Pure & Appl. Chem., Vol. 65, No. 11, pp. 2313–2318, 1993.
Printed in Great Britain.

Replication and assembly

Cecilia Andreu, René Beerli, Neil Branda, Morgan Conn, Javier de Mendoza*, Amalia Galán, Ivan Huc, Yoko Kato, Maria Tymoschenko, Carlos Valdez, Edward Wintner, René Wyler and Julius Rebek, Jr.

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139
*Universidad Antonoma de Madrid, Cantoblanco 28049 Madrid, Spain

Abstract - Molecules bearing surfaces capable of recognition phenomena can be covalently attached to each other to give self-complementary structures. Such molecules can reveal autocatalytic effects as replicators, and they are able to assemble in a variety of superstructures. Closed-shell assemblies, capable of inclusion phenomena are particularly desirable targets for synthetic and nucleation studies.

INTRODUCTION

Four years ago, in my first article in this Journal, I described model studies in molecular recognition using new molecular shapes[1]. At that time, the notion of self-complementarity and its significance for molecules capable of self-replication was introduced. In the current article I intend to expand this idea and to trace a connection between *how* molecules fit together and *why* molecules fit together. Before beginning, some cautious words on vocabulary are in order. A difficulty frequently arises with the shared words of biology and chemistry; words like regulation, transport, recognition and replication. They have well-defined meanings in biology, but contain no structural information for model builders like myself in bioorganic chemistry. The challenge is to devise structures that express the desired molecular behaviors, if molecules can be said to behave. The larger problem is to reconcile the actual activities of molecules with the limited number of words we have to describe them.

REPLICATORS

The starting point for the research group was the molecular recognition of nucleic acid components, and because adenine has received its fair share of literature, we use here thymine as the example. Macrocyclic receptors for this heterocyclic nucleus were originally developed by Hamilton[2], and our own constructs involve the derivatives of a xanthene diacid. The diacid is now commercially available, and we were able to elaborate aromatic stacking surfaces and the diaminotriazene function within a single molecule[3]. One of the motivations was to generate realistic models for photolyases, and these have now been developed by Rose[4]. We turned instead to constructs in which the receptor for thymine was *covalently attached* to thymine. That is, a self-complementary molecule was synthesized.

The actual reaction involved the amine nucleophile and an active ester of the thymine to give the structure shown in Fig. 1. This reaction indeed featured autocatalysis; addition of the product resulted in increased rates of coupling between the components[5]. We were able to relate this behavior to the formation of a productive termolecular complex. The product could collect on its surface the two components from which it was made and position them for a relatively easy reach to the transition state for the reaction, thence product. The dimeric product can dissociate and give rise to the increasing concentrations of catalyst that characterize autocatalytic phenomena. These systems, as well as our earlier adenine-based replicators follow the "square root law" devised by von Kiedrowski[6] for self-replicating molecules based on nucleic acids.

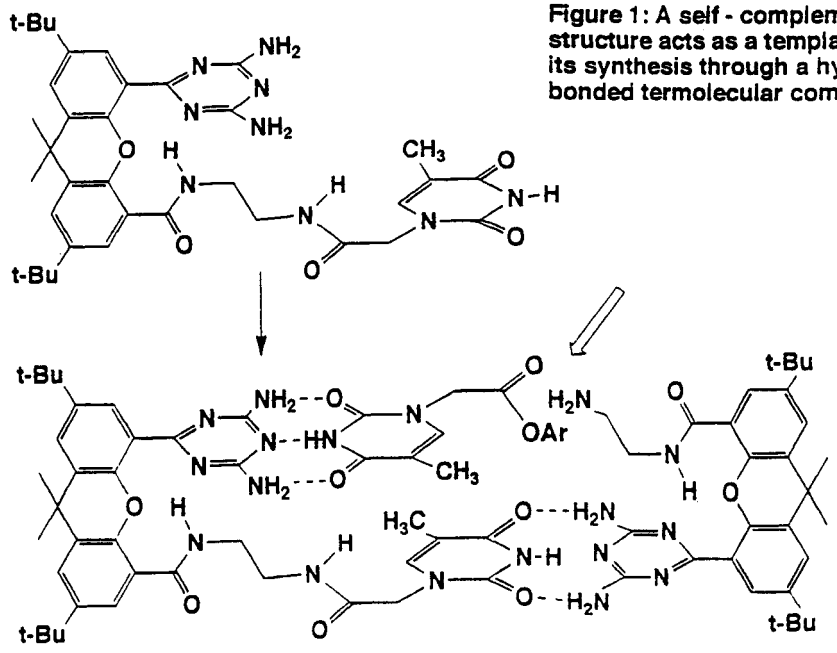

Figure 1: A self - complementary structure acts as a template for its synthesis through a hydrogen-bonded termolecular complex.

With two structurally different self-replicating molecules in hand, it was possible to take advantage of their common feature, the formation of an amide bond, to generate new hybrids. Specifically, a recombination experiment could be devised in which a receptor for adenine could be coupled covalently to a receptor for thymine[7]. The other recombinant involved the covalent attachment of the adenine nucleophile to the thymine electrophile (Fig. 2). This is a chemical version of the genetic algorithm that had been so successfully used in computer programming. These chemical experiments along with those of others[8] became the "wetware" counterparts for the software and hardware constructs at the Artificial Life Conference in Santa Fe last summer[9].

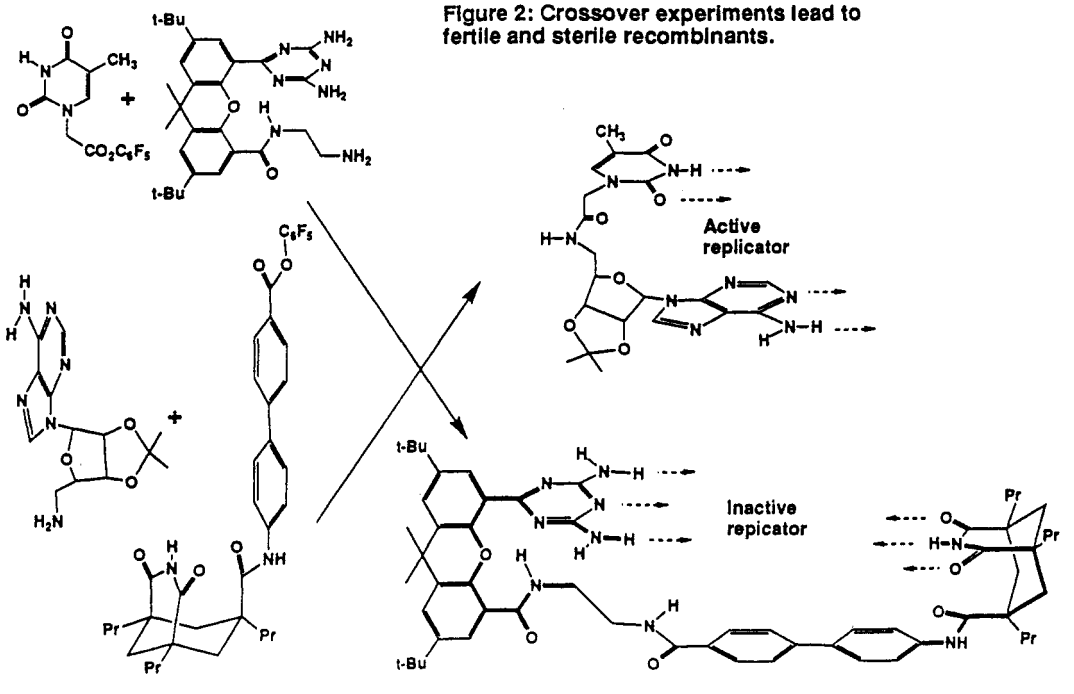

Figure 2: Crossover experiments lead to fertile and sterile recombinants.

At first glance, the two self-complementary hybrids might be expected to be replicators themselves. The gross molecular shape becomes an overriding consideration and limits their behavior. The adenine-thymine recombinant which now resembles a nucleic acid with a peptide backbone, can assume a conformation in which replication is possible. The recognition surfaces are oriented in such a way that a productive termolecular complex can be formed. The product, the cyclic hydrogen bonded dimer, features a *cavity*, and we will pursue the uses of cavity- forming molecules further below.

For the present, consider the shape of the other hybrid. It is composed of two U-shaped modules , the xanthene diacid and the Kemp triacid. Both of these modules feature bonding arrays that fold back on themselves, and the overall shape of the recombinant must be C-shaped, (as shown) in which case there is not enough room in the interior to form a productive catalytic complex, or S-shaped, (not shown) in which a termolecular complex can be formed, but it is not productive. The amine nucleophile and the active ester electrophile are in divergent conformations. This molecule is not a replicator and it has to express its self-complementarity in polymeric forms, in the same way as do molecules of interest in materials science.

Self-complementarity appears then to be a necessary, but not sufficient structural feature of replicators. The structural landscape for replicators has a fitness built into it, and the recombination result is instructive from an evolutionary point of view. A relatively small pool of molecules can give rise to a family tree of replicators; one branch of which is an evolutionary dead-end. It would be an appropriate fate for the the unsuccessful molecule to be metabolized and have its pieces recycled by incorporation into the successful replicators.

All of our synthetic replicators take advantage of template effects, that is, the product molecules act to reduce the entropy of activation for the coupling reaction. Otherwise they are quite passive. Can some more active form of catalysis be built into these? In the language of biology can the genotype be simultaneously a useful phenotype? We are making some progress on this goal. The idea is to position catalytically useful functions along the mechanistic pathway of replication; to place acids or bases on the molecular skeleton where they are likely to enhance the replication step. The current construct being pursued is a bipyridyl derivative (in place of the biphenyl spacer) in which the nitrogens could act as general bases for the coupling of species gathered on their surface; an intramolecular general-base catalytic step.

Another shortcoming, constantly and gleefully trotted out by our critics, is that our replicators are not informational. While the structures themselves carry information, there is a question of how these might be strung together in a longer sequence. We haved made some desultory attempts at lengthier replicators using peptides with self-complementary structures, but none of these showed autocatalytic behavior. The lack of conformational control in such molecules is the likely culprit, and we have turned to conformational restraints in peptides to overcome this problem. The β-strand structure, in which α-α dialkyl amino acids enforce a linear, rigid conformation[10], suggest that molecules as shown in Fig. 3 may have reasonable success as informational replicators. The peptide backbone fixes the conformation, and hydrogen bonding from a chiral hybrid molecules, part amino acid and part nucleic acid, can provide the base-pairing information along the sequence. Whether these types of molecules have the added charm of prebiotic significance is debatable, but a synthesis of these molecules from relatively well established prebiotic fragments[11] (formaldehyde, adenine and cyanide) is conceivable.

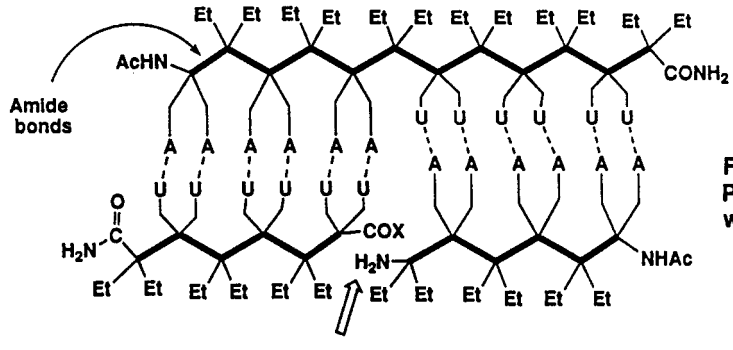

Fig. 3
Proposed informational replicator with β- strand peptide backbone

ASSEMBLIES

The self-assembly of synthetic structures has become such a popular undertaking among practitioners of molecular recognition that the nonspecialist might feel intimidated by the sheer volume of activity. So many systems have been proposed that new constructs are hard to imagine. Ribbons, tapes, sheets, helices and interlocking rings are all well-represented, and the reigning paradigm appears merely to specify two complementary components which can selectively associate [12-15]. Examples include melamine-cyanuric acid and other hydrogen-bonded pairs; metal-bipyridyl and aryl-aryl charge transfer interactions. While some spectacular three-dimensional assemblies have been devised[16, 17] generally the systems put forth have been two-dimensional. Moreover, with few exceptions[18-20], self-complementary (single component) molecules are underrepresented.

Our own constructs (as are those of many others) are inspired by natural assemblies: membranes, duplex nucleic acids and viral coat proteins. The appearance of identical subunits in many of these biological examples, suggest that self-complementarity is both useful and advantageous. Allosteric enzymes and viral coat proteins all reveal an *economy* of information; the instructions for assembly are written within the shapes of the subunits and the orientation of their recognition surfaces. The binding information that is contained within the structure becomes expressed during the assembly.

The G-quartet (Fig. 4) known for more than 30 years is taken as an example. The hydrogen bonding donor and acceptor sites on the Watson-Crick edge of guanine find their complements on the Hoogsteen edge [21]. The orientation of these edges with respect to one another is approximately 90°. In the presence of nucleating ions such as potassium[22] the assembly takes place in two-dimensions as a tetramer. *It cannot do otherwise.*

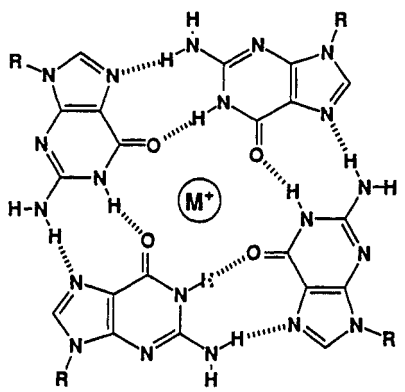

Fig 4: Four guanines assemble around a nucleating metal ion

Our own entry into this enterprise came with the realization that self-complementarity, the key feature of self-replicating molecules is also ideal for assemblies. After all, the best of the self-replicating molecules we encountered involved a dimeric structure that featured a cavity, so we have pursued the ideas of closed-shelled surface assemblies from small, self-complementary pieces.

A tennis ball provides a simple macroscopic analogy for the notional and structural features. Cut along its seam the ball gives two identical pieces. In shape, the ends are complementary to the middle and the subunits feature curvatures that dictate the overall spherical shape of the dimer. A similar spacing and appropriate curvatures in molecular subunit structures are required if their assembly is to present an overall cavity. In collaboration with the group in Madrid, we have reduced this to ball-shaped target molecules.

The simplest structure that presents the desired features at the molecular level is shown in Fig. 5. The lactam functions of the structure provide self-complementary hydrogen bond donors and acceptors indicated by the arrows. The concavity of the molecule along its length is a consequence of the folding caused by each of the two 7-membered rings when all of the aryl subunits are on the same face of the structure. The concavity along its width is caused by the *cis* fusion of the 5-membered rings. The compound was prepared by condensing two molecules of glycouril[23] with durene tetrabromide.

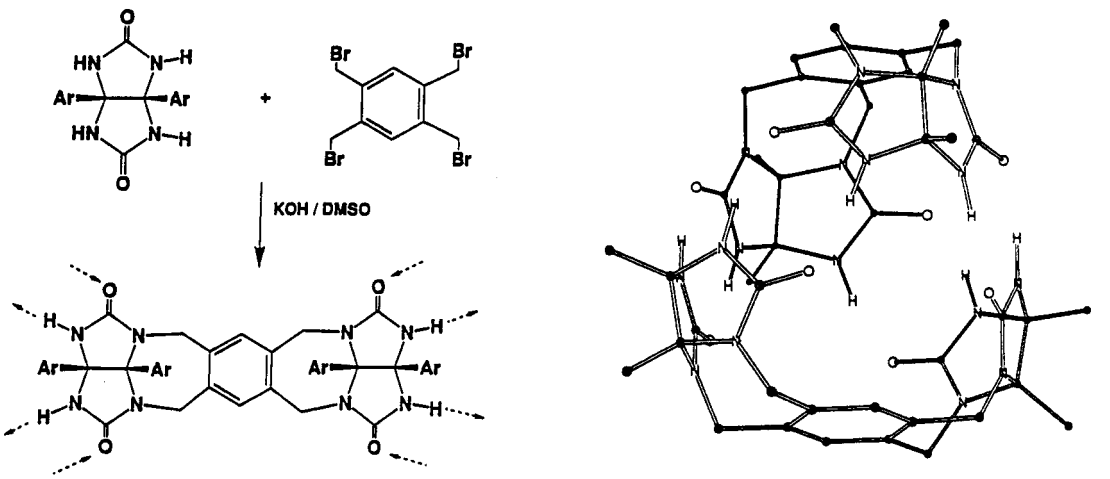

Fig 5 Synthesis of a hemisphere and its dimerization.

Dimerization of this self-complementary structure can form a closed-shell, three-dimensional surface through a network of hydrogen bonds. The dimeric nature of this system was established by physical measurements including vapor pressure osmometry, mass spectrometry, and NMR. A crystalline sample has been prepared, and x-ray crystallographic studies are underway. The larger version of this structure, a softball, has also been devised.

Another module that features self-complementary and curvature and permits a number of closed-shelled assemblies is an unusual trilactam. Four of these modules fit comfortably at the corners of a tetrahedron, and an even better fit is observed for an octamer in which the subunits are at the corners of a cube. One face of such a cube is shown in the partial structure (Fig. 6). An energy minimized structure of the cube shows that hydrogen bond angles are greater than 170°, and the cavity is spacious enough to accommodate molecules such as tetramethyl adamantane. Of course, many other oligomers might be formed as well as polymeric, extensive three-dimensional arrays. However, a particularly beautiful closed shell assembly involves twenty subunits at the corners of a dodecahedron. The hydrogen bonds of this assembly are perfectly rectilinear and an enormous cavity results.

Another molecule capable of assembly into a cube is the tetraamide shown in Fig. 7. In this, each subunit represents a face of the cube. A modification of this theme is given by the C_3 structure also shown, which might also assemble into closed-shell surfaces.

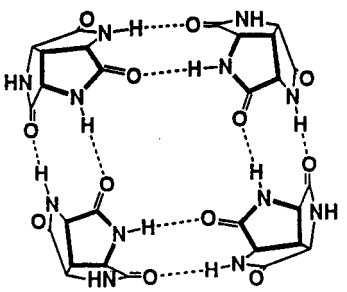

Fig. 6 One face of a cube assembled from a convex trilactam.

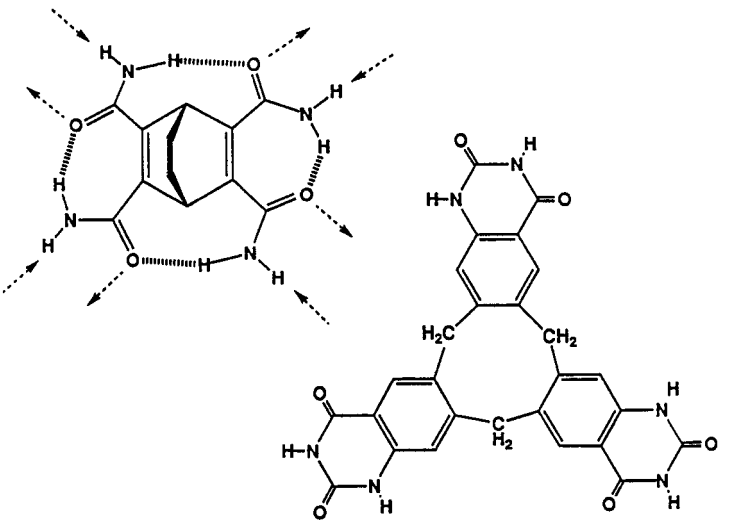

Fig. 7 Modules capable of assembly into closed-shell surfaces

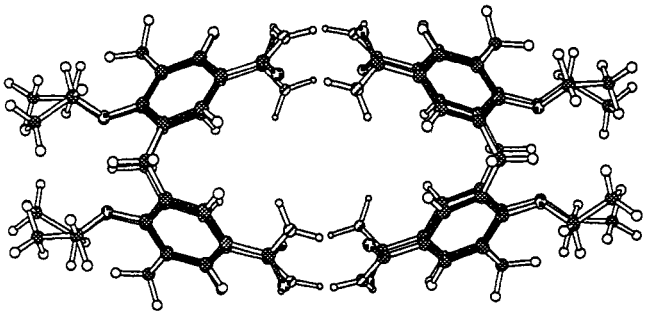

Figure 8. Proposed dimeric structure of Calix[4]arene tetraamide

Finally, we even consider a self-complementary macrocycle for such assemblies. A system currently under investigation involves the calixarene tetraamide of Fig. 8 This shows a dimerization constant in the millimolar range in organic solvent mixtures, and modeling shows that the cavity of the dimer could easily accommodate molecules such as 2-butyne. We will report on the synthesis and characterization of these self-complementary structures in due course.

Acknowledgements

We are grateful to our talented coworkers whose efforts in molecular replication may be found in the original literature citations. and to our generous collaborators. The National Science Foundation, the National Institutes of Health, and Hoechst-Celanese provided financial support for this research.

References

1. J. Rebek, Jr., Pure and Appl. Chem., 61, 1517-1522 (1989).
2. A. D. Hamilton and D. Van Engen, J. Am. Chem. Soc., 109, 5035-5036 (1987).
3. T. K. Park, J. Schroeder and J. Rebek, Jr., J. Am. Chem. Soc., 113, 5125-5127 (1991).
4. M. S. Goodman and S. D. Rose, J. Am. Chem. Soc., 113, 9380-9382 (1991); S. C. Hirst and A. D. Hamilton, Tetrahedron Lett, 31, 2401-2404 (1990).
5. T. K. Park, Q. Feng and J. Rebek, Jr., J. Am. Chem. Soc., 114, 4529-4532 (1992).
6. G. von Kiedrowski, B. Wlotzka, J. Helbing, M. Matzen and S. Jordan, Angew. Chem.,Int. Ed. Engl., 30, 423 (1991).
7. Q. Feng, T. K. Park and J. Rebek, Jr., Science, 256, 1179-1180.
8. P. A. Bachmann, P. Walde, P. L. Luisi and J. Lang, J. Am. Chem. Soc., 113, 8204 (1991).
9. See for example, Artificial Life II, C. G. Langton, Ed., Addison-Wesley, Reading, MA (1992).
10. C. Toniolo and E. Benedetti, Macromolecules, 24, 14, 4004-4009 (1991).
11. A. Eschenmoser and E. Lowenthal, Chem. Soc. Rev., 1, (1992).
12. G. M. Whitesides, J. P. Mathias and C. T. Seto, Science, 254, 1312-1319 (1991); J. A. Berkowski, C. T. Seto, D. A. Wierda and G. M. Whitesides, J. Am. Chem. Soc., 112, 9025-9026 (1990).
13. C. Fouquey, J.-M. Lehn and A. M. Levelut, Adv. Mater, 2, 254 (1990); J.-M. Lehn, A. Rigault, J. Siegel, J. Harrowfield, B. Chevrier and D. Moras, Proc. Nat. Acad. Sci. USA, 84, 2565-2569 (1987); U. Koert, M. M. Harding and J. -M. Lehn, Nature, 346, 339 (1990).
14. F. Garcia-Tollado, S. J. Geib, S. Goswami and A. D. Hamilton, J. Am. Chem. Soc., 113, 9265-9269 (1991).
15. P. L. Anelli, et. al., J. Am. Chem. Soc., 114, 193-218 (1992).; J. L. Sessler, D. Magda and H. Furuta, J. Org. Chem, 57, 818-826 (1992).
16. C. T. Seto and G. M. Whitesides, J. Am. Chem. Soc., 113, 712 (1991); P. Baxter, J. -M. Lehn, A. DeCian and J. Fischer, Angew. Chem. Int. Ed. Engl., 32, 69-72 (1993).
17. M. Simard, D. Su and J. D. Wuest, J. Am. Chem. Soc., 113, 4694-4698 (1991).
18. R. P. Bonar-Law and J. K. M. Sanders, Tetrahedron Lett., 34, 1677-1680 (1993).
19. Y. Durcharme and J. D. Wuest, J. Org. Chem., 53, 5787-5789 (1988).
20. S. C. Zimmerman and B. F. Duerr, J. Org. Chem., 57, 2215-2217 (1992).
21. M. Gellert, M. N. Lesett and D. R. Davies, Proc. Nat. Acad. Sci. USA, 48, 2013 (1962).
22. R. G. Barr and T. J. Pinnavaia, J. Chem. Phys., 90, 328-334 (1986).
23. J. W. Smeets , P. P. Sijbesma, L. van Dalen, A. L. Speks, W. J. J. Smeets and R. J. M. Nolte, J. Org. Chem., 54, 3710-3717 (1989).
24. R. Wyler, unpublished.