Intermolecular Interactions

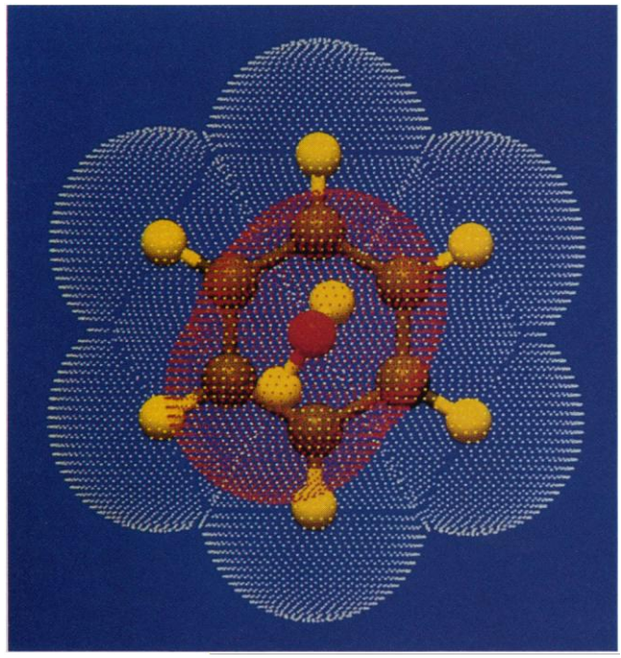
William Klemperer

Intermolecular interactions pervades virtually all areas of science. The physical interaction between chemically bonded species is an appealing intuitive picture; molecules have their own structural integrity that they maintain as they serve as building blocks for the formation of aggregates. The vast body of structural data that exists for valence forces allows ready prediction of the structures of molecules made up of light elements, such as hydrogen, carbon, nitrogen, and oxygen. The question of how to view weak interactions is an area of experimental and theoretical activity. It is of particular interest to learn what standard types of binary interactions exist and to see if there are essential local interactions. The spectroscopic study of weak, nonvalence interactions between molecules provides many new insights (*1*).

The widespread use of the seeded supersonic jet has permitted virtually every conceivable method of molecular spectroscopy to be applied to a variety of van der Waals molecules. The species are gases, almost always in their lowest vibrational levels corresponding to a system with an equivalent temperature of 10 K. The ubiquitous nature of intermolecular forces allows the observer to choose systems on the basis of molecular personality, or essentially whim.

For those interested in elementary geometric structure, perhaps the simplest problem is that of an inert gas atom and linear molecule, that is, a ball and a rod. A large number of systems of this type have been structurally characterized, particularly with argon (Ar). Sometimes a linear array is the minimum energy, sometimes a close-packed arrangement obtains, and sometimes the system can barely make up its mind, which results in large amplitude excursions between the two structurally inequivalent geometries. The structural characterization of inert gas binding in ArHF and ArHCl has been executed and analyzed with a precision and thoroughness as great as that achieved for chemically bonded species. In addition, the species Ar_nHF and Ar_nHCl with n up to 4 have been structurally characterized by rotational spectroscopy (*2*).

Although the study of inert gas binding is certainly intellectually enjoyable and provides benchmarks for the broad testing of the electron structure theory of chemistry, this field is sometimes viewed as esoteric. Pressures therefore exist to study real world systems. It is clear that the interactions of water have a primary importance.



Fast friends. A molecular graphics image of the hydrogen-bonding interaction between a benzene molecule and a water molecule, as determined by rotational-vibrational spectroscopic studies of clusters formed in a supersonic molecular beam. The van der Waals surfaces are depicted in white dots for the benzene molecule and in red dots for the water molecule. [Figure: S. Suzuki]

A large number of species bound to water have now been well characterized. Monomeric water exhibits two characteristic structures in its weak bonds, hydrogen donation (hydrogen bonding) and bonding by oxygen. Both are common. It is rare that both types of bonds occur in a single system. For the hydrogen-bonded water systems that have been studied, only the single hydrogen-bonded form has been observed; no example of a bidentate structure has yet been observed. The structure of water dimer is of this type; with one unit is the hydrogen donor and the other is the acceptor. Perhaps one of the most interesting structures is that of carbon dioxide and water dimer. Although the structure of the binary H_2O-CO_2 is as written, with the water oxygen bonding to the carbon in a planar symmetric form, the structure of $(H_2O)_2$-CO_2 is essentially cyclic, with both the bonding that obtains in $(H_2O)_2$ and in H_2O-CO_2 but also containing a new bond, namely, a hydrogen bond to the oxygen of CO_2 (*3*). It is noteworthy that the structure of the trimer of water has been most recently been shown to be cyclic, as anticipated, with each monomeric unit acting as both a hydrogen-bond donor and acceptor (*4*).

The structure of a number of hydrocarbons bonding with water are known and exhibit the variability of bonding described above. Acetylene is acidic enough to donate hydrogen to water, but ethylene, cyclopropane, and, most importantly, benzene (*5*) appear to be nucleophilic and accept a hydrogen from water. The hydrogen-bonded structures of these species are quite similar to those formed with the simple acids HF and HCl. It is interesting that the hydrogen bond appears to be directed to the center of benzene rather than toward the carbon ring.

In addition to structural characterization of weakly bonded species by means of their rotation, vibration, and hyperfine structure energies, there exists the possibility of observing a rich variety of dynamical behaviors. Consider a dimeric species made up of two identical monomeric units. If the dimer is polar, then perforce an isoenergetic structure with the dipole moment reversed must exist. The stationary states of the system will then be those in which quantum tunneling between isoenergetic structures occurs in a phased manner. A general description of weakly bonded systems is obtained by considering the structures generated by all permutations of identical nuclei that do not break the chemical bonds. A considerable body of work exists for the classification of the vibrational, rotational, and hyperfine levels under the permutation-inversion group that provides a most convenient technique for the quantitative discussion of the energy levels, degeneracies, and selection rules of nonrigid species (*6*).

The structural character of the complexes of water, namely, the minimum energy or equilibrium geometry, is just the beginning of the characterization of these systems. In virtually all cases, irrespective of whether water

Department of Chemistry, Harvard University, Cambridge, MA 02138.

attaches by oxygen or hydrogen bonding, slightly hindered internal rotation about the weak bond will occur. In the event that the water hydrogen-bonds, an unusual molecular motion, that of hydrogen-bond interchange, will also occur. Because all of the examples are at very low effective temperatures, these motions occur by tunneling through the barrier rather than through classical, over-the-barrier motion (*7*).

The structural characterization of binary gas-phase complexes of common, important species is by no means well balanced. There are numerous interesting complexes of monomeric water and a few containing the water dimer. Of the common atmospheric species, complexes of molecular oxygen appear conspicuously absent. Whether this insufficiency is of material significance, in view of the large number of condensed-phase magnetic resonance spectroscopic studies, is unclear.

The nature of intermolecular interactions is being explored empirically. Questions such as the uniqueness of hydrogen bonding as the directional weak interaction will in all likelihood be answered as relatively nonunique. It is likely that stereospecificity will be the rule in intermolecular interactions.

REFERENCES AND NOTES

1. See, for example, *Chemical Reviews*, vol. 88, no. 6 (1988).
2. H. S. Gutowsky *et al.*, *J. Chem. Phys.* **88**, 2919 (1988); *ibid.* **91**, 63 (1989).
3. K. I. Peterson, R. D. Suenram, F. J. Lovas, *ibid.*, **94**, 414 (1991).
4. N. Pugliano and R. J. Saykally, *Science*, in press.
5. S. Suzuki *et al.*, *ibid.* **257**, 942 (1992).
6. T. R. Dyke, *J. Chem. Phys.* **66**, 492 (1977); J. T. Hougen and N. Ohashi, *J. Mol. Spectrosc.* **109**, 134 (1984).
7. P. A. Block, M. D. Marshall, L. G. Pedersen, R. E. Miller, *J. Chem. Phys.* **96**, 7321 (1992).

Splicing Takes a Holliday

Joan Argetsinger Steitz

A central problem in pre–messenger RNA (pre-mRNA) splicing is to determine how the ends of an intron are juxtaposed for cutting and subsequent exon ligation. Small nuclear ribonucleoprotein particles (snRNP's) containing U1, U2, U4, U5, and U6 RNA's, as well as many protein factors, are essential to the splicing process (*1*). Also, a two-step mechanism has been described in which the 2′-hydroxyl of an adenosine residue located upstream of the 3′ splice site first attacks the 5′ splice site to form a lariat intermediate, and then the two exons are joined with release of the lariat intron. Consensus sequences at the 5′ splice site and at the branch point are recognized by base pairing with the U1 and U2 snRNP's, respectively (Fig. 1A). But the contributions of the U4, U5, and U6 RNA's, which assemble later as a tri-snRNP complex to form a splicing-competent body called the spliceosome, have remained obscure.

Three recent elegant studies that utilized yeast genetics suggest that the U5 RNA collaborates with U1 to bring the splice sites together in the newly assembled spliceosome. (i) Newman and Norman (*2*) discovered that point mutations within an evolutionarily invariant nine-nucleotide loop sequence in U5 RNA allowed use of novel 5′ splice sites when the normal 5′ splice site in a *Saccharomyces cerevisiae* pre-mRNA was mutated. (ii) Equally unanticipated were results of subsequent suppression analyses, implicating the same conserved U5 loop sequence in 3′ splice site activation (*3*). Specifically, splicing of defective introns was restored when positions 5 or 6 of the invariant U5 loop (see Fig. 1B) were mutated so that they were complementary to the nucleotides at positions 2 and 3 upstream of the novel 5′ splice site, or when positions 3 or 4 of the U5 loop sequence were mutated to allow pairing with the first two nucleotides of the 3′ exon. (iii) Reich, VanHoy, Porter, and Wise (*4*) found that appropriate substitutions in the invariant nucleotides 9 and 10 of U1 RNA could suppress splicing-defective changes in the **ag** at the 3′ splice site of a *Schizosaccharomyces pombe* pre-mRNA—but just for the first step of the reaction. In summary, these new observations suggest first that U1 base pairs with intron nucleotides at the 3′ as well as the 5′ splice site, as in an earlier crossover model (*5*), and second that U5 can pair with nucleotides in both exons to specify the exact points of cleavage at the 5′ and 3′ splice sites.

Evidence that all these base pairs between the pre-mRNA, U1, and U5 form simultaneously is currently lacking; but the interactions are not incompatible with one another. I therefore propose that a structure mimicking a Holliday junction (*6*)—a well-characterized intermediate in homologous recombination of DNA molecules—may exist in a newly assembled spliceosome and serve to juxtapose the 5′ and 3′ splice sites. Clearly, formation of every one of the base pairs shown in Fig. 1, B and C, is not essential since only the U1 and U5 sequences shown are absolutely conserved, whereas pre-mRNA 5′ and 3′ splice sites each conform to a consensus. Moreover, the first step of splicing can occur on some pre-mRNA's that lack the **ag** dinucleotide at the 3′ splice site (*7*) and therefore would lack branch **2** of the structure. Evidence that the U5 loop shifts between the first and second steps of splicing (*3*) argues that the structure is poised to undergo conformational change and that all four arms of the crossover may not simultaneously be comprised of duplex RNA. Nonetheless, the model suggests how the U1, U5, and pre-RNA molecules may be positioned relative to one another immediately after spliceosome assembly.

To form the proposed Holliday structure, U1 nucleotides 9 and 10 would exchange their initial pairing with 5′ exon nucleotides −1 and −2 (Fig. 1A) for pairing with the **ag** at the 3′ end of the intron (*4*) (forming branch **2** in Fig. 1, B and C). Simultaneously, the U5 invariant loop would contact the 5′ exon (forming branch **3**). Exon sequences are not highly conserved. Therefore, the multiple U residues in the U5 loop may have been selected during evolution for their ability to pair nondisruptively with all other nucleotides, as in mitochondrial decoding (*8*). Holliday structures are characterized by their potential for isomerization and strand exchange (branch migration). The release of U1 RNA sometime prior to the first step of splicing would destroy branches **1** and **2** and allow U5 to establish closer contacts with the 3′ exon (branch **4**) by branch migration of the U at the top of the conserved loop, as suggested by Newman and Norman (*3*). Additional conformational changes seem likely to occur at this point. Biochemical data support re-recognition of the **ag** dinucleotide at the 3′ end of the intron (*9*) and a change in the environment of the 3′ splice site (*10*) as the spliceosome proceeds from the first to the second step of the reaction.

If a Holliday-like structure is critical for defining splice sites, it should be formed by all types of spliceosomes. Trans splicing is a special circumstance in which a 5′ exon carried on an snRNP-like RNA (called an SL RNA) is joined to a 3′ exon on a separate cellular transcript. Organisms such as trypanosomes, which engage exclusively in trans splicing, curiously have only U2, U4, and U6 RNA's, whereas organisms like nematodes that carry out both normal and

Department of Molecular Biophysics and Biochemistry, Howard Hughes Medical Institute, Yale University School of Medicine, 295 Congress Avenue, New Haven, CT 06536–0812.