Restriction Enzymes

Mala Mani, *Dana Farber Cancer Institute, Boston, Massachusetts, USA*
Karthikeyan Kandavelou, *Johns Hopkins University, Baltimore, Maryland, USA*
Srinivasan Chandrasegaran, *Johns Hopkins University, Baltimore, Maryland, USA*

Type II restriction enzymes are the molecular scissors that catalyse the double-strand cleavage of deoxyribonucleic acid (DNA) at specific base sequences. They are essential tools for manipulating DNA including but not limited to cloning, analysis and sequencing. Recent advances have made it possible to design and engineer chimaeric nucleases to target specific DNA sites within a genome thus making them useful tools to carry out gene therapy.

Advanced article

Article Contents

doi: 10.1002/9780470015902.a0000973.pub3

Introduction

The type II restriction endonucleases, also commonly known as restriction enzymes, are molecular 'scissors' that bind to specific sequences in deoxyribonucleic acid (DNA) and cut within or adjacent to these sites. Their discovery in the late 1960s ushered in a recombinant DNA technology revolution in molecular biology. For their work in restriction enzymes, three scientists were awarded the Nobel Prize in Medicine (1978): Werner Arber of the Basel University, for experiments showing the existence of restriction endonucleases; Hamilton Smith, from Johns Hopkins University, for the discovery of a restriction enzyme; and Daniel Nathans, also of the Johns Hopkins University, for demonstrating the utility of restriction enzymes for analysing DNA. **See also**: Bacterial Restriction–Modification Systems

Restriction Enzymes: Essential Tools for Manipulating DNA

Restriction enzymes are endonucleases that recognize specific DNA sequences and make double-strand cleavages. Restriction endonucleases are divided into three groups, type I, type II and type III based on their subunit composition, cofactor requirement and enzymatic mechanism (**Table 1**). They all require Mg^{2+} as cofactor. Type I and type III also require adenosine triphosphate (ATP) and *S*-adenosyl-L-methionine (AdoMet or SAM) as activators and their cleavage sites are situated at a distance from their recognition site. Cleavage sites for type II enzymes are located at or close to their recognition site. The common type II enzymes recognize specific sequences with a dyad axis of symmetry, called palindromes, and cleave within or adjacent to these sequences (Mani *et al*., 2003). Some enzymes cleave at the axis of symmetry to yield 'flush' or 'blunt' ends, while others make staggered cuts to yield overhanging single-stranded 3′ or 5′ ends, known as cohesive termini. The phosphodiester bond cleavage results in 3′ hydroxyl and 5′-phosphate termini. DNA fragments with compatible (i.e. complementary) ends may be ligated to each other using DNA ligases to produce recombinant DNA molecules. There are also numerous enzymes that recognize an asymmetric sequence and cleave a short distance from that sequence. These are termed as type IIs enzymes, where 's' stands for shifted cleavage. These enzymes do not recognize any specific sequence at the site cut.

The ability to manipulate DNA in defined ways has led to many discoveries in molecular biology since the 1970s. The essential tools of a genetic engineer are the enzymes that catalyse specific reactions on DNA molecules. The restriction enzymes that cleave DNA at discrete nucleotide sequences are critical for carrying out the most important reactions involved in recombinant DNA technology. The discovery of restriction enzymes has made it possible to purify homogeneous DNA fragments of defined length by molecular cloning. The DNA fragments generated by restriction enzymes are used as substrates for a wide variety of other enzymatic manipulations. Furthermore, the specific cleavage sites provide unique molecular landmarks for obtaining a physical map of DNA. Thus, restriction enzymes have proved to be essential tools for analysing and manipulating DNA.

Cleavage Specificity of Restriction Enzymes

Over 3000 restriction–modification enzymes have been identified; these come from widely diverse organisms, but mostly from bacteria. These enzymes fall into 'isoschizomer' (identically cleaving) groups with about 200 sequence specificities. REBASE, the restriction enzyme database, provides information regarding restriction

Table 1 Prototypes of type II and type I is restriction endonucleases and methylases

Enzyme	Recognition site/cleavage site(↓)	Recognition site/methylation site (m)
*Alu*I	5′ --- A G ↓ C T ---- 3′ 3′ --- T C ↑ G A ---- 5′	5′--- A G $\overset{m}{C}$ T --- 3′ 3′--- T $\underset{m}{C}$ G A --- 5′
*Bam*HI	5′ – G ↓ G A T C C --- 3′ 3′ - C C T A G ↑ G --- 5′	5′ – G G A T $\overset{m}{C}$ C --- 3′ 3′ - C $\underset{m}{C}$ T A G G --- 5′
*Eco*RI	5′ --- G ↓ A A T T C --- 3′ 3′ --- C T T A A ↑ G --- 5′	5′--- G $\overset{m}{A}$ A T T C --- 3′ 3′--- C T T $\underset{m}{A}$ A G --- 5′
*Pst*I	5′ ---- C T G C A ↓ G ---- 3′ 3′ ---- G ↑ A C G T C ---- 5′	5′--- C T G C $\overset{m}{A}$ G ---- 3′ 3′--- G $\underset{m}{A}$ C G T C --- 5′
*Not*I	5′ ----- G C ↓ G G C C GC ---- 3′ 3′ ----- C G C C G G ↑ CG ---- 5′	ND
*Pac*I	5′ ----- T T A A T ↓ T A A ---- 3′ 3′ ----- A A T ↑ T A A T T ---- 5′	ND
*Ple*I	5′ ---- G A G T C (N_4) ↓ ----- 3′ 3′ ---- C T C A G (N_5) ↑ ----- 5′	ND
*Fok*I	5′ ----- G G A T G (N_9) ↓ ------ 3′ 3′ ----- C C T A C (N_{13}) ↑ ------ 5′	5′--- G G $\overset{m}{A}$ T G (N_9) --- 3′ 3′--- C C T $\underset{m}{A}$ C (N_{13}) --- 5′

ND, not determined; N_n, nonspecific sequence of length *n* bases

enzymes, DNA methyltransferases, isoschizomers, neoschizomers, recognition sequence, commercial availability and references. Most restriction enzymes typically recognize DNA sites that are 4–8 base pairs (bp) in length. Prototypes of type II and type IIs restriction endonucleases and modification methylases are shown in **Table 1**. Enzymes recognizing simple tetranucleotide and hexanucleotide palindromic sequences are quite prevalent. There are 16 possible palindromic tetranucleotide sites and 64 possible hexanucleotide sites. Members of almost all of the possible sequences of these types have been found. The discovery of new enzymes involves tedious and time-consuming efforts that entail extensive screening of bacteria and other microorganisms. Even when a new enzyme is found, more often than not it falls into one of the already discovered isoschizomer groups. It has become increasingly difficult to discover new specificities by random screening (**Table 2**).

Restriction enzymes appear to have a key biological function: to protect cells from infection by foreign DNA that would otherwise destroy them. A corresponding cognate methylase protects the host genome from cleavage by the restriction enzymes. Viral genomes are usually small and are unlikely to carry the methylation pattern necessary to render them resistant to cleavage by the restriction enzymes. Thus it is likely that bacteria select for restriction–modification systems with small recognition sites (4–6 bp) because these sites occur more frequently in the viral genome.

Biochemical Properties of Restriction Enzymes

The type II enzymes are homodimers with subunits of about 25–50 kDa. A divalent cation, preferably Mg^{2+}, is required for cleavage. Since most type II enzymes recognize palindromic sequences with 2-fold rotational symmetry, it is expected that two enzyme subunits arranged symmetrically bind the recognition sites. Crystal structures of several enzyme–cognate site complexes have shown this to be the case. In contrast, the type IIs enzymes that recognize an asymmetric sequence appear to bind DNA as monomers.

How do the restriction enzymes recognize and cleave DNA? *In vitro* studies suggest that restriction enzymes initially bind DNA nonspecifically and then slide along the DNA to find their recognition site. Once they locate their binding sites, conformational changes in the DNA and the enzyme occur to form the specific DNA–enzyme complex that triggers DNA cleavage. Catalysis is mediated by a hydroxyl ion that is formed as a result of abstraction of a proton from a water molecule or through activation of a water molecule by complexation with Mg^{2+}. The restriction enzymes discriminate their substrate from specific and nonspecific DNA sequences with several direct and indirect sequence-recognition mechanisms. Direct recognition is mediated through specific hydrogen bonds with purine and pyrimidine bases within the recognition site and van der

Table 2 Salient features of restriction enzymes

Type I	Type II	Type III
Requires Mg^{2+}, ATP, AdoMet for DNA cleavage	Requires Mg^{2+} for DNA cleavage	Requires Mg^{2+}, ATP, AdoMet for DNA cleavage
Requires Mg^{2+} and AdoMet for methylation	Requires AdoMet for methylation	Requires Mg^{2+} and AdoMet for methylation
Cleaves as far as 1000 bp away from recognition site	Cleaves at or near recognition site	Cleaves 25–27 bp away from recognition site
Consists of three different subunits catalysing both restriction and methylation of DNA	Restriction and methylation carried out by separate enzymes	Consists of two subunits catalysing both restriction and methylation of DNA

Waals interactions with pyrimidines. Indirect interactions are mediated through sequence-specific hydrogen bonds to the phosphate backbone of the DNA substrate and sequence-dependent DNA distortions.

Star activity

Some restriction endonucleases relax or alter their recognition specificity under sub-optimal reaction conditions. This altered specificity is termed as 'star' activity. Star activity may be an inherent property of some enzymes. Conditions that influence star activity are high glycerol concentrations, low ionic strength, high enzyme concentration, high pH, trace organic solvents and substitution of Mg^{2+} by other divalent metals. *Eco*RI cleaves at the canonical site G↓AATTC, but at high pH and low ionic strength it cleaves the sequence N↓AATTN. Recent studies have shown that *Eco*RI star activity results in the cleavage of any site that differs from the canonical recognition site by a single base substitution. Star activity can be avoided by following the optimal buffer conditions recommended by the manufacturer.

Restriction Enzyme Production by Many Diverse Bacteria

Type II restriction enzymes are widespread in nature. Many thousands of bacterial species have been examined for the presence of restriction endonucleases, and they have been found in all genera examined. One-quarter of all bacterial species examined appear to contain one or more type II restriction–modification systems. They have been characterized in 11 of 13 phyla of bacteria and archaea. Most of them are derived from bacilli or proteobacteria. The four common sequence specificities, namely CGCG, GGCC, CCGG and GATC, have been found throughout the phylogenetic tree, including the archaea and proteobacteria.

Barany and co-workers have cloned and purified *Taq*I endonuclease isoschizomers from *Thermus* species obtained from different regions of the globe. These grow in hot spring temperatures ranging from 53 to 85°C. It appears that, although the protein sequences are generally conserved and amino acid residues important for catalytic activity are highly conserved, their thermostability can vary significantly.

Van Etten and co-workers have shown that *Chlorella* viruses that infect certain unicellular, eukaryotic, chlorella-like green algae encode multiple DNA methyltransferases and endonucleases. The recognition sites of these virus-encoded systems vary from 2 to 4 bp, in contrast to bacterial enzymes, which recognize sequences of 4–8 base-pairs. The functional significance of the *Chlorella* virus endonucleases remains a mystery.

Restriction–Modification Systems

Bacterial species contain restriction–modification systems with genes that encode both a restriction endonuclease and a methyltransferase that recognizes the same sequence. The host DNA is fully protected from the action of the restriction enzyme by the methyltransferase. However, invading or infecting DNA from a plasmid or a phage is not likely to carry the appropriate methylation and so it is susceptible to cleavage by the restriction enzyme. If there are one or more recognition sites on the incoming DNA, a single cut by the restriction enzyme is likely to incapacitate the invading DNA. This leads to the destruction of the DNA.

More than 200 restriction–modification systems have been cloned. Although some of the type II restriction systems are encoded on the plasmid, others have been found on the chromosome. The most striking feature of the gene organization of the restriction–modification systems is that both the restriction enzyme and the methyltransferase genes lie close to one another. In some cases, the two genes are separated by a single, small open reading frame that appears to control the expression of the system. The proximity of the methyltransferase and restriction enzyme genes may be necessary for some kind of coordinate expression at the transcriptional level or at the translational level, or this may be important to prevent their separation by recombination.

Sequences are now known for more than 100 restriction enzymes and 150 methyltransferase genes. There is no

significant sequence similarity between the restriction endonuclease gene and the methyltransferase gene of any cognate systems. Protein sequence alignment of several methyltransferases indicates that the m5C-methyltransferases contain several core motifs that are highly conserved, along with a variable target recognition domain. While the N6A-methyltransferases and N4C-methyltransferases also show similarities, they contain only a few conserved protein motifs.

Other bacterial restriction–modification systems include type I and type III enzymes. The former comprises three subunits: the R subunit, which contains the nuclease; the M subunit, which contains the methylase; and the S subunit, which contains the DNA sequence-specificity determinant. The cleavage can occur as far as–1000 bp away from the recognition site. Adenosine triphosphate (ATP) and AdoMet are required for the cleavage activity. Type III enzymes comprise two subunits: a methylation subunit (Mod) and a restriction subunit (Res). They do not require ATP for cleavage and the cut site occurs fairly close to the recognition site.

Modification Methylases

Each type II restriction enzyme has a counterpart modification methylase. This binds to the same recognition sequence and methylates one nucleotide within that sequence (**Table 1**). The methyltransferase uses AdoMet as a substrate and transfers the methyl group from AdoMet to either a cytosine or an adenine residue within the recognition sequence. This methylation renders the site insensitive to cleavage by the restriction endonuclease. Three types of methylations are used to provide protection against the cognate restriction enzyme: N_6-methyladenine, N_4-methylcytosine and 5-methylcytosine. Many laboratory strains of *Escherichia coli* contain three site-specific DNA methylases: Dam methylase, Dcm methylase and *Eco*KI methylase.

The crystal structure of the *Hha*I methyltransferase complexed with its cognate-binding site has revealed that the enzyme flips its target base out of the DNA helix. The structure of the enzyme suggests that it comprises two domains: a small domain that has evolved the sequence-specific DNA-binding apparatus and a large domain that contains the catalytic apparatus. It appears that flipping a base out of the DNA helix is a common feature employed by other proteins that chemically modify bases in DNA.

Methylation is found in C residues of CG sequences in animals and CNG sequences in plant cells. Although the biological role of methylation in bacteria is fairly clear, its importance in eukaryotes is still unclear. After eukaryotic DNA replication, a maintenance methylase ensures that all of the sites that were methylated in parental DNA are methylated in daughter DNA. Methylation may be responsible for tissue-specific inactivation of genes during development. Bestor and co-workers have suggested that the primary function of mammalian DNA methyltransferase may be to suppress the parasitic sequence elements that are present in the genome, including endogenous retroviruses and transposable elements.

Restriction–Modification System *in vivo*

An important characteristic of restriction enzymes is their ability to discriminate their recognition sites from all other sites. This is especially true *in vivo*. A restriction enzyme would be lethal if it cleaved DNA readily at any sequence other than its recognition site. In a cell containing a restriction–modification system, all the recognition sites on the chromosome are methylated by the methyltransferases. Sequences that differ from the recognition site by one base pair remain unmethylated. Under certain conditions, restriction enzymes cleave DNA both at their recognition sites and at a limited number of additional sites that generally differ from the recognition site by one base pair. Under optimal conditions, the ratio of activities at cognate versus noncognate sites is very large. Reactions at noncognate sites proceed via two successive single-strand breaks. Under similar conditions, the same enzyme produces double-strand breaks at the cognate sites. It appears that the single-strand nicks that arise as a result of cleavage at noncognate sites are readily sealed by the DNA ligase that is present in the cell with no physiological consequences.

A hemimethylated DNA, which has a methyl group on one strand only, is a preferred substrate for the methylase but not for the restriction enzymes, which generally cleave only when the recognition site is unmethylated on both strands. This is another important aspect of restriction *in vivo*. Immediately after replication, there is hemimethylated DNA within a cell. The genome needs to be resistant to double-strand cleavage by the restriction endonuclease that is present in the cell. In some cases, one strand may be nicked. In these cases, the DNA ligase can readily seal the nicks. The ability of ligase to repair the nicks can be critical, especially when bacterial cells express high levels of restriction endonucleases and the cognate methylase fails to provide complete protection.

Type II restriction–modification systems are not easily lost from their host cell. The progeny of cells that lose a restriction–modification system are unable to modify a sufficient number of recognition sites in their chromosome to protect them from lethal attack by the remaining molecules of restriction enzyme. Kobayashi and co-workers have hypothesized that this capacity of restriction–modification systems to act as a selfish genetic element may have contributed to the spread and maintenance of restriction–modification systems. This hypothesis remains to be tested.

Although in most restriction–modification systems the expression of the restriction enzyme and that of methyltransferase are tightly linked, it is still possible for a strain

to have lost its restriction endonuclease but have retained its methyltransferase intact.

Changing the Sequence Specificity of Restriction Enzymes

Although the type II enzymes are very useful in manipulating recombinant DNA, they are not suitable for producing large DNA segments or for genome engineering. For example, restriction enzymes that recognize six base pairs result in cuts as often as every 4096 bases. Even rare cutters (enzymes that recognize 8 bp-long sequences) cut DNA once every 65 536 bases on average. So far, only a limited number of restriction enzymes that recognize sequences longer than six base pairs have been identified. In many instances, it is preferable to have fewer but longer DNA strands, especially during genome mapping. Therefore, a long-term goal in the field of restriction–modification enzymes has been to generate novel restriction endonucleases with longer restriction sites by mutating or engineering existing enzymes.

Before restriction enzymes can cleave DNA, they must bind to the correct DNA sequence. Thus, they have a dual function, namely DNA recognition and DNA cleavage. In the case of type II enzymes, these functions overlap each other. Several methods, including the bacteriophage P22 challenge-phage system, have been applied for the selection of mutations that alter sequence specificity in restriction–modification enzymes. However, attempts by genetic manipulation of the existing type II enzymes to generate new specificities, particularly longer recognition sites, have not been successful. This may simply be because multiple mutations are needed before a change in specificity can be achieved. Alternatively, since the DNA recognition and catalytic functions overlap each other in type II enzymes, it is possible that attempts to change amino acid residues within the DNA recognition domain that are responsible for the sequence specificity may also affect the catalytic domain. Changes in the DNA-binding domain may alter the geometry of catalytic site; this is probably accompanied by a drop in cleavage activity over several orders of magnitude. Type II enzymes may simply not be suitable subjects for changing sequence specificity.

Researchers have tried to generate universal restriction enzymes by combining type IIs enzymes like *Fok*I with properly designed oligonucleotide adaptors. However, this method is not as useful as chimaeric nucleases (see later), since the target needs to be single-stranded DNA.

The type IIs enzymes, for example *Fok*I, have been shown to contain two domains: one responsible for DNA sequence recognition and the other for DNA cleavage (**Figure 1**). An elegant study of the crystal structures of *Fok*I with and without DNA by Aggarwal and co-workers has shown this model to be correct. Based on the modular nature of type IIs endonucleases, Chandrasegaran and co-workers postulated that these enzymes probably evolved

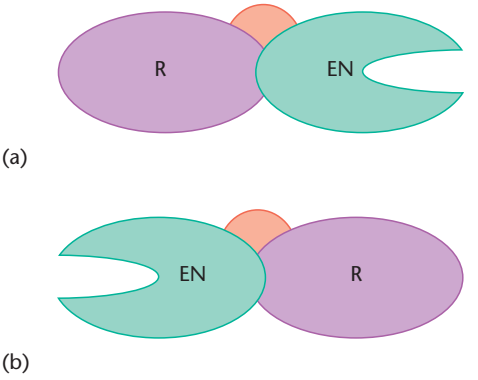

Figure 1 Representation of multimodular enzymes. (a) *Fok*I (a type IIs restriction enzyme); (b) I-*Tev*I (a homing endonuclease). R, recognition domain; EN, endonuclease domain.

by random fusions of the DNA-binding domains to nonspecific endonucleases. Over time, these fusions were further refined into sequence-specific type IIs restriction enzymes by acquiring allosteric interaction between the recognition domain and the catalytic domain. Recent studies suggest that the modular architecture of type IIs enzymes is much more common in nature than previously thought. I-*Tev*I, an intron-encoded homing endonuclease, appears to have a similar bipartite structure (**Figure 1**). Unlike *Fok*I, in which the recognition domain is at the N-terminus and the cleavage domain is at the C-terminal third of the molecule, the homing endonuclease I-*Tev*I appears to be an enzyme with an N-terminal catalytic domain and C-terminal DNA-binding domain connected by a flexible linker. Recent studies suggest that similar multimodular endonuclease fusions may be much more prevalent. R2 retrotransposon endonuclease, *Drosophila* P1 transposase and Rec BCD enzyme involved in recombination may fall into this category.

The type IIs enzymes appear to be ideal candidates for changing sequence specificities. The modular structure of *Fok*I endonuclease suggested that it might be feasible to construct chimaeric nucleases with novel sequence specificities by linking other DNA-binding proteins to the cleavage domain of *Fok*I. Chandrasegaran and co-workers have created the first chimaeric nucleases by fusing the isolated nuclease domain of *Fok*I to other sequence-specific DNA-binding proteins. These include the three common eukaryotic DNA-binding motifs, namely the helix–turn–helix motif, the zinc-finger motif and the basic helix–loop–helix protein (b-HLH) containing a leucine zipper motif (**Figure 2**). These hybrid enzymes cleave the DNA at the binding site preferred by the DNA-binding proteins. Increased levels of ligase within cells have been utilized for the production of chimaeric nucleases. Since there are no counterpart methylases available for the chimaeric nucleases, production of these enzymes *in vivo* is lethal to cells. By increasing the levels of the DNA ligase within the cells, the clones carrying the chimaeric nucleases are made more viable.

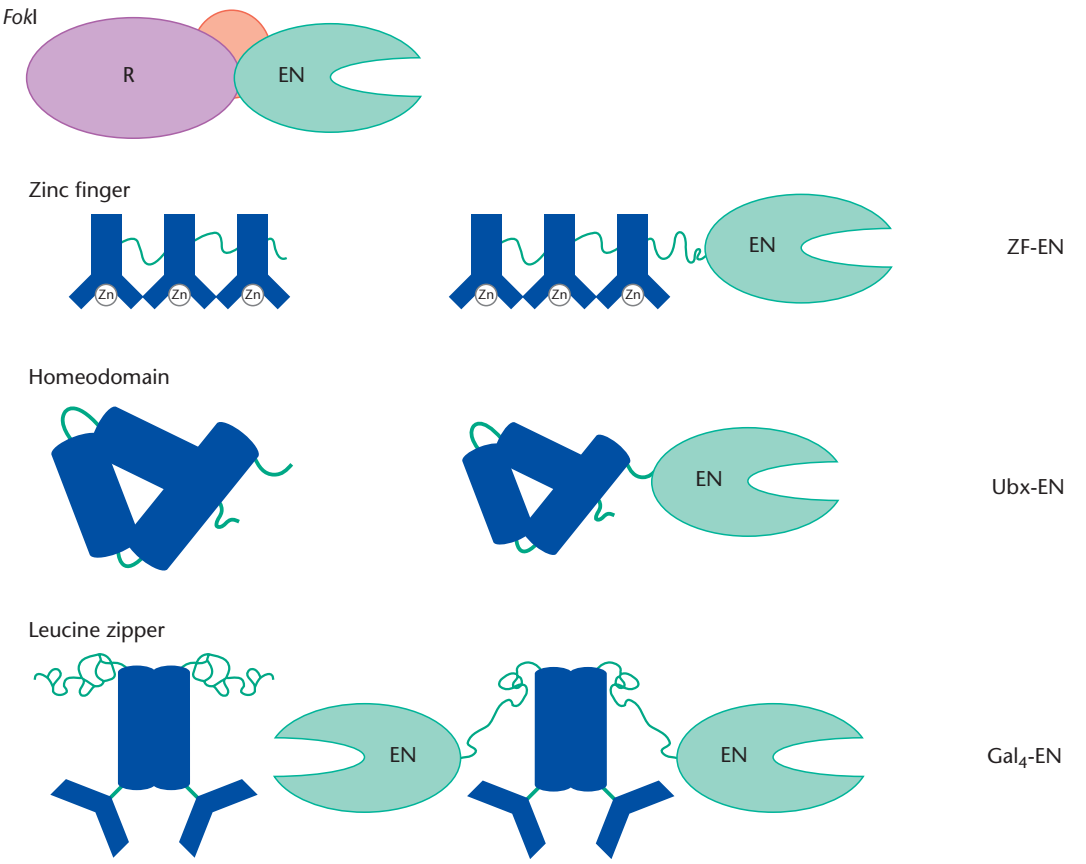

Figure 2 Representation of various chimaeric nucleases. ZF-EN, Ubx-EN and Gal_4-EN were made by fusing the isolated nuclease domain (EN) of *FokI* to the zinc-finger motif, helix–turn–helix motif and Gal_4, respectively. R, recognition domain; EN, endonuclease domain.

The most important chimaeric restriction endonucleases are those based on zinc-finger DNA-binding proteins. Each individual zinc finger, a peptide of about 30 amino acids, recognizes three bases along the DNA. These proteins, like many sequence-specific DNA-binding proteins, bind to the DNA by inserting an α helix into the major groove of the double helix. The crystallographic structure of the three zinc-finger proteins bound to cognate sites reveals that each finger interacts with a triplet within the DNA substrate. Each finger, because of variations of certain key amino acids from one zinc finger to the next, makes its own unique contribution to DNA-binding affinity and specificity. Because they appear to bind as independent modules, the zinc fingers can be linked together in a peptide designed to bind a predetermined DNA site. In theory, one can design a zinc finger for each of the 64 possible triplets; by using a combination of these fingers, one could design a protein for sequence-specific recognition of any segment of DNA. Studies attempting to understand the rules relating to zinc-finger sequences as well as their DNA-binding preferences and redesigning of DNA-binding specificities of zinc-finger protein are under way. An alternate approach to the design of zinc-finger proteins with new specificities involves the selection of desirable mutants from a library of randomized zinc fingers displayed on phage. The ability to design or select zinc-finger proteins with desired specificity implies that DNA-binding proteins containing zinc fingers will be made to order. Therefore, one could design zinc-finger nucleases (ZFNs) that will cut DNA at any preferred site by making fusions of zinc-finger proteins to the cleavage domain of *Fok*I endonuclease. Zinc-finger proteins, because of their modular nature, offer an attractive framework for chimaeric nucleases with tailor-made sequence specificities.

Chandrasegaran and co-workers have fused three zinc fingers to the nuclease domain and achieved cleavage at the predicted 9-bp recognition site. It is immediately obvious that, by combining different zinc fingers together, numerous new DNA-binding specificities and cleavage patterns could be achieved. Pabo and co-workers have reported the design of poly zinc-finger proteins that bind to 18-bp DNA sites with high affinity. These proteins could be converted into site-specific cleavage enzymes by linking them to the *Fok*I cleavage domain.

How might these ZFNs be used in genome engineering? One approach would be to recruit the preexisting cellular machinery to achieve this goal. In somatic or vegetative cells of many different organisms, homologous recombination is

used to repair DNA damage, especially double-strand breaks (Porteus *et al*., 2006). Carroll and co-workers have shown that making a targeted double-strand break would greatly stimulate homologous recombination between the exogenous DNA and a chromosomal sequence in frog oocytes. Such experiments have been performed using intron-encoded homing endonucleases in yeast, cultured mammalian cells and plant cells. In collaboration with Dana Carroll's laboratory, we have successfully used ZFNs to stimulate homologous recombination in frog oocytes. Carroll's group has extended this approach to show specific chromosomal cleavage and mutagenesis of the yellow (Y) gene by ZFNs in *Drosophila melanogaster*. More recently, Porteus and Baltimore have described a gene targeting system based on the correction of a mutated green fluorescent protein (GFP) gene to show that chimaeric nucleases can stimulate gene targeting in human somatic cells. Recent studies have reliably used ZFNs to introduce chromosomal double-strand breaks and stimulate homologous recombination (HR) in *C. elegans*, cultured human cells (Moehle *et al*., 2007; Urnov *et al*., 2005) and *Arabidopsis*. Gene targeting using ZFNs is an emerging new technology. These developments indicate that ZFNs could be used to correct human genetic defects in the near future.

Summary

Since their discovery, type II restriction enzymes have played a crucial role in the development of biotechnology and the field of molecular biology. The restriction and modification enzymes, which protect bacteria from phage or foreign DNA, are important because they provide reagents for recombinant DNA technology. They are the essential tools for manipulating DNA. Restriction enzymes play an important role in the cloning and sequencing of numerous DNA fragments. Restriction mapping using restriction enzymes involves the cleavage of DNA at specific sites followed by the determination of the length of the DNA fragments by gel electrophoresis. The physical map is constructed after digesting the DNA of interest with a variety of enzymes. Thus, restriction enzymes have played a key role in the assembly of a physical map of the human genome.

Although the type II enzymes are useful in manipulating recombinant DNA, they are not suitable for producing large DNA fragments or for genome engineering. Restriction enzymes that recognize 15–18 bp sites would be invaluable in genome engineering experiments. Because the DNA recognition and catalytic functions of type II enzymes overlap each other, attempts to change the sequence specificity of these enzymes have not been successful. The type IIs enzymes appear to be ideal candidates for change sequence specificity. The modular nature of *Fok*I restriction endonuclease has made it possible to construct chimaeric nucleases by linking other DNA-binding proteins to the cleavage domain of *Fok*I. The modular nature of zinc-finger proteins enables the design or selection of peptides that will bind DNA at any predetermined site. The convergence of these two areas of research makes it possible to create artificial nucleases that will cut DNA near a predetermined site. By using these ZFNs, one can make targeted double-strand breaks within a chromosome, and thereby stimulate homologous recombination of exogenous DNA with a chromosomal sequence.

A silent revolution is taking place in the field of restriction–modification enzymes. Scientists are gathering important information about structure, and about the mechanism of DNA recognition and DNA cleavage by many restriction endonucleases. There is real excitement about the possibility of making 'artificial' restriction enzymes that will recognize a particular site within a genome and cleave near that site. We may be able to generate many novel enzymes with tailor-made sequence specificities that are desirable for various applications. Ultimately, we might be able to target specific genes for cleavage within cells. In the future, ZFNs are likely to provide the second generation of the molecular scissors that are such important diagnostic and therapeutic reagents for the research community. Also, the complete nucleotide sequence of the genomes of many organisms, including the human genome, are now known. The availability of chimaeric nucleases that target a specific site within a genome should make it feasible to carry out gene therapy.

Acknowledgement

The work on chimaeric nucleases in Prof Chandrasegaran's lab. is funded by a grant from NIH (GM 53923).

References

Mani M, Kandavelou K, Wu J and Chandrasegaran S (2003) Restriction enzymes. In: Cooper DN (ed.) *Nature Encyclopedia of the Human Genome*, vol. 5, pp. 31–35. London: Nature Publishing Group.

Moehle EA, Rock JM, Lee YL *et al*. (2007) Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proceedings of the National Academy of Sciences of the USA* **104**: 3055–3060.

Porteus MH, Connelly JP and Pruett SM (2006) A look to future directions in gene therapy research for monogenic diseases. *PLoS Genetics* **2**: e133.

Urnov FD, Miller JC, Lee YL *et al*. (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**: 646–651.

Further Reading

Aggarwal AK and Wah DA (1998) Novel site-specific DNA endonuclease. *Current Opinion in Structural Biology* **8**(1): 19–25.

Berg JM (1995) Zinc finger domains: from prediction to design. *Accounts of Chemical Research* **28**: 14–19.

Bibikova M, Beumer K, Trautman JK and Carroll D (2003) Enhancing gene targeting with designed zinc finger nucleases. *Science* **300**: 764.

Bibikova M, Carroll D, Segal DJ *et al*. (2001) Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Molecular and Cellular Biology* **21**: 289–297.

Bibikova M, Golic M, Golic KG and Carroll D (2002) Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc finger nucleases. *Genetics* **161**: 1169–1175.

Carroll D (1996) Homologous genetic recombination in *Xenopus*: mechanisms and implications for gene manipulation. *Progress in Nucleic Acid Research and Molecular Biology* **54**: 101–125.

Choo Y and Klug A (1997) Physical basis of a protein–DNA recognition code. *Current Opinion in Structural Biology* **7**: 117–125.

Durai S, Mani M, Kandavelou K *et al*. (2005) Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Research* **33**: 5978–5990.

Gormley NA, Watson MA and Halford SE (2005) Bacterial restriction–modification systems, doi.wiley.com/10.1038/npg.els.0003897.

Kandavelou K, Mani M, Durai S and Chandrasegaran S (2004) Engineering and applications of chimeric nucleases. In: Pingoud AM (ed.) *Nucleic Acids and Molecular Biology*, vol. 14, pp. 413–434. Berlin: Springer-Verlag.

Kandavelou K, Mani M, Durai S and Chandrasegaran S (2005) "Magic" scissors for genome surgery. *Nature Biotechnology* **23**: 686–687.

Kandavelou K, Mani M, Reddy SPM and Chandrasegaran S (2005) Ligation: theory and practice, doi.wiley.com/10.1038/npg.els.0003838.

Kim J-S and Pabo CO (1998) Getting a handhold on DNA: design of poly zinc finger proteins with femtomolar dissociation constants. *Proceedings of the National Academy of Sciences of the USA* **95**: 2812–2817.

Kim Y-G, Cha J and Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusion to *Fok*I cleavage domain. *Proceedings of the National Academy of Sciences of the USA* **93**: 1156–1160.

Li L, Wu LP and Chandrasegaran S (1992) Functional domains in FokI restriction endonuclease. *Proceedings of the National Academy of Sciences of the USA* **89**: 4275–4279.

New England Biolabs (2000–2001) *New England Biolabs Catalog and Technical Reference*. Beverly, MA: New England Biolabs.

Porteus MH and Baltimore D (2003) Chimeric nucleases stimulate gene targeting in human cells. *Science* **300**: 763.

Roberts RJ and Cheng X (1998) Base flipping. *Annual Review of Biochemistry* **67**: 181–198.

Roberts RJ and Halford SE (1993) Type II restriction endonuclease. In: Linn SM, Lloyd RS and Roberts RJ (eds) *Nuclease*, pp. 35–38. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Roberts RJ, Vincze T, Posfai J and Macelis D (2003) REBASE: restriction enzymes and methyltransferase. *Nucleic Acids Research* **31**: 418–421(http://rebase.neb.com/rebase/rebase.html).

Samuelson JC, Morgan RD, Benner JS *et al*. (2006) Engineering a rare-cutting restriction enzyme: genetic screening and selection of NotI variants. *Nucleic Acids Research* **34**: 796–805.

Sistla S and Rao DN (2004) *S*-Adenosyl-L-methionine-dependent restriction enzymes. *Critical Reviews in Biochemistry and Molecular Biology* **39**: 1–19.

Szybalski W, Kim SC, Hasan N and Podhajska AJ (1991) Class-IIs restriction enzymes – a review. *Gene* **100**: 13–26.

Williams RJ (2001) Restriction endonucleases and their uses. *Methods in Molecular Biology* **160**: 409–429.