From the origin of life to self-replicating peptides

Qi Liu*

October 2011

Abstract

In this essay, synthetic self-replicating systems construted from nucleotides, peptides, and artificial compounds are investigated. The mechanism behind these self-replicating molecules can help us understand how life emerged from simple chemical contents in the prebiotic earth. In the process of improving these manmade systems to better mimic living organism, useful theoretic models as well as techniques have been built, which not only helps in understanding the mechanism and searching for novel self-replicating molecules, but also creates new fields of chemistry research.

Contents

1 Introduction

One essential and fascinating characteristic of life is its ability to replicate itself. In modern form of living organisms, genetic information is stored in DNA or RNA molecules, and duplicated through

*Top Master Programme in Nanoscience, Zernike Institute for Advanced Materials, University of Groningen. Email address: liu.2@student.rug.nl. Supervisor: Prof. Sijbren Otto.

efficiently organizing different kinds of highly developed enzymes. However, this kind of modern machinery is thought to be too sophisticated to be involved in the very beginning stage of life. Although it was demonstrated early that some of the amino acids can be formed in the atmosphere on primitive earth [1], nucleotide chemists are unable to show that relatively pure oligonucleotides can accumulate into substantial amounts on primitive earth [2]. It is thus believed that simpler forms of self-replicating molecules were likely to have taken part in the origin of life. Over the years, people have come up with a hypothesis that a RNA world might exist on prebiotic earth, where RNA served both as the storage of genetic information, and the enzyme to process the replication of genome. This was proved to be possible when scientists first discovered an enzyme-like RNA in 1989 [3]. After this, more self-replicating systems have been discovered by mainly making use of oligonucleotides, peptides and some artificial compounds, and implementing the mechanism of template-directed reactions.

In a typical self-replicating system, the self-replicator molecule serves as a catalyst in the synthesis reaction of itself from precursor molecules. Due to similar in components with DNA and RNA, oligonucleotides were first investigated as a candidate for synthetic self-replicators. Later, after several examples of oligonucleotide self-replicators were demonstrated, potential self-replicating peptides were inspected considering that they are building blocks of highly efficient and effective proteins as well as the capability to transfer information by specific pairing like the pairing of base pairs in DNAs and RNAs. After many successful demonstration of self-replicating peptides, non-biological molecules that make use of peptides were also introduced to the study. All these studies give us better understanding of how did the emergence of life happen in a chemical level; additionally, they also raise some questions that provide us new viewing angles of chemistry and biology.

In the following sections, we will discuss the models used to analyze self-replicating processes, and investigate some of the examples achieved using different materials, showing the special properties as well as limitations of these synthetic self-replicating systems.

2 Analysis of a self-replicating system

2.1 Autocatalysis model

People studying self-replicating systems often use a three step model to elucidate the replication process, as shown in Figure 1.

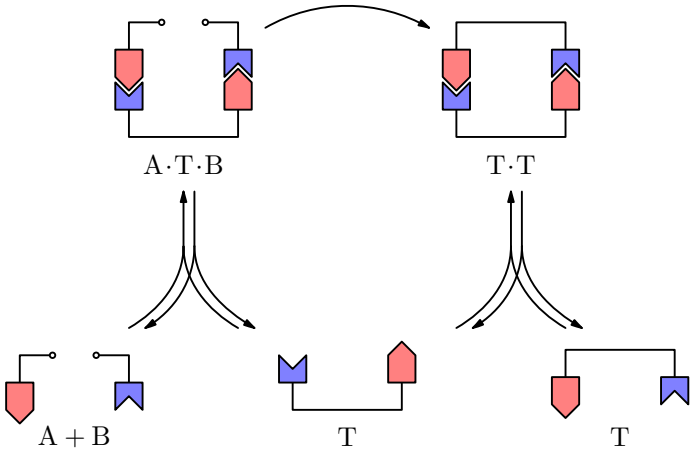


Figure 1: Schematic drawing of a self-replicating system.

In this model, a template molecule T, which is self-complementary, clones itself via an autocatalytic replication cycle, given its building blocks, molecules A and B. In the first step, the template T binds A and B reversibly by binding the corresponding recognition sites, for example by hydrogen bonding, forming a trimolecular complex A·B·T. In this complex, the reactive sites of A and B are brought into close proximity, which facilitates their binding reaction. In the second step, molecules A and B need to link to each other by forming a covalent bond between the reactive sites, resulting in the same molecule as the template T. Thus the trimolecular complex A·B·T is irreversibly transformed into a dimer T·T. In the last step, the dimer T·T dissociates into two template molecules reversibly, finishing one cycle of replication. Then the next cycle can continue with the newly produced template molecules. We can also describe the reaction process using formulas shown in Equation 1.

$$\mathrm{A} + \mathrm{B} + \mathrm{T} \rightleftharpoons \mathrm{A{\cdot}B{\cdot}T} \; ; \tag{1a}$$

$$\mathrm{A{\cdot}B{\cdot}T} \rightarrow \mathrm{T{\cdot}T} \; ; \tag{1b}$$

$$\mathrm{T{\cdot}T} \rightleftharpoons \mathrm{T} + \mathrm{T} \; . \tag{1c}$$

This reaction is an autocatalytic reaction, where the presence of the template molecules facilitates the formation of the template itself. Through copying itself, the template molecule transmits its structural information to the new copies. We know that such behavior of information transcription is essential in gene replication of all living organism.

For efficient self-replication processes, the second step is crucial, where the dimer T·T is formed from the trimolecular complex A·B·T irreversibly. This means the trimolecular complex is more stable than the dimer. But in many cases, this is normally not true, because of entropic reasons. The excessively stability of the dimer product will lead to inefficient replication. This is the problem of product inhibition, the biggest obstacle of many self-replicating systems.

Also important in this model is that the direct binding of the two building blocks A and B should be inhibited compared to the template-directed binding. This makes the auto-catalytic cycle the main contribution of the production of template T.

Besides auto-catalytic mechanisms, Self-replication can also be achieved using cross-catalytic systems. In a cross-catalytic system, instead of one self-complementary template, we have a pair of templates which are complementary to each other. One template acts as a catalyst for the formation of the other, and vice versa. The reaction process can be described as formulas below.

$$\mathrm{A} + \mathrm{B} + \mathrm{T_{CD}} \rightleftharpoons \mathrm{A{\cdot}B{\cdot}T_{CD}} \; , \qquad \mathrm{C} + \mathrm{D} + \mathrm{T_{AB}} \rightleftharpoons \mathrm{C{\cdot}D{\cdot}T_{AB}} \; ;$$

$$\mathrm{A{\cdot}B{\cdot}T_{CD}} \rightarrow \mathrm{T_{AB}{\cdot}T_{CD}} \; , \qquad \mathrm{C{\cdot}D{\cdot}T_{AB}} \rightarrow \mathrm{T_{AB}{\cdot}T_{CD}} \; ;$$

$$\mathrm{T_{AB}{\cdot}T_{CD}} \rightleftharpoons \mathrm{T_{AB}} + \mathrm{T_{CD}} \; .$$

Cross-catalytic self-replications can also be achieved by utilizing three or more template/product molecules that each of them can catalyze the formation of another, forming a multi-step replicating cycle. These are of course more complex systems which are more difficult to control and predict, but they are largely involved in making for example synthetic molecular networks.

2.2 Reaction order

As shown in the model, after each cycle the number of templates is doubled. Thus we would expect an exponential growth of the product concentration, at the beginning of the auto-replication reaction or in case of enough resources (A and B in the model). Then after all the resource reactants are consumed, the concentration of the product would reach a maximum. Thus in principle, we would have a sigmoidal (or "S" shaped) concentration-time profile for an auto-replication reaction.

In real self-replicating systems, exponential growth is often not observed, instead a parabolic growth usually arises. We will describe these two important kinds of amplification processes briefly as follows.

We consider a simple reproduction process as shown below

$$\mathrm{A} + \mathrm{S} \to 2\mathrm{A} + \mathrm{W},$$

where A is a replicator, S and W are the source and waster, respectively.

Note that there are two types of order: the catalytic order r and the reaction order p. The catalytic order r describes the net chemical equation of the form:

$$\mathrm{A} + \mathrm{B} + r\,\mathrm{T} \to (1+r)\mathrm{T}.$$

And the reaction order p describes the reaction rate of product formation

$$\frac{\mathrm{d}[\mathrm{T}]}{\mathrm{d}t} \propto [\mathrm{T}]^p.$$

In general we have $p \leq r$[4].

In an exponentially growing system, the autocatalytic reaction order $p = 1$. The concentration of the replicator A, denoted by c, satisfies the differential equation

$$\frac{\mathrm{d}c}{\mathrm{d}t} = kc. \tag{2}$$

Notice that this could only be true when provided infinite amount of source, or at the initiating stage of the reaction. Then the growth of A is exponential in time characterized by a reaction rate constant k, assuming enough source is provided,

$$c(t) = c(0)\,\mathrm{e}^{kt}. \tag{3}$$

For two species A and B, with different reaction rate constant k_A and k_B respectively, the ratio between there concentrations at time t, is given by

$$\frac{c_\mathrm{A}(t)}{c_\mathrm{B}(t)} = \frac{c_\mathrm{A}(0)\,\mathrm{e}^{k_\mathrm{A}t}}{c_\mathrm{B}(0)\mathrm{e}^{k_\mathrm{B}t}} = C\mathrm{e}^{(k_\mathrm{A}-k_\mathrm{B})t},$$

where $C = \frac{c_\mathrm{A}(0)}{c_\mathrm{B}(0)}$ is a constant. Assuming $k_\mathrm{A} > k_\mathrm{B}$, we would have

$$\frac{c_\mathrm{A}(t)}{c_\mathrm{B}(t)} \to \infty, \quad \text{as } t \to \infty.$$

Thus the concentration of A will eventually be infinitely large with respect to the concentration of B, regardless of the ratio of the initial concentrations of the two. This means species A will be selected to survive over species B in the end, in the progression of evolution. This is the well-known case of survival of the fittest, or Darwinian selection. Here in this model system, the fittest species is the one with the largest reaction constant k.

If we take into account the finite amount of available source materials for the replication, we will get a sigmoidal shape of the concentration-time profile. This can be simply shown as follows. For the same reaction $\mathrm{A} + \mathrm{S} \to 2\mathrm{A} + \mathrm{W}$, the equation will be different from above after some time, when the amount of source S is not enough for all the replicator molecules A to run a replication cycle. Suppose this critical time is t_1, and in the beginning, fixed amounts of replicator A and source S are given, denoted by $c_\mathrm{A}(0)$ and $c_\mathrm{S}(0)$ respectively, with $c_\mathrm{A}(0) \ll c_\mathrm{S}(0)$. Then the time dependence of the concentration $c_\mathrm{A}(t)$ should be rewritten as

$$\frac{\mathrm{d}c_\mathrm{A}(t)}{\mathrm{d}t} = \begin{cases} k\,c_\mathrm{A}(t), & \text{if } t \leq t_1; \\ k\,c_\mathrm{S}(t), & \text{if } t > t_1. \end{cases}$$

The concentration of source molecules $c_{\mathrm{S}}(t)$ follows a similar equation. For simplicity, we suppose at moment t_1, we have $c_{\mathrm{A}}(t_1) = c_{\mathrm{S}}(t_1)$. Then the solution is found to be

$$c_{\mathrm{A}}(t) = \begin{cases} c_{\mathrm{A}}(0)\,\mathrm{e}^{kt}, & \text{if } t \le t_1; \\ c_{\mathrm{S}}(0)\left(1 - 1/2\,\mathrm{e}^{k(t-t_1)}\right), & \text{if } t > t_1. \end{cases}$$

The lineshape of the concentration as a function of time of such a simple model is shown in Figure 2. It is clear from the figure that the maximum reaction rate occurs in the middle of the a reaction

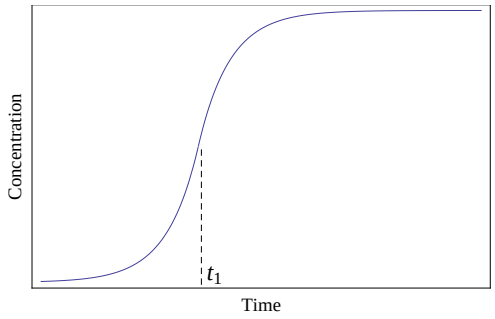


Figure 2: Sigmoidal lineshape of the concentration as a function of time in exponential growth system.

process (t_1 in this model). Before this maximum, the system undergoes a relatively slow initiation because of low concentration of the template.

Whereas in parabolic growth, we'll show as follows that survival of everyone is the evolution result. For parabolic growth, reaction order $p = 0.5$, the concentration of a replicator A follows the "square root law":

$$\frac{\mathrm{d}c}{\mathrm{d}t} = kc^{1/2}, \tag{4}$$

which gives

$$c(t) = \left(\frac{1}{2}kt + \sqrt{c(0)}\right)^2. \tag{5}$$

It follows that as time goes on, the relative concentration of two parabolically replicating species becomes

$$\frac{c_{\mathrm{A}}(t)}{c_{\mathrm{B}}(t)} = \frac{\left(\frac{1}{2}k_{\mathrm{A}}t + \sqrt{c_{\mathrm{A}}(0)}\right)^2}{\left(\frac{1}{2}k_{\mathrm{B}}t + \sqrt{c_{\mathrm{B}}(0)}\right)^2} \to \frac{k_{\mathrm{A}}^2}{k_{\mathrm{B}}^2}, \quad \text{as } t \to \infty.$$

Thus there won't be infinite difference between species with different reaction constants. This means that every replicator will possibly both survive, leading to an indefinite coexistence, which is not Darwinian kind of selection.

We can also look at systems with a general reaction order p ($0 < p \le 1$). The concentration variation becomes

$$\frac{\mathrm{d}c}{\mathrm{d}t} = kc^p,$$

and the solution is given by

$$c(t) = \left((1-p)kt + c(0)^{1-p}\right)^{\frac{1}{1-p}}.$$

It also follows that

$$c(t) \to \left((1-p)k\right)^{\frac{1}{1-p}} t^{\frac{1}{1-p}}, \quad \text{as } t \to \infty.$$

From the order of the dependence on time of the concentration function ($c(t) \to t^{\frac{1}{1-p}}$), we can see that in a $p-$order system, as the reaction order p becomes larger (closer to 1), the difference between the limit concentrations of species with different reaction constants k will become larger. But as

long as $p < 1$, such differences are not infinite, but will diverge when p converges to 1. Also can be seen is that, between two self-replicating species with different reaction orders, the one with larger p will be infinitely larger than that with a smaller p. Thus in the presence of several self-replicating pathways with different orders, the species with the largest reaction order will survive over the others, regardless of the relation between the rate constants k, provided enough resources for its replication. This is also the case where only the fittest species survives, with the "fittest" here meaning the one with the largest reaction order. This tendency towards larger reaction order makes rational sense since in the end of evolution, what remains in modern organisms are only the efficient mechanisms.

From the discussion above, we now understand that in a self-replicating system, one need to show the concentration of the replicator is growing exponentially to prove it the Darwinian type of evolution. And the closer the reaction order p gets to 1, better selectivity of the fittest species will be achieved. Thus people studying synthetic self-replication molecules are making their efforts to get larger reaction order systems, getting closer to exponential growth.

2.3 Proof of self-replication

In order to confirm a reaction to be self-replicating, we need to do some characterization. First, one needs to show the template effect of the self-replicator by for example comparing the reaction rate with or without pre-existing templates. Second, to obtain the reaction order p, one needs to measure the production rate variation in time at the beginning stage, which should have the lineshape of a exponential function $[\mathrm{T}]^p$ with [T] the template concentration.

3 Experimental achievements

The discovery of RNA as an enzyme-like catalyst in the early 1980s [3] and as a "self-replicating ligase ribozyme"[5] not only supports the hypothesis of a RNA world, but also inspires people to synthesize more enzyme-free self-replicating molecules. Until now, people have succeeded in making many self-replicating systems, based on nucleotides, peptides and even artificial compounds.

3.1 Self-replicating oligonucleotides

One of the first designs of non-enzyme nucleic acid replication implementing an autocatalytic process was reported by von Kiedrowski in 1986[6]. In this work, a hexadeoxynucleotide (**3**) can catalyze its own formation from a trideoxynucleotide (**1**) and a complementary trideoxynucleotide (**2**) as precursors, structures shown in Figure 3. As shown in the figure, the 5'-terminus-protected trimer **1** d(Me-CCG-p) (denoting a deoxynucleotide sequence C(cytosine)-C-G(guanine) with 5' terminal attached with methyl group and 3' terminal with phosphate group) and the 3'-terminus-protected trimer **2** d(CGG-p') (p' here denoting the protecting group) can react in the presence of EDC (1-ethyl-3-(3'-dimethylaminopropyl)carbodiimide, which is generally used to activate phosphate groups), linking by a phosphodiester linkage to yield a self-complementary hexamer **1** d(Me-CCGCGG-p'). The sequence was chosen to make the product **3** could act as a template to facilitate its own formation.

In the auto-catalytic cycle, a trimolecular complex is formed from **1**, **2**, and **3**, in which the reactive 3' terminus of **1** and 5' terminus of **2** are close in space, thus ready to be ligated. Then the 3' phosphate group of **1** attacks the 5' hydroxyl group of **2**, forming a 3'-5' internucleotide bond, resulting in a duplex of two **3** molecules. In the final step, the duplex can dissociate to form two copies of the template **3**, initiating another replicating cycle.

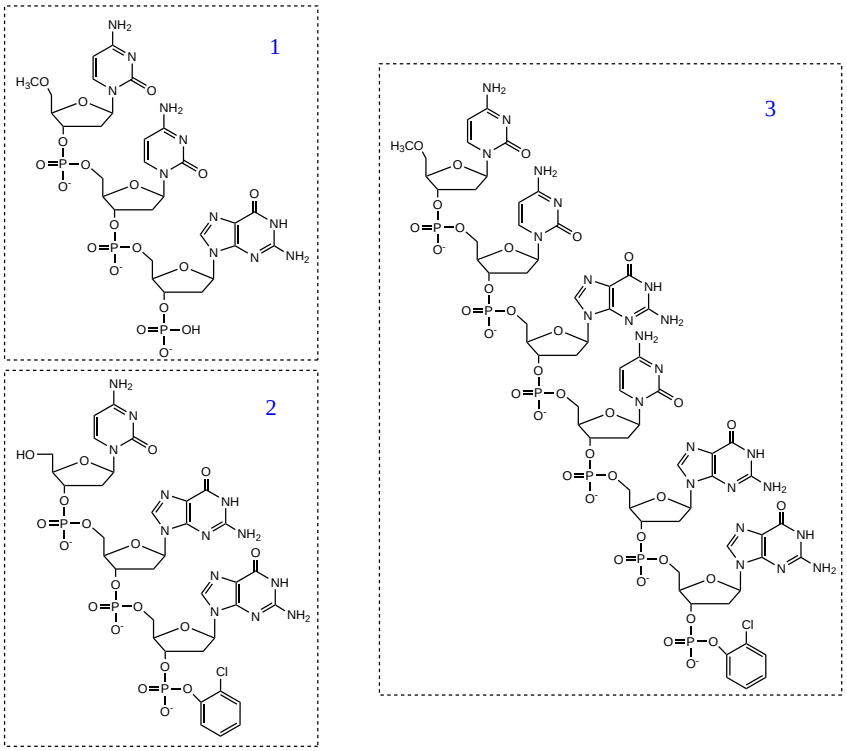


Figure 3: The first nucleotide direvative self-replication.

In this system, the formation of **3** consists of two different contributions, one with template-directed auto-catalytic replication, the other is the direct ligation of precursors **1** and **2** with template. The initial rate of the product followed an empirical rate equation Eq.6

$$\frac{\mathrm{d}c}{\mathrm{d}t} = a\sqrt{c} + b, \tag{6}$$

where the synthesis of **3** via a template-dependent reaction channel is described as square root law with a slope a, and via a template-independent channel denoted by a constant b. The initial reaction rate increased when adding more template, in a square root way. This was the confirmation of the autocatalytic pathway contributing to the formation of the product.

The reaction order of the this self-replicating system was found to be 1/2, due to the stability of the duplex. Most of the template molecules stay in the duplex form (a double helix), leaving them in an inactive state, inhibiting the auto-catalytic cycle. It is also significant in this system that the non-template pathway the predominant contribution of the reaction.

In later works, the rate of the template-induced hexamer formation was greatly increased, the sequence selectivity in the self-replication of haxadeoxynucleotides was also demonstrated, giving further evidence for the auto-catalytic nature of this kind of reaction [7]. But exponential growth was not demonstrated using similar systems.

3.2 Overcoming the product inhibition problem

Several methods are adopted to overcome the obstacle of production inhibition, getting closer and closer towards exponential growth. Some of these methods are: replicator design, SPREAD procedure, and mechanically induced dissociation.

An exponential non-enzymatic replication of oligonucleotides was realized in 1998 [8] on the surface of a solid support using a method called SPREAD (Surface Promoted Replication and Exponential Amplification of DNA analogues), shown in Figure 4. In this procedure, a template oligonucleotide

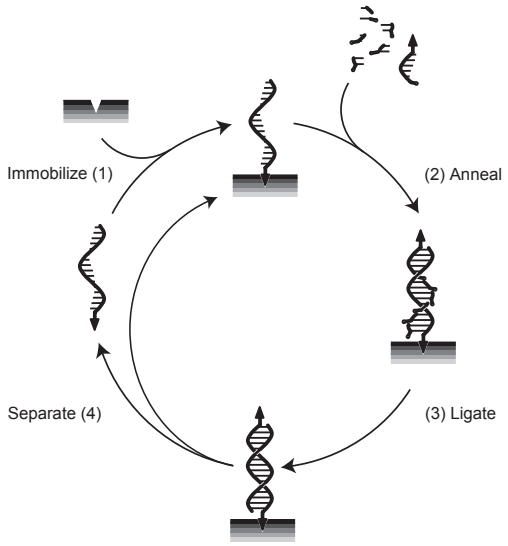


Figure 4: General scheme of the SPREAD procedure. Figure taken from Ref[8].

(a 14-mer in the paper) are immobilized on a solid support by using the irreversible disulphide exchange reaction. Then the fragments bind to complementary sites of the template by base pair recognition. Then EDC is added into the system to ligate these fragments which are already brought into proximity. Finally the copies are liberated from the solid support by rinsing with 0.1 M NaOH. And these copies are re-immobilized again to become new templates for the next replication cycle.

The SPREAD method successfully demonstrates exponential amplification of the replication. The product inhibition doesn't exist here, because the dimer product is separated manually after the ligation step, so the relative stability of the dimer doesn't influence the production. However, this stepwise procedure not autonomous, thus by definite far away from a natural form of life. Nevertheless, this method may be useful in search for autonomous variants which may lead to self-sustaining chemical systems displaying exponential growth.

Another way to solve the self-inhibition problem to approach exponential growth is by make using of mechanically induced dissociation. Here the template molecules don't stop at forming a dimer, instead they would form a fiber-like structure with each template molecule stacking on top of another. With each of such fibers, one end of the fiber can serve as a catalytic site, bringing precursor molecules into proximity. Then precursors can form another template molecule, elongating the fiber by one unit of template, at the same time forming a new catalytic end, ready for next cycle of replication. Then a different story happens after one fiber grows long enough: it can be broken by external mechanical forces, into two short segments each with a catalytic end ready for replication.

This procedure also generates nearly exponential growth. In this method, the number of catalytic units is not doubled after every new template molecule is formed, instead it's doubles after the formation of several template products. By doing this, the stability of the product dimer is no longer a problem, instead one only need to have relative unstable long fibers, which is the case for normal systems. Examples using this procedure will be discussed in detail later.

3.3 Self-replicating peptides

Considering that of amino acids and peptides possibly existing on prebiotic earth, and the ability of molecular recognition, poly-peptides seem a strong candidate in the search for non-enzymatic self-replicating systems. But unlike nucleic acid bases, which provide a clear framework to establish complementary molecular interactions, information replication can be more complex for polypeptides, because their functionalities might dependent on both the amino acid sequence and the folded three-dimensional structure, like the more complex proteins.

3.3.1 First demonstration

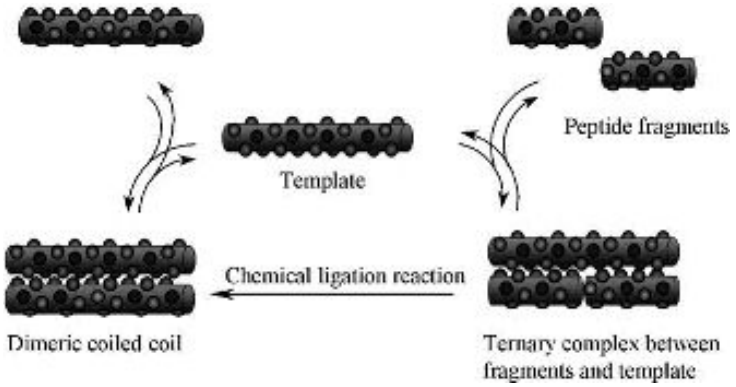


Figure 5: An autocatalytic coiled-coil self-replication cycle. Figure taken from Ref[9].

The first self-replicating peptide system was demonstrated by Lee and Ghadiri *et al* in 1996 [10] with an α-helical peptide.

α-helix is the most common type of secondary structure of proteins (thus can also describe polypeptides), with backbone atoms coiled into a right handed rotational helix (in the direction of the sequence, i.e. from N-terminus (free amino group) to C-terminus (free carboxyl group)). In many situations α-helix is energetically favorable because it allows hydrogen bonds formed between carboxyl oxygen atoms at position i and amide hydrogen atoms at position $i + 4$.

In this experiment, the self-replication process is based on a 32-residue α-helical peptide, similar in structure to the leucine-zipper homodimerization domain of an enzyme, the yeast transcription factor GCN4. The leucine zipper is a super-secondary structure which can function as a dimerization domain, generating adhesion force between two parallel α-helices. The homodimerization ability of the leucine-zipper structure makes this peptide helix a good choice as a self-replicating template. It can autocatalytically templates its own synthesis by accelerating the amide-bond condensation of a 15mer fragment and a 17mer fragment forming the 32mer itself, in neutral, dilute aqueous solutions.

This design of peptide replicator is based on one simple form of natural protein tertiary structure: the α-helical coiled coil. Figure 5 shows the pathway of an autocatalytic self-replicating peptide in the conformation of α-helical coiled coil, which most self-replicating peptide systems adopted. This kind of structure is marked by its amphiphilic primary sequence of heptad (7mer) repeats $(abcdefg)_n$. The first (a) and the fourth (d) amino acids are the hydrophobic sites which defines the complementary inter-helical recognition surface, as shown in Figure 6, showing also the sequences of the peptides.

The ligation site (the break point of the 15mer and 17mer) is chosen to lie on the solvent exposed surface, avoiding possible interference with the hydrophobic recognition sites. These two peptide fragments are reactive to each other, with the 15mer being nucleophilic and the 17mer electrophilic.

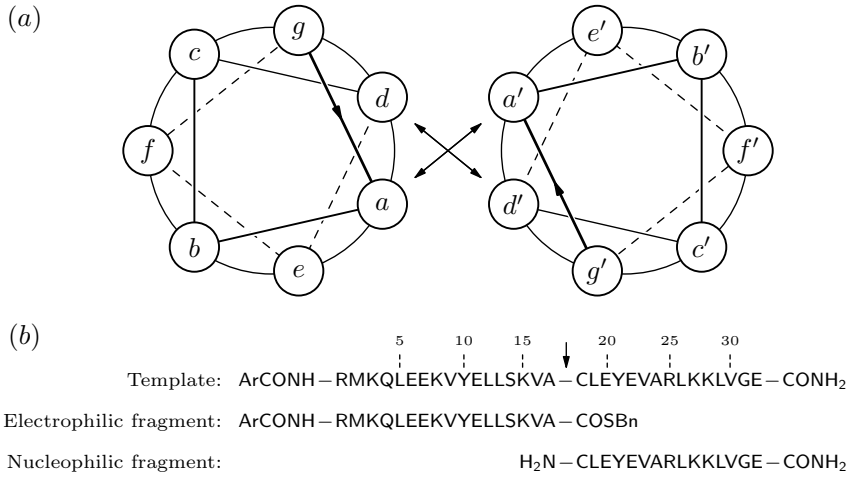


Figure 6: (a) A helical wheel diagram of the template peptide in dimeric α-helical coiled-coil configuration, emphasizing the heptad-repeat motif. The diagram of one α-helical peptide can be viewed as a projection of the helical structure, with one cycle of 7 amino acids, connected in the order of $a_1b_1c_1d_1e_1f_1g_1 - a_2b_2c_2d_2e_2f_2g_2 - a_3 \cdots$. In each heptad unit, element a and d are hydrophobic. In a dimerization structure, e.g. a leucine zipper structure, two parallel (with same rotating direction) helices can be chained together at the hydrophobic sites. (b) Amino acid sequences of the peptides, with ligation site marked by an arrow. (Ar: aryl group; Bn: benzyl group.)

The formation of the amide bond between the two parts makes use of thioester-promoted peptide fragment condensation strategy introduced by Kent in 1994, which is called native chemical ligation or Kent ligation [11]. It's "native" in such a way that one does not need to add in any external coupling reagents for the ligation reaction, thus it's a convenient method for chemical synthesis of proteins. In the reaction model, the separate single stranded peptides may be partially folded or randomly coiled, then a template, a 15mer and a 17mer are brought together by inter-helical hydrophobic interactions to form nicely coiled α-helices termolecular complex. In this trimer complex, the C-terminal thiobenzyl ester COSBn of the 17mer and the N-terminal cysteine residue are in proximity, ready for the formation of an amide bond through native chemical ligation. After the ligation, a new template is formed out of one 15mer and 17mer. A self-replication cycle is finished after the separation of the template complex.

To corroborate the autocatalytic self-replication, several characterizing measurements were carried out. The concentration of the template and thus the reaction rate of product formation was calculated using reversed phase HPLC (high-performance liquid chromatography). The initial production rate is proportional to the square root of in the initial concentration of the template, or reaction order 0.5. Thus production inhibition also plays a role in this reaction cycle, same as expected since the template dimer complex (similar to a leucine zipper) would be rather stable.

Special in this experiment, there are various plausible pathways in template production, auto-catalytic and non-autocatalytic. The catalytic channel is composed of three pathways (one ternary complex, two binary complexes), where one of the electrophilic fragment and the nucleophilic fragment or both the two are associated with the template by inter-helical interactions. In non-autocatalytic channel, the fragments are partially folded or randomly coiled, with no inter-helical association with the template helical. To probe the contribution of two binary complexes to overall autocatalytic process, two "crippled" templates were used, each of which disrupts inter-helical recognition with one

of the fragments. Then reactions were performed as control experiments, results of which strongly suggest the binary complex pathways contribution is insignificant, thus the observed autocatalysis must in fact proceed through the self-replicative ternary intermediate. Control experiments also show implicitly that sequence-selective molecular recognition plays a critical role in peptide self-replication.

This is the first experimental demonstration of the chemical feasibility of peptide self-replication. Different designs of self-replicators using α-helical as well as β-sheet structures came out, with reaction order getting closer to 1. The self-replicating peptides may have played a role in the molecular origin of life in prebiotic earth.

3.3.2 Approaching exponential growth with self-replicating peptides

Parabolic growth is the result of product inhibition in dimeric self-replicating systems, which in the absence of product inhibition, exponential growth should be achieved. The increase of reaction order was accomplished by different strategies of destabilizing the template duplex: shortening the coiled coil structure, or creating a kink in the helical by inserting a proline amino acid.

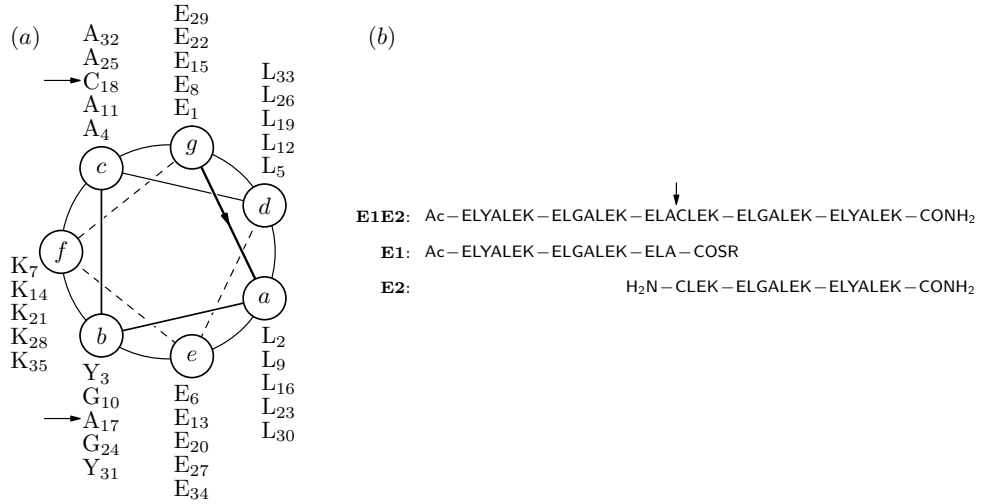


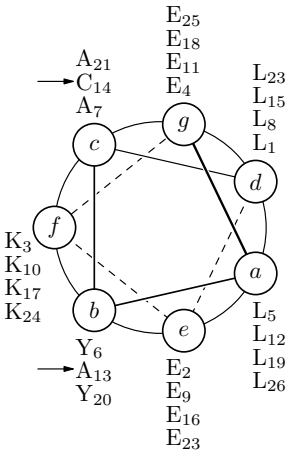
Figure 7: (a) The helical wheel diagram of peptide E1E2 showing the positions of the coiled-coil, heptad repeat ($a \sim g$). (b) Peptide sequence of the peptides, ligation residues Ala17 and Cys18 are located at the solvent exposed b and c positions. (R = $(CH_2)_2CONH_2$)

Soon after the first demonstration, in 1997, Yao *et al* discovered a peptide sequence capable of autocatalytic self-replicating from two peptide fragments, in a pH-dependent manner [12]. They designed a 35-amino acid peptide E1E2, shown in Figure 7. E1E2 was designed to form coiled-coil structure under acidic conditions due to the relief of repulsion between Glu sites at e and g positions because of protonation. While under neutral pH, E1E2 cannot form coiled-coil helix because of the negatively charged Glu residues destabilize the coiled coil. E1E2 should adopt a randomly coiled conformation.

Since the conformation of E1E2 can be modulated by changing pH, one can easily investigate the dependence of the autocatalytic behavior upon the conformation of the peptide. Experiments showed that the self-replicating peptide E1E2 promotes its own production under acidic conditions, while in neutral pH the autocatalysis is suppressed and adding template does not have indistinguishable effect.

Although the E1E2 system had a relatively low catalytic efficiency, it was the first display of self-replication combined with environmental control in peptide autocatalysis processes.

In the effort to improve the catalytic efficiency of the self-replicating peptide E1E2, Chmielewski and coworkers designed another peptide with a very high reaction order $p = 0.91$ [13]. This peptide RI-26 contains 3 full heptad repeats within the coiled coil, one cycle shorter than the original E1E2 peptide, as shown in Figure 8. RI-26 maintained the design principle of E1E2, with glutamate residues



RI-26: Ac−LEKELYALEKELACLEKELYALEKEL−$CONH_2$

RI-26a: Ac−LEKELYALEKELA−COSR

RI-26b: CLEKELYALEKEL−$CONH_2$

Figure 8: Helical wheel diagram and sequence of peptide RI-26 and its fragments.

also positioned at the e and g positions of the helical to achieve pH-based control over helicity. The two fragments RI-26a and RI-26b can undergo thioester mediated ligation to form RI-26.

Similar to E1E2, peptide RI-26 also tends to adopt a helical conformation in acidic environment. At pH 7.0, the helical content was only 28%, while at pH=4.0, this rate increased to 87%*. Interestingly, the peptide RI-26 was found to exist in forms of tetramers, at pH 4.0, different from E1E2 which is a dimer under similar conditions..

The reaction order of this self-replicating peptide was calculated by finding the best fit for the catalytic and noncatalytic reaction rates, and was found to be about 0.91. The initial rate was also found to be linearly proportional to the concentration of the template to the power of 0.91. This is higher than expected, since if the ligation complex adopts the form RI-26a·RI-26b·$(\text{RI-26})_3$, the reaction order can only reach 0.75 because of product inhibition. Thus they called this process "weakly exponential". This high reaction order was caused by a reduced stability of the tetramer complex of RI-26 as compared to the dimer complex in E1E2 self-replication, as well as a decrease in the non-catalytic background condensation from its fragments. This was confirmed by the fact of lower melting temperature of RI-26 than E1E2 (by 20 °C) and a 90-fold smaller non-catalytic reaction efficiency than that of E1E2.

This work on peptide RI-26 was the first demonstration of self-replicating system with a reaction order larger than the product inhibition.

Besides shortening the peptide, Chmielewski and coworkers came up another way to enhance the self-replication process, namely incorporating proline in the peptide to induce a kind in the helix, hence reducing the stability of the coiled coil [14].

The limitation of the method of weakening the product complex is that the interactions between peptide fragments and template peptide may be also weakened at the same time. Whereas introducing

* The percentage of the peptide in α-helical conformation can be determined by circular dichroism (CD) spectroscopy, a frequently-used methods which uses the ultraviolet absorption of peptide bonds to estimate the secondary structure of proteins and polypeptides. α-helix, β-sheet and randomly coiled coil each has a characteristic shape of CD spectrum. The absorption spectra of peptides in different conformations have different profiles, for example, a typical CD spectrum of one β-sheet peptide has maximum at 200 nm and minimum at 216 nm.

a proline residue at the center of the template appeared to be an ideal technique, because the fragments should maintain their interactions with template on both sides of the kink, but after ligation the product would reduce the affinity within template complex.

In this work, two self-replicating coiled-coil peptides, XL-1 and XL-2, were designed, also based on the self-replicating peptide E1E2, with each containing one proline residue. XL-1 was made from E1E2 by replacing Leu19 of E1E2 with a proline, XL-2 made from E1E2 replacing Glu20 with a proline. Two peptide fragments was also synthesized for each peptide, one fragment bearing a C-terminal thioester, the other containing an N-terminal cysteine. Ligation of two fragments to produce a template was also performed by native chemical ligation reaction. The sequences of the two peptides and their fragments are shown below (sequence starting point is in g site of the heptad repeat).

$$\begin{aligned}
&\mathsf{E1E2}: \mathsf{Ac-ELYALEKELGALEKELACL^{19}E^{20}KELGALEKELYALEK-CONH_2}\\
&\mathsf{XL\text{-}1}: \mathsf{Ac-ELYALEKELGALEKELACP^{19}E^{20}KELGALEKELYALEK-CONH_2}\\
&\mathsf{XL\text{-}2}: \mathsf{Ac-ELYALEKELGALEKELACL^{19}P^{20}KELGALEKELYALEK-CONH_2}\\
&\mathsf{XL\text{-}3}: \mathsf{Ac-ELYALEKELGALEKELA-COSR}\\
&\mathsf{XL\text{-}4}: \qquad\qquad\qquad\qquad\quad \mathsf{CPEKELGALEKELYALEK-CONH_2}\\
&\mathsf{XL\text{-}5}: \qquad\qquad\qquad\qquad\quad \mathsf{CLPKELGALEKELYALEK-CONH_2}
\end{aligned}$$

The replacement of Leu19 or Glu20 with proline should affect the stability of the coiled coil structure due to induced distortions in the helices. This was confirmed by evaluating the α-helical content of the peptides using circular dichroism spectroscopy. XL-1 and XL-2 showed helical contents of 45% and 55% respectively, while for E1E2, this number is 85% under identical conditions. Moreover, the helical contents of peptide fragments XL-3, XL-4, and XL-5 increased distinctly after adding corresponding templates (XL-3 and XL-4 increased 39% and 42% respectively upon addition of XL-1; and XL-3 and XL-5 increased 38% and 19% after adding their template XL-2), which indicates the possibility of the catalysis of the templates. Besides different helical contents, the melting temperatures of the modified peptides measured by using thermal denaturation were lower than E1E2, with XL-1's melting temperature 30 °C lower than XL-2.

The self-assembly status of these two peptides are also different from the original E1E2. Analytical ultra-centrifugation measurement showed that XL-1 aggregated as an octamer, while XL-2 existed as a tetramer.

The self-replicating properties are also different for XL-1 and XL-2. The ligation of fragments XL-3 and XL-5 to form template XL-2 was accelerated by the presence of the template: the initial reaction rate of template formation was increased with increasing amount of template in the reaction mixture. On the contrary for XL-1, the reaction between fragments XL-3 and XL-4 was very slow, and adding template had no influence on product formation. This shows that XL-2 is an autocatalytic self-replicator, where as XL-1 is not. The reaction order of self-replication of XL-2 was found to be 0.91. This value is much higher than the predicted value 0.75 considering product inhibition of tetrameric XL-2. Chmielewski also classified this reaction as weakly exponential.

It is interesting that on one hand both peptides are capable of templating coiled coil formation of the corresponding fragments to increase their helical contest; on the other hand, XL-2 displays a high catalytic efficiency of self-replication while fragments of XL-1 do not tend to ligate, even in the presence of template. The different behaviors of two peptides were thought to be caused by the different positions of prolines in the helices, namely d site proline for XL-1 and e site for XL-2. This difference should induce different directions of 30 ° bending of the helical segments on one side of the proline away from segments on the other side. For XL-1 this bent would break the hydrophobic surface of the coiled coil since the original hydrophobic d site is disconnected by the bent caused by proline, placing the termini of the two fragment peptides in a bad orientation for ligation reaction.

Whereas for XL-2, this bent at position e would not break the hydrophobic surface but would reduce the repulsion between original negative charged e residues, resulting in a more favorable positioning for ligation of the fragments.

This experiment showed us that introducing prolines into the helices seems a effective strategy for designing a self-replicating α-helical peptides. The position of the proline kink in the helix is critical in this design.

In conclusion of this part, two cases of high reaction order self-replicating α-helical peptides are explained, one with shortened length of 26 amino acids, the other with the insertion of one proline amino acid leading to kink in the helix.

3.3.3 β-sheet peptides

We have seen that by shortening the length of the self-replicating peptides, scientists succeeded in improving the catalytic efficiency of the self-replicating process. Moreover, short peptides are able to form well-defined structures. Bearing these facts in mind, one may speculate that with even shorter and properly designed peptides, highly efficient self-replication could probably be achieved. Following that speculation, people have successfully demonstrated some short peptides, which form β-sheet assemblies, can also catalysis formation of themselves.

β-sheet is another major type of secondary structure of proteins, comprised of several individual β-strands. In each β-strand, the backbone atoms of protein (or polypeptide) shape into a zigzag pattern, making each sides of the strands have alternative carboxyl group oxygen atoms and amide group hydrogens. This also allows energetically favorable hydrogen bonds to form between two β-strands at either side of the strand. Normally two adjacent strands have antiparallel directions. A β-sheet often has a hydrophilic side and a hydrophobic side.

A preliminary construction of self-replicating β-sheet peptides was given by Mihara and coworkers in 2003 [15]. In their demonstration, they designed a 17-amino acid β-strand peptide EKCEK (sequence Ac-YGGALEQKLGCLEQKLA-NH_2) which can self-assemble into β-sheet amyloid-like fibrils. This macromolecular amyloid fibril recognizes the same single polypeptide molecules that composes the fibril itself, and organizes the peptide fragments EK-SBn and CEK (sequences Ac-YGGALEQKLG-SBn and CLEQKLA-NH_2 respectively) into a β-sheet to form a new unit of the fibril. The fibril thus extends the length of itself, but maintains the active sites for the next circle of reaction.

A schematic view of such self-replicating systems is shown in Figure 9. The reaction cycle can be

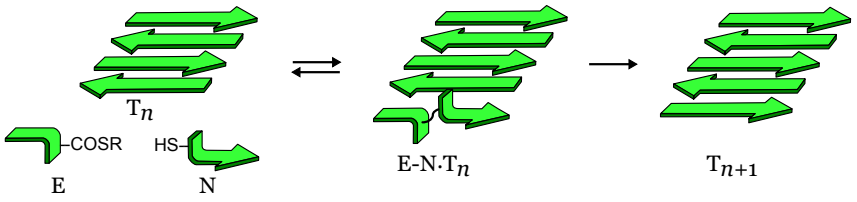


Figure 9: Schematic show of self-replicating β-sheet peptides. An electrophilic fragment peptide and a nucleophilic fragment are organized by the β-sheet template; native chemical ligation is carried out in the condensation of the fragments, the production of this ligation is a β-sheet peptide which is the same as the unit that composes the fibril. The generated peptide then acts as the new templating site for the next round of association.

expressed by Eq.7. Notice that when the template complex gets to a critical length, if can dissociate

into two smaller assemblies, thus increasing the number of catalytic sites.

$$\mathrm{E} + \mathrm{N} + \mathrm{T}_n \rightleftharpoons \mathrm{E{\cdot}N{\cdot}T}_n \ ; \tag{7a}$$

$$\mathrm{E{\cdot}N{\cdot}T}_n \rightarrow \mathrm{T}_{n+1} \ ; \tag{7b}$$

$$\mathrm{T_L} \rightleftharpoons \mathrm{T_S} + \mathrm{T_S} \ . \tag{7c}$$

Back to the self-replicating amyloid-like fibril experiment. The authors argued that the self-replication of the β-sheet template peptide from conformationally disordered fragments can be viewed as a replication of both spatial conformation and chemical composition. In other words, the template fibrils not only catalyze the formation of the peptides, but also help to fold the peptides to the fibril form. This leads to the self-amplification of the fibrils. The problem with this design is that the catalytic efficiency of this design was not high though.

Inspired by the fibril replication work, or maybe not, Ashkenasy and coworkers gave the demonstration of a high efficient self-replicating amphiphilic β-sheet peptide **1** in 2009 [16]. In this work, two relative simple peptides **1** and **2** were designed, their structure analogous to the synthetic amphiphilic Glu-(Phe-Glu)$_n$ peptides. These two peptides are able to form one-dimensional β-sheets in water, while only **1** is able to accelerate ligation reaction of its fragments and thus accomplishing self-replication.

The design was based on the fact that peptides with alternative hydrophilic and hydrophobic amino acid residues (e.g. in this case hydrophilic Glu and hydrophobic Phe) tend to adopt β-sheet conformation. Also noticing the effect of a terminal proline residue on breaking a β-sheet, peptide **2** bearing the sequence Aba-Glu-(Phe-Glu)$_5$-Pro was designed and synthesized ("Aba" in short for acetamidobenzoate). Further, in order to employ native chemical ligation for condensation reaction, peptide **2** was modified into peptide **1**, sequence Aba-(Glu-Phe)$_2$-Ala-Cys-(Glu-Phe)$_2$-Glu-Pro. Thus peptide **1** can be from two fragment peptides, electrophilic fragment $\mathrm{E_1}$ (Aba-(Glu-Phe)$_2$-Ala-COSR, where R=$\mathrm{CH_2CH_2CONH_2}$) and nucleophilic $\mathrm{N_1}$ (Cys-(Glu-Phe)$_2$-Glu-Pro).

The β-sheet structures formed by peptides **1** and **2** at air-water interface and in aqueous solutions was examined by several techniques. Grazing incidence X-ray diffraction of layers of peptides **1** and **2** on deionized water surface showed two-dimensional lattice patterns. Lattice constants were calculated from diffraction spectra as $(a_1, a_2) = (58.1\,\text{Å}, 4.8\,\text{Å})$ for peptide **1** and $(51.9\,\text{Å}, 4.8\,\text{Å})$ for **2**. The 4.8 Å spacing corresponds to the bond length inter-strand hydrogen bond, while the 58 Å and 52 Å distances agree with the projected length of β-sheet on water surface, with peptide **1** less alternation because the change of sequence from **2**. And Langmuir surface pressure vs. area (π-A) isotherms of peptides **1** and **2** were measured, both peptides having limiting area of $\sim 250\text{Å}^2$, indicating they form stable monolayers. Moreover, circular dichroism spectra confirmed the major β-sheet contents of both peptides, and also some α-helical content for peptide **1**. Additionally, cryogenic transmission electron microscopy (cryo-TEM) showed long and entangled fibril structure of peptide **1** formed in $100\,\mu\text{M}$ aqueous solution.

The self-replicating properties of peptide **1** was investigated. In the ligation reaction with equal concentration of fragments $\mathrm{E_1}$ and $\mathrm{N_1}$, the reaction rate increased significantly after initial several hours. While a control experiment on a modified peptide 1^{gg} (sequence Aba-Glu-Phe-Gly-Phe-Ala-Cys-Glu-Phe-Gly-Phe-Glu-Pro), which comes from peptide **1** by introducing two glycine residues to replace two glutamate residues thus does not form nice β-sheet, showed a stable low reaction rate. Furthermore, when seeded with different amount of **1** in the ligation reaction, the production rate enhanced with increasing seeding concentration; while seeding with peptide 1^{gg} or **2**, the production of **1** is relatively slow (although in the beginning 3 hours the catalytic effects of 1^{gg} and **2** are more

significant, which may comes from cross-catalytic behaviors between **1** and 1^{gg} or **2**). This suggest that peptide **1** can auto-catalyze the ligation reaction of itself specifically.

The reaction order of self-replication of peptide **1** was found to be $p = 1.2$ *, calculating from the linear fit of the logarithm of the initial reaction rate and the logarithm of concentration. This reflects an exponential type of growth of the self-replication process, suggesting the highly efficient auto-catalysis of peptide **1**.

The nonlinearity of the catalysis, $p > 1$, reflects that production rate only depends on the number of the β-sheet aggregates but independent of the length of the fiber-like assemblies. It was postulated that the underlying mechanism is that after reaching a critical size, the β-sheet aggregates tend to dissemble into two smaller aggregate. Thus, the reaction is intensified more at higher concentration of template peptide, since under high concentration the β-sheet assembly tends to break more easily because of more collision between the aggregates, as a result the number of catalytic sites grows faster than the concentration of the peptide. This hypothesis was supported by the fact that sonication, thus breaking more aggregates, apparently increased the reaction rate in the first hour of ligation reaction seeded with mediate and high concentration of **1**. To further support the theory, dynamic light scattering measurements were performed to show the size distribution of template aggregates at different concentrations. The results revealed that mainly one type of large aggregates (e.g. $200 \sim 300$ nm in water) are formed, meanwhile two types of smaller aggregates ($15 \sim 20$ nm and $60 \sim 80$ nm) also exist; increasing amounts of smaller aggregates form as lowering the concentration. Thus, it is most likely that after reaching a critical size, large aggregates much dissemble into smaller pieces.

This study of amphiphilic β-sheet peptides expands the repertoire of synthetic self-replicating molecules, and opens up new possibilities of achieving high reaction order self-replication by using oligomeric form of template molecules instead of just the monomer. Remarkably, the reaction order of this self-replication process using oligomeric template easily achieved exponential, which is much larger than the cases of using monomer templates. Ashkenasy argued by using a simplified mathematical model that this phenomenon may relate to the fact that oligomeric type of templates can have more asymmetry between binding and dissociating of the template-product complexes [4].

Also notice that one may classify this self-replication process as cross-catalytic, considering the product of the reaction is a monomer while the template is not the monomer peptide itself but the oligomer of the peptide, which is of course not equal to the monomeric peptide.

3.3.4 From β-sheet peptides to non-biological self-replicating molecules

After the successful demonstration of self-replication in β-sheet conformation peptides, non-biological molecules that form similar fiber-like aggregates of β-sheets are also shown capable of carrying out self-replication. An interesting example was shown by Otto and coworkers in 2010 on Science [17].

In this work, a self-binding peptide derivative (**1** in Figure 10) featuring a peptide sequence with alternating hydrophilic (lysine) and hydrophobic (leucine) amino acids was used. This building block can generate a small dynamic combinatorial library†, in which several macrocycles held together covalently by different number of **1** molecules through oxidative disulfide formation from thiol groups. It was shown that different modes of agitation during reaction, either stirring, shaking or without agitation, will give different distributions with different dominating species of macrocycles, indicating a kinetically controlled dynamic combinatorial library. Without agitation, the final composition

*Note that this value has a large uncertainty since it's from the linear fit of 4 data points, error bar not shown in the paper though.

†A group of molecules assembled from building blocks reversibly, giving a distribution of inter-converting library members often under thermodynamic control.

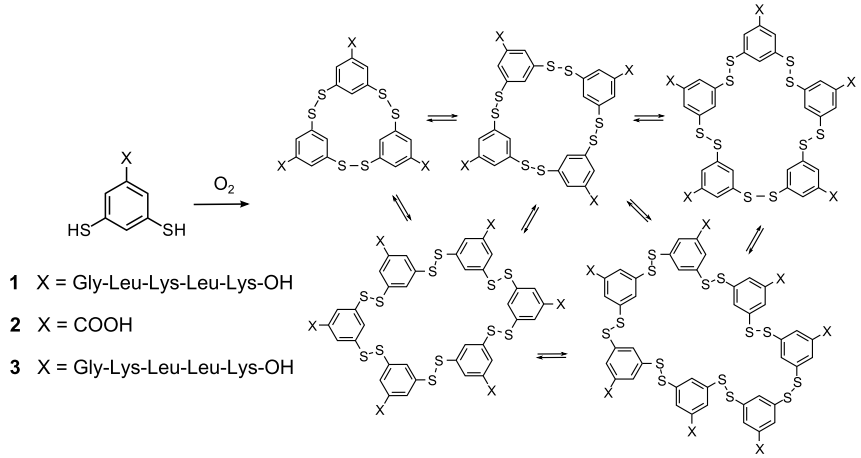


Figure 10: Schematic illustration of a small dynamic combinatorial library made from dithiol building blocks.

is a mixture of cyclic trimer and tetramer mostly; while under stirring, the cyclic heptamer rapidly becomes the dominant product and in the meantime consuming the population of timer and tetramer; remarkably when repeating the experiment using shaking, the cyclic hexamer is preferentially induced instead of heptamer.

The self-replicating properties of the production of hexamer and heptamer were examined by analyzing the time variance of the production rate. The sigmoidal shape growth of hexamer and heptamer suggests the template effect of them in their own formation. When seeding with a small amount of hexamer or heptamer in the beginning stage of each reactions, they observed a clear inducing of more corresponding macrocycles.

The hexamer $\mathbf{1}_6$ and heptamer $\mathbf{1}_7$ can form amyloid fiber-like structures which can possibly catalyze the formation of the macrocycle molecules themselves, just like the case of self-replicating β-sheets shown above. (The active templating effect of the fibers on the formation of the macrocycles were not examined in detail though.) The formed fiber structures were apparent under cryo-TEM, with lengths up to 1 to 2 μm and diameters about 4.7 to 4.9 nm, in agreement with the diameter of a single macrocycle assuming peptide chains in β-sheet conformation. Additionally, the β-sheet components in hexamer and heptamer macrocycles were further verified by circular dichroism spectroscopy. The spectrum of solution containing mainly hexamer and heptamer shows typical features of β-sheet peptides, while the spectrum of solution containing mainly trimer and tetramer exhibits characters typical for random coils. Further evidence of β-sheet formation was given by fluorescence microscopy using fluorophores that show characteristic change in spectra when bound to extended β-sheets.

To verify the role of β-sheet structure on the preference of hexamer and heptamer production, control experiments were carried out using different building blocks **2** (Figure 10) which does not have a peptide chain, or **3** (Figure 10) which has a peptide chain of different sequence thus doest not form β-sheets. As a result, building blocks **2** and **3** both gave a mixture dominated by cyclic trimer and tetramer, when **2** was stirred and **3** in any the three agitation modes. The absence of large amounts of hexamer or heptamer for **3** indicates that the formation of these larger macrocycles is a sequence selective process.

Based on these observations, a self-organizing mechanism of formation of hexamer and heptamer macrocycles was proposed as schematically shown in Figure 11. Stacks of cyclic hexamers are brought

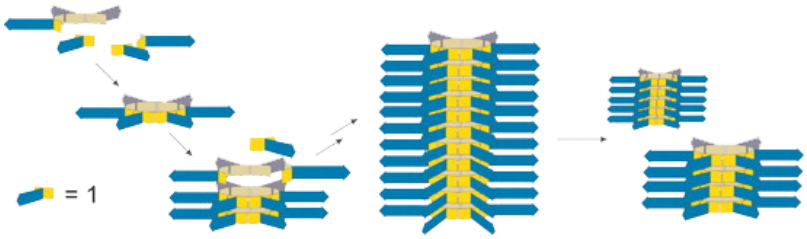


Figure 11: Schematic representation of the proposed formation of fibers of $\mathbf{1}_6$. The benzenedithiol cores of building block **1** in yellow and the peptide chains in blue. Figure taken from Ref[17].

together by self-assembly of the β-sheet peptide tails, forming elongated cross-β sheets*. The fiber of stacks grows longer by attachment of new stacks to both ends of the fiber. Then the fragmentation of the stacks (breaking of the fibre) creates more active ends, enabling exponential growth. Since the fragmentation of the stacks is a mechanosensitive process, different modes of agitation would result in different size distribution of the fibers, thus creating different numbers of active sites for binding of single macrocycles. This should have an influence on the reaction rate of the hexamer and heptamer production: faster reaction in more strongly agitated sample. This is confirmed by investigating length distribution of fibres by cryo-TEM, and by comparing the reaction rates of samples using seeds processed with different agitation modes. The result showed similar distributions in length for the hexamer and the heptamer, with shaking giving shorter fibers. And sample seeded with agitated hexamer seed indeed had a faster growth rate than non-agitated sample.

To further investigate the impact of different agitation modes on the final dominating species, equal amounts of hexamer and heptamer were added to one solution containing mostly trimer and tetramer as their sources, and the winning species were inspected after 7 days reaction under different agitating conditions. The result is: under stirring or without agitation, the heptamer out-competes the hexamer, and only under shaking does the hexamer have a competitive reaction rate. The differences of the hexamer and the heptamer are listed below.

- The heptamer is more efficient in linear growth process of fiber formation.
- The hexamer fiber is more easily fragmented under violent agitation (shaking) since it has fewer β-sheets formed, although both the two have similar degree of fragmentation under mild agitation (stirring).

Thus, under stirring conditions, the heptamer dominates because it's more efficient elongating the fiber; and under shaking conditions, hexamer wins because there are larger amounts of fragments.

To conclude, this work is a nice example of dynamic combinatorial libraries under kinetic control, in which several macrocycles can exist in a distribution that is dependent on the means of agitation. It also demonstrated nonbiological molecules derived from β-sheet peptide have the potential to undergo self-replications through elongation and breaking of β-sheet like fibers served as templates in the self-replicating process. Such type of dynamic combinatorial libraries featuring self-assembly under kinetic control opens up new opportunities in systems chemistry, an emerging branch of chemistry, and provides possible answers to the chemical origin of life.

*A quaternary structure characterizing amyloid fibers, in which several β-sheets stack together forming the breadth of the fiber, and large numbers of β strands align side to side elongating the fiber.

4 Conclusions and discussions

The origin of life is one of the most significant and challenging problem facing modern chemistry and biology. Unlike traditional biology researches, which follow a top-down fashion to unveil the physical and chemical mechanisms behind biological activities inside one cell, the researches of synthetic molecules, which bear similar functions of proteins among other parts of a living cell, give us a bottom-up view of how those functionalities emerge from prebiotic environment. Self-replication, one of the distinctive characteristics of organisms has attracted many scientists to try to develop synthetic replicators over the past 20 years. So far, various types of self-replicating molecules have been synthesized in the lab, from oligonucleotides, to α-helical and β-sheet peptides, to nonbiological molecules. Over the years, manmade self-replicating systems have become more and more complicated, bringing us closer to the answer of how life emerged in prebiotic earth, providing chemists novel systems to study and benefit from at the same time.

5 Acknowledgement

The origin of life has always been a fascinating problem to me. Thanks to the course NS190 and professor Sijbren Otto, I've gained more knowledge concerning the chemical mechanism of the emergence of life, and more importantly I have more confidence of the ability of modern science in answering this ultimate question.

References

[1] S. Miller. A production of amino acids under possible primitive earth conditions. *Science*, 117(3046):528–529, 1953.

[2] Andr Brack. From interstellar amino acids to prebiotic catalytic peptides: A review. *Chemistry and Biodiversity*, 4(4):665–679, 2007.

[3] G. Joyce. Rna evolution and the origins of life. *Nature*, 338(6212):217–224, 1989.

[4] N. Wagner and G. Ashkenasy. Symmetry and order in systems chemistry. *Journal of Chemical Physics*, 130(16), 2009.

[5] N. Paul and G. Joyce. A self-replicating ligase ribozyme. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20):12733–12740, 2002.

[6] G. von Kiedrowski. A self-replicating hexadeoxynucleotide. *Angewandte Chemie-International Edition in English*, 25(10):932–935, 1986.

[7] V. Patzke and G. von Kiedrowski. Self replicating systems. *ARKIVOC*, (Part 5):293–310, 2007.

[8] A. Luther, R. Brandsch, and G. von Kiedrowski. Surface-promoted replication and exponential amplification of DNA analogues. *Nature*, 396(6708):245–248, 1998.

[9] Hana Robson Marsden and Alexander Kros. Self-Assembly of Coiled Coils in Synthetic Biology: Inspiration and Progress. *Angewandte Chemie-International Edition*, 49(17):2988–3005, 2010.

[10] D.H. Lee, J.R. Granja, J.A. Martinez, K Severin, and M.R. Ghadiri. A self-replicating peptide. *Nature*, 382(6591):525–528, 1996.

[11] P. Dawson, T. Muir, I. Clarklewis, and S. Kent. Synthesis of proteins by native chemical ligation. *Science*, 266(5186):776–779, 1994.

[12] S. Yao, I. Ghosh, R. Zutshi, and J. Chmielewski. A ph-modulated, self-replicating peptide. *Journal of the American Chemical Society*, 119(43):10559–10560, 1997.

[13] R. Issac and J. Chmielewski. Approaching exponential growth with a self-replicating peptide. *Journal of the American Chemical Society*, 124(24):6808–6809, 2002.

[14] X. Q. Li and J. Chmielewski. Peptide self-replication enhanced by a proline kink. *Journal of the American Chemical Society*, 125(39):11820–11821, 2003.

[15] Y. Takahashi and H. Mihara. Construction of a chemically and conformationally self-replicating system of amyloid-like fibrils. *Bioorganic and Medicinal Chemistry*, 12(4):693–699, 2004.

[16] B. Rubinov, N. Wagner, H. Rapaport, and G. Ashkenasy. Self-replicating amphiphilic beta-sheet peptides. *Angewandte Chemie-International Edition*, 48(36):6683–6686, 2009.

[17] J. M. A. Carnall, C. A. Waudby, A. M. Belenguer, M. C. A. Stuart, J. J. P. Peyralans, and S. Otto. Mechanosensitive self-replication driven by self-organization. *Science*, 327(5972):1502–1506, 2010.