and G. F. Joyce, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 12733–12740.

26 J. Rogers and G. F. Joyce, *RNA*, 2001, **7**, 395–404.

27 D. H. Lee, J. R. Granja, J. A. Martinez, K. Severin and M. R. Ghadri, *Nature*, 1996, **382**, 525–528.

28 E. K. O'Shea, J. D. Klemm, P. S. Kim and T. Alber, *Science*, 1991, **254**, 539–544.

29 P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. Kent, *Science*, 1994, **266**, 776–779.

30 K. Severin, D. H. Lee, A. J. Kennan and M. R. Ghadiri, *Nature*, 1997, **389**, 706–709.

31 S. Yao, I. Ghosh, R. Zutshi and J. Chmielewski, *J. Am. Chem. Soc.*, 1997, **119**, 10559–10560.

32 S. Yao, I. Ghosh, R. Zutshi and J. Chmielewski, *Angew. Chem. Int. Ed. Engl.*, 1998, **37**, 478–481.

33 R. Issac and J. Chmielewski, *J. Am. Chem. Soc.*, 2002, **124**, 6808–6809.

34 X. Li and J. Chmielewski, *J. Am. Chem. Soc.*, 2003, **125**, 11820–11821.

35 (a) R. H. Yun, A. Anderson and J. Hermans, *Proteins: Struct., Funct. Genet.*, 1991, **10**, 219–228; (b) M. W. MacArthur and J. M. Thornton, *J. Mol. Biol.*, 1991, **218**, 397–412.

36 Z. Dadon, M. Samiappan, E. Y. Safranchik and G. Ashkenasy, *Chem. Eur. J.*, 2010, **16**, 12096–12099.

37 (a) C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace and S. Altman, *Cell*, 1983, **35**, 849–857; (b) K. Kruger, P. J. Grabowski, A. K. Zaug, J. Sands, D. E. Gottschling and T. R. Cech, *Cell*, 1982, **39**, 2281–2285.

38 (a) M. P. Robertson and G. F. Joyce, *Cold Spring Harb. Perspect. Biol.*, 2012, **4**:a003608; (b) G. F. Joyce, *Nature*, 2002, **418**, 214–221; (c) G. F. Joyce and L. E. Orgel, *Prospects for understanding the origin of the RNA world. In The RNA world*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1993, pp 1–25.

39 (a) A. Eschenmoser, *Orig. Life Evol. Biosph.*, 2004, **34**, 277–306; (b) K.-U. Schöning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy and A. Eschenmoser, *Science*, 2000, **290**, 1347–1351; (c) A. Eschenmoser and R. Krishnamurthy, *Pure Appl. Chem.*, 2000, **72**, 343–345; (d) A. Eschenmoser, *Science*, 1999, **284**, 2118–2124.

40 P. E. Nielsen, *Acc. Chem. Res.*, 1999, **32**, 624–630.

41 (a) P. E. Nielsen, *Chem. Biodivers.*, 2007, **4**, 1996–2002; (b) U. J. Meierhenrich, G. M. Muñoz Caro, J. H. Bredehöft, E. K. Jesseberger and W. H.-P. Thiemann, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 9182–9186; (c) K. E. Nelson, M. Levy and S. L. *Miller*, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 3868–3871.

42 *Protocells—Bridging Nonliving and Living Matter*, S. Rasmussen, M. A. Bedau, L. Chen, D. Deamer, D. C. Krakauer, N. H. Packard and P. F. Stadler, Eds., MIT Press, Cambridge, London, 2009.

43 C. Bohler, P. E. Nielsen and L. E. Orgel, *Nature*, 1995, **376**, 578–581.

44 T. A. Plöger and G. von Kiedrowski, *Org. Biomol. Chem.*, 2014, **12**, 6908–6914.

45 A. Singhal and P. E. Nielsen, *Org. Biomol. Chem.*, 2014, **12**, 6901–6907.

46 T. Tjivikua, P. Ballester and J. Rebek, Jr., *J. Am. Chem. Soc.*, 1990, **112**, 1249–1250.

47 D. S. Kemp and K. S. Petrakis, *J. Org. Chem.*, 1981, **46**, 5140–5143.

48 F. M. Menger, A. V. Eliseev and N. A. Khanjin, *J. Am. Chem. Soc.*, 1994, **116**, 3613–3614.

49 D. N. Reinhoudt, D. M. Rudkevich and F. de Jong, *J. Am. Chem. Soc.*, 1996, **118**, 6880–6889.

50 V. Rotello, J.-I. Hong and J. Rebek, Jr., *J. Am. Chem. Soc.*, 1991, **113**, 9422–9423.

51 (a) R. J. Pieters, I. Huc and J. Rebek, Jr., *Tetrahedron*, 1995, **51**, 485–498; (b) E. A. Wintner, M. M. Conn and J. Rebek, Jr., *J. Am. Chem. Soc.*, 1994, **116**, 8877–8884; (c) R. J. Pieters, I. Huc and J. Rebek, Jr., *Angew. Chem. Int. Ed. Engl.,* 1994, **33**, 1579–1581; (d) J. I. Hong, Q. Feng, V. Rotello and J. Rebek, Jr., *Science*, 1992, **255**, 848–850.

52 (a) S. Kamioka, D. Ajami and J. Rebek, Jr., *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 541–544; (b) T. K. Park, Q. Feng and J. Rebek, Jr., *J. Am. Chem. Soc.*, 1992, **114**, 4529–4532; (c) Q. Feng, T. K. Park and J. Rebek, Jr., *Science*, 1992, **256**, 1179–1180.

53 B. Wang and I. O. Sutherland, *Chem. Commun.*, 1997, 1495–1496.

54 M. Kindermann, I. Stahl, M. Reimold, W. M. Pankau and G. von Kiedrowski, *Angew. Chem. Int. Ed.*, 2005, **44**, 6750–6755.

55 F. Garcia-Tellado, S. Goswami, S.-K. Chang, S. J. Geib and A. D. Hamilton, *J. Am. Chem. Soc.*, 1990, **112**, 7393–7394.

56 A. Dieckmann, S. Beniken, S. Lorenz, N. L. Doltsinis and G. von Kiedrowski, *J. Syst. Chem.*, 2010, **1**:10.

57 (a) E. Kassianidis and D. Philp, *Chem. Commun.*, 2006, 4072–4074; (b) R. J. Pearson, E. Kassianidis, A. M. Z. Slawin and D. Philp, *Chem. Eur. J.*, 2006, **12**, 6829–6840; (c) E. Kassianidis, R. J. Pearson and D. Philp, *Chem. Eur. J.*, 2006, **12**, 8798–8812; (d) E. Kassianidis, R. J. Pearson and D. Philp, *Org. Lett.*, 2005, **7**, 3833–3836; (e) R. J. Pearson, E. Kassianidis, A. M. Z. Slawin and D. Philp, *Org. Biomol. Chem.*, 2004, **2**, 3434–3441; (f) R. J. Pearson, E. Kassianidis and D. Philp, *Tetrahedron Lett.*, 2004, **45**, 4777–4780.

58 J. M. Quayle, A. M. Slawin and D. Philp, *Tetrahedron Lett.*, 2002, **43**, 7229–7233.

59 V. C. Allen, D. Philp and N. Spencer, *Org. Lett.*, 2001, **3**, 777–780.

60 E. Kassianidis and D. Philp, *Angew. Chem. Int. Ed.*, 2006, **45**, 6344–6348.

61 D. Sievers, G. Von Kiedrowski, *Chem. Eur. J.*, 1998, **4**, 629–641.

62 D. Sievers, G. von Kiedrowski, *Nature*, 1994, **369**, 221–224.

63 D.-E. Kim and G. F. Joyce, *Chem. Biol.*, 2004, **11**, 1505–1512.

64 T. A. Lincoln and G. F. Joyce, *Science*, 2009, **323**, 1229–1232.

65 A. C. Ferretti and G. F. Joyce, *Biochemistry*, 2013, **52**, 1227–1235.

66 M. P. Robertson and G. F. Joyce, *Chem. Biol.*, 2014, **21**, 238–245.

67 (a) C. Olea and G. F. Joyce, *Methods Enzymol.*, 2015, **550**, 23–29; (b) J. T. Sczepanski and G. F. Joyce, *Nature*, **515**, 440–442; (c) C. Olea, D. P. Horning and G. F. Joyce, *J. Am. Chem. Soc.*, 2012, **134**, 8050–8053; (d) B. J. Lam and G. F. Joyce, *Nat. Biotechnol.*, 2009, **27**, 288–292.

68 S. Yao, I. Ghosh, R. Zutshi and J. Chmielewski, *Nature*, 1998, **396**, 447–450.

69 D. H. Lee, K. Severin, Y. Yokobayashi and M. R. Ghadiri, *Nature*, 1997, **390**, 591–594.

70 K. Severin, D. H. Lee, J. A. Martinez, M. Vieth and M. R. Ghadiri, *Angew. Chem. Int. Ed. Engl.*, 1998, **37**, 126–128.

71 (a) J. M. Ribó, J. Crusats, Z. El-Hachemi, A. Moyano and D. Hochberg, *Chem. Sci.,* 2017, **8**, 763–769; (b) J. M. Ribó, C. Blanco, J. Crusats, Z. El-Hachemi, D. Hochberg and A. Moyano, *Chem. Eur. J.*, 2014, **20**, 17250–17271; (c) D. G. Blackmond, *Cold Spring Harb. Perspect. Biol.*, 2010, **2**:a002147; (d) D. G. Blackmond, *Chem. Eur. J.*, 2007, **13**, 3290–3295; (e) L. Plasson, D. K. Kondepudi, H. Bersini, A. Commeyras and K. Asakura, *Chirality*, 2007, **19**, 589–600; (f) R. Plasson, H. Bersini and A. Commeyras, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 16733–16738; (g) K. Soai, T. Shibata, H. Morioka and K. Choji, *Nature*, 1995, **378**, 767–768.

72 A. Saghatelian, Y. Yokobayashi, K. Soltani and M. R. Ghadiri, *Nature*, 2001, **409**, 797–801.

73 J. Rivera Islas, J.-C. Micheau and T. Buhse, *Orig. Life Evol. Biosph.*, 2004, **34**, 497–512.

74 G. Ashkenasy, R. Jagasia, M. Yadav and M. R. Ghadiri, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 10872–10877.

75 G. Ashkenasy and M. R. Ghadiri, *J. Am. Chem. Soc.*, 2004, **126**, 11140–11141.

76 (a) L. Cardelli, R. D. Hernansaiz-Ballesteros, N. Dalchau and A. Csikász-Nagy, *PLOS Comput Biol.*, 2017, **13**: e1005100; (b) W. Kolch, M. Halasz, M. Granovskaya and B. N. Kholodenko, *Nat. Rev. Cancer*, 2015, **15**, 515–527; (c) L. Cardelli and A. Csikász-Nagy, *Scientific Reports*, 2012, **2**: 656; (d) M. Freeman, *Nature*, 2000, **408**, 313–319.

77 E. Kassianidis, R. J. Pearson, E. A. Wood and D. Philp, *Faraday Discuss.*, 2010, **145**, 235–254.

78 T. Kosikova and D. Philp, *J. Am. Chem. Soc.*, 2017, **139**, 12579–12590.

79 (a) M. Samiappan, Z. Dadon and G. Ashkenasy, *Chem. Commun.*, 2011, **47**, 710–712; (b) G. Ashkenasy, Z. Dadon, S. Alesebi and N. Wagner, *Isr. J. Chem.*, 2011, **51**, 106–117; (c) V. C. Allen, C. C. Robertson, S. M. Turega and D. Philp, *Org. Lett.*, 2010, **12**, 1920–1923.

80 (a) A. Vidonne, T. Kosikova and D. Philp, *Chem. Sci.*, 2016, **7**, 2592–2603; (b) T. Kosikova, N. I. Hassan, D. B. Cordes, A. M. Z. Slawin and D. Philp, *J. Am. Chem. Soc.*, 2015, **137**, 16074–16083; (c) A. Vidonne and D. Philp, *Tetrahedron*, 2008, **64**, 8464–8475.

81 (a) J.-M. Lehn, *Angew. Chem. Int. Ed.*, 2015, **54**, 3276–3289; (b) J. Li, P. Nowak and S. Otto, *J. Am. Chem. Soc.*, 2013, **135**, 9222–9239; (c) M. E. Belowich and J. F. Stoddart, *Chem. Soc. Rev.*, 2012, **41**, 2003–2024; (d) R. A. R. Hunt and S. Otto, *Chem. Commun.*, 2011, **47**, 847–858; (e) S. J. Rowan, S. J. Cantrill, G. R. L. Cousins, J. K. M. Sanders and J. F. Stoddart, *Angew. Chem. Int. Ed.*, 2002, **41**, 898–952.

82 (a) M. Mondal and K. H. Hirsch, *Chem. Soc. Rev.*, 2015, **44**, 2455–2488; (b) S. Otto, *Acc. Chem. Res.*, 2012, **45**, 2200–2210; (c) N. Giuseppone, *Acc. Chem. Res.*, 2012, **45**,

2178–2188; (d) P. T Corbett, J. K. M Sanders and S. Otto, *J. Am. Chem. Soc.*, 2005, **127**, 9390–9392; (e) J. D. Cheeseman, A. D. Corbett, J. L. Gleason and R. J. Kazlauskas, *Chem. Eur. J.*, 2005, **11**, 1708–1716; (f) B. Brisig, J. K. M. Sanders and S. Otto, *Angew. Chem. Int. Ed.*, 2003, **42**, 1270–1273; (g) P. T. Corbett, S. Otto and J. K. M. Sanders, *Org. Lett.*, 2004, **6**, 1825–1827.

83 (a) J. J. Armao IV and J.-M. Lehn, *Angew. Chem. Int. Ed.*, 2016, **55**, 13450–13454; (b) Q. Ji and O. S. Miljanić, *J. Org. Chem.*, 2013, **78**, 12710–12716; (c) M. Sakulsombat, Y. Zhang and O. Ramström, *Top. Curr. Chem.*, 2012, **322**, 55–86; (d) P. Vongvilai, M. Angelin, R. Larsson, O. Ramström, *Angew. Chem. Int. Ed.*, 2007, **46**, 948–950; (e) J. D. Cheeseman, A. D. Corbett, R. Shu, J. Croteau, J. L. Gleason and R. J. Kazlauskas, *J. Am. Chem. Soc.*, 2002, **124**, 5692–5701.

84 (a) E. Moulin and N. Giuseppone, *Top. Curr. Chem.*, 2012, **322**, 87–106; (b) V. del Amo and D. Philp, *Chem. Eur. J.*, 2010, **16**, 13304–13318.

85 Z. Dadon, M. Samiappan, N. Wagner and G. Ashkenasy, *Chem. Commun.*, 2012, **48**, 1419 –1421.

86 Z. Dadon, N. Wagner, S. Alasibi, M. Samiappan, R. Mukherjee and G. Ashkenasy, *Chem. Eur. J.*, 2015, **21**, 648–654.

87 J. M. A. Carnall, C. A. Waudby, A. M. Belenguer, M. C. A. Stuart, J. J.-P. Peyralans and S. Otto, *Science*, 2010, **327**, 1502–1506.

88 G. Leonetti and S. Otto, *J. Am. Chem. Soc.*, 2015, **137**, 2067–2072.

89 A. Pal, M. Malakoutikhah, G. Leonetti, M. Tezcan, M. Colomb-Delsuc, V. D. Nguyen, J. van der Gucht and S. Otto, *Angew. Chem. Int. Ed.*, 2015, **54**, 7852–7856.

90 J. W. Sadownik, E. Mattia, P. Nowak and S. Otto, *Nat. Chem.*, 2016, **8**, 264–269.

91 M. Malakoutikhah, J. J.-P. Peyralans, M. Colomb-Delsuc, H. Fanlo-Virgós, M. C. A. Stuart and S. Otto, *J. Am. Chem. Soc.*, 2013, **135**, 18406–18417.

92 P. W. J. M. Frederix, J. Idé, Y. Altay, G. Schaeffer, M. Surin, D. Beljonne, A. S. Bondarenko, T. L. C. Jansen, S. Otto and S. J. Marrink, *ACS Nano*, 2017, **11**, 7858–7868.

93 E. Mattia, A. Pal, G. Leonetti and S. Otto, *Synlett*, 2017, **28**, 103–107.

94 For specific examples, see: (a) K. Ikeda and M. Nakano, *Langmuir*, 2015, **31**, 17–21; (b) M. D. Hardy, J. Yang, J. Selimkhanov, C. M. Cole, L. S. Tsimring and N. K. Devaraj, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112,** 8187–8192; (c) M. Li, X. Huang and S. Mann, *Small*, 2014, **10**, 3291–3298; (d) J. Li, I. Cvrtila, M. Colomb-Delsuc, E. Otten and S. Otto, *Chem. Eur. J.*, 2014, **20**, 15709–15714; (e) I. Nyrkova, E. Moulin, J. J. Armao, M. Maaloum, B. Heinrich, M. Rawiso, F. Niess J. J. Cid, N. Jouault, E. Buhler, N. G. Semenov and N. Giuseppone, *ACS Nano*, 2014, **8**, 10111–10124; (f) K. Adamala and J. W. Szostak, *Nat. Chem.*, 2013, **5**, 495–501; (g) B. Rubinov, N. Wagner, H. Rapaport and G. Ashkenasy, *Angew. Chem. Int. Ed.*, 2009, **48**, 6683–6686;

95 For review materials, see: (a) A. J. Bissette and S. P. Fletcher, *Orig. Life Evol. Biosh.*, 2015, **45**, 21–30; (b) J. C. Blain and J. W. Szostak, *Annu. Rev. Biochem.*, 2014, **83**, 615–640; (c) N. Giuseppone, *Acc. Chem. Res.*, 2012, **45**, 2178–2188; (d) P. Razeto-Barry, *Orig. Life Evol. Biosph.*, 2012, **42**, 543–567.

96 A. Terfort and G. von Kiedrowski, *Angew. Chem. Int. Ed.*, 1992, **31**, 654–656.

97 S. Xu and N. Giuseppone, *J. Am. Chem. Soc.*, 2008, **130**, 1826–1827.

98 V. del Amo, A. M. Z. Slawin and D. Philp, *Org. Lett.*, 2008, **10**, 4589–4592.

99 (a) T. Kosikova, H. Mackenzie and D. Philp, *Chem. Eur. J.*, 2016, **22**, 1831–1839; (b) J. W. Sadownik and D. Philp, *Org. Biomol. Chem.*, 2015, **13**, 10392–10401.

100 J. W. Sadownik and D. Philp, *Angew. Chem. Int. Ed.*, 2008, **47**, 9965–9970.

101 (a) S. N. Semenov, L. J. Kraft, A. Ainla, M. Zhao, M. Baghbanzedeh, V. E. Campbell, K. Kang, J. M. Fox and G. M. Whitesides, *Nature*, 2016, **537**, 656–660; (b) P. Schuster, K. Sigmund and R. Wolf, *J. Math. Anal. Appl.*, 1980, **78**, 88–112.

102 (a) S. Soh, M. Byrska, K. Kandere-Grzybowska and B. A. Grzybowski, *Angew. Chem. Int. Ed.*, 2010, **49**, 4170–4198; (b) B. Hess, *Naturwissenschaften*, 2000, **87**, 199–211.

103 (a) J. Slingo and T. Palmer, *Philos. Trans. A, Math Phys. Eng. Sci.*, 2011, **369**, 4751–4767; (b) E. N. Lorenz, *J. Atmos. Sci.*, 1963, **20**, 130–141.

104 (a) S. Kondo and T. Miura, *Science*, 2010, **329**, 1616–1620; (b) S. Kondo, *Genes Cells*, 2002, **7**, 535–541; (c) P. Ball, *The Self-made Tapestry: Pattern Formation in Nature*, Oxford University Press Inc., New York, 2001.

105 (a) S. C. Müller and J. Ross, *J. Phys. Chem. A*, 2003, **107**, 7997–8008; (b) R. E. Liesegang, *Naturwissenschaftliche Wochenschrift*, 1896, **11**, 353–362.

106 (a) K. Showalter and I. R. Epstein, *Chaos*, 2015, **25**, 097613; (b) *Chemical waves and patterns*, R. Kapral and K. Showalter, Eds., Kluwer Academic Publishers, Berlin, 2012; (c) B. A. Grzybowski, *Chemistry in Motion*, John Wiley & Sons, Ltd.: New York, 2009; (d) V. K. Vanag and I. R. Epstein, *Int. J. Dev. Biol.*, 2009, **53**, 673–681; (e) M. C. Cross and P. C. Hohenberg, *Rev. Mod. Phys.*, 1993, **65**, 851–1112.

107 (a) R. J. Field, E. Körös and R. M. Noyes, *J. Am. Chem. Soc.*, 1972, **94**, 8649–8664; (b) R. M. Noyes, R. J. Field and E. Körös, *J. Am. Chem. Soc.*, 1972, **94**, 1394–1395; (c) A. N. Zaikin and A. M. Zhabotinsky, *Nature*, 1970, **225**, 535–537; (d) A. M. Zhabotinsky, *Biofizika*, 1964, **9**, 306–311; (e) B. P. Belousov, *Sb. Ref. po Radiatsionni Meditsine*, 1958, 145.

108 E. Boga, G. Peintler and I. Nagypál, *J. Am. Chem. Soc.*, 1990, **112**, 151–153.

109 (a) J. S. McCaskill and G. J. Bauer, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 4191–4195; (b) G. J. Bauer, J. S. McCaskill and H. Otten, *Proc. Natl. Acad. Sci. U. S. A.*, 1989, **86**, 7937–7941.

110 A. Padirac, T. Fujii, A. Estévez-Torres and Y. Rondelez, *J. Am. Chem. Soc.*, 2013, **135**, 14586–14592.

111 T. Fujii and Y. Rondelez, *ACS Nano*, 2013, **7**, 27–34.

112 A. S. Zadorin, Y. Rondelez, J.-C. Galas and A. Estévez-Torres, *Phys. Rev. Lett.*, 2015, **114**, 068301.

113 K. Montagne, R. Plasson, Y. Sakai, T. Fujii and Y. Rondelez, *Mol. Syst. Biol.,* 2011, **7**:466.

114 J. H. Merkin, A. J. Poole, S. K. Scott, J. Masere, J. and K. Showalter, *J. Chem. Soc., Faraday Trans.*, 1998, **94**, 53–57.

115 I. Bottero, J. Huck, T. Kosikova and D. Philp, *J. Am. Chem. Soc.*, 2016, **138**, 6723–6726.

116 (a) J. F. Stoddart, *Angew. Chem. Int. Ed.*, 2017, **56**, 11094–11125; (b) J.-P. Sauvage, *Angew. Chem. Int. Ed.*, 2017, **56**, 11080–11093; (c) B. L. Feringa, *Angew. Chem. Int. Ed.*, 2017, **56**, 11060–11078; (d) J.-M. Lehn, *Angew. Chem. Int. Ed.*, 1988, **27**, 89–112; (e) D. J. Cram, *Angew. Chem. Int. Ed.*, 1988, **27**, 1009–1020; (f) C. J. Pedersen, *Angew. Chem. Int. Ed.*, 1988, **27**, 1021–1027.

117 (a) R. Pascal and A. Pross, *Synlett.*, 2017, **28**, 30–35; (b) R. Pascal and A. Pross, *Chem. Commun.*, 2015, **51**, 16160–16165; (c) R. Pascal and A. Pross, *J. Syst. Chem.*, 2014, **5**:3; (d) A. Pross, *Curr. Org. Chem.*, 2013, **17**, 1702–1703; (e) A. Pross, *J. Syst. Chem.*, 2011, **2**:1.

118 (a) J. W. Taylor, S. A. Eghtesadi, L. J. Points, T. Liu and L. Cronin, *Nat. Commun.*, 2017, **8**, 237; (b) S. Matsumura, Á. Kun, M. Ryckelynck, F. Coldren, A. Szilágyi, F. Jossinet, C. Rick, P. Nghe, E. Szathmáry and A. D. Griffiths, *Science*, 2016, **354**, 1293–1296; (c) J. W. Szostak, D. P. Bartel and P. L. Luisi, *Nature*, 2001, **409**, 387–390; (d) F. J. Ghadessy, J. L. Ong and P. Holliger, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **98**, 4552–4557.

119 (a) A. Dieckmann and K. N. Houk, *Chem. Sci.*, 2013, **4**, 3591–3600; (b) P. V. Coveney, J. B. Swadling, J. A. D. Wattis and H. C. Greenwell, *Chem. Soc. Rev.*, 2012, **41**, 5430–5446.