Journal of Inclusion Phenomena and Molecular Recognition in Chemistry 7 (1989), 7–17.

New Molecular Shapes for Recognition and Catalysis

JULIUS REBEK, JR
Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, U.S.A.

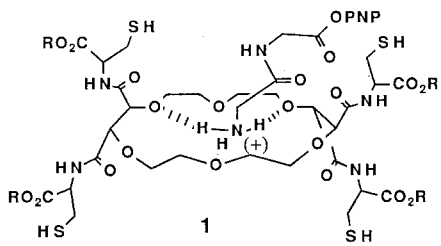
(Received: 1 February 1988)

Abstract. Progress in molecular recognition is reviewed with special emphasis on the advantages offered by molecular clefts. These new structures are rapidly assembled from readily accessible starting materials and feature functional groups that converge on smaller species that present complementary surfaces. The sizes and shapes of the clefts are controlled by the use of appropriate spacer elements. The selective binding of acids, amines, amino acids, metal ions, heterocyclic compounds and nucleosides is described. Their special applicability to problems involving concerted catalysis is also introduced.

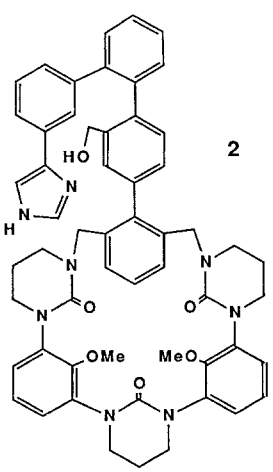
Key words. Recognition, hydrogen bonding, molecular clefts, convergent functionality, concerted catalysis, stereoelectronic effects.

1. Introduction

With the award of the 1987 Nobel Prize in chemistry to Cram, Lehn and Pedersen, model systems for bioorganic chemistry have become respectable. In the not-too-distant past, biochemical phenomena such as catalysis, recognition and transport were believed to be unique properties of macromolecules. Recent successes in imitating such phenomena using much smaller compounds has shown that chemical behaviour can be engineered into simple molecules. Crown ethers, for example, are satisfactory models for allosteric cooperativity [1] and they have gained much popularity as enzyme models [2]. In particular, the crown ether **1** described by Lehn [2c] and the macrocyclic cavitand **2** described by Cram [3] have been successful in showing large rate enhancements in their reactions with p-NO_2-phenylester

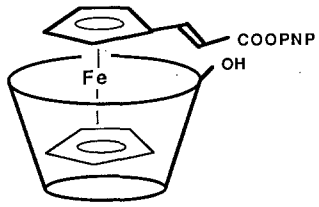


1

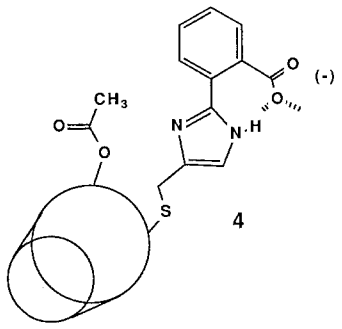


2

derivatives. The cyclodextrin systems explored by Bender and Breslow [4] have shown similar rate enhancements for the hydrolysis of other esters such as **3** and **4**.



3



4

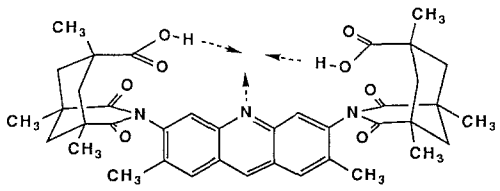
2. Structural and Synthetic Considerations

These models stress the use of binding to enhance reaction rates, and the advantage of readily available binding forces provided by these systems is the primary reason for their widespread use. The difficulty in attaching catalytically useful functionality to the parent macrocyclics is their principal disadvantage. For example, Lehn's system **1** requires four functional arms to provide a nucleophile near the bound ester, and a heroic synthetic effort is required in the Cram protease model **2**. Even so, only two of the three functional groups of the catalytic triad are in position to react with the substrate. Similar problems face cyclophane-based structures [5]. The use of such model systems to demonstrate concerted catalysis is even more difficult [6]. We have turned to new molecular shapes to overcome these disadvantages, and while development of macrocylic structures continues unabated, a few recent reports [7] suggest our views are gaining some acceptance.

The premise is that well-placed functionality can create a unique micro-environment for rate enhancements, and we are applying our recent discoveries [8] concerning molecular clefts to such situations. Specifically, we are developing systems in which functional groups converge (as in **5**) on substrates bound within. This structural strategy permits functional groups to act on several sides of a substrate molecule simultaneously and thereby enhance both recognition and catalysis.

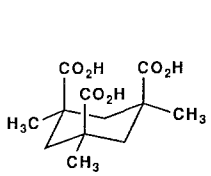
Our departure from the classical macrocyclic shapes was made possible by access to the unusual triacid **6**, first described by Kemp [9], in which three equatorial methyl groups force the carboxyl functions to a triaxial arrangement. The resulting U-shaped relationship that exists between any two carboxyl functions provides an opportunity to reverse the direction of bonds within a given molecule. For example, condensation of **6** with most primary amines gives imides, but with diamines derived from aromatics (e.g. *m*-xylidinediamine) an unusual C-shaped structure **7** results. The *ortho* substituents serve to prevent rotation about the C_{aryl}—N_{imide} bond, and enforce the convergent conformation of two carboxyl groups as shown. Indeed, **7** shows all the spectroscopic earmarks of a conventional hydrogen-bonded dimer of carboxylic acids, but is unable to dissociate into the monomer.

Convergent functionality

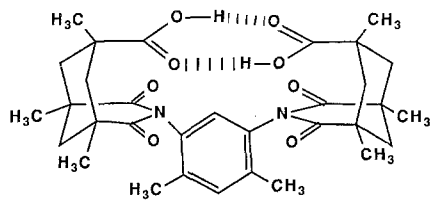


5

Scheme 1



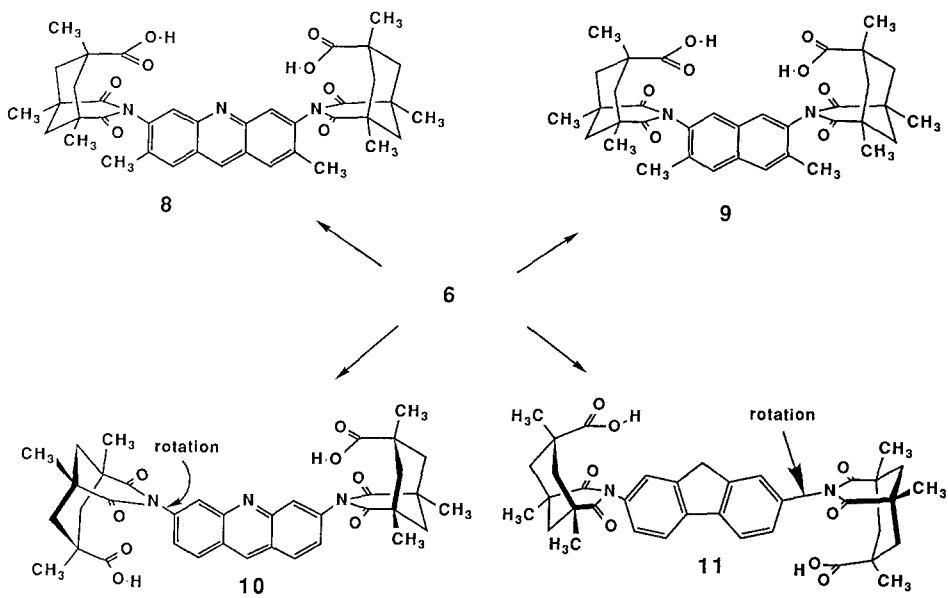
Kemp triacid 6



7

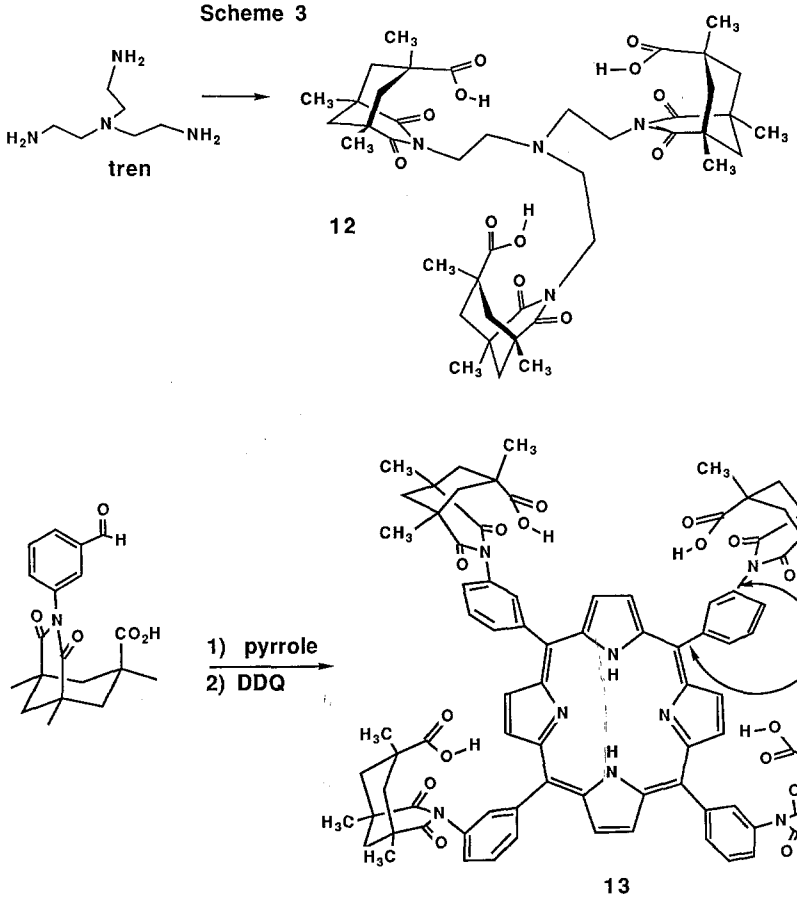
Using larger aromatic systems as spacers the distances between the acid functions can be increased (Scheme 2). The diacid derived from acridine yellow gives a system (**8**) in which ~ 8.5 Å separates the opposing carboxyl oxygens, while the naphthalene spacer **9** provides a distance of 5.6 Å. Other suitable diamines are also commercially available; e.g. 2,7-diaminoacridine gives **10** in which rotation permits several conformations [10]. The diaminofluorene derivative **11** features slightly different angles and distances; these subtle changes are reflected in its altered binding selectivities.

Scheme 2



These systems all permit conformations in which two carboxyl groups converge, but the condensation of **6** with primary amines is sufficiently general that it may be used as an architectural cliche. A number of sizes and shapes for model molecular receptors can be constructed with a minimal synthetic investment. For example, with tren, a new triacid **12** is obtained and four convergent carboxyls e.g. **13** may be prepared through Lindsey's [11] porphyrin synthesis, using the appropriate aldehyde (Scheme 3).

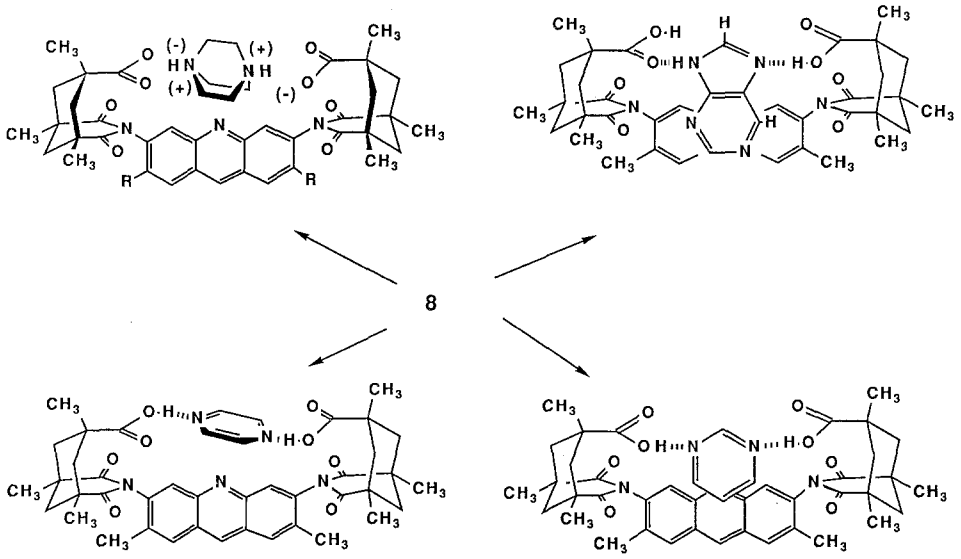
Scheme 3



3. Complexation of Complementary Structures

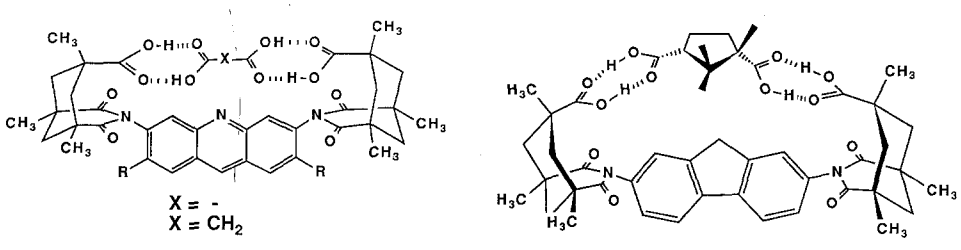
The model receptors **7–13** provide a range of sizes and shapes with which the rules for binding selectivity to smaller molecules can be laid out. Substrates with complementary basicity are ideal and we have studied the complexation of heterocyclic diamines extensively. The acridine **13** binds tenaciously to pyrazine, DABCO and other heterocycles that bridge the gap between the carboxyl functions [12]. With imidazole and its derivatives 2 : 1 complexes are formed; both the acid and base character of the heterocycle is expressed in the binding event (Scheme 4).

Scheme 4



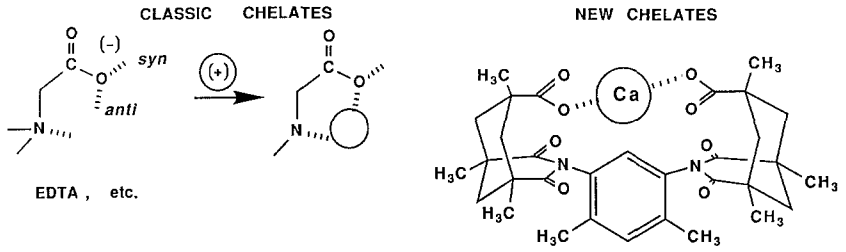
Dicarboxylic diacids offer complementary functionality in the form of hydrogen bonding that gives rise to the dimerization of most carboxylic acids in non-competing media. The acridine spaces of **8** and **10** accommodate oxalic and malonic acids but reject longer diacids such as succinic or glutaric acid [13]. The fluorene spacer of **11** shows affinity for glutaric and camphoric acids but eschews the smaller diacids (Scheme 5).

Scheme 5



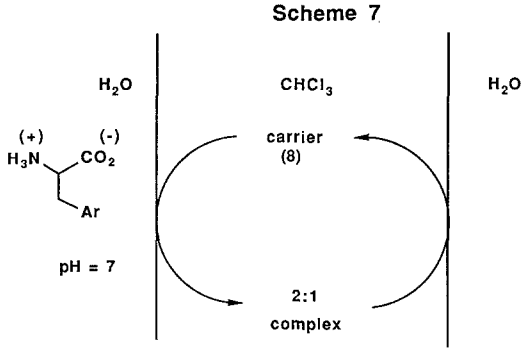
The chelation of small molecules described above may be extended to metal ions. The convergence of the carboxyls within the molecular clefts provides a microenvironment ideal for divalent metals. A special structural feature of the new ligands involves stereoelectronic effects at carboxyl oxygen. Classical chelates such as EDTA present the less *anti* lone pairs to the metal ion, but the new structures offer the more basic *syn* lone pairs. Metals such as Ca^{++} and Mg^{++} are tightly bound and readily transported across liquid membranes (Scheme 6). In addition, the mode of binding within the new ligands is exclusively *trans*, a feature which is likely to lead to altered reactivity of the bound metal ions as catalysts.

Scheme 6



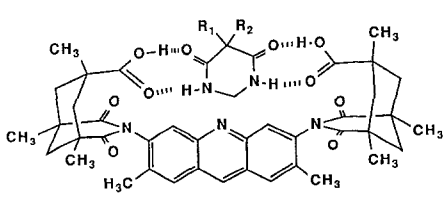
Binding of certain amino acids, as neutral (zwitterionic) substrates within the cleft of the acridine substrates has also been observed [14]. For such complexes, an additional element of recognition exists: stacking interactions between the acridine surface and the aromatic side chains of tryptophan, phenylalanine and tyrosine are detected by NMR methods. This specific contact – which is also seen with other β-phenethylamines [15] – results in selective transport of the β-aryl amino acids across liquid membranes with these carriers as 2 : 1 complexes (Scheme 7).

Scheme 7

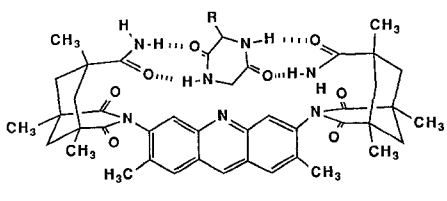


Other neutral heterocyclics can be bound within these molecular receptors provided that appropriate matching of functionality is arranged. For example, the acridine **8** can form cooperative hydrogen bonds with cyclic diamides such as primidone **14**, whereas its diamide derivative **15** provides an ideal microenvironment for diketopiperazines [16] (Scheme 8). A variety of heterocycles of differing size, shape and functionality were used to delineate the 'promiscuity' of the synthetic molecular clefts.

Scheme 8



14

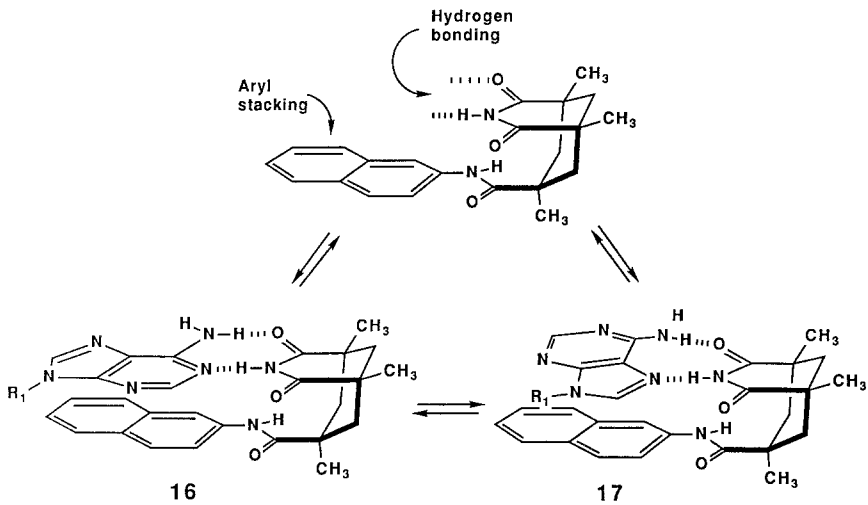


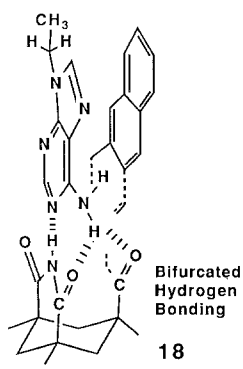
15

4. Nucleic Acids

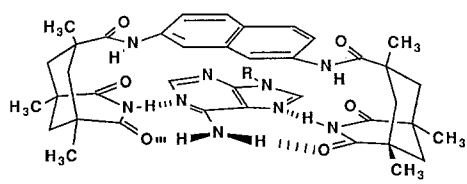
Another recent target for practitioners of molecular recognition involves nucleic acid components. For this, the Kemp triacid is used to prepare simple imides having aromatic surfaces near and parallel to the plane of the imide function. This arrangement of functionality presents a surface complementary to adenine derivatives; base-pairing and aryl stacking forces converge from perpendicular directions to provide a complementary surface for adenine [17]. Watson-Crick (**16**), Hoogsteen (**17**), and even bifurcated hydrogen bonding (**18**) is detected by NMR methods [8]. Simultaneous binding in both case-pairing senses can be achieved by the bisimide (**19**), a molecular chelate for adenine derivatives (Scheme 9).

Scheme 9





18

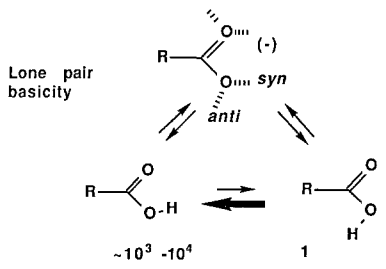


19

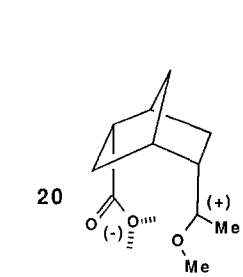
5. Catalytic Applications

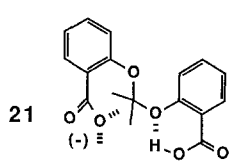
A unique advantage of the new structures relevant to their significance as enzyme models is the possibility of exploring stereoelectronic effects at carboxyl oxygen. Unlike the situation with acyl carbon, reaction trajectories and lone pair orientation at carboxyl oxygen have been difficult to assess. Gandour [19] has pointed out that in most enzymes where the carboxylate appears at the active site, the more basic *syn* lone pair is involved in catalysis. (Scheme 10).

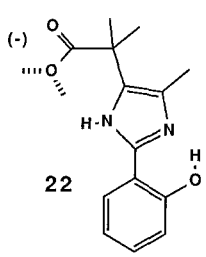
Scheme 10

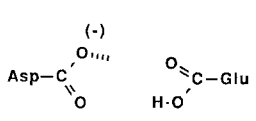


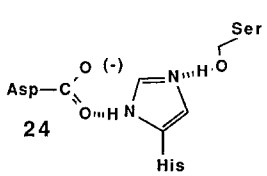
In previous models for lysozyme or the serine proteases this orientation effect has been neglected. The Loudon [20] structure **20** involves the less basic *anti* lone pair for stabilizing the developing carbonium ion, while the Fife [21] case **21** involves both the less basic lone pair and the *anti* form of the carboxylic acid. The Bruice [22] system **22** is typical of serine protease models. These may be contrasted with the convergent arrangement of the diacids in lysozyme **23**, or the involvement of the Asp *syn* lone pair in the serine proteases **24**.







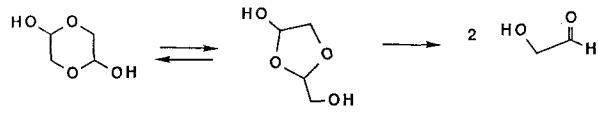


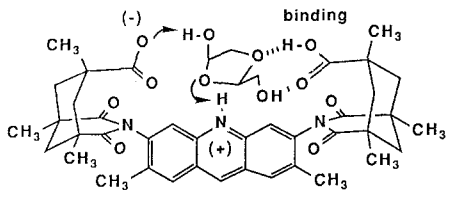


The molecular clefts are able to orient carboxyl groups in a specified direction and we have observed unusual reactivity with hemiacetals that fit within its confines [23]. The convergent diacid **8** causes the rapid dissociation of the dimer of

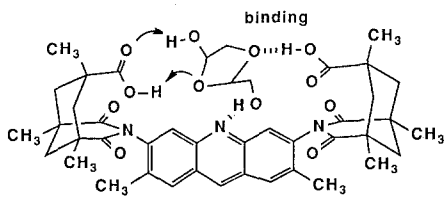
glycolaldehyde (Scheme 11). Initial complexation followed by catalysis of hemiacetal cleavage is the most probable scenario for this reaction structures **25** and **26** offer two possibilities. Pyridone or simple carboxylic acids, much admired for their ability to catalyze glucose mutarotation [24], are ineffective in this dissociation reaction.

Scheme 11





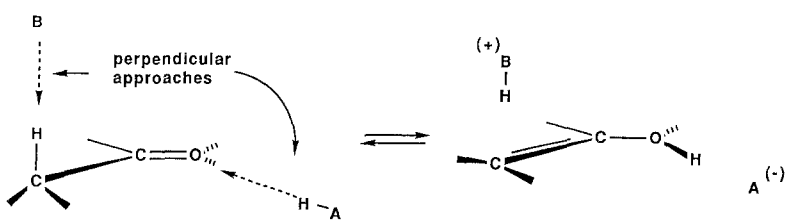
25

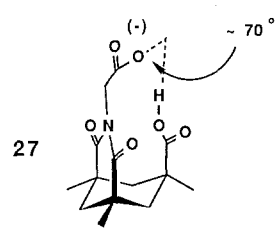


26

Another promising system involves the concerted acid/base catalysis of enolization of ketones. Stereoelectronic considerations [25] indicate that an optimal catalyst for this reaction requires that acid and base components converge from perpendicular directions on the ketone (Scheme 12). The structure **27** (a glycine derivative of Kemp's triacid) exhibits considerable activity in the enolization of phenylacetone. It should be possible to engineer additional points of contact between ketone (substrate) and diacid (catalyst) to enhance the enolization process, and we are working toward this goal.

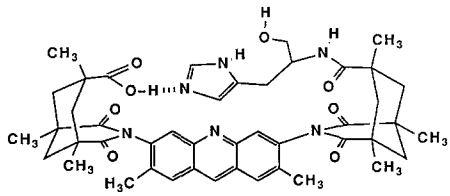
Scheme 12



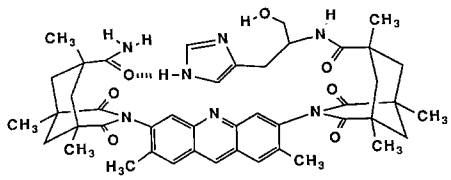


27

It is possible to 'line' the sides of the cleft with a number of functional groups. For example, derivatization of the acridine diacid **8** with histidino [26] gives **28** (Scheme 13). This structure represents the first model system in which the appropriate lone pair of the carboxylate is directed toward an imidazolium ion. We have also prepared the corresponding amide **29**; its reactivity will be compared to that of **28** in the same way as the mutant chymotrypsin was recently compared with the wild-type enzyme [27].



28

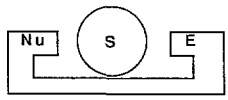


29

6. Future Directions

Perhaps the greatest advantage of the new systems is a consequence of their rigidity. This provides an opportunity to place both electrophilic and nucleophilic centers on a single substrate. This notion was first suggested by Swain [28] as ideal for concerted catalysis (Scheme 14), but flexible structures are unable to prevent intramolecular reactions of the catalytic acid-base pairs. We are pursuing these notions with suitable spacers and functional groups and will report on developments in due course.

Scheme 14



Substrate between nucleophilic and electrophilic centers

Acknowledgements

Financial support for this research was provided by the National Institutes of Health and the National Science Foundation. Year Laboratories supported work on metal ion chelation. I am most grateful for the outstanding contributions of my co-workers, whose names may be found in the literature citations.

References

1. J. Rebek, Jr., T. Costello, L. Marshall, R. Wattley, R. C. Gadwood and K. Onan: *J. Am. Chem. Soc.* **107**, 7481 (1985).
2. (a) J.-M. Lehn: *Science* **227**, 846 (1985); (b) D. J. Cram: *Ibid.* **219**, 1177 (1983); (c) J.-M. Lehn and C. Sirlin: *J. Chem. Soc., Chem. Commun.* 949 (1978).

3. D. J. Cram, P.Y.-S. Lam and S. P. Ho: *J. Am. Chem. Soc.* **108**, 839 (1986).
4. G. Trainor and R. Breslow: *ibid.* **103**, 154 (1981); V. T. D'Souza and M. L. Bender: *Acc. Chem. Res.* **20**, 146 (1987).
5. H. Stetter and E.-E. Roos: *Chem. Ber.* **88**, 1390, 1985; K. Odashima, A. Itai, Y. Iitaka, and K. Koga: *J. Am. Chem. Soc.* **102**, 2504 (1980); S. P. Miller, and H. W. Whitlock, Jr.: *ibid.* **106**, 1492 (1984); J. Winkler, E. Coutouli-Argyropopoulou, R. Leppkes and R. Breslow: *ibid.* **105**, 7198 (1983), F. Diederich and D. Griebel: *ibid.* **106**, 8037 (1984); C. D. Gutsche: *Acc. Chem. Res.* **16**, 161 (1983).
6. See for example, R. Breslow, P. Bouy and C. L. Hersh: *J. Am. Chem. Soc.* **102**, 2115 (1980) and references therein.
7. C. S. Wilcox, L. M. Greer and V. Lynch: *J. Am. Chem. Soc.* **109**, 1865 (1987); T. R. Kelly and M. P. Maguire: *ibid.* **109**, 6549 (1987).
8. J. Rebek, Jr.: *Science* **235**, 1478 (1987); J. Rebek, Jr., L. Marshall, R. Wolak, K. Parris, M. Killoran, B. Askew, D. Nemeth and N. Islam: *J. Am. Chem. Soc.* **107**, 7476 (1985).
9. D. S. Kemp and K. S. Petrakis: *J. Org. Chem.* **46**, 5140 (1981).
10. J. Rebek, Jr., B. Askew, M. Killoran, D. Nemeth and F.-T. Lin: *J. Am. Chem. Soc.* **109**, 2426 (1987).
11. J. S. Lindsey, I. C. Schreiman, H. C. Hsu, P. C. Kearney and A. M. Marguerattaz: *J. Org. Chem.* **52**, 827 (1987).
12. J. Rebek, Jr. and D. Nemeth: *J. Am. Chem. Soc.* **108**, 5637 (1986).
13. J. Rebek, Jr., D. Nemeth, P. Ballester and F.-T. Lin: *ibid.* **109**, 3474 (1987).
14. J. Rebek, Jr., B. Askew, D. Nemeth and K. Parris: *ibid.* **109**, 2432 (1987).
15. J. Rebek, Jr., B. Askew, P. Ballester and A. Costero: *ibid.* **109**, 6866 (1987).
16. J. Rebek, Jr., B. Askew, N. Islan, M. Killoran, D. Nemeth and R. Wolak: *ibid.* **107**, 6736 (1985).
17. J. Rebek, Jr., B. Askew, P. Ballester, C. Buhr, S. Jones, D. Nemeth and K. Williams: *ibid.* **109**, 5033 (1987). For another system using macrocyclic structures see A. D. Hamilton, and D. Van Engen: *ibid.* **109**, 5035 (1987).
18. J. Rebek, Jr., B. Askew, P. Ballester, C. Buhr, A. Costero, S. Jones and K. Williams: *ibid.* **109**, 6866 (1987).
19. R. Gandour: *Bioorg. Chem.* **10**, 169 (1981).
20. D. E. Ryono and G. M. Loudon: *J. Am. Chem. Soc.* **98**, 1889 (1976); G. M. Luodon and D. E. Ryono: *ibid.* **19**, 1900 (1976).
21. T. H. Fife and T. J. Przystas: *J. Am. Chem. Soc.* **102**, 292 (1980).
22. G. A. Rogers and T. C. Bruice: *ibid.* **96**, 2473 (1974).
23. J. Wolfe, D. Nemeth, A. Costero and J. Rebek: *ibid.* **110**, 983 (1988).
24. (a) C. G. Swain and J. F. Brown: *ibid.* **74**, 2534 (1952); (b) K. A. Engdahl, H. Birehed, O. Bohman, U. Obenius, and P. Ahlberg: *Chimica Scripa* **18**, 176 (1981); (c) P. R. Rony and R. O. Nef: *J. Am. Chem. Soc.* **95**, 2896 (1973) and earlier work by these authors. For other studies of the dissociation, see R. P. Bell, and J. P. H. Hirst: *J. Chem. Soc.* 1777 (1939), P. W. Wertz, J. C. Garver, and L. Anderson: *J. Am. Chem. Soc.* **103**, 3916 (1981) and references therein.
25. (a) E. J. Corey and R. A. Sneen: *J. Am. Chem. Soc.* **78**, 6279 (1956); (b) B. Capon, A. K. Siddhanta, and C. Zucco; *J. Org. Chem.* **50**, 3580 (1985); (c) B. Capon and A. K. Siddhanta; *ibid.* **49**, 255 (1984); (d) P. Deslongchamps: *Stereoelectronic Effects in Organic Chemistry* (Pergamon Press, Oxford) Chapter 2 (1983); E. S. Hand and W. P. Jencks: *J. Am. Chem. Soc.* **97**, 6221 (1975).
26. Reported at the 191st ACS meeting, New York City, 1986.
27. C. S. Craik, S. Roczniak, C. Largman and W. J. Rutter: *Science* **237**, 909 (1987).
28. C. G. Swain and J. F. Brown, Jr: *J. Am. Chem. Soc.* **74**, 2538 (1952).