VOLUME 23 NUMBER 12

ACCOUNTS OF CHEMICAL RESEARCH®

DECEMBER 1990

Molecular Recognition and Biophysical Organic Chemistry

JULIUS REBEK, JR.

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

Received June 20, 1990 (Revised Manuscript Received September 5, 1990)

Recent work in molecular recognition shows a shift of attention from the seamless curvature of macrocyclic polyethers to synthetic receptors in a variety of shapes, sizes, and chemical linings. Supramolecules, assemblies, devices, and even self-replicating systems have been concocted. The question is no longer *if* something can be built, but *what* to build and *why*. There are some technical issues. How rigid or flexible should these synthetic receptors be? Is it desirable, or even possible, to control their shapes and sizes to tenths of angstroms? What are the optimal complements to a given functional group? Some time ago I had the idea that cleft-like structures offered advantages for such studies. In this Account, it becomes clear that now the idea has me.

Preorganization

How well things fit together depends on their predisposition to do so, a matter frequently referred to as preorganization.[1] While the term conjures up entropic effects, it is hard to separate these from enthalpic effects. Take, for example, the comparison of 2,2′-bipyridyl and *o*-phenanthroline (Figure 1) as metal-chelating ligands. The latter has a higher affinity for metals; its ground-state conformation resembles those of its complexes. Is the enhanced affinity of *o*-phenanthroline due to the loss of rotational freedom of the bond between the two pyridyl nuclei or to the destabilized ground state in which the dipoles (lone-pair electrons) converge, or both? Can these be distinguished? A similar question can be raised for 18-crown-6 (1) vs a typical spherand **2**. Recent work on polyethers has made important progress on this question. Still[2] prepared **3**, in which conformational restraints are imposed by remote steric effects involving the heterocyclic subunits. The conformation **4**, calculated to be near the global energy minimum, is preorganized for metal binding. Its affinity for lithium ions was consequently very high, comparable to those of macrocycles such as **1**.

Our own approach began with the premise that distinguishing between enthalpic and entropic effects was more likely to succeed through using large structures rather than through further investigations of small molecules. If the ligating sites were moved far enough apart, their ability to sense each other would diminish and entropic factors would dominate and be assessable. Our initial studies involved the acridine system shown in Figure 2. In **5**, the condensation product of Kemp's triacid and proflavin, the ortho hydrogens permit rotation about the C_{aryl}–N_{imide} bond. As a result, three conformations exist: the "in–out" one shown, a divergent one, and a convergent one. The ortho methyl groups of **6**, however, lock the system into the convergent mode, that is, the OH bonds of the acid converge toward the center of the structure.

Titrations of **5** and **6** with diazabicyclooctane (DABCO) gave the respective chelated complexes **7**. Diacid **6** had a higher affinity than did **5**; the ratio of association constants was 12:1 indicating about 1.5 kcal/mol difference in affinities.[3] This value is larger than might be expected from that based on statistics. Rather surprisingly, **5** bound quinoxaline somewhat better than did **6**, the values of K_a being 2.9×10^4 M^{-1} vs 2.3×10^4 M^{-1}. This change in preference may be due to the larger

Julius Rebek, Jr., was born in Hungary in 1944. He received the Ph.D. degree from the Massachusetts Institute of Technology in 1970 for studies in peptide chemistry with Professor D. S. Kemp. He has held faculty positions at UCLA and the University of Pittsburgh. In 1989 he returned to MIT as Professor of Chemistry. His current research interests are in mechanistic organic chemistry, enzyme models, prebiotic chemistry, and molecular recognition.

(1) Cram, D. J. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 1009–1020.
(2) Iimori, T.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 3439–3440. Iimori, T.; Erickson, S. D.; Rheingold, A. L.; Still, W. C. *Tetrahedron Lett.* **1989**, *30*, 6947–6950.
(3) For a convenient synthesis, see: Rebek, J., Jr.; Askew, B.; Killoran, M.; Nemeth, D.; Lin, F.-T. *J. Am. Chem. Soc.* **1987**, *109*, 2426–2431.

0001-4842/90/0123-0399$02.50/0

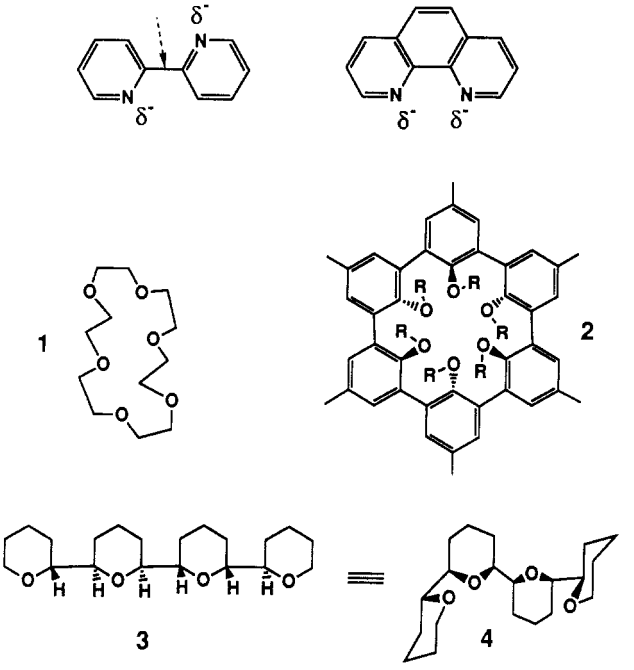


Figure 1. Preorganization with enthalpic destabilization.

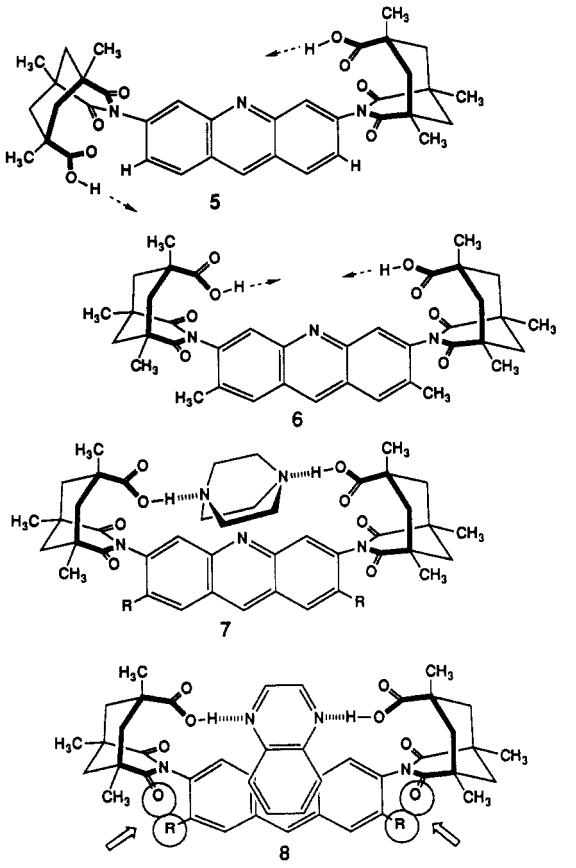


Figure 2. Convergent, divergent, and chelated complexes of a diacid.

length of pyrazine vs DABCO. As Jorgensen et al.[4] have pointed out, a pyrazine nucleus is slightly too large for true chelation by the structure bearing the ortho methyls and prefers to float above the diacid plane. The ortho hydrogens would permit more rotation and freedom from buttressing effects, allowing **5** to accommodate the heterocycle in a chelated manner as indicated in **8**.

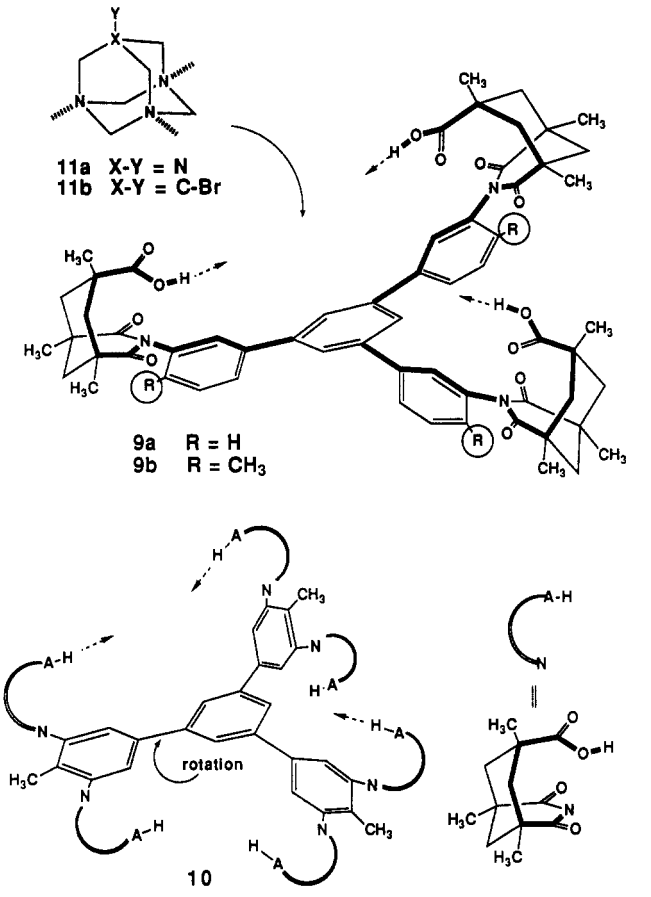


Figure 3. Preorganization of multiple contact points.

It seemed that even larger structures were required, lest the conformational restrictions that are introduced inadvertently change the size or the shape of the receptor and mask the issue. Accordingly, we moved to the triarylbenzenes as spacers to hold apart three carboxyl groups provided by Kemp triacid subunits (Figure 3).[5] The molecules are readily prepared by acid-catalyzed trimerization[6] of the appropriate nitroacetophenones followed by catalytic reduction. The amines are then condensed with the anhydride acid chloride derivative of Kemp's triacid.

Direct comparisons could be made in this system between ortho hydrogen an ortho methyl, e.g., structures **9a** and **9b**. Complexation of hexamethylene tetramine **11a** titrated in CD_3CN showed the ortho methyl compound to be some 2 times more effective than the ortho hydrogen compound.[7] With the triamine **11b** the ratio of association constants was 3.5:1. In either **9a** or **9b**, rotation about the aryl–aryl bonds permits conformations in which all three carboxyls are on the same side, or another in which two are one side and one is on the other. Thus **9b** has two conformations whereas the ortho hydrogen molecule **9a** has eight. The seemingly passive methyl groups have, at best, the 4-fold effect anticipated on purely entropic grounds. We also examined the hexaacid **10** with other azaadamantane

(4) Jorgensen, W. L.; Boudon, S.; Nguyen, T. B. *J. Am. Chem. Soc.* **1989**, *111*, 755–757.

(5) Kemp, D. S.; Petrakis, K. S. *J. Org. Chem.* **1981**, *46*, 5140. Commercially available from the Aldrich Chemical Co.

(6) Ebmeyer, F.; Vögtle, F. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 79–81.

(7) Horiuchi, N.; Huff, J.; Rebek, J., Jr. *Tetrahedron Lett.* **1990**, *31*, 5121–5124.

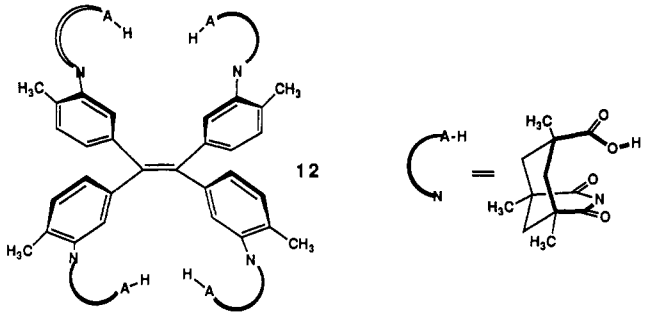


Figure 4.

Table I
Binding of Amines to 12

guest	$K_{1:1}$, M^{-1}	$K_{1:2}$, M^{-1}
hydrazine	630	35000
ethylenediamine	9	1700
DABCO	150	360
2-aminopyrimidine	1400	90

derivatives. Its preorganization is as complete as we can make it. Neither surprises nor allosteric cooperativity was observed in its binding behavior; rather, statistical effects characterized its affinity.

The allosteric effect could be observed in a system based on the tetraarylethylene skeleton[8] (Figure 4). The results of the binding studies with **12** and several diamines are shown in Table I. Very large cooperativity is observed, e.g., nearly 200-fold for the smaller guests. Presumably, intramolecular hydrogen bonds exist in the uncomplexed host species, but upon the formation of a 1:1 complex, these are destroyed, forcing the two nonparticipating carboxyls to the opposite face of the structure. These carboxyls are then exposed and more available for amine binding.

Allosteric effects may be one means by which nature addresses the problem of preorganization. Binding at one site is used to organize (or disorganize) the remote site and turn on (or off) the enzyme's activity. The model systems shown here demonstrate that this can be an effective strategy.

Another means of preorganization involves the use of intramolecular hydrogen bonding in the same sense in which they are used to establish secondary structure in proteins: to position functional groups for substrate binding. Our observations[9] were with the xanthene diacid **13**, a molecule that features the U-turn but offers larger spaces in its spacer-linked derivatives, e.g., **14** (rotations around the indicated bonds in Figure 5 lead to cleft-like shapes). The intramolecular hydrogen bonds preorganize the cavity for binding of quinoxaline-2,3-dione as shown in **15**. The arrangement is inappropriate for diketopiperazine, which is not bound in solid–liquid extraction experiments.

Subkilocalorie Effects in Hydrogen Bonding

We have already reported extensively on the use of model compounds derived from Kemp's triacid for the study of base pairing.[10] The advantage they offer is that both hydrogen bonding and aryl stacking interactions can be observed in the relatively noncompeting solvent $CDCl_3$. This solvent acts as a magnifying glass permitting associations that would not be observable in a more competitive environment such as H_2O. Binding in such synthetic receptors has led to a profoundly useful analysis of secondary interactions in hydrogen-bonding arrays.[11] These can be stabilizing or destabilizing and can therefore affect association constants quite dramatically. For example, in the cytosine–guanine base pair **16** (Figure 6) these secondary interactions—involving intermolecular forces between partially charged atoms—cancel, leaving three hydrogen bonds to stabilize the system. In the uracil–diaminopyridine complex **17** the secondary interactions are all repulsive and reduce the stability to that typically observed for only two hydrogen bonds.

We have evaluated these secondary effects in three systems. In the first, the relative hydrogen-bonding affinities of imides and lactams were examined (Figure 7).[12] Such functional groups have been studied in the past, but their self-association or dimerization is quite small in $CDCl_3$. However, with the U-turn inherent in Kemp triacid derivatives, it was possible to arrange the acidic and basic sites to converge in an intramolecular manner. Three sets of structures were prepared: the diimides **18**, the racemic dilactams **19**, and the hybrid imide–lactam **20**. In each set the functions were separated by flexible tethers of varying length. Their tendencies to form cyclic (intramolecular) arrays of hydrogen bonds are summarized in Table II.

The lactams consistently showed a higher tendency to self-associate than did the imide functions. In accord with Jorgensen and Pranata's analysis,[11] the spectator oxygens of the imides destabilize nearby hydrogen bonds. The effect appears to involve ~0.4 kcal/mol in these systems per destabilizing interaction, as shown in **21** and **22**. The enhanced acidity of the imide is of no advantage in hydrogen bonding to other imides or lactams; however, it does appear to be an important factor in base pairing to adenines.

Our second system involved the use of adenines as probes. The specific cases involved were the naphthyl derivatives shown in Figure 8. The imide **23a** showed *higher* affinity (by approximately 1 kcal/mol) than the lactam **23b** for 9-ethyladenine (**24**), despite the presence of the spectator carbonyl.[13] In another comparison, the flexible diimide with a five-carbon spacer **25a** and the corresponding (meso) dilactam **25b** were evaluated. Again, the imides outperformed the dilactams by a very large margin in the formation of complexes **25**. Finally, the derivatives **26** featuring a rigidified naphthalene spacer were compared (the propyl derivatives enhance solubility). Here the effect was quite large, involving a difference of several kilocalories/mole in binding affinity. The adenine binding results are summarized in Table III.

That nature chose imides rather than primary amides to base pair with adenines carries with it some message, but what is it? It might be found in a recent calcula-

(8) Ebmeyer, E.; Rebek, J., Jr. *Angew. Chem., Int. Ed. Engl.*, in press.
(9) Nowick, J. S.; Ballester, P.; Ebmeyer, J., Jr. *J. Am. Chem. Soc.*, in press.
(10) Askew, B.; Ballester, P.; Buhr, C.; Jeong, K. S.; Jones, S.; Parris, K.; Williams, K.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1989**, *111*, 1082–1090. Williams, K.; Askew, B.; Ballester, P.; Buhr, C.; Jeong, K. S.; Jones, S.; Rebek, J., Jr. *Ibid.* **1989**, *111*, 1090–1094.
(11) Jorgensen, W.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008–2010.
(12) For a preliminary account of these results, see: Jeong, K. S.; Tjivikua, T.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1990**, *112*, 3215–3217.
(13) Jeong, K.; Tjivikua, T.; Muehldorf, A.; Deslongchamps, G.; Famulok, M.; Rebek, J., Jr. *J. Am. Chem. Soc.*, in press. For the use of macrocycles in selective recognition of barbiturates, see: Chang, S. K.; Hamilton, A. D. *J. Am. Chem. Soc.* **1988**, *110*, 1318–1319.

Table II
Cyclization Data for the Systems of Figure 7 in $CDCl_3$ at 25 °C

	imide–imide 18			lactam–lactam 19		lactam–imide 20	
n	% cyclic	K_{obsd}	K_{calcd}(corr)	% cyclic	K_{calcd}	% cyclic	K_{calcd}
2	63	1.7	0.85	86	6.1	75	3
3	83	5.2	2.6	96	24	87–90	8
4	32	0.45	0.22	48	0.92	27–44	0.43–0.78
5	39	0.64	0.32	55	1.2	39–45	0.66–0.82

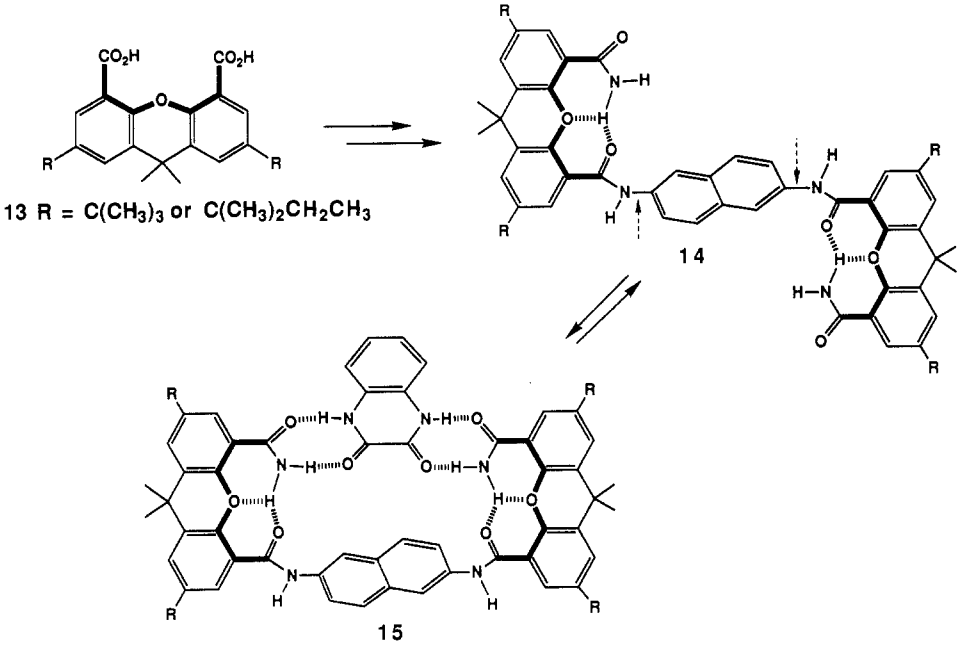


Figure 5. Synthesis and complexation of large molecular clefts.

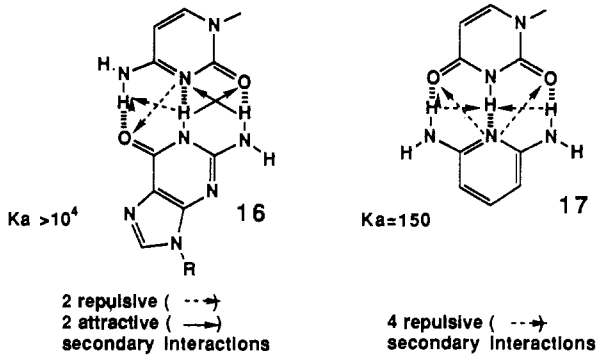
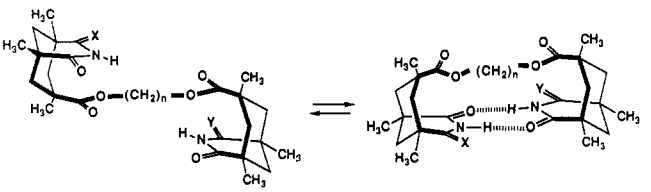


Figure 6. Secondary interactions in hydrogen-bonding patterns.

tion[14] in which the calculated structure of the base pair shows a great deal of polarization of the C_2-H. Thus a sort of hydrogen bond exists between the "spectator" carbonyl and this polarized hydrogen as shown in **27**. Presumably a similar polarization occurs on the Hoogsteen edge (C_8-H) in the chelated derivatives. The interpretation draws support from the unconventional hydrogen bonds to CH functions in caffeine observed previously in the crystalline state.[15] Thus model systems provide a means by which such subtle but intrinsic factors in hydrogen bonding can be assessed at the subkilocalorie level.

The third system in which secondary hydrogen-bonding effects were encountered involved the binding of peptides. It has been possible to parlay the self-affinity of lactams into highly selective receptors for optically active diketopiperazines and related heterocycles. For example, one enantiomer of a dilactam (resolved on a Pirkle column)[16] formed a complex with cyclo(L-



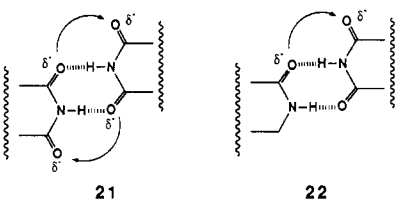


Figure 7. Relative hydrogen-bonding affinities of imides and lactams.

Table III
Association Constants ($CDCl_3$, 24 °C) for Imides vs Lactams in Binding of Adenines

imides		lactams	
compd	assoc const, M^{-1}	compd	assoc const, M^{-1}
23a	110	**23b**	15
25a	2400	**25b**	52
26a	50000	**26b**	290

(14) Jorgensen, W.; Severance, D. *J. Am. Chem. Soc.*, in press.
(15) For a discussion, see: Donohue, J. In *Structural Chemistry and Molecular Biology*; Rich, A., Davidson, N., Eds.; W. H. Freeman: New York, 1968; pp 459–465.

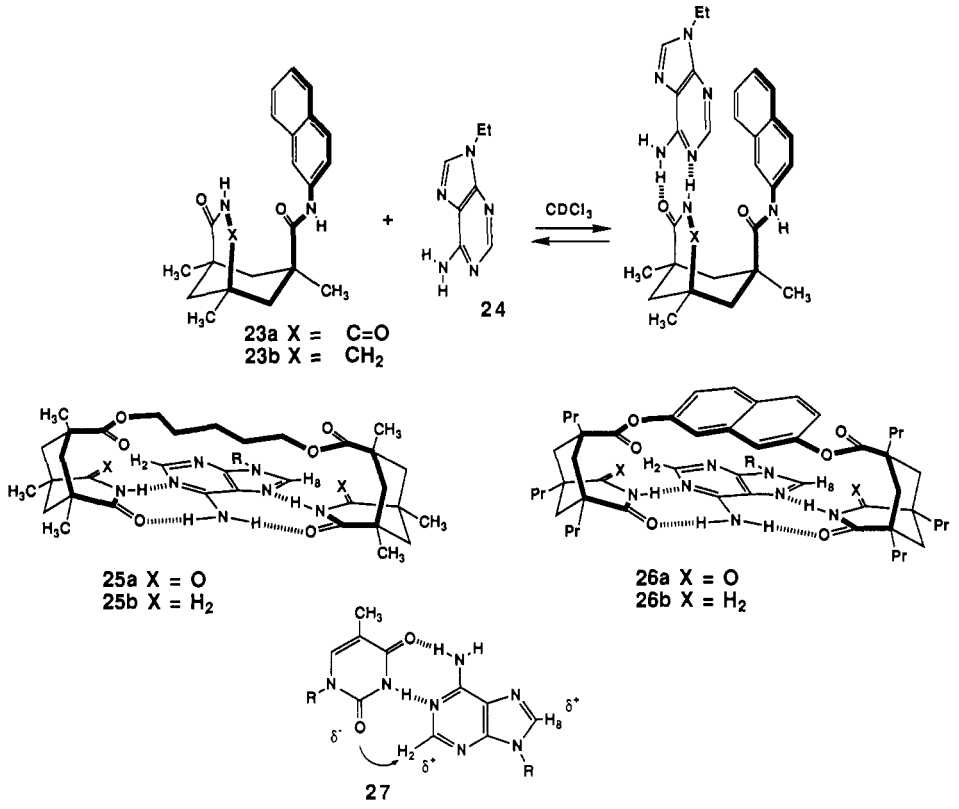


Figure 8. Complexation of adenines by imides and lactams.

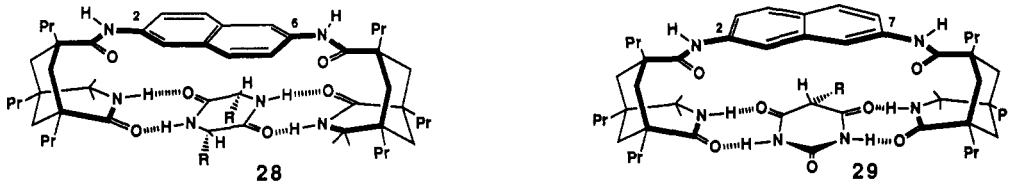


Figure 9. Selective microenvironments for diketopiperazines and barbiturates.

leucine)$_2$ **28**, Figure 9, which showed >2.5 kcal/mol higher stability than did its enantiomer. The meso compound **29** showed excellent complementarity to barbiturates. The corresponding diimides were inferior receptors for either heterocyclic guest.

The fit of diketopiperazines into the dilactam in complex **28** is an expression of *shape* complementarity; both components feature C_2 symmetry, and their affinity to each other is not surprising. What is surprising, and to us even bewildering, is the widespread use of C_2-symmetric receptors for enantioselective recognition of substrates bearing single asymmetric centers. This resembles the clumsiness and misfit of a person stumbling through a turnstile.[17] In the few cases where comparisons are possible,[18] abandoning C_2 symmetry has paid dividends for the adventuresome. Moreover, the most successful chromatographic resolutions[16] involve chiral surfaces featuring simple asymmetric structures of C_1 symmetry.

Merging Recognition and Catalysis

One of the long-standing ambitions of biomimetic chemistry has been the design and synthesis of enzyme models. Molecular recognition plays a key role in the binding step, that is, the formation of the reactive complex. Macrocyclic structures have been useful in this regard, and recent advances by Kelly[19] indicate that functional groups in the complexes can be arranged to promote actual chemical reactions in acyclic settings as well. We presumed that convergent functionality would provide an advantage in that activity could be "focused" in a highly localized manner at an active site. The ultimate goal was to merge the recognition and reaction steps *in space and in time* such that maximum binding would occur to transition states, as was anticipated by Pauling so many years ago.[20]

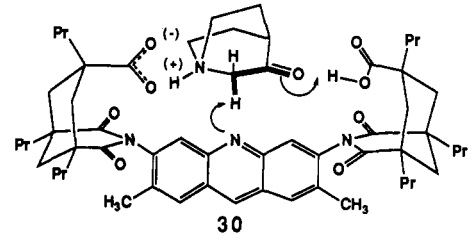


Figure 10. Acid and base converge from perpendicular directions.

(16) Pirkle, W. H.; Pochapsky, T. *Chem. Rev.* **1989**, *89*, 347–362.
(17) For discussions and leading references, see: Topiol, S.; Sabio, M. *J. Am. Chem. Soc.* **1989**, *111*, 4109–4110. Rebek, J., Jr.; Askew, B.; Ballester, P.; Doa, M. *J. Am. Chem. Soc.* **1987**, *109*, 4119–4120. Reference 16 of this Account.
(18) Castro, P. P.; Georgiadis, T. M.; Diederich, F. *J. Org. Chem.* **1989**, *54*, 5835–5838.
(19) Kelly, T. R.; Zhao, C.; Bridger G. J. *J. Am. Chem. Soc.* **1989**, *111*, 3744–3746. For recent examples of catalysis within macrocyclic complexes, see: Hosseini, M. W.; Lehn, J. M.; Jones, K. C.; Plute, K. E.; Mertes, K. B.; Mertes, M. P. *J. Am. Chem. Soc.* **1989**, *111*, 6330–6335. Anslyn, E.; Breslow, R. *J. Am. Chem. Soc.* **1989**, *111*, 5972–5973.
(20) Pauling, L. *Chem. Eng. News* **1946**, *24*, 1375–1377.

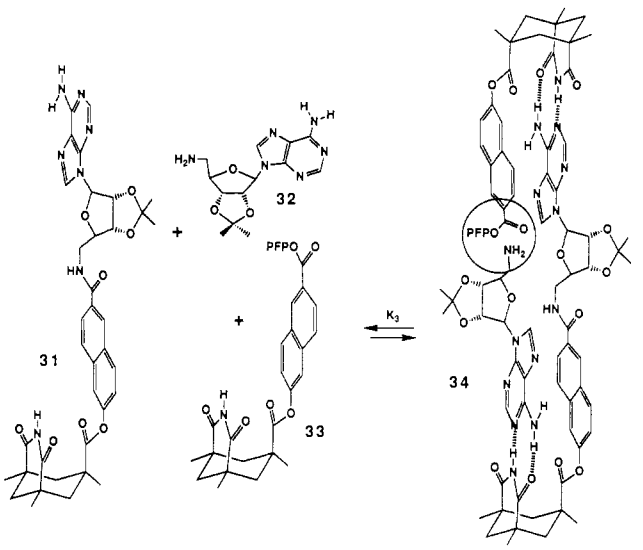


Figure 11. Autocatalysis in a self-replicating system.

We have made some desultory progress in two areas. The first involves enolization of the quinuclidinone **30** (Figure 10). Conventional acids and bases, and even their mixtures, are quite poor in catalyzing the α-H exchange of this material in deuterated solvents. However, the acridine diacid is excellent in this regard.[21] The structure fits nicely within the cleft where a constellation of functional groups is poised to affect the bond-breaking and -making steps. In particular, acid and base functions converge from perpendicular directions as is expected[22] to be optimal for enol formation directly from the quinuclidinone. The result is a 10^4 rate enhancement for the exchange of the α-protons in $CDCl_3$ saturated with D_2O. The system shows many of the earmarks of enzyme catalysis including selectivity, saturation kinetics, and competitive inhibition. The specificity for substates of complementary size was also observed for hemiacetal cleavage catalyzed by related structures.[23]

The rate enhancement is, of course, a function of what reference reaction or catalyst is chosen, and such comparisons are almost always unsatisfactory. For example, in bimolecular vs intramolecular systems, the orientation of the functional groups—the stereoelectronics—is inevitably different and structural similarities fade quickly. Fitting substrates into existing receptors is a time-honored, but meretricious, activity in bioorganic chemistry. The honest goal is to develop *general* enolization catalysts and thereby address the questions of concerted vs stepwise mechanisms.

Prebioorganic Chemistry

For the second application of recognition in catalysis, we developed an intriguing system capable of *self-replication.*[24] The structure **31** (Figure 11) is self-complementary because it features an adenine covalently bound to a receptor for adenine. While it can (and does) associate to form a dimer, it is unable to fold shut upon itself in an intramolecular manner. However, it is capable of gathering both of the two components from which it is made (**32** and **33**) on its surface as a transient termolecular complex. This arrangement promotes an otherwise bimolecular reaction to a unimolecular one. Adding the template product to a mixture of the active ester and the amine nucleophile enhances their coupling rate. From the known association constants, the relative concentrations of the various species can be calculated; the termolecular complex **34** is typically present as only a few percent of the total concentration, but its efficiency is quite high.

This system shows the unusual feature of autocatalysis. While this has been observed in nucleic acid systems, both enzymatically and nonenzymatically,[25,26] the case at hand involves the formation of an amide or peptide bond enhanced by base-pairing events elsewhere in the molecule. Thus it provides a tenuous bridge between the nucleic acid and peptide worlds.[27] It is possible that self-complementary peptides will be capable of self-replication, a property that has previously been associated exclusively with nucleotides. If so, some rethinking will be required on the subject of the ancestral molecule. For the moment, we are pursuing the construction of sturdier bridges between these two systems in the form of adaptor molecules, and we will report on them in due course.[28]

Outlook

The study of weak intermolecular forces emerged as one of the more exciting research topics of the 1980s. In this Account model systems have been emphasized, but enviable contributions to the evaluation of hydrogen bonding and hydrophobic effects have also come from site-directed mutagenesis of proteins[29] and the study of DNA–drug interactions.[30] Now, the pace of research is being further accelerated through studies using monoclonal antibodies.[31] Accordingly, research in molecular recognition appears destined to continue well into the next decade.

I am most grateful for the capable experimental assistance and intellectual stimulation provided by my co-workers; their names appear in the original literature citations. I thank Prof. W. Jorgensen for helpful discussions and for sharing his insights with us. Financial support was provided by the National Science Foundation through their Special Creativity Grant program, the National Institutes of Health, and the Ciba Foundation.[32] *This Account is dedicated to the memory of my collaborator and friend, Professor Francisco Gaviña, deceased May 1990.*

(21) Wolfe, J., unpublished observations.
(22) Rebek, J., Jr. *J. Inclusion Phenom. Mol. Recognit. Chem.* **1989**, *7*, 7–17.
(23) Wolfe, J.; Nemeth, D.; Costero, A.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1988**, *110*, 983–984.
(24) Tjivikua, T.; Ballester, P.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1990**, *112*, 1249–1250.
(25) Zielinski, W. S.; Orgel, L. *Nature* **1987**, *328*, 346–347.
(26) von Kiedrowski, G.; Woltzka, B.; Helbing, B. *J. Am. Chem. Soc.* **1989**, *28*, 1235–1237.
(27) Orgel, L. E. *Cold Spring Harbor Symp. Quant. Biol.* **1987**, *52*, 9–16.
(28) Rebek, J., Jr. *Experientia*, in press.
(29) Fersht, A. R. *Trends Biochem. Sci.* **1987**, *12*, 301–304.
(30) Maher, L. J.; Wold, B.; Dervan, P. B. *Science* **1989**, *245*, 725.
(31) See, for example: Huse, W. D.; Sastry, L.; Iverson, S. A.; Kang, A. S.; Alting-Mees, M.; Burton, D.; Benkovic, S. J.; Lerner, R. A. *Science* **1989**, *246*, 1275–1281.
(32) This Account is based on a lecture given by the author at Ciba Symposium 158 in London, July 1990.