Supramolecular Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713649759

Recognition, replication and extrabiotic chemistry

Julius Rebek Jr.[a]
[a] Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA

To cite this Article Rebek Jr., Julius(1993) 'Recognition, replication and extrabiotic chemistry', Supramolecular Chemistry, 1: 3, 261 — 266
To link to this Article: DOI: 10.1080/10610279308035169
URL: http://dx.doi.org/10.1080/10610279308035169

SUPRAMOLECULAR CHEMISTRY, Vol. 1, pp. 261–266
Reprints available directly from the publisher
Photocopying permitted by license only

Recognition, replication and extrabiotic chemistry

JULIUS REBEK, JR.

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

(Received July 28 1992)

INTRODUCTION

The subject is how molecules fit together – molecular recognition – and the intriguing specific issue is *why* molecules fit together. Complementarity and self-complementarity are central to these problems, and by complementarity is meant that of molecular size, shape, and charge that gives rise to the reciprocal, weak intermolecular forces that bind host and guest, however temporarily, to one another. One of the unexpected dividends from the study of forces is concerned with 'extrabiotic'[1] chemistry. This term is intended to define chemical or other systems which show properties reminiscent[2] of living systems, yet have little or no structural relationship to what is regarded as biological. In our own research hydrogen bonding and aryl stacking interactions have provided the intermolecular forces of recognition. Self-replicating systems appear to be an inevitable consequence of such forces.[3]

One rather random way of generating self-replicating molecules is through reciprocal templating effects. For example, consider a scheme in which a concave surface, lined with a variety of chemical functional groups for binding and catalysis, acts as a mould for the assembly of a complementary structure of a convex nature (Fig 1). When the complex is fully formed and it dissociates, these roles can be reversed; the convex surface acts as a template to force together the components needed to make the concave structure. the result of the two catalytic events is a replication cycle (or more exactly, a bicycle). The replication of nucleic acids is an example of such a cycle, and it can operate with the aid of enzymes, ribozymes[4] or even without added catalysts.[5]

Suppose now a covalent bond is accidently formed between the two complementary molecules of the cycle and they become linked to one another at some peripheral site that keeps their respective recognition surfaces mutually accessible. Such a unit is a minimalist replicator, it is 'self-complementary', and can make identical copies of itself by acting as a template (Fig 2). Of course, the tether that connects them must enjoy some flexibility and most importantly, the orientation of the recognition surfaces with respect to one another must permit the formation of a productive termolecular collection as shown. Such opportunities presented themselves in a desultory way during our studies of molecular recognition.

We came upon minimalist replicators through an incremental approach. By knowing in detail how two molecules fit together in a complex, it was possible to arrange reactions *within* the complex that were appropriate for replication. We made use of the same

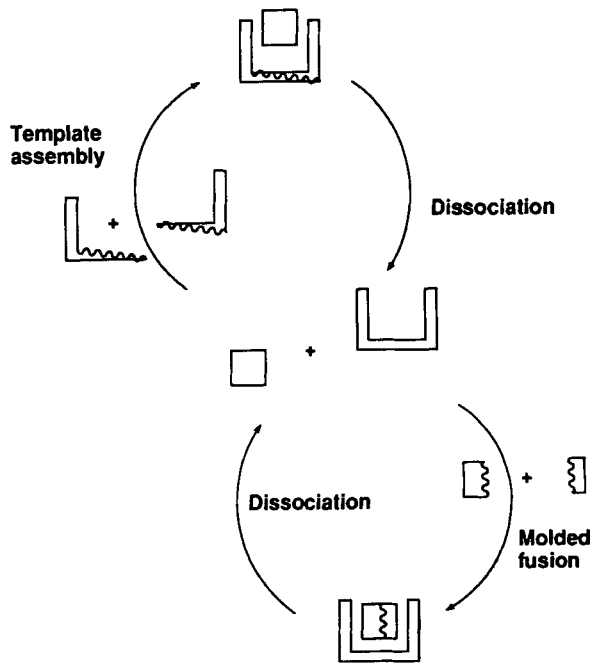

Figure 1 Recognition and a replication bicycle.

* To whom correspondence should be addressed.

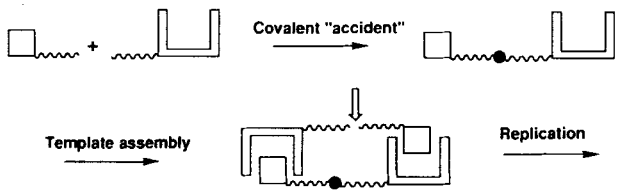

Figure 2 Self-complementary minimalist replicators.

intermolecular forces that stabilize double-stranded nucleic acids: hydrogen bonds and aryl stacking interactions, but our departure was in the choice of solvent. In contrast to other studies in water, we used organic solvents in which these forces are magnified and intermolecular associations are considerably enhanced. Reaction times are accordingly contracted to the human (rather than evolutionary) timescales. These synthetic molecules and solvents are surely not those of prebiotic or early biotic earth, and we make no such claims, yet the properties they show – replication, reciprocity and mutation – must also have been features of those molecules that were.

STRUCTURE AND RECOGNITION

We prepared the minimalist replicators through the covalent coupling of adenosine derivatives to imides of a rather peculiar shape. These molecules, which involve structures that fold back on themselves, resemble the hydrogen bonding surface of thymine. They are readily accessible as derivatives of Kemp's triacid,* [6] and we have explored a number of these, both with rigid and flexible attachments. One of the trends that emerged is shown in Figure 3. The increase in aromatic surface size leads to enhanced stacking interactions. These are not conventional hydrophobic effects, since the binding studies are performed in organic solvents such as $CHCl_3$. Rather, the forces represent the response of large, polarizable aromatic surfaces to the nearby dipole of adenine, pinned there by its base pairing to the imide. In the series shown, the monotonic increase in affinity (0.4 kcal/mol) on moving from the phenyl to naphthyl to anthracyl series provides a direct measure of the relative polarizability of these aromatic surfaces.[7] Other synthetic receptors developed in a number of laboratories have now provided a comprehensive picture of intermolecular forces involved in the recognition of nucleic acid components.[8]

The hydrogen bonding surfaces could be tailored for complementarity to cytosine and guanosine

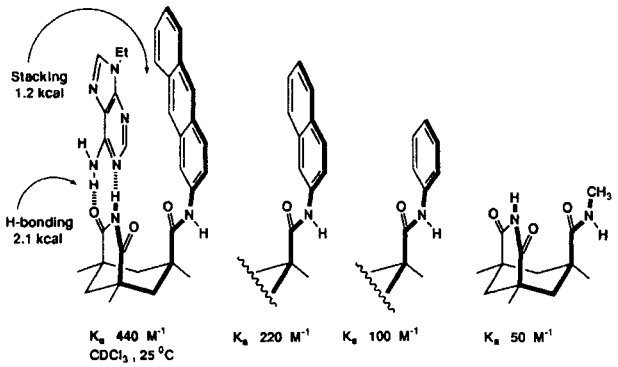

Figure 3 Complexation of adenines in $CDCl_3$.

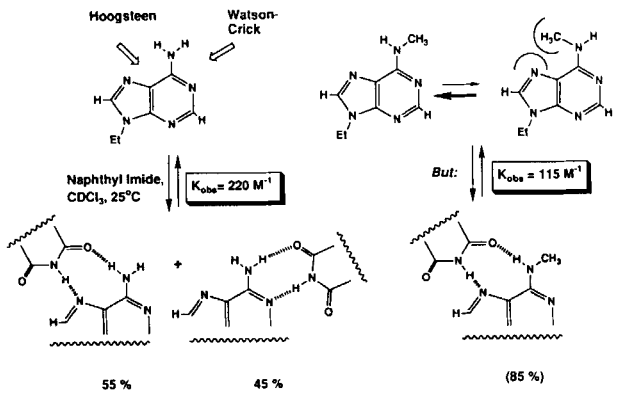

Figure 4 Base pairing in substituted adenine derivatives.

derivatives,[9] and even to thymine derivatives.[10] For this last case, a xanthene skeleton was devised with the same U-turn as provided by the Kemp triacid, but giving rise to much larger cavities.* [11] As shall be described, the U-turn feature results in a peculiar consequence for self-replicating molecules.

For the adenine derivatives we were able to show how Watson–Crick vs Hoogsteen base pairing could be controlled through the use of peripheral bulky groups that provided steric effects,[12] and such steric effects operate on the purine skeleton as well. For example, any *N*-substituents prefer to be directed away from the 5-membered ring and expose the adenine skeleton to almost exclusive Hoogsteen hydrogen bonding[13] (Fig 4). Thus, *N*-methyl derivatives are at a statistical disadvantage for base pairing to typical imides, and intermolecular NOE experiments could establish to what extent the numerical disadvantage operates.[14]

By placing two imide functions on a suitable aromatic 'spacer' we were able to generate molecules with extraordinary affinity for adenines – so much so

* Commercially available from Aldrich Chemical Company.

* The xanthene diacid is commercially available from Aldrich Chemical Company.

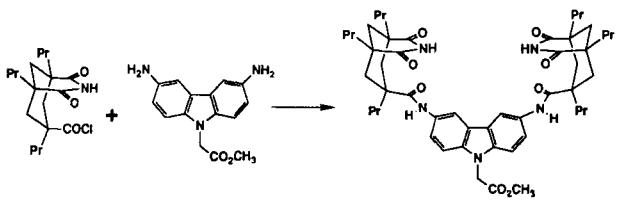

Figure 5 A chelating agent for adenines.

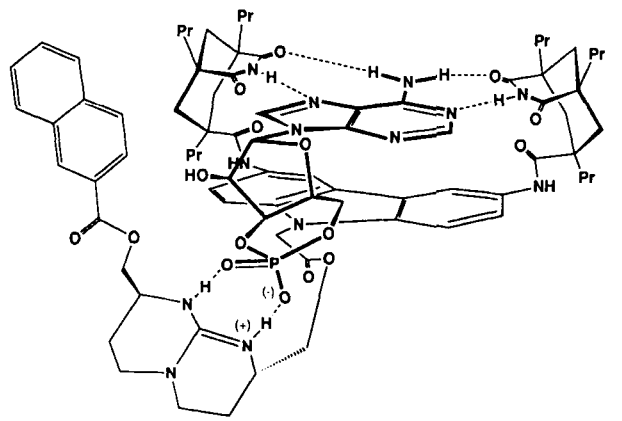

Figure 6 Complex of a synthetic receptor with cyclic AMP.

that they operated as vehicles for selective adenine transport across simple liquid membranes.[15] To date, our most refined synthetic adenine receptor is due to a collaborative effort with Javier de Mendoza in Madrid.[16] Using carbazole as a spacer, we were able to match, and provide for the gentle asymmetry required of adenines chelated by simultaneous Watson–Crick and Hoogsteen base pairing (Fig 5).

Moreover, the carbazole nitrogen provided a chemical handle by which additional recognition elements could be placed onto the skeleton. Specifically, an ionic recognition site, the guanidinium function, developed by de Mendoza,[17] Lehn,[18] and Schmidtchen,[19] could be easily incorporated. These molecules showed sufficient affinity for phosphorylated derivatives that stoichiometric amounts of cyclic AMP could be extracted from aqueous solution into organic solvents[16] (Fig 6). Further elaboration has even permitted the quantitative extraction of dinucleotide derivatives.[20]

REPLICATION

As mentioned above, the replicators arose through the formation of a covalent bond between adenine[21] and its receptor. The self-complementary unit was obtained by the reaction of an amino adenosine with an activated ester attached to the imide surface.[22] The product is in rapid equilibrium with a dimeric form that is held together by the familiar forces of adenine–thymine base pairing and stacking interactions. Once it has dissociated, a monomeric form of the product acts as a template; it can gather on its surface the two components from which it is assembled in a productive termolecular complex (Fig 7). The reactive centres in it are poised for relatively effortless *trans* amide bond formation. Intramolecular reaction follows and the product is an exact replica of the template. Dissociation of the dimeric product then generates the ever increasing number of template molecules that result in the autocatalytic effect. With suitably spaced functional groups, related molecules can show the characteristic sigmoidal product growth expected for autocatalytic systems.[23] Parallel studies with nucleic acid derivatives in water have also expressed sigmoidal product growth[24] and provided detailed accurate mathematical models for self-replicating systems.[25] A very diifferent replicator, based on thymine-diaminotriazene recognition was also prepared. The electrophilic component was thymine acetic acid, moderately activated as a phenyl ester, and the replication reaction involved its coupling with an amine nucleophile (Fig 8). Again, the rate enhancement can be attributed to the ability of the

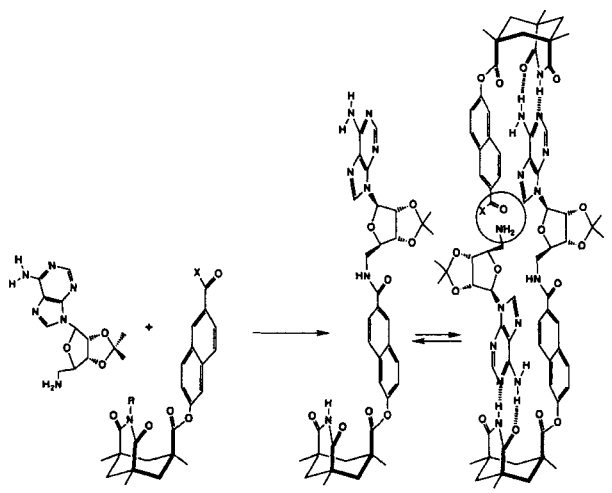

Figure 7 Template effects resultin in a self-replicating system.

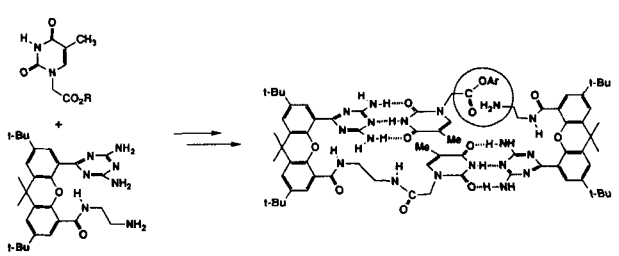

Figure 8 Replication of thymine derivatives.

product to act as a template: both reactants can assemble on its surface where they are coupled in an intracomplex reaction that leads to an exact replica of the template catalyst as a dimer. Dissociation of the dimer then exposes more template molecules, etc.

From these experiments we have arrived at the premise that any recognition event involving weak, intermolecular forces can lead to self-replicating systems; all that is required is that the two complementary entities become covalently attached.* [26] If the recognition surfaces are oriented in a manner that permits dimerization, the molecule becomes a minimalist replicator, since it can act as a template for its own formation. If the recognition surfaces diverge, the units assemble in chains as polymers or as other repeating mosaics of significance to materials science.[27] The reason *why* molecules fit together then becomes apparent: to replicate.

The binding forces need not be limited to hydrogen bonds in organic solvents. For example, cyclodextrin[28] or cyclophane[29] inclusion of aromatics in water might be used to generate a self-replicating system through covalent attachment of guest and host. We are involved in collaborative efforts with Prof. Diederich of the ETH to achieve this goal.

EVOLUTION

Can such structures show further signs of life? For evolution to occur at the molecular level, replicators are expected to make mistakes[1] or respond to environmental effects that select the more competitive species at the expense of the less competitive ones. Both of these features have now been demonstrated in the context of our synthetic structures. In one experiment[30] two replicators were subjected to a classical competition experiment (Fig 9). These molecules had such structural similarity that they catalysed not only their own formation, but each others! For the molecules at hand, in the language of organic chemistry, there is a lack of selectivity in replication. The template-catalysed reaction is insensitive to whether R_1 and R_2 (Fig 9) are the same; the nature of these two substituents on the N_6 of adenine is not as important as their mere presence, since any groups on that position favour Hoogsteen vs Watson–Crick base-pairing.[13]

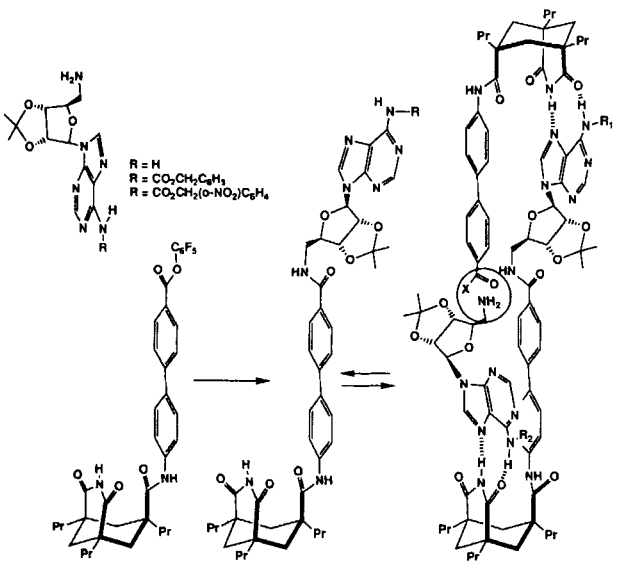

Figure 9 Replicators competing and reciprocating.

The substituents on the purine nuclei of the replicators were not, of course, chosen at random. Rather, they were intended to show that heritable, chemical 'mutations' could be observed in these replicators. We chose photochemical irradiation as a possibility for mutation because it is certain to have been a primary source of energy in a prebiotic world, and its consequences are dramatic.

In the experiment, a competition was staged between two urethane-substituted adenines for a limited amount of active ester. When the ester was consumed, the product replicator and amine solution was irradiated.[31] One of the replicators (and its precursor amine) bore photochemically labile functions, the *ortho*-NO_2-benzyloxycarbonyl group, and it was cleaved on irradiation. The resulting new unsubstituted species were more effective at replication than the original. The superiority of N_6-unsubstituted derivatives is due to their ability to base-pair in productive complexes more frequently than is possible for the competitor. This results in the mutant's enhanced reproductive success. The photochemical mutation is inherited, that is, it catalyses its own formation in subsequent generations. When more ester is added the new, unsubstituted adenosine and its template rapidly take over the system's resources; its replication is at the direct expense of its competitors. The success of the mutant as a replicator is also due to its rapid initial reaction with the ester; this can take place through the Watson–Crick base pair, where aryl stacking interactions position the two reacting functions near each other. Despite its efficiency, the mutant is not selfish; it provides effective catalysis for the formation of its competitors as well.

* Ref 26(a) gives the original formulation of self-complimentarity in replicating structures. While we have been unable to locate this source, we take comfort in the fact that so many other central notions of bioorganic chemistry have been traced to Linus Pauling. For pairwise and otherwise assembly of complementarity, if not self-complementary molecules, see refs 26(b)–26(e).

HYBRIDS

A different sort of competition experiment was staged in which hybrid (crossover) products could be produced.[32] Specifically, coupling of the adenine nucleophile to the thymine ester gave a dinucleotide analogue with a peptide/ribose linkage (Fig 10). The corresponding reaction of the triazene with the Kemp derivative gave the second recombinant structure.

At first glance, both hybrids might be expected to replicate. They bear self-complementary recognition surfaces, and can gether their respective reaction components in termolecular complexes. In fact, the adenine–thymine hybrid does so. It is an unusually fertile hybrid. When compared with our other synthetic replicators, it shows the largest autocatalytic effects observed to date. This is not the case for the other hybrid. It does not catalyse its own formation; this is a sterile hybrid.

The differences in behaviour of the recombinants can be attributed to the orientations of their respective recognition surfaces (Fig 11). In the adenine–thymine combination the hydrogen bonding surfaces can achieve a parallel U-shaped arrangement which can collect reactants in a productive termolecular complex. The other hybrid is composed of two U-shaped modules – the Kemp triacid and the xanthene diacid. Its overall skeleton or configuration is either C-shaped, in which its recognition surfaces converge (not shown), or S-shaped, in which its recognition surfaces diverge. In neither case can a productive termolecular complex be assembled. Nor can it form a hydrogen-bonded cyclic dimer; instead, its self-complementarity is expressed by forming chains.

Self-complementarity is a common feature of macromolecules. The orientation of recognition surfaces within these structures determines the nature of the assemblies formed. When they diverge, the molecular assemblies lead to one-dimensional polymer chains, two-dimensional mosaics or three-dimensional

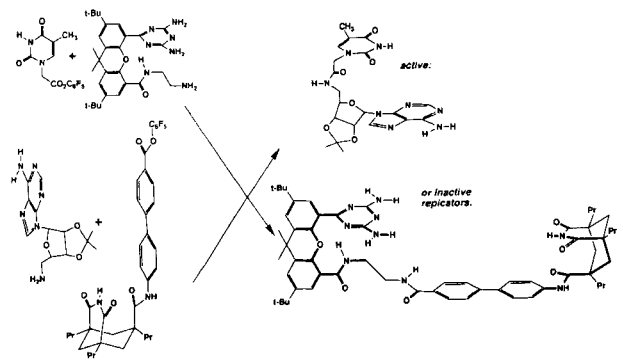

Figure 10 Recombinants from crossover reactions as active or inactive replicators.

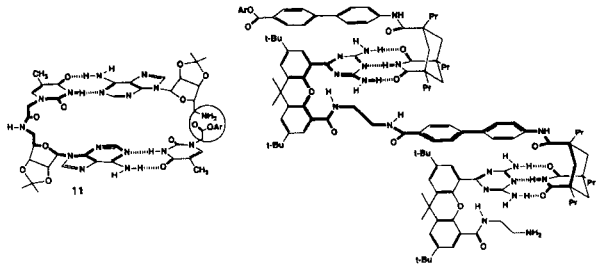

Figure 11 A U-shaped molecule that can replicate, (left) and an S-shaped molecule that cannot (right).

structures. Examples of the latter drawn from molecular biology include multisubunit enzymes, leucine zippers and viral coat proteins. With carefully fixed recognition elements, the assembly of synthetic self-complementary structures into predicable, closed surfaces that encapsulate molecules—or events of a complementary scale should also be possible.[33]

MOLECULAR BEHAVIOUR

In closing, I feel some comment on the vocabulary of this manuscript is necessary, given the frequently conflicting uses of the language by chemists versus biologists and model systems versus biochemistry. Words such as replication, mutation, adaptation and hybridization have focused well-defined meanings in biology and, to be sure, considerable understanding of these phenomena has been achieved using nucleic acids. But chemists are now reducing these phenomena to expression with entirely synthetic molecules, having little or no resemblance to biostructures. Parallel developments are taking place in computer science. Why then should biochemists feel uncomfortable when chemists use these terms? For example, 'mutation' carries a specific biological sense, tightly defined in terms of frequency and mode of transition. These strict constructionists bristle when I use the word in relation to the studies described in this article.

The chemist's agenda trascends biology, and includes molecular systems which express properties never before seen and therefore undefined. One example involves synthetic 'helicates'[34] which remotely resemble nucleic acid double helices turned inside-out. Another, 'homo-DNA',[35] is a ladder-like structure that shows many properties of replicating systems. A spectacular example in supramolecular chemistry is the autopoetic replication of micelles.[36] None of these systems are prebiotic as defined[37] by practitioners. Instead, 'extrabiotic' is proposed as a term that describes synthetic molecules or even computer constructs[2] that show life-like characteristics. We hope that the term is enough of a disclaimer to calm our more excitable colleagues in biology.

ACKNOWLEDGMENTS

I am deeply indebted to the superb coworkers whose names appear on the original publications. Funds were provided by the NSF and NIH.

REFERENCES

1 Rebek, Jr., J.; *Chem. Ind.* **1992**, 171.
2 Ray, T.S.; in *Artificial Life II* (Langton, C.G.; ed.), Addison-Wesley, **1992**, p. 371.
3 Famulok M.; Nowick, J.S.; Rebek, Jr., J.; *Act. Chem. Scand.* **1992**, *46*, 315.
4 Cech, T.R.; *Science* **1987**, 236, 1532. Joyce, G.F.; *Nature* **1989**, *338*, 217. Doudna, J.A.; Szostak, J.W.; *Nature* **1989**, *339*, 519. Doudna, J.A.; Couture, S.; Szostak, J.W.; *Science* **1991**, *251*, 1605.
5 von Kiedrowski, G.; *Angew. Chem., Int. Ed. Engl.* **1986**, *98*, 932. Zielinski, W.S.; Orgel, L.E.; *Nature* **1987**, *327*, *346*, 26.
6 Kemp, D.S.; Petrakis, K.S.; *J. Org. Chem.* **1981**, *46*, 5140.
7 Williams, K.; Askew, B.; Ballester, P.; Buhr, C.; Jeong, K.S.; Jones, S.; Rebek, Jr., J.; *J. Am. Chem. Soc.* **1989**, *111*, 1090.
8 Goswami, S.; Van Engen, D.; Hamilton, A.D.; *J. Am. Chem. Soc.* **1989**, *111*, 3425. Adrian, Jr. J.C.; Wilcox, C.S.; *J. Am. Chem. Soc.* **1989**, *111*, 8055. Zimmerman, S.C.; Wu, W.; *J. Am. Chem. Soc.* **1989**, *111*, 8054. Galán, A.; Pueyo, E.; Salmeron, A.; de Mendoza, J.; *Tetrahedron Lett.* **1992**, *32*, 1827.
9 Jeong, K.S.; Rebek, Jr. J.; *J. Am. Chem. Soc.* **1988**, *110*, 3327. Park, T.K.; Schroeder, J.; Rebek, Jr. J.; *Tetrahedron* **1991**, *47*, 2507.
10 Park, T.K.; Schroeder, J.; Rebek, Jr. J.; *J. Am. Chem. Soc.* **1991**, *113*, 5125.
11 Nowick, J.S.; Ballester, P.; Ebmeyer, F.; Rebek, Jr. J.; *J. Am. Chem. Soc.* **1990**, *112*, 8902.
12 Rebek, Jr. J.; Williams, K.; Parris, K.; Ballester, P.; Jeong, K.-J.; *Angew. Chem. Int. Ed. Eng.* **1987**, *26*, 1244.
13 Dodin, G.; Dreyfus, M.; Dubois, J.-E.; *J. Chem. Soc. Perkin 2*, **1979** 439.
14 Askew, B.; Ballester, P.; Buhr, C.; Jeong, K.S.; Jones, S.; Parris, K.; Williams, K.; Rebek, Jr. J.; *J. Am. Chem. Soc.* **1989**, *111*, 1082.
15 Benzing, T.; Tjivikua, T.; Wolfe, J.; Rebek, J., Jr.; *Science* **1988**, *242*, 266.
16 Deslongchamps, G.; Galán, A.; de Mendoza, J.; Rebek, J., Jr.; *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 61.
17 Galán, A.; Pueyo, E.; Salmeron, A.; de Mendoza, J.; *Tetrahedron Lett.* **1991**, *32*, 1827.
18 Echavarren, A.; Galan, A.; Lehn, J.-M.; de Mendoza, J.; *J. Am. Chem. Soc.* **1989**, *111*, 4994.
19 Schmidtchen, F.P.; *Tetrahedron Lett.* **1989**, *30*, 4493. Kurtzmeier, H.; Schmidtchen, F.P.; *J. Org. Chem.* **1990**, *55*, 3749.
20 Galán, A.; de Mendoza, J.; Toiron, C.; Bruix, M.; Deslongchamps, G.; Rebek, J., Jr.; *J. Am. Chem. Soc.* **1991**, *113*, 9424.
21 Feibush, B.; Figueroa, A.; Charles, R.; Onan, K.D.; Feibush, P.; Karger, B.L.; *J. Am. Chem. Soc.* **1986**, *108*, 3310. Hamilton, A.D.; Van Engen, D.; *J. Am. Chem. Soc.* **1987**, *109*, 5035. Muehldorf, A.V.; Van Engen, D.; Warner, J.C.; Hamilton, A.D.; *J. Am. Chem. Soc.* **1988**, *110*, 6561.
22 Tjivikua, T.; Ballester, P., rebek, J., Jr.; *J. Am. Chem. Soc.* **1990**, *112*, 1249. Nowick, J.; Feng, W.; Tjivikua, T.; Ballester, P.; Rebek, Jr., Jr.; *J. Am. Chem. Soc.* **1991**, *113*, 8831.
23 Rotello, V.; Hong, J.-I.; Feng, Q.; Rebek, Jr., Jr.; *J. Am. Chem. Soc.* **1991**, *113*, 9422.
24 von Kiedrowski, G.; Wlotzka, B.; Helbing, J.; Matzen, J.M.; Jordan, S.; *Angew. Chem., Int. Ed. Engl.* **1991**, *30*, 423. Terfort, A.; von Kiedrowski, G.; *Angew. Chem., Int. Ed. Eng.* **1992**, *31*, 654.
25 von Kiedrowski, G.; Woltzka, B.; Helbing, B.; *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 1235.
26 (a) Pauling, L.; *Molecular Architecture and the Processes of Life*, Jesse Boot Foundation, Nottingham, 1948. (b) Fouquey, C.; Lehn, J.-M.; Levelut, A.-M.; *Adv. Mater.* **1990**, *2*, 254. (c) Geib, S.J.; Hirst, S.C.; Vincent, C.; Hamilton, A.D.; *J. Chem. Soc., Chem. Commun.* **1991**, 1283. (d) Zerkowski, J.A.; Seto, C.T.; Wierda, D.A.; Whitesides, G.M.; *J. Am. Chem. Soc.* **1990**, *112*, 9025. (e) Seto, C.T.; Whitesides, G.M.; *J. Am. Chem. Soc.* **1990**, *112*, 6409.
27 Barr, R.G.; Pinnavaia, T.J.; *J. Chem. Phys.* **1986**, *90*, 328. Etter, M.C.; Urbanczyk-Lipkowska, Z.; Jahn, D.A.; Frye, J.S.; *J. Am. Chem. Soc.* **1987**, *108*, 5871. Simard, M.; Su, D.; Wuest, J.D.; *J. Am. Chem. Soc.* **1991**, *113*, 4696. Zimmerman, S.C.; Duerr, B.F.; *J. Org. Chem.* **1992**, *57*, 2215.
28 Breslow, R.; Greenspoon, N.; Guo, T.; Zarzyck, R.; *J. Am. Chem. Soc.* **1989**, *111*, 8296.
29 Ferguson, S.B.; Sanford, E.M.; Seward, E.M.; Diederich, F.N.; *J. Am. Chem. Soc.* **1991**, *113*, 5410.
30 Hong, J.I.; Feng, Q.; Rotello, V.; Rebek, Jr. J.; *Science* **1992**, *255*, 848.
31 Patchornik, A.; Amit, A.; Woodward, R.B.; *J. Am. Chem. Soc.* **1970**, *92*, 6333.
32 Feng, Q.; Park, T.K.; Rebek, Jr. J.; *Science* **1992**, *256*, 1179.
33 Branden, C.; Tooze, J.; *Introduction to Protein Structure*, Garland, New York, **1991**, Chapt. 11. Rebek, Jr. J.; in *Molecular Recognition II*, Royal Chemical Society Special Publication III, **1992**, 65–73.
34 Lehn, J.-M.; Rigault, A.; *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 1095.
35 Eschenmoser, A.; *Nach. Chem. Tech. Lab.* **1991**, *39*, 795.
36 Bachmann, P.A.; Walde, P.; Luisi, P.L.; Lang, J.; *J. Am. Chem. Soc.* **1991**, *113*, 8204. Bachmann, P.A.; Walde, P.; Luisi, P.L.; Lang, J.; *J. Am. Chem. Soc.* **1990**, *112*, 8200. Bachmann, P.A.; Luisi, P.L.; Lang, J.; *Nature* **1992**, *357*, 57.
37 Orgel, L.E.; *Nature* **1993**, in press.