

# Molecular structure as a blueprint for supramolecular structure chemistry in confined spaces

Nicholas J. Turro\*

Department of Chemistry, Columbia University, 3000 Broadway, MC 3119, New York, NY 10027

Edited by Jack Halpern, University of Chicago, Chicago, IL, and approved June 20, 2005 (received for review March 17, 2005)

Chemistry is an ever expanding “universe” at the microscopic level requiring mastery of the invisible at ever increasing levels of complexity. Starting with atoms, assemblies of a single nucleus and orbiting electrons, the chemist has learned to understand atomic structure through mastery of the principles of electronic configurations. Proceeding to molecules, assemblies of two or more atoms, the chemist has learned to understand molecular structure through mastery of the principles of the covalent bond. It is quite natural that the next level of complexity will be supermolecules, assemblies of two or more molecules, whose structure will be understood through mastery of the principles of the intermolecular bond, which are at the heart of supramolecular chemistry. We shall see that the principles of weak intermolecular “bonds,” while perhaps a bit nonintuitive, obeys the same rules of chemistry as the strong intramolecular covalent bonds.

Chemists (and nature) have used <100 atoms on which they have designed the assembly of millions of molecules by using the concept of the covalent bond between atoms. The number of supramolecular systems that can be assembled is stunning and infinite. A holy grail of supramolecular chemistry is to put the “syntheses” of supramolecular systems based on intermolecular bonds on the same level of understanding as the current state of the syntheses of molecular systems based on covalent bonds.

## The Supramolecular Paradigm

Chemists have grown so familiar with the paradigm of molecular structure that its validity and centrality to all of chemistry are no longer seriously debated. However, chemists are now seeking the development of a paradigm at a level beyond that of molecular structure, namely at the level of supramolecular level. Supramolecular structures are related to molecules in the same way that molecular structures are related to atoms. Although atoms play a critical role in the paradigms of chemistry, it is molecular structure that is generally at the center of the chemistry of many important technological and biological materials. The “intellectual glue” of molecular structure is the concept of the covalent

bond and the connectivity of atoms and the stereochemistry of atoms in space that are associated with covalent bonds. The paradigm of the covalent chemical bond provides the rules governing the structures, dynamics, physical characteristics, and chemical transformations of molecules. We therefore expect that an effective paradigm of noncovalent, intermolecular bonds will provide the rules governing the structures, dynamics, characteristics, and transformations of supramolecular assemblies. The development of such a paradigm is an activity that has received considerable attention over the past two decades and will continue to be of increasing importance.

Just as the level of atomic structure is inadequate for understanding aspects of chemistry where molecular aspects dominate, the level of molecular structure is inadequate for understanding aspects of chemistry where supramolecular aspects dominate. The emerging paradigm of supramolecular systems emphasizes the chemistry of the intermolecular bond that binds molecules together into assemblies we can term supermolecules, rather than the covalent bond, which binds atoms together into assemblies we term molecules. Noncovalent intermolecular bonds are more varied and complex than covalent intramolecular bonds. In general, intermolecular, noncovalent bonds are held together by much weaker forces than intramolecular, covalent bonds. For example, the important forces holding molecules together may consist of a mixture of dispersion forces, hydrogen bonds, hydrophobic bonds, etc. Although these forces, per atom contact, are small, in many supramolecular systems many such weak bonds are formed in a cooperative manner, thereby causing the supramolecular structure to be energetically stable; however, because each “bonding contact” is relatively weak, the bonds may be broken relatively easily so that the structure may have a certain amount of flexibility and over time the molecular connections and stereochemical relationships will vary. Thus, a common feature of supramolecular systems is a “supervalency” of supramolecular structures, in which a supramolecular complex of many molecules is rendered stable by the summation of a large number of weak intermolecular

bonds. Micelles and the double helix of DNA are examples of supervalent supramolecular systems.

In one of its most important forms, supramolecular chemistry is concerned with the structure and dynamics of a small molecule (termed a guest) that is noncovalently bound to a larger molecule (termed a host). The symbol guest@host has its greatest chemical informational value when the distinction between the guest and host is clear. In many guest@host complexes, the guest is a molecule that is relatively small compared with the host and the host may either be a single large molecule or an assembly of molecules that behave as a unit. Typically, the host provides a cage that completely surrounds the guest (guest@cage) or the host provides a cavity (guest@cavity) that partially surrounds the guest. However, there are also examples for which the notion of guest and host are blurred. For example, in a micelle composed of surfactant monomers, a single monomer may be considered as a guest in a host in its own supramolecular assembly. In turn, a molecule that is not a monomer surfactant is a clear-cut guest in the micelle assembly of surfactants.

A useful term for the formation of an intermolecular bond between guest and host is “complexation,” a term that preserves the idea of chemically relevant “looseness” expected of noncovalent bonds between molecules. In this regard, the host is the partner of the guest@host complex whose property design and structural variations determine the types of guests that will be bound in the complex. The other partner, the complexed guest, is usually the partner whose chemistry or physical properties are being investigated. These concepts are derived from the paradigms of enzyme chemistry and provide a familiar and useful working basis for the discussion of supramolecular chemistry. A guest@host complex may be considered as a supermolecule or supramolecular assembly depending on the

This paper was submitted directly (Track II) to the PNAS office.

\*E-mail: njt3@columbia.edu.

© 2005 by The National Academy of Sciences of the USA

complexity of the supramolecular structure under discussion.

### Supramolecular Chemistry: Structure and Dynamics at the Topological, Geometric, and Chemical Level

Supramolecular chemistry has been defined (1, 2) in various ways such as the chemistry of molecular assemblies and the intermolecular bond, chemistry beyond the molecule, and chemistry of the noncovalent bond. Each of these definitions is useful as are others, but each has limitations and exceptions when the definitions are pushed to specific examples. Indeed, supramolecular chemistry can be defined to encompass the chemistry of molecular assemblies from a molecule in a molecular solvent cage to the constellation of molecular assemblies (composed of proteins, lipids, DNA, RNA, etc.) that constitute the enormous chemical complexity of a living cell.

Because the bonding between molecules in guest@host complexes is often a mixture of many weak electrostatic and dispersion interactions, it is difficult to define and classify the nature of the noncovalent bonds precisely. As a starting point for discussing supramolecular assemblies it is convenient to consider the “neighborhood” relationships between certain atoms of the guest with those of the host that can control the physical and chemical properties of the guest@host complex. Often it is the properties of the guest that are most significantly modified by the neighborhood relations imposed by binding to the host that are of interest, but the combined supramolecular properties of the complex and modification of the host structure are also of considerable interest and importance.

By neighborhood relationships in supramolecular chemistry we mean that two molecules are in the proximity of one another during a certain time period. During this time period the two molecules may be considered to be bonded irrespective of the nature of the bonding and the reason the atoms are close to one another. Thus, the idea of neighborhood relationships allows the acceptance of any type of interactions that keep molecules in each other's neighborhood. For example, a molecule that is contained as a guest inside a host fullerene has a clear neighborhood relationship to the internal cage of the fullerene. Similarly, a molecule that is contained as a guest in the cavity of a host such as a cyclodextrin or a cavitand has a clear neighborhood relationship to the cavity of the host. A molecule that is contained as a guest in a crystal and is surrounded by molecules of a crystal-



Fig. 1. Topological representations of three common supramolecular motifs: guest@cage (G@cage) (a), guest@cavity (G@cavity) (b), and guest plus cavity (G + cavity) (c).

line host has a clear neighborhood relationship to the surrounding molecules of the crystal. Finally, a guest molecule that is intercalated in a host DNA double-helix has a chemical relationship to a small set of specific bases that are in its neighborhood.

### A Topological Approach to Classification of Supramolecular Structures

Topology (3) is a field of mathematics that deals with neighborhood relationships. It is therefore natural to use the concepts of topology as a very general starting point to describe the structure of supramolecular systems.

Fig. 1 shows a schematic representation of an approach to thinking about supramolecular systems that emphasizes the topological or neighborhood relationships of simple guest@host supramolecular structures for three common motifs. Fig. 1a is a guest@cage motif in which the guest is completely surrounded by the neighborhood of the host; Fig. 1b is a guest@cavity motif in which the guest is only partially surrounded by the neighborhood of the host; and Fig. 1c is a guest plus host motif in which the guest and host are separated and are not in each other's neighborhoods.

In Fig. 1a for the guest@cage complex, the structure of the host is represented by a closed continuous boundary (shown as a circle in two dimensions, but easily extended to a sphere in three dimensions), which separates the internal space of the host from the external space outside of the host. The specifics of the geometric and chemical nature of the host is left to the imagination of the supramolecular chemist. At the geometric level any 3D space with a closed boundary, irregardless of its shape, is topologically equivalent to the surface of a sphere (3). Thus, a spherical shape contains all of the topological information that is available so it can be used as an exemplar for a host that serves as a cage. At the chemical level the space corresponding to the topological sphere may be as simple as the “soft” solvent cage formed around a guest provided by an ordinary solvent such as benzene, or as the soft supercage that is the internal hydrophobic core of a micelle made by aggregation of surfactants in aqueous solution. Other host spaces may be the

“hard” cages provided by the internal surface of a porous solid such as porous silica or zeolites, or by the internal cavity of a fullerene or the periodic array of molecules about a selected guest molecule in a crystalline solid.

In Fig. 1b for the guest@cavity complex, the host does not completely surround the guest and there is a path for the guest to leave the neighborhood of the host and exit into the space outside the host. In chemical terms the cavity that surrounds the guest may be free volume produced by special displacement of the host molecules (solvent, micelle) or the inherent void space of a host (porous solid, fullerene, crystalline solid, DNA double-helix fluctuations).

The dynamics of binding and dissociation of guest@host complexes is of critical importance to supramolecular chemistry because both the equilibrium properties of the bound complex and the time scale for which the guest is bound to the host are associated with these dynamics. Fig. 1 provides a topological representation of the dynamics of a generalized bound guest@host complex and the unbound guest plus host components. The cage completely surrounding the guest (Fig. 1a) may be in equilibrium with a cavity that only partially surrounds the guest (Fig. 1b). The guest@cavity complex may be in equilibrium with guest plus cavity free components (Fig. 1c).

In moving from topological representations to a geometric representation of the guest@host in 3D space, metric issues such as the dimensions of the guest and host become important. If the size of the guest is of the order of the size of the host cavity, then the guest and host are more or less in constant significant contact. If the size of the host cavity is very large relative to the size of the guest, then the surface of the cavity will appear to be flat to the guest and the contact cross section of the guest to the host may be small. The latter corresponds to an important situation of the absorption of molecules on flat surfaces, but is out of the scope of this Perspective. We will be concerned only with systems for which the size of the host cavity is of the same order as the size of the guest molecule.

Fig. 1a and b becomes related if the closed structure of Fig. 1a can open its

**Table 1. Topological and geometric features of molecular and supramolecular structure**

Structural feature	Mathematical level	Molecular structure	Supramolecular structure
Composition	Topological	Numbers and kinds of atoms in a molecular structure	Numbers and kinds of molecules in a supramolecular structure
Constitution	Topological	Connections between the atoms in a molecular composition	Connections between the molecules in a supramolecular composition
Configuration	Geometric	Orientation of neighboring atoms about a given atom in 3D space	Orientation of neighboring molecules about a given molecule in 3D space
Conformation	Geometric	Instantaneous global molecular shape of a structure in 3D space	Instantaneous global shape of a supramolecular structure in 3D space

boundary or if the open structure of Fig. 1*b* can close its structure. In general, the issue of whether a guest@host complex is closed or open will depend on time scales, with the guest having a certain probability of dissociating from the host and entering the space outside the boundary of the host. As one moves to shorter and shorter time scales, even soft solvent cages become hard with respect to the ability of the guest to move or to escape from the cage. For example, for physical or chemical events that occur on the time scale of  $10^{-13}$  s, an ordinary solvent cage is “frozen” and a molecule in a solvent cage may be best represented by guest@cage, because the walls of the solvent completely surround the guest during this short time period. However, on longer time scales, the solvent molecules separate to produce free volume through which the guest can escape to the space outside of the solvent cage.

### Using the Tools of Molecular Structure as a Blueprint for Supramolecular Structure

The paradigm of molecular structure is at the heart of virtually all current chemical thinking. Chemists have built their ideas of molecular structure first on topology and then on 3D geometry, two purely intellectual fields of mathematics (Table 1). Chemical ideas (atoms) can be introduced at either the topological or geometric level. Molecular composition is defined as the number and kinds of atoms that make up a molecule and provides the most fundamental level of information about a molecular structure. At the level of composition only the number and kinds

of “elements” or atoms of the structure are of chemical interest.

The next highest level of molecular structure is molecular constitution, which is concerned with the way the atoms of a given composition are connected to one another. At the level of constitution, not only is the composition of chemical interest, but also the number of possible ways in which the atoms can be connected to one another following the usual rules of valence. A connection may be viewed mathematically as a neighborhood relationship; a connection or a bond between atoms means that if we know where a selected atom is in space we are guaranteed that bonded atoms will be in the neighborhood of the selected atom. A unique molecular structure at this level may be identified as one for which all of the atoms of the composition possess a unique connectivity or neighborhood relationship.

The configuration of a molecular constitution is the way the atoms of the molecular structure are oriented in 3D space, in particular the orientation of neighboring bonded atoms about a given atom. At the level of configuration, in addition to the composition and constitution, the number of possible dispositions in 3D space of neighboring atoms about a central atom is of chemical interest and is termed molecular stereochemistry. Finally, the highest level of (static) molecular structure is that of conformation that is defined as the instantaneous shape of a molecular structure in 3D that results from a particular assembly of configurations. At the level of conformation, the possible shapes of a given composition, constitution, and specific set of configurations are of chemical interest. A unique structure at

this level may be identified as one that differs from others in terms of shapes resulting from rotations about single bonds.

Each of these molecular structural concepts of composition, constitution, configuration, and conformation at the topological/geometric level may be immediately and directly imaged onto supramolecular structures (Table 1).

Molecular structures possess a time dependence in which one or more of the fundamental topological or geometric properties change with time. These time-dependent changes are termed molecular dynamics. The information implied in the dynamics of molecular structures can be transferred to supramolecular systems. For example, for a given molecular composition, different constitutions can exist, leading to molecular isomers that differ because of the ways that atoms are connected. Similarly, for a given supramolecular composition, different ways of connecting the molecules are generally possible, resulting in supramolecular constitutional isomers. As another example for a given molecular constitution, different configurations can exist, so that we immediately deduce that for a given supramolecular constitution supramolecular stereoisomers are possible. From this brief comparison of the structural features of molecular and supramolecular structure, we can expect that there maybe a strict one-to-one mapping of the structural features of molecular structure onto the structural features of supramolecular structures. Although the level of complexity is much greater for supramolecular systems, the mapping process provides a means of guiding expectations and classifications and eventually mastering the rules of noncovalent bonding in determining supramolecular structure to a level comparable to the current mastery of the rules of covalent bonding in determining molecular structure.

### From the Solvent Cage to the Supramolecular Cage

To put some concrete chemistry into the relatively abstract discussion of supramolecular guest@host systems discussed so far, we consider the example of the solvent cage that is perhaps the simplest example of a supramolecular system. In this case the molecule is a guest in a cage of solvent molecules, i.e., a guest@solvent\_cage complex. The concept of solvent molecules surrounding a pair of reacting radicals was used in the interpretation of the lowering of the quantum yield for photochemical production of radicals upon going from the gas to the liquid phase (4). Excitation of certain molecules possessing “continu-

ous absorption” in the gas phase was known to cause dissociation to occur with a quantum yield of unity, i.e., each absorbed photon causes the dissociation of the molecule into a pair of radicals. However, photolysis of the same molecules in solution resulted in quantum yields of much less than unity. Franck and Rabinovitch (4) provided an explanation of why the quantum yield of free radicals,  $\Phi_{FR}$ , resulting from photodissociation of a molecule, is lower in liquids than in the gas phase. They proposed that in a liquid the radical pair resulting from photodissociation is born in a solvent cage, which will encourage “primary recombination” of the radical pair before the individual radicals of the pair can separate and escape from the solvent cage to become free radicals (which are easily trapped and can initiate polymerization). The lowering of the quantum yield upon going from the gas to the liquid phase has become known as the “cage effect” and is now commonly attributed to any chemical or physical effect that can be attributed to a radical<sub>pair</sub>@solvent<sub>cage</sub> complex. The Franck–Rabinovitch (4) explanation of the cage effect is an example of a proposal of a supramolecular structure to explain a chemical phenomenon.

The solvent cage, rather than the gas phase, is an appealing benchmark for comparison with other supramolecular systems. When the behavior of a cage of a supramolecular system deviates significantly from that expected from the benchmark solvent cage, we can term the supramolecular system as a “super cage.” In a more general way, when experiments produce results that cannot be understood on the basis of molecular structure alone one should look for a supramolecular explanation, i.e., the environment must be explicitly considered when trying to understand the results.

To calibrate expectations of the behavior of a solvent cage, let’s examine the Franck–Rabinovitch (4) hypothesis of a cage effect and its influence on bimolecular chemical reactions in a bit more detail. An important feature of bimolecular chemical reactions is the number of collisions experienced by a pair of reacting molecules in a unit of time. A pair of molecules can be considered to be in a state of collision as long as their potential energy of interaction is of the order of magnitude of  $kT$ , which is  $\approx 1$  kcal $\cdot$ mol $^{-1}$  at room temperature. Such a definition brings with it the concept that a molecule is in the state of constant collision with the solvent walls and that a pair of molecules that find themselves in a solvent cage are neighbors and therefore might collide numerous times before separating and escap-

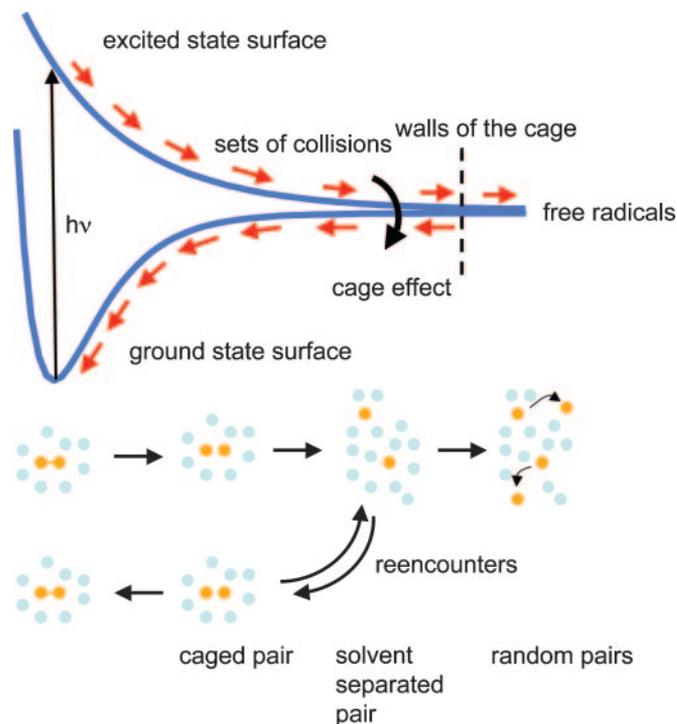


Fig. 2. Energy surface description of the cage effect and the influence of the solvent walls on sets of collisions.

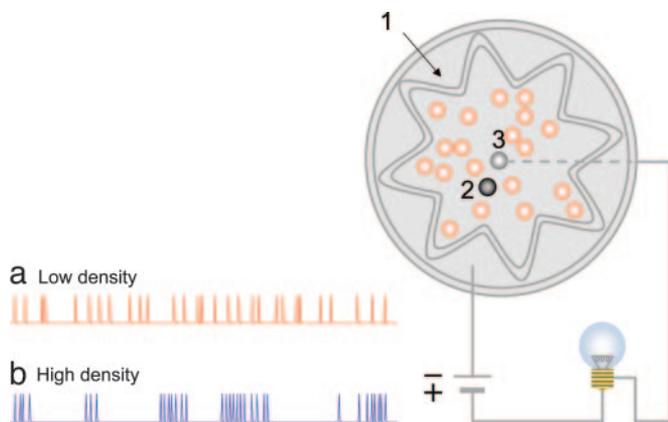
ing irreversibly from the solvent cage. A set of collisions that occur is termed an encounter. The mean lifetime of an encounter of a collision pair is roughly the time it takes for the partners to separate  $\approx 1$  or 2 molecular diameters apart. For a simple random walk for a nonviscous solvent at room temperature it will take  $\approx 10^{-11}$  s or less. This time is  $\approx 100$  times longer than the time scale of a collision of a pair of molecules in the gas phase.

The concept of the molecular solvent cage (and by extension the super cage) can be expressed in terms of the energy surface shown in Fig. 2. In Fig. 2 the guest@solvent<sub>cage</sub> system is taken from the ground-state surface to an excited-state surface as the result of the absorption of a photon,  $h\nu$ . The molecule guest then undergoes photochemical dissociation into a geminate radical pair (two radicals are born together) in a solvent cage, i.e., radical<sub>pair</sub>@solvent<sub>cage</sub>. On the excited-state surface the pair of radicals begin to separate from each other until they reach the walls of the solvent cage. At this point the radicals bounce off the wall and undergo a collision. The process of separation of the radical pair and collision occur repeatedly until there is an opening in the solvent cage that allows one of the radicals to escape from the solvent cage and diffuse into the bulk solvent. The opening of the solvent

cages corresponds to a radical<sub>pair</sub>@solvent<sub>cage</sub>  $\rightarrow$  radical<sub>pair</sub>@solvent<sub>cavity</sub> transformation (Fig. 1) and opens the possibility of escape of the radicals into the bulk solvent to form free radicals, radical<sub>pair</sub>@solvent<sub>cavity</sub>  $\rightarrow$  free radicals plus solvent cavity.

The basic idea of the cage effect and its control of the collisions between a pair of molecules in a solvent cage was demonstrated by Rabinovitch and Wood (5) who cleverly designed a simple 2D device that served as a mechanical model of the molecules in a liquid and convincingly demonstrated that collisions between two solute molecules occur in sets that increase in size as the density of the solvent particles is increased.

A model of the mechanical device demonstrating the cage effect is shown in Fig. 3. It consists of a flat brass plate with a zig-zag border (indicated by 1 in Fig. 3) and a metal knob (indicated by 3 in Fig. 3) that is isolated from the brass plate. The metal knob is connected to one of the poles of a battery and the brass plate itself is connected to the other pole. A number of nonconducting wooden balls (open circles in Fig. 3) are distributed on the plate and a single conducting metal ball (black ball indicated by 2 in Fig. 3) is added to the set of wooden balls. The nonconducting wooden balls rep-



**Fig. 3.** Demonstration of the cage effect on collisions in a mechanical simulation of a solvent cage. 1 indicates a zig-zag border over a conducting metal plate. 2 (black ball) indicates a metal knob; and 3 indicates a metal ball that is isolated from the metal plate. When the black ball makes contact with the metal knob the circuit is closed and a light bulb in the circuit flashes on. (a) The frequencies of electrical impulses caused by collision when 25 insulated wooden balls are on the plate. (b) The frequencies of electrical impulses caused by collision when 50 insulated wooden balls are on the plate.

represent the molecules of the solvent, and the single conducting ball and the central metal knob represent two dissolved molecules that can be tracked in time because every time the conducting ball and the central knob come into contact the electrical circuit is closed through the battery and the collision is registered electrically (by a lamp flashing or through a recording device). The experiment demonstrating the cage effect consists of placing the brass plate on a shaking machine capable of producing strong agitation of the plate. The zig-zag border of the device serves to transform the regular motion of the shaking machine into a chaotic agitation of the balls, simulating random thermal motion of the solute and solvent.

Fig. 3 *a* and *b* shows the results of the experiment as a function of different densities of added insulated balls. Fig. 3*a* (low density) indicates the frequencies of electrical impulses caused by collision when 25 insulated wooden balls are on the plate. Fig. 3*b* (high

density) indicates when 50 wooden balls are on the plate. Comparison of the results in Fig. 3 *a* and *b* demonstrates an important feature of the cage effect: at sufficiently high densities, collisions between solute molecules occur in sets. Depending on the density, the sets contained an average of about one collision per set at low density (Fig. 3*a*), whereas at high solvent density (Fig. 3*b*) the sets contained an average of 10 or more collisions per set. The important conclusion is that if two molecules do become nearest neighbors and collide in a solvent cage they can remain as colliding neighbors for a significant period.

The mechanical origin of the sets of collisions in a liquid is caused by the fact that the colliding pair of solute particles is surrounded by molecules of the solvent, and the pair can escape from the solvent cage only by a process of diffusion during which pairs undergo collisions many times before effective escape actually occurs. In other words, the reason for the increase in the size

of the number of collisions in a set, compared with the situation in the gas phase, is the reduction of the probability that one member of the pair, after separating from a collision with its partner, will find a “free volume” in the wall of solvent forming the cage without being reflected back into collision with its original partner. The final result is the expectation that, no matter what the detailed model of pair in a liquid, two neighboring molecules will exist in such a relation for a time in an encounter that is long compared with molecular vibrational frequencies.

## Conclusion

Perhaps the most important consequences of thinking supramolecularly is the ability to see apparently disparate chemical systems as the same thing in the sense that techniques and strategies for mechanistic investigation can be mapped from a well established supramolecular system onto a new one, once the topological supramolecular connections have been established. In this respect the move from molecular chemistry to supramolecular chemistry has the flavor of moving from static molecular structures to conformationally flexible structures. The rewards for adopting such an approach should be as great as those achieved when considering molecular conformational structures.

Chemists have mastered the concept of the covalent bond and demonstrated this mastery through the synthesis of remarkably complex molecules; chemists now seek to obtain a mastery of the intermolecular bond and apply this knowledge to an understanding of the profound complexity of the cell and chemical biology.

This work was supported by National Science Foundation Grants DUE-02-05781 and CHE-04-15516.

1. Lehn, J.-M. (1995) *Supramolecular Chemistry* (VCH, New York).
2. Weiss, R. G., Ramamurthy, V. & Hammond, G. S. (1993) *Acc. Chem. Res.* **26**, 530–536.
3. Turro, N. J. (1988) *Angew. Chem. Int. Ed. Engl.* **25**, 882–901.
4. Franck, J. & Rabinowitch, E. (1934) *Trans. Faraday Soc.* **30**, 120–131.
5. Rabinowitch, E. & Wood, W. C. (1936) *Trans. Faraday Soc.* **32**, 1381–1387.