

# THE DESIGN OF MOLECULAR HOSTS, GUESTS, AND THEIR COMPLEXES

Nobel Lecture, 8 December 1987

by

DONALD J. CRAM

Department of Chemistry and Biochemistry, University of California at Los Angeles, Los Angeles, California 90024, U.S.A.

## *Origins*

Few scientists acquainted with the chemistry of biological systems at the molecular level can avoid being inspired. Evolution has produced chemical compounds exquisitely organized to accomplish the most complicated and delicate of tasks. Many organic chemists viewing crystal structures of enzyme systems or nucleic acids and knowing the marvels of specificity of the immune systems must dream of designing and synthesizing simpler organic compounds that imitate working features of these naturally occurring compounds. We had that ambition in the late 1950's. At that time, we were investigating *pi*-complexes of the larger [m.n.]paracyclophanes with  $(\text{NC})_2\text{C}=\text{C}(\text{CN})_2$ , and envisioned structures in which the *pi*-acid was sandwiched by two benzene rings. Although no intercalated structures were observed [1,2], we recognized that investigations of *highly* structured complexes would be central to simulation of enzymes by relatively simple organic compounds.

In 1967, Pedersen's first papers appeared [3,4] which reported that alkali metal ions bind crown ethers to form highly structured complexes. We immediately recognized this work as an entree into a general field. The 1969 papers on the design, synthesis, and binding properties of the cryptands by J.-M. Lehn, J.-P. Sauvage, and B. Dietrich [5,6] further demonstrated the attractions and opportunities of complexation chemistry. Although we tried to interest graduate students in synthesizing *chiral crown ethers* from 1968 on, the efforts were unsuccessful. In 1970 we insisted that several postdoctoral co-workers enter the field. During 1973, we published five Communications on the subject [7,11]. In 1974 with Jane M. Cram, we published a general article entitled "Host-Guest Chemistry", which defined our approach to this research [12].

Aeschylus, the Athenian Poet-Dramatist, wrote 2 500 years ago, "Pleasantist of all ties is the tie of host and guest" [13]. Our research of the past 17 years has dealt with the pleasant tie between host and guest and the organic molecular level. The terms *host*, *guest*, *complex*, and their binding forces were defined in 1977 as follows [14]. "Complexes are composed of two or more molecules or

ions held together in unique structural relationships by electrostatic forces other than those of full covalent bonds . . . molecular complexes are usually held together by hydrogen bonding, by ion pairing, by *pi*-acid to *pi*-base interactions, by metal to ligand binding, by van der Waals attractive forces, by solvent reorganizing, and by partially made and broken covalent bonds (transition states) high structural organization is usually produced only through multiple binding sites a highly structured molecular complex is composed of at least one host and one guest component . . . a host-guest relationship involves a complementary stereoelectronic arrangement of binding sites in host and guest ... the host component is defined as an organic molecule or ion whose *binding sites converge* in the complex the guest component is defined as any molecule or ion whose *binding sites diverge* in the complex..." In these definitions, hosts are synthetic counterparts of the receptor sites of biological chemistry, and guests, the counterparts of substrates, inhibitors, or cofactors. These terms and concepts have gained broad international acceptance [15]. A new field requires new terms which, if properly defined, facilitate the reasoning by analogy on which research thrives.

From the beginning, we used Corey-Pauling-Koltun (CPK) molecular models [16], which served as a compass on an otherwise uncharted sea full of synthesizable target complexes. We have spent hundreds of hours building CPK models of potential complexes, and grading them for desirability as research targets. Hosts were then prepared by my co-workers to see if they possessed the anticipated guest-binding properties. Crystal structures of the hosts and their complexes were then determined to compare what was anticipated by model examination with what was experimentally observed. By the end of 1986, Drs. K. N. Trueblood, C. B. Knobler, E. F. Maverick, and I. Goldberg, working at UCLA, had determined the crystal structures of over 50 complexes, and those of another 25 hosts. These crystal structures turned our faith into confidence. Chart I traces the steps involved in linking the structures of *biotic complexes* of evolutionary chemistry with our *abiotic complexes* designed with the aid of CPK molecular models [17].

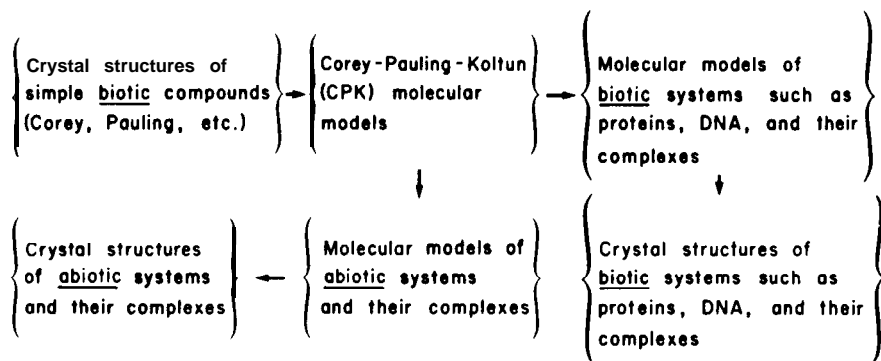


Chart I. Crystal structures of biotic compounds are correlated with those of abiotic compounds through CPK models.

In molecular modeling, we made extensive use of the self-evident principle of complementarity: "to complex, hosts must have binding sites which cooperatively contact and attract binding sites of guests without generating strong nonbonded repulsions" [18]. Complexes were visualized as having three types of common shapes: 1) perching complexes, resembling a bird perching on a limb, an egg protruding from an egg cup, or a scoop of ice cream sitting on a cone; 2) nesting complexes, similar to an egg resting in a nest, a baby lying in its cradle, or a sword sheathed in its scabbard; 3) capsular complexes, not unlike a nut in its shell, a bean in its pod, or a larva in its cocoon. Chart II provides a comparison of CPK models of the three types of complexes (1, 2, and 3) and their actual crystal structures [19,20].

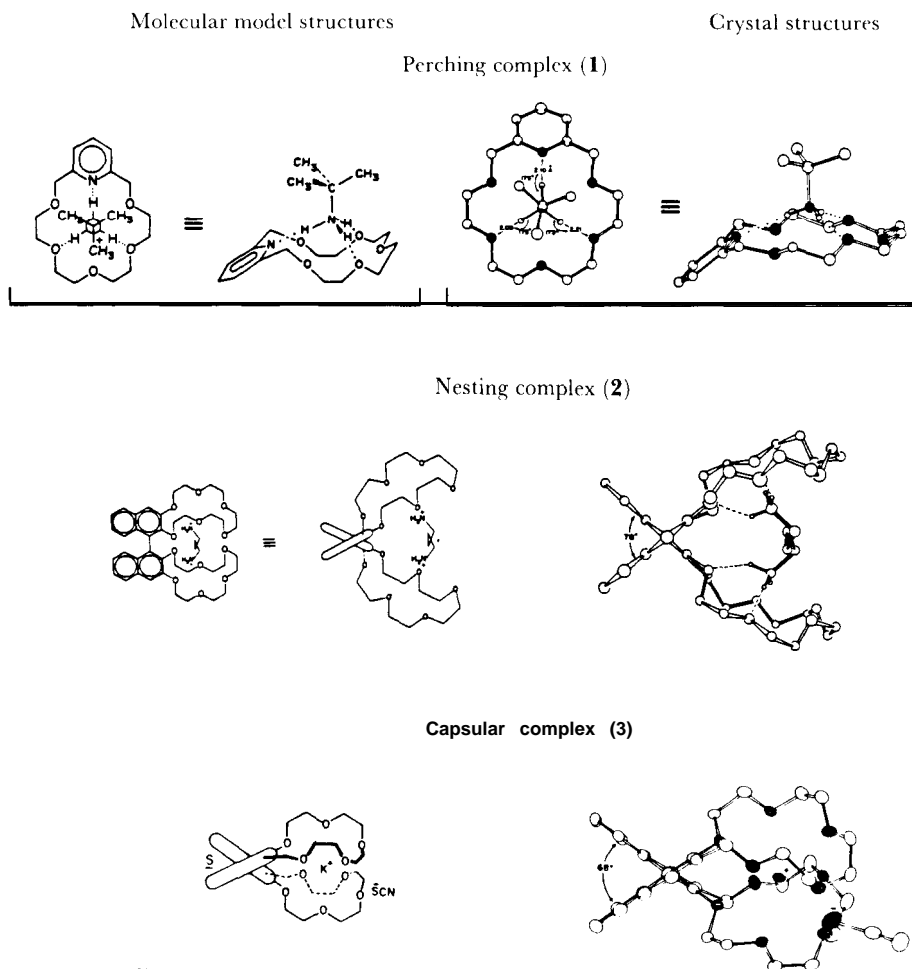
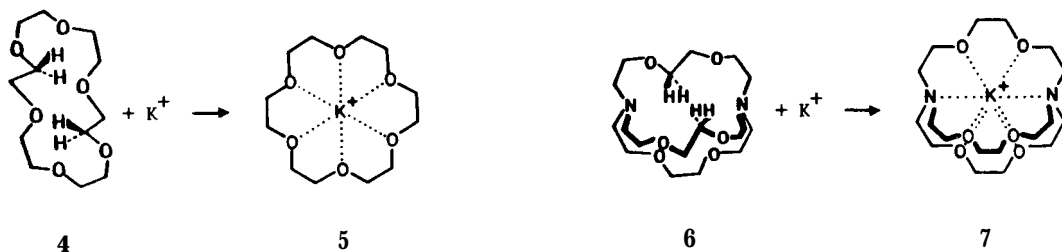


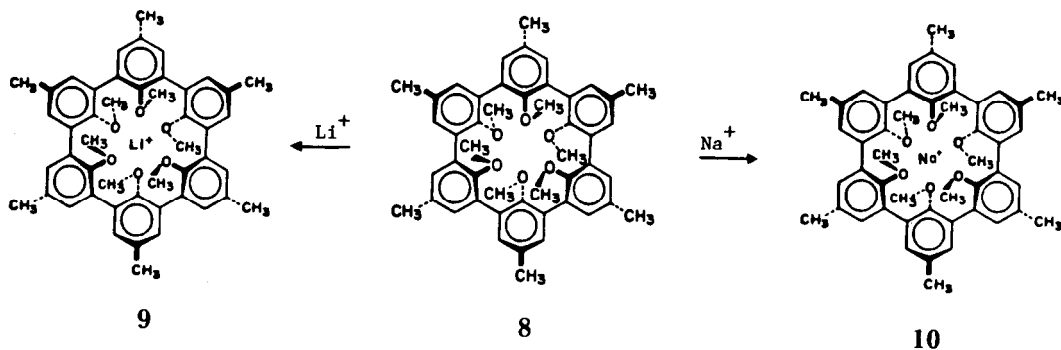
Chart II. Three types of complexes.

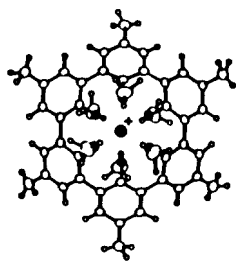
### Principle of Preorganization

Crystal structures of Pedersen's 18-crown-6 [21] and Lehn's [2.2.2] cryptand [22,23] show that in their uncomplexed states, they contain neither cavities nor convergently-arranged binding sites. Comparisons of the crystal structure of host **4** with that of its  $K^+$  complex **5**, and of host **6** with that of its  $K^+$  complex **7** indicate that the complexing act must be accompanied by host reorganization and desolvation.

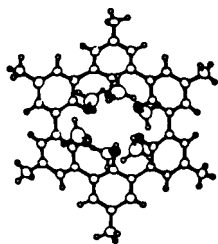


With the help of CPK molecular models, we designed ligand system **8**, whose oxygens have no choice but to be octahedrally arranged around an enforced spherical cavity complementary to  $Li^+$  and  $Na^+$  ions. We have given the family name, *spherand*, to completely preorganized ligand systems, and the name, *spheraplex*, to their complexes, which like **7**, are capsular [24]. The syntheses and crystal structures of **8**, **9** and **10**, have been reported [25]. As expected, the crystal structure of **11** contains a hole lined with 24 electrons, which are shielded from solvation by six methyl groups. The snowflake-like structures of **11** and of spheraplexes **12** and **13** are nearly identical. Thus **8** is the first ligand system to be designed and synthesized which was completely organized for complexation during synthesis, rather than during complexation.

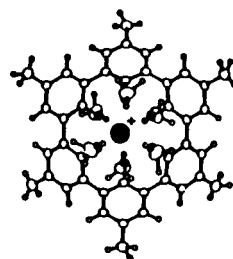




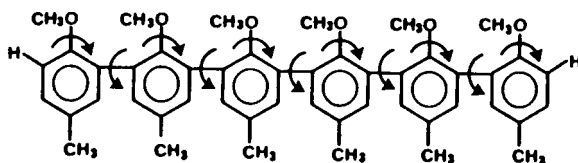
12



11

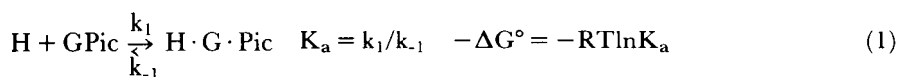


13



14

A method was developed of determining the binding free energies of lipophilic hosts toward guest picrate salts of  $\text{Li}^+$ ,  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Rb}^+$ ,  $\text{Cs}^+$ ,  $\text{NH}_4^+$ ,  $\text{C}_2\text{H}_5\text{NH}_3^+$ , and  $t\text{-BuNH}_3^+$ . The guest salts were distributed between  $\text{CDCl}_3$  and  $\text{D}_2\text{O}$  at  $25^\circ\text{C}$  in the presence and absence of host. From the results,  $K_a$  ( $\text{mol}^{-1}$ ) and  $-\Delta G^\circ$  values ( $\text{kcal mol}^{-1}$ ) were calculated (equations (1)). This method was rapid and convenient for obtaining



$-\Delta G^\circ$  values at  $25^\circ\text{C}$  ranging from about 6 to 16  $\text{kcal mol}^{-1}$  in  $\text{CDCl}_3$  saturated with  $\text{D}_2\text{O}$  [26]. Higher values (up to 22  $\text{kcal mol}^{-1}$ ) were obtained by equilibration experiments between complexes of known and those of unknown  $-\Delta G^\circ$  values [18, 27, 28]. Others were determined from measured  $k_{-1}$  and  $k_1$  values, all in the same medium at  $25^\circ\text{C}$  [18]. Spherand **8** binds  $\text{LiPic}$  with  $>23$   $\text{kcal mol}^{-1}$ ,  $\text{NaPic}$  with 19.3  $\text{kcal mol}^{-1}$ , and totally rejects the other standard ions, as well as a wide variety of other di- and trivalent ions [18]. The open-chain counterpart of **8**, podand **14**, binds  $\text{LiPic}$  and  $\text{NaPic}$  with  $-\Delta G^\circ < 6$   $\text{kcal mol}^{-1}$  [29]. Podand is the family name given to acyclic hosts [15].

Podand **14** differs constitutionally from spherand **8** only in the sense that **14** contains two hydrogen atoms in place of one Ar-Ar bond in **8**. The two hosts differ radically in their conformational structures and states of solvation. The spherand possesses a single conformation ideally arranged for binding  $\text{Li}^+$  and  $\text{Na}^+$ . Its oxygens are deeply buried within a hydrocarbon shell. The orbitals of their unshared electron pairs are in a microenvironment whose dielectric properties are between those of a vacuum and of a hydrocarbon. No solvent can approach these six oxygens, which remain unsolvated. The free energy costs of

organizing the spherand into a single conformation and of desolvating its six oxygens were paid for during its synthesis. Thus spherand **8** is preorganized for binding [30]. The podand in principle can exist in over 1 000 conformations, only two of which can bind metal ions octahedrally. The free energy for organizing the podand into a binding conformation and desolvating its six oxygens must come out of its complexation free energy. Thus the podand is not preorganized for binding, but is randomized to maximize the entropy of mixing of its conformers, and to maximize the attractions between solvent and its molecular parts.

The difference in  $-\Delta G^\circ$  values for spherand **8** and podand **14** binding  $\text{Li}^+$  is  $>17 \text{ kcal mol}^{-1}$ , corresponding to a difference in  $K_a$  of a factor of  $>10^{12}$ . The difference in  $-\Delta G^\circ$  values for **8** and **14** binding  $\text{Na}^+$  is  $>13 \text{ kcal mol}^{-1}$ , corresponding to a difference in  $K_a$  of a factor of  $>10^{10}$ . These differences are dramatically larger than any we have encountered that are associated with other effects on binding power toward alkali metal ion guests. We conclude that *preorganization is a central determinant of binding power*. We formalized this conclusion in terms of what we call the principle of preorganization, which states that "the more highly hosts and guests are organized for binding and low solvation prior to their complexation, the more stable will be their complexes." Both enthalpic and entropic components are involved in preorganization, since solvation contains both components [29]. Furthermore, binding conformations are sometimes enthalpically rich. For example, the benzene rings in spherand **8** and spheraplexes **9** and **10** are somewhat folded from their normal planar structures to accommodate the spacial requirements of the six methoxyl groups [30]. The anisyl group is an intrinsically poor ligand [31, 32]. That **8** is such a strong binder provides an extreme example of the power of preorganization.

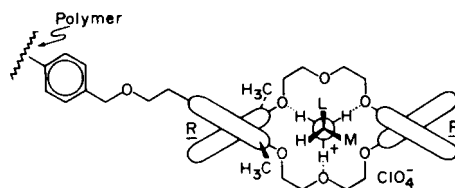
Families of hosts generally fall into the order of their listing in Chart III when arranged according to their  $-\Delta G^\circ$  values with which they bind their most complementary guests: spherands  $>$  cryptaspherands  $>$  cryptands  $>$  hemispherands  $>$  corands  $>$  podands. Corand is the family name given to modified crown ethers [33]. Spheraplex **8**  $\cdot \text{Li}^+$  provides a  $-\Delta G^\circ$  value, of  $>23 \text{ kcal mol}^{-1}$ . Cryptaspheraplexes **15**  $\cdot \text{Na}^+$ , **16**  $\cdot \text{Na}^+$ , and **17**  $\cdot \text{Cs}^+$  [34] give values of 20.6, 21.0, and 21.7  $\text{ kcal mol}^{-1}$ , respectively [27]. Cryptaplexes **18**  $\cdot \text{Li}^+$ , **19**  $\cdot \text{Na}^+$ , and **6**  $\cdot \text{K}^+$  give respective values of 16.6, 17.7, and 18.0  $\text{ kcal mol}^{-1}$  [27]. Hemispheraplexes **20**  $\cdot \text{Na}^+$ , **21**  $\cdot \text{Na}^+$ , and **22**  $\cdot \text{K}^+$  are bound by 12.2, 13.5, and 11.6  $\text{ kcal mol}^{-1}$  [35, 36]. Coraplex **23**  $\cdot \text{K}^+$  has a  $-\Delta G^\circ$  value of 11.4 [26, 31] and podaplexes **14**  $\cdot \text{M}^+$  values of  $<6 \text{ kcal mol}^{-1}$  [29]. Although the numbers of binding sites and their characters certainly influence these values, the degree of preorganization appears to be dominant in providing this order.











27

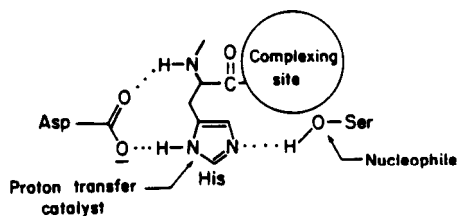
active site. This material (the host part of **27**) was used to give complete enantiomeric resolution of several amino acid salts. The behavior in the chromatographic resolution paralleled that observed in the extraction and transport experiments, and was useful both analytically and preparatively. Separation factors ranged from 26 to 1.4, the complexes of the (*R,R*)-(*D*)- or (*S,S*)-(*L*)-configurations always being the more stable. The structure envisioned for the more stable complex is formulated in **27** [41].

#### Partial Transacylase Mimics

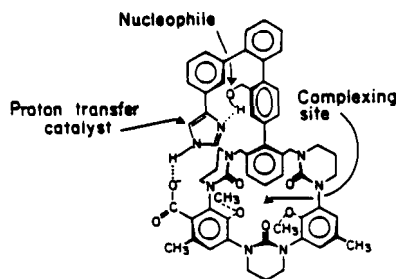
The design and synthesis of enzyme-mimicking host compounds remains one of the most challenging and stimulating problems of organic chemistry. We chose to examine transacylase mimics first because the mechanism of action of these enzymes had been so thoroughly studied.

The active site of chymotrypsin combines a binding site, a nucleophilic hydroxyl, an imidazole, and a carboxyl group in an array preorganized largely by hydrogen bonds as indicated in **28**. With the help of molecular models, we designed **29** as an "ultimate target" host possessing roughly the same organization of groups as that of **28**.

Compound **29** is much too complicated to synthesize without getting encouragement from simpler model compounds. An incremental approach to **29** was employed. We first prepared **30**, and found that it binds *t*-BuNH<sub>3</sub>Pic in CDCl<sub>3</sub> saturated with D<sub>2</sub>O with  $-\Delta G^\circ = 13.2 \text{ kcal mol}^{-1}$ . The complex, **30**·(CH<sub>3</sub>)<sub>3</sub>CNH<sub>3</sub><sup>+</sup>, had the expected crystal structure [42]. Accordingly, **31** was prepared, and found to bind CH<sub>3</sub>NH<sub>3</sub>Pic and NaPic under our standard



28



29





demonstrates that complexation of the transition states for transacylation can greatly stabilize those transition states to produce large rate factor increases over comparable noncomplexed transition states [46]. Others have shown that the imidazole of chymotrypsin is acylated first by esters of nonspecific substrates [47].

These investigations demonstrate that totally synthetic systems can be designed and prepared which mimic the following properties of enzymes: the ability to use complexation to vastly enhance reaction rates and the sensitivity to competitive inhibition. In a different, chiral system, we demonstrated that a synthetic host was capable of distinguishing between enantiomeric reactants [48, 49]. We anticipate that as the field matures, many of the other remarkable properties of enzyme systems will be observed in designed, synthetic systems. Our results illustrate some of the strategies and methods that might be applied in this expanding field of research.

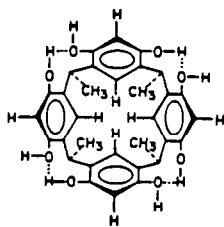
#### *Cavitands-Synthetic Molecular Vessels*

Although enforced cavities of molecular dimensions are frequently encountered in enzyme systems, RNA, or DNA, they are almost unknown among the seven million synthetic organic compounds. In biological chemistry such cavities play the important role of providing concave surfaces to which are attached convergent functional groups which bind substrates and catalyze their reactions. If synthetic biomimetic systems are to be designed and investigated, simple means must be found of synthesizing compounds containing enforced concave surfaces of dimensions large enough to embrace simple molecules or ions. We applied the name *cavitand* to this class of compound [50].

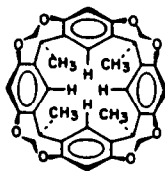
Cavitands designed and studied include compounds **42-45**, many of which were prepared from **41**. The structure and conformational mobility of **41** had been established by A. G. S. Högberg [51]. The substance is prepared in good yield by treatment of resorcinol with acetaldehyde and acid. We rigidified **41** and its derivatives by closing four additional rings to produce **42-45** [50, 52].

As anticipated by molecular model examinations, **42-45** crystallize only as solvates because these rigid molecules taken alone are incapable of filling their voids either intermolecularly or intramolecularly. They are shaped like bowls of differing depth supported on four methyl "feet." Compound **42** forms crystallates with  $\text{SO}_2$ ,  $\text{CH}_3\text{CN}$ , and  $\text{CH}_2\text{Cl}_2$ , molecules to which it is complementary (molecular model examination). Cavitand **43**, whose cavity is deeper, crystallizes with a mole of  $\text{CHCl}_3$ . Crystal structures of **42-CH<sub>2</sub>Cl<sub>2</sub>** and **43-CHCl<sub>3</sub>** show they are caviplexes, as predicted [53]. Cavitand **44** is vase-shaped. It crystallizes with one mole of  $(\text{CH}_3)_2\text{NCHO}$ , which is just small enough to fit into the interior of **44** in models. Although the amide cannot be removed at high temperature and low pressure, it is easily displaced with  $\text{CHCl}_3$ , one and one-half moles of which appear to take the place of the  $(\text{CH}_3)_2\text{NCHO}$  in the crystallate [50].

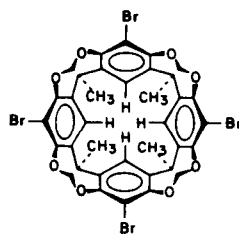
Treatment of octal **41** with  $\text{R}_3\text{SiCl}_2$  gave a series of cavitands, of which **45** is typical. In molecular models, **45** has a well-shaped cavity, defined by the bottoms of four aryls and by four inward-turned methyl groups. In molecular



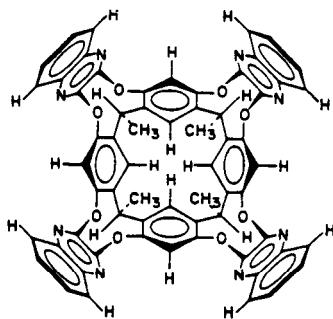
41



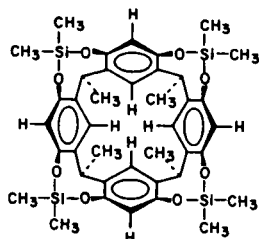
42



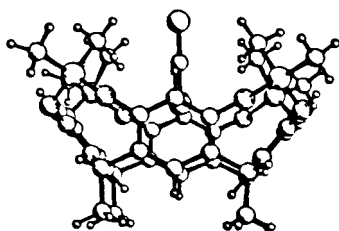
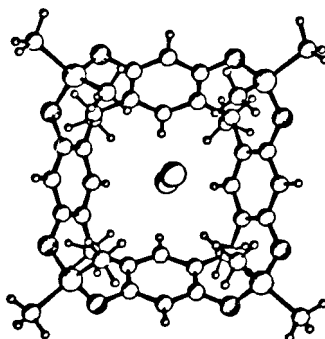
43



44



45

45 · CS<sub>2</sub> (side view)45 · CS<sub>2</sub> (top view)

models, this well is complementary to small, cylindrical molecules such as  $S=C=S$ ,  $CH_3C=CH$ , and  $O=O$ , but not to larger compounds such as  $CDCl_3$  or  $C_6D_6$ . Cavitan **45** and its analogues when dissolved in  $CDCl_3$  or  $C_6D_6$  complex guests such as those mentioned above, whose external surfaces are complementary to the internal surface of the host cavity. Association constants were determined for **45** and its analogues binding  $S=C=S$ . Values of  $-\Delta G^\circ$  as high as  $2 \text{ kcal mol}^{-1}$  have been observed. A crystal structure of **45**· $CS_2$  shows that  $CS_2$  occupies the well in the expected manner. Compound **45** in  $CDCl_3$  was also shown to bind dioxygen reversibly [52]. Dissolution of **45** in solvents such as  $CDCl_3$  or  $C_6D_6$  is the equivalent of dissolving "holes" in a medium into

which appropriately shaped solutes fall. The discrimination shown by the holes for the guests exemplifies the principle of complementarity as applied to cavitand complexation.

The next steps in research on these cavitands is to append to them water-solubilizing and catalytic groups. The former will provide them with hydrophobic driving forces to complex nonpolar guests, and the latter to catalyze reactions of such guests.

### *Carcerands-Synthetic Molecular Cells*

Absent among the millions of organic compounds hitherto reported are closed-surface hosts with enforced interiors large enough to imprison behind covalent bars, guests the size of ordinary solvent molecules. After much thought and molecular model examination, we chose **48** as the target for synthesis of the first molecular cell. The term *carcerand* was applied to this class of compound. The synthesis involved treating  $\text{Cs}_2\text{C O}_3$  with a solution in  $(\text{CH}_3)_2\text{NCHO}-(\text{CH}_2)_4\text{O}$  of equal molar amounts of cavitands **46** and **47** under an atmosphere of argon. The first question to be answered was: what guest compounds would be trapped inside during the shell closure? This question is akin to asking whether two soup bowls closed rim-to-rim under the surface of a kettle of stew would net any stew. The answer was that **48** "contained" essentially every kind of component of the medium present during ring closure [54].

The product (**48** and guests) was very insoluble in all media, and was purified by extracting it with the most powerful solvents of each type. The remaining material was subjected to elemental analysis for C, H, S, O, N, Cl, and Cs. Nitrogen analysis and an IR spectrum of the substance revealed that  $(\text{CH}_3)_2\text{NCHO}$  had been entrapped. The presence of equivalent amounts of Cs and Cl demonstrated that one or the other ion or both had to be encapsulated in the host.

A fast atom bombardment mass spectrum of **48**·G showed the presence of the following host-guest combinations, the species trapped in the interior of **48** being enclosed by parentheses:

**48**·no guest; **48**· $(\text{Cs}^+) \cdot \text{Cl}^-$ ; **48**· $((\text{CH}_3)_2\text{NCHO})$ ; **48**· $(\text{Cs}^+ + \text{H}_2\text{O}) \cdot \text{Cl}^-$ ;  
**48**· $((\text{CH}_2)_4\text{O} + \text{H}_2\text{O})$ ; **48**· $((\text{CH}_3)_2\text{NCHO} + \text{Cs}^+) \cdot \text{Cl}^-$ ; **48**· $(\text{Cs}^+ + \text{Ar}) \cdot \text{Cl}^-$ ;  
**48**· $(\text{Cs}^+ + \text{H}_2\text{O} + \text{Cs}^+) \cdot \text{Cl}_2^-$ ; **48**· $(\text{Cs}^+ + \text{Cl}^-)$ , and **48**· $(\text{Cs}^+ + \text{Cs}^+ + \text{Cl}^-) \cdot \text{Cl}^-$ .

