

# Restriction enzymes

## Definition:

A **restriction enzymes** (or restriction endonucleases), are bacterial enzymes that cleave both strands of DNA at or near specific recognition nucleotide sequences (4 -8 bp ) known as **restriction sites**.

## 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 term restriction enzyme originated from the studies of phage  $\lambda$  and the phenomenon of host-controlled restriction and modification of a bacterial virus. The phenomenon was first identified in work done in the laboratories of Salvador Luria and Giuseppe Bertani in early 1950s. It was found that a bacteriophage  $\lambda$  that can grow well in one strain of *Escherichia coli*, for example *E. coli* C, when grown in another strain, for example *E. coli* K, its yields can drop significantly, by as much as 3-5 orders of magnitude. The *E. coli* K host cell, known as the restricting host, appears to have the ability to reduce the biological activity of the phage  $\lambda$ . If a phage becomes established in one strain, the ability of that phage to grow also becomes restricted in other strains. In the 1960s, it was shown in work done in the laboratories of Werner Arber and Matthew Meselson that the restriction is caused by an enzymatic cleavage of the phage DNA, and the enzyme involved was therefore termed a restriction enzyme.

The three dimensional structure of the restriction enzyme allows it to fit perfectly in the groove formed by the two strands of DNA molecule. When attached to the DNA, the enzyme slides along the double helix until it recognizes a specific sequence of base pairs which signals the enzyme to stop sliding. The enzyme, then digests the DNA at that site. If the a specific site occurs in more than one location on a DNA molecule, a restriction enzyme will make a cut at each of those sites, resulting in multiple fragments of DNA.

Each restriction enzyme recognizes a specific nucleotide base sequence, termed **palindrome**, in the DNA called **restriction site**.



Restriction enzymes recognize a specific sequence of nucleotides and produce a double-stranded cut in the DNA. The recognition sequences usually vary between **4** and **8** nucleotides, and many of them are palindromic, meaning the base sequence reads the same backwards and forwards. In theory, there are two types of palindromic sequences that can be possible in DNA. The mirror-like palindrome is similar to those found in ordinary text, in which a sequence reads the same forward and backwards on a single strand of DNA strand, as in **GTAATG**. The **inverted repeat palindrome** is also a sequence that reads the same forward and backwards, but the forward and backward sequences are found in complementary DNA strands (i.e., of double-stranded DNA), as in **GTATAC** (**GTATAC** being complementary to **CATATG**).

Inverted repeat palindromes are more common and have greater biological importance than mirror-like palindromes.

**EcoRI** digestion produces "**sticky**" ends,



**Note:** The staggered termini are called sticky ends because they can reassociate with another by hydrogen bonding.

whereas **SmaI** restriction enzyme cleavage produces "**blunt**" ends:



Recognition sequences in DNA differ for each restriction enzyme, producing differences in the length, sequence and strand orientation (5' end or the 3' end) of a sticky-end "overhang" of an enzyme restriction.

Different restriction enzymes that recognize the same sequence are known as **neoschizomers**. These often cleave in different locales of the sequence.

Different enzymes that recognize and cleave in the same location are known as **isoschizomers**.

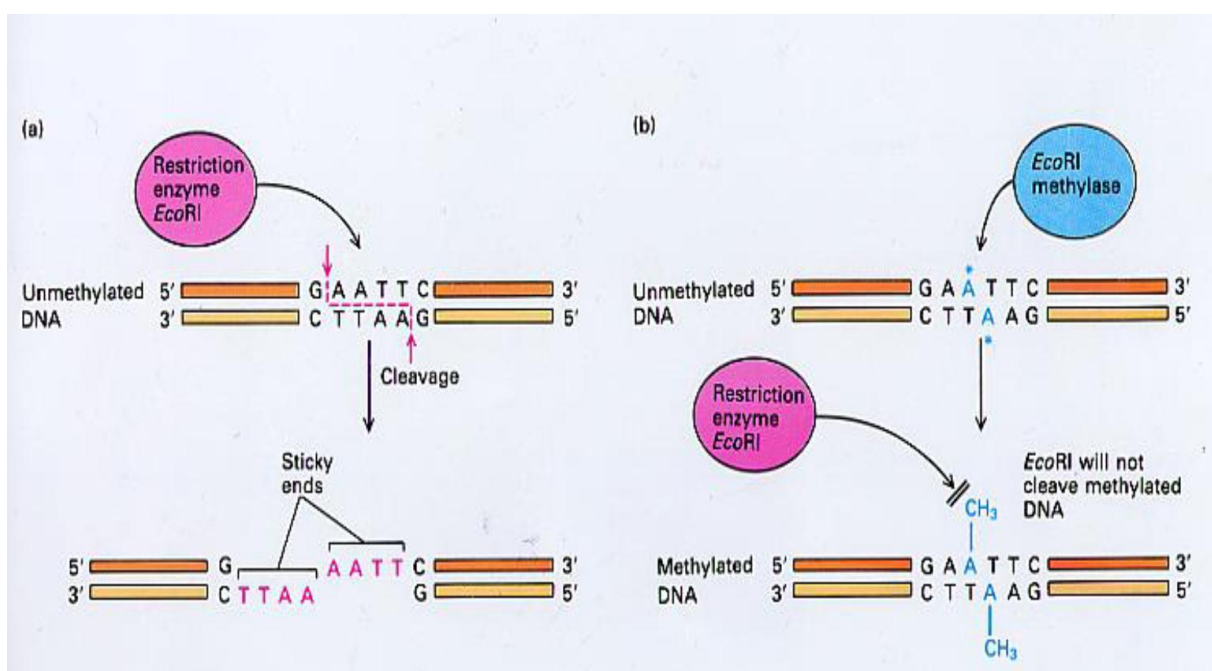
## Methylation

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**.

Restriction enzymes are part of a **pathogen resistance pathway** known as **host-controlled modification**. In restriction modification systems, bacteria modify their own DNA to protect against cleavage by endogenous restriction enzymes. However, the bacterial restriction enzymes cleave phage DNA to prevent infection.

Restriction enzymes have partner **methyltransferases**, which methylate DNA. In bacteria, this serves to protect the bacterial genome from digestion by its own endogenous enzymes.

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.



## Types

Naturally occurring restriction endonucleases are categorized into four groups (Types I, II III, and IV) based on their composition and enzyme cofactor requirements, the nature of their target sequence, and the position of their DNA cleavage site relative to the target sequence. All types of enzymes recognize specific short DNA sequences and carry out the endonucleolytic cleavage of DNA to give specific fragments with terminal 5'-phosphates.

They differ in their recognition sequence, subunit composition, cleavage position, and cofactor requirements, as summarised below:

**Type I** enzymes cleave at sites remote from recognition site; require both ATP and S-adenosyl-L-methionine to function; multifunctional protein with both restriction and methylase activities.

**Type II** enzymes cleave within or at short specific distances from recognition site; most require magnesium; single function (restriction) enzymes independent of methylase.

**Type III** enzymes cleave at sites a short distance from recognition site; require ATP (but do not hydrolyse it); S-adenosyl-L-methionine stimulates reaction but is not required; exist as part of a complex with a modification methylase.

**Type IV** enzymes target modified DNA, e.g. methylated, hydroxymethylated and glucosyl-hydroxymethylated DNA.

**Note:** Type II restriction enzymes cannot recognize methylated DNA, leading to much frustration if the site you want to cut is methylated. However, Type IV restriction enzymes can recognize methylated DNA, providing you with an option when you have methylated DNA.

**Note:** Restriction enzymes do not discriminate between prokaryotic and eukaryotic DNA.

## Nomenclature

Since their discovery in the **1970s**, more than **100** different restriction enzymes have been identified in different bacteria. Each enzyme is named after the bacterium from which it was isolated using a naming system based on bacterial genus, species and strain.

**For example**, the name of the **EcoRI** restriction enzyme was derived from:

<b>E</b>	Escherichia (genus )
<b>co</b>	coli (species )
<b>R</b>	RY13(strain )
<b>I</b>	First identified order of identification in the bacterium

**The nomenclature of restriction endonucleases follows a general pattern:**

(1) The first letter of the name of genus in which a given enzyme is discovered is written in capital.

(2) This is followed by the first two letters of species name of the organism. These three letters are generally written in italics, e.g., *Eco* from *Escherichia Coli*, *Hin* from *Haemophilus influenzae*, *Hpa* from *Haemophilus parainfluenzae*, etc.

(3) Strain or type identification is depicted as subscript, e. g., *Eco<sub>k</sub>*, if the enzyme is encoded by a plasmid, the plasmid name is written as a subscript, e. g., *EcoRI*.

(4) When an organism produces more than one enzyme, they are identified by sequential Roman numerals, e.g., the different enzymes produced by *H. influenzae* strain Rd are named *HindII*, *HindIII*, etc.

(5) All restriction enzymes are designated by the general symbol R, which is prefixed to their names, e.g., *REcoRI*, *RHindIII*, *RBamHI*, etc. (this is to distinguish them from the corresponding methylases isolated from the same strains; the methylases are prefixed by M).

## Applications

Isolated restriction enzymes are used to manipulate DNA for different scientific applications. They are used to assist insertion of genes into plasmid vectors during gene cloning and protein expression experiments. For optimal use, plasmids that are commonly used for gene cloning are modified to include a short *polylinker* sequence (called the **multiple cloning site**, or **MCS**) rich in restriction enzyme recognition sequences.

**Examples of restriction enzymes: (للاطلاع فقط)**

Enzyme	Source	Restriction sequence
<i>EcoRI</i>	<i>Escherichia coli</i>	5'GAATTC 3'CTTAAG
<i>EcoRII</i>	<i>Escherichia coli</i>	5'CCWGG 3'GGWCC
<i>BamHI</i>	<i>Bacillus amyloliquefaciens</i>	5'GGATCC 3'CCTAGG
<i>HindIII</i>	<i>Haemophilus influenzae</i>	5'AAGCTT 3'TTCGAA
<i>TaqI</i>	<i>Thermus aquaticus</i>	5'TCGA 3'AGCT
<i>NotI</i>	<i>Nocardia otitidis</i>	5'GCGGCCGC 3'CGCCGGCG
<i>HinfI</i>	<i>Haemophilus influenzae</i>	5'GANTCA 3'CTNAGT
<i>Sau3A</i>	<i>Staphylococcus aureus</i>	5'GATC 3'CTAG
<i>PvuII*</i>	<i>Proteus vulgaris</i>	5'CAGCTG 3'GTCGAC
<i>SmaI*</i>	<i>Serratia marcescens</i>	5'CCCGGG 3'GGGCCC
<i>HaeIII*</i>	<i>Haemophilus aegyptius</i>	5'GGCC 3'CCGG
<i>HgaI</i>	<i>Haemophilus gallinarum</i>	5'GACGC 3'CTGCG
<i>AluI*</i>	<i>Arthrobacter luteus</i>	5'AGCT 3'TCGA
<i>EcoRV*</i>	<i>Escherichia coli</i>	5'GATATC 3'CTATAG

Key: \* = blunt ends

N = C or G or T or A

W = A or T