Mechanisms of Autocatalysis
Andrew J. Bissette and Stephen P. Fletcher*

Keywords:
asymmetric catalysis - autocatalysis -
molecular recognition -
origins of life - self-replication
Self-replication is a fundamental concept. The idea of an entity that can repeatedly create more of itself has captured the imagination of many thinkers from von Neumann to Vonnegut. Beyond the sciences and science fiction, autocatalysis has found currency in economics and language theory, and has raised ethical fears memorably summed up by the “gray goo” trope. Autocatalysis is central to the propagation of life and intrinsic to many other biological processes. This includes the modern conception of evolution, which has radically altered humanity’s image of itself. Organisms can be thought of as imperfect self-replicators which produce closely-related species, allowing for selection and evolution. Hence, any consideration of self-replication raises one of the most profound questions of all: what is life? Minimal self-replicating systems have been studied with the aim of understanding the principles underlying living systems, allowing us to refine our concepts of biological fitness and chemical stability, self-organization and emergence, and ultimately to discover how chemistry may become biology.

1. Introduction

The concept of self-replication has in various guises inspired science, philosophy, literature, and even fear.[1–5] Autocatalytic chemical reactions have been studied for over a century, and it is widely accepted that they must have played a key role in the emergence of life.[6] In turn, studies of the origins of life impact upon synthetic biology and the definition of life itself. The RNA world hypothesis is a classic example of an origins of life scenario which gives a central role to autocatalysis.[8] This “replicators-first” scenario posits that prebiotic chemical processes generated autocatalytic oligomers of RNA. A population of closely-related oligomers would then arise through imperfect replication, allowing for “unnatural selection”[9] to act upon the population. This would lead to evolution[10] and then presumably to life as we know it.

As an alternative to roles in “replicators-first” scenarios, autocatalytic processes have been proposed as a source of prebiotic “building blocks”, a cause of biological homochirality, and a kind of prebiotic metabolism.[11] Szathmáry has postulated that evolution of autocatalytic reaction networks preceded polymer-based replicators.[7] Conversely, Pross has proposed that prebiotic “metabolism” arose from “replicator-first” scenarios.[12]

1.1. Scope

Here we survey chemical systems known to be capable of autocatalysis, and organize them by their mechanism of operation. Mechanism is central both to understanding the scope, limits, and applications of these systems, and to elucidating the principles underlying autocatalysis in the lab and in nature. The chemistry ranges from the mundane to the exquisite, from acid-catalyzed ester hydrolysis (Section 2) to complex networks of self-replicating peptides (Section 3.3). Despite the diversity of chemistries and mechanisms at work, many of these reactions share common principles and themes. We hope that this will provide a useful resource and offer insight into the design of systems operating by these, or novel, mechanisms.

This Review is broadly split into systems operating by a template mechanism and those using less well-defined physical processes. The special case of absolute asymmetric autocatalysis is discussed separately. Special attention is paid to prebiotically relevant processes. Although many exciting enzyme-based systems are known, such as the network of ribozymes recently reported by Vaidya et al.,[13] we do not discuss these here and refer readers to the review by Meyer et al.[14]

1.2. Definitions

An autocatalytic reaction is one in which the product acts as the catalyst for its own formation (Scheme 1). Autocatalysis is usually demonstrated in two ways (Scheme 2): by the presence of an exponential product/time curve, and by a positive correlation between initial product concentration and time. In the case of Scheme 2, a positive correlation indicates autocatalysis. The general form of an autocatalytic reaction is shown in Scheme 1.

Scheme 1. General autocatalytic reaction. Compound C catalyzes a chemical reaction in which a second molecule of C is formed.

[^] A. J. Bissette, Dr. S. P. Fletcher Chemistry Research Laboratory, University of Oxford 12 Mansfield Rd., Oxford, OX1 3TA (UK) E-mail: stephen.fletcher@chem.ox.ac.uk Homepage: http://fletcher.chem.ox.ac.uk
and reaction rate. Both effects arise because the rate of reaction is proportional to initial concentration of product. Often, autocatalytic efficiency is limited and fully exponential growth is not realized, but rather parabolic growth is seen. Only very inefficient systems are expected to show linear growth.

In this Review we use the term “exponential” to refer to very efficient systems, and “parabolic” to refer to those with limited efficiency. Where there is ambiguity we use the term “sigmoidal”. We mean this qualitatively, as both exponential and parabolic plots share an S-shaped profile. Refer to Section 3 for a more detailed discussion of these terms.

Additionally, other related forms of catalysis are relevant (Scheme 3). Cross-catalysis involves multiple species which catalyze each other’s formation. This is often a symmetrical relationship, called mutual catalysis, with two species cross-catalyzing the formation of each other. However, it is not necessarily symmetrical: one species may catalyze a competitor’s formation without its own formation being mutually catalyzed by the competitor. A species which catalyzes only its own formation, and not that of a closely-related competitor, is termed a “selfish” autocatalyst. The importance of cross-catalysis in the origins of life has recently been emphasized.[7,15]

Two autocatalysts which are mutually catalytic form a hypercycle. It has been proposed by Eigen and Schuster that hypercycles had a key role to play in the origins of life and that vestiges of hypercyclic structure exist in modern biology.[16] The stability of hypercycles under prebiotic conditions has been disputed.[17]

Blackmond has distinguished between true autocatalysis and autoinduction (Scheme 4). Here, “autoinduction” refers to processes in which a reaction product “accelerates the rate of a kinetically meaningful step of a reaction sequence without directly producing more of itself”. [18] Only a truly autocatalytic cycle can persist independently of secondary reaction cycles and their catalysts. We will focus on truly autocatalytic reactions.

Autocatalysis has been implicated in the origin of biological homochirality, most notably by the Frank model[19] (Scheme 5). This postulates mutual antagonism...
between enantiomers of an autocatalyst. If each enantiomer is autocatalytically active, but heterochiral dimers of the autocatalyst are inactive, then very small \( e e \) values can be amplified to high \( e e \). Small \( e e /C29s \) may arise stochastically or from a variety of physical sources. [20] The Frank model has been experimentally realized by Soai and co-workers (Section 5.1).

### 2. Examples of Autocatalysis

The literature is rich with diverse examples of autocatalysis in common chemical transformations. We will not attempt to cover these comprehensively, but rather to give a brief overview of the major themes.

Several methods of signal amplification rely upon autocatalysis (reviewed by Scrimin and Prins [22]). Minute quantities of an analyte generate autocatalytic products, which multiply independently of analyte concentration. These processes typically produce a fluorophore or chromophore with a distinct signal, which the autocatalytic cycle amplifies to measurable levels (Scheme 6).

Decomposition processes often exhibit autocatalysis. [23] Indeed, some of the earliest reports of autocatalysis involve such reactions. [24] A common example is acid-catalyzed ester hydrolysis, where the product of the hydrolysis is itself an acid. Autocatalysis has also been implicated in synthetically-useful reactions, including esterification of alcohols and anhydrides. [25] Oxidative addition of aryl bromides to \( Pd^0 \) complexes, [26] and C-H activation by Ru \( ^{III} \). [29]

Inorganic redox chemistry commonly involves autocatalysis. The famous Belousov-Zhabotinsky chemical oscillator is a notable example. [30] Heterogeneous processes often exhibit autocatalysis due to nucleation. Similarly, reactions at surfaces often depend upon surface area, and reactions which increase surface area can proceed autocatalytically. [31] Mondloch et al. have reviewed autocatalysis in nanoparticle formation. [32]

For similar reasons, autocatalysis occurs in crystal growth, as the rate of growth depends upon the surface area of the crystal. Importantly, this can lead to spontaneous resolution of racemic solutions by asymmetric autocatalysis. [33] As this involves a phase change rather than bond-forming or bond breaking processes, we will not discuss this in detail and refer the reader to Weissbuch and Lahav. [34] For an interesting application see Schulman et al. [35]

#### 2.1. The Formose Reaction

Perhaps the best-known autocatalytic reaction is the formose reaction, discovered by Butlerow in 1861. [36] At high pH, formaldehyde (1) condenses to form a complex mixture of sugars (Scheme 7). The reaction is unselective, giving products with varying numbers of carbon atoms and stereochemistries. Eventually “browning” occurs and an intractable polymeric mixture is obtained.

This reaction has long been favored as a prebiotic source of sugars. Due to its poor selectivity, other prebiotic sources of sugars that offer higher selectivity have been proposed, notably Eschenmoser’s glyoxylate scenario. [39]
Hein and Blackmond recently discussed the role of the formose reaction in the emergence of biomolecular homochirality and the prebiotic synthesis of sugars and nucleotides.

Breslow found that the reaction is initiated by the autocatalytic conversion of formaldehyde 1 to glycolaldehyde 2. Two observations suggest an autocatalytic mechanism: first, the rate of the reaction has a long initiation period where little reaction occurs, followed by rapid conversion. Second, this lag period is eliminated by addition of an initiator, including products of the reaction such as 2.

The initial slow conversion of 1 to 2 was proposed to involve a formyl anion. This was contested by Socha et al. who showed that small contaminants are required to initiate the reaction, and that pure formaldehyde is unreactive. In this view, dimerization of 1 does not occur directly, but is catalyzed by higher sugars. However, in the presence of minerals such as montmorillonite, freshly-distilled formaldehyde does undergo the formose reaction. Despite these observations, the view that the reaction is initiated by the formyl anion remains common in the literature.

Once the autocatalytic cycle is established, the product 2 can either react further in the autocatalytic pathway or go on to produce higher sugars. These in turn can react directly with 1 and 2. At long reaction times, 2 is consumed. Autocatalysis by 2 is therefore ultimately self-destructive.

The selectivity of the reaction can be improved somewhat using potentially prebiotically-relevant methods. Examples include carrying out the reaction in vesicles or in the presence of silicates or borates.

3. Template Replication

The formose reaction can be thought of as autocatalysis out of control: while some selectivity can be imposed, the result is always a complex mixture of products. In search of selective autocatalytic reactions, researchers have been inspired by the replication of biological polymers.

Template-based replicators share a common mechanism. A template molecule recognizes two or more fragments and binds reversibly to them (Scheme 8). This brings the fragments together, increasing their effective concentration and catalyzing their coupling. This coupling produces a second template molecule and hence the reaction is autocatalytic.

Equation (1) shows the general form of a rate equation for template-based autocatalysis, where \( c \) is the concentration of product, \( b \) is the non-autocatalytic term, and \( p \) is the order of reaction with respect to product:

\[
\frac{dc}{dt} = ac^p + b
\]

Ideally template autocatalysis will lead to an exponential increase in product concentration, that is, \( p = 1 \). We refer to this condition as “exponential growth”. Drawing on the work of Szathmáry, von Kiedrowski proposed that exponential replication is a prerequisite for natural selection.

Template-based self-replication is complicated by product inhibition, wherein template molecules associate into dimers or multimers that are catalytically inactive. The reduced catalytic efficiency of the template, reduce the efficiency of replication by a complementary trinucleotide (7 and 8) is catalyzed by a complementary hexamer (9), which is itself the product of the reaction. This
work was preceded by studies of template synthesis from monomers by Orgel,\[51\] which showed low efficiency and did not demonstrate a full cycle of replication. The use of trimers overcame this problem.

Study of analogous systems\[52\] found that efficiency is highest when the trimers and hexamer are fully complementary, strongly suggesting that a template mechanism is operating. Orgel developed a similar system using dinucleotides 10 and 11 to form a tetramer 12 (Scheme 10).\[53\] Unlike von Kiedrowski’s first system, this reaction used an amine nucleophile to form a phosphoramidate backbone.

In these early systems the expected parabolic growth was masked by strong product inhibition and inefficient catalysis.\[49\] The first demonstration of parabolic growth adapted Orgel’s phosphoramidate chemistry to give a 75-fold increase in efficiency\[54\] (Scheme 11).

A self-replicating oligonucleotide which proceeds through a slightly different mechanism was reported by Nicolaou and Li (Scheme 12).\[55\] A short palindromic oligonucleotide associates into a duplex, which acts as the template for the formation of a third strand from two fragments. Subsequent work on self-replicating peptides (Section 3.2) has shown parallels to this mechanism, but we are not aware of any other oligonucleotide-based replicators that proceed through a triplex.

Reaching exponential replication in oligonucleotide-based systems has proven difficult,\[57\] but was achieved using a manual cycling protocol termed SPREAD (Scheme 13).\[58\] By attaching oligonucleotides to a solid support and manually melting and annealing the strands, product inhibition could be avoided, enabling exponential replication. This manual cycling of a self-replicating system has a parallel in the work of Szostak\[59\] on the manual cycling of vesicle self-reproduction (Section 4.2). Exponential self-replication without enzymes has not yet been realized in free oligonucleotides.

Efforts to polymerize RNA under prebiotic conditions have often been limited by poor selectivity between 2'-5' and

---

**Scheme 9.** First non-enzymatic self-replicating system. 9 catalyzes the ligation of 7 and 8.

**Scheme 10.** Orgel’s system uses phosphoramidate chemistry to drive the autocatalytic formation of 12.

**Scheme 11.** The use of phosphoramidate chemistry allows more efficient growth.

**Scheme 12.** Nicolaou’s self-replicating oligonucleotide. In this system the template is duplex DNA, offering a different mechanism than other oligonucleotide replicators. From Ref. [56] with permission.
3′-5′ linkages. Szostak recently reported that this heterogeneity actually reduces product inhibition and frees the strands for further replication. This demonstrates the relevance of studies of self-replication to ongoing work and prebiotic chemistry.

Further studies also have biological relevance. Von Kiedrowski demonstrated selection between competing replicators (Scheme 14). The template was split into three parts: a trimer (A), a dimer (B), and a monomer (C). These recombined into five products: a trimer (BC), two tetramers (AC and BB), a pentamer (AB), and the original hexamer (ABC).

Products AC, AB, and ABC were autocatalysts, and AC was the most active. However, no other species promote the formation of AC. When all components are present, cross-catalysis between the self-replicating pentamer AB and hexamer ABC allow these species to promote their own formation at the expense of AC. Ghadiri suggested that this forms a minimal hypercycle. This intriguing example of primitive selection convincingly demonstrates that autocatalysts can alter product distributions in their own favor.

An unusual feature of this system is the ability of AB5 to act as a catalyst for its own reaction with C1 to produce hexamer ABC5. This is not an autocatalytic process, but is unusual and interesting.

Cross-catalysis was also addressed (Scheme 15). Four hexamers AB', BB', AA', and BA' were synthesized from four common starting materials, creating competition. AB' and BA' were selfish autocatalysts, while AA' and BB' were autocatalytically inert but capable of mutual cross-catalysis. When all starting materials were combined, all four hexamers formed at similar rates. These studies suggest that cross-catalytic replication can co-exist with, and be as efficient as, autocatalysis.
3.2. Helical Peptides

Several groups have reported self-replicating peptides (typically around 30 residues long) with superhelical “coiled-coil” structures consisting of two intertwined α-helices. Hydrophobic forces create molecular recognition, driving templated ligation of two shorter peptides (Scheme 16). The parabolic growth shown by these systems suggests that product inhibition occurs. Several reports have suggested that in some of these systems, dimers or even trimers of product may be catalytically active.\cite{65,66} This resembles the unusual nucleotide-based system reported by Nicolaou et al. (Section 3.1).

The first self-replicating peptide was reported by Ghadiri et al.\cite{67} and two closely-related peptides were soon reported by Chmielewski et al.\cite{65,68} The latter are pH sensitive: one is only active at low pH, when key glutamate residues are protonated (Scheme 17), and the other is active at high pH or high ionic strength, due to deprotonation of key lysine residues.

Environmental control was also reported by Ashkenasy et al.\cite{70} A photolabile group was bound to a lysine residue, rendering the peptide catalytically inactive until freed by photolysis (Scheme 18). Analogues that act as a NAND gate,\cite{71} or direct a dynamic combinatorial library (DCL) of replicators, have also been reported.\cite{72}

Matsumura et al. incorporated nucleobase-substituted amino acids in an attempt to increase selectivity (Scheme 19). Peptides with different patterns of nucleobase incorporation were prepared, allowing the secondary structure to be modified (e.g. parallel vs. antiparallel coils). This achieved a 3- to 5-fold increase in efficiency over unmodified replicators due to base pairing, although it also increased product inhibition somewhat.

Chmielewski reported impressive steps towards exponential replication by shortening the peptide\cite{74} or introducing a proline “kink”.\cite{75} The reaction order for one peptide, described as “weakly exponential”, was calculated as 0.91 ± 0.04 (Scheme 20).\cite{74} These changes destabilized product dimers and hence reduced product inhibition. Exponential peptide replication has since been realized using a different class of peptides (Section 3.3).

Both Ghadiri and Chmielewski have incorporated self-replicating peptides into more complex systems. This section aims to provide general insight, rather than an exhaustive discussion of each system.

Ghadiri’s replicator is sensitive to changes in even a single residue. While it presents a challenge for the de novo design of similar systems, this sensitivity can be employed as an error correction mechanism (Scheme 21).\cite{76} Three “mutant” pep-
tides were prepared, each with a single residue altered (two, T9A and T26A, with one mutant fragment each, and one, T9/26A, with both mutant fragments). The mutant peptides were incapable of autocatalysis, but T9A and T26A were able to catalyze the formation of the native peptide T. T did not catalyze the formation of any mutants, and the double mutant T9/26A was entirely inactive. Due to these relationships, the native template consistently dominated reaction mixtures.

This selectivity has also been used as a means of chiral selection (Scheme 22).[77] Ghadiri et al. originally used homochiral peptides, in which all the amino acids shared a common absolute configuration. When two homochiral fragments of opposite chirality were combined, a heterochiral peptide was produced. Only homochiral peptides were autocatalytic; heterochiral peptides did not promote the formation of any species.

Even the substitution of a single L-amino acid for its D-enantiomer impaired autocatalysis. Peptides with one L/D substitution could catalyze the formation of the homochiral peptide at their own expense. Consequently the homochiral products could be selectively produced from a racemic mixture of fragments, and “stereoechemical mutations” at single residues were selected out. The end point of this system was a racemic mixture of two dominant, enantiomerically enriched homochiral peptides.[77]

This mechanism presents an alternative role for autocatalysis in the origin of biological homochirality other than that posited by Frank (Section 1.2). It may be that life before the appearance of biological polymers was not homochiral, and error correction mechanisms like those discussed above led to the evolution of homochiral polymers from a weakly enantiomerically enriched population of autocatalytic polymers.[21]

Ghadiri’s peptides have been incorporated into a hypercycle (Scheme 23).[78] A replicator and a more efficient analogue compete over a common starting material. However, the two are mutually cross-catalytic, creating a hypercycle (Scheme 3). This prevents the more efficient competitor from dominating the system.

A more complicated system was reported by Chmielewski (Scheme 24).[79] The cross-catalytic relationships of four autocatalytic peptides that share starting materials were elucidated. Several unexpected relationships were identified. Cross-catalysis allowed the coexistence of active and inactive replicators, and pH control allowed selective production of pH-activated replicators.

Ghadiri et al. have reported networks of even greater complexity.[80,81] The auto- and cross-catalytic capabilities of 81 peptides were predicted theoretically. The 25 expected to be active were incorporated into a network (Scheme 25). A subset of 9 peptides was studied experimentally; interestingly,
some predicted pathways were shown to occur in isolation, but to be suppressed when the subset was considered as a whole. Sections of the network were also shown to behave as Boolean logic gates when “programmed” with fragments and templates. [83]

This study directly supports the notion that mixtures of closely-related autocatalysts can form organized networks with features similar to those found in biological systems. Behavior of this kind has been implicated in models of chemical evolution, including the RNA world. [11]

3.3. Non-Helical Peptides

Systems of self-replicating peptides that form β-sheets have been reported. Here product inhibition does not occur as the newly-ligated peptide does not need to dissociate from the template in order to act as a catalyst. This mechanism resembles autocatalysis in crystal growth (Section 2).

Takahashi and Mihara reported self-replicating amyloid fibrils (Scheme 26). [82] The surface of these regularly-aligned β-sheets acts as a template for the ligation of smaller fragments. As the ligation only occurs at the ends of fibrils, the area available for catalysis remains constant and efficiency is limited. As the newly-synthesized peptide adopts the amyloid conformation, this process is described as conformational replication.

A related system was reported by Ashkenasy et al. [83] A 12-residue peptide made of alternating pairs of hydrophobic and hydrophilic residues assembles into one-dimensional β-sheets in water (Scheme 27). These are capable of catalyzing the ligation of fragments into fresh sheets. Other mechanisms such as non-specific binding and molecular crowding were ruled out with an analogous, inactive peptide.

Further study found that the β-sheets assemble into larger supramolecular structures including fibers and nanotubes (Scheme 28). [84] The fibers, which are believed to be the catalytically-active species, were shown to exist only transiently during this process. This is the first self-replicating peptide (that we know of) that exhibits exponential growth.
Otto reported an unusual system\textsuperscript{[85]} (Scheme 29) where a DCL of amphiphilic pentapeptides forms a mixture of macrocyclic oligomers through disulfide bond formation. These macrocycles aggregate into $\beta$-sheets, which grow as fibers. Remarkably, fibers composed of different-sized macrocycles had different responses to mechanical stimulation. Shaking the solution led to exponential growth in hexamer-based fibers. Shaking selectively fragments these fibers, increasing the area available for growth. However, stirring the solution disrupts fibers composed of either hexamers or heptamers. As the heptamers react faster at fiber ends than the hexamers do, the heptamers dominate the system when stirred. Thus despite the system being under thermodynamic control, the product distribution is determined kinetically. Otto described this as "a first step toward the far-from-equilibrium character of life."\textsuperscript{[86]}

### 3.4. Small Molecules

The design of non-natural molecules capable of templated self-replication presents a great challenge to chemists. The recognition sites must be sufficiently separate to prevent the template from simply binding to itself, but close enough to allow efficient ligation. The template must bind strongly to the fragments to allow efficient catalysis, but weakly to itself to prevent product inhibition. These are demanding constraints on the design of self-replicating molecules.

Non-enzymatic self-replicating molecules were reported by Rebek and co-workers beginning in 1990. Two classes of replicators were reported, differentiated by the recognition chemistry driving autocatalysis. The first includes adenine-derived recognition sites, and the second uses thymine-derived sites. In both classes, the key reaction is amide bond formation.

The first class used complementary adenine and Kemp’s imide groups, and went through two generations of development. The first system had a naphthyl spacer between the recognition and reaction sites (Scheme 30, top).\textsuperscript{[87]} The naphthyl spacer was small enough that, in the absence of template, 14 and 15 could pre-associate and catalyze the background reaction.\textsuperscript{[88]} By increasing the size of the spacer (Scheme 30, bottom), this pathway was inhibited sufficiently to reveal parabolic growth (Scheme 31).\textsuperscript{[89]}

A second generation of this system was developed with the aim of further restricting the preassociative pathway (Scheme 32).\textsuperscript{[90]} The spacer was expanded and a second imide was added to strengthen binding. While this did eliminate the preassociative pathway, it also increased product inhibition, and exponential growth was not achieved.

The second class of systems uses recognition between a thymine derivative and a diaminotriazine to drive amide bond formation.\textsuperscript{[91]} This system showed parabolic growth (Scheme 33).

The proposed template mechanism was disputed by Menger et al.,\textsuperscript{[92,93]} who proposed that the amide bond in the product could be catalytically active without the need for molecular recognition. Kinetic analysis by Reinhoudt\textsuperscript{[94]}
identified five reactions: the background bimolecular reaction, the desired template autocatalysis, pre-association between the starting materials, and two pseudo-bimolecular reactions involving binding of one or the other fragment to the template followed by ligation. The relative importance of each pathway was shown to be concentration-dependent, explaining the differing results of Rebek and Menger’s work. Crucially, the pathway proposed by Menger was found to be negligible at the concentrations used by Rebek to demonstrate autocatalysis, supporting the template mechanism over simple amide catalysis.

The crossover products of the two classes showed interesting behavior (Scheme 34). One product, 22, was the most active replicator known at the time, while the other, 25, was inactive due to the unusual S-shaped conformation of the termolecular complex [23-24-25].

Rebek’s replicators have been incorporated into more complex systems. Variants of the thymine-derived replicator were used as a very simple model of hypothesized “RNA


Scheme 32. Rebek’s second-generation self-replicating system.

Scheme 33. Rebek’s second class of self-replicating molecules.

Scheme 34. The cross-over products of Rebek’s systems generate an active replicator 22 and inactive product 25. The termolecular complex of 23, 24 and 25 is S-shaped, preventing catalysis.
world” self-replicating ribozymes. These variants can catalyze both their own synthesis and orthogonal reactions (Scheme 35). The ligation chemistry was modified to create a new catalytically-active functional group in the product. Two systems were reported, both capable of catalyzing 1,4-conjugate additions; the first contains a Jacobsen-type catalyst (26) and the second contains an imidazolidinone analogue (27). The adenine-derived repli-
cators were used to demon-
strate biologically-relevant processes including cross-catalysis[98,99] mutation, and competition[100] (Scheme 36). In the latter study, the adenine group was modified by a phenyl (28c) or nitrophenyl (28b/30b) group, creating inefficient replicators due to restricted hydrogen bonding with the imide. Photolysis of 28b/30b to the more efficient native replicator 28a/30a simulates mutation; 30a consumes the common starting material 29 and dominates the reaction, simulating competition.

A number of autocatalytic cycloaddition reactions have been reported. The first, a Diels–Alder reaction reported by Wang and Sutherland,[101] was later studied in depth by von Kiedrowski using closely related analogues (Scheme 37).[102,103] These reactions are efficient, approaching exponential replication.

Here, the enantiomers of the diastereoselective replicators 36a and 36b form a hypercycle: each autocatalyst cross-catalyzes the formation of its enantiomer. This precludes asymmetric autocatalysis: adding either enantiomer to racemic starting material will give racemic product by cross-catalysis. Asymmetric autocatalytic reactions are a special class of self-replicators discussed in Section 5.

Replicators with a fulvene skeleton simplified analysis by avoiding the creation of enantiomeric products (Scheme 38).[104]
There are four possible cycloaddition products of this reaction: two endo (37a and 37b) and two exo (37c and 37d) products, differing in the cis/trans orientation of the aminopyridine group relative to the newly-formed bridge. The exo products are catalytically inactive. One endo product (37a) is autocatalytic, while the other (37b) is both autocatalytic and cross-catalytic: it is capable of promoting the formation of 37a as well as itself. As a result, 37a dominates the product distribution at extended reaction times. These studies highlight the sensitivity of replicators to small changes in structure.

Philp and co-workers have reported numerous replicators based on Diels–Alder and 1,3-dipolar cycloadditions. These have been reviewed recently,[105] and here we highlight particularly interesting examples.

An extreme example of product inhibition is found in the 1,3-dipolar cycloaddition of an azide and a maleimide (Scheme 39).[106] The system did not behave autocatalytically as expected; X-ray analysis of the product revealed that it exists as a dimer in the solid state, suggesting strong product inhibition.

One system involves a product which is designed to eliminate itself from a competing pair of reactions (Scheme 40). The addition of nitrones and maleimides gives endo and exo products; the exo diastereoisomer is folded so that strong intramolecular interactions make it catalytically inert. This gave improved yields and diastereoselectivity over previous systems.[107]

This reaction was adapted as an OR logic gate, where the addition of either of two templates promotes the formation of one species (Scheme 41).[108] Three components react to produce two pairs of diastereomeric products. As before, of each diastereomeric pair the endo product (T1 and T2) is autocatalytically active, and the exo product (T1' and T2') is inert. The result is essentially a broken hypercycle: while T2 is a catalyst for the formation of T1, T1 weakly inhibits the formation of T2. As a consequence, the addition of either T1 or T2 promotes the formation of T1.

Philp and co-workers examined autocatalytic aspects of the Diels–Alder reaction between furans and maleimides in depth,[109–112] (Scheme 42). This system is highly sensitive to structural variations: for example, a methylene group can be the difference between efficient autocatalysis and systems that suffer from competitive preassociative reactions.[109]
bond formation, these are dynamic systems under thermodynamic control.

Von Kiedrowski and Terfort reported the first such system in 1992 (Scheme 43). Recognition here is between a carboxylate and amidinium group. This system displayed first-order kinetics with respect to the template, rather than following the square-root law, suggesting the possibility of exponential growth. However, as far as we are aware subsequent studies were never reported, and the consequences of reversibility in this system were not discussed.

Philp and co-workers explored a similar reaction (Scheme 44). Due to the reversibility of the system, addition of product increased the reaction rate but reduced the yield. This was overcome by reducing the imine to an amine. Unfortunately this could not be done in situ, as the amine precipitated under the reaction conditions. This simulated freezing the autocatalyst into a related cross-catalyst.

Philp and Sadownik have incorporated replicators into a DCL (Scheme 45). Two aldehydes, an amine, and a hydroxylamine equilibrate to form two imines and two nitrones. The nitrones react irreversibly and diastereoslectively with a maleimide to form four products, of which one (trans-38) is autocatalytic. As nitrone formation is reversible, and trans-38 is an autocatalyst, trans-38 is amplified in the product pool at the expense of other products. This demonstrates that under thermodynamic control, autocatalysts can amplify their equilibrium concentration.

Another example of this behavior has been reported by Giuseppone and Xu. A DCL comprising three aldehydes and two amines produces six imines, of which one is an analogue of the original Rebek replicator (Scheme 46). The replicator enhances its equilibrium concentration relative to the predicted distribution in the absence of replication. As the replicator consumes most of its starting materials, other imines sharing these components are depleted. Imines without any components in common with the replicator are slightly enriched relative to the theoretical background distribution. Like Rebek’s original system, the replicator efficiency is strongly limited by product inhibition.

4. Physical Reproduction

Dyson and subsequent authors have distinguished physical replicators from template replicators on the grounds that the autocatalytic species is an aggregate of reaction
products which may not have a defined size or shape. They are often referred to as “self-reproducing” (as opposed to self-replicating) systems, as the product may not be an identical replica of the catalyst. These typically have much less stringent requirements than template replicators in terms of molecular recognition and stoichiometry and are not affected by product inhibition.

These systems operate on a common principle of phase behavior (Scheme 47). Two reactants occupying separate phases react at the interface to produce an amphiphilic product. The reaction proceeds slowly until the product concentration reaches a critical aggregation concentration, leading to the formation of micelles, vesicles, or other supramolecular assemblies. The consequent increase in interfacial surface area increases the rate of reaction, and is therefore autocatalytic. Luisi noted that this represents an unconventional variety of catalysis based not on the lowering of the activation energy of the reaction, but on “physical catalysis”.[119] However, these mechanisms may not be as simple as described here (Section 4.1).

A difference between template and physical replicators is seen in seeding experiments. While template replicators are expected to show a linear relationship between initial rate and initial product concentration, physical replicators are expected to show little increase in initial rate until the initial concentration of product exceeds the critical aggregation concentration.

4.1. Self-Reproducing Micelles

The first examples of autocatalytic reverse micelles,[120,121] and autocatalytic micelles,[122,123] were reported by Luisi and
co-workers in the early 1990s. Micelles composed of octanoate anions were found to catalyze the formation of additional octanoate (Scheme 48).

Octanoate was produced by ester hydrolysis, either under strongly basic conditions\cite{120,123} or through enzyme catalysis\cite{122} by lipases. The oxidation of 1-octanol to octanoate by permanganate\cite{122} was also shown to be autocatalytic (Scheme 49). Similar behavior was observed by Kust and Rathman in the oxidation of amines to N-oxides by hydrogen peroxide.\cite{124} All of these reactions are autocatalytic with respect to octanoate, and Luisi interpreted this behavior as arising from the mechanism described above.

These experiments were attempts to develop minimal autopoietic chemical systems (Scheme 50).\cite{125,126} Autopoiesis describes a system of chemical reactions that are constrained by a boundary (such as a cell membrane), which draws upon energy and material from the environment to continuously produce all the components of the system, including the boundary.\cite{126} These criteria may allow us to determine whether a synthetic system is or is not living. This point is controversial even amongst its proponents: Varela has suggested that autopoietic systems meet the minimum criteria for life,\cite{127} while Luisi holds that autopoiesis is a necessary but not sufficient condition for life.\cite{119}

Luisi’s ester hydrolysis reactions have been modeled by several groups.\cite{128} Mavelli\cite{128} and Coveney\cite{125,126} treated the reaction as classical micelle catalysis, while Buhse et al.\cite{131,132} modeled the reaction as phase transfer catalysis (Scheme 51).

In the case of long-chain (including C₈) alkanoates, Buhse’s work confirms the broad strokes of Luisi’s proposal: mixed micelles composed of product and ester form during the reaction and solubilize the ester in the aqueous phase, thus catalyzing the reaction. However, there is a crucial mechanistic distinction: the hydrolysis reaction itself does not occur at or in the micelles, but rather in bulk solution. The micelles act as a phase-transfer catalyst.

This model accounts for the experimental data well, and explains some unusual behavior. The ethanol byproduct acts as a cosolvent, increasing the solubility of the ester and allowing the reaction to proceed faster. This leads to autocatalysis even when the alkanoate is too short (C₄) to form micelles (Scheme 52). A degree of generality for this mechanism is indicated by three other experiments.

Autocatalysis is observed in biphasic ester hydrolysis if the product is hydrotropic, meaning it is capable of solubilizing hydrophobic molecules in the aqueous phase without forming...
Autocatalysis

Scheme 50. Autopoiesis is a proposed set of minimal criteria for life. A bounded system contains all the components needed to generate all of the components of the system, including the boundary. Adapted from Ref. [119].


Scheme 52. Hydrotropic autocatalysis in ester hydrolysis.

A third, rather complex example involves the biphasic hydrolysis of tri-butyl phosphate (Scheme 54). During the reaction a third phase comprising six species develops. It is predominantly composed of the product, but includes the reagents and solvents. The rate of hydrolysis is mildly increased in the third phase, representing a weak autocatalytic effect.

These observations point to a degree of generality for physical autocatalysis: any biphasic reactions which generate co-solvents, hydrotropes, surfactants, or phase transfer catalysts may exhibit autocatalytic behavior. There are likely unrecognized examples of this behavior present in the literature, although the relevance of these examples to prebiotic chemistry is not immediately obvious.

Perhaps the most sophisticated example of micellar autocatalysis is reported by Giuseppone (Scheme 55). A DCL comprising a hydrophobic aldehyde and eight hydrophilic amines reversibly condenses into eight imine surfactants. Micelles of these imines exhibited varying degrees of autocatalytic efficiency.

This system was used to demonstrate competition (Scheme 56). When amines 39 and 40 were allowed to react with aldehyde A, two imines 39A and 40A were formed. Despite the greater thermodynamic stability of 39A, the higher autocatalytic efficiency of 40A allowed it to out-compete 39A.

While a full kinetic model of this system has not been reported, evidence was provided that it operates by micellar catalysis rather than by the phase transfer mechanism proposed for Luisi’s systems. A DOSY NMR study (Scheme 57) of 40A at equilibrium indicated that the hydrophobic aldehyde was fully associated with the micelles, and the hydrophilic amine was free in solution. This is consistent with the proposed autopoietic mechanism. If the micelles do act as phase transfer catalysts, the quantity of free aldehyde in
the aqueous phase is below the detection limit in this experiment. This may represent an experimental proof of principle for the “lipid world” model, which proposes that populations of self-reproducing lipid assemblies played a role in the origins of life.[137] This model uses differential autocatalytic efficiency correlated with lipid composition as the basis for prebiotic selection, a theory which has mostly been tested with computational studies.

Scheme 56. In MeCN, imine 39A has greater thermodynamic stability than 40A. When transferred to D2O (a, right in plot) or when the starting materials are mixed in D2O (b), 40A dominates at the expense of 39A due to its higher autocatalytic efficiency. From Ref. [136] with permission.

4.2. Self-Reproducing Vesicles

Vesicles have a special place in prebiotic chemistry as models for the study of primitive cell membranes and as components of proposed prebiotic systems. Like modern cell membranes, vesicles may have served to isolate molecules from the prebiotic environment, enable prebiotic and early biological reactions by concentrating reagents, and protect the biological apparatus from hostile environmental conditions.

Crucial to these scenarios is the coupling of vesicle reproduction to the replication of functional molecules within the vesicle,[138] often called core-and-shell reproduction.[139] Without this, the concentrations of essential molecules cannot be maintained at the appropriate range, presumably leading to death by dilution or over-concentration.

The focus of this section is on vesicle self-reproduction driven by chemical reactions, rather than growth and division prompted by other effects.

The mechanisms by which vesicles self-reproduce have been the focus of many experimental and theoretical studies. Vesicles have been shown to divide, fuse together, or undergo other phase behavior that may be relevant to primitive prebiotically relevant processes.[119,140–142] It appears that the mechanism by which vesicles divide varies depending upon the source of new amphiphiles (e.g. reactions vs. pre-formed micelles), the size and composition of the vesicles, and other factors. For a thorough discussion of this subject, see Ref. [119] and [143].

Autocatalytic vesicles which use pre-formed micelles as starting materials will not be discussed here, as these represent a supramolecular rearrangement rather than a chemical reaction. An exciting example of this kind was reported by Szostak, who demonstrated several cycles of self-reproduction in giant vesicles (GVs) (Scheme 58).[59] Vesicle growth led to deformation of spherical vesicles into long threads. Mechanical agitation fragmented the weak threads into smaller daughter vesicles, which could repeat the cycle of growth and division multiple times. This can be seen as analogous to von Kiedrowski’s SPREAD protocol for the manual cycling of self-replicating oligonucleotides (Section 3.1). Further, alongside work by Otto (Section 3.3),[85] it highlights the potential importance of considering mechanical forces in origins of life scenarios.

Early attempts to develop self-reproducing vesicles using enzyme catalysis to drive the autocatalytic hydrolysis of either octanoic or oleic anhydrides proved difficult.[144,145] Unfortu-nately, these reactions proceeded slowly and vesicle division

Scheme 57. DOSY spectrum of Giuseppone’s self-reproducing imine 40A. Aldehyde A is associated with the imine micelles, while hydrophilic amine 40A is free in solution. From Ref. [136] with permission.

Scheme 58. Cycles of vesicle growth and division are driven by mechanical agitation. From Ref. [59] with permission.
was not unambiguously observed. In these experiments, the enzymes were contained within the vesicles, rather than bound to the membrane.

Non-enzymatic methods proved more successful: alkaline hydrolysis of fatty acid anhydrides allowed efficient self-reproduction[^146][^147] as did a later system driven by base-catalyzed deprotection of a phospholipid (Scheme 59).[^146] Anhydride-based systems were successfully combined with enzymatic reactions contained within the vesicles,[^139][^149][^150] including enzymatic RNA replication. The latter experiments represent a model of “core-and-shell” reproduction, a simple model of a cell. However, replication of the “genetic” component and reproduction of the vesicle are not coupled in a controlled manner in these systems.

An interesting system was reported by Conde-Frieboes and Blöchliger[^151] (Scheme 60). Micelles composed of 42 react with 41 to form a two-tailed lipid 43, which forms vesicles rather than micelles. Inorganic phosphate and 41 could generate 42 in situ, although the reaction was very slow, and did not proceed appreciably within one month. The addition of a cationic surfactant (CTAB) allowed significant conversion to occur within three days. The phase behavior of these systems is complicated and was not fully analyzed by the authors; further, they do not claim that it is autocatalytic. However, it seems probable that micelles of the starting material and vesicles of the product could both potentially catalyze the reaction.

As with self-reproducing micelles, in these vesicle-based systems sigmoidal growth is observed and the presence of pre-formed vesicles reduces the induction phase. A number of kinetic models have been developed; for a thorough discussion see Ref. [143].

A single example of autopoietic homeostasis has been reported, in which the synthesis and destruction of vesicle components are balanced.[^152] Vesicles of oleate are produced autocatalytically by alkaline hydrolysis of oleic anhydride. Simultaneously, oleate is oxidized to 44 (Scheme 61) which does not form vesicles. A point of interest is that the OsO₄ oxidant was artificially bound to the vesicle membrane, to ensure that the destruction reaction took place at the vesicle surface rather than in bulk solution. The two processes could be balanced to give growth, death, or homeostatic stages.

This system has been modeled theoretically (Scheme 61).[^153] Notably, simulations of populations of self-reproducing vesicles including both production and destruction of vesicle components were shown to be subject to primitive selection, based on their stability to stochastic fluctuations.[^154]

More recently, Sugawara and co-workers have used novel bolaamphiphiles to develop several generations of autopoietic GVs. These systems demonstrate innovative methods of controlling the location of the reaction within the vesicle, and culminated in a striking example of core-and-shell reproduction.

An early system used a membrane-bound catalyst to deprotect an aldehyde, ensuring that condensation with an amine to produce the surfactant occurred inside the vesicle (Scheme 62).[^155]

Several systems based on imine hydrolysis were used to study the factors influencing growth and division (Scheme 63).[^156][^158] The interaction of non-amphiphilic organic ions with the charged membrane was shown to

---

[^147]: Conde-Frieboes and Blöchliger
[^148]: Scheme 59. Self-reproducing vesicles reported by Luisi demonstrate efficient growth. From Ref. [146] with permission.
[^149]: Scheme 60. A reaction reported by Conde-Frieboes and Blöchliger may be autocatalytic, but analysis is complicated by the phase behavior of the system.
induce division. By including a phosphate moiety in the surfactant precursor, the reaction was localized to the membrane. As a consequence, new vesicles formed at the membrane, a process described as “peeling”.

The system was modified to tolerate the conditions necessary for the polymerase chain reaction, allowing the development of an ambitious system (Scheme 64). Newly-formed DNA synthesized by PCR inside vesicles associates with the charged membrane. This interaction induces division of the vesicles, just as simpler organic ions were shown to do previously. Consequently, vesicles containing larger amounts

Scheme 61. Homeostatic vesicles based on the balance between synthetic and destructive reactions. From Ref. [152] with permission.

Scheme 62. Imine formation drives vesicle self-reproduction. Adapted from Ref. [119].

Scheme 63. “Peeling” driven by surfactant structure. Adapted from Ref. [159].

Scheme 64. DNA replication in a replicating vesicle. From Ref. [160] with permission.
of DNA divided faster. This creates a one-way coupling between DNA replication and vesicle reproduction: only after DNA replication reaches a critical threshold do the vesicles divide. The study represents a large step towards full “core-and-shell” self-replication, and thus a model of a minimal cell. As the catalyst for surfactant formation and the PCR enzymes are not replicated, this groundbreaking model stops just short of that ideal. The significant implications of this for synthetic biology have been discussed elsewhere.[161]

5. Asymmetric Autocatalysis

5.1. The Soai Reaction

Asymmetric autocatalysis is of special interest as it may be related to the origin of biomolecular homochirality.[20,21,162] This topic has fascinated scientists since Pasteur, is of significant contemporary interest, and has been extensively reviewed from several perspectives (for example, see Ref. [163]). While asymmetric autocatalysis has long been known in crystals,[33,162,164] the Soai reaction represents the first example in organic chemistry and the first realization of Frank’s model.[19] There is no suggestion that this chemistry is prebiotically plausible; subsequent research has looked for prebiotically-relevant examples of this behavior.

In 1990 Soai showed that the alkylation of aldehydes by organozincs can behave autocatalytically (Scheme 65).[165] Initial studies used catalysts with high ee values to give products with modest ee’s, but asymmetric autocatalysis with amplification of ee relative to that of the catalyst was soon demonstrated.[166]

Studies have mostly used 2-alkynyl-5-pyrimidyl alcohols alkylated with diisopropyl zinc, as they have proven to be sensitive substrates. The method has been refined over time, from consecutive rounds of reaction and purification[166] to a one pot procedure without purification of the product between reactions[167]

Early studies varied the alkylating agent,[165,168] and other substrates have been studied, including ferrocene derivatives,[169,170] quinolyl alcohols,[171–173] chiral diols,[174,175] pyridine carbaldehydes,[165] and nicotinamides.[176,177]

The reaction is highly sensitive to miniscule sources of chirality: even tiny ee’s in the product can be amplified to optical purity (Scheme 66). A striking demonstration of this effect amplified an initial catalyst ee of 10⁻⁵% to >99.5% in three rounds of reaction.[174]

An exhaustive variety of sources of chirality have been used to induce ee’s in the product, covering everything from simple chiral molecules to chiral crystals and circularly polarized light (Scheme 67). The absolute configuration of all of these initiators has been shown to be reproducibly correlated with the absolute configuration of the product.

The reaction is sensitive to chirality arising from isotopic substitution (Scheme 68). H/D substitution,[180,181] ¹³C/¹²C,[182] ¹⁶O/¹⁸O,[175] and partially deuterated methyl groups[183] are all capable of inducing reproducible enantioselectivity to the reaction.

The high sensitivity of the Soai reaction has been used to identify unknown sources of asymmetry present in the

Scheme 65. The Soai reaction (without amplification).

Scheme 66. Amplification from tiny ee values.

Scheme 67. A small selection of chiral initiators of the Soai reaction.

Scheme 68. Sensitivity to isotopic substitution.
Murchison and Murray meteorites\cite{184} to assign the absolute configuration of challenging compounds\cite{185,186} or compounds with ee’s as low as 0.1 \%.\cite{187}

A related reaction was reported by Carreira et al. in the synthesis of efavirenz.\cite{188} Here, a combination of product 46 and chiral ligand 45 gave 46 in higher ee than did the use of 45 alone (Scheme 69).

\begin{center}
\textbf{Scheme 69.} Soai-related reaction in the synthesis of efavirenz.
\end{center}

Spontaneous symmetry breaking\cite{189} in the absence of any chiral additive was reported by Singleton,\cite{190} Brown,\cite{191} and Soai.\cite{192–194} This is attributed either to unknown chiral impurities in the solvent\cite{192} or to tiny stochastic ee’s.\cite{193} This finding is of great significance as it supports arguments that biological homochirality is a statistical inevitability\cite{20,162} rather than a biological “invention” or the product of deterministic physical forces.

The mechanism of the Soai reaction has been investigated by several groups.\cite{163,195–199} While the exact details remain under investigation, the work of Blackmond and Brown\cite{191,200–210} presents a consistent picture (Scheme 70). The reactions produce complex mixed aggregates of product and other species, and some, but not all of the aggregates are catalytically active. The active species is thought to be an aggregate of (at least) two homochiral product molecules, two molecules of the aldehyde starting material, and multiple organozinc species. The assembly of these components is fast and reversible, with alkylation of the aldehyde by disopropyl zinc acting as the rate-limiting step.

5.2. Mannich and Aldol Reactions

In 2007, Tsogoeva et al. reported asymmetric autocatalytic Mannich and aldol reactions without amplification of ee\cite{211} (Scheme 71). The absolute configurations of the products were reproducibly correlated with those of the catalyst, and subsequently spontaneous symmetry breaking was reported.\cite{212}

Tsogoeva and co-workers have speculated on the mechanism of asymmetric autocatalysis in these organocatalytic systems.\cite{212,214–215} Blackmond argued that the proposed mechanism violates the principle of microscopic reversibility,\cite{216,217} which Tsogoeva and co-workers deny.\cite{214} Detailed studies of these systems have not yet been reported, but Wang et al. have examined a “pseudo-autocatalytic” variant of the Mannich procedure using catalysts of high structural similarity to the products.\cite{218}

Amedjkouh and Brandberg have also reported a similar asymmetric autocatalytic Mannich reaction which takes place in the presence of water (Scheme 72).\cite{219} Modest ee’s were reported, attributed to racemization of the newly-formed product and less selective competing processes.

While these organocatalytic autocatalytic reactions are still under investigation they would be of great interest from a prebiotic perspective if the results are verified. The Soai reaction does not operate under prebiotically plausible conditions, while the Mannich reaction would appear to be much more relevant. The discovery of a second asymmetric autocatalytic reaction may hint at generality, that this behavior may be realized under diverse conditions, rather than being a quirk of the Soai reaction.
6. Summary and Outlook

The past two decades have seen the proliferation of numerous non-natural autocatalytic molecules based on diverse chemistry. This work has drawn inspiration from biology, and we now have numerous models for the replication of genetic molecules and cellular boundaries. The stage is set for the development of increasingly ambitious “core-and-shell” self-reproducing systems which model cell replication. As the line between chemical models of life and synthetic biology begins to blur, work in this area can be expected to offer increasing insight into the origins and definition of life, the source of biological homochirality, and the likelihood of discovering life beyond the earth.

Received: May 3, 2013
Published online: October 11, 2013


