electrophiles than is possible with pure FSO_3H.

The highest known acidities have been achieved with HF/SbF_5 (*6*). Anhydrous HF is a relatively weak acid, but at near $HF/SbF_5 = 3/1$, the acidity is enormous. However, this mixture is rather unsuitable for many applications requiring a superacid: SbF_5 is strongly oxidizing, the mixture is somewhat of a vegetable soup of complex ions, it is poorly soluble in and reacts with many cosolvents, and HF is a very hazardous chemical. This system shows clearly that for many applications of superacids, more is needed than just very high acidity. A user-friendly conjugate base is equally important, allowing strong electrophiles to be studied in solution and the solid state without strong interactions with the solvent or the conjugate base of the superacid. The study of acid-base reactions in the gas phase also requires a superacid capable of existing as a molecular species in the gas phase.

Research toward these general goals has evolved in two related but different directions: new molecular acids in which the conjugate base would be more stabilized and less reactive and very weakly coordinating anions as the conjugate bases of hypothetical superacids.

One route to potential new molecular superacids involved moving away from oxyacids. Using the powerful electron-withdrawing group CF_3SO_2, very strong acids based on nitrogen and carbon can be prepared. Examples are $(CF_3SO_2)_2NH$ and $(CF_3SO_2)_3CH$, the parent members of the more general families $(R_fSO_2)_2NH$ and $(R_fSO_2)_3CH$, where R_f are a wide variety of fluorinated groups. In the conjugate bases of these acids, the negative charge is delocalized over more atoms than in the related oxyacid CF_3SO_3H. The high acidity of these species has been demonstrated in the gas phase, where they are stronger acids than CF_3SO_3H. The strongest acid yet measured in the gas phase is the sulfonimide $(C_4F_9SO_2)_2NH$ (*7*). The sulfonimides are a particularly versatile class of acids. They can be prepared in a large variety and are easily incorporated into polymeric systems. They are currently of interest in many applications including lithium batteries, polymer membrane fuel cells, photoacid generators, ionic liquids, and catalysis (*8*).

Another route to very weakly coordinating anions uses bulky electron-withdrawing groups around a typical Lewis acid central atom. Here, the parent superacid HA might not exist as an isolable compound or might not be a superacid in the usual sense. The forerunner to this approach is the tetraphenylborate anion $B(Ph)_4^-$. Incorporating fluorine atoms and fluorinated groups onto the phenyl rings, as in $B(C_6F_5)_4^-$ and $B[3,5\text{-}C_6H_3(CF_3)_2]_4^-$, lowers the coordinating ability. Very weakly coordinating anions were realized with the large electron-withdrawing substituent TeF_5O (*9*). For example, the anions $Nb(OTeF_6)_6^-$ and $Sb(OTeF_6)_6^-$ were prepared based on the strong Lewis acids NbF_5 and SbF_5. The negative charge of these large spherical anions is distributed over the 36 fluorine atoms, which form a protective shield around the central atoms and lead to high kinetic stability of the anions. Salts of electrophiles, such as $AgSb(OTeF_5)_6$, are remarkably soluble in low-polarity solvents like dichloromethane. The resultant naked Ag^+ is then forced to become less "naked" by coordinating to the dichloromethane, a phenomenon not previously observed.

However, even these large anions show weak interactions with electrophiles in the solid state through the fluorine atoms of the TeF_5O groups. To obtain even more weakly coordinating anions, chemists turned to carboranes as a potential building block for the ultimate superacid.

The carborane anion $CB_{11}H_{12}^-$ was recognized as having high potential for this purpose. The challenge was how to modify it into the conjugate base of a superacid. Several approaches have been pursued by substituting the hydrogen by halogens (F, Cl, and Br) and fluorinated groups like trifluoromethyl. These new directions in superacids are exciting and hold promise for both fundamental science and novel applications. The late Robert W. Taft dreamed of a superacid so strong that it would undergo spontaneous ionization into a plasma of protons in the gas phase. This dream may be coming closer to reality.

References

1. N. F. Hall and J. B. Conant, *J. Am. Chem. Soc.* **49**, 3047 (1923).
2. G. A. Olah, *Angew. Chem. Intl. Ed. Engl.* **12**, 173 (1973).
3. R. J. Gillespie and T. E. Peel, *J. Am. Chem. Soc.* **95**, 5173 (1973).
4. C. A. Reed, K.-C. Kim, R. D. Bolskar, L. J. Mueller, *Science* **289**, 101 (2000).
5. G. A. Olah and G. K. S. Prakash, *Supeacids* (Wiley-Interscience, New York, 1985).
6. R. J. Gillespie and J. Liang, *J. Am. Chem. Soc.* **110**, 6053 (1988).
7. I. A. Koppel *et al.*, *J. Am. Chem. Soc.* **1166**, 3047 (1994).
8. D. D. DesMarteau, *J. Fluorine Chem.* **72**, 203 (1995).
9. S. Strauss, *Chem. Rev.* **93**, 927 (1993).

PERSPECTIVES: MOLECULAR BIOLOGY

Transposase Team Puts a Headlock on DNA

Tanya L. Williams and Tania A. Baker

Transposable DNA elements (transposons) populate the genomes of numerous organisms, from bacteria to humans. These genetic elements "jump" from old to new DNA locations, a process that mutates genes, rearranges chromosomes, and transmits genetic information between cells. Transposons, along with their retroviral cousins (for example, HIV), relocate through a common set of DNA cleavage and joining reactions (*1*). The proteins catalyzing these reactions (transposases and integrases, respectively) are encoded in the DNA of the mobile element and share structurally related catalytic regions (*2*). How do these proteins cleave both ends of the linear element (which may be far apart) and then join them to a new, distant DNA site? The first structure of a transposase-DNA complex, described by Davies *et al.* (*3*) on page 77 of this issue, begins to reveal the molecular details of this process.

Transposition and retroviral integration start with assembly of a protein-DNA complex. By binding to sequences near each end of the mobile element, transposases and integrases pair both DNA ends within the complex. Only after formation of this complex can these enzymes catalyze their reactions. In the bacterial Tn5 transposase-DNA complex reported by Davies *et al.* (*3*), two transposase molecules (subunits) pair the two DNA ends in a manner poised for catalysis. How does assembly of this DNA-protein complex control catalytic activity? One possibility is that the enzyme's active sites are constructed only upon formation of the complex. The Tn5 structure shows, however, that each active site is composed entirely of amino acids from one subunit. The geometry of the catalytic amino acids closely resembles that found in the active site of a non–DNA-bound version of the protein (*4*). This similarity suggests that, at least for Tn5 transposase, most features of the active site are not dependent on complex assembly.

Given that each transposase molecule comes armed with an intact active site, why does enzyme activity require complex assembly? Within the Tn5 complex, each DNA end is positioned within an active site by extensive protein-DNA contacts made by both transposase subunits. Amino acids near the

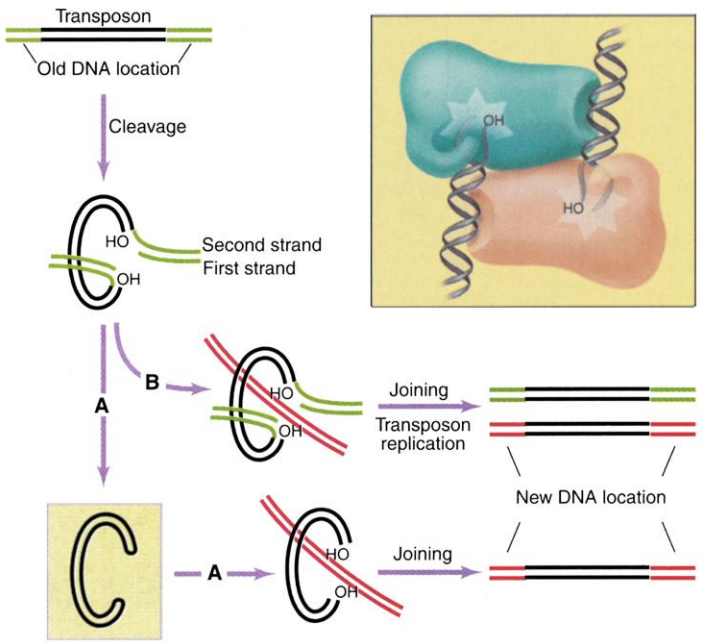
T. L. Williams is in the Department of Chemistry, Office 68-517, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: tlwillia@mit.edu
T. A. Baker is in the Department of Biology, Building 68, Room 523, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: tabaker@mit.edu

active site of one subunit, in conjunction with a protein "arm" that extends from this catalytic region, clamp onto the DNA end near the terminal nucleotide of the mobile element. This same DNA molecule is locked in place through additional interactions with a distinct region in the second protein subunit (shown as a crescent in the figure). This pattern of subunit-DNA interactions organizes the protein and DNA components such that the subunit bound to one DNA end through the "crescent" catalyzes cleavage and joining of the second DNA end.

Could other transposase and integrase family members adopt a similar organizational structure? The three-dimensional structure of the Tn5 transposase catalytic region is conserved throughout this family, although the "arm" and "crescent" DNA-binding regions are not. In spite of these differences, biochemical studies (*5*) of the transposase-DNA complex of bacteriophage Mu (a virus that uses transposition to infect bacteria) indicate a similar architecture: A subunit bound to one DNA end cleaves and joins the other end. These two examples hint at a common arrangement of protein-DNA interactions within this protein family that provide a framework for coupling cleavage and joining between the ends of these mobile elements.

Once assembled into an active complex, how do these proteins catalyze the chemical reactions at the ends of the mobile element? Each active site contains three strictly conserved acidic amino acids (the DDE motif) (*6*) and catalyzes a series of reaction steps at one DNA end (see the figure). These acidic amino acids are thought to promote catalysis by acting as ligands for essential divalent metal ions. The Davies *et al.* structure (*3*) gives the first view of DNA in a transposase active site; it contains a DNA end that has already been cleaved and one Mn^{2+} ion positioned near two of the DDE motif amino acids. The single Mn^{2+} ion is close to the 3′-OH of the terminal nucleotide, where it could activate this 3′-OH to attack a properly positioned DNA phosphodiester bond.

Although Davies *et al.* (*3*) observe only one divalent metal ion per active site, the number of metal ions directly involved in catalysis by this protein family is debatable. Structures of the catalytic regions of other family members show either one or two metal ions bound by the DDE motif in the absence of DNA (*7*). The similarity of these catalytic domains to those of other DNA and RNA processing enzymes has led to the previous suggestion of a two-metal-ion mechanism (*1, 8*). Solving the structure of complexes caught at different reaction steps and therefore containing additional DNA components should help to clarify the catalytic mechanisms used by this protein family.



Transposon "Jumping." A transposon (black) embedded within an old DNA location (green) moves to a new DNA location (red) through one of two pathways. Both pathways (A and B) contain two identical reaction steps: (1) cleavage of the first strand at each DNA end to yield a 3′-OH and (2) joining of this 3′-OH to a new DNA site. Cut-and-paste transposons (pathway A) cleave both strands at each end of the element before the joining step. The yellow box depicts the reaction step mimicked in the Tn5 structure (*3*). **(Inset)** Organization of the "hairpin" intermediate of the Tn5 transposase-DNA complex. Each end of the mobile element DNA (black) is locked in place through contacts with both transposase subunits (blue and orange). One DNA end interacts with amino acids within and near the active site (star) of one subunit and with the amino-terminal region (crescent) of the second subunit.

Some transposons (like Tn5) "jump" using a mode of transposition in which both DNA strands at each end of the element are cleaved, releasing the element completely from its old location, before joining it to the new DNA site (cut-and paste transposition). For a subset of these transposases, the same active site cleaves the first strand, then cleaves the second strand, before joining the first strand to a new DNA site (see the figure). How does the same active site catalyze all of these reactions?

The structure reported by Davies *et al.* captures a mimic of the DNA intermediate of the second cleavage step. Transposase promotes second-strand cleavage by catalyzing the direct attack of the phosphodiester backbone of this strand by the free 3′-OH of the first strand. This attack simultaneously cleaves the second strand and forms a new phosphodiester bond between the first and second strands (*9*). To form this "DNA hairpin," the backbone of the second strand is distorted within the active site, looping around to place the 5′ end of this strand next to the 3′-OH. This DNA loop is held in place by extensive electrostatic and stacking interactions with several amino acid side chains (some belonging to the YREK motif of known hairpin-forming transposases). Thus, amino acids of the DDE and YREK motifs cooperate to cleave both strands at a DNA end.

Although the Tn5 structure provides a detailed snapshot of the architecture of a transposase-DNA complex, several questions remain. The structure lacks two important DNA components: the old and new DNA sequences (see the figure). What regions of transposase interact with these DNAs? Within the Tn5 complex, the two active sites lie at the bottom of a large, positively charged groove that appears capable of housing DNA. Can this groove bind both old and new DNA sequences? Could other transposases and retroviral integrases that do not cleave both strands at each end of the element before joining to a new DNA site use a similar binding pocket? How are the movements of these DNA pieces with respect to the active sites controlled? Discovery of the strategies used by different members of this transposase/integrase family to bind and coordinate these DNA components at each reaction step is eagerly awaited.

References

1. For example, reviewed in K. Mizuuchi, *Annu. Rev. Biochem.* **61**, 1011 (1992); N. L. Craig, *Science* **270**, 253 (1995).
2. P. Rice, *Curr. Opin. Struct. Biol.* **6**, 76 (1996).
3. D. R. Davies, I. Y. Goryshin, W. S. Reznikoff, I. Rayment, *Science* **289**, 77 (2000).
4. D. R. Davies *et al.*, *J. Biol. Chem.* **274**, 11904 (1999).
5. T. L. Williams *et al.*, *Genes Dev.* **13**, 2725 (1999); S.-Y. Namgoong and R. M. Harshey, *EMBO J.* **17**, 3775 (1998); H. Aldaz *et al.*, *Cell* **85**, 257 (1996); H. Savilahti and K. Mizuuchi, *Cell* **85**, 271 (1996).
6. J. Kulkosky *et al.*, *Mol. Cell. Biol.* **12**, 2331 (1992).
7. S. Maignon *et al.*, *J. Mol. Biol.* **282**, 359 (1998); G. Bujacz *et al.*, *J. Biol. Chem.* **272**, 18161 (1997); G. Bujacz *et al.*, *Structure* **4**, 89 (1996); K. Katayanagi *et al.*, *Proteins* **17**, 337 (1993).
8. For example, see K. Mizuuchi, *Genes Cells* **2**, 1 (1997); Y. Wang and T. A. Steitz, *Structure* **3**, 131 (1995); L. S. Beese and T. A. Steitz, *EMBO J.* **10**, 25 (1991).
9. A. K. Kennedy *et al.*, *Cell* **95**, 125 (1998); A. Bhasin *et al.*, *J. Biol. Chem.* **274**, 37021 (1999).