Review Article

Self-Replicating Systems†

Michael Famulok, James S. Nowick and Julius Rebek, Jr.*

Department of Chemistry, Massachusetts Institute of Technology, Cambridge MA, 02139, USA


Som considerable time ago, chemistry and physics spawned biology and it is likely that a single molecule existed at that interface. While the attributes of that molecule have stimulated much debate and speculation, one property is agreed upon: its ability to self-replicate. Many researchers believe that by looking at today’s molecules, RNA and DNA, it is possible to trace back to the structures of the earliest replicators. Others believe that the road is ‘washed out’, while still others place their faith in synthetic systems as models. In this article, we review some of the work with nucleic acid components as replicating systems and then describe in some detail our own studies.

Self-replicating systems based on oligonucleotides

The past five years have seen increasing efforts to devise nucleotides that replicate without the aid of enzymes. The finding that RNA has the capability to function both as a carrier of genetic information and as a catalyst has inspired much of this research. This discovery has also made RNA the most likely candidate for the first prebiotic self-replicating molecule.1

Early model studies by Orgel and coworkers support this theory.2 When guanosine 5’-phosphorimidazole I, a polycytidylic acid template, and Zn2+ catalyst are allowed to react, 3’-5’-linked oligonucleotides longer than (polyG)10 form.3 Under the same conditions, Pb2+ efficiently catalyzes the formation of exclusively 2’-5’-linked oligomers.2e When poly(C) was incubated with a mixture of equimolar amounts of the activated guanosine G, adenosine A, uracil U and cytosine C monomers, G was incorporated 200 times more efficiently than the ‘incorrect’ nucleotides. The pentahomopolymer poly(C), facilitated the synthesis of mainly G-containing oligonucleotides from a mixture of activated C and G whereby the complementary polyguanylic acid [poly(G)] was formed in 17% yield.2e When random copolymer templates with C as the major component were used, nucleotides were incorporated into the product when its complement was present. With poly(CU), for example, only I-A and I-G oligomerized. The products formed had a mean chain length of six nucleotides and the incorporation of the nucleotides occurred with very high fidelity.2d Limited regions of each template were found to be copied more accurately than others, leading to the conclusion that only a restricted range of oligomers were efficient templates. In these examples, the newly synthesized strand does not separate from the template strand, so template molecules are not used more than once.

![Diagram of 1 Base = U, C, A or G]

An example of self-replicating oligonucleotides was given by von Kiedrowski, who showed the self-replicating capability of a protected hexadeoxynucleotide,4 the palindromic template, 5’-H2CO-d(CpCpGp-CpGpGp)-O-(2-CIC3H7) 2, was formed in presence of the carbodiimide ethyl(dimethylaminopropyl)carbodiimide (EDC) 3 [CH4.CH2.N=C=N-(CH3)2-N(CH3)2] from the complementary trideoxynucleotides 5’-H2CO-d(CpCpGp) 4 and 5’-HO-d(CpGpGp)-O-(2-Cl-C6H5) 5 to a small extent via an autocatalytic pathway. The reaction order for 2 in the kinetics of its own formation was ½, and the amount of product synthesized increases with the square root of the template concentration. The major by-product was the pyrophosphate 3’p-GGC>5’pp>‘CGGp>”, resulting from the self-condensation of 5. When the more nucleophilic 5’-H2N-d(CpGpGp)-O-(2-Cl-C6H5) 6 was used instead of 5, together with the template 2, the related aminotemplate was found to be the only product. Kinetic studies of this system revealed that the time course of the formation of 2 follows a sigmoidal curve.2e

1 Presented as a plenary lecture at the third European Symposium on Organic Reactivity in Göteborg, Sweden, July 7–12, 1991.
* To whom correspondence should be addressed.


315
By using an oligonucleotide template based on pyrophosphate linkages, the synthesis of the corresponding pyrophosphate product strand was achieved by Visscher et al. The oligomerization of the monomer 10 was studied in the absence and in the presence of the template poly-d (CpC)₄ 11. The cyclization of the monomer was the primary reaction in the absence of the template. When 11 was present, the total yield of oligomers doubled, including the formation of pyrophosphate-linked oligomers of at least nine nucleotides in length. Nucleic acid analogs with a pyrophosphate backbone were proposed as the possible precursors of the first RNA molecules, which makes this result especially relevant for theories on the origins of life.

An approach towards a self-replicating RNA was described by Doudna and Szostak, who demonstrated the synthesis of a complementary strand RNA on an external template, catalyzed by the Tetrahymena ribozyme. The key step was the separation of the template region of the ribozyme (the stem loop of the isolated internal guide sequence) from the catalytic center (the part of the RNA that catalyzes the ligation reaction). The reactants were short oligonucleotides, nine to ten bases in length. These oligonucleotides line up on the external template, and are ligated by the catalytic part. The consequence of the separation of the template region from the catalytic center is a weakened interaction of the substrates with the catalytic region. After ligation, the complementary strand pairs with the template, and the guanosine formed during the ligation is rapidly released from the enzyme. This enables another catalytic cycle to begin. The full-length 40 nucleotide product strand, however, was obtained in only 1% yield. Such low efficiency makes it impossible to realise self-replication of a replicase as large as the Tetrahymena ribozyme (413 bases in length).

A remarkable improvement of this reaction was achieved with a modified version of the self-splicing sunY intron from bacteriophage T4. For this study, the intron was divided into three short RNA fragments, 59, 75 and 43 bases in length. When five oligonucleotide substrates complementary to one of the mutant sunY subunits were used, the multisubunit ribozyme catalyzed the synthesis of the subunit's own antisense strand. Ligation of the substrates even occurred when no additional template beyond that in the enzyme complex itself was added. Thus, the enzyme complex was sufficiently stable to function as a ligase, in
the presence of the synthesized antisense strand, while some fraction was unfolded enough to allow the annealing of the substrates. These experiments suggest that it might be possible to assemble RNA oligomers of modest length, which can assemble to give a RNA replicase.

Replication in synthetic model systems

In our laboratory, we have chosen to forsake oligonucleotides in favor of simpler synthetic molecules of our own design. We use Nature as an example of what is possible, rather than as a blueprint. Part of the challenge is taking words such as regulation, catalysis, recognition and replication and working out what structural features might result in these behaviors. These words have no structural content, and its up to the imagination of the researcher to create systems or molecular devices that express such phenomena. Our approach is minimalist, one that defines the simplest requirements. We have identified that replication invariably involves molecules that are complementary. One could even say that a principle of self-complementarity exists. We take comfort in the possibility that this idea, like so many others of bio-organic chemistry, may be traced to Linus Pauling.10

In Scheme 1, we show a simple cartoon, a two-dimensional example of how self-complementarity operates. Within the circular figure, the sigmoid line represents a weak intermolecular contact between the two self-complementary (and here, identical) components. If these units are broken along the jagged line, the fusion of any of the new pieces leads to a shape that can act as a template for the assembly of an identical subunit. At first glance, then, such simple notions should be easily reduced to real structures and we are somewhat embarrassed by the amount of time that it took us to do so. Let us begin with the intermolecular forces and the general types of molecules involved.

Like others, we have used base pairing to provide the recognition and the organization of the two components with respect to one another. The binding event involves both hydrogen bonding and aryl stacking interactions in organic solvents.11 We make no apologies for working in non-aqueous media. Indeed, it has been quite useful in exposing the subtle and intrinsic, such as secondary, effects12 in hydrogen bonding patterns that are so difficult to observe in aqueous solution. Moreover, the magnification of association constants available in non-competing solvents such as chloroform ensure high levels of association. As a result, it contracts the timescale of our experiments to manageable lengths. To be sure, chloroform is not likely to be the prebiotic soup, but so what? We attribute the difficulties encountered by previous workers to their insistence on the use of aqueous media. There, association effects are generally so small that only very long reaction times and large structures could reveal their expressions in self-replication.

The specific structures involve derivatives of Kemp's triacid 1213 (Scheme 2) and feature both hydrogen bonding edges and aryl stacking interactions. These converge from perpendicular directions in receptor 14 and provide a complementary microenvironment to adenine derivatives. We have spent considerable effort in mapping how the two components fit together and these are described in experimental detail elsewhere.14 For the present purposes, structure 15 represents these systems adequately. Base pairing is shown in the Watson–Crick sense, and a bifurcated hydrogen bond is involved in binding the amino function of adenine. These complexations are very rapid on the NMR timescale at room temperature. Association constants are on the order of 100 M⁻¹ and involve contributions from not only Watson–Crick, but also Hoogsteen base-pairing modes.

The complementary binding surfaces provide the free energy for association and it should have been an easy
matter to position complementary reactive functionality in both components that could lead to a covalent bond to be formed within the complex. Nonetheless, our early attempts were dismal failures. For example, base pairing does take place between the 9-(aminopropyl)adenine 17 and receptor 16, and intramolecular aminolysis does occur. However, the resulting product 19 remains folded shut (Scheme 3). The short aryl spacer can accommodate a perfectly stable trans amide bond while maintaining base pairing, specifically as shown in the Hoogsteen sense. These intramolecular contacts were established by NOE experiments.

Our second system also failed, and it seemed like the whole project would face early extinction. Here, the product molecules remained stuck together in a bimolecular fashion as the dimer. Though the spacer was longer, the self-affinity of the product was so high that its ability to act as a template was severely limited. Nonetheless, the general shape of structures (shown in Scheme 4) suggested that building a large bulge in the center of the molecules might diminish their self-affinity. Specifically, the 5'-aminoadenosine derivative 20 was prepared and coupled to an active ester bearing a naphthalene surface 21a. The expectation was that the ribose acetonide subunits of the resulting adduct 22a would provide spherical, lumpy bulges, and destabilize the dimer 22a · 22a.

This indeed proved to be the case. For example, the dimerization constant of 22a was a manageable 630 M⁻¹. Even simpler systems established that base pairing for individual components was on the order of 60 M⁻¹. Since the two ends of the structure were not expected to communicate, one might have anticipated an association constant for the dimer of greater than (60)². That is, the chelate effect would enhance the association, since translational entropy involved in each base pairing would not have to be paid for twice. Apparently, the bulge was doing its job. As for geometry, NOE experiments showed no particular bias in the arrangement of the dimer; both Watson–Crick (Scheme 5, 22a · 22a) and Hoogsteen (not shown) senses were represented among the populations of dimeric molecules. For
the active ester component, the pentafluorophenyl derivative gave coupling rates that were on a compassionate and humane timescale. HPLC provided reasonable accuracy to determine the rate of product formation.

Preliminary studies established that added product template 22a indeed speeded up the reaction of 20 and 21a, whereas addition of 2,6-(diacetylamino)pyridine slowed the coupling rate down. Moreover, the corresponding N-methyl imide derivative 21b, incapable of base pairing, showed greatly reduced coupling rates. The initial rates of product formation were linear. The results are plotted in Fig. 1.

These results indicate that self-replication is occurring. Preassociation of the reactants and template to form the complex 20 \(\cdot\) 21a \(\cdot\) 22a leads to intramolecular acyl transfer in the coupling event (shown in Scheme 6). This is the act of replication, since the new structure is an identical copy of the template.

In actuality, three distinct mechanisms generate template 22a: (1) a background bimolecular reaction, in which aminolysis takes place without base pairing; (2) a base-paired bimolecular reaction, in which the amine 20 and ester 21a form a hydrogen-bonded dimer 20 \(\cdot\) 21a prior to reaction; (3) the termolecular or template-catalyzed process (shown in Scheme 6), in which the amine 20, ester 21a and template 22a form a trimeric complex 20 \(\cdot\) 21a \(\cdot\) 22a prior to reaction. The following equations represent these pathways. Only the latter pathway is responsible for self-replication and sigmoidal product growth. The observed kinetic curves are a result of the interplay of all three reactions.

\[
\begin{align*}
20 + 21a & \rightarrow 22a \\
20 + 21a & \rightarrow 20 \cdot 21a \rightarrow 22a \\
20 + 21a + 22a & \rightarrow 20 \cdot 21a \cdot 22a \rightarrow 22a \cdot 22a \rightarrow 22a \rightarrow 22a
\end{align*}
\]
The base-paired mechanism involves complex 20·21a, in which the amino and pentafluorophenyl groups are brought into proximity and react readily to generate the template. Because of the length of the aromatic spacer (the naphthalene), the product is initially formed in the less stable cis-amide conformation. Isomerization to the trans conformation follows. Scheme 7 summarizes these processes. In contrast, the termolecular process initially generates the template trans-amide conformation.

From the experiments carried out with varying amounts of added template, it is possible to extract kinetic parameters for all three reactions. Analysis of the data presented in Fig. 1 afforded rate constants that fit the data quite well. The background bimolecular reaction occurs with $k_1 = 0.023 \text{ M}^{-1} \text{ min}^{-1}$, the bimolecular base-paired process with $k_2 = 0.0036 \text{ min}^{-1}$, and the template-catalyzed process with $k_3 = 0.070 \text{ min}^{-1}$. This permits a calculation of the effective molarity of the reacting groups in the various complexes. For example, the effective molarity in the termolecular complex is actually higher (3.0 M) than in the
bimolecular complex (0.16 M). This 19-fold difference probably reflects the higher activation energy involved in the formation of the cis-amide bond.

Although the effective molarity in the template termo-
molecular process is considerably higher than in the base-
paired bimolecular process, a very large fraction of the product is formed by the bimolecular pathway. This occurs because of the difference in molecularity of the two pathways. For example, at 10 millimolar total concentration of 20 and 21a, and 3 millimolar total concentration of template, 67% of the amine and ester are free, 27% are associated, and only 1.5% are present as a termolecular complex. Under these conditions, the relative contributions of the three reactions are 1.0, 4.3, and 4.6, for the background, base-paired and termolecular components, respectively.

Although the reaction of 20 and 21a is autocatalytic, it does not exhibit a sigmoidal product-growth curve. Longer reaction times and higher concentrations of the reactants failed to show the expected sigmoidal behavior. For exam-
ple, Figs 2 and 3 show the results of increased time and concentration. Perhaps this is not surprising; von Kiedrowski\(^2\) and Orgel\(^4\) have also found it quite difficult to observe sigmoidal growth in related systems. Nonetheless, we were convinced that it was possible to devise a system that exhibited sigmoidal growth.

The sigmoid behavior characteristic of an autocatalytic pathway is expected only in systems where *most of the product* is formed by way of autocatalysis. Calculations suggested that increasing the concentration of reactants would increase the contribution of autocatalysis. However, the limiting solubility of these systems prevented us from reaching the appropriate concentrations. If the base-paired bimolecular pathway were not operative a noticeably sigmo-
dial growth curve would be predicted. Fig. 4 shows the calculated effect of shutting down this reaction pathway. Curve a corresponds to the present system, whereas curve b corresponds to a system in which the base-paired pathway is lacking.

Accordingly, we have restructured the system to reduce the bimolecular base-paired pathway. We used the simple artifice of replacing the hydroxynaphthyl spacer with an even longer aminobiphenyl group. Structures 23 and 24 represent the perfluorophenyl ester and template com-
ponents in this system. Although the ester can still base-
pair with aminoadenosine 20, the reactive amino and pen-
tfluorophenyl ester functions reach each other only with difficulty (structure 20⋅23a vs. structure 20⋅21a). The tri-
propyl analog\(^19\) of Kemp's tricatic was used as the structural scaffold, to enhance further the solubility of these systems.

This system exhibits modest sigmoidal growth curves.\(^15\) Fig. 5 illustrates the formation of template upon reaction of aminoadenosine 20 with the biphenyl imide 23a or the N-methyl imide 23b. Curve a corresponds to the formation
of template 24a, and is distinctly sigmoidal. Curve b represents the formation of 24b, and is typical of a simple bimolecular reaction. Only the NH derivatives, which are able to participate in base-pairing, exhibits sigmoidal behavior. Furthermore, addition of 0.20 equivalents of the template enhances the coupling rate by more than a factor of nearly two (Fig. 6). The system exhibits efficient autocatalysis!

It is interesting to note that the structural modifications on going from the naphthyl to the biphenyl spacer do not substantially affect the rates of the background bimolecular and template termolecular processes. The biphenyl spacer does change the position of the ribose bulges, however, and the dimerization constant of the biphenyl template 24a is considerably larger than that of the naphthyl template 22a (80000 m$^{-1}$ vs. 630 M$^{-1}$). This observation has led us to reconsider the effect of a large dimerization constant upon the efficacy of replication. The dimerization constants of the replicating oligonucleotides of von Kiedrowski and Orgel are obviously quite large. The same forces that result in better binding of the template will also lead to better binding of the transition state. We are currently developing self-replicating systems that employ Hoogsteen and Watson–Crick binding simultaneously, in order to develop more efficient replicators.

Fig. 5. Plots of appearance of 24a and 24b vs. time as determined by HPLC. Initial concentrations of 20, 23a and 23b were 50 mM in CHCl$_3$ with 9 equiv. Et$_3$N added. (a) Reaction of 20 and 23a; (b) reaction of 20 and 23b. Boxes represent approximate uncertainty (± 0.1 mmol).

Fig. 6. Plots of initial product formation of 24a vs. time as determined by HPLC. Initial concentrations of 20 and 23a were 25 mM in CHCl$_3$ with 9 equiv. Et$_3$N added. (a) Reaction of 20 and 23a with 0.2 equiv. of template 24a added. (b) Reaction of 20 and 23a without added template.
**Prospects for the future**

Can such systems show evolution? The ability to make mistakes and mutate into more or less effective replicators appears to be one of the requirements for evolution, and we have made some preliminary progress on this goal. For example, using the N-methyladenine derivative 25, a second replicator 26 has been developed. In this compound, the N-alkyl group serves to limit the number and quality of base-pairs to largely Hoogsteen senses. The unsubstituted adenine derivative 22a can also catalyze the formation of the other replicator 26. However, in experiments in which both 25 and 20 compete in the same flask for the active ester 21a, the unsubstituted adenine 20 is the more effective replicator; its population grows at the expense of the N-methylated product. It can catalyze its own formation in both Watson–Crick and Hoogsteen senses, whereas the N-methyladenine 25 is restricted.

Our current plans involve the introduction of mutational-type events that alter the efficiency of replicating systems. In principle, changes in temperature, pH, salt concentration, etc., could be used to trigger such an event. Our plan is to use radiation. Specifically, a photochemically removable blocking group has been incorporated into an adenine nucleophile, and the competition of the two replicators is quite evenly matched. However, shining light on the system and cleaving the photo group introduces a new, unsubstituted adenine which can rapidly take over because of its superior base-pairing abilities.

We conclude that designing replicators is easier than once believed and we point out a simple principle that can be used to generate replicators: covalent attachment of a host to its guest is all that is required. Tethering a crown ether to an ammonium ion, a cyclophane to an aromatic, or a cyclodextrin to an aryl all give self-complementary systems that could act as templates for their own assembly. The situation holds even for metals 27 and their ligands 28; the facile rise of a replicator 20 is depicted in Scheme 9.

Other research in the group has as its goal the use of the informational content of base pairing to help drive acyl-transfer reactions. The intent is to catalyze the formation of peptides with more recognizable shape and functions, i.e. to develop a simple genetic code. We will report on these in due course.

**Acknowledgements.** We are pleased to acknowledge the superb experimental and intellectual contributions of the co-workers whose names appear in the original literature citations. The National Science Foundation is thanked for financial support of this work.

**References**


Abbreviations: In the oligonucleotides discussed here the bases are abbreviated as follows: \( G = \) guanine, \( A = \) adenine, \( T = \) thymine, \( C = \) cytosine, \( U = \) uracil; polyhomonucleotides = (poly\(N\)), \( [N = A, G, C \text{ or } T; \ x = \) number of bases]; polyheteronucleotides = (poly\(NpN\)), deoxyheterooligonucleotides = d(poly\(NpN\)), polypyprophosphate linked deoxyoligomers = poly-d(poly\(NpN\)).


(10) Pauling, L. Molecular Architecture and the Processes of Life; Jesse Boot Foundation, Nottingham, UK 1948.


Received July 24, 1991.