Biochemistry

 Volume 41, Number 47 November 26, 2002

Current Topics

Catalytic Mechanisms of Restriction and Homing Endonucleases

Eric A. Galburt and Barry L. Stoddard*

Fred Hutchinson Cancer Research Center and Graduate Program in Biomolecular Structure and Design, University of Washington, 1100 Fairview Avenue North, A3-023, Seattle, Washington 98109

Received July 10, 2002; Revised Manuscript Received September 23, 2002

ABSTRACT: The catalytic mechanisms of type II restriction endonucleases and homing endonucleases are discussed and compared. Brief reviews of the chemistry of phosphoryl transfers and canonical one-metal and two-metal endonucleolytic mechanisms are provided along with possible future directions in the study of endonuclease active sites. The discussion of type II restriction endonucleases is comprised of a description of the general architecture of the canonical active site structural motif followed by more in-depth examples of one- and two-metal mechanisms. The homing endonuclease section is comprised of four sections describing what is known regarding the cleavage mechanisms of the four group I intron homing endonuclease families: LAGLIDADG, His-Cys box, H-N-H, and GIY-YIG.

Enzymes have long been appreciated for their ability to efficiently and specifically catalyze chemical transformations. One of the most important chemical transformations catalyzed in nature is phosphoryl transfer. "Phosphate esters and anhydrides dominate the living world" (*1*) as phosphoryl transfer plays crucial roles in most basic cellular functions, including metabolism, signal transduction, DNA replication, and transcription. Phosphoryl transfer reactions also allow for the manipulation of nucleic acids and form the basis for modern molecular biology. Kinases catalyze the transfer of the γ-phosphate from ATP to a variety of molecules, while phosphatases catalyze the removal of phosphoryl groups. Mutases transfer phosphoryl groups between two atoms on a single molecule. Polymerases catalyze the extension of nucleic acid polymers by transferring the α-phosphate of nucleotide triphosphates from pyrophosphate to the terminal 3′-hydroxyl, and endonucleases are able to cleave phosphodiesters by transferring a phosphoryl group from a bridging oxygen atom to an activated water molecule.

* To whom correspondence should be addressed. Phone: (206) 667-4031. Fax: (206) 667-6877. E-mail: bstoddar@fhcrc.org.

Both homing and restriction endonucleases have been extensively studied biochemically and structurally. A great deal of information regarding the specific structural mechanisms of these enzymes has been obtained, yet there are still unresolved issues. Here we discuss the chemistry of phosphoryl transfer reactions, similarities and differences between the catalytic mechanisms of homing endonucleases and type II restriction endonucleases, and potential future directions that may shed more light on the details of these mechanisms.

Phosphoryl Transfer

Formally, the direct transfer of a phosphoryl group can take place according to a variety of mechanisms depending on the position of the attacking nucleophile relative to the leaving group and the degree of bond formation at different points along the reaction coordinate (*2*). The transition state geometry of phosphoryl transfers is a trigonal bipyramid with three equatorial and two axial oxygens. Oxygen atoms entering or leaving this configuration must always be axial. The extremes with respect to bond formation are associative mechanisms (S_N2) where the nucleophile is completely

10.1021/bi020467h CCC: $22.00
Published on Web 10/12/2002

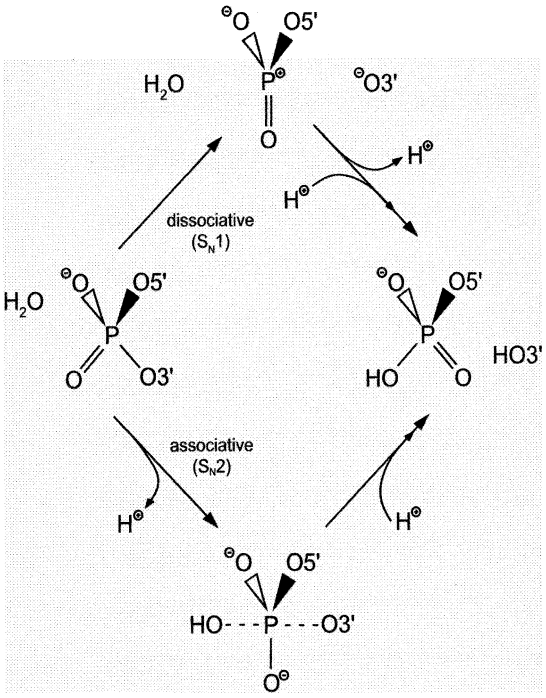

FIGURE 1: Schematic showing associative (S_N2) and dissociative (S_N1) pathways to phosphodiester bond cleavage.

bonded to the phosphorus atom before the leaving group is eliminated and dissociative mechanisms (S_N1) in which the elimination of the leaving group occurs first and the nucleophile attacks an unstable metaphosphate intermediate (PO_3) (Figure 1). In solution, S_N1 reactions result in the racemization of products, but an enzyme that does not release metaphosphate constrains the position and orientation of the intermediate such that chirality is either maintained or inverted. Associative mechanisms can be classified as in-line or adjacent. When the nucleophile and leaving group are both apical at the same time along the reaction coordinate, the mechanism is termed in-line and the reaction results in an inversion of configuration. In an adjacent mechanism, the leaving group is equatorial when the nucleophile attacks and the nucleophile is equatorial when the leaving group is released. Thus, adjacent mechanisms are also characterized by a *pseudorotation* somewhere along the reaction coordinate that reorganizes the bipyramid so that oxygen atoms are switched from apical to equatorial and vice versa. It has been shown experimentally that phosphoryl transfer enzymes use in-line mechanisms. The products of some transfers do retain configuration relative to substrate, but a two-step reaction where a phospho−enzyme intermediate is formed accounts for this. This leads to a retention of configuration through an even number of in-line attacks.

At physiological pH, phosphate ester bonds have large barriers to cleavage even though they are thermodynamically unstable. This is mostly due to the fact that around neutral pH, the phosphoryl group is negatively charged and electrostatically repels potential attacking nucleophiles. This duality of kinetic stability and thermodynamic instability can explain nature's choice of phosphate for such a variety of roles (*1*). To efficiently catalyze the cleavage of phosphate esters, certain chemical elements are required. These include a nucleophilic group to which to transfer the phosphoryl group, a basic moiety to activate and position the nucleophile,

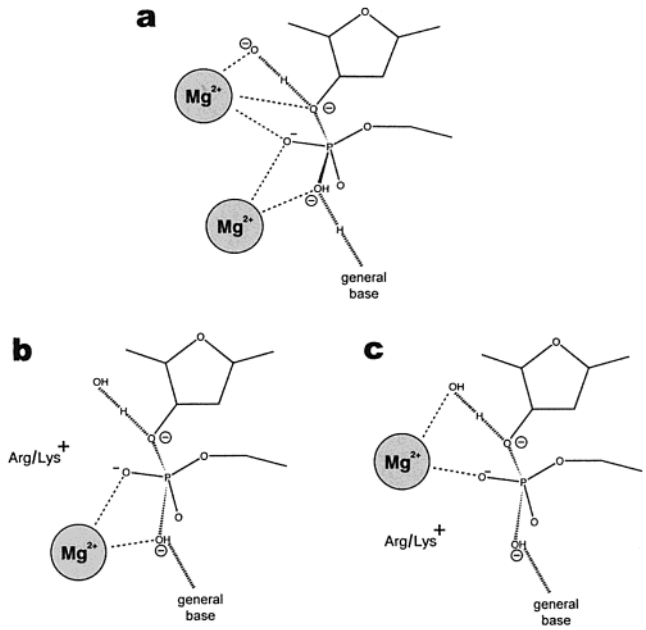

FIGURE 2: Canonical transition states of one- and two-metal endonucleolytic mechanisms. In each mechanism, the metal ions stabilize the negative charge that develops on the dianionic transition state. (a) The two-metal mechanism uses metals to position both the nucleophile and the proton donor water molecules. (b) One-metal mechanism observed in restriction endonucleases in which the metal ion is responsible for positioning and activating the nucleophile. (c) One-metal mechanism observed in I-*Ppo*I in which the metal ion positions the proton donor, but does not make any contacts with the nucleophilic water. Both one-metal mechanisms include additional positive charge from protein side chains to complement the divalent metal ion and stabilize the dianionic transition state.

a general acid to protonate the leaving group, and the presence of one or more positively charged groups to stabilize the phosphoanion transition state. The diversity of chemical groups available to proteins has made it possible for evolution to arrive at many diverse strategies that satisfy the above requirements. A common feature of many nuclease catalysts (and other phosphoryl transfer enzymes), including restriction and homing endonucleases, exonucleases, and some ribozymes, is the use of bound metal ions as cofactors. Endonucleases utilizing one, two, and even three bound metals per active site have been observed and described crystallographically. Metal ions can act as Lewis acids by lowering the pK_a of their directly coordinated water molecules. A resulting hydroxide may then serve as either a nucleophile or a general base. Alternatively, an acidic, metal-bound water molecule can efficiently protonate the leaving group. Perhaps most importantly, the positive charge of the divalent metal ion can stabilize the −2 charge that accumulates at the phosphoanion transition state relative to the −1 charge of the ground state. All three of these functions could lower the free energy of the transition state and accelerate the reaction. It has long been understood that divalent metals are required for endonuclease function and that even a high concentration of monovalent cations does not restore activity. General one- and two-metal mechanisms are illustrated in Figure 2. Exceptions are ribonucleases such as barnase and ribonuclease A that cleave phosphodiester bonds in RNA (*2*, *3*). They do not use metal ions as cofactors, but instead use protein side chains to provide the necessary positive

charge and to act as proton donors and acceptors. Likewise, studies of the hairpin and hammerhead ribozymes indicate that metals are needed for the proper folding of the ribozyme, but do not seem to play a direct role in the catalytic mechanism (*4*, *5*). A possible explanation for the lack of metals in some ribonucleolytic active sites could be the relative instability of an RNA phosphodiester as compared to that of DNA. These enzymes have an easier job than deoxyribonucleases due to the presence of 2′-hydroxyls that can be positioned and activated for nucleophilic attack and formation of a 2′,3′-cyclic phosphodiester.

The catalytic mechanisms of type II restriction endonucleases have been extensively studied. More recently, our understanding of homing endonuclease catalysis has also been growing rapidly. Here, we aim to compare and contrast what is known and what is yet to be determined about the catalytic mechanisms of both type II restriction and homing endonucleases.

Restriction Endonucleases

There are more than 3000 known type II restriction endonucleases. Restriction enzymes combined with methyltransferase enzymes form the basis of RM (restriction−modification) systems of bacteria and other prokaryotes that protect host organisms against the invasion of foreign DNA. Restriction endonucleases are used ubiquitously as reagents for the manipulation of DNA in the laboratory, but have long been a field of study unto themselves in which the fundamental biochemical, structural, kinetic, and catalytic features of protein−DNA interactions are considered. The first crystal structure of a restriction endonuclease−DNA complex, corresponding to *Eco*RI bound to its cognate DNA site, was published in 1986 at a resolution of 3 Å (*6*); the connectivity and topology of that structure were corrected in a subsequent study in 1990 (*73*). Since then, the protein−DNA complex structures of 11 other members of this family have been determined, and the resolution for many of the complexes extends well past 2 Å (*7*−*18*, *74*). These high-resolution structures have allowed a detailed examination of both DNA binding and catalysis. The resulting wealth of information is summarized well by a recent review (*19*).

Despite the plethora of biochemical and structural data, the exact mechanisms used by type II restriction endonucleases are still a matter of debate. Here, we will review the general structure of the type II active site and discuss different mechanistic proposals.

Type II Restriction Endonuclease Active Site Architecture. The active sites of type II restriction endonucleases are defined by the conserved PD...(D/E)XK motif where the gap ranges between eight (*Nae*I) and 25 residues (*Bgl*I) and X represents any hydrophobic amino acid. Different positions within the motif are conserved to varying degrees (Table 1). The first aspartic acid is the only residue to be absolutely conserved among the enzymes that have structural data. The proline has been observed as an isoleucine, valine, asparagine, and threonine. The second acidic residue has been observed as a serine. The hydrophobic residue has been observed as a cysteine. The lysine has been observed as a glutamic acid or a glutamine. The motif serves as a general guideline, but does not guarantee the existence of an endonuclease active site. For example, the motif was observed in the sequence of the His-Cys box homing endonuclease (*20*), I-*Ppo*I, but the structure revealed that the motif residues do not constitute the active site (*21*).

Table 1: Summary of Active Site Motifs for Type II Endonucleases[a]

enzyme (year)	D/E	$P_{x-1}D_x$	...	$(D/E)_{x-2}Z_{x-1}{}^{c}K_x$
*Eco*RI (1990)	D_{59}	PD_{91}	X_{19}	EAK_{113}
*Eco*RV (1993)	E_{45}	PD_{74}	X_{15}	DIK_{92}
*Bam*HI (1994)	E_{77}	**I**D_{94}	X_{16}	EF**E**$_{113}$
*Pvu*II (1994)	E_{55}	**N**D_{58}	X_1	ELK_{70}
*Fok*I (1997)		PD_{450}	X_{16}	DTK_{469}
*Bgl*I (1998)	E_{87}	PD_{116}	X_{25}	DIK_{144}
*Cfr*10I (1998)	E_{71}	PD_{134}	X_{53}	**S**VK_{190}**(E$_{204}$)**
*Mun*I (1999)		PD_{83}	X_{14}	EIK_{100}
*Bgl*II (2000)	N_{69}	**I**D_{84}	X_9	DV**Q$_{95}$**
*Ngo*MIV (2000)	E_{70}	PD_{140}	X_{44}	**S**CK_{187}**(E$_{201}$)**
*Bso*BI[b] (2001)		**V**D_{212}	X_{27}	ELK_{242}
*Nae*I (2001)	E_{70}	**T**D_{86}	X_8	DCK_{97}
*Bse*634I (2002)	E_{80}	PD_{146}	X_{49}	**G**LK_{198}**(E$_{212}$)**

[a] Bold text denotes variations from the "conserved" sequence motif and ligands for the second metal ion or residues that align with them. The observed metal ion is Gd^{3+} and was used as a heavy atom derivative. [b] *Bso*BI contains a $D_{251}E_{252}H_{253}$ motif, and H253 is believed to be the general base. [c] Z represents any hydrophobic group.

The active sites of type II restriction endonucleases have been observed crystallographically in many different metal-bound states. The extreme case is the 13 different crystal structures that have been described for *Eco*RV. There appear to be two main varieties of type II active sites with regard to their metal binding capabilities. The enzymes *Bgl*II and *Eco*RI have been observed to contain a single metal binding site per subunit (*16*). Importantly, the *Eco*RI structure (1QPS) is an active complex demonstrating that a single metal is sufficient for DNA cleavage. In contrast, the enzymes *Bam*HI, *Bgl*I, *Pvu*II, and *Ngo*MIV have been observed to contain two metal binding sites per subunit (*13*, *15*, *22*, *23*) (Figure 3). The structure of *Bam*HI has been determined in both substrate and product complexes, indicating a two-metal mechanism is responsible for endonuclease activity (*22*). One of the two metal-binding sites is homologous to the single site observed in *Bgl*II and *Eco*RI (*16*) (M1, Figure 3). This site is formed by the two conserved acidic residues, the backbone carbonyl of the hydrophobic residue [PD...(D/E)XK], and a nonbridging oxygen from the scissile phosphate. The second metal site (M2) is created by the first conserved aspartic acid, a substrate nonbridging oxygen, and a complement of other acidic residues depending on the specific enzyme. In *Pvu*II, the backbone carbonyl of G56 completes the second metal site, while in *Bam*HI, it is the side chain of E77.

Lying outside the clearly defined one-metal and two-metal categories is *Eco*RV. The structure of *Eco*RV has been determined in a variety of different states with many divalent cations (Mg^{2+}, Mn^{2+}, Co^{2+}, Ca^{2+}, Mg^{2+}/Ca^{2+}, and Mn^{2+}/Ca^{2+}) by both cocrystallization and soaking experiments (*24*−*28*). The data available indicate that *Eco*RV uses one to three metal ions to cleave DNA, but the number of metal ions involved in the catalytic mechanism is still in question. Clearly, the precise mechanism by which these enzymes catalyze DNA cleavage depends on the minimal number and positions of metal ions in the active site. For example, the proposed catalytic mechanism of the LAGLIDADG homing endonuclease, I-*Cre*I, was revised after a high-resolution

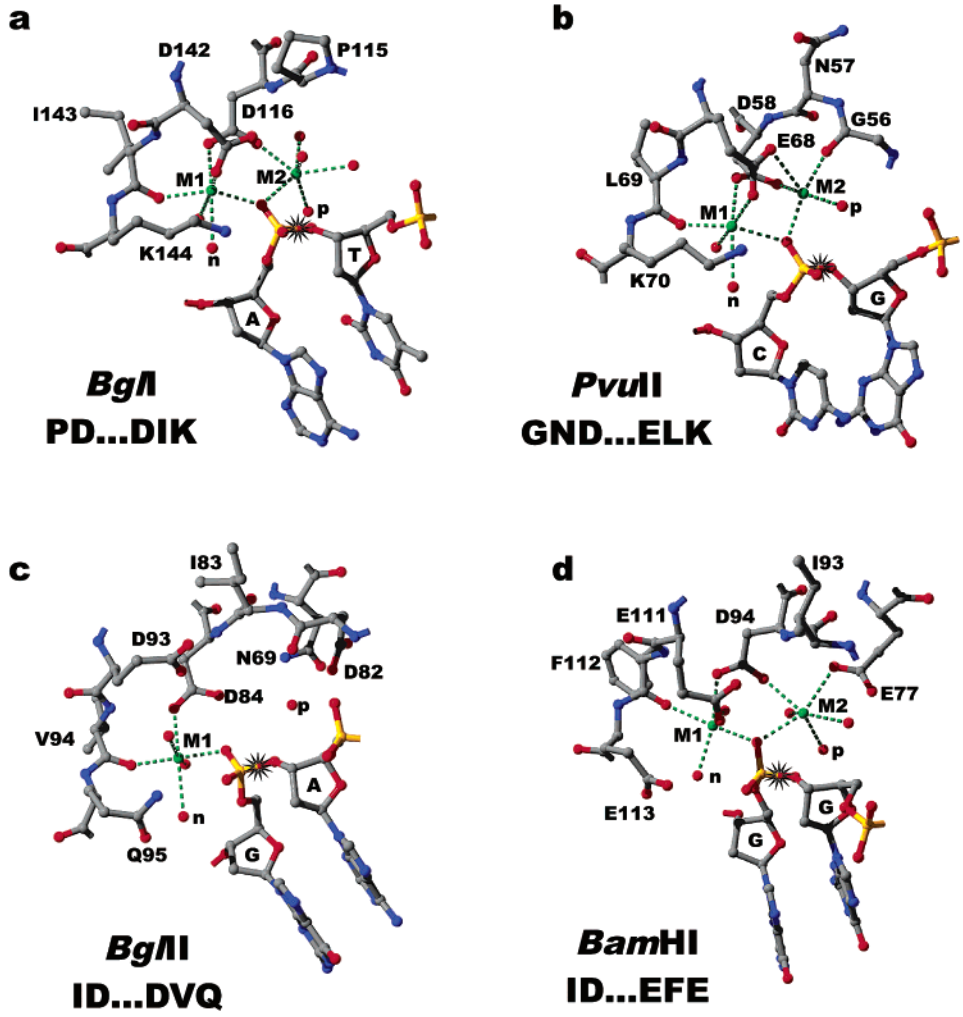

FIGURE 3: Ball and stick representations of the active sites of four restriction endonuclease enzyme−substrate complexes: *Bgl*II (a), *Pvu*II (b), *Bgl*I (c), and *Bam*HI (d). Each active site is labeled along with its sequence motif and is shown in the same orientation for ease of comparison. The metal ion (green) site that is conserved in all four structures is labeled M1, and the second site is labeled M2. Metal ligands are indicated with green dotted lines. Water molecules are labeled according to their proposed function: nucleophile (n) or proton donor (p). The scissile bond is denoted with an asterisk.

crystal structure of a metal-bound product complex was determined to have three metals bound per protein dimer instead of two (*29*). The following sections describe the potential one-metal and two-metal mechanisms that have been proposed for type II restriction endonucleases.

One-Metal Mechanism. The one-metal mechanism involves a metal bound in site 1 as described above (Figure 3c). A metal-coordinated water molecule is positioned and activated for in-line nucleophilic attack of the electrophilic phosphorus atom. In *Bgl*II, the nucleophilic water molecule also makes hydrogen bonds to Q95, both bridging and nonbridging oxygen atoms from the scissile phosphate, and a water molecule bound to a nonbridging oxygen atom from the 3′-adjacent phosphate (*16*). A second metal-bound water molecule is positioned by N69, D82, and D84, and makes a hydrogen bond to the leaving group oxygen. This water molecule appears to be well-positioned to donate a proton to stabilize the leaving group. In *Eco*RI, the nucleophilic water is not contacted by a protein residue, but does make hydrogen bonds to both a 3′-phosphate-bound water and a nonbridging oxygen from the scissile phosphate (*19*). A putative leaving group protonating water molecule is positioned by R145. The single-metal mechanism uses an in-line, metal-bound nucleophile with a pK_a that is lowered through metal association. However, neither structure has an obvious general base, which would be required to abstract the proton from the nucleophilic water molecule in an associative nucleophilic attack. Possible features that could complete a single-metal mechanism include substrate-assisted catalysis where the general base is a nonbridging phosphate oxygen (*30*), a bulk solvent-supplied metal-bound hydroxide ion that does not need to be deprotonated, or a dissociative transition state. Any of these possibilities would abrogate the requirement for an efficient general base supplied by the protein.

Two-Metal Mechanism. The two-metal mechanism is similar to the one-metal mechanism in that the metal ion in site 1 is responsible for positioning the nucleophilic water for in-line attack. In addition, a lysine residue in both *Pvu*II (K70) and *Bgl*I (K144) contacts the nucleophile (*13*, *23*). The second metal is located on the other side of the scissile phosphate from the nucleophilic water molecule and is well positioned to activate a water molecule for protonation of the leaving group oxygen and to help stabilize the negative charge that develops in an associative transition state (Figure 3). Intriguingly, a calcium-bound, prereactive complex of

*Bam*HI indicates that E113 makes a hydrogen bond to the nucleophilic water molecule (2.9 Å) and represents a possible general base (*22*). However, neither *Pvu*II nor *Bgl*I has a potential general base positioned close to the nucleophilic water, which results in mechanistic issues similar to those in the single-metal mechanism. Furthermore, a mutation of E113K in *Bam*HI kills the enzyme despite the fact that E113 is located in the lysine position of the conserved active site motif (*31*). This divergence of enzyme active sites even between enzymes that belong to a putative (two-metal) subgrouping of type II enzymes illustrates a crucial point. At our current level of understanding, it seems completely possible, if not likely, for type II enzymes to use different catalytic mechanisms. Even enzymes that have both been observed to bind two metal ions might differ in their reaction mechanism and reaction coordinate. For example, *Bam*HI (with a putative general base) might proceed along a fully associative path, while *Pvu*II (that lacks a putative general base) could catalyze the reaction through a more dissociative path.

Homing Endonucleases

Homing endonucleases are proteins encoded by open reading frames contained within mobile introns (*32*−*34*). The first homing endonuclease was discovered in yeast, but they are now known to exist in all branches of life. Homologous endonucleases are also found as optional, independently folded domains in self-splicing protein introns termed "inteins". The catalytic activity of these endonucleases promotes the lateral transfer of the intervening sequences by a targeted transposition mechanism termed homing. These enzymes loosely recognize 14−40 DNA base pairs in contrast to the strict recognition of four to eight base pairs exhibited by restriction endonucleases. While type II restriction enzymes are generally intolerant of base substitutions in their recognition sequence, homing endonucleases can accommodate a variety of polymorphisms in their target sites and, in some cases, have regions of their recognition site that are completely degenerate. There are four group I, intron-encoded, homing endonuclease families that are characterized by well-conserved sequence motifs. They are the LAGLIDADG, His-Cys box, H-N-H, and GIY-YIG families. In the following section, we will focus on the catalytic mechanisms of each of these enzyme families.

LAGLIDADG. The LAGLIDADG enzymes comprise the largest homing endonuclease family with many more than 200 known members and contain one or two copies of the conserved motif. Enzymes that contain one LAGLIDADG sequence like I-*Cre*I and I-*Ceu*I are ∼18−25 kDa in mass and form obligate homodimers. Those with two copies in the same chain like PI-*Sce*I and I-*Dmo*I have about twice the mass of their homodimeric cousins and fold into monomers, with two topologically similar domains related by approximate dyad symmetry and connected by a highly variable peptide linker. The first seven residues of the motif form the core of an α-helix which forms the center of the dimer (or pseudodimer) protein interface and are fairly well conserved (Table 2). The last three residues are almost completely conserved and are composed of the metal binding catalytic aspartic acid (sometimes a glutamic acid) flanked by glycine residues which facilitate a tight β-turn structure and terminate the LAGLIDADG helix. Outside this conserved sequence signature, the LAGLIDADG enzymes are widely divergent in sequence. Most of the crystal structures that have been determined have been of apoenzymes (*35*−*38*), so the following description of the catalytic mechanism for LAGLIDADG enzymes is based on the recent high-resolution cocrystal structure between I-*Cre*I and its DNA homing site in both substrate and product complexes and the superposition of other known structures using I-*Cre*I's DNA binding mode (*29*).

Table 2: Summary of the Conserved Motifs for LAGLIDADG Homing Endonucleases

enzyme	LAGLIDAD$_n$G	active site
I-*Cre*I	LAGFVDGD$_{20}$G	K_{98} R_{51} E_{47}
I-*Dmo*I	LLGLIIGD$_{21}$G	K_{120} K_{43} Q_{42}
	IKGLYVAE$_{117}$G	− K_{130} N_{129}
PI-*Sce*I	LLGLWIGD$_{218}$G	K_{301} R_{231} D_{229}
	LAGLIDSD$_{326}$G	K_{403} H_{343} T_{341}
I-*Pfu*I	LAGFIAGD$_{149}$G	L_{220} − D_{173}
	IAGLFDAE$_{250}$G	K_{322} − M_{263}

The I-*Cre*I active site appears to conform to a canonical two-metal mechanism similar to those described above. One metal positions and activates the nucleophilic water, and the second metal stabilizes both the transition state phosphoanion and the oxygen leaving group. The most striking and unique feature of this active site is that the second metal is used jointly by both active sites. This results in three metals shared between two active sites in a relatively small space (Figure 4). They are coordinated by the two aspartic acids, the oxygens on the scissile phosphates, the conserved glycine's backbone carbonyl, and water molecules (two in the substrate complex and one in the product complex). In the substrate structure, a metal-bound water appears to be well-positioned for nucleophilic attack; however, there are no direct protein contacts to this water, which once again begs the question of what is functioning as the general base. Instead of a protein contact, there is a well-ordered network of water molecules such that hydrogen-bonded water molecules link the nucleophilic water to the leaving group oxygen or to bulk solvent. This raises the interesting possibility that the reaction proceeds by the concerted transfer of hydrogen atoms along this chain of water molecules and that these exchanges are responsible for both nucleophilic activation and leaving group protonation. Another possibility that arose in the discussion of type II restriction endonucleases is that catalysis is substrate-assisted. The substrate-assisted mechanism assigns a nonbridging phosphate oxygen as the general base. The pK_a for the first ionization of phosphoric acid is ∼2 so it does not initially seem like a good candidate, but the pK_a of H_3O^+ is −1.7 so it might be a better base than nearby water molecules. If the general base oxygen is on the scissile phosphate, a further advantage for the substrate-assisted mechanism is that protonation of the phosphate oxygen neutralizes the charge repulsion between the phosphate and the negatively charged nucleophilic water. The p21^{ras} GTPase is similar to I-*Cre*I (and some restriction endonucleases as we will discuss later) in that the general base had not been located by mutagenesis studies and is not clearly defined by the structure of the active site. Molecular dynamics simulations (*39*) and linear free energy relationships (*40*) have indicated that substrate-assisted catalysis is used by p21^{ras} and perhaps other GTPases.

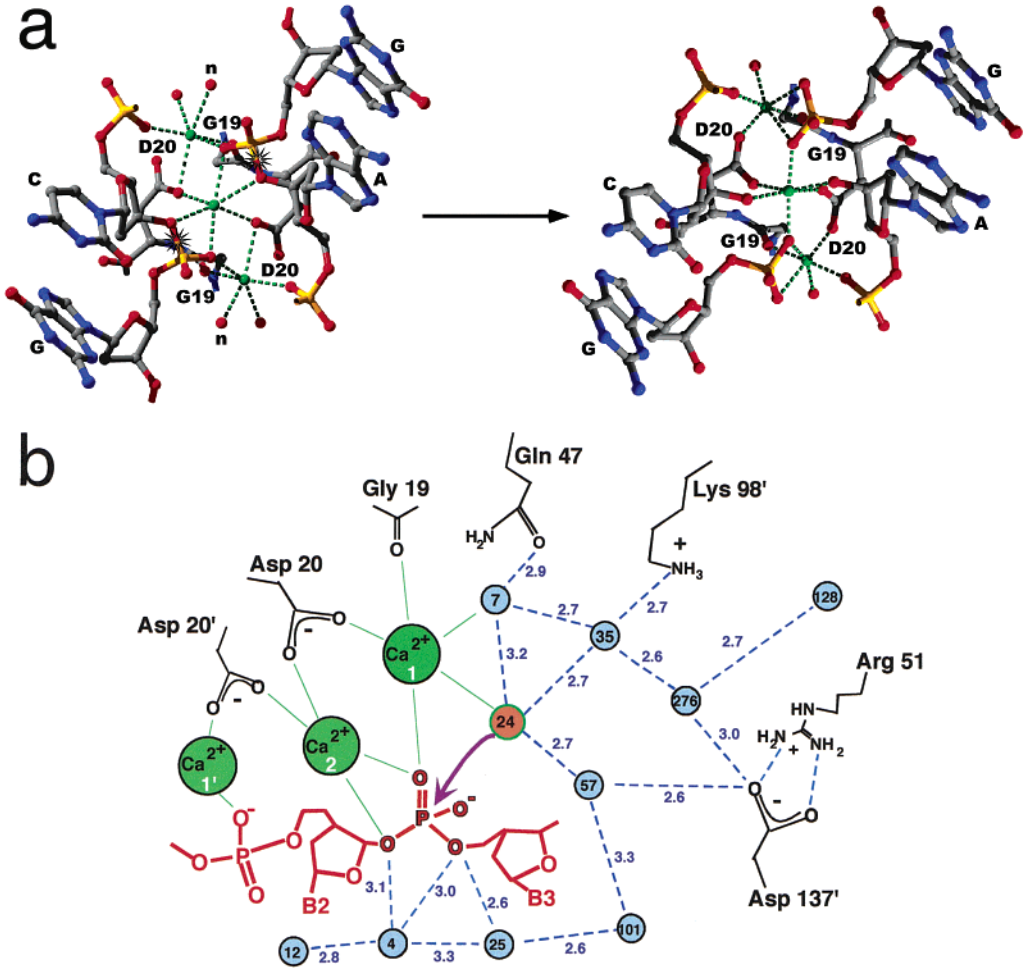

FIGURE 4: (a) Ball and stick representation of the active site of I-*Cre*I in both substrate (left) and product (right) complexes. The coloring and labeling scheme is the same as in Figure 2. (b) Cartoon of the mechanism of nucleophilic attack proposd for I-*Cre*I and homologous enzymes. Asp 20 and 20′ are the same as D20 in panel a above. A solvent pocket described in the text is lined with a series of nonconserved residues (Gln 47, Lys 98, and Arg 51) as listed in Table 2.

In either the water network or the substrate-assisted mechanism, the environment created by the protein−DNA complex must stabilize any high-energy intermediates along the reaction coordinate. Three other residues in I-*Cre*I that have been shown to be important for catalysis likely contribute to this stabilization: lysine 98, arginine 51, and glutamine 47 (*41*). Their corresponding residues in I-*Ceu*I have also been shown to be required for activity (*42*). However, these residues are only moderately conserved in the other LAGLIDADG enzymes (*34*) (Table 2). The two copies of K98 in I-*Cre*I are K120 and nothing in I-*Dmo*I, and L220 and K322 in PI-*Pfu*I. The two copies of R51 in I-*Cre*I are nonexistent in PI-*Pfu*I, and R231 and H343 in PI-*Sce*I. Finally, the two copies of Q47 in I-*Cre*I are D173 and M263 in PI-*Pfu*I, and D229 and T341 in PI-*Sce*I. One aspect of these residues that is well-conserved is that most have the capacity to either donate or accept one or more hydrogen bonds. Another conserved aspect is the presence of basic residues. In the complex between I-*Cre*I and DNA, these residues appear to be responsible for ordering the network of water molecules that surrounds the scissile bond. It is possible that these peripheral active site residues are responsible for positioning and polarizing this water network in an optimal catalytic configuration. The lack of strong conservation of these residues among family members could indicate that there are many different combinations and configurations of chemical groups that are able to fulfill this role. Alternatively, it is still not clear that all LAGLIDADG enzymes utilize the same endonucleolytic mechanism. Different LAGLIDADG enzymes might use different catalytic strategies for which I-*Cre*I is not a viable model. I-*Cre*I is a perfect example of an enzyme for which there are biochemical data on active site mutants, both substrate and product complexes have been determined to high resolution, and yet the precise chemical mechanism is still a matter of debate. We will revisit this point at the end of the review where we discuss possible future directions in the study of the endonuclease mechanism.

His-Cys Box. The His-Cys box homing endonuclease family is defined by two clusters of histidine and cysteine that form structural zinc binding sites. Zinc is required for protein folding, but does not directly play a role in catalysis. I-*Nja*I (*43*), I-*Nan*I, and I-*Nit*I (*44*) are isoschizomeric members of this family, but the most well studied member of the His-Cys box family is I-*Ppo*I (*21*, *45*−*47*). The catalytic mechanism of I-*Ppo*I has been studied structurally (*48*) and biochemically (*49*, *50*), and both studies support similar conclusions. I-*Ppo*I uses a single-metal ion mechanism in which a magnesium ion accelerates the reaction by stabilizing the phosphoanion transition state, positioning and activating a water molecule to donate a proton to the leaving group oxygen, and introducing strain in the substrate complex

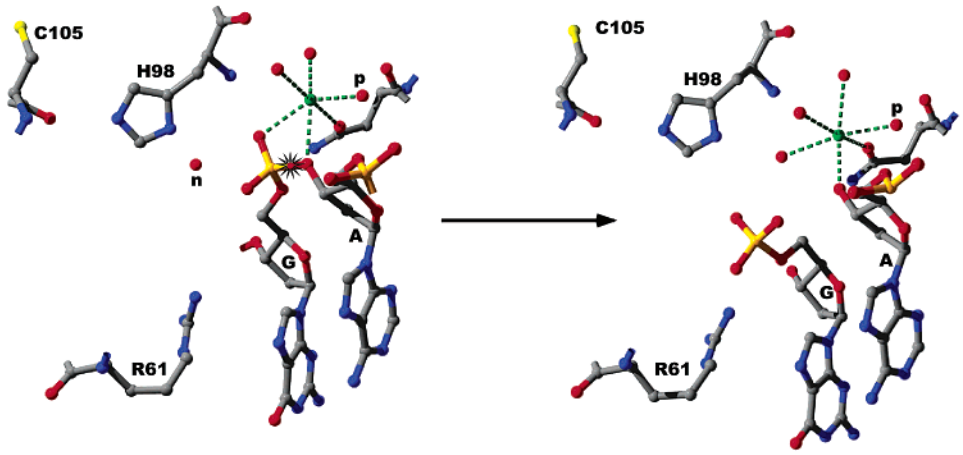

FIGURE 5: Ball and stick representation of the active site of I-*Ppo*I in both substrate (left) and product (right) complexes. The coloring and labeling scheme is the same as in Figure 2.

that is relieved in the product complex. The bound metal is coordinated by asparagine 119, the leaving group oxygen, a nonbridging oxygen on the scissile phosphate, and three water molecules in the substrate complex (Figure 5). In the product complex, the now terminal 5′-phosphate moves away from the metal and toward arginine 61. The leaving group is stabilized by the metal, and a fourth water molecule now replaces the nonbridging oxygen as the sixth metal ligand. The nucleophilic water has been observed in the crystal structures and is positioned nicely for in-line attack. Histidine 98 is the general base, and a hydrogen bond is observed between the nucleophile and δN of H98. The backbone carbonyl of a zinc-coordinating cysteine (C105) stabilizes the correct tautomer of H98 by hydrogen bonding to the ϵNH group. Mutational analysis has indicated that H98A eliminates nearly all activity, N119A and R61A depress activity by 3−4 orders of magnitude, while H78A and H101A decrease activity by a factor of ~10 (*49*).

This class of enzymes is novel in their use of a histidine residue as the general base in a DNA endonuclease reaction. Furthermore, it is noteworthy that the nucleophilic water is not a metal ligand and so is not being activated directly by the metal ion. The extent of the interaction between the scissile phosphate and R61 increases significantly in the product so it is hypothesized that R61 acts to stabilize the transition state relative to the substrate to accelerate the forward reaction and the product relative to the transition state to decelerate the reverse reaction, but it is not known what this interaction is like in the transition state. Also, it is not clear what the precise roles of H78 and H101 are. Interestingly, the nonspecific *Serratia* nuclease uses the exact same active site as I-*Ppo*I (*50*−*52*). *Serratia* nuclease is able to cleave both DNA and RNA and is the most well-studied member of a family of nonspecific nucleases of which NucA is another structurally characterized member. Homology modeling has shown that NucA shares the same active site architecture with I-*Ppo*I, but the arginine (R61) that is conserved in *Serratia* (R57) is an aspartic acid (D95) in NucA (*53*). The substitution of the positively charged arginine with a negatively charged aspartate is in conflict with the proposed mechanism, and the authors have proposed the asparate binds a second metal ion, thus reconstituting the electrophilic nature of that position in the active site in NucA.

H-N-H. The H-N-H family of endonucleases is characterized by a defining motif that spans 25 residues of well-conserved sequence. The homing endonucleases I-*Cmoe*I, I-*Tev*III, I-*Hmu*I, and I-*Hmu*II (*54*, *55*) are members of the H-N-H family, and I-*Cmoe*I has been well-characterized biochemically (*56*). No H-N-H homing endonuclease structure has been determined. However, the H-N-H motif has been structurally characterized by studies on the bacterial colicins E7 and E9. Bacterial colicins are nonspecific endonucleolytic weapons used by bacteria to compete with one another for limited resources. The endonucleases are expressed by a bacterium that also expresses a gene encoding an "immunity protein" that binds to the colicin and inactivates it. Once the endonuclease is outside the cell, it loses the immunity protein as it is actively transported through a target cell's membrane and becomes active. The structures of colicins E7 and E9 have both been determined in complex with their respective immunity proteins (Im7 and Im9) (*57*, *58*). Interestingly, the immunity proteins do not inactivate the colicins by directly occluding the active sites, but are believed to prevent DNA from binding productively to the colicin (*58*). The H-N-H motif folds into a conserved structural motif that has been called the $\beta\beta\alpha$-Me motif (*59*). It is composed of two antiparallel β-strands followed by an α-helix and is stabilized by the binding of a metal ion between the first strand and the helix. This active site structural motif is also observed in the His-Cys box family, the *Serratia* nuclease family, and the T4 endonuclease VII (*60*). The His-Cys box enzymes, T4 endonuclease, and colicins have very different tertiary structures overall, so the reoccurrence of the $\beta\beta\alpha$-Me motif is probably best explained by the convergent evolution of enzyme active sites between evolutionarily unrelated families. The conservation of the $\beta\beta\alpha$-Me motif also suggests the catalytic mechanisms for these endonucleases might be related.

None of the H-N-H enzymes have been studied crystallographically in DNA complexes, but the combination of biochemical results and the structures of the active sites under several different conditions allow for the formation of hypotheses regarding the mechanism of phosphodiester hydrolysis used by the H-N-H family of enzymes. The colicin structures have been determined under three different conditions. The active site of colicin E9 has been determined with nickel bound and with no metal bound (Figure 6), while the

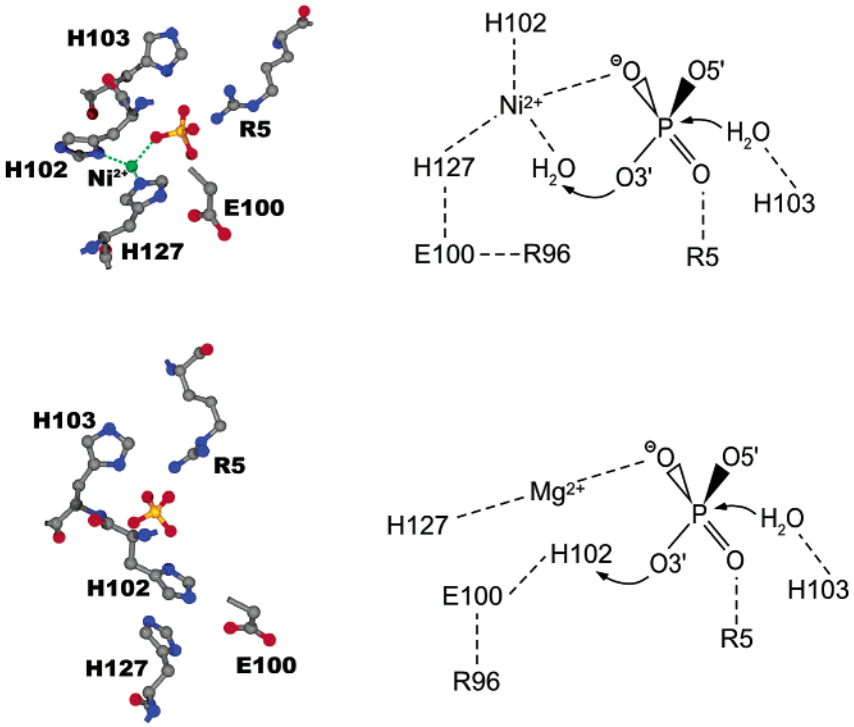

FIGURE 6: Ball and stick representations of two active site structures from bacterial colicin E9 (left). The top structure shows the active site with a bound nickel ion and phosphate (PDB entry 1BXI), while the lower structure contains only a bound phosphate (PDB entry 1EVW). Potential catalytic mechanisms are also shown in schematic form (right). The top scheme is proposed for nickel-supported catalysis and is based on the nickel- and phosphate-bound structure, while the bottom scheme is proposed for magnesium-supported catalysis and is based on the structure of the active site with no metal bound.

active site of colicin E7 has been visualized with a bound zinc ion (*57*, *58*). In both metal-bound structures, the cations are tetrahedrally coordinated. The zinc structure clearly shows the metal coordinated by three histidine residues (H544, H569, and H573) and a water molecule. The nickel structure revealed the metal coordinated by two histidine residues (H102 and H127) and an oxygen from a bound phosphate molecule. NMR experiments verified the fourth ligand as a third histidine residue (H131) as seen in the zinc structure (*61*). The phosphate in the nickel structure is also contacted by a fourth histidine residue (H103). R96, E100, and H127 have all been shown by mutagenesis experiments to be important for catalysis (*62*). These residues form a triad in both colicin structures with E100 in the middle forming a salt bridge with R96 and accepting a hydrogen bond from H127.

Activity assays have revealed that while zinc is able to stabilize the colicins against proteolytic digestion, the enzymes are not active against DNA (*63*). The colicins have displayed activity against double-stranded DNA, single-stranded DNA, and single-stranded RNA (*63*). Magnesium is preferred for dsDNA cleavage, while nickel is favored over magnesium for ssDNA cleavage. Interestingly, the colicins have been shown to cleave ssRNA with no requirement for metal cofactors, although nickel does accelerate the reaction. In the metal free active site, putative catalytic residues are observed to rearrange and take on different conformations. This brings up the intriguing possibility that the basic H-N-H catalytic motif is highly adaptable and capable of taking on slightly different configurations depending on the enzyme and substrate (*63*).

On the basis of these experimental observations, researchers have proposed two possible mechanisms (*63*) (Figure 6). They both describe the activation of the nucleophilic water by the general base H103 and the stabilization of the negative charge of the transition state by a bound metal ion. In the "magnesium" mechanism, magnesium is coordinated by H127 and H102, which are stabilized by the E100−R96 salt bridge. The metal-activated H102 then serves as a proton donor for the leaving group. In the "nickel" mechanism, a nickel-bound water molecule functions as the proton donor. These mechanisms have a great deal in common with the well-characterized mechanisms for I-*Ppo*I and *Serratia* nuclease described above. One significant difference between the experimentally observed colicin structures and those of I-*Ppo*I or *Serratia* nuclease is that the catalytic metal coordination is tetrahedral instead of octahedral, although one large unsettled point is how magnesium will bind in the active site. Since magnesium prefers octahedral coordination spheres, it is possible that in an active complex with magnesium and DNA bound that the active site reorganizes to afford such a coordination. Magnesium also prefers oxygen ligands, so it will be interesting to see if the histidine coordination observed in the current structures is maintained in a magnesium complex. In the structure of T4 endonuclease VII, a calcium ion is predicted to be coordinated tetrahedrally by the oxygen atoms of N62 and D40 and two oxygens from a bound phosphate molecule (*60*).

A recent mutagenic and biochemical study has reinforced the concept that the H-N-H motif is the site of DNA binding, that at least two mechanisms of DNA cleavage are utilized by these enzymes, and that Mg^{2+}-dependent cleavage of double-stranded DNA is responsible for bacterial cell death (*75*). More active site residues are required for catalysis in the presence of Mg^{2+} ions than transition metals, consistent with the hypothesis that the E9 DNase hydrolyzes DNA by

two distinct, cation-dependent catalytic mechanisms. While the structures of DNA-bound complexes of H-N-H and other $\beta\beta\alpha$-Me enzymes are required to verify these conclusions, it seems not only that the H-N-H motif is adaptable but also that the $\beta\beta\alpha$-Me scaffolding itself is adaptable and is used in a variety of ways by different classes of nucleases.

GIY-YIG. The fourth family of homing endonucleases is the GIY-YIG family represented by I-*Tev*I (*64*−*66*) and I-*Bmo*I (*67*). These enzymes have an N-terminal catalytic domain that contains the GIY-YIG sequence motif and a C-terminal DNA binding domain. The binding domain−DNA cocrystal structure of I-*Tev*I has been determined by crystallography (*66*). The binding domain is a combination of helix−turn−helix, α-helix, and Zn-finger DNA binding subdomains, although the Zn finger has recently been shown to have a novel function as a distance determinant for cleavage (*68*). The catalytic domain has been studied by NMR (*69*) and crystallography (P. Van Roey, M. Belfort, and V. Derbyshire, personal communication). It consists of conserved GIY and YIG triplets separated by 11 amino acids and folds into a novel α/β-fold with a three-stranded antiparallel β-sheet core. The primary role of the conserved motif appears to be structural, but there is no clear identification of an active site. Residues that have been implicated by mutagenesis to be catalytic (R27 and E75) (*64*) are separated by at least 6.5 Å and do not form a well-defined active site. It is also not clear how a single catalytic domain of I-*Tev*I performs the sequential nicking that generates the double-strand break. Perhaps a transient protein dimer is formed on DNA and generates two competent active sites. Alternatively, the catalytic domain of I-*Tev*I could catalyze both strand cleavages in series. Clearly, there are still many unknowns regarding the catalytic mechanism of this family. The structure of the full-length enzyme bound to DNA should go a long way toward elucidating what appears to be a novel endonuclease mechanism.

Future Prospects

To provide the best possible description of endonuclease catalytic mechanisms, each enzyme must be studied with a combination of crystallographic and biochemical techniques. The structures of a particular enzyme−DNA complex in a product complex and preferably many substrate complexes that have been trapped via different methods (mutation, thiol substitution, Ca^{2+} ion, etc.) must be solved at resolutions that are sufficient for accurate determination of the positions of metal ions and solvent molecules. Single-turnover activity assays must be carried out with active site mutants, with metal substitutions, and across a wide range of pH values to assay the effects on the rate of the chemical step. Through the application of these multiple experimental techniques, a particular enzyme's catalytic mechanism can be determined. However, as described above, even after a variety of experimental results have been gathered, the exact molecular mechanism of catalysis often remains elusive.

A way to test mechanistic proposals that is beginning to come of age is the computational modeling of reaction paths. These approaches can test the validity of a particular mechanism by comparing the experimentally derived rates and crystal structures to those predicted by the calculation of energies along a proposed path (*70*, *71*). New techniques under development may eventually be able to determine the best possible reaction coordinate for a catalyzed reaction (without having to define a small subset of important coordinates) given the crystallographically determined substrate and product complex structures (*72*). Calculations with modified active site components will give an estimate for the energetic contributions of these components and allow researchers to tease apart the potentially multiple functions of individual groups. Combined with biochemical experiments and the structural determination of trapped substrate, intermediate, and product complexes, these rapidly developing computational methods should bring us closer to a complete understanding of how endonucleases and other enzymes work.

Summary

We have described the catalytic mechanisms of type II restriction and group I intron-encoded homing endonucleases. All of the enzymes that are discussed herein must supply the necessary catalytic elements for DNA cleavage: a properly positioned nucleophilic water molecule, a general base to activate the nucleophile, a general acid to protonate the leaving group, and localized positive charge to stabilize the phosphoanion transition state. In some cases, all of these elements have been determined, while in other systems, the identity of only a subset are agreed upon. Interestingly, mechanisms of endonucleolytic catalysis cannot be segregated according to the pathway function of a particular enzyme (i.e., restriction or homing). For example, the apparent catalytic mechanism used by I-*Cre*I is more related to the mechanisms that are used by type II restriction enzymes than to the His-Cys box homing endonuclease, I-*Ppo*I. Much has been learned about the specific catalytic mechanisms of both restriction and homing endonucleases, but there is still much that is unknown. Open issues include the mechanisms of DNA cleavage used by H-N-H and GIY-GIY homing endonucleases, whether all endonucleases that share active site motifs use the same mechanism, and the quantitative energetic contributions to catalysis made by individual components of an active endonucleolytic complex along the reaction coordinate. The continued study of endonuclease mechanisms by biochemical, structural, and computational methods should reveal answers to these questions.

REFERENCES

1. Westheimer, F. H. (1987) *Science 235*, 1173−1178.
2. Fersht, A. (1999) *Structure and Mechanism in Protein Science: a guide to enzyme catalysis and protein folding*, W. H. Freeman and Co., New York.
3. Silverman, R. B. (2000) *The organic chemistry of enzyme-catalyzed reactions*, Academic Press, San Diego.
4. Rupert, P. B., and Ferre-D'Amare, A. R. (2001) *Nature 410*, 780−786.
5. Murray, J. B., and Scott, W. G. (2000) *J. Mol. Biol. 296*, 33−41.
6. McClarin, J. A., Frederick, C. A., Wang, B. C., Greene, P., Boyer, H. W., Grable, J., and Rosenberg, J. M. (1986) *Science 234*, 1526−1541.
7. Winkler, F. K., Banner, D. W., Oefner, C., Tsernoglou, D., Brown, R. S., Heathman, S. P., Bryan, R. K., Martin, P. D., Petratos, K., and Wilson, K. S. (1993) *EMBO J. 12*, 1781−1795.
8. Grazulis, S., Deibert, M., Rimseliene, R., Skirgaila, R., Sasnauskas, G., Lagunavicius, A., Repin, V., Urbanke, C., Huber, R., and Siksnys, V. (2002) *Nucleic Acids Res. 30*, 876−885.

9. Cheng, X., Balendiran, K., Schildkraut, I., and Anderson, J. E. (1994) *EMBO J. 13*, 3927−3935.
10. Newman, M., Strzelecka, T., Dorner, L. F., Schildkraut, I., and Aggarwal, A. K. (1994) *Structure 2*, 439−452.
11. Wah, D. A., Hirsch, J. A., Dorner, L. F., Schildkraut, I., and Aggarwal, A. K. (1997) *Nature 388*, 97−100.
12. Bozic, D., Grazulis, S., Siksnys, V., and Huber, R. (1996) *J. Mol. Biol. 255*, 176−186.
13. Newman, M., Lunnen, K., Wilson, G., Greci, J., Schildkraut, I., and Phillips, S. E. (1998) *EMBO J. 17*, 5466−5476.
14. Deibert, M., Grazulis, S., Janulaitis, A., Siksnys, V., and Huber, R. (1999) *EMBO J. 18*, 5805−5816.
15. Deibert, M., Grazulis, S., Sasnauskas, G., Siksnys, V., and Huber, R. (2000) *Nat. Struct. Biol. 7*, 792−799.
16. Lukacs, C. M., Kucera, R., Schildkraut, I., and Aggarwal, A. K. (2000) *Nat. Struct. Biol. 7*, 134−140.
17. Huai, Q., Colandene, J. D., Topal, M. D., and Ke, H. (2001) *Nat. Struct. Biol. 8*, 665−669.
18. van der Woerd, M. J., Pelletier, J. J., Xu, S., and Friedman, A. M. (2001) *Structure 9*, 133−144.
19. Pingoud, A., and Jeltsch, A. (2001) *Nucleic Acids Res. 29*, 3705−3727.
20. Wittmayer, P. K., and Raines, R. T. (1996) *Biochemistry 35*, 1076−1083.
21. Flick, K. E., Jurica, M. S., Monnat, R. J., Jr., and Stoddard, B. L. (1998) *Nature 394*, 96−101.
22. Viadiu, H., and Aggarwal, A. K. (1998) *Nat. Struct. Biol. 5*, 910−916.
23. Horton, J. R., and Cheng, X. (2000) *J. Mol. Biol. 300*, 1049−1056.
24. Horton, N. C., and Perona, J. J. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 5729−5734.
25. Horton, N. C., Newberry, K. J., and Perona, J. J. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 13489−13494.
26. Kostrewa, D., and Winkler, F. K. (1995) *Biochemistry 34*, 683−696.
27. Martin, A. M., Sam, M. D., Reich, N. O., and Perona, J. J. (1999) *Nat. Struct. Biol. 6*, 269−277.
28. Perona, J. J., and Martin, A. M. (1997) *J. Mol. Biol. 273*, 207−225.
29. Chevalier, B. S., Monnat, R. J., Jr., and Stoddard, B. L. (2001) *Nat. Struct. Biol. 8*, 312−316.
30. Jeltsch, A., Alves, J., Wolfes, H., Maass, G., and Pingoud, A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8499−8503.
31. Xu, S. Y., and Schildkraut, I. (1991) *J. Biol. Chem. 266*, 4425−4429.
32. Belfort, M., and Perlman, P. S. (1995) *J. Biol. Chem. 270*, 30237−30240.
33. Belfort, M., and Roberts, R. J. (1997) *Nucleic Acids Res. 25*, 3379−3388.
34. Chevalier, B. S., and Stoddard, B. L. (2001) *Nucleic Acids Res. 29*, 3757−3774.
35. Duan, X., Gimble, F. S., and Quiocho, F. A. (1997) *Cell 89*, 555−564.
36. Heath, P. J., Stephens, K. M., Monnat, R. J., Jr., and Stoddard, B. L. (1997) *Nat. Struct. Biol. 4*, 468−476.
37. Silva, G. H., Dalgaard, J. Z., Belfort, M., and Van Roey, P. (1999) *J. Mol. Biol. 286*, 1123−1136.
38. Ichiyanagi, K., Ishino, Y., Ariyoshi, M., Komori, K., and Morikawa, K. (2000) *J. Mol. Biol. 300*, 889−901.
39. Schweins, T., Langen, R., and Warshel, A. (1994) *Nat. Struct. Biol. 1*, 476−484.
40. Schweins, T., Geyer, M., Scheffzek, K., Warshel, A., Kalbitzer, H. R., and Wittinghofer, A. (1995) *Nat. Struct. Biol. 2*, 36−44.
41. Seligman, L. M., Stephens, K. M., Savage, J. H., and Monnat, R. J., Jr. (1997) *Genetics 147*, 1653−1664.
42. Turmel, M., Otis, C., Cote, V., and Lemieux, C. (1997) *Nucleic Acids Res. 25*, 2610−2619.
43. Elde, M., Haugen, P., Willassen, N. P., and Johansen, S. (1999) *Eur. J. Biochem. 259*, 281−288.
44. Elde, M., Willassen, N. P., and Johansen, S. (2000) *Eur. J. Biochem. 267*, 7257−7266.
45. Ellison, E. L., and Vogt, V. M. (1993) *Mol. Cell. Biol. 13*, 7531−7539.
46. Argast, G. M., Stephens, K. M., Emond, M. J., and Monnat, R. J., Jr. (1998) *J. Mol. Biol. 280*, 345−353.
47. Wittmayer, P. K., McKenzie, J. L., and Raines, R. T. (1998) *Gene 206*, 11−21.
48. Galburt, E. A., Chevalier, B., Tang, W., Jurica, M. S., Flick, K. E., Monnat, R. J., Jr., and Stoddard, B. L. (1999) *Nat. Struct. Biol. 6*, 1096−1099.
49. Mannino, S. J., Jenkins, C. L., and Raines, R. T. (1999) *Biochemistry 38*, 16178−16186.
50. Friedhoff, P., Franke, I., Krause, K. L., and Pingoud, A. (1999) *FEBS Lett. 443*, 209−214.
51. Friedhoff, P., Kolmes, B., Gimadutdinow, O., Wende, W., Krause, K. L., and Pingoud, A. (1996) *Nucleic Acids Res. 24*, 2632−2639.
52. Miller, M. D., Cai, J., and Krause, K. L. (1999) *J. Mol. Biol. 288*, 975−987.
53. Meiss, G., Gimadutdinow, O., Haberland, B., and Pingoud, A. (2000) *J. Mol. Biol. 297*, 521−534.
54. Eddy, S. R., and Gold, L. (1991) *Genes Dev. 5*, 1032−1041.
55. Goodrich-Blair, H., and Shub, D. A. (1996) *Cell 84*, 211−221.
56. Drouin, M., Lucas, P., Otis, C., Lemieux, C., and Turmel, M. (2000) *Nucleic Acids Res. 28*, 4566−4572.
57. Ko, T. P., Liao, C. C., Ku, W. Y., Chak, K. F., and Yuan, H. S. (1999) *Structure 7*, 91−102.
58. Kleanthous, C., Kuhlmann, U. C., Pommer, A. J., Ferguson, N., Radford, S. E., Moore, G. R., James, R., and Hemmings, A. M. (1999) *Nat. Struct. Biol. 6*, 243−252.
59. Kuhlmann, U. C., Moore, G. R., James, R., Kleanthous, C., and Hemmings, A. M. (1999) *FEBS Lett. 463*, 1−2.
60. Raaijmakers, H., Toro, I., Birkenbihl, R., Kemper, B., and Suck, D. (2001) *J. Mol. Biol. 308*, 311−323.
61. Hannan, J. P., Whittaker, S. B., Davy, S. L., Kuhlmann, U. C., Pommer, A. J., Hemmings, A. M., James, R., Kleanthous, C., and Moore, G. R. (1999) *Protein Sci. 8*, 1711−1713.
62. Garinot-Schneider, C., Pommer, A. J., Moore, G. R., Kleanthous, C., and James, R. (1996) *J. Mol. Biol. 260*, 731−742.
63. Pommer, A. J. (2001) *J. Mol. Biol. 314*, 735−749.
64. Derbyshire, V., Kowalski, J. C., Dansereau, J. T., Hauer, C. R., and Belfort, M. (1997) *J. Mol. Biol. 265*, 494−506.
65. Kowalski, J. C., Belfort, M., Stapleton, M. A., Holpert, M., Dansereau, J. T., Pietrokovski, S., Baxter, S. M., and Derbyshire, V. (1999) *Nucleic Acids Res. 27*, 2115−2125.
66. Van Roey, P., Waddling, C. A., Fox, K. M., Belfort, M., and Derbyshire, V. (2001) *EMBO J. 20*, 3631−3637.
67. Edgell, D. R., and Shub, D. A. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 7898−7903.
68. Dean, A. B., Stanger, M. J., Dansereau, J. T., Van Roey, P., Derbyshire, V., and Belfort, M. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 8554−8561.
69. Van Roey, P., Meehan, L., Kowalski, J. C., Belfort, M., and Derbyshire, V. (2002) *Nat. Struct. Biol. 9* (11) (in press).
70. Glennon, T. M., Villa, J., and Warshel, A. (2000) *Biochemistry 39*, 9641−9651.
71. Dinner, A. R., Blackburn, G. M., and Karplus, M. (2001) *Nature 413*, 752−755.
72. Henkelman, G., Johannesson, G., and Jonsson, H. (2000) in *Progress on Theoretical Chemistry and Physics* (Schwartz, S. D., Ed.) pp 269, Kluwer Academic, Dordrecht, The Netherlands.
73. Kim, Y. C., Grable, J. C., Love, R., Greene, P. J., and Rosenberg, J. M. (1990) *Science 249* (4974), 1307−1309.
74. Horton, N. C., Dorner, L. F., and Perona, J. J. (2002) *Nat. Struct. Biol. 9* (1), 42−47.
75. Walker, D. C., Georgiou, T., Pommer, A. J., Walker, D., Moore, G. R., Kleanthous, C., and James, R. (2002) *Nucleic Acids Res. 30* (14), 3225−3234.

BI020467H