# Studies in Molecular Replication

EDWARD A. WINTNER, M. MORGAN CONN, AND JULIUS REBEK, JR.*

*Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

*Received January 5, 1994*

It was not until we had self-complementary molecules in hand, misbehaving, that it occurred to us that the misbehavior was a message. Interpreting the message led us to an encounter with replication, then to other phenomena generally regarded as biological. This Account views these phenomena at the molecular level from the perspective of structural organic chemistry.

At the outset we distinguish between self-replication and other forms of autocatalysis. There are many examples of the latter, the bromination of acetone and the formose reaction[1] being among the oldest examples. Self-replication is a special subset of autocatalytic reactions in which molecular recognition plays a role. *Self* refers to the mutual recognition event; the product of the reaction is complementary to the starting materials and serves to bring them together to accelerate the reaction. *Replication* implies the directed production of a copy of a molecule. Two different types of systems with these characteristics have been developed: those based on the replication of physical structures and those based on template effects.

The replication of physical structures has been extensively explored by Luisi.[2a] The autocatalytic generation of micelles or reverse micelles in both aqueous and organic media has been observed. The recognition event in these types of systems is the preferential binding of a substrate to the micelle whereby the exposure of the starting material to reagents, which are typically biphasic, is enhanced. The autocatalytic product in these systems, the micelle, is an aggregate of the reaction products with a loosely defined size and structure. This system differs from template-based replication, which is more strictly defined in its requirements of a complementary fit and stoichiometry in the recognition event. The difference is sufficiently great that "self-reproduction" has been proposed as the term for the behavior of the micelle-like systems.[2b]

This Account is based on template effects: the enhancement of reactions by complementary surfaces. These effects lie at the heart of many chemical[3] and biological processes, the most relevant of which were revealed in Watson and Crick's structure of double-stranded DNA. It was clear to them that, during replication, one strand of DNA acted as a template for the other. This feature led to the inspired experiments of Leslie Orgel and his co-workers at the Salk Institute,[4] and there, in 1986, Gunther von Kiedrowski showed that a short, self-complementary segment of DNA could act as a template for its own formation, even without the aid of enzymes. For decades, such experiments have defined the reigning paradigm for prebiotic chemistry: aqueous solutions at gentle temperatures and neutral pH with only those reagents likely to have been available on primitive earth. There are possibilities for molecular replication in organic chemistry that are far outside these constraints, and we have proposed "extrabiotic" as a more appropriate term for them.[6]

Our tools were the weak intermolecular forces which give rise to molecular recognition phenomena: hydrogen bonds, aryl stacking, and hydrophobic effects. These forces cause complexes to form and dissipate rapidly; the short lifetimes of the complexes provided a dynamic environment for our experiments. While tampering with the effects of chemical complementarity, we learned that the very molecules which recognize each other and are stabilized or protected by each other's presence can, with only small modification, become self-replicating structures: replicators.

## Replication

The general mechanism of our replicating systems is depicted schematically in Figure 1, where two complementary components (A and B) react in an intermolecular fashion to form a template (T).

Because of the *self-complementary* nature of the template, two additional units of A and B can form a termolecular complex with the template; weak intermolecular forces permit the template to anchor on its surface and components from which it is made. On the template, the reactants find each other with greater probability than they do in the bulk solution; the

Edward A. Wintner was born outside Philadelphia, PA, son of two Haverford College professors. He survived high school despite breakfast-table tutelage in English, German, and chemistry and went on to study at Yale, where he received a B.S. degree in chemistry. He graduated from Yale in 1991 and has since been a graduate student under Professor Rebek as an NSF predoctoral fellow. Still not tired of academic life, Ed hopes to pursue a career in organic chemistry as a university professor.

M. Morgan Conn was born in 1968 in Ramapo, NY, and raised in Canada. He received his B.Sc. in chemistry and biochemistry from the University of Toronto in 1989. He joined the group of Prof. Julius Rebek, Jr., at MIT as an NSF predoctoral fellow and, subsequently, an NSERC predoctoral fellow and received his Ph.D. in January 1994. Dr. Conn is currently a Miller research fellow in the laboratory of Professor Peter G. Schultz at the University of California, Berkeley, working on the *in vitro* selection of catalytic RNA molecules.

Julius Rebek, Jr., was born in Hungary and received his degrees from the University of Kansas (B.A.) and the Massachusetts Institute of Technology (Ph.D.). He has been on the faculty of UCLA and the University of Pittsburgh and is currently Camille Dreyfus Professor of Chemistry at MIT. His research interests include molecular recognition, replication, and self-assembly.

(1) Mason, S. F. *Chemical Evolution: origins of the elements, molecules and living systems*; Oxford: Clarendon Press, 1991. Weiss, A. H.; Socha, R. F.; Likholobov, V. A.; Sakharov, M. M. *CHEMTECH* **1980**, 643–647. (The formose reaction involves the polymerization of formaldehyde.)

(2) (a) Bachmann, P. A.; Walde, P.; Luisi, P. L.; Lang, J. *J. Am. Chem. Soc.* **1991**, *113*, 8204–8209. (b) P. L. Luisi, Personal Communication.

(3) For a timely review, see: Anderson, S.; Anderson, H. L.; Sanders, J. K. M. *Acc. Chem. Res.* **1993**, *26*, 469–475. See also: Kelly, T. R.; Bridger, G. J.; Zhao, C. *J. Am. Chem. Soc.* **1990**, *112*, 8024–8034. Goodwin, J. T.; Lynn, D. G. *Ibid.* **1992**, *114*, 9197–9198.

(4) Orgel, L. E. *Cold Spring Harbor Symp. Quant. Biol.* **1987**, *52*, 9–16. Joyce, G. F. *Ibid.* **1987**, *52*, 41–51. For recent advances with RNA hairpins, see: Wu, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1992**, *114*, 7963–7969.

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

(6) Rebek, J., Jr. *Chem. Ind. (London)* **1992**, 171–174.

0001-4842/94/0127-0198$04.50/0 

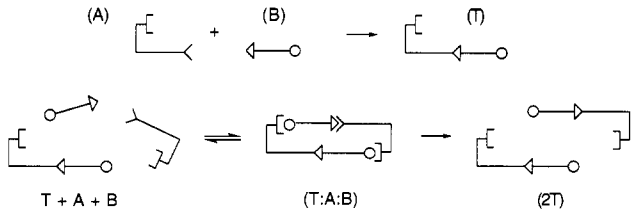


**Figure 1.**

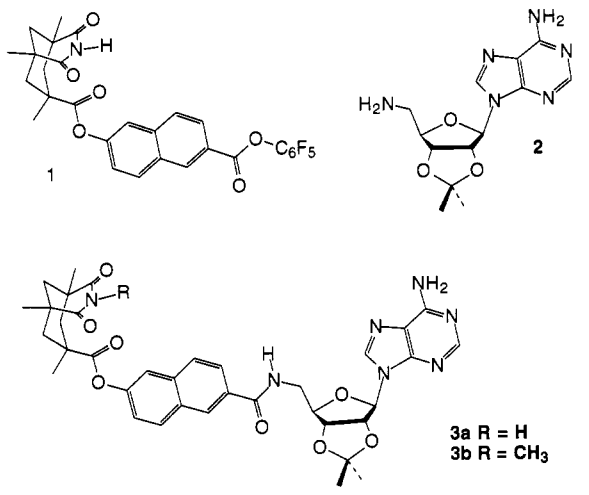


**Figure 2.**

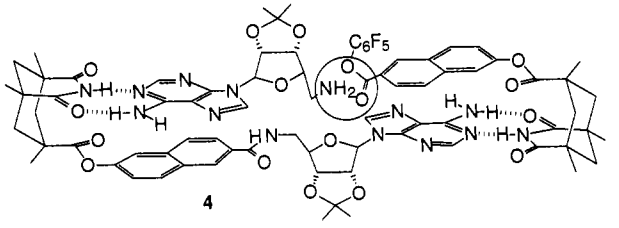


**Figure 3.**

template reduces the entropy of the process. Intracomplex reaction forms a dimer (2T), and when the fusion of the components is complete, the weak intermolecular forces allow dissociation of the dimer. This results in increased template concentration: the reaction is thus autocatalytic.

Our first self-replicating system was based on the hydrogen-bonding recognition of adenine and an imide of Kemp's triacid.[7] The naphthoyl active ester 1 reacts with 5′-amino-5′-deoxy-2′,3′-isopropylideneadenosine (2) to form the self-complementary autocatalytic template **3a** (Figure 2). The autocatalytic nature of the reaction is evident from the rate acceleration caused by seeding the reaction with its product (Figures 4 and 5). We proposed that the autocatalysis observed is the result of the product's ability to gather on its framework the two components of which it is formed (Figure 3).

The many geometric possibilities for a productive complex are simplified in the figure. Whatever the details of base-pairing (Watson–Crick, Hoogsteen, reverse Watson–Crick, etc.), the amine and activated carbonyl are then within relatively easy reach of each other, and covalent coupling gives a replica of the template catalyst in dimeric form.

(7) Tjivikua, T.; Ballester, P.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1990**, *112*, 1249–1250. Nowick, J. S.; Feng, Q.; Tjivikua, T.; Ballester, P.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 8831–8839.

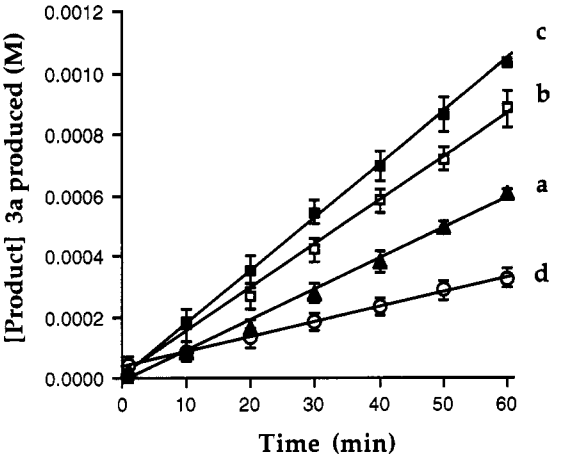


**Figure 4.** Reaction of 1 and 2 in $CHCl_3$ as followed by HPLC; 8.2 mM initial concentrations of 1 and 2, 1% TEA base added: (a) base-line reaction (1 + 2); (b) base-line reaction plus 0.2 equiv of product (**3a**); (c) base-line reaction plus 0.5 equiv of product (**3a**); (d) base-line reaction plus 1.0 equiv of 8.

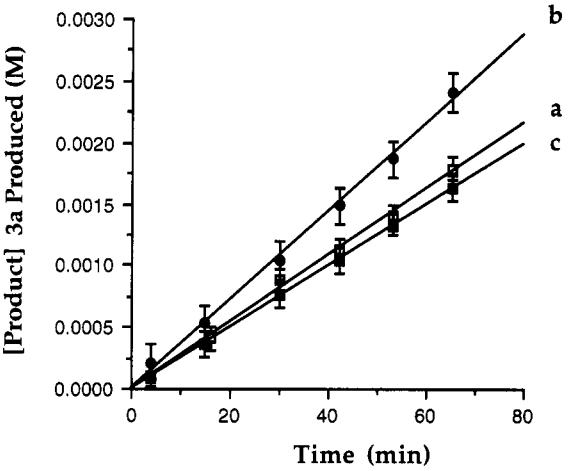


**Figure 5.** Reaction of 1 and 2 in $CHCl_3$ as followed by HPLC; 16.0 mM initial concentrations of 1 and 2, 1% TEA base added: (a) base-line reaction (1 + 2); (b) base-line reaction plus 0.2 equiv of product (**3a**); (c) base-line reaction plus 0.2 equiv of N-methylated product (**3b**).

Lest the reader cavil at the significance of template effects in these reactions, we dwell here at some length on the experimental evidence. To begin, the system shows autocatalysis. The initial rates of the reaction were studied at 16.0, 8.2, and 2.2 mM concentrations of starting materials 1 and 2 in $CHCl_3$ at ambient temperature. Under conditions of 8.2 mM, addition of 0.2 equiv of product produced a 43% increase in the reaction rate (Figure 4). With varied amounts of added product, the rate enhancements are not directly proportional to product concentration, but rather to its square root.[7] Thus, under the same conditions, addition of 0.5 equiv produced a 73% increase in the reaction rate. This "square root law" was described by von Kiedrowski[5,8] to characterize nucleic acid replicators in which the autocatalytic entity exists largely in dimeric form.

But what is it about this system that makes us call it self-replicating rather than another case of chemical autocatalysis? After all, **3a** bristles with functional

(8) von Kiedrowski, G.; Wlotzka, B.; Helbing, J.; Matzen, M.; Jordan, S. *Angew. Chem., Int. Ed. Engl.* **1991**, *30*, 423–426.
(9) Submitted for publication.

**Table 1.[9] Effects of Various Additives on the Reaction of 1 + 2 in $CHCl_3$ as Followed by HPLC[a]**

| entry | additive (0.5 equiv) | av initial rate of prod formatn ($\mu$M/min) ± 5% | % of base-line rate |
|---|---|---|---|
| 1 | | 0.54 | |
| 2 | **3a** | 0.81 | 150 |
| 3 | **5** | 0.55 | 102 |
| 4 | **6a** | 0.50 | 93 |
| 5 | **7** | 0.52 | 96 |
| 6 | **3b** | 0.55 | 102 |
| 7 | **9** | 0.56 | 104 |
| 8 | **6b** | 0.56 | 104 |
| 9 | **10** | 0.63 | 117 |

[a] 2.2 mM initial concentrations of 1 and **2**, 22 ± 1 °C, 1% TEA base added.

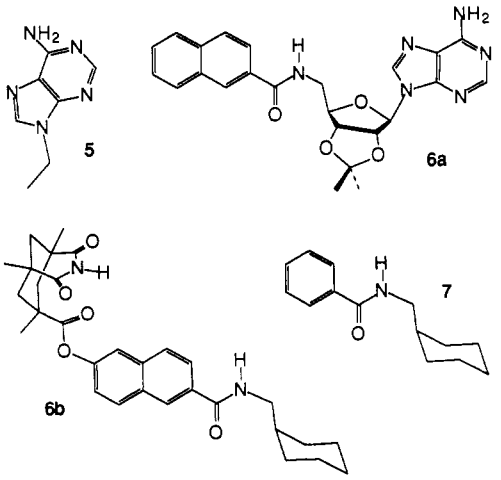


**Figure 6.**

groups. The imide, amide, ribose, and purine functionalities must all be considered as possible explanations for the autocatalysis observed. For example, imidazole is a well-known catalyst for acylation reactions, and the purine contains such a subunit. Could not this functionality be the cause? A claim of replication must be backed up. The potentially catalytic functions of the product molecule had to be individually tested, tested *in the structural context of* **3a**, and under the conditions where **3a** acts as an autocatalyst. We proceeded, through overlapping control experiments, to exclude each individual function of the product molecule as a source of simple chemical catalysis. The results are summarized in Table 1 (2.2 mM concentration) and Figures 4 and 5 (8.2 and 16.0 mM concentrations, respectively). At 2.2 mM, a 50% increase in initial rate was observed when the reaction was seeded with 0.5 equiv of product **3a** (Table 1, entry 2). Our control experiments, detailed below, show that this 50% increase is the result of replication, *i.e.*, a template effect as shown in Figure 3.

The question of purine catalysis was answered by the addition of 9-ethyladenine (**5**) and the naphthoylated ribosyl derivative **6a** (Figure 6) (Table 1, entries 3 and 4). The absence of catalysis in these experiments excluded the purine nucleus (and the ribose) as the sources of catalysis in entry 2. Additional control experiments with added *trans* secondary amides such as **7** (and **6a** and **6b**) showed that an external, secondary amide function, presented in the steric environment of **3a**, was unable to catalyze the reaction (Table 1, entries 4, 5, and 8). A most telling control experiment involved

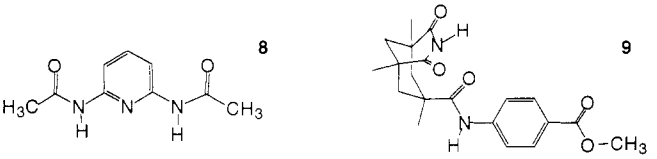


**Figure 7.**

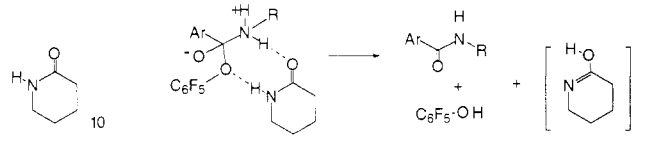


**Figure 8.**

the N-methylated imide **3b**. In the presence of **3b**, no rate enhancement in product formation was seen (within the 5% experimental error) at either high or low concentrations of reactants (Figure 5 and Table 1, entry 6). N-Methylation of the imide shuts down autocatalysis. The remarkable effect of this singular change[7]—substitution of a methyl group for the imide hydrogen—*excludes the purine, the ribose, the naphthoyl, and the secondary amide connecting them as the seat of the autocatalytic effect.*

These studies pointed to the imide as an essential ingredient in the replication of these systems, with base-pairing between imide and adenine as the likely mode of action. Support for this notion was obtained with the use of 8 (Figure 7). Bis(acylamino)pyridines such as 8 are much admired as hydrogen-bonding complements to imides,[10] and addition of bis(amide) 8 *inhibited* the replication reaction (Figure 4).[7] This result further implicates base-pairing between adenine and the imide as necessary for autocatalysis. But could there be a different, more active catalytic role for the imide functional group? Not according to further control experiments. Addition of structurally similar imides such as **9** and **6b** did not catalyze the coupling reaction (Table 1, entries 7 and 8). Furthermore, as neither **6a** nor **6b** is a catalyst under these conditions, one may conclude that the full template is necessary for catalysis: merely positioning one of the two substrates on the template backbone is not sufficient. The isolated, individual features and functionalities of **3a**, are, therefore, unable to account for the autocatalysis observed. Rather, the whole product molecule is more effective than the sum of its parts. The most economical explanation for these results is Figure 3: *template-catalyzed replication is the source of autocatalysis.*

To be sure, catalysts other than the template **3a** can be found for this system. The action of bifunctional catalysts in acylation reactions and glucose mutarotation is well-known,[11] and indeed, it was found that the *cis* amide valerolactam (**10**) increases the initial rate of formation of **3a** by 17% when added to the reaction of 1 and **2** (Table 1, entry 9). Figure 8 depicts a possible role for valerolactam in catalysis. The acidic and basic sites on valerolactam could interact with the tetrahedral intermediate to facilitate the required proton transfers for product release; subsequent tautomerization of the iminol would regenerate the catalyst.

(10) Hamilton, A. D.; Van Engen, D. *J. Am. Chem. Soc.* **1987**, *109*, 5035–5036.
(11) Rony, P. R. *J. Am. Chem. Soc.* **1969**, *91*, 6090–6096. Swain, C. G.; Brown, J. F., Jr. *J. Am. Chem. Soc.* **1952**, *74*, 2538–2543.

Some 20 years ago, Su and Watson[12] showed that small amides (even tertiary amides such as dimethyl acetamide) can hydrogen bond to the tetrahedral intermediates in related reactions and catalyze their breakdown to products. However, in light of the failure of secondary amides **6a**, **6b**, **7**, **9**, and **3b** to catalyze the reaction, the *trans* amide of **3a** cannot be a significant contributor to the autocatalysis observed under these conditions.

The most plausible explanation, then, for the autocatalytic nature of **3a** remains its ability to act as a template for its parts. The forces of hydrogen bonding and aryl stacking position the amine and active ester on the template's surface in a termolecular complex (**4**), and subsequent formation of the amide bond results in a template dimer. The weak intermolecular forces which stabilize the dimer also permit its dissociation (a dimerization constant of 630 $M^{-1}$ was measured for the product **3a** in $CDCl_3$), and monomeric template is generated. Thus, the template, through specific noncovalent contacts, has produced a copy of itself: the molecule replicates. The template effect is admittedly modest, and its magnification in more effective replicators (about which more later) is our constant goal. Understanding the details of the autocatalytic step is the likeliest path to this goal, but even without this knowledge, the mere experimental phenomenon of self-replication stands.

## Explorations

We had proposed that replication at the molecular level, as seen in these systems, was a primitive sign of life, into which mutation and even evolution might be incorporated to produce new molecules. Eyebrows were raised. We were, after all, dealing only with small rate enhancements, and were contemplating only small structural changes in our synthetic molecules. Rather than making attempts at definitions of life, mutation, or evolution, we shall simply recount several engaging aspects of our imide–adenine type replicators. The reader is left to judge the relevance of these model systems to prebiotic chemistry or biology.

As our experiments hinged on a comparison of rates of template catalysis, it was desirable to enhance the template effect which we saw in the reaction of **1** with **2**. To increase the rate of the termolecular, template-catalyzed step at the expense of the desultory but inevitable background reactions, a longer spacer was inserted into the replicator. By exchanging the 2,6-substituted naphthalene spacer of **1** for a 4,4′-biphenyldiyl spacer in **11**, the relative rate of product formation through the termolecular replicative complex was increased.[7] Condensation of **11** with **2** gave **14** (Figure 9), and following the time course of the reaction revealed the gentle sigmoidal product growth curve expected of an efficient self-replicating system,[13] a feature which had already been observed for nucleic acid replicators.[8]

Next, we modified the adenine component. The exocyclic amine of the purine was outfitted with urethane-type blocking groups. The benzyloxycarbonyl was attached in **12**, and the *o*-nitrobenzyloxycarbonyl

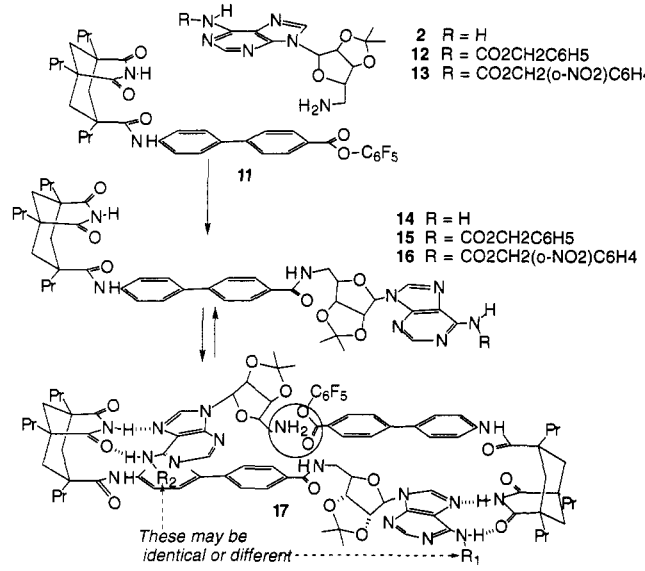


**Figure 9.**

was attached in **13**. These changes were known to reduce base-pairing possibilities; the blocking groups protrude on the Watson–Crick hydrogen-bonding edge of the purine (Figure 9) and force molecules which bear them to base-pair on the Hoogsteen edge.[14] Both **12** and **13** coupled to the biphenyl **11** in the familiar manner, giving the respective self-complementary products (**15** and **16**). They were each replicators; hobbled as they were in their base-pairing capacities, they could not be terribly efficient, but they did catalyze their own formation. They also made mistakes; as shown in the generic **17**, there is no means by which either template can distinguish between those molecules bearing the benzyloxycarbonyl group and those bearing the *o*-nitrobenzyloxycarbonyl group. The behavior of these molecules was reciprocal; one catalyzed the formation of the other and vice versa.[15]

These molecules were, of course, *contrived* to make a point. We intended to show that a change in environment could alter the system such that a third replicator was produced which was more efficient than either **15** or **16**. Our change in environment was the introduction of light to the system; when the molecules bearing the *o*-nitrobenzyloxycarbonyl group are irradiated, the photolabile blocking group is cleaved.[16]

In the experiment, amines **12** and **13** were allowed to compete for a limited amount of active ester **11**. The result (Figure 10) was that the nitro derivative was a slightly more effective replicator; more **16** was formed than **15**. The solution was then irradiated, causing the photolabile nitro groups to be removed (the photocleavage product of **13** is **2**, and that of **16** is **14**). The new replicator **14** generated in this manner had a profile which permitted Watson–Crick as well as Hoogsteen base-pairing; it had a statistical advantage over either **15** or **16**. Now, when more active ester was added to the medium, the new molecule was more effective at replication, and it took over the resources of the system. The permanent change in structure caused by light

(12) Su, C.; Watson, J. W. *J. Am. Chem. Soc.* **1974**, *96*, 1854–1857.
(13) Rotello, V.; Hong, J. I.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 9422–9423.
(14) Dodin, G.; Dreyfus, M.; Dubois, J.-E. *J. Chem. Soc., Perkin Trans.* **1979**, 438–442.
(15) Hong, J.-I.; Feng, Q.; Rotello, V.; Rebek, J., Jr. *Science* **1992**, *255*, 848–850.
(16) Patchornik, A.; Amit, B.; Woodward, R. B. *J. Am. Chem. Soc.* **1970**, *92*, 6333–6335.

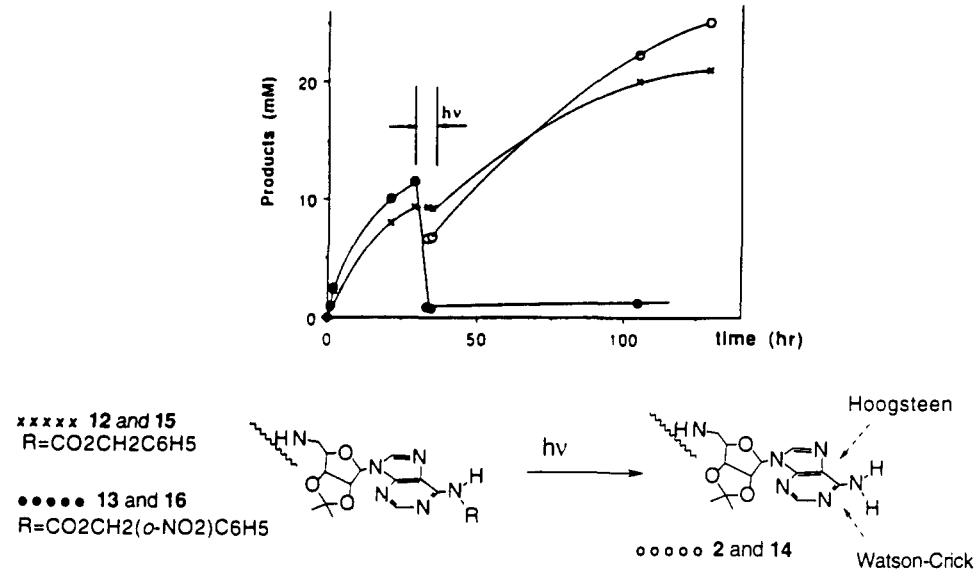


**Figure 10.**

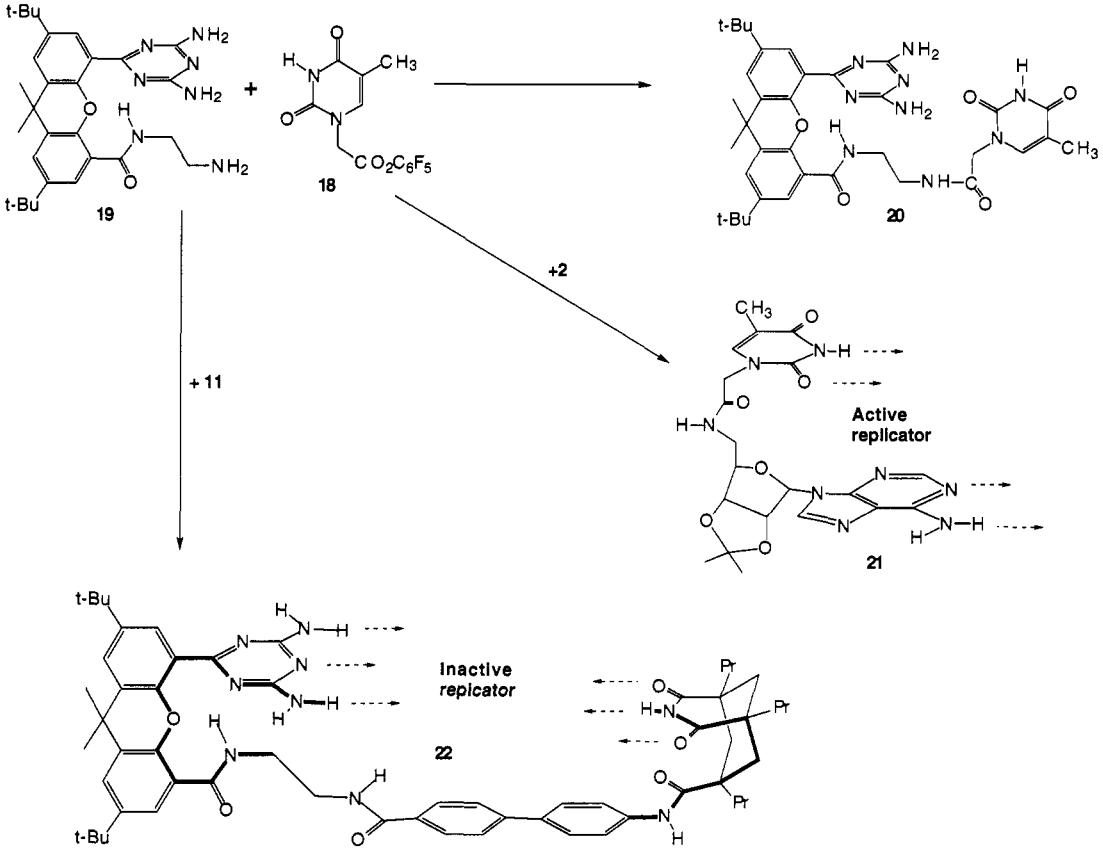


**Figure 11.**

created a self-replicating molecule more fit to the environment of the experiment.

Inspired by the rapidly emerging model systems of artificial life,[17] we were confronted with the question, could synthetic replicators shuffle their components to generate new, hybrid replicators? To answer this, we developed a second self-replicating system (Figure 11), based on the molecular recognition of thymine derivative **18** by diaminotriazine **19**. Again, coupling of the components through an amide bond gave a self-complementary structure **20**. The new structure showed autocatalytic behavior,[18] just as the adenine-based systems described above.

The selection of functional groups on the components was deliberate and was dictated by the premise that self-complementarity of structure was sufficient for replication. After all, had not all our replicators (and those of others[19]) shared this feature? All possible combinations of **2**, **11**, **18**, and **19** were duly synthesized (Figure 10), and their behavior taught us a lesson

(17) Levy, S. *Artificial Life: The Quest for a New Creation*, 1st ed.; Pantheon: New York, 1992; p 390.

(18) Park, T. K.; Feng, Q.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1992**, *114*, 4529–4532.

(19) Terfort, A.; von Kiedrowski, G. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 654–656. See, however: Persico, F.; Wuest, J. D. *J. Org. Chem.* **1993**, *58*, 95–99.

concerning molecular shape. One of the shuffled replicators, the adenine–thymine product **21**, closely resembles a modern nucleic acid, but with an amide backbone. It turned out to be the most effective synthetic replicator we have encountered to date (perhaps this is not mere coincidence). The other shuffled replicator **22** was unable to catalyze its own formation.[20]

The reasons for the differing behaviors of **21** and **22** are implicit in their molecular shapes. The adenine–thymine product can present its recognition surfaces (arrows on structure **21**) in such a way that a productive termolecular complex can be assembled with its precursors. No such conformation is available to the hapless **22**. Because it is made up of two U-shaped components (xanthene and Kemp triacid skeletons), the overall conformation of **22** is either a C-shape (shown) or an S-shape (not shown). In the C-shaped conformation, recognition surfaces converge (arrows on structure **22**), and the component parts cannot fit within the cleft. In the S-shaped conformation, a termolecular complex can form, but it is not productive; the recognition surfaces diverge, and reactive centers are too far apart from each other to form a covalent bond. Accordingly, self-complementarity is insufficient for replication; it is necessary to find a cyclic, dimeric conformation.

## Outlook

Our program for the next generation of synthetic replicators is faced with several challenges. The first deals with enhancing the replication event by chemical catalysis in addition to template effects. Can replicators be devised which feature catalytically useful functionality within their structures as do ribozymes?[21] Such functions might act as general acids, general bases, nucleophiles, or even bifunctional catalysts. We have identified some likely candidates, based on the well-known action of bifunctional catalysts in acylation reactions and glucose mutarotation.[11] Other functionalities that can present hydrogen-bonding sites to the tetrahedral intermediates involved in the reaction should also be catalysts (provided they are compatible with the reaction conditions). The success of these plans will depend on how precisely functional groups can be trained on the tetrahedral intermediates. New carbazole-based complements to adenines[22] show promise in this regard, and we are elaborating their structures accordingly.

The second challenge involves information. Can synthetic systems be devised which provide sequential information? Catalysis, of the sort discussed immediately above, is a very simple form of information: the structure *is* the message, and the message is faithfully transmitted from copy to copy. However, with the present structures, severe synthetic difficulties, perhaps insurmountable, would be expected if attempts to serialize them—string them together to increase their informational content—were to be made. We are therefore looking at other structures[23] which are more amenable to such modifications and will report progress as it is made.

*We are grateful to the National Science Foundation for financial support. E.A.W. thanks the NSF and M.M.C. the NSERC for predoctoral fellowships. In addition, we thank a number of former co-workers for their contributions to the work described in this Account: postdoctoral fellows P. Ballester, M. Famulok, J.-I. Hong, J. Nowick, and V. Rotello and graduate students T. K. Park, Q. Feng, and T. Tjivikua, in whose theses much of this work is detailed. We also thank Gerald Joyce and JoAnne Stubbe for their comments on the manuscript.*

(20) Feng, Q.; Park, T. K.; Rebek, J., Jr. *Science* **1992**, *256*, 1179–1180.
(21) For a recent review, see: Pyle, A. M. *Science* **1993**, *261*, 709–714.
(22) Conn, M. M.; Deslongchamps, G.; de Mendoza, J.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1993**, *115*, 3548–3557.
(23) Andreu, C.; Beerli, R.; Branda, N.; Conn, M. M.; de Mendoza, J.; Galán, A.; Huc, I.; Kato, Y.; Tymoschenko, M.; Valdez, C.; Wintner, E.; Wyler, R.; Rebek, J., Jr. *Pure Appl. Chem.* **1993**, *65*, 2313–2318.