Restriction Endonucleases

Classification, Properties, and Applications

Raymond J. Williams

Abstract

Restriction endonucleases have become a fundamental tool of molecular biology with many commercial vendors and extensive product lines. While a significant amount has been learned about restriction enzyme diversity, genomic organization, and mechanism, these continue to be active areas of research and assist in classification efforts. More recently, one focus has been their exquisite specificity for the proper recognition sequence and the lack of homology among enzymes recognizing the same DNA sequence. Some questions also remain regarding in vivo function. Site-directed mutagenesis and fusion proteins based on known endonucleases show promise for custom-designed cleavage. An understanding of the enzymes and their properties can improve their productive application by maintaining critical digest parameters and enhancing or avoiding alternative activities.

Index Entries: Restriction endonucleases; R/M systems; star activity; single-stranded cleavage; site-specific nickases.

1. Introduction

Restriction endonucleases, which cleave both strands of DNA in a site-specific manner, are a fundamental tool of molecular biology. Discovery of endonucleases began in the 1960s and led to commercial availability in the early 1970s. The number of characterized enzymes continues to grow as does the number of vendors and the size of their product lines. Although many similarities exist among endonucleases in terms of structures, mechanisms, and uses, important differences remain. Now a staple of molecular biology, restriction endonucleases remain an area of active research regarding their cleavage mechanism, in vivo function, evolutionary origins, and as a model for site-specific DNA recognition. New native enzymes continue to be discovered, known enzymes cloned, and new endonuclease activities developed by using protein engineering and fusions to produce novel polypeptides.

1.1. Diversity and In Vivo Function

Although primarily found in bacterial genomes and plasmids, restriction endonucleases also exist in archaea, viruses, and eukaryotes. It is estimated that 1 in 4 bacteria examined contain one or more ***(1)***. *Neisseria* and *Helicobacter pylori* are particularly rich sources for multiple enzymes in a single strain. Respectively, as many as 7 and 14 endonuclease genes have been discovered in individual strains, although some of the genes are not actively expressed ***(2,3)***. Including all types, >3500 restriction enzymes that recognize 259 different DNA sequences are now known. The vast majority of these, approx 3460 enzymes recognizing 234 DNA sequences, are classified as orthodox Type II or Type II subclasses. These are the common tools of molecular biology with more than 500 enzymes comprising over 200 specificities commercially available. In addition, 58 homing endonucleases, so-called because they are encoded by genes that

*Author to whom all correspondence and reprint requests should be addressed: Protein Purification Dept., Promega Corp., 2800 Woods Hollow Road, Madison, WI 53711-5399. E-mail:RWilliam@Promega.com.

are mobile, self-splicing introns or inteins, each with a unique recognition site, have been discovered. A total of 16 site-specific nickases are currently known as well. In all, 297 restriction enzymes have been cloned and sequenced. A database of all known endonucleases is updated monthly by Dr. Richard J. Roberts and Dana Macelis and is available at http://www.neb.com/rebase. A number of formats are available, references given, and statistics maintained.

Restriction endonucleases were originally named for their ability to restrict the growth of phage in a host bacterial cell by cleavage of the invading DNA. In this manner, they may be acting as bacterial protection systems. The DNA of the host is protected from restriction by the activity of a methylase(s), which recognizes the same sequence as the restriction enzyme and methylates a specific nucleotide (4-methylcytosine, 5-methylcytosine, 5-hydroxymethylcytosine, or 6-methyladenine) on each strand within this sequence. Once methylated, the host DNA is no longer a substrate for the endonuclease. Because both strands of the host DNA are methylated and even hemi-methylated DNA is protected, freshly replicated host DNA is not digested by the endonuclease.

The role of restriction endonucleases as a protection system may be oversimplified however. Various characteristics lower an enzyme's protective potential. There would be no effect on phages without at least a dsDNA intermediate or those for whom the DNA was also modified at the critical bases. A small number of phage may be methylated by the host before restriction can occur, and thus be able to propagate protectively methylated copies of themselves. Also, large enzyme recognition sites would tend to be rare in small phage genomes. Restriction site avoidance appears to be more important in a group of bacteria rather than a corresponding group of phage. The endonucleases generally have a longer half-life than the corresponding methylases, a potentially lethal problem for the host if the methylase is not properly maintained. For these reasons, it has also been proposed that restriction-methylase systems may be mobile, selfish genetic elements that become essential for host survival once acquired ***(4,5)***.

2. Nomenclature and Genomic Organization

Individual enzymes are named in accordance with the proposal of Smith and Nathans ***(6)***. Briefly, three letters in italics are derived from the first letter of the genus and the first two letters of the microbial species from which the enzyme was derived. An additional letter without italics may be used to designate a particular strain. This is followed by a roman numeral to signify the first, second, and so on, enzyme discovered from the organism. As may be deduced from the large number of enzymes and the limited number of different DNA sequences they recognize, many enzymes from different biological sources recognize the same DNA sequence and are called isoschizomers. A subset wherein two enzymes recognize the same DNA sequence but cleave at a different position is referred to as neoschizomers.

An important point to emphasize as a result of cloning and sequence comparison is that little if any sequence homology exists between the endonuclease and methyltransferase recognizing the same DNA sequence. Furthermore, even restriction isoschizomers may show little or no homology, including the amino acids involved in recognition, and as such are excellent candidates for a comparative study of protein–DNA interaction. For example, the enzymes *Hha*II and *Hinf*I are both isolated from strains of *Haemophilus*, recognize GANTC, and cleave between the G and A. However, they share only 19% identity in their amino acid sequence ***(7)***. Endonuclease/methylase systems recognizing the same sequence may also exhibit different methylation patterns and restriction sensitivity. Only a limited common amino acid motif, PD...D/EXK, has been proven by mutational or structural analysis to participate in catalysis for 10 endonucleases. However, the 10 enzymes include members that are classified as Type II, IIe, IIs, IV, or intron encoded endonucleases ***(8)***. In contrast, general motifs have been

found for 30 6-methyladenine, 4-methylcytosine, and 5-methylcytosine methylases ***(9)***.

Frequently referred to as an R/M system, the restriction endonuclease and modification methylase genes lie adjacent to each other on bacterial DNA and may be oriented transcriptionally in a convergent, divergent, or sequential manner. The proximity of these genes appears to be universal and is utilized in a common cloning method sometimes referred to as the "Hungarian Trick" ***(10)***. Basically, an endonuclease is used to digest the genomic DNA from the bacteria containing the R/M system of interest and create a library of clones. The expression vector used must contain the recognition site of the R/M system. Purified plasmids from the clones are then subjected to the restriction enzyme of interest in vitro. If a plasmid contains the expressed methylase gene, it will be resistant to cleavage. Often, the endonuclease is expressed as well without the need for subcloning.

It is assumed that methylation must occur before restriction activity to protect the host DNA. One approach bacteria use to limit the possibility of self-restriction is to significantly reduce the number of recognition sites in their genomes. Alternatively, methylase expression may precede that of the endonuclease. One manner in which this may be accomplished is through an open reading frame located upstream of the endonuclease gene encoding a "C" or control protein in some R/M systems. This C protein positively regulates the endonuclease gene and allows for the activity of the constitutively expressed methylase to precede expression of the endonuclease ***(11)***. Such C genes are frequently found in situations where the methylase and endonuclease genes are in divergent or convergent transcriptional orientations. Using cloned R/M systems with disrupted C genes for *Bam*HI, *Sma*I, *Pvu*II, and *Eco*RV R/M, various C genes were provided on a separate plasmid. *Bam*HI restriction activity was equally stimulated by the *Sma*I C and the *Bam*HI C gene and only one order of magnitude less by the *Pvu*II C gene. The *Eco*RV C gene provided no stimulation. The *Bam*HI C gene stimulated *Pvu*II restriction activity as well as the *Pvu*II C gene ***(12)***. Why some C genes stimulate expression of alternate endonucleases is not fully understood, but the phenomenon may have evolutionary implications for R/M systems.

3. Structure, Specificity, and Mechanism

3.1. Classification and General Mechanism

Restriction endonucleases are classified according to their structure, recognition site, cleavage site, cofactor(s), and activator(s). Sets of these criteria are used to define the different types (I, II, III, and IV) and subclasses (IIe, IIf, IIs, etc.), which are explained in detail in **Table 1**. Multiple subunit and holoenzyme assemblies are possible to achieve the needed restriction, methylase, and specificity domains. These three domains may be present on three separate polypeptides, two polypeptides, or a single polypeptide. At a minimum, all R/M systems share an absolute requirement of Mg^{2+} for endonuclease activity and AdoMet (also referred to as S-adenosyl methionine) as the methyl donor for methylase activity. In general, Type I restriction requires Mg^{2+}, AdoMet, and ATP (which becomes hydrolyzed). Type II restriction requires only Mg^{2+}, although a second recognition site or AdoMet may be stimulatory. Type III restriction requires Mg^{2+} and ATP (which is not hydrolyzed) and may be stimulated by AdoMet and a second recognition site. Type IV restriction requires Mg and AdoMet, and also has the unusual property of cleaving both DNA strands on both sides of its recognition site, effectively excising the site. Homing endonucleases are a diverse group with several differences from Types I–IV. It should be noted that *Eco57*I and like enzymes, previously classified as Type IV ***(25,26)***, have been reclassified as Type IIg ***(16)***. In addition, it is newly proposed in this article that the enzymes previously classified Type IIb, including *Bcg*I, *Bsp*24I, *Bae*I, *Cje*I, and *Cje*PI, be moved into the vacated Type IV classification due to their unique properties as stated above.

The majority of recognition sites are four, six, or eight bases long and palindromic. Some enzymes recognize sites with a limited degree of

Table 1
Restriction Enzyme Types and Classification[a]

Type	Example(s)	Subunit Structure of Endonuclease[1]	Cofactors[2] and Activators	Recognition Site	Cleavage Site	Methylase Properties
I (EC 3.1.21.3)	*Eco*KI, *Eco*AI, *Eco*BI, *Cfr*AI, *Sty*SPI, etc.	Usually a pentameric complex (2 R, 2 M, and 1 S)	Mg^{2+}, AdoMet, ATP (hydrolyzed)	Interrupted Bipartite	Distant and variable from recognition site, for example, *Eco*KI: AAC(N_6)GTGC($N_{>400}$)↓ TTG(N_6)CACG($N_{>400}$)↑	May be heterodimer (1 M, 1 S) or heterotrimer (2 M, 1 S)
Orthodox II (EC 3.1.21.4)	*Eco*RI, *Bam*HI, *Hind*III, *Kpn*I, *Not*I, *Pst*I, *Sma*I, *Xho*I, etc.	Homodimer (2 R-S)	Mg^{2+}	Palindromic or interrupted palindrome, ambiguity may be allowed	Defined, within recognition site, may result in a 3' overhang, 5' overhang, or blunt end, for example, *Eco*RI: G↓A A T T C C T T A A↑G	Separate, single, monomeric (M-S) methyltransferase, a few systems contain 2 methyltransferases
IIe [3]	*Nae*I, *Nar*I, *Bsp*MI, *Hpa*II, *Sac*II, *Eco*RII, *Atu*BI, *Cfr*9I, *Sau*BMKI, and *Ksp*632I	Homodimer (2 R-S) or monomer (R-S), similar to Type II or Type IIs	Mg^{2+}, A second recognition site, acting in *cis* or *trans*, binds to the endonuclease as a allosteric affector	Palindromic, palindromic with ambiguities, or nonpalindromic	Cuts in defined manner within the recognition site or a short distance, needs activator DNA containing a recognition site for complete cleavage, for example, *Nae*I: GCC↓GGC CGG↑CCG	Separate, single, monomeric (M-S) methyltransferase
IIf	*Sfi* I, *NgoM* IV, *Cfr* 10 I, *Aat* II	Homotetramer (4 R-S)	Mg^{2+}	Palindromic or interrupted palindrome, 2 cleavable recognition sites must be bound for activity	Defined, within recognition site, may result in a 3' or 5' overhang, for example, *NgoM* IV: G↓C C G G C C G G C C↑G	Separate, single, monomeric (M-S) methyltransferase

IIg (formerly Type IV)	*Eco57I, Bce83I, Hae*IV, *Mme*I, *Bsp*LU11III, *Bse*MII	R-M-S monomer	Mg^{2+}, (AdoMet)*	Nonpalindromic	Cuts in a defined manner a short distance away from recognition site, may not cut to completion, for example, *Eco*57I: CTGAAG(N) $_{16}\downarrow$ GACTTC(N) $_{14}\uparrow$	Separate, single, monomeric (M-S) methyltransferase (methylase activity of restriction monomer only methylates one strand)
IIm	*Dpn*I	Homodimer (2 R-S)	Mg^{2+}	Palindromic	Cuts within the recognition site to leave a blunt end, recognition site must be methylated	None
IIs	*Fok*I, Alw26I, BbvI, *Bsr*I, *Ear*I, *Hph*I, *Mbo*II, *Ple*I, *Sfa*NI, *Tth*111I, etc.	Monomeric (R-S)	Mg^{2+}	Nonpalindromic, nearly always contiguous and without ambiguities	Cuts in defined manner with at least one cleavage site outside of the recognition site, rarely leaves blunt ends, for example, *Fok*I: GGATG(N) $_{9}\downarrow$ CCTAC(N) $_{13}\uparrow$	May be 1 monomeric (M-S) which methylates one or both strands, or 2 separate monomeric (M-S) methyltransferases, one for each strand, may also methylate different nucleotides
IIt	*Bpu*10 I *Bsl* I	Heterodimer (α, β) or Heterotetramer (2 α, 2β)	Mg^{2+}	Interrupted bipartite or interrupted palindrome	Defined, within recognition site or a short distance away, resulting in a 3' overhang, for example, *Bsl* I: C C N N N N N$\downarrow$N N G G G G N N$\uparrow$N N N N N C C	May be 1 monomeric (M-S) which methylates both strands, or 2 separate monomeric (M-S) methyltransferases, one for each strand
III (EC 3.1.21.5)	*Eco*P15I, *Eco*PI, *Hinf*III, and *Sty*LTI	Both R and M-S required	Mg^{2+}, (AdoMet)*, ATP (not hydrolyzed)[4], May require a second unmodified site in opposite orientation, variable distance away[5]	Nonpalindromic	Cuts approx 25 bases away from recognition site, may not cut to completion, for example, *Eco*P15I: CAGCAG(N) $_{25–26}\downarrow$ GTCGTC(N) $_{25–26}\uparrow$	Methylates adenines, only on one strand, in an independent manner

(continued)

Table 1 *(continued)*

Type	Example(s)	Subunit Structure of Endonuclease[1]	Cofactors[2] and Activators	Recognition Site	Cleavage Site	Methylase Properties
IV (formerly Type IIb)	*Bcg*I, *Bsp*24I, *Bae*I, *Cje*I, and *Cje*PI	Heterotrimer[6] (2 R-M, 1 S)	Mg^{2+}, AdoMet	Interrupted bipartite	Cuts both strands on both sides of recognition site a defined, symmetric, short distance away and leaves 3' overhangs, for example, *Bcg*I: ↓ 10(N)CGA$(N)_6$TCG(N)12↓ ↑ $_{12}$(N)GCT$(N)_6$ACG$(N)_{10}$↑	Same heterotrimer (2 R-M, 1 S) only methylates symmetric adenines of recognition site
Intron or Intein encoded	I-*Ppo*I, I-*Ceu*I, I-*Hmu*I I-*Sce*I, I-*Tev*I, PI-*Psp*I, F-*Sce*II, etc.	Monomer, homodimer, other protein or RNA may be required	Mg^{2+}, may also bind Zn^{2+}	12–40 bp, tolerance for base pair substitutions exists	3' and 5' overhangs from 1-10 bases, a few not yet determined, may cleave 1 strand preferentially or in the absence of Mg^{2+} , 2 enzymes cleave only one strand, for example, I-*Ppo*I: C T C T C T T A A↓ G G T A G C G A G A G↑A A T T C C A T C G	none

[a]The first five columns list examples and properties of the restriction endonuclease. The recognition and cleavage site of the first example is given under the column "Cleavage Site." The sequence of the top strand is given from 5' to 3'. Cleavage is indicated by the arrows. The last column refers to the methyltransferase activity. AdoMet, also referred to as S-adenosyl methionine, is always required for methylation. It should be noted that endonucleases previously classified as Type IV have been reclassified as Type IIg ***(16)***. Also, it is proposed for the first time that endonucleases previously classified as Type IIb be moved to the vacated Type IV classification due to their significant differences from other Type II enzymes. Information presented represents the knowledge known to date and future discoveries may provide exceptions.

[*] Components in parentheses stimulate activity but are not required.

For reviews, see the following references: Type I ***(13,14)***, Type II ***(1,15,16)***, Type IIe ***(17,18)***, Type IIs ***(19)***, Type III ***(13)***, Type IV ***(20,21)***, and homing endonucleases (intron or intein encoded) ***(22).***

[1] R, M, and S refer to restriction, methyltransferase, and substrate specificity domains which may exist as separate subunits (R, M, S) or be combined (R-S, M-S, R-M) in a single polypeptide. In the case of Type II systems, the primary sequence of the restriction endonuclease and methyltransferase specificity domains demonstrate little, if any, homology.

[2] Although showing a strong preference for Mg^{2+}, other divalent metals may substitute, usually Mn^{2+} but also Ca^{2+}, Co^{2+}, Fe^{2+}, Ni^{2+}, and Zn^{2+}. However, specificity may be relaxed and cleavage rates significantly decreased.

[3] Many isoschizomers exist which are common Type II. There is evidence to suggest that *Eco*57I could also be classified as Type IIe ***(17)***.

[4] ATPase activity has been previously reported as <1% compared to Type I restriction activity and therefore ATP was regarded as a cofactor rather than a substrate. However, more recent evidence with *Eco*P15I suggests a need to investigate more closely possible ATPase activity of Type III restriction activities ***(23)***.

[5] In the host protection mechanism for *Eco*P15I, DNA is hemi-methylated in the fully protected state and freshly replicated DNA is protected by the fact that a second, convergently orientated, and also totally unmodified site is required for cleavage. This host protection mechanism may be true for the other type III systems as well (*Eco*PI, *Hin*fIII, and *Sty*LTI) ***(23,24)***.

[6] This tertiary structure has only been shown for *Bcg*I while the structures of the other 4 systems of this type (*Bsp*24I, *Bae*I, *Cje*I, and *Cje*PI) are unknown.

ambiguity or those consisting of interrupted palindromes. When the specificity domain allows ambiguities, the possible nucleotide substitutions at a particular position are defined and others are strictly excluded. This results in palindromic and partially palindromic sites that are recognized and cleaved by Type II endonucleases. For example, the recognition site for *Sty*I is listed as CCWWGG. Therefore, the substrate sequences for *Sty*I can be palindromic (CC**TA**GG or CC**AT**GG) or partially palindromic (CC**TT**GG or CC**AA**GG) ***(27)***. This flexibility of recognition is not currently understood. Particularly interesting are the situations where allowed nucleotides can be either purine or pyrimidine or when only a single nucleotide is excluded. The single letter code for these ambiguities is as follows:

R = A or G	Y = C or T	M = A or C
K = G or T	S = G or C	W = A or T
B = not A (C or G or T)		D = not C (A or G or T)
H = not G) (A or C or T)		V = not T (A or C or G)
N = A or C or G or T		

The generalized mechanism for site-specific cleavage of DNA by restriction enzymes involves several steps. First, water begins to be excluded as the enzyme binds to DNA in a nonspecific manner that usually only involves interaction with the phosphate backbone. The enzyme then moves along the DNA by linear diffusion. For *Eco*RV, it has been estimated the enzyme is capable of scanning 2×10^6 base pairs at the rate of 1.7×10^6 bp s^{-1} during one binding event ***(28)***. When the specific recognition site is found, additional water is excluded and hydrogen bonds (typically 15–20) are formed with the recognition site bases in addition to van der Waals base contacts and hydrogen bonds to the backbone ***(16)***. The sequence flanking the recognition site may also influence specific binding. For *Bam*HI, binding increases 5400-fold as oligonucleotide length increases from 10 to 14 bp and varies 30-fold based on the best to worst flanking triplets ***(29)***. Some differences exist as to whether an enzyme binds cognate DNA with this greatly enhanced affinity in the absence of Mg^{2+}. Most enzymes, such as *Eco*RI, are able to bind specifically without Mg^{2+} but do not cleave. Another group binds cognate and noncognate DNA with relatively similar affinity in the absence of a divalent metal cation although some controversy remains regarding its most studied member, *Eco*RV ***(16,29)***. As the specific complex forms, structural shifts occur in both the enzyme and DNA. In one crystallographic study on *Eco*RV, the two DNA-binding/catalytic domains of the enzyme rotated 25° with respect to each other and the cognate DNA was bent 57° and 42° in two differently obtained lattices ***(30)***.

After the specific enzyme:DNA complex is formed in the presence of Mg^{2+}, a specific phosphodiester bond in each strand is cleaved. Briefly, a general base produces a hydroxide ion that acts as a nucleophile to attack the scissile phosphorous. The resulting negatively charged pentacovalent transition state is stabilized by a Lewis acid and a general acid provides the proton for the leaving group, the 3' DNA hydroxyl ***(15)***. A variation of this mechanism may also occur, with water acting as a weaker attacking nucleophile and thereby requiring stronger stabilization of the transition state and leaving group ***(16)***. In either case, the phosphorous retained on the 5' end of the DNA becomes inverted. The cleavage position may generate a blunt end or a single-stranded 3' or 5' overhang of one to four bases. It should be noted that enzymes with ultimately different recognition sites may still produce overhangs that are complementary and therefore suitable for ligation, although the recognition site for one or both enzymes may be lost in the ligation product. For example, NarI, MspI, AcyI, TaqI, ClaI, Csp45I, HpaII, and AccI all produce a 5'-CG overhang, although each has a different recognition sequence. More specific information regarding a few of the enzyme types and subclasses is given below.

3.2. Orthodox Type II Endonucleases

Generally, the common Type II endonucleases are homodimers (most between 25 and 35 kDa for the monomeric subunit), require only Mg^{2+}, and cleave within palindromes, partial palindromes, or interrupted palindromes. Despite dissimilar primary sequence, Type II endonucleases have a

similar 3D structure: a "U" shaped dimeric holoenzyme with each of the identical subunits contributing recognition and catalytic domains on the sides and bridging domains at the bottom. Monomers lack activity by themselves. Crystal structures of Type II enzymes including the different subclasses appear to have a common core consisting of five β-sheets flanked on each side by an α-helix, similar to the enzymes *Mut*H and α-exonuclease ***(31–33)***. It appears Type II enzymes could also be categorized based on additional structural homology between those producing 5' overhangs, which approach DNA from the major groove, and those producing 3' overhangs or blunt ends, which approach DNA from the minor groove. For the known crystal structures, Type II endonucleases producing 5' overhangs appear to use an α-helix to distinguish their specific site and have been tentatively labeled α-helix recognition. For example, *Bam*HI recognizes GGATCC while *Bgl*II recognizes the closely related site AGATCT. Both generate the same 5', four base overhang of GATC. However, the protein-base contacts for the common internal four bases and the distortion of the DNA in the specific complex are significantly different for the two enzymes ***(34)***. Conversely, the 3' and blunt end producing enzymes seem to rely on a β-strand and are tentatively labeled as β-strand recognition. There are differences in the polarity of the β-sheets between the two groups as well ***(35)***.

3.3. Type IIe Endonucleases

The Type IIe endonucleases are similar to the common Type II or Type IIs in their structure, recognition patterns, and mechanism. However, they are distinct in being activated to cleave slow or resistant sites by the binding of a second recognition sequence to a distal, noncatalytic site on the enzyme. Typically, these enzymes cleave incompletely at a subset of recognition sites. Isoschizomers of Type IIe endonucleases cleave completely. For *Eco*RII and pBR322 (six recognition sites per DNA molecule) the ratio of enzyme to recognition sites in a reaction mix for optimal activity is 0.25–0.5 or two to four recognition sites per enzyme dimer. This suggests each enzyme dimer binds a recognition sequence at its catalytic site and a second at the allosteric site ***(36)***. Based on observed cleavage at particular sites, an original classification of Type IIe endonuclease activity in a 1 h digest is as follows: cleavable sites, >90% cleavage with 1- to 5-fold excess enzyme; slow sites, 5–90% cleavage with 5-fold excess enzyme and additional cleavage with 10- to 30-fold excess; resistant sites, <5% cleavage with fivefold excess enzyme and no additional cleavage with a 10–30-fold excess. A Type IIe enzyme may cleave one DNA site slowly while another site in the same or on a different DNA molecule is resistant to cleavage ***(18)***.

The Type IIe enzymes can be separated into two classes in a more descriptive manner based on the change in cleavage kinetics upon binding of an affector sequence, which may be an oligonucleotide, linear phage, or supercoiled DNA. In the K class of enzymes (*Nar*I, *Hpa*II, *Sac*II), activator DNA binding decreases the K_m without altering the V_{max} of cleavage, indicating that cooperative binding induces a conformational shift that increases the affinity of the enzyme for the substrate. In the V class (*Nae*I, *Bsp*MI), binding of activator DNA increases V_{max} without changing K_m, indicating that the increased catalytic activity is not related to the affinity of the enzyme for substrate ***(18)***. It is assumed that the cleavage kinetics of different recognition sites is influenced by the flanking sequences for Type IIe enzymes. The flanking sequence preferences are not presently understood. However, sequences including a readily cleaved site and its flanking regions are a starting point to determine good activator sequences.

The incomplete digestion by Type IIe enzymes that often occurs can make interpretation of banding patterns and subsequent applications difficult. Adding activators may improve cleavage. For example, oligos containing the recognition site for *Eco*RII that are uncleavable due to specific methylation or the presence of nucleotide analogs, can bind to the allosteric site and stimulate cleavage of refractory sites in pBR322 ***(37)***. A similar

approach, developed by Topal and coworkers, used an oligonucleotide containing the recognition site for *Nae*I with a phosphorothioate at the scissile bond ***(38)***. Complete substrate cleavage is achieved without consuming the activator oligonucleotide as the sulfur prevents hydrolysis by *Nae*I. The same strategy has also been used successfully for *Nar*I, demonstrating utility of this approach for both V and K class enzymes. Some of these enzymes are available commercially with the activating oligo premixed in the provided reaction buffer (e.g., Promega's Turbo™ *Nae*I and Turbo™ *Nar*I). The presence of the oligo does not interfere with ligation or random primer labeling and a one-step purification yields a cleaved DNA suitable for end labeling ***(39)***.

*Nae*I contains a slight variant, TD...DCK, of the endonuclease motif in the N-terminal region and a 10 amino acid motif, ^{39}TLDQLYDGQR^{48} in the N-terminal region similar to a motif in human DNA ligase I ***(35)***. The leucine at position 43 in *Nae*I is a lysine in the ligase motif that is involved in the adenylated intermediate and is essential for ligation. A mutant of the endonuclease, *Nae*I L43K, exhibits type I topoisomerase activity (cleavage, strand passage, and reunion). This suggests a possible origin for the activator DNA binding site in the C-terminal region and a potential link between this endonuclease and topoisomerases and recombinases ***(40,41)***. In addition, based on mutational analysis, it has been proposed that residues 182–192 are involved in communication between the endonuclease and topoisomerase (*Nae*I L43K) or activator DNA (*Nae*I) domains ***(42)***.

3.4. Type IIs Endonucleases

Type IIs endonucleases are monomeric, 45–110 kD, require only Mg^{2+}, recognize non-palindromic sequences, and cleave at least one of the two strands outside the recognition site. The majority of structural information available for these endonucleases is based on the crystal structure of one member, *Fok*I, bound to DNA. The amino terminal portion contains the DNA recognition domain and the carboxy terminal portion contains the cleavage domain. In the absence of Mg^{2+}, the crystal structure of *Fok*I bound to a 20 bp fragment containing its recognition site revealed two apparent anomalies. First, the cleavage domain was not in contact with the cleavage site. This observation has also been substantiated by footprinting studies. The cleavage domain is positioned away from the DNA while the enzyme searches for its recognition site. When bound to its site, and in the presence of Mg^{2+}, the *Fok*I cleavage domain swings into an active position through a series of intramolecular shifts ***(43)***. However, there is only a single cleavage domain per monomer. In order to cleave both strands, the next step involves transient dimerization of the catalytic domain of a second monomer at the cleavage site. Structural similarity to the catalytic and bridging domains of the homodimeric Type II enzyme *Bam*HI further substantiates this model although the dimer interface is smaller for *Fok*I, supporting its existence as a monomer in free solution ***(44)***. It has also been found that the second *Fok*I molecule must also be bound to cognate DNA for cleavage of the initial substrate. At this time it is not known if the second DNA duplex is parallel to the first, which would allow protein–protein interaction and stabilization or antiparallel, which places the protein molecules in a more symmetrical orientation ***(45)***. Sequestering the nonspecific cleavage domain and requiring multiple, specific conditions to be met before catalytic activation is likely important for maintaining a degree of fidelity similar to that of other Type II enzymes.

3.5. Type IIf, IIg, and IIt Endonucleases

Type IIf endonucleases are similar to orthodox Type II in most respects. The two differences are that they exist as homotetramers, two typical dimers in a back-to-back orientation, and that cognate DNA must be bound to both catalytic clefts for cleavage to occur. Examples of this subclass are *Sfi*I ***(46)***, *Aat*II ***(47)***, *Cfr*10I ***(48)*** and *Ngo*MIV ***(49)***. Because they need two copies of their recognition site for cleavage, Type IIf enzymes are similar to Type IIe enzymes in that hydrolysis of the last few sites in a reaction can be problematic even when the enzyme is in excess relative to sub-

strate. For *Sfi*I, it has been shown that the homotetramer must interact with two intact recognition sites containing cleavable phosphodiester bonds as opposed to one DNA segment containing a nonhydrolyzed phosphorothioate as in the activator sequences explained for Type IIe enzymes ***(50)***. Owing to higher effective concentration, having the two sites in *cis* rather than *trans* is preferred and both sites are cleaved in a single turnover ***(51,52)***. An additional observation regarding the interrupted palindrome recognized by *Sfi* I is a 70-fold difference in reaction rate based on the spacer sequence, which contains the scissile phosphates. It has been proposed that a certain amount of initial DNA rigidity imposed by the spacer sequence results in additional backbone strain after enzyme-induced bending, which contributes to catalysis ***(53)***.

Type IIg endonucleases were previously classified as Type IV ***(15,25,26)*** but have recently been reassigned based on the only absolute requirement for cleavage being Mg^{2+}, although AdoMet is stimulatory ***(16)***. Additional enzymatic properties are also shared with other Type II endonucleases. Cleavage outside nonpalindromic recognition sites mimics Type IIs enzymes. Reactions may not proceed to completion similar to Types IIe and IIf. In addition to the contribution of AdoMet, Type IIg is distinguished by its founding member, *Eco*57I, which exists as a monomer containing recognition, cleavage, and methylase activities. A gene expressing a separate methylase exists as well ***(25)***.

A relatively new subclass containing the enzymes *Bpu*10I and *Bsl*I has been designated Type IIt ***(16)***. Although both have interrupted recognition sites and cleave within the nonspecified region, the *Bpu*10I site is non-palindromic and the *Bsl*I site is palindromic. The defining characteristic of Type IIt restriction is the requirement for both α and β polypeptides. The association between subunits for *Bpu*10I appears to be weak as they separate easily during purification and require reconstitution for activity ***(54)***. In studies with *Bsl*I, DNA mobility shifts occur only with subunit mixtures and the cloned α and β genes can be singly expressed in the absence of methylase without killing the host. It has been proposed that the active form is a $\alpha_2\beta_2$ heterotetramer although heterodimers and oligomers also exist in solution ***(55)***.

3.6 Type IV Endonucleases

Type IV endonucleases were previously classified as Type IIb ***(15,16,26)***. It is newly proposed in this article to move them into the recently vacated classification of Type IV as they require AdoMet as well as Mg^{2+} for restriction activity. However, the most unique characteristic of this group is the cleavage of four DNA strands, a double stranded break on both sides of their recognition sites, resulting in excision of the site. The subunit assembly and the relationship between restriction and methylation are unmatched as well. The only holoenzyme model proposed thus far is for *Bcg*I and this is not yet based on crystallographic data. In solution, the molecular weight determined by gel filtration suggests a heterohexamer consisting of two identical working units, each of which is capable of binding a separate recognition site ***(20)***. The working unit of this model, derived from sequence motifs, mutational and truncation analysis, and subunit stoichiometry, is a heterotrimer consisting of one specificity polypeptide plus two identical polypeptides containing restriction and methylation domains. The restriction-methylation subunits are bound one on each side of the specificity subunit, positioning them both upstream and downstream of the recognition site. Double-stranded cleavage by both restriction-methylation subunits of each heterotrimer thereby excises its recognition site. Substrates containing a single site are cleaved at a much lower rate than those with two, suggesting that both recognition domains of the complete heterohexamer must be occupied ***(56)***. A host recognition site that is hemi-methylated, such as after recent replication, is preferentially methylated on the other strand rather than restricted. Conversely, a recognition site unmethylated on both strands, such as foreign DNA, is cleaved ***(57)***.

3.7. Homing Endonucleases

The homing endonucleases, sometimes referred to as intron and intein (protein intron) encoded

endonucleases, are different from the standard restriction enzymes in several respects. They may be monomers or dimers and may require other proteins or RNA for activity. They tolerate some base substitutions in their large recognition sequences, especially the outside regions, with only small changes in cleavage rates. They may also retain significant activity and relative fidelity after substituting other divalent metals for Mg^{2+}. For one member, I-*Ppo*I, crystal studies indicate a histidine residue is responsible for the more spatially precise activation of the attacking nucleophilic water and the metal ion is only involved in the less restrictive stabilization of the transition state, which may explain this metal tolerance ***(58)***. They have been found in archaea and bacteria and, unlike typical endonucleases, even occur in eukaryotes. Their genomic location can be mitochondrial, chloroplast, chromosomal, or extrachromosomal. They can be subdivided into four groups based on sequence motifs. To date, 58 have been identified and characterized to varying degrees ***(22)***.

4. Altered Specificities, Fusion Proteins, and Specialized Applications

4.1. Star Activity

Although endonucleases bind DNA nonspecifically, they exhibit a very high preference catalytically for their recognition site over sites with even a one base pair difference. A partial relaxation of specificity under suboptimal digest conditions is an inherent property of some enzymes that is commonly referred to as "star activity." Depending on the enzyme, star activity is most influenced by volume excluders (glycerol, ethylene glycol) or substitution of Mg^{2+} with another metal and, to a lesser degree, by pH ***(15)***. The number of water molecules normally present at the protein–DNA interface for *Eco*RI at noncognate sites is reduced at high osmotic pressure due to volume exclusion and the tighter binding of the enzyme results in the active conformation being more easily achieved at star sites ***(59)***. For example, *Eco*RI cleaves its recognition site (5'-GAATTC-3') at a rate 10^5 times faster than the next best sequence (5'-TAATTC-3') under optimal conditions ***(60)***. Complexes with this next best sequence and general non-specific sequences both contain approx 110 more water molecules than the specific complex at low osmotic pressure ***(61)***. With increasing ethylene glycol concentrations, cleavage rates decrease at the cognate site but increase at the next best site until the rates approach equivalence at 4 *M* ethylene glycol ***(59)***. At higher pH, the high $[OH^-]$ may reduce the need for activated water formed at the catalytic site as the attacking nucleophile ***(15)***. Alternately, pH and ionic strength may alter the dissociation of nonspecifically bound protein rather than influencing the specific/nonspecific equilibrium or being directly involved in catalysis ***(61)***. All restriction endonucleases prefer Mg^{2+} for activity. A few can use a different divalent metal, usually Mn^{2+}, but occasionally Ca^{2+} Fe^{2+}, Co^{2+}, Ni^{2+}, and Zn^{2+}. However, cleavage with these ions is usually less specific and slower ***(19)***. Mn^{2+} bound H_2O may be better than Mg^{2+} bound H_2O at providing the proton necessary for the leaving group 3' OH. For EcoRV, activity at the cognate site is 10^6 times higher than at the star site with Mg^{2+}, but only six-fold higher with Mn^{2+} ***(62)***. The type of salt ions, trace organic solvents, and high enzyme to DNA ratios may also result in star activity. Read the information sent with commercial preparations to avoid star activity for those enzymes that are susceptible.

4.2 Single Stranded Cleavage by Restriction Enzymes and Nickases

All restriction endonucleases cleave double-stranded DNA, but a few enzymes hydrolyze ssDNA at significantly reduced rates. Two theories exist regarding the mechanism of apparent ssDNA cleavage. Although cleavage of actual ssDNA has been reported ***(63)***, in other cases the enzyme may really act on transiently formed double stranded DNA ***(64)***. One method for cleaving ssDNA uses an oligonucleotide adaptor and a Type IIs enzyme where the recognition site and cleavage site are significantly separated such as *Fok*I. The oligonucleotide contains a hairpin loop, a double-stranded region with the recognition site of the enzyme, and a single-stranded tail extending past the recognition site. This single-stranded

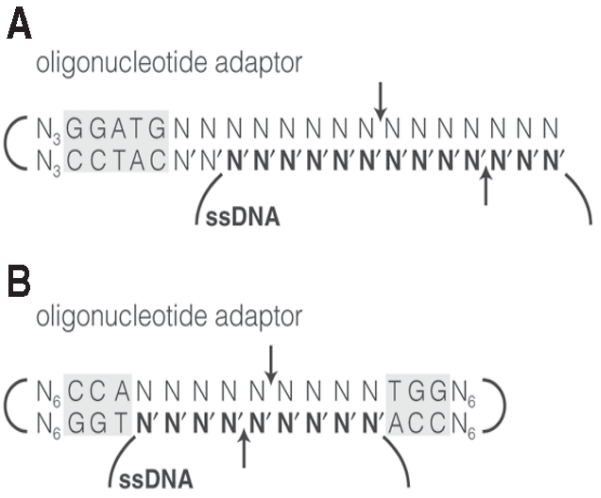

Fig. 1. Use of oligonucleotide adaptors for cleaving ssDNA with the endonucleases *Fok*I (***A***) and *Xcm*I (**B**). The ssDNA appears in bold, the endonuclease recognition site contained within the oligonucleotide is highlighted, and the cleavage positions are indicated the arrows.

region of the oligonucleotide protruding past the recognition site is complementary to the ssDNA substrate. The endonuclease is then able to recognize its cognate site on the double-stranded region of the oligonucleotide and cleave in the region formed by the oligo:ssDNA hybrid as illustrated in **Fig. 1A**. After cleavage, the oligonucleotide fragment bound to the ssDNA can be heat denatured. This method cannot be used to cleave ssRNA, however ***(19)***. A similar approach for cutting ssDNA uses the enzyme *Xcm*I, which recognizes the longest (9 nt) degenerative sequence known (5'-CCANNNNN/NNNNTGG-3'). An oligo is designed with two hairpin loops and adjacent inside regions of dsDNA containing the recognition nucleotides but leaving the center, nonspecific nucleotides single-stranded. A complementary ssDNA that hybridizes to these nonspecific nucleotides will be cleaved as illustrated in **Fig. 1B** ***(65)***.

A small group of naturally occurring nickases are now known that mimic restriction enzymes in recognizing a double-stranded DNA site but cleaving only one specific phosphate diester bond. The two best characterized nickases, N.*Bst*SE ***(66)*** and N.*Bst*NBI ***(67)***, are isoschizomers recognizing the nonpalindromic sequence GAGTC and cleaving the top strand four bases from the 3' end of the recognition site. N.*Bst*NBI has been cloned and shows a high degree of sequence similarity with the Type IIs enzymes *Ple*I and *Mly*I, which also recognize GAGTC. Furthermore, *Ple*I cleaves the top strand in the same position as N.*Bst*NBI. Based on gel filtration with bound DNA, DNA cleavage, and mutational analysis, a structure and mechanism similar to that of *Fok*I without the need for dimerization at the cleavage site has been proposed for N.BstNBI ***(68)***.

4.3. Specificity Alteration Through Protein Engineering

Despite the complete sequence and 3D structural data available, site-directed mutagenesis of restriction enzymes to alter the DNA sequence they recognize has generally resulted only in relaxed specificity and/or decreased cleavage rates. Simple mutations that easily alter the recognition site for restriction would be lethal to the host without the same alteration in methylase activity; therefore, the endonuclease has evolved to be highly specific and redundant in recognition. Mutations to accept a modified base or increasing specificity of the site so far have proven to be easier. Mutants of *Eco*RV have been constructed that cleave recognition sites with a uracil instead of thymine by more than two orders of magnitude over wild-type (N188Q) ***(69)*** or with a methylphosphonate in one of the phosphate backbone positions by three orders of magnitude over wild-type (T94V) ***(70)***. In addition, *Eco*RV A181K and A181E mutants preferentially cleave sites with a purine or thymine, respectively, 5' to the recognition site ***(71)***. A directed evolution approach has produced a N97T/S183A/T222S mutant with a 20-fold preference for an oligonucleotide with a GC-rich flanking region and a K104N/A181T mutant with a seven-fold preference for an AT-rich flanking region ***(72)***. Heterodimers of EcoRV containing a catalytically inactive mutant subunit act as site-specific nickase ***(73)***. To facilitate additional enzyme engineering, more structural information is needed regarding the large number of

enzyme–DNA contacts and intra/intermolecular protein shifts. However, several other approaches to achieve custom designed, sequence-specific cleavage have also been investigated such as metal-catalyzed cleavage, Achilles heel cleavage, and fusion proteins.

4.4. Metal Catalyzed Cleavage

One approach to custom-designed specificity and cleavage uses either an oligonucleotide capable of forming a triple helix ***(74)*** or a DNA-binding protein to provide the specificity. Covalently attached to the oligonucleotide or protein is a metal complex, usually EDTA-iron or *o*-phenanthroline-copper, which catalyzes phosphodiester cleavage in the presence of a reducing agent. Proteins successfully used to supply specificity include Cro ***(75)***, the catabolite activator protein "CAP" ***(76)***, and the Msx-1 homeodomain ***(77)***. However, cleavage at more than one phosphodiester bond in each strand results in a mixed population of overhangs and cleavage does not proceed to completion.

4.5. Achilles Heel Cleavage

A technique to achieve more precise and complete cleavage but at a less frequent number of sites than standard endonucleases is known as "Achilles heel cleavage." First, a target sequence is protected by a bound *Rec*A/oligo complex ***(78,79)*** or triple helix formation ***(80)***. A methylase modifies all sites except the protected target. The methylase is removed by purification, followed by the protecting group. The target sequence is then specifically cleaved by a methyl sensitive restriction enzyme.

4.6. Fusion Proteins

Hybrid enzymes can be constructed by fusing recognition and cleavage domains from different proteins. In one such example, the recognition domain of the Type IIs enzyme *Alw*I was fused to the catalytic domain of the nicking enzyme N.*Bst*NBI, which does not contain the dimerization potential necessary for double stranded cleavage. The resultant chimeric enzyme, N.*Alw*I, cuts only the top strand four bases downstream from the *Alw*I recognition site, GGATC ***(81)***.

Using various spacers and constructs, the Type IIs *Fok*I catalytic domain has been combined with DNA binding domains from the *Drosophila Ubx* homeodomain ***(82)***, the zinc finger region of the eukaryotic transcription factor Sp1, the designed zinc finger consensus sequence protein QQR ***(83)***, and the zinc finger region of the yeast transcription factor Gal4 ***(84)***. Interestingly, the fusions with Sp1 and QQR will also bind and cleave the DNA strand of DNA–RNA hybrids. Since the recognition sequence of Gal4 is palindromic and the protein dimerizes at the site, the strands are cleaved on opposite sides of the recognition site. This results in a ≥24 base, 5' overhang, which includes the recognition nucleotides. Sites for the *Ubx*, Sp1, and QQR hybrids are nonpalindromic, the fusion proteins may act as monomers, and both strands are cleaved to one side of the recognition site at high chimeric enzyme concentration although generally at more than one position.

This approach has recently been further refined. A new chimeric enzyme was created with the cleavage domain of *Fok*I fused to QNK, another designed zinc finger protein. However, the cleavage domain of the fusion protein was again not under the same level of allosteric regulation as in the native *Fok*I and therefore low levels of hydrolysis occurs at nearby phosphates when enzyme concentrations are sufficient to produce double-stranded cuts ***(85)***. Cleavage efficiency was greatly enhanced, and alternate site cleavage reduced, by positioning two of the nonpalindromic consensus sequences close together in a tail-to-tail orientation. The greatest fidelity of cleavage for QQR at a single phosphate bond in each strand occurred when the intervening sequence was 12 bp. In this manner, each strand was cleaved eight bases from their respective recognition site, approximately one helical turn, and a 5' overhang of four bases was generated as in native *Fok*I. In addition, the substrate DNA could consist of one site each of QQR and QNK in similar tail-to-tail orientation, and requiring both respective chimeric enzymes for cleavage. When the dimerization-deficient mutants of the *Fok*I catalytic domain were used, cleavage did not occur. Therefore, it appears likely that catalytic

domain dimerization in the spacer DNA of approximately one helical turn is necessary for efficient and specific cleavage ***(86)***.

Taken a step further, a DNA containing appropriately spaced and orientated QQR sites was injected into oocytes and allowed to assemble into chromatin. Subsequent injection with the chimeric QQR enzyme and incubation yielded nearly 100% homologous recombination when the spacer was 8 bp in length. Significant homologous recombination was also observed when the hybrid QQR/QNK substrate and both chimeric enzymes were injected. Since each consensus sequence is 9 nucleotides in length, the predicted occurrence of such an 18-bp site would be 4^{18} or 6.9×10^{10}. Use of additional zinc fingers would expand the consensus binding sequence. Given that the human genome is approx 3×10^9 base pairs, continued development of this technique may hold promise for stimulating targeted homologous recombination in vivo ***(87)***.

5. Commercially Prepared Restriction Endonucleases

5.1. Unit Definition and Application to Other Substrates

Commercial vendors of restriction endonucleases use standard assays for unit activity definition with only minor variations. The products of digestion are generally separated by electrophoresis in agarose gels and detected by ethidium bromide staining. An activity unit is defined as the amount of enzyme necessary to completely digest 1 µg of the defined substrate, usually lambda DNA, in a 50 µL reaction volume in 1 h at the specified temperature. If the number of sites is small (≤3), lambda pre-digested with another enzyme (e.g., *Eco*RI) may be used to improve gel resolution of the fragments resulting from digestion with the enzyme in question. A different DNA such as Adenovirus 2 is used if there is a single or no sites in lambda.

The reaction conditions and enzyme units needed to digest a given substrate must be chosen carefully to ensure performance. The buffer systems provided with commercially obtained enzymes are designed to balance optimal individual enzyme performance and limiting the number of different buffers. Very few provided reaction buffers are specifically optimized for a single enzyme, nor is there a true "universal" buffer. As with any group of similar enzymes, endonucleases are all unique to some degree in their preferences for buffer components and concentration such as cation (Na^+ or K^+), anion (Cl^- or acetate), pH (7.2–8.5), stabilizer (BSA, detergent, or spermidine), and reducing agent. The storage buffer of the enzyme may adversely affect use when it comprises an unusually large amount of the total reaction volume (volume exclusion of glycerol causing star activity, chelators for stability binding Mg in reaction, and so on).

As it may also constitute a large percentage of the reaction volume, substrate preparation is critical for enzyme performance. Unit definitions are generally given for high purity linear phage or viral DNA, which is not necessarily the situation in many applications. Enzymes vary in their resistance to proteases, interference by DNA binding proteins, competitive inhibition from RNA, and tolerance for EDTA, PEG, SDS, CsCl, phenol, chloroform, and alcohols. Extra caution should also be used for cutting near the end of a DNA substrate. Endonucleases require contact with the DNA backbone for several bases adjacent to the recognition sequence. In general, the recognition site must lie 3 bp from the end to give good cleavage. Tables have been developed for a limited list of enzymes based on cutting a short oligo ***(88)***, cutting a PCR fragment near one end ***(89)***, and double digests of adjacent sites in a polylinker ***(90)***. One has to keep in mind the number of pmoles of cut sites used to define a unit vs a substrate of interest. The following table suggests the theoretical enzyme units needed for complete cutting with *Bam*H I based strictly on the number of cut sites and optimal conditions. Although no other parameters are taken into account, this approach can be a useful approximation.

5.2. Quality Control Assays

An overdigest or nonspecific exo- and endonuclease assay is performed in the same fashion as an activity assay except that a large excess of enzyme

DNA	Base pairs	Picomoles in 1 µg*	Cut sites (*Bam*HI)	Picomoles cut sites	Enzyme units needed
Unit Definition (ex. lambda)	48,502	0.0312	5	0.156	1
Plasmid	3,000	0.5	1	0.5	3–4**
PCR Fragment	700	2.16	1	2.16	12–15
Oligonucleotide	25	60.6	1	60.6	400–380

*Based on 660 Daltons per bp of DNA.
**Enzymes differ in their ability to cut supercoiled vs linearized substrates.

units is used (5–100 U) and incubation times are long (generally 16 h). Often the result is reported as (X-) fold-overdigestion (units × hours) in spite of the fact that the activity of the enzyme typically diminishes significantly over this length of time at the reaction temperature. Although star activity is an inherent property of the enzyme itself and not a separate, distinct contaminant, it does result in cleavage at noncognate sites, which will interfere with downstream applications, and therefore must be considered in determining the endpoint of this assay. Some suppliers only consider the absence of discrete contaminants in their specifications, which can be misleading.

The cut–ligate–recut assay is more sensitive with regard to contaminating exonuclease and can additionally detect phosphatase activity. DNA is slightly overdigested, ligated with T4 DNA ligase, and then recut. Dephosphorylated ends will not be ligated and staggered ends that have been blunted by single-stranded exonuclease activity will exhibit less efficient ligation. Loss of any of the original terminal nucleotides after cleavage will almost always also result in the loss of the recognition site for recutting even if the substrate religated. Although ligation of a one base overhang is still efficient enough to be useful in T-vector cloning of PCR fragments, for unknown reasons the efficiency of ligation, as indicated by transformation efficiencies of plasmids, can be as much as two orders of magnitude lower than that observed for blunt end ligation. In contrast, a four-base G-C overhang is stable enough to transform well even without in vitro ligation.

Some vendors also test for the presence of nickases by incubating the enzyme and a supercoiled substrate (RF I form) that does not contain a recognition site to determine the amount of substrate converted to a relaxed, open circle (RF II form). However, the impact of nicking activity on the major applications of cloning and mapping are minimal. One application that would be influenced by nicking is the generation of nested deletions with exonuclease III. However, improperly purified DNA is a far more likely source of nicks than contaminant activity present in the endonuclease. Current purification methods and the sensitivity of other assays are such that rarely, if ever, is a nickase test warranted.

Another test used for the detection of exonucleases is to digest ^{3}H-ds and -ssDNA, TCA precipitate the remaining DNA, and detect released nucleotides by scintillation counting of the supernatant. Be aware that if the enzyme cleaves the substrate within 30 bp of the DNA end, inefficient precipitation of the resultant small fragments may lead to an incorrect interpretation that suggests exonuclease activity. Labeling the 5' or 3' ends of DNA with ^{32}P is a sensitive way to detect contaminating phosphatase or exonuclease. However, this method requires frequent preparation of substrate.

The Blue/White assay combines excellent sensitivity and a verification of performance. A cloning plasmid is used that contains a multiple cloning site flanked by RNA polymerase promoter(s) within a coding sequence for the *lacZ* gene α-peptide and a separate selectable marker such as ampicillin resistance. The plasmid is several-fold overdigested with an enzyme having a single site within the multiple cloning region. The DNA is ligated (without insert) and then transformed into cells lacking the α-peptide region of *lacZ*. An agarose gel of cut, ligated, and recut DNA is also examined. If the integrity of the cut ends is perfectly maintained, ligation will produce mostly higher molecular weight concatamers and a lesser amount of circularized monomer. Upon transformation and α-peptide expression, the functional *lacZ* gene product β-galactosidase is produced through α-complementation. When plated in the presence of IPTG and X-Gal, blue colonies will result. Expected transformation efficiency will be 1–2 orders of magnitude lower than

with control supercoiled plasmid. Both phosphatase and exonuclease contamination will lower ligation efficiency and thereby decrease transformation efficiency. More importantly, loss of nucleotides at the cut site, even if ligatable, yields a mixed population of clones containing frame shifts and codon deletions in the *lacZ* gene α-peptide. These cause white colonies, i.e., false positives in a cloning experiment with an insert ***(91)***. A special case is an as yet unidentified contaminant that is found in native and cloned preparations of endonuclease that removes a single 3' nucleotide from the end of DNA. The resultant colonies are then able to use an alternative start codon that shifts to become in-frame. However, it produces weak translation initiation and/or improper complementation for fully active β-galactosidase and the colonies develop a faint blue color, which is easily mistaken for white. This is especially problematic with blunt end cutting enzymes. Not all commercial vendors specifically assay for and remove this contaminant ***(92)***.

Acknowledgment

The author would like to thank Michael Slater, Mark Klekamp, Terri Sundquist, and Isobel Maciver for their review of the manuscript and many helpful suggestions.

References

1. Roberts, R. J. and Halford, S. E. (1993), Type II Restriction Endonucleases, in Nucleases, 2nd Edition (Linn S. M., Lloyd, S. R., and Roberts, R. J. eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 35–88.
2. Stein, D.C., Gunn, J.S., Radlinska, M., and Piekarowicz, A. (1995) Restriction and modification systems of *Neisseria gonorrhoeae*. *Gene* **157,** 19–22.
3. Lin, L-F. Posfai, J., Roberts, R. J., and Kong, H. (2001) Comparative genomics of the restriction-modification systems in *Helicobacter pylori*. *Nuc. Acids Res.* **98,** 2740–2745.
4. Rocha, E. P. C., Danchin, A., and Viari, A. (2001) Evolutionary Role of Restriction/Modification Systems as Revealed by Comparative Genome Analysis. *Genome Res.* **11,** 946–958.
5. Nobusato, A., Uchiyama, I., and Kobayashi, I. (2000) Diversity of restriction-modification gene homologues in *Helicobacter pylori Gene* **259,** 89–98.
6. Smith, H. O. and Nathans, D. (1973) A suggested nomenclature for bacterial host modification and restriction systems and their enzymes. *J. Mol. Biol.* **81,** 419–423.
7. Wilson, G. G. and Murray, N. E. (1991) Restriction and Modification Systems. *Annu. Rev. Genet.* **25,** 585–627.
8. Stahl, F., Wende, W., Jeltsch, A., and Pingoud, A. (1998) The Mechanism of DNA Cleavage by the Type II Restriction Enzyme *Eco*RV: Asp36 Is Not Directly Involved in DNA Cleavage but Serves to Couple Indirect Readout to Catalysis. *Biol. Chem.* **379,** 467–473.
9. Smith, H. O., Annau, T. M., and Chandrasegaran, S. (1990) Finding sequence motifs in groups of functionally related proteins. *Proc. Natl. Acad. Sci. USA* **87,** 826–830.
10. Szomolanyi, E., Kiss, A., and Venetianer, P. (1980) Cloning the modification methylase gene of *Bacillus sphaericus* R in *Escherichia coli*. *Gene* **10,** 219–225.
11. Tao, T., Bourne, J. C., and Blumenthal, R. M. (1991) A Family of Regulatory Genes Associated with Type II Restriction-Modification Systems. *J. Bact.* **173,** 1367–1375.
12. Ives, C. L., Sohail, A., and Brooks, J. E. (1995) The Regulatory C Proteins from Different Restriction-Modification Systems Can Cross-Complement. *J. Bact.* **177,** 6313–6315.
13. Bickle, T. A. (1993) *The ATP-dependent Restriction Enzymes, in Nucleases,* 2nd ed. (Linn S. M., Lloyd, S. R., and Roberts, R. J. eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 35–88.
14. Murray, N. E. (2000) Type I restriction systems: sophisticated molecular machines. *Microbiol. Mol. Biol. Revs.* **64,** 412–434.
15. Pingoud, A. and Jeltsch, A. (1997) Recognition and cleavage of DNA by type-II restriction endonucleases. *Eur. J. Biochem.* **246,** 1–22.
16. Pingoud, A., Jeltsch, A. (2001) Structure and function of type II restriction endonucleases. *Nuc. Acids Res.* **29,** 3705–3727.
17. Reuter, M., Kupper, D., Pein, C. D., Petrusyte, M., Siksnys, V., Frey, B., and Kruger, D. H. (1993) Use of Specific Oligonucleotide Duplexes to Stimulate Cleavage of Refractory DNA Sites by Restriction Endonucleases. *Anal. Biochem.* **209,** 232–237.
18. Oller, A. R., Broek, W. V., Conrad, M., and Topal, M. (1991) Ability of DNA and Spermidine To Affect the Activity of Restriction Endonucleases from Several Bacterial Species. *Biochemistry* **30,** 2543–2549.
19. Szybalski, W., Kim, S. C., Hasan, N., and Podhajska, A. J. (1991) Class-IIS restriction enzymes - a review. *Gene* **100,** 13–26.
20. Kong, H. (1998) Analyzing the Functional Organization of a Novel Restriction Modification System, the *Bcg*I System. *J. Mol. Biol.* **279,** 823–832.
21. Sears, L. E., Zhou, B., Aliotta, J. M., Morgan, R. D., and Kong, H. (1996) BaeI, another unusual *Bcg*I-like restriction endonuclease. *Nucleic Acids Res.* **24,** 3590–3592.
22. Belfort, M. and Roberts, R. J. (1997) Homing endonucleases: keeping the house in order. *Nuc. Acids Res.* **25,** 3379–3388.

23. Meisel, A., Mackeldanz, P., Bickle, T. A., Kruger, D. H., and Schroeder, C. (1995) Type III restriction endonucleases translocate DNA in a reaction driven by recognition site-specific ATP hydrolysis. *EMBO J.* **14,** 2958–2966.
24. Kruger, D. H., Kupper, D., Meisel, A., Reuter, M., and Schroeder, C. (1995) The significance of distance and orientation of restriction endonuclease recognition sites in viral DNA genomes. *FEMS Microbiol. Rev.* **17,** 177–184.
25. Janulaitis, A., Petrusyte, M., Maneliene, Z., Klimasauskas, S., and Butkus, V. (1992) Purification and properties of the *Eco*57I restriction endonuclease and methylase — prototypes of a new class (type IV). *Nucl. Acids Res.* **20,** 6043–6049.
26. Williams, R. J. (2001), Restriction Endonucleases and Their Uses, in *The Nucleases* (Schein, C. H., ed.), Humana Press, New York, NY, pp. 409–429.
27. Mise, K. and Nakajima, K. (1985) Purification of a new restriction endonuclease, StyI, from Escherichia coli carrying the hsd+ minplasmid. *Gene* **33,** 357–361.
28. Jeltsch, A., and Pingoud, A. (1998) Kinetic Characterization of Linear Diffusion of the Restriction Endonuclease *Eco*RV on DNA. *Biochemistry* **37,** 2160–2169.
29. Engler, L. E., Sapienza, P., Dorner, L. F., Kucera, R., Schildkraut, I., and Jen-Jacobson, L. (2001) The Energetics of the Interaction of *Bam*HI Endonuclease with its Recognition Site GGATCC. *J. Mol. Biol.* **307,** 619–636.
30. Horton, N. C. and Perona, J. J. (2000) Crystallographic snapshots along a protein-induced DNA-bending pathway. *Proc. Natl. Acad. Sci.* **97,** 5729–5734.
31. Kovall, R. A. and Matthews, B. W. (1998) Structural, functional, and evolutionary relationships between λ-exonuclease and the type II restriction endonucleases. *Proc. Natl. Acad. Sci.* **95,** 7893–7897.
32. Kovall, R. A. and Matthews, B. W. (1999) Type II restriction endonucleases: structural, functional and evolutionary relationships. *Curr. Opin. Chem. Biol.* **3,** 578–583.
33. Kubareva, E. A., Thole, H., Karyagina, A. S., Oretskaya, T. S., Pingoud, A., and Pingoud, V. (2000) Identification of a base-specific contact between the restriction endonuclease *Sso*II and its recognition sequence by photocross-linking. *Nucl. Acids Res.* **28,** 1085–1091.
34. Lukacs, C. M., Kucera, R., Schildkraut, I., and Aggarwal, A. K. (2000) Understanding the immutability of restriction enzymes: crystal structure of *Bgl*II and its DNA substrate at 1.5 Å resolution. *Nature Struc. Biol.* **7,** 134–140.
35. Huai, Q., Colandene, J. D., Chen, Y., Luo, F., Zhao, Y., Topal, M. D., and Ke, H. (2000) Crystal structure of *Nae*I-an evolutionary bridge between DNA endonuclease and topoisomerase. *EMBO J.* **19,** 3110–3118.
36. Reuter, M., Kupper, D., Meisel, A., Schroeder, C., and Kruger, D. H. (1998) Cooperative Binding Properties of Restriction Endonuclease *Eco*RII with DNA Recognition Sites. *J. Biol. Chem.* **273,** 8294–8300.
37. Pein, C. D., Reuter, M., Meisel, A., Cech D., and Kruger, D. H. (1991) Activation of restriction endonuclease *Eco*RII does not depend on the cleavage of stimulator DNA. *Nuc. Acids Res.* **19,** 5139–5142.
38. Conrad, M., and Topal, M. (1992) Modified DNA fragments activate *Nae*I cleavage of refractory DNA sites. *Nuc. Acids Res.* **20,** 5127–5130.
39. Senesac, J. H. and Allen, J. R. (1995) Oligonucleaotide Activation of the Type IIe Restriction Enzyme *Nae*I for Digestion of Refractory Sites. *BioTechniques* **19,** 990–993.
40. Jo, K. and Topal, M. D. (1995) DNA Topoisomerase and Recombinase Activities in *Nae*I Restriciton Endonuclease. *Science* **267,** 1817–1820.
41. Jo, K. and Topal, M. D. (1998) Step-wise DNA relaxation and decatenation by *Nae*I-43K. *Nucleic Acids Res.* **26,** 2380–2384.
42. Colandene, J. D. and Topal, M. D. (2000) Evidence for Mutations That Break Communication between the Endo and Topo Domains in *Nae*I Endonuclease/ Topoisomerase. *Biochemistry* **39,** 13703–13707.
43. Wah, D. A., Hirsch, J. A., Dorner, L. F., Schildkraut, I., and Aggarwal, A. K. (1997) Structure of the multimodular endonuclease *Fok*I bound to DNA. *Nature* **388,** 97–100.
44. Bitinaite, J., Wah, D. A., Aggarwal, A. K., and Schildkraut, I. (1998) *Fok*I dimerization is required for DNA cleavage. *Proc. Natl. Acad. Sci.* **95,** 10570–10575.
45. Scheuring, E. S., Santagata, S., and Aggarwal, A. K. (2001) FokI Requires Two Specific DNA Sites for Cleavage. *J. Mol. Biol.* **309,** 69–78.
46. Wentzell, L. M., Nobbs, T. J. and Halford, S. E. (1995) The *Sfi*I restriction endonuclease makes a 4-strand DNA break at two copies of its recognition sequence. *J. Mol. Biol.* **248,** 581–595.
47. Sato, H., Suzuki, T., and Yamada, Y. (1990) Purification of Restriction Endonuclease from *Acetobacter aceti* IFO 3281 (*Aat*II) and its Properties. *Agric. Biol. Chem.* **54,** 3319–3325.
48. Siksnys, V., Skirgaila, R., Sasnauskas, G., Urbanke, C., Cherny, D., Grazulis, S. (1999) The *Cfr*10I restriction enzyme is functional as a tetramer. *J. Mol. Biol.* **291,** 1105–1118.
49. Deibert, M., Grazulis, S., Sasnauskas, G., Siksnys, V., and Huber, R. (2000) Struture of the tetrameric restriction endonuclease *Ngo*MIV in complex with cleaved DNA. *Nature Struct. Biol.* **7,** 792–799.
50. Nobbs, T. J., Williams, S. A., Connolly, B. A., and Halford, S. E. (1998) Phosphorothioate Substrates for the *Sfi*I Restriction Endonuclease. *Biol. Chem.* **379,** 599–604.
51. Nobbs, T. J. and Halford, S. E. (1995) DNA cleavage at two recognition sites by the *Sfi*I restriction endonuclease: salt dependence of cis and trans interactions between distand DNA sites. *J. Mol. Biol.* **252,** 399–411.

52. Embleton, M. L., Sidsnys, V., and Halford, S. E. (2001) DNA Cleavage Reactions by Type II Restriction Enzymes that Require Two Copies of their Recognition Sites. *J. Mol. Biol.* **311,** 503–514.
53. Williams, S. A. and Halford, S. E. (2001) SfiI endonuclease activity is strongly influenced by the non-specific sequence in the middle of its recognition site. *Nucl. Acids Res.* **29,** 1476–1483.
54. Stankevicius, K., Lubys, A., Timinskas, A., Vaitkevicius, D., and Janulaitis, A. (1998) Cloning and analysis of the four genes coding for *Bpu*10I restriction-modification enzymes. *Nucl. Acids Res.* **26,** 1084–1091.
55. Hsieh, P.-C., Xiao, J.-P., O'Loane, D., and Xu, S.-Y. (2000) Cloning, Expression, and Purification of a Thermostable Nonhomodimeric Restriction Enzyme, *Bsl*I. *J. Bact.* **182,** 949–955.
56. Kong, H. and Smith, C. L. (1998) Does *Bcg*I, a Unique Restriction Endonuclease, Require Two Recognition Sites for Cleavage? *Biol. Chem.* **379,** 605–609.
57. Kong, H. and Smith, C. L. (1997) Substrate DNA and cofactor regulate the activities of a multi-functional restriction-modification enzyme, *Bcg* I. *Nucl. Acids Res.* **25,** 3687–3692.
58. Galburt, E. A., Chevalier, B., Tang, W., Jurica, M. S., Flick, K. E., Monnat, R. J., and Stoddard, B. L. (1999) A novel endonuclease mechanism directly visualized for I-*Ppo*I. *Nature Struct. Biol.* **6,** 1096–1099.
59. Robinson, C. R. and Sligar, S. G. (1998) Changes in solvation during DNA binding and cleavage are critical to altered specificity of the *Eco*RI endonuclease. *Proc. Natl. Acad. Sci. USA* **95,** 2186–2191.
60. Lesser, D. R., Kurpiewski, M. R., and Jen-Jacobson, L. (1990) The Energetic Basis of Specificity in the *Eco*RI Endonuclease-DNA Interaction. *Science* **250,** 776–786.
61. Sidorova, N. Y. and Rau, D. C. (2001) Linkage of *Eco*RI Dissociation from its Specific DNA Recognition Site to Water Activity, Salt Concentration, and pH: Separating their Roles in Specific and Non-specific Binding. *J. Mol. Biol.* **310,** 801–816.
62. Vermote, C. L. M. and Halford, S. E. (1992) *Eco*RV restriction endonuclease: Communication between catalytic metal ions and DNA recognition. *Biochemistry* **31,** 6082–6089.
63. Yoo, O. J. and Agarwal, K. L. (1980) Cleavage of single strand oligonucleotides and bacteriophage phiX174 DNA by Msp I endonuclease. *J. Biol. Chem.* **255,** 10559–10562.
64. Blakesley, R. W., Dodgson, J. B., Nes, I. F., Wells, R. D. (1977) Duplex regions in "single-stranded" phiX174 DNA are cleaved by a restriction endonuclease from Haemophilus aegypius. *J. Biol. Chem.* **252,** 7300–7306.
65. Shaw, P. C. and Mok, Y. K. (1993) XcmI as a universal restriction enzyme for single-stranded DNA. *Gene* **133,** 85–89.
66. Abdurashitov, M. A., Belichenko, O. A., Shevchenko, A. V. and Degtyarev, S. K. (1996) N.*Bst*SE-site-specific nuclease from *Bacillus stearothermophilus* SE-589-restriction endonuclease production. *Mol. Biol. (Mosk.)* **30,** 1261–1267.
67. Morgan, R. D., Calvet, C., Demeter, M., Agra, R. and Kong, H. (2000) Characterization of the specific DNA nicking activity of restriction endonuclease N.BstNBI. *Biol. Chem.* **381,** 1123–1125.
68. Higgins, L. S., Besnier, C. and Kong, H. (2001) The nicking endonuclease N.*Bst*NBI is closely related to Type IIs restriction endonucleases *Mly*I and *Ple*I. *Nucl. Acids Res.* **29,** 2492–2501.
69. Wenz, C., Selent, U., Wende, W., Jeltsch, A., Wolfes, H., and Pingoud, A. (1994) Protein engineering of the restriction endonuclease *Eco*RV: replacement of an amino acid residue in the DNA binding site leads to an altered selectivity towards unmodified and modified substrates. *Biochim. Biophys. Acta.* **1219,** 73–80.
70. Lanio, T., Selent, U., Wnez, C., Wende, W., Schulz, A., Adiraj, M., Katti, S. B., and Pingoud, A. (1996) *Eco*RV-T94V: a mutant restriction endonuclease with an altered substrate specificity towards modified oligodeoxynucleotides. *Protein Eng.* **9,** 1005–1010.
71. Schottler, S., Wenz, C., Lanio, A., Jeltsch, A., and Pingoud, A. (1998) Protein engineering of the restriction endonuclease *Eco*RV: Structure-guided design of enzyme variants that recognize the base pairs flanking the recognition site. *Eur. J. Biochem.* **258,** 184–191.
72. Lanio, T., Jeltsch, A., and Pingoud, A. (1998) Towards the Design of Rare Cutting Restriction Endonucleases: Using Directed Evolution to Generate Variants of *Eco*RV Differing in Their Substrate Specificity by Two Orders of Magnitude. *J. Mol. Biol.* **283,** 59–69.
73. Stahl, F., Wende, W., Jeltsch, A., and Pingoud, A. (1996) Introduction of asymmetry in the naturally symmetric restriction endonuclease *Eco*RV to investigate intersubunit communication in the homodimeric protein. *Proc. Natl. Acad. Sci. USA* **93,** 6175–6180.
74. Dervan, P. B. (1992) Reagents for the site-specific cleavage of megabase DNA. *Nature* **359,** 87–88.
75. Ebright, Y. W., Chen, Y., Pendergrast P. S., and Ebright, R. H. (1992) Incorporation of an EDTA-metal complex at a rationally selected site within a protein: application to EDTA-iron DNA affinity cleaving with catabolite gene activator protein (CAP) and Cro. *Biochemistry* **31,** 10664–10670.
76. Pendergrast, P. S., Ebright, Y. W., and Ebright, R. H. (1994) High-Specificity DNA Cleavage Agent: Design and Application to Kilobase and Megabase DNA Substrates. *Science* **265,** 959–962.
77. Shang, Z., Ebright, Y. W., Iler, N., et al. (1994) DNA affinity cleaving analysis of homeodomain-DNA interaction: Identification of homeodomain consensus

sites in genomic DNA. *Proc. Natl. Acad. Sci. USA* **91,** 118–122.
78. Koob, M., Burkiewicz, A., Kur J., Szybalski, W. (1992) RecA-AC: single-site cleavage of plasmids and chromosones at any predetermined restriction site. *Nucl. Acids. Res.* **20,** 5831–5836.
79. Schoenfeld, T., Harper, T., and Slater, M. (1995) RecA Cleavage and Protection for Genomic Mapping and Subcloning. *Promega Notes* **50,** 9–14.
80. Strobel, S. A. and Dervan, P. B. (1991) Single-site enzymatic cleavage of yeast genomic DNA mediated by triple helix formation. *Nature* **350,** 172–174.
81. Xu, Y., Lunnen, K. D., and Kong, H. (2001) Engineering a nicking endonuclease N.*Alw*I by domain swapping. *Proc. Natl. Acad. Sci.* **98,** 12990–12995.
82. Kim, Y. G. and Chandrasegaran, S. (1994) Chimeric restriction endonuclease. *Proc. Natl. Acad. Sci. USA*. **91,** 883–887.
83. Kim, Y. G., Shi, Y., Berg, J. M., and Chandrasegaran, S. (1997) Site-specific cleavage of DNA-RNA hybrids by zinc finger/*Fok*I cleavage domain fusions. *Gene* **203,** 43–49.
84. Kim, Y. G., Smith, J., Durgesha, M., and Chandrasegaran, S. (1998) Chimeric Restriction Enzyme: Gal4 Fusion to FokI Cleavage Domain. *Biol. Chem.* **379,** 489–495.
85. Smith, J., Berg, J. M., Chandrasegaran, S. (1999) A detailed study of the substrate specificity of a chimeric restriction enzyme. *Nucl. Acids Res.* **27,** 674–681.
86. Smith, J., Bibikova, M., Whitby, F. G., Reddy, A. R., Chandrasegaran, S. and Carroll, D. (2000) Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucl. Acids Res.* **28,** 3361–3369.
87. Bibikova, M., Carroll, D., Segal, D. J., Trautman, J. K., Smith, J., Kim, Y-G. and Chandrasegaran, S. (2001) Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases. *Mol. and Cell. Biol.* **21,** 289–297.
88. New England BioLabs 2000·01 Catalog (2000). New England BioLabs, Inc., 210–211.
89. Dallas-Yang, Q., Jiang, G., and Sladek, F. M. (1998) Digestion of Terminal Restriction Endonuclease Recognition Sites on PCR Products. *BioTechniques* **24,** 582–584.
90. Moreira, R. F. and Noren, C. J. (1995) Minimum Duplex Requirements for Restriction Enzyme Cleavage Near the Termini of Linear DNA Fragments. *BioTechniques* **19,** 57–59.
91. Hung, L., Murray, E., Murray, W., Bandziulis, R., Lowery, R., Williams, R., Noble, R. (1991) A Blue/White Cloning Assay for Quality Control of DNA Restriction and Modifying Enzymes. *Promega Notes* **33,** 12–13.
92. Murray, E., Singer, K., Cash, K., and Williams, R. (1993) Cloning-qualified blunt end restriction enzymes: Causes and cures for light blue colonies. *Promega Notes* **41,** 1–5.