Restriction endonucleases and their recognition sequences

THOMAS R. GINGERAS
Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, U.S.A.

Any survey of Type-II restriction endonucleases requires a consideration of the various specificities that are the hallmark of these enzymes. The latest compilation of the restriction-enzyme list contains more than 200 enzymes, which exhibit more than 50 unique specificities (Roberts, 1980). Because these enzymes have demonstrated their greatest value as biochemical tools, classification of these enzymes into groups based on common themes inherent within their varied specificities provides an attractive scaffolding for this presentation.

The unique specificities of the currently known restriction endonucleases are listed in Table 1 based on the size of their recognition sites. These recognition sequences vary in length from four to six base pairs. An important addition to the tetranucleotide group of enzymes has been made with the recent discovery of enzyme RsaI (GTAC) by J. F. Gardner & S. Kaplan (unpublished work). With the description of this recognition site, eight of the 16 possible tetranucleotide combinations have now been discovered. Another restriction endonuclease of interest listed within the tetranucleotide group is the enzyme MnlI (CCTC). This enzyme, as well as four others from the pentanucleotide group (Hgal, Hphl, MboII and SfaI), do not demonstrate a property which is common to all other restriction endonucleases thus far observed. These enzymes do not exhibit a twofold rotational symmetry with the sites of DNA-strand scission being symmetrically disposed relative to the twofold axis. Rather, these enzymes introduce a double-stranded break several nucleotides away from their recognition sites. It is tempting to speculate that symmetry is important for recognition and/or cleavage of DNA templates by restriction enzymes.

Several other endonucleases listed within the tetranucleotide group of enzymes possess properties which are unique. Since the initial description of the first Type-II endonuclease (HindII) by Smith and his co-workers (Smith & Wilcox, 1970; Kelly & Smith, 1970), it has been known that the activity of such enzymes can be inhibited by modification within the sequence recognized by each of these enzymes. It subsequently has been shown that either 6-methyladenosine or 5-methylcytosine, when present within its recognition sequence, can inhibit such Type-II enzymes. Enzyme DpnI from *Diplococcus pneumoniae* is unusual in this regard in that it is the only known restriction endonuclease which absolutely requires a methylated base (6-methyladenosine) to be present within its recognition sequence (GGATC) in order for cleavage to occur. Surprisingly, these enzymes do not exhibit a twofold rotational symmetry with the sites of DNA-strand scission being symmetrically disposed relative to the twofold axis. Rather, these enzymes introduce a double-stranded break several nucleotides away from their recognition sites. It is tempting to speculate that symmetry is important for recognition and/or cleavage of DNA templates by restriction enzymes.

Several other endonucleases listed within the tetranucleotide group of enzymes possess properties which are unique. Since the initial description of the first Type-II endonuclease (HindII) by Smith and his co-workers (Smith & Wilcox, 1970; Kelly & Smith, 1970), it has been known that the activity of such enzymes can be inhibited by modification within the sequence recognized by each of these enzymes. It subsequently has been shown that either 6-methyladenosine or 5-methylcytosine, when present within a recognition site, can inhibit such Type-II enzymes. Enzyme DpnI from *Diplococcus pneumoniae* is unusual in this regard in that it is the only known restriction endonuclease which absolutely requires a methylated base (6-methyladenosine) to be present within its recognition sequence (GGATC) in order for cleavage to occur. Surprisingly, these enzymes do not exhibit a twofold rotational symmetry with the sites of DNA-strand scission being symmetrically disposed relative to the twofold axis. Rather, these enzymes introduce a double-stranded break several nucleotides away from their recognition sites. It is tempting to speculate that symmetry is important for recognition and/or cleavage of DNA templates by restriction enzymes.

Several other endonucleases listed within the tetranucleotide group of enzymes possess properties which are unique. Since the initial description of the first Type-II endonuclease (HindII) by Smith and his co-workers (Smith & Wilcox, 1970; Kelly & Smith, 1970), it has been known that the activity of such enzymes can be inhibited by modification within the sequence recognized by each of these enzymes. It subsequently has been shown that either 6-methyl adenosine or 5-methylcytosine, when present within a recognition site, can inhibit such Type-II enzymes. Enzyme DpnI from *Diplococcus pneumoniae* is unusual in this regard in that it is the only known restriction endonuclease which absolutely requires a methylated base (6-methyladenosine) to be present within its recognition sequence (GGATC) in order for cleavage to occur. Surprisingly, these enzymes do not exhibit a twofold rotational symmetry with the sites of DNA-strand scission being symmetrically disposed relative to the twofold axis. Rather, these enzymes introduce a double-stranded break several nucleotides away from their recognition sites. It is tempting to speculate that symmetry is important for recognition and/or cleavage of DNA templates by restriction enzymes.

Several other endonucleases listed within the tetranucleotide group of enzymes possess properties which are unique. Since the initial description of the first Type-II endonuclease (HindII) by Smith and his co-workers (Smith & Wilcox, 1970; Kelly & Smith, 1970), it has been known that the activity of such enzymes can be inhibited by modification within the sequence recognized by each of these enzymes. It subsequently has been shown that either 6-methyladenosine or 5-methylcytosine, when present within a recognition site, can inhibit such Type-II enzymes. Enzyme DpnI from *Diplococcus pneumoniae* is unusual in this regard in that it is the only known restriction endonuclease which absolutely requires a methylated base (6-methyladenosine) to be present within its recognition sequence (GGATC) in order for cleavage to occur. Surprisingly, these enzymes do not exhibit a twofold rotational symmetry with the sites of DNA-strand scission being symmetrically disposed relative to the twofold axis. Rather, these enzymes introduce a double-stranded break several nucleotides away from their recognition sites. It is tempting to speculate that symmetry is important for recognition and/or cleavage of DNA templates by restriction enzymes.

Several other endonucleases listed within the tetranucleotide group of enzymes possess properties which are unique. Since the initial description of the first Type-II endonuclease (HindII) by Smith and his co-workers (Smith & Wilcox, 1970; Kelly & Smith, 1970), it has been known that the activity of such enzymes can be inhibited by modification within the sequence recognized by each of these enzymes. It subsequently has been shown that either 6-methyladenosine or 5-methylcytosine, when present within a recognition site, can inhibit such Type-II enzymes. Enzyme DpnI from *Diplococcus pneumoniae* is unusual in this regard in that it is the only known restriction endonuclease which absolutely requires a methylated base (6-methyladenosine) to be present within its recognition sequence (GGATC) in order for cleavage to occur. Surprisingly, these enzymes do not exhibit a twofold rotational symmetry with the sites of DNA-strand scission being symmetrically disposed relative to the twofold axis. Rather, these enzymes introduce a double-stranded break several nucleotides away from their recognition sites. It is tempting to speculate that symmetry is important for recognition and/or cleavage of DNA templates by restriction enzymes.

Several other endonucleases listed within the tetranucleotide group of enzymes possess properties which are unique. Since the initial description of the first Type-II endonuclease (HindII) by Smith and his co-workers (Smith & Wilcox, 1970; Kelly & Smith, 1970), it has been known that the activity of such enzymes can be inhibited by modification within the sequence recognized by each of these enzymes. It subsequently has been shown that either 6-methyladenosine or 5-methylcytosine, when present within a recognition site, can inhibit such Type-II enzymes. Enzyme DpnI from *Diplococcus pneumoniae* is unusual in this regard in that it is the only known restriction endonuclease which absolutely requires a methylated base (6-methyladenosine) to be present within its recognition sequence (GGATC) in order for cleavage to occur. Surprisingly, these enzymes do not exhibit a twofold rotational symmetry with the sites of DNA-strand scission being symmetrically disposed relative to the twofold axis. Rather, these enzymes introduce a double-stranded break several nucleotides away from their recognition sites. It is tempting to speculate that symmetry is important for recognition and/or cleavage of DNA templates by restriction enzymes.
the enzyme DpnII, produced by another strain of D. pneumoniae, recognizes the same sequence, but is inhibited by this same methylated base.

The significance of the occurrence of modified nucleotides in eukaryotic DNA is not as yet understood, although a relationship between modification and gene expression has been proposed by several workers (Waalwijk & Flavell, 1978; Bird et al., 1979). By using appropriate restriction endonucleases a simple method can be used in order to probe and map the location of such modified nucleotides (Bird & Southern, 1978). Several pairs of restriction endonucleases would be useful for this purpose: MboI (GATC) (Gelinas et al., 1977) and Sau3A (CCATG) (Sussenbach et al., 1976), HpaII (CCGG) (Mann & Smith, 1977; Garfin & Goodman, 1974) and MspI (CCGG) (Waalwijk & Flavell, 1978); EcoRII (CC(A/T)GG) and BstI [CC(A/T)GG] (J. Brooks, personal communication). Each of these pairs of restriction endonucleases could be used to investigate the modification pattern present within a genome.

Two recent additions to the pentanucleotide class of restriction enzymes are Ddel from Desulfitobrio desulfuricans (R. A. Manula & R. Meagher, unpublished work) and Fnu4HI from Fusobacterium nucleatum 4H (Leung et al., 1979). The recognition sites of enzymes Ddel (CTNAG) and Fnu4HI (GCNGC) are similar, owing to the degeneracy present in the central nucleotide of the recognition sites. These two enzymes, along with HinfI (GANTC) and Asul (GGNCC), constitute a class of restriction endonucleases which may be useful in identifying specific portions of structural genes.

The Ddel enzyme also provides an example of the newer methods for the determination of recognition sites. A computer-assisted strategy which results in a unique and small number of predictions for the recognition sequence of Type-II enzymes has been expanded to include enzymes which recognize degenerate sequences. Their use has led to the elucidation of two new recognition sequences: enzymes XhoI (RGATCY) and Ddel (CTNAG).

The hexanucleotide group of endonucleases contains three new enzymes: Aevl (GRCYG) from Anabaena cylindrica (DeWaard et al., 1978), SphI (GCAATGC) from Streptomyces phaeochromogenes (Fuchs et al., 1980), and XmaIII (CGCCCG) from Xanthomonas malvacearum (Kunkel et al., 1979). These and other enzymes belonging to the hexanucleotide group of restriction endonucleases are useful in constructing recombinant DNA molecules.

Finally, a strategy which makes use of the varied specificities of each of the restriction endonucleases listed in Table 1 has been developed for the purpose of checking the accuracy of newly derived primary DNA sequences.

I thank R. J. Roberts for his valuable suggestions and discussion concerning the topics presented. I also thank Marie Moschitta for her efforts in the preparation of this abstract. T. R. G. is a National Institutes of Health Postdoctoral Fellow.


Sequence determination of restriction-endonuclease recognition sites

NIGEL L. BROWN
Department of Biochemistry, University of Bristol, University Walk, Bristol BS8 1TD, U.K.

Restriction endonucleases are widely used in the study of the physical structure of DNA molecules, and in the genetic manipulation of DNA molecules in vitro. They also provided interesting and useful model systems for the study of protein–DNA interactions. A knowledge of the DNA sequence that is recognized by a restriction endonuclease, and of the structure of the termini of the DNA fragments, is essential to the prediction of the uses to which the enzyme may be put.

Class-II restriction endonucleases recognize a specific sequence or a family of related sequences in DNA, and specifically cleave DNA within or near that recognition sequence. The DNA fragments produced will have identical termini only if the enzyme recognizes a unique sequence in DNA and cleaves within that sequence. Endonucleases that cleave within a family of related sequences or cleave away from their recognition sequence will generate DNA fragments with non-identical termini. The relationship between the cleavage sites in each strand of duplex DNA is not the same for all restriction endonucleases, and the DNA fragments produced may have fully base-paired termini, or may have 5’ or 3’ single-stranded extensions. All class-II restriction endonucleases so far examined cleave DNA to generate 5’-phosphoryl and 3’-hydroxyl termini. Many of the methods used to determine the recognition and cleavage specificities of class-II restriction endonucleases have depended on 32P-phosphate end-labelling of a population of DNA fragments produced by cleavage with the restriction endonuclease, and analysis of 32P-labelled oligonucleotides produced by non-specific endonucleolytic cleavage of the end-labelled fragments (see Roberts, 1976). Different end-labelling methods must be used for restriction endonucleases which cleave to give 5’ or 3’ terminal extensions; and end-labelling methods are of limited use for enzymes which cleave DNA away from their recognition sequence, as the terminal sequences of the DNA fragments are not related to the recognition sequence of the enzyme. Computer methods are available for predicting the recognition sequences of class-II restriction endonucleases, in which the cleavage pattern of the enzyme on a fully sequenced DNA molecule is determined, and is used to predict the recognition sequence. Such methods give no information on the cleavage specificity of the enzyme and on the type of DNA fragment termini produced.

A method has been described whereby individual cleavage sites of any class-II restriction endonuclease can be sequenced, and the recognition sequence can be predicted from the analysis of several cleavage sites (Brown & Smith, 1977). This method has subsequently been modified (Brown & Smith, 1980) for use in conjunction with the chemical method (Maxam & Gilbert, 1977) or the chain-termination method (Sanger et al., 1977) for DNA sequence analysis. However, the method requires reasonably large amounts of DNA (approx. 5 pmol), and requires that