Replicators

From molecules to organisms

István Zachar[1], Ádám Kun[2,3], Chrisantha Fernando[4,5,6] and Eörs Szathmáry[1,2,3,4]

[1] Department of Plant Taxonomy and Ecology, Research Group of Ecology and Theoretical Biology, Eötvös University and The Hungarian Academy of Sciences, Pázmány P. sétány 1/C, H-1117 Budapest, Hungary

[2] Parmenides Center for the Study of Thinking, Kirchplatz 1, D-82049 Munich/Pullach, Germany

[3] Department of Plant Taxonomy and Ecology, Institute of Biology, Eötvös University, Pázmány Péter sétány 1/C, 1117 Budapest, Hungary

[4] Collegium Budapest, Institute for Advanced Study, Szentháromság utca 2, 1014 Budapest, Hungary

[5]National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA

[6]Center for Computational Neuroscience and Robotics, Sussex University, Brighton, Falmer, UK.

Abstract

Here we review chemical and biological replicators that were either engineered (artificially) or evolved (either naturally or artificially). They are automata even though they need not be electro-mechanical machines or computer programs. Gánti has described the class of fluid automata (Gánti, 2003a) into which almost (but perhaps not) all biological replicators fall. Replicators are very special because they are the foundation of evolution by natural selection.

Evolution by natural selection occurs whenever there are units of evolution. Units of evolution must be capable of replication (i.e. multiplication), variation and heredity (Maynard Smith 1987; Szathmáry and Maynard Smith 1993, 1995). Without selection the relative frequency of variants changes by neutral drift, but if an environment tends to allow some variants to replicate faster than others, then these fitter variants can dominate.

An exceptional class of replicator is capable of exponential (Malthusian) growth, and it can be simply shown mathematically that this results in survival of the fittest rather than survival of the first (sub-exponential growth) or survival of the common (hyperbolic growth) (Szathmáry 1991). Exponential growth is an important condition for populations of replicators to produce adaptation.

We hypothesize that any non-trivial self-reproducing system must either rest on chemistry or must emulate processes of chemistry.

Replicators are Autocatalysts

Autocatalysis is a well-known concept originating from chemistry. As Orgel has pointed out: "*all replicating systems are, by definition, autocatalytic and all autocatalytic systems result, in some sense, in replication*" (Orgel 1992, p. 203). This is in close coherence with what Dawkins said when he established the notion of the replicator: "*I define a replicator as anything in the universe of which copies are made*" (Dawkins 1982a, p. 83).

Autocatalysis is a process during which an entity facilitates the appearance of more entities of the same kind (cf. Fig. 1). In a physical system this process needs input material, and may produce other entities, but the important fact is that the entity in focus is produced in surplus. Notice that in the above definitions there is no mention of heredity or variation at all: autocatalysis does not require that the entity be capable of hereditary variation, only that it be capable of the multiplication of a single variant. Still there can be variation between parent and offspring, but unless this difference has no phenotypic effect (i.e. selection cannot discriminate them better than random) and cannot be inherited, parent and offspring are equivalent.

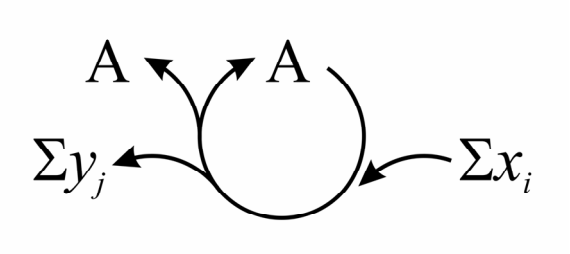

Figure 1. Abstract representation of an autocatalytic cycle. For each turn of the cycle, a new A is created, consuming input materials Σx_i, and producing waste materials Σy_j.

An example of a simple autocatalytic chemical molecule is glycolaldehyde in the formose reaction (Butlerow 1861, Fig. 2). The formose reaction is a network of sugar-related reactions, in which resides an autocatalytic core. This autocatalytic core is a cycle that starts with the spontaneous formation of glycolaldehyde from two molecules of formaldehyde. All the reactions happen spontaneously (without enzymes), and with each turn of the cycle a new molecule of glycolaldehyde is created, igniting more cycles, therefore the synthesis speeds up. The important point here is that one does not need anything else than one molecule of glycolaldehyde (and of course a supply of formaldehyde molecules), to kick-start the system. The main principle of an autocatalytic entity is therefore that it "*can arise only if there is a pre-existing structure of the same kind in the vicinity*" (Maynard Smith and Szathmáry 1995), and usually no bootstrapping is possible (i.e. synthesis from scratch).

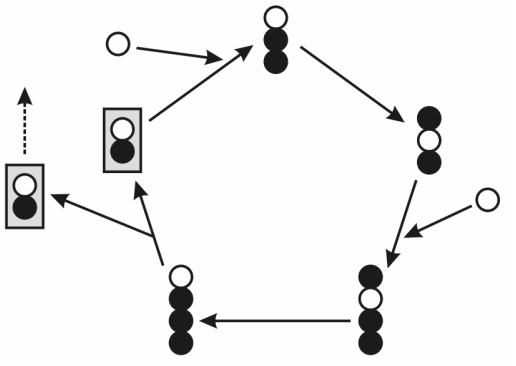

Figure 2. The autocatalytic core of the formose reaction. The two-carbon-atom molecule in the box is glycolaldehyde. Each circle denotes a carbon atom, the white one representing the carbonyl (C=O) group.

Simple autocatalytic chemical entities without heredity (like glycolaldehyde in the formose reaction) are not units of evolution. A wide range of complex entities can be autocatalytically multiplied. Zachar and Szathmáry (2010) showed that reproducing organisms also qualify as autocatalysts. In addition, heterocatalytic products that aid the autocatalytic cycle producing them (such as proteins in case of DNA replication, see later) can also be considered to be autocatalysts in the sense that the pre-existence and the multiplication principles hold for these entities. Parent organisms are needed to produce offspring, and proteins are needed to replicate DNA: they act as catalysts.

Autocatalysis is not enough for evolution

The insufficiency of autocatalysis for evolution is worth emphasizing. Take glycolaldehyde in the formose reaction: no spontaneously arising change of the molecule would end up producing alternative autocatalysts. Because neither glycolaldehyde nor proteins can inherit information they are *non-informational replicators* (sensu Orgel 1992). *Informational replicators* are able to pass on changes they acquired as mutations. For non-informational replicators, either no change is allowed (because the structure would be completely ruined), or no change is heritable (the cycle will go on producing unchanged molecules). These attributes may produce the same phenomenological dynamics, although the causes are not identical (cf. Zachar and Szathmáry 2010). The bottom line is that genes remain the main (but not only, cf. epigenetic inheritance systems, Jablonka and Lamb 1995; Jablonka and Lamb 2006) units of evolution capable of yielding novel evolutionary adaptations.

It is true that even informational replicators cannot transmit *every kind of* change. They have a dedicated part (or rather a function of the entity, in the mathematical sense), which, if changed, passes on changes to offspring. This part is called the genotype (cf. Zachar and Szathmáry 2010). For the DNA, the genotype is the base sequence (as usually no other changes are inherited, e.g. isotope substitutes), for cells it is the genome. The genotype of the replicator is responsible for the hereditary potential of the replicator; therefore it contributes to the evolutionary potential as well.

But is it enough to have a modular structure and the ability to transmit changes in the genotype for fully fledged evolution? Not quite. Oligonucleotides (very short DNA

sequences) for example can pass on mutations during replication, but the number of possible sequences is small. Because of this Szathmáry and Maynard Smith have introduced the terms *limited and unlimited hereditary* (Szathmáry and Maynard Smith 1993 p. 201.). Limited hereditary replicators, due to their structure, are unable to encode for a practically infinite set of varieties, therefore their evolution is restricted to a fixed domain of the search space. The smaller this domain is, the more limited the range of opportunities to come up with new evolutionary adaptations. In contrast, unlimited hereditary replicators can explore a vast search space, which is usually much larger than the actual space covered by the replicators present in a population.

Note that non-informational replicators cannot be subjects to evolution directly, only as members of a more complex assembly (see later). Limited hereditary replicators may undergo only limited evolutionary changes. Only unlimited hereditary replicators are capable of *open-ended evolution*.

It is important to be precise about how informational replicators (with limited or unlimited heredity) comply with the criteria of *units of evolution* (Maynard Smith 1987; Szathmáry and Maynard Smith 1993, 1995): any entity that is able to multiply (**A** → 2 **A**), shows variability in traits (i.e. entities are not identical: **A**, **B**, **C**, ...), and can stably inherit the traits (like begets like: A creates more **A**, **B** creates more **B**, etc.) are units of evolution (cf. Table 1). If entities are not identical but inheritance is not possible (i.e. new changes are not passed on), entities cannot evolve. Still, selection can prefer those variants that are fitter in a given environment. If multiplication is not present then there can be no iterated selection, as only one sorting event would cause the loss of unfit entities, ultimately leading to the extinction of the population (cf. Zachar and Szathmáry 2010).

In pursuit of replicators with evolutionary potential, our concept of simple autocatalytic cycles must be extended to be able to cope with mutations and therefore hereditary information. In turn, we will investigate several of these systems, from well-known to more complicated ones, to provide the reader with an overview of replicators and some of the implications for natural and artificial evolution.

Table 1. Extending the classification of Maynard Smith (Maynard Smith 1987).

	Variation	Multiplication	Heredity
units of sorting	X		
units of selection	X	X	
units of evolution	X	X	X

Genes and DNA

The best known informational replicators are genes. The meme of the *selfish gene* (Dawkins 1976) proved to be quite successful. But what is a gene? We mostly think of a gene as something that codes for some hereditary trait in living organisms. The existence of genes was already suggested by Gregor Mendel, the term itself coined by Hugo de Vries and Wilhelm Johannsen. Later this crude (albeit still valid) view of the gene was elaborated as a region in the genome (the total genetic material of a cell) that codes for a protein (see a historical overview of the gene concept in Gerstein et al. 2007). However, very little of our own DNA codes for peptides. The rest was mostly thought to be either junk DNA or part of the gene regulating system, like the *lac* operon of bacteria (Jacob and Monod 1961). It turns out that most of our DNA is actually transcribed, but does not code for protein (The ENCODE Project Consortium 2007). That is, an RNA molecule is transcribed (produced) from the DNA template, but no protein is translated from the RNA. RNA molecules have a much more varied use, than just being messengers between DNA and protein (mRNA), and coding for some highly important molecules of the translational machinery (e.g. rRNA and tRNA). Even in the translational machinery, the role of RNA has been underestimated. In the ribosome, the macromolecular complex catalyzing peptide synthesis, ribosomal-RNA was thought to be mere structural support for the peptide enzyme. It came as a surprise that at the heart of the ribosome (at its catalytic core) lies RNA (Moore and Steitz 2002). Thus peptide synthesis is done by an RNA enzyme in all of us (Steitz and Moore 2003).

A huge variety of other functional RNA have been discovered in recent years (Meli et al. 2001; Spirin 2002; Dieci and Fiorino G 2007; Collins et al. 2009). For example, small

nucleolar RNAs (snoRNA) play a role in the modification of other RNAs (many RNA molecules have non standard nucleotides, which are generated by modification after transcription); other small nuclear RNAs are involved in splicing (cutting certain parts out of an RNA strand) and regulation; micro RNAs regulate gene expression (Boross et al. 2009; Ghildiyal and Zamore 2009); and the list can go on.

All this evidence forced us to reconsider the definition of gene to "*a locatable region of genomic sequence, corresponding to a unit of inheritance, which is associated with regulatory regions, transcribed regions, and or other functional sequence regions*" (Pearson 2006; Pennisi 2007). It seems that genes are not necessary positioned continuously in the genome, not even if we disregard introns, thus the definition can be further refined to: "*the gene is a union of genomic sequences encoding a coherent set of potentially overlapping functional products*" (Gerstein et al. 2007).

Irrespective of the exact definition, the main storage of hereditary information of a living being is its DNA, in the form of long sequences of 4 nucleobase-pairs. Mutations change DNA, producing variation. The new variant can be passed on to the next generation, and the variant having the highest growth rate can become dominant or take over the population. Genes are evolutionary units. Traits, either coded in a single locus or having a complex genetic background, have been selected for by natural or artificial selection. Humankind has harnessed the force of evolution in agricultural and animal husbandry for thousands of years (Diamond 1997).

Some replicators and units of evolution are poorly understood compared to genes. Some of these seem to have paved the way to the sophisticated Weismannian inheritance system of DNA and proteins. In the following we will review some of them.

Enzymes and RNA

While DNA is the predominant hereditary material of living beings, the tools that actually build the intricate and complex organic system we call life are the enzymes. Enzymes catalyze the myriad reactions in metabolism. Enzymes, as parts of a reproducing organism, are autocatalytic, as they are required to produce more enzymes. Enzymes are also required to replicate the genome, as they exhibit a heterocatalytic effect (i.e. 'aid') by driving and controlling DNA replication. So if the genome is a replicator,

then every protein (and ultimately the whole cell or body as the vehicle, cf. Dawkins 1982b) is a replicator as well. Nevertheless, there is a major difference between the DNA genome and the protein enzymes: information only travels from the DNA to the protein (via translation), therefore proteins can never inherit changes they gathered in their lifetime. There is no way a mutated protein can write back its changed sequence into the genome, according to the central dogma. Protein enzymes are non-informational replicators, whereas genes are informational replicators. This difference makes a huge impact on the evolutionary potential of peptide enzymes.

We were careful to mention peptide enzymes, hinting that there are other enzymes as well, that do not necessarily have this limitation. As we have seen, the ribosome is basically an RNA enzyme (even though it needs proteins for proper functionality). The first purely RNA enzymes were discovered in the beginning of the ’80 (well before the structure of active site of the ribosome had been glimpsed). Cech and coworkers (Kruger et al. 1982) and Altman and coworkers (Guerrier-Takada *et al.* 1983) described an RNA molecule that was capable of catalyzing a chemical reaction. In the first case, a ribosomal RNA intron of *Tetrahymena* was spliced from the RNA chain without any peptide enzyme present. In the second case it became evident that the catalysis was done by the RNA part of RNase P enzyme. Five more natural RNA enzymes (ribozymes) had been described since (Doudna and Cech 2002): group II intron (Peebles et al. 1986); hammerhead ribozyme (Forster and Symons 1987); hairpin ribozyme; Hepatitis delta virus (Sharmeen et al. 1988); and the Neurospora Varkund satellite ribozyme (Saville and Collins 1990). All natural ribozymes cleave RNA, albeit by different mechanisms (Westhof 1999; Doherty and Doudna 2000).

The possibility of RNA catalysis led to the formulation of the RNA world hypothesis (Gilbert 1986): there was an era when information was stored in RNA, and reactions were also catalyzed by RNA. RNA enzymes are informational replicators: any change in their sequence is propagated as the genome and the enzyme is not separated by the one directional translation, but simply the complementary sequence of an RNA enzyme can be considered to be its gene, acting as a template. The gene is replicated to produce the enzyme, and the enzyme is replicated to produce the gene.

Ribozymes can be evolved to perform novel catalytic functions (cf. Szathmáry, 1989, 1990). A pool of RNA molecules is generated, in which at least some part of the sequence is varied. Usually, except for a few nucleotides at the ends of the strands, the rest is randomly generated. The starting pool can have 10^{14}-10^{15} molecules, which is still a small portion of the possible 4^L sequences (where L is the length of the RNA molecule). Molecules are usually between 100-200 nucleotides long. This pool then undergoes iterative rounds of selection. The most commonly used method is SELEX (Systematic Evolution of Ligands by Exponential Enrichment) (Joyce 2007), where selection is done on a chromatographic column. During all chemical reactions, a so-called transition state complex forms. This is a chemical entity that is somewhere halfway between the starting materials and the end products. Enzymes bind the transition state complex of the reaction they catalyze. Anything that binds the transition state complex well could be a catalyst. These complexes however are seldom stable, but it is possible to synthesize stable analogues of them. The transition state analogue is immobilized on the inside surface (Fig. 3) of a column. The pool of RNAs is then slowly flushed through the column. Those that bind to the analogues remain on the surface of the column, while all others are washed out. Binding does not need to be strong in the beginning, just stronger than the binding of the others. The retained RNA molecules are amplified by mutagenic PCR, a technique capable of replicating RNA or DNA molecules and introduce some random mutation to the sequence as well. This variation is essential for evolution to work. The technique is quite powerful and a large number of artificially evolved ribozymes have been produced with it (Joyce 1998; Landweber et al. 1998; Joyce 2002; Spirin 2002). The emerging repertoire of ribozymes is well capable of catalyzing all the important reactions a protocell might need (Jeffares et al. 1998; Joyce 2002). Let us note one interesting ribozyme here, the shortest to date: it catalyses the addition of an amino acid to the end of a substrate RNA chain. The ribozyme is only 5 nucleotides long (Chumachenko et al. 2009)!

There is one catalytic function that has not yet been successfully evolved: template based RNA replication. This is undoubtedly the most important catalytic function an RNA world needs to posses. It would allow *self replication* of the replicase and replication of all the other enzyme molecules in the system. There were attempts to

evolve such an enzyme from a ligase ribozyme (Johnston et al. 2001). Ligases can bind two strands of RNA together. Replication of RNA can be viewed as a template-directed, successive ligation of nucleotides to a strand. The evolved ribozyme is able to ligate at most 20 nucleotides based on a template (Unrau and Bartel 1998; Johnston et al. 2001; Zaher and Unrau 2007). The ribozyme itself is around 200 nucleotides long, thus self-replication is still far off. Researchers are now looking at the 3D structure of the ribozyme (Robertson and Scott 2007; Shechner et al. 2009) to obtain insight of how it can be improved.

One problem usually encountered when haunting enzymes evolved via SELEX is the lack of processivity, i.e. the enzyme can catalyze the reaction once, but not more. An enzyme should be able to catalyze many reactions throughout its lifetime. The problem arises from the fact that these molecules were selected for binding and not high turnover catalysis. The hurdles are even more tedious for protein enzymes that can be selected with the same procedure. Selection acts on the peptide (phenotype), but it is DNA that has to be replicated (genotype). In these experiments randomized DNA sequences are linked to the peptide they code. Bound peptides are separated, and their DNAs are amplified to begin the new round of selection.

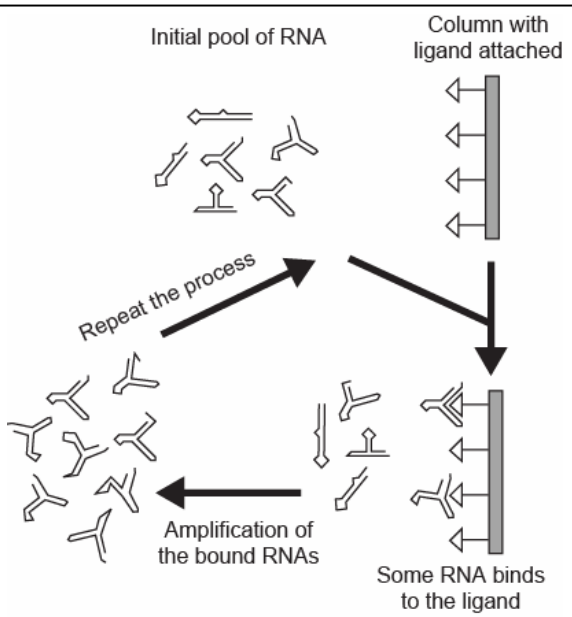

Figure 3. Schematic representation of the SELEX procedure. See details in the text. Please note, that the procedure can work with proteins also, then the DNAs of the selected proteins are amplified, and then transcribed to obtain the protein pool for the next round.

One way to escape this problem is to evolve the enzymes in an *in vitro* compartmentalized system (Tawfik and Griffiths 1998; Griffiths and Tawfik 2000).

Here microdroplets containing all ingredients for replication, enzyme synthesis and the reaction itself are formed in a water-in-oil emulsion (Griffiths and Tawfik 2006). Using in vitro compartmentalization (IVC) genotype and phenotype are linked, similarly to compartmentalization of genes in cells as in nature, by compartmentalization in aqueous microdroplets in water-in-oil emulsions. IVC can be used to select large gene libraries (~10^9). It has been used to select a range of proteins (Tawfik and Griffiths 1998; Ghadessy et al. 2001; Lee et al. 2002; Griffiths and Tawfik 2003; Cohen et al. 2004; Doi et al. 2004; Ghadessy et al. 2004; Aharoni et al. 2005; Mastrobattista et al. 2005) and RNAs (Agresti et al. 2005; Levy et al. 2005) for catalysis, and has also be used to select peptides and proteins for ligand binding (Sepp et al. 2002; Yonezawa et al. 2003; Bertschinger and Neri 2004; Yonezawa et al. 2004; Sepp and Choo 2005) and for regulatory activity (Bernath et al. 2005). IVC selects for all enzymatic features simultaneously - substrate recognition, product formation, rate acceleration and turnover, which have enabled, for example, selection of a variant phosphotriesterase which is one of the most efficient enzymes ever described (Griffiths and Tawfik 2003).

Molecular replicators

Molecular replicators represent an attempt to model living systems in a simpler chemical system. The quest for such replicators gave rise to the field of systems chemistry, which deals with the emergent properties of interacting chemical systems or networks (Ludlow and Otto 2008). The simplest forms of molecular replicators are self-replicators, molecules that can directly catalyze the synthesis of an identical molecule. In contrast there are replicators that, replicating autocatalytically, are not able to do it by themselves, but require ancillary machinery. DNA is one such replicator, which codes for its own machinery (the replicase enzyme). Here we briefly review the rich literature on simpler molecular replicators (for review see also Bag and von Kiedrowski 1996; Isaac et al. 2001; Li and Chmielewski 2003; Paul and Joyce 2003, 2004; Patzke and von Kiedrowski 2007). Networks of molecular replicators will be discussed in the section on metabolism.

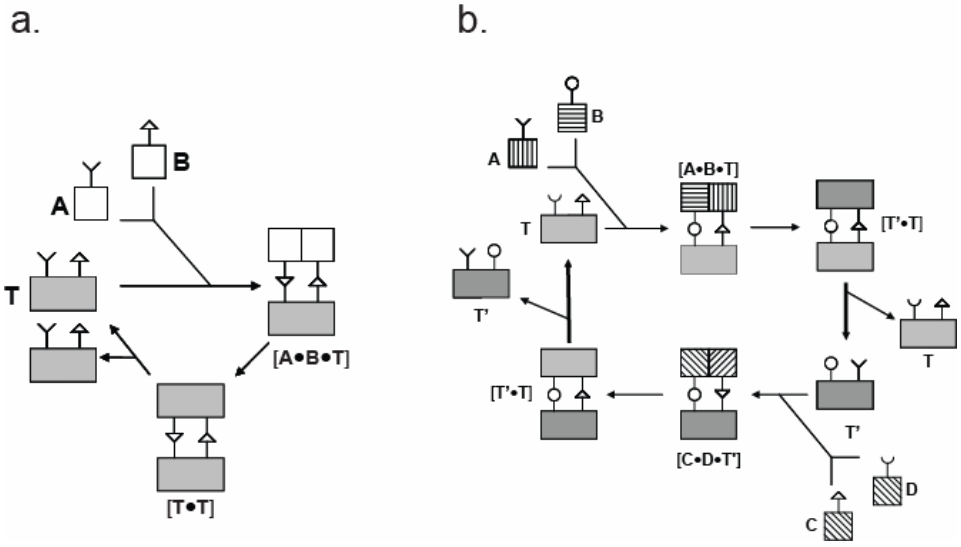

Figure 4. Basic scheme of (a) self-replication and (b) cross catalytic replication. See details in the text.

The basic scheme of all self-replicating molecules is as follows (Fig. 4.a). First, an uncatalyzed reaction between molecule **A** and **B** yields a template **T**. **T** has complementary binding sites for molecules **A** and **B**, and thus can reversibly bind them. The ternary complex [**A•B•T**] forms. Then bond formation occurs between **A** and **B** to give a product duplex [**T•T**]. The duplex dissociates to provide 2 molecules of **T**, and the autocatalytic cycle can start anew. The first realization of such a self-replicating system was provided by Günter von Kiedrowski (1986), where a 6 membered oligonucleotide served as the template (**T**) and two 3 membered nucleotides were the starting material (**A** and **B**). The reaction employed the natural templating capabilities of polynucleotides (like DNA and RNA). The Watson-Crick base pairs bring together the template and the two component oligonucleotides, which leads to the formation of a reaction product that is both complementary and identical to the template. A number of similar systems have been designed in the past two decades (Zielinksi and Orgel 1987). The template can form from more than two molecules, for example **A** and **B** forms an **AB** molecule, which with **C** can bind to T ([**AB•C•T**]) and then give 2 **T**-s (Achilles and Von Kiedrowski 1993).

Another possibility of autocatalytic replication is the one employed by contemporary DNA, where a complementary copy is made (Fig. 4.b). Molecules **A** and **B** yield template **T**, and molecules **C** and **D** yield template **T'**. **T** and **T'** are complementary. **T** has

complementary binding sites for **C** and **D**, and thus can reversibly bind them. The ternary complex [**C•D•T**] forms. Then bond formation occurs between **C** and **D** to give a product duplex [**T'•T**]. This dissociates to give **T** and **T'**. **T'** has complementary binding sites for **A** and **B**, and thus can reversibly bind them. The ternary complex [**A•B•T'**] forms. Then bond formation occurs between **A** and **B** to give a product duplex [**T'•T**]. This set of reactions is autocatalytic as was demonstrated by Sievers and von Kiedrowski (1994) (further example can be found in Kassianidis and Philp 2006a).

An interesting example of nucleotide self-replication is the self-replicating ligase ribozyme evolved artificially by Paul and Joyce (2002). The original R3C ligase ribozyme catalyzes the joining of two RNA molecules (Rogers and Joyce 2001). The evolved ribozyme catalyzes the ligation of two RNA fragments which results in an RNA molecule identical to the template. The two ribozymes then disassociate, and the autocatalytic cycle continues. Amplification of the product can be sustained indefinitely (Lincoln and Joyce 2009). Furthermore, if there is variation in the constituent molecules of **A** and **B** (e.g. **A'** and **B'**) then selection can act on them.

Not only nucleotides can form complementary structures. Smaller molecules can have complementary 3D structures, which allows for the same minimal self-replication scheme to work (Tjivikua et al. 1990; Nowick et al. 1991; Terfort and von Kiedrowski 1992; Conn et al. 1994; Allen et al. 2001; Kassianidis and Philp 2006b; del Amo et al. 2008). For example von Kiedrowski and colleagues designed a pair of self-replicating molecules which can reversibly bind each other via amidinium – carboxylate salt bridge (ionic binding between the two molecules) (Terfort and von Kiedrowski 1992).

Furthermore, peptides can have complementary 3-dimensional structures. Coiled-coil structure forms by hydrophobic and electrostatic interactions between the amino acids of the two peptides. This allows for the reversible binding of two smaller complementary peptides and by forming a peptide bond and a new peptide is made. Both self-replicating and cross-catalytic systems have been demonstrated (Yao et al. 1998b; Isaac et al. 2001).

For the time being molecular replicators are mostly proofs of concept that such a chemical feat is possible. As the art progresses, molecular replicators will find application. Replicators have the unique potential to contribute to the novel biomaterial application of peptides as hydrogels, tapes, and self healing materials (Li and

Chmielewski 2003), if the replication process can be controlled. A change in the environment can cause the self-replication process to start or stop. Chmielewski and colleagues designed a self-replicating peptide that replicates only in acidic conditions (Yao et al. 1997) or at high salt concentration (Yao et al. 1998a). Thus amplification of the product can start or stop when triggering conditions in the environment are met. This is a very simple, albeit important, form of control.

A novel finding is that the autocatalyst can, apart from amplifying itself, act as a catalyst for another reaction (Kamioka et al. 2010). This demonstrates that not only are full fledged enzymes capable of replication and catalytic enhancement of reactions, but so are much simpler chemical systems. By the study of these systems we have the hope to glimpse the pre-RNA world.

Catalytic networks and metabolism

Gánti (1971; 2003a) showed that in any autocatalytic cycle, not just the molecule in focus acts as a replicator, but every other intermediate is a replicator on its own. Furthermore, as the products can ignite new cycles the whole cycle is replicated as well, thus he proposed that the metabolism should be autocatalytic in itself. We have seen that information replication (DNA synthesis) and the enzymes themselves are autocatalytic, but until recently no such insight was available for the intermediate metabolism.

Small molecular metabolism converts food molecules to the chemical constituents of a living being. We know that certain chains of reactions are autocatalytic, for example the Calvin cycle, which produces sugars by fixing CO_2; or ATP production via glycolysis. Here we need to make an important note: an autocatalytic cycle is not considered to be one on the level of the system if there are other pathways producing the autocatalytic molecule. An important characteristic of autocatalytic cycles is that one or some of its constituents needs to be present in order for the cycle to start.

An analysis of the metabolism of 8 Eubacteria including a photosynthetic bacterium, an Archea and a Eukaryote suggest that metabolism is universally autocatalytic (Kun et al. 2008). It also means, that metabolism could not be kick started just from metabolites taken up from the outside. At least one molecular species is required to be present. This universal metabolite is ATP (or any other molecule that can yield ATP), the universal

energy molecule. Other autocatalytic molecules have also been reported, like NAD, Coenzyme A, tetrahydrofuran (THF), quinones and sugars (Kun et al. 2008). Sugars proved to be autocatalytic in the photosynthetic bacterium (*Synechocystis* sp.), thus the Calvin cycle is autocatalytic not only by itself, but also embedded in a network of reactions. The rest of the molecules are cofactors, biochemical substances that help in the transfer of certain chemical groups (NAD transfers hydrogen ion; Coenzyme A acetyl-group; THF methyl-, formyl- and methylene-groups; and quinones transfer electrons). The presence of an obligate autocatalytic cycle can be condition dependent. For example, the Calvin cycle is autocatalytic only if the organism cannot take up sugar (e.g. glucose); once the environment contains some form of sugar that the organism can take up, the Calvin cycle is no longer autocatalytic. Similarly, for *Escherichia coli* in rich medium only the universally autocatalytic ATP is obligatory. But on a minimal medium consisting of only sugar and inorganic substances, other autocatalytic cycles (those for NAD+, CoA and quinones) are revealed. Actually, the metabolic pathways are usually present in the organisms, but they are not obligatorily autocatalytic in many of them. For instance, enzymes of CoA biosynthesis are found in all studied species. Either due to the possibility to uptake certain intermediates from the environment or due to the presence of enzymatic reactions leading to key intermediates, these metabolic routes do not operate as autocatalytic sub-networks.

There exist different autocatalytic pathways for the synthesis of an autocatalytic molecule in different organism. For instance, NAD^+ is an autocatalytic metabolite in both *Methanosarcina barkeri* and *Geobacter sulfurreducens*, but NAD^+ (or NADH) is required for its own synthesis in different biochemical reactions in the two organisms (see Kun et al. 2008), hence providing evidence for the existence of alternative forms of metabolic replicators.

More sophisticated control of artificial network of reactions can be achieved. A self organized, synthetic peptide network consisting of 5 template peptides was shown to be able to exhibit Boolean logic functionality.

Encapsulated metabolism and templates: the Chemoton

Intricate metabolic networks, discussed above, are heavily exposed to environmental fluctuations of input components: changing the environment may ruin the autocatalytic process. If a complex autocatalytic system is therefore to be maintained stably, it must be enclosed in a membrane, to provide the stable local milieu, where the membrane growth is connected to that of the internal autocatalytic cycle, preventing the system choking itself. Gánti has devised a theoretical model, the chemoton (Gánti 2003b), of such a chemical supersystem that explicitly stands on the pillar of autocatalysis, and also fulfils all the criteria of a minimal living organism. The chemoton is a fluid chemical system capable of growing, reproducing and stably maintaining itself, being therefore the minimal model of (cellular) life.

The first version of Gánti's model involved only two subsystems: a metabolic cycle coupled to template replication (Gánti 1971). Later, Gánti included a third subsystem in his model: membrane growth (see Gánti 2003b for review). All the subsystems (membrane, metabolism and template) are chemically coupled, and the template process regulates the other two (just like in real cells). What is important here is that by the introduction of an information carrier template molecule in the system, at least limited heredity can be achieved: splitting microspheres are able to pass on changes in their template molecules to offspring. If the template molecule is long enough, there is no theoretical objection against unlimited heredity and open-ended evolution.

It must be emphasized of course, that the chemoton is a theoretical entity, and no successful physical manifestation has been done in vitro. Nevertheless, there is an increasing amount of theoretical and experimental work dealing with such microspheres, called protocells, to simulate the origin of life and early cellular organisms on Earth (Fernando and Di Paolo 2004;Rasmussen et al. 2008; Solé 2009). One particular line of research focuses on the self-assembly and replication of the boundary subsystem only. There is an ongoing debate whether membranes really can code for and inherit information. In turn we will discuss the possibility of membrane heredity in the context of reflexively autocatalytic sets.

Replication as part of a whole

An important, but not necessary aspect of replicators is self-assembly. DNA is not able to directly self-assemble itself from nucleotides without enzymes, although it is the DNA itself which codes for these enzymes. Nevertheless, an interesting field of prebiotics is to devise such replicating and evolving systems that are able to self-assemble themselves without (external) enzymatic aid. One such model of lipid vesicle replication was conceived by Lancet and co-workers, and is called the Graded Autocatalysis Replication Domain (GARD) (Segré et al. 2000; Lancet and Shenhav 2009).

The GARD model hypothesizes a set of membranogenic molecules, which cross-catalyze the inclusion of other molecules of the set into the membrane, being therefore mutually catalytic. Given a specific interaction matrix of these molecules it can be ensured that a specific composition of the lipid vesicle (an assembly) is maintained even after successive replications (note however, that, contrarily to DNA, the information is not coded in the spatial arrangement of elements but in the composition of assemblies). In pursuit of evolution, it was claimed that the lipid assemblies of the GARD model exhibit evolutionary potential, and the specific assemblies were named therefore compositional genomes, composomes (Segré et al. 2000). In theory, such compositional replicators can be stably maintained, and if mutations are allowed (accidental changes during inclusion), new information can be incorporated into the assembly, which can be passed on subsequently. However, it turned out, that the replication of compositional genomes is so inaccurate (due to the non-specific mutual effects, i.e. each molecule catalyzes the inclusion of many others), that selection cannot maintain fitter genotypes (Vasas et al. 2010). Therefore, composomes are not even capable of limited heredity!

In general, the GARD model is one representative of the concept of reflexively autocatalytic molecule sets. These were first hypothesized for proteins (Eigen 1971; Dyson 1985; Kauffman 1986; Ruiz-Mirazo et al. 2008), but later the idea was extended to the lipid world (GARD) as it seemed more promising to be able to realize *in vivo*. A set of cross-catalytic molecules is reflexive if the synthesis of each member of the set is catalyzed by at least one other member of the set. If this is the case then the whole set therefore grows autocatalytically (given food molecules) even if the members themselves are not autocatalytic individually. If each element is autocatalytic in itself, then we arrive

at the theoretical concept of the hypercycle (Eigen 1971; Eigen and Schuster 1977, 1978). The question of whether such autocatalytic sets can increase in complexity and accumulate adaptations has not been fully answered.

We just briefly mention that membrane heredity does exist. The various genetic membranes (Cavalier-Smith 1995), like membranes of the organelles in a eukaryotic cell, are growing autocatalytically as lipids already built into the membrane facilitate the further inclusion of lipid molecules. Also inheritance is present, as specific protein components in the lipid bilayer catalyze the inclusion of further membrane-specific proteins, maintaining therefore the functional identity of the membrane. Due to these properties, genetic membranes qualify as limited hereditary replicators (Szathmáry 1999).

Again, it must be emphasized that both the GARD model and the reflexively autocatalytic protein networks of Kaufmann are hypothetical systems. There is membrane heredity in cells, although it scarcely leads to novel evolutionary information, as it is basically defined by the genetic information behind membrane proteins and membranogenic processes. However, if a novel membrane layer is created or an old one is lost, this evolutionary change can be propagated indefinitely, as it happened a few times during the evolution of cell organelles, such as plastids.

Prions: real evolutionary replicator?

Prions are proteins that may exist in several distinct conformations, and at least one of these conformations is capable of replication by forcing its alternate conformation (and therefore alternate function) on the normal physiological form of the protein (both having the same amino-acid sequence). Prion proteins are present in both animals, where they are associated with fatal diseases (e.g. transmissible spongiform encephalopathies, i.e. mad-cow disease and scrapie in sheep; Creutzfeldt-Jakob syndrome, kuru, etc. in humans), and in fungi, where their role is not unequivocally malicious. Some prions can even infect other species (cf. human infections of the BSE).

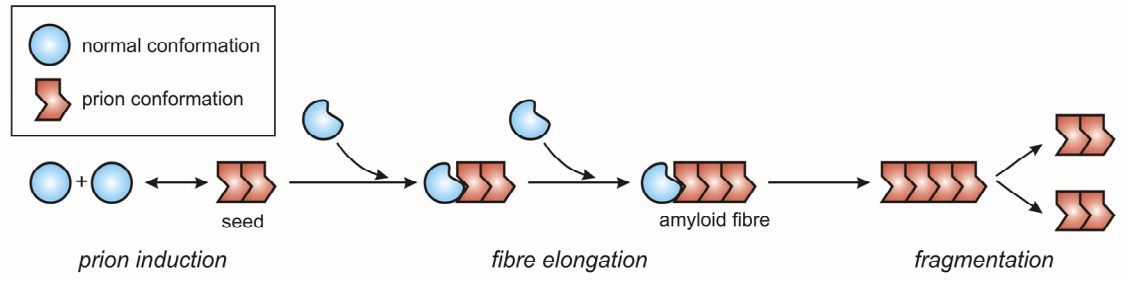

Figure 5. Replication of prion proteins. The prion conformation appears after a spontaneous conversion of native proteins. The prion seed converts further proteins by joining them to its two ends, thus the amyloid fiber starts to grow. As fibres fragment, new seeds are introduced, which further increase the speed of conversion (based on Shorter and Lindquist 2005).

Prion replication involves the formation of amyloid fibres, which are specific structures of aggregated proteins. Due to the specific arrangement of the main polypeptide chain the amyloid fiber can grow by binding further proteins of the same kind on both ends of the fiber for practically unlimited size. Amyloidogenesis is a general phenomenon; many proteins have amyloid conformers that can be seeds of aggregation. Although it is usually unwanted in living cells, therefore selection preferred globular protein structures where the main polypeptide chain is hidden behind a shield of interacting sidechains. Nevertheless, the potential to form amyloids is there, but since it is a bit dangerous to allow accidental aggregation, the 'genie' is enclosed in its globular bottle.

Amyloid fibres thus can seed their own assembly and due to fragmentation can self-propagate (Shorter and Lindquist 2005), therefore they count as replicators (cf. Fig. 5). It must be noted that if there is no supply of native prion proteins via gene expression, no invasive conformation can propagate (although early theories hypothesized a template-based synthesis of new proteins on existing proteins, cf. Root-Bernstein 1983). Since self-propagation is not enough for transmission to other hosts or to other species; amyloids alone are not infective (i.e. the prion has to find the same, or at least similar, proteins in the other species to be infective, and also has to overcome the countermeasures the alien body poses). From the viewpoint of replication the fact that prions may replicate not just in one host, but can be transmitted between species may

mean that they can convert different protein substrates as well (Shorter and Lindquist 2005).

For amyloid fibres, or more precisely prion proteins to be real evolutionary replicators they have to comply with the phenomenological criteria of units of evolution. Multiplication is obviously fulfilled, as the amyloid structure in itself propagates the alternative conformation and by fragmentation causes the appearance of new seeds (Paushkin et al. 1996; Ness et al. 2002). The question is whether changes are possible (i.e. not ruining/curing the infective conformation) and whether these changes are heritable at all. Two things can be changed in proteins: the amino acid sequence and the conformation. Random shuffling of the prion domain (the part of the sequence of the protein where point mutations prevent the prion to propagate) indicates that it is the amino acid composition of this domain, rather than the exact sequence, that determines the prion-quality (Ross et al. 2004), i.e. the infective phenotype is quite robust against changes in its 'genotype', the amino-acid sequence. This means that it is only the infective conformation that is transmitted, and not the altered sequence: information is therefore coded in the conformation, rather than in the sequence (as in case of nucleic acids). Thus no change in the sequence can be inherited - just as anyone would have expected it according to the central dogma. But are there more variants of conformations for the same protein that can be selectively inherited?

It turned out that there can be multiple conformations for a given amino-acid sequence (and not just two: wild type and infectious one), which can stably exist, called strains (Safar et al. 1998; Prusiner 2004). These strains are known to have different phenotypes, i.e. possess differences on incubation time, caused symptoms, etc. (Bruce et al. 1991). If prions have various phenotypes, and these phenotypes cause differential survival, then prions are units of selection. Li et al. (2009) have demonstrated in a series of experiments that different selective regimes (e.g. presence or absence of a prion inhibitor) cause the propagation of different prion strains, effectively demonstrating that prions are units of selection.

The next step is to ascertain that at least some mutations are inherited stably during replication, i.e. during successive transfers of the conformation. The specifically arranged β-sheets inside the amyloid structure exposes the main polypeptide chain which may very

well act as a template being modular, inheriting any change in the template. This template is assumed to be responsible for the transmission of conformation, i.e. the specific structure of the amyloid (Wickner et al. 2007). Li et al. (2009) have also found that new variants appeared *de novo* during replication in the prion population (instead of being there in an initially heterogeneous population), indicating that mutations do affect prions. Even more stunning is that they found different phenotypic properties and therefore different conformations to be heritable. Thus we are dealing with at least limited heredity of prion proteins.

The final step to establish would be to measure the specificity and fidelity of prion conformers. We have seen that prions can convert proteins of other amino-acid sequences (either due to shuffling, or due to alien body). But specificity should refer to conformation only. A high range of specificity means that a certain prion protein may change a narrow selection of non-prion conformations that can be recognized by the prion. Cross-seeding is a proposed, possible mechanism among different prion proteins (Vitrenko et al. 2007). For a high fidelity it must be ensured that if a conformation A is changed to conformation B or C, then they will generate more B or C, respectively, in most of the time, otherwise one can say that B and C have identical phenotypes. According to Li et al. (2009), it seems that different strains can stably inherit phenotypic differences, which means that the conformation is inherited stably. Prions are therefore existing supramolecular hereditary replicators.

Neuronal Replicators

Since William James (1890) the idea has been around that processes of thinking and problem solving are analogous to evolution by natural selection. Natural selection is a subset of Markov search processes that use populations of replicators. Whereas non-replicative search processes such as reinforcement learning (Sutton and Barto 1998) synaptic selectionism (Changeux 1985) and the misleadingly named “Neural Darwinism” (Edelman 1987) have been proposed, there has been much skepticism about the possibility of a replicative processes occurring in the brain, with some notable exceptions (Calvin 1996; Adams 1998; Aunger 2002).

The recently proposed Neuronal Replicator Hypothesis (NRH) which states that there exist replicators in the brain (Fernando et al. 2008). It claimed that patterns of neuronal connectivity were capable of being copied from one part of the brain to another via a topographic map. Spikes occurring in the parental layer are transmitted through the topographic connections to another layer, establishing a template matching in the same sense that hydrogen bonds establish complementary links between nucleotides. Then a process of spike-time dependent plasticity (STDP) takes place in the offspring layer to which copying of the pattern of connectivity is to occur. STDP changes synaptic strengths according to the following rule. If a post-synaptic neuron fires after a pre-synaptic neuron, the synapse is strengthened. If on the other hand the post-synaptic neuron fires before the pre-synaptic neuron, the synapse is weakened. Thus, this asymmetric Hebbian type rule is capable of a simple kind of causal inference. Replication can take place due to causal inference by STDP processes acting in offspring layer of the correlated spike patterns it receives from the parental layer. If combined with activity reverberation limitation to prevent the explosion of Markov equivalent patterns of connectivity, indefinitely large neuronal networks can be copied. Note there is no replication of neurons here. *What is being replicated is a pattern of connections between neurons.*

In addition to structural replicators we also propose electrical replicators; replicators that are patterns of bistable neuronal activity, and spatiotemporal patterns of spikes (Fernando and Szathmáry 2009). These would be capable of generation times in the order of seconds or even milliseconds, allowing millions of generations of evolution by natural selection overnight as one slept.

Neuronal replication would permit a much more powerful kind of cognitive search than is possible by reinforcement learning. The capacity for neuronal search through a space of structured representations may be essential for human generative creativity, insight problem solving, and language learning during infancy (Steels and Szathmáry 2008). In fact, many cognitive architectures implicitly assume the capacity for informational replication of variables, rules and data structures (Hofstadter and Mitchell 1995). What is certain is that neuroscience is severely lacking an understanding of how

search in the space of structured representations is implemented, connectionist models being severely limited when it comes to explaining such search (Marcus 2001).

The same *leitmotif* of the evolution of unlimited heredity from attractor based heredity may have played itself out in the evolution of mechanisms for information transmission within brains. The origin of symbol processing in neuronal systems may have been analogous to the origin of unlimited template replication in genetics. One interesting possibility is that neuronal natural selection was an exaptation of a copying mechanism originally evolved for memory storage. Recently it was shown that the capacity to copy and retrieve actor-critic networks allows multiple reinforcement-learning controllers to exist, thereby helping to solve the stability-plasticity dilemma in a robotic learning task (Fernando 2010).

Outlook

It is perhaps remarkable that the replicators (and reproducers, such as the chemoton) that we have been dealing with either realized in chemistry or kept and manipulated in general-purpose information-processing devices (brains and computers). Is there any other mean to achieve self-reproduction or replication? The answer is affirmative, but with severe limitations. Penrose's (1959) replicators are purely mechanical, whereas a more recent example of non-chemical replicators is mechano-electro-magnetic (Zykov et al. 2005). Upon inspection, it is easy to see that such artifacts could never spontaneously arise, in contrast to replicator chemistry that did arise without intelligent intervention (we believe). These non-chemical replicators do not metabolize beyond simple incorporation of carefully made building blocks that are fed to them in an equally careful manner. Replication thus can be regarded as fairly trivial (despite the ingenuity of their designers).

Let us imagine self-reproducing space probes whereby one could imagine the colonization of the Galaxy by some intelligent civilization. It is easy to see that such a space probe must be a self-reproducing factory. It must have metabolism, because its chemical materials are unlikely to be found ready-made in the environment of the visited alien planets. Thus not only mechanical but also chemical work must be performed. One could say that in a general sense it will be very similar to a cell, with its chemical and mechanochemical devices. It will be the combinatorics of chemical transformations that

will render the system capable of non-trivial replication. Thus we hypothesize that *any non-trivial self-reproducing system must either rest on chemistry or must emulate processes of chemistry.*

In this regard it is useful to contrast open-endedness with selectability. Open-endedness results from powerful combinatorial systems, such as chemistry or language. Natural selectability requires stable propagation of variants produced using the combinatorial system. Open-ended evolvability requires both.

Acknowledgements

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 225167. E.Sz. and Á.K. acknowledge funding from the Hungarian Scientific Research Funds (OTKA NK73047). C.F. is a funded by a Marie Curie Fellowship at Collegium Budapest. This work was carried out as part of EU COST action CM0703 "Systems Chemistry".

References

Achilles T, Von Kiedrowski G (1993) A self-replicating system from three starting materials. Angewandte Chemie International Edition in English 32(8): 1198-1201.

Adams P (1998) Hebb and Darwin. Journal of Theoretical Biology 195(4): 419-38.

Aunger R (2002) The Electric Meme: A New Theory of How We Think. New York, The Free Press.

Agresti JJ, Kelly BT, Jaschke A, Griffiths AD (2005) Selection of ribozymes that catalyse multiple-turnover Diels-Alder cycloadditions by using in vitro compartmentalization. Proceedings of the National Academy of Sciences of the USA 102(45): 16170-16175.

Aharoni A, Amitai G, Bernath K, Magdassi S, Tawfik DS (2005) High-throughput screening of enzyme libraries: thiolactonases evolved by fluorescence-activated sorting of single cells in emulsion compartments. Chemical Biology 12(12): 1281-1289.

Allen VC, Philp D, Spencer N (2001) Transfer of stereochemical information in a minimal self-replicating system. Organic Letters 3(5): 777-780.

Bag BG, von Kiedrowski G (1996) Templates, autocatalysis and molecular replication. Pure & Applied Chemistry 68(11): 2145-2152.

Bernath K, Magdassi S, Tawfik DS (2005) Directed evolution of protein inhibitors of DNA-nucleases by in vitro compartmentalization (IVC) and nano-droplet delivery. Journal of Molecular Biology 345(5): 1015-1026.

Bertschinger J, Neri D (2004) Covalent DNA display as a novel tool for directed evolution of proteins in vitro. Protein Engineering, Design and Selection 17(9): 699-707.

Boross G, Orosz K, Farkas IJ (2009) Human microRNAs co-silence in well-separated groups and have different predicted essentialities. Bioinformatics 25: 1063-1069.

Bruce ME, McConnell I, Fraser H, Dickinson AG (1991) The disease characteristics of different strains of scrapie in sinc congenic mouse lines: implications for the nature of the agent and host control of pathogenesis. Journal of General Virology 72: 595–603.

Butlerow A (1861) Formation synthetique d'une substance sucree. Compt Rend Acad Sci 53: 145–147.

Calvin WH (1996) The cerebral code. Cambridge, MA., MIT Press.

Cavalier-Smith T (1995) Membrane heredity, symbiogenesis, and the multiple origins of algae. In: Arai R, Kato M, Y. D, editors. Biodiversity and Evolution. Tokyo, Japan: The National Science Museum Foundation. pp. 75–114.

Changeux JP (1985) Neuronal Man: The Biology of Mind, Princeton University Press.

Chumachenko NV, Novikov Y, Yarus M (2009) Rapid and simple ribozymic aminoacylation using three conserved nucleotides. Journal of American Chemical Society 131(14): 5257-5263.

Cohen HM, Tawfik DS, Griffiths AD (2004) Altering the sequence specificity of HaeIII methyltransferase by directed evolution using in vitro compartmentalization. Protein Engineering, Design and Selection 17(1): 3-11.

Collins LJ, Kurland CG, Biggs P, Penny D (2009) The modern RNP world of eukaryotes. Journal of Heredity 100(5): 597-604.

Conn MM, Wintner EA, Rebek J, Jr. (1994) Template effects in new self-replicating molecules. Angewandte Chemie International Edition in English 33(15): 1577-1579.

Consortium TEP (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447: 799-816.

Dawkins R (1976) The Selfish Gene. Oxford: Oxford Univ. Press.

Dawkins R (1982a) The Extended Phenotype. Oxford: Oxford University Press.

Dawkins R (1982b) Replicators and Vehicles. In: Brandon R, Burian R, editors. Genes, Organisms, Populations. Cambridge: The MIT Press. pp. 161-180.

del Amo V, Slawin AMZ, Philp D (2008) Manipulating replication processes within a dynamic covalent framework. Organic Letters 10(20): 4589-4592.

Dieci G, Fiorino G CM, Teichmann M, Pagano A. (2007) The expanding RNA polymerase III transcriptome. Trends Genetics 23(12): 614-622.

Doherty EA, Doudna JA (2000) Ribozyme structure and mechanism. Annual Review of Biochemistry 69: 597-615.

Doi N, Kumadaki S, Oishi Y, Matsumura N, Yanagawa H (2004) In vitro selection of restriction endonucleases by in vitro compartmentalization. Nucleic Acid Research 32(12): e95.

Doudna JA, Cech TR (2002) The chemical repertoire of natural ribozymes. Nature 418(6894): 222-228.

Dyson F (1985) The origin of life. Cambridge: Cambridge University Press.

Edelman GM (1987) Neural Darwinism. The Theory of Neuronal Group Selection New York, Basic Books.

Eigen M (1971) Selforganization of matter and the evolution of biological macromolecules. Naturwissenscaften 10: 465-523.

Eigen M, Schuster P (1977) The hypercycle. A principle of natural self-organisation Part A: emergence of the hypercycle. Naturwissenscaften 64: 541-565.

Eigen M, Schuster P (1978) The Hypercycle. A principle of natural self-organization. Part C: The realistic hypercycle. Naturwissenscaften 65(7): 341–369.

Fernando C (2010) Neuronal Replicators Solve the Stability-Plasticity Dilemma. GECCO 2010 Portland, Oregon.

Fernando C, Karishma KK, Szathmáry E (2008) Copying and Evolution of Neuronal Topology. PLoS ONE 3(11): e3775.

Fernando C, Szathmáry E (2009) Chemical, neuronal and linguistic replicators. Towards an Extended Evolutionary Synthesis. M. P. a. G. Müller. Cambridge, MA, MIT Press.

Fernando C, Di Paolo EA (2004) The Chemoton: A model for the origin of long RNA templates. Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems, ALIFE'9 Boston, September 12th-15th, 2004. MIT Press.

Fernando C, Szathmáry E (2009) Natural selection in the brain. Toward a Theory of Thinking. V. G. a. A. v. M. B. Glatzeder. Berlin, Springer.

Forster AC, Symons RH (1987) Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active site. Cell 49: 211-220.

Gánti T (1971) The Principle of Life (in Hungarian) Budapest: Gondolat.

Gánti T (2002) On the early evolutionary origin of biological periodicity. Cell Biology International 26(8): 729–735.

Gánti T (2003a) The principles of life. Oxford: Oxford University Press.

Gánti T (2003b) Chemoton Theory. New York: Kluwer Academic/Plenum Publishers.

Gerstein MB, Bruce C, Rozowsky JS, Zheng D, Du J et al. (2007) What is a gene, post-ENCODE? History and updated definition. Genome Research 17: 669-681.

Ghadessy FJ, Ong JL, Holliger P (2001) Directed evolution of polymerase function by compartmentalized self- replication. Proc Natl Acad Sci U S A 98(8): 4552-4557.

Ghadessy FJ, Ramsay N, Boudsocq F, Loakes D, Brown A et al. (2004) Generic expansion of the substrate spectrum of a DNA polymerase by directed evolution. Nature Biotechnology 22(6): 755-759.

Ghildiyal M, Zamore PD (2009) Small silencing RNAs: an expanding universe. Nature Reviews 10: 94-108.

Gilbert W (1986) The RNA world. Nature 319: 618.

Griffiths AD, Tawfik DS (2000) Man-made enzymes-from design to in vitro compartmentalisation. Current Opinion in Biotechnology 11(4): 338-353.

Griffiths AD, Tawfik DS (2003) Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization. Embo J 22(1): 24-35.

Griffiths AD, Tawfik DS (2006) Miniaturising the laboratory in emulsion droplets. Trends in Biotechnology 24(9): 395-402.

Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell 35: 849-857.

Hofstadter DR, Mitchell M (1995) The Copycat Project: A Model of Mental Fluidity and Analogy-Making. Fluid Concepts and Creative Analogies: computer models of the fundamental mechanisms of thought. New York, Basic Books: 205-267.

Isaac R, Ham T-W, Chmielewski J (2001) The design of self-replicating helical pepetides. Current Opinion in Structural Biology 11(4): 458-436.

Jablonka E, Lamb RM (1995) Epigenetic Inheritance and Evolution. Oxford: Oxford University Press.

Jablonka E, Lamb MJ (2006) The evolution of information in the major transitions. J Theor Biol 239(2): 236–246.

Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol 3: 318–356.

Diamond, J. (1997) Guns, Germs, and Steel: The Fates of Human Societies: W.W. Norton & Co.

Jeffares DC, Poole AM, Penny D (1998) Relics from the RNA world. J Mol Evol 46(1): 18-36.

Johnston WK, Unrau PJ, Lawrence MS, Glasen ME, Bartel DP (2001) RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. Science 292(5520): 1319-1325.

Joyce GF (1998) Nucleid acid enzymes: playing with a fuller deck. Proc Natl Acad Sci USA 95: 5845-5847.

Joyce GF (2002) The antiquity of RNA-based evolution. Nature 418(6894): 214-220.

Joyce GF (2007) Forty years of in vitro evolution. Angewandte Chemie International Edition 46(34): 6420 - 6436.

Kamioka S, Ajami D, Rebek J, Jr. (2010) Autocatalysis and organocatalysis with synthetic structures. Proc Natl Acad Sci U S A 107(2): 541-544.

Kassianidis E, Philp D (2006a) Reciprocal template effects in a simple synthetic system. Chemical Communications(39): 4072-4074.

Kassianidis E, Philp D (2006b) Design and Implementation of a highly selective minimal self-replicating system Angewandte Chemie International Edition 45(38): 6344-6348.

Kauffman SA (1986) Autocatalytic sets of proteins. J Theor Biol 119: 1-24.

Kruger K, Grabowski P, Zaug AJ, Sands J, Gottschling DE et al. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell 31: 147-157.

Kun Á, Papp B, Szathmáry E (2008) Computational identification of obligatorily autocatalytic replicators embedded in metabolic networks. Genome Biology 9: R51.

Lancet D, Shenhav B (2009) Compositional Lipid Protocells: Reproduction without Polynucleotides. In: Rasmussen S, Bedau MA, Chen L, Deamer D, Krakauer DC et al., editors. Protocells. Cambridge: The MIT Press. pp. 233–252.

Landweber LF, Simon PJ, Wagner TA (1998) Ribozyme engineering and early evolution. BioScience 48(2): 94-103.

Lee YF, Tawfik DS, Griffiths AD (2002) Investigating the target recognition of DNA cytosine-5 methyltransferase HhaI by library selection using in vitro compartmentalisation. Nucleic Acid Research 30(22): 4937-4944.

Levy M, Griswold KE, Ellington AD (2005) Direct selection of trans-acting ligase ribozymes by in vitro compartmentalization. RNA 11(10): 1555-1562.

Li J, Browning S, Mahal SP, Oelschlegel AM, Weissman C (2009) Darwinian evolution of prions in cell culture. Science: DOI: 10.1126/science.1183218.

Li X, Chmielewski J (2003) Challenges in the design of self replicating peptides. Organic and Biomolecular Chemistry 1: 901-901.

Lincoln TA, Joyce GF (2009) Self-sustained replication of an RNA enzyme. Science 323(5918): 1229 - 1232.

Ludlow RF, Otto S (2008) Systems chemistry. Chemical Society Reviews 37: 101-108.

Marcus G F (2001) The Algebraic Mind: Integrating Connectionism and Cognitive Science, MIT Press.

Mastrobattista E, Taly V, Chanudet E, Treacy P, Kelly BT et al. (2005) High-throughput screening of enzyme libraries: In vitro evolution of a beta-galactosidase by fluorescence-activated sorting of double emulsions. Chemical Biology 12(12): 1291-1300.

Maynard Smith J (1987) How to model evolution. In: Dupré JA, editor. The Latest on the Best. Cambridge: MIT Press.

Maynard Smith J, Szathmáry E (1995) The Major Transition in Evolution. Oxford, UK: W.H. Freeman.

Meli M, Albert-Fournier B, Maurel MC (2001) Recent findings in the modern RNA world. International Microbiology 4(1): 5-11.

Moore PB, Steitz TA (2002) The involvement of RNA in ribosome function. Nature 418(6894): 229-235.

Ness F, Ferreira P, Cox BS, Tuite MF (2002) Guanidine hydrochloride inhibits the generation of prion "seeds" but not prion protein aggregation in yeast. Molecular and Cellular Biology 22(15): 5593–5605.

Nishida N, Katamine S, Shigematsu K, Nakatani A, Sakamoto N et al. (1997) Prion protein is necessary for latent learning and long-term memory retention. Cellular and Molecular Neurobiology 17(5): 537–545.

Nowick JS, Feng Q, Tjivikua T, Ballester P, Rebek J, Jr. (1991) Kinetic studies and modeling of a self-replicating system. Journal of American Chemical Society 113(23): 8831–8839.

Orgel LE (1992) Molecular replication. Nature 358(6383): 203-209.

Patzke V, von Kiedrowski G (2007) Self replicating systems. ARKIVOC: 293-310.

Paul N, Joyce GF (2002) A self-replicating ribozyme. Proc Natl Acad Sci USA 99(20): 12733-12740.

Paul N, Joyce GF (2003) Self-replication. Curr Biol 13(2): R46-R46.

Paul N, Joyce GF (2004) Minimal self-replicating systems. Current Opinion in Chemical Biology 8: 634-639.

Paushkin SV, Kushnirov VV, Smirnov VN, Ter-Avanesyan MD (1996) Propagation of the yeast prion-like [*psi*$^+$] determinant is mediated by oligomerization of the sup35-encoded polypeptide chain release factor. The EMBO Journal 15(12): 3127–2134.

Pearson H (2006) Genetics: What is a gene? Nature 441: 398-401.

Peebles CL, Perlman PS, Mecklenburg KL, Pertillo ML, Tabor JH et al. (1986) A self-splicing RNA excises an intron lariat. Cell 44: 213-223.

Pennisi E (2007) DNA study forces rethink of what It means to be a gene. Science 316: 1556–1557.

Penrose L.S. (1959) Self-reproducing machines. Sci. Am. 200(6), 105-114.

Prusiner SB (2004) Prion biology and diseases: Cold Spring Harbour Laboratory Press.

Rasmussen S, Bedau MA, Chen L, Deamer D, Krakauer DC et al., editors (2008) Protocells: Bridging nonliving and living matter. Cambridge: MIT Press.

Robertson MP, Scott WG (2007) The structural basis of ribozyme-catalyzed RNA assembly. Science 315: 1549-1553.

Rogers J, Joyce GF (2001) The effect of cytidine on the structure and function of an RNA ligase ribozym. RNA 7: 395-404.

Root-Bernstein RS (1983) Protein replication by amino acid pairing. J Theor Biol 100(1): 99–106.

Ross ED, Baxa U, Wickner RB (2004) Scrambled prion domains form prions and amyloid. Molecular and Cellular Biology 24(16): 7206–7213.

Ruiz-Mirazo K, Umerez J, Moreno A (2008) Enabling conditions for 'open-ended evolution'. Biology and Philosophy 23(1): 67–85.

Safar J, Wille H, Itri V, Groth D, Serban H et al. (1998) Eight prion strains have prpsc molecules with different conformations. Nature Medicine 4: 1157–1165.

Saville BJ, Collins RA (1990) A site-specific self-cleavage reaction performed by a novel RNA in Neurospora mitochondria. Cell 61: 685-696.

Segré D, Ben-Eli D, Lancet D (2000) Compositional genomes: Prebiotic information transfer in mutually catalytic noncovalent assemblies. Proc Natl Acad Sci USA 97(8): 4112–4117.

Sepp A, Choo Y (2005) Cell-free selection of zinc finger DNA-binding proteins using in vitro compartmentalization. Journal of Molecular Biology 354(2): 212-219.

Sepp A, Tawfik DS, Griffiths AD (2002) Microbead display by in vitro compartmentalisation: selection for binding using flow cytometry. FEBS Lett 532(3): 455-458.

Sharmeen L, Kuo MYP, Dinner-Gottlieb G, Taylor J (1988) Antigenomic RNA of human Hepatitis delta viruses can undergo self-cleavage. Journal of Virology 62: 2674-2679.

Shechner DM, Grant RA, Bagby SC, Koldobskaya Y, Piccirilli JA et al. (2009) Crystal structure of the catalytic core of an RNA-polymerase ribozyme. Science 326(5957): 1271-1275.

Shorter J, Lindquist S (2005) Prions as adaptive conduits of memory and inheritance. Nature Reviews Genetics 6: 435–450.

Sievers D, Von Kiedrowski G (1994) Self-replication of complementary nucleotide-based oligomers. Nature 369: 221 - 224.

Solé RV (2009) Evolution and self-assembly of protocells. The International Journal of Biochemistry & Cell Biology 42(2): 274–284.

Spirin AS (2002) Omnipotent RNA. FEBS Letters 530(1-3): 4-8.

Steels L, Szathmáry E (2008) Replicator Dynamics and Language Processing. The Evolution of Language: Proceedings of the 7th International Conference (EVOLANG7). A. D. M. Smith, Smith, K. and Ferrer i Cancho, R. Singapore, World Scientific Press: 503.

Steitz TA, Moore PB (2003) RNA, the first macromolecular catalyst: the ribosome is a ribozyme. Trends in Biochemical Sciences 28, 411-418.

Sutton SR, Barto AG (1998). Reinforcement Learning: An Introduction. Cambridge, MA, MIT Press.

Szathmáry E. (1989) The emergence, maintenance, and transitions of the earliest evolutionary units. Oxf. Surv. Evol. Biol. 6, 169-205.

Szathmáry, E. (1990) Towards the evolution of ribozymes. Nature 344, 115.

Szathmáry E (1991) Simple growth laws and selection consequences. Trends in Ecology and Evolution, 6, 366-370.

Szathmáry E (1999) Chemes, genes, memes: A revised classification of replicators. Lectures on Mathematics in the Life Sciences 26: 1-10.

Szathmáry E, Maynard Smith J (1993) The origin of genetic systems. Abstracta Botanica 17(1-2): 197–206.

Szathmáry E, Maynard Smith J (1995) The major evolutionary transitions Nature 374: 227–232.

Tawfik DS, Griffiths AD (1998) Man-made cell-like compartments for molecular evolution. Nat Biotechnol 16(7): 652-656.

Terfort A, von Kiedrowski G (1992) Self-replication by condensation of 3-aminobenzamidines and 2-formylphenoxyacetic acids. Angew Chem Int Ed Engl 31(5): 654-656.

Tjivikua T, Ballester P, Rebek J, Jr. (1990) Self-replicating system. Journal of American Chemical Society 112(3): 1249–1250.

Tompa P, Friedrich P (1998) Prion proteins as memory molecules: an hypothesis. Neuroscience 86(4): 1037–1043.

Unrau PJ, Bartel DP (1998) RNA-catalysed nucleotide synthesis. Nature 395(6699): 260-263.

Vasas V, Szathmáry E, Mauro S (2010) Lack of evolvability in self-sustaining autocatalytic networks: A constraint on metabolism-first path to the origin of life. Proc Natl Acad Sci U S A: doi: 10.1073/pnas.0912628107.

Vitrenko YA, Gracheva EO, Richmond JE, Liebman SW (2007) Visualization of aggregation of the Rnq1 prion domain and cross-seeding interactions with Sup35NM. J Biol Chem 282: 1779–1787.

von Kiedrowski G (1986) A self-replicating hexadeoxy nucleotide. Angew Chem Int Ed Engl 25: 932-935.

Westhof E (1999) Chemical diversity in RNA cleavage. Science 286(5437): 61-62.

Wickner RB, Edskes HK, Shewmaker F, Nakayashiki T (2007) Prions of fungi: Inherited structures and biological roles. Nature Reviews Microbiology 5(8): 611–618.

Yao S, Ghosh I, Zutshi R, Chmielewski J (1997) A pH-modulated, self-replicating peptide. Journal of American Chemical Society 119: 10559-10560.

Yao S, Ghosh I, Zutshi R, Chmielewski J (1998a) A self-replicating peptide under ionic control. Angewandte Chemie International Edition 37(4): 478-481.

Yao S, Ghosh I, Zutshi R, Chmielewski J (1998b) Selective amplification by auto- and cross-catalysis in a replicating peptide system. Nature 396: 447-450.

Yonezawa M, Doi N, Higashinakagawa T, Yanagawa H (2004) DNA display of biologically active proteins for in vitro protein selection. Journal of Biochemistry (Tokyo) 135(3): 285-288.

Yonezawa M, Doi N, Kawahashi Y, Higashinakagawa T, Yanagawa H (2003) DNA display for in vitro selection of diverse peptide libraries. Nucleic Acid Research 31(19): e118.

Zachar I, Szathmáry E (2010) A New Replicator: A general framework for analyzing replication. Unpublished.

Zaher HS, Unrau PJ (2007) Selection of an improved RNA polymerase ribozyme with superior extension and fidelity. RNA 13(7): 1017-1026.

Zielinksi WS, Orgel LE (1987) Autocatalytic synthesis of a tetranucleotide analogue. Nature 327: 346 - 347.

Zykov V, Mytilinaios E, Adams B, Lipson H (2005) Self-reproducing machines. Nature 435: 163-164.