TP1: Restriction-modification Systems

Fundamentos de Biologia Molecular

Curso de Licenciatura em Biologia
2º Ano, 1º Semestre
Ano Letivo 2018/2019

Componente Teórico-Prática

Docente Responsável: Rita Zilhão
Docente TPs: Andreia Figueiredo
TP1: Restriction-modification Systems

- R-M systems discovery
- Nobel prizes
- R-M classification, action mechanisms and function
- R-M type I, II and III
- Other R-M mechanisms
- Practical applications
- 1915: virology was in its infancy.
- Frederick Twort, an English physician, was attempting unsuccessfully to propagate vaccinia virus, the primary component of the smallpox vaccine, on agar plates. Rather than vaccinia, however, the only things growing on Twort’s plates were contaminating bacteria.

**Despite the lack of progress towards his original goal, Twort soon observed that something else was happening:** mysterious “glassy and transparent” spots, which turned out, upon closer examination, to be zones of dead bacteria, would occasionally materialize on his plates.
Bacterial colonies

Area where the phage is dominant. Each plaque represents the lysis of a phage-infected bacterial culture and can be designated as a plaque-forming unit (PFU)

Lysis plaques
TP1: Restriction-modification Systems

1. **Bacteria**
   - Initial state
   - One phage infecting a bacterial cell

2. **Bacterial lawn**
   - After infection
   - Time elapsed
   - Infection spread

3. **Plaque**
   - Final state
   - Cloning of phages
   - Multiple plaques formed

Diagram illustrates the process of phage infection and cloning in bacterial cultures.
TP1: Restriction-modification Systems

A phage infects a single cell in a lawn of sensitive, growing bacteria.

Phage reproduces in the cell, typically yielding about 50 progeny phage per infected cell.

Lysis of the cell releases phage into the medium. The phage diffuses through the medium and infects adjacent cells.

Phage reproduces in these cells, releasing 50 additional progeny phage per infected cell.

These cells lyse, releasing more phage which can then diffuse outward and infect the surrounding cells. Lysis of the cells results in a circular clearing in the lawn of bacteria. This region of visible lysis is called a plaque.
Outcomes of phage infection: Lytic cycle and lysogenic cycle
TP1: Restriction-modification Systems

R-M system

REase - Restriction endonuclease
MTase - Methyltransferase
In 1962, Werner Arber's studies showed that restriction involved changes in the DNA of the phage and was accompanied by degradation of the DNA. It appeared that bacteria could somehow limit phage infection by breaking down the virus DNA.

Several years later, Arber's lab confirmed that the bacterial host was able to modify or restrict viral DNA, as well as foreign bacterial DNA. Arber theorized that each particular bacterial strain generated an endonuclease (enzyme) that could recognize specific sequences of nucleotides in foreign DNAs, and cut them there. Each bacterium would protect its own DNA from the enzyme by adding methyl groups to the specific sites, so the enzyme would not recognize them.

By 1968, two such enzymes had been identified in strains of E. coli, generating much excitement among biochemists. However, it turned out that while they recognized specific DNA sequences, they did not always cut the DNA close to those sites. Since the fragments produced were of random size and character, the enzymes (later known as Type I restriction enzymes) were unsuitable for mapping or sequencing.
• Restriction-modification (R-M) system: consist of two distinct enzymatic activities:
  - a restriction endonuclease that cleaves DNA at a specific recognition sequence
  - a DNA methyltransferase that methylates DNA at the same site and thus prevents cleavage by the cognate restriction enzyme.

• Restriction enzyme and its cognate modification system constitutes the R-M system
In 1972, he purified the first site-specific "Type II" restriction enzyme, known as \textit{Hind II}. The crucial discovery came by chance: Incubating bacteria and phage together, Smith happened to notice that the phage DNA degraded over time. He and his colleagues were successful in purifying the enzyme and they were able to identify the short sequence of 6 base pairs in phage P22 that \textit{Hind II} recognized and specifically cleavage.

Daniel Nathans applied this discovery within genetics. Among other applications, Daniel Nathans used different restriction enzymes on the DNA of the SV40 virus and studied its components, which allowed him to map the virus' genome.

- Smith's enzyme cut it into eleven specific fragments-consistent, manageable pieces, onto which individual genetic activities could be mapped.
- New Type II enzymes were rapidly discovered, and during the next several years, Nathans' group and used to delineate the SV40 genome. They deduced the size and physical order of the fragments in the genome, and created the first cleavage maps of a viral DNA, showing where each restriction enzyme cut it.
TP1: Restriction-modification Systems

The Nobel Prize in Physiology or Medicine 1978

"for the discovery of restriction enzymes and their application to problems of molecular genetics"
The restriction modification (R-M) system is used by bacteria as a way to protect themselves from foreign DNA (e.g., phage DNA) – it is a part of a prokaryote “rudimental” immune system.

This phenomenon was noticed in the 1950 decade – some bacteria strain inhibited virus growth and some didn’t.

Presence of restriction enzymes that cut specific sequences

Endonucleases are restriction enzymes that cut double DNA fragments in specific sequences - these fragments are then destroyed by other endonucleases – protection mechanism.

Restriction enzymes cleave phosphodiester internal, non terminal bonds originating 5’P and 3’OH extremities (except NcI1: originates 3’P and 5’OH extremities)
Restriction enzymes usually cut 4-6 bp long sequences that are frequently **palindromes**\(^1\). They are classified according to subunit composition, cleavage position, sequence specificity and cofactor requirements.

To prevent the restriction of its own DNA, bacteria modifies its DNA by adding methyl (CH\(_3\)) groups.

**TP1: Restriction-modification Systems**

A methyl group is added to the 5-carbon position of cytosine residues in DNA.

---

\(^1\)A **palindrome** is a word, phrase, number, or other sequence of characters which reads the same backward or forward.
• Approximately $\frac{1}{4}$ of known bacteria present R-M systems and $\frac{1}{2}$ of those present more than one R-M system.
### TP1: Restriction-modification Systems

**R-M types** : The majority of R-M systems can be classified as type I, II, or III on the basis of enzyme structure, cofactor requirements, structure of the DNA recognition site, and location of DNA cleavage relative to the recognition site.

<table>
<thead>
<tr>
<th>Type</th>
<th>Protein structure</th>
<th>Restriction &amp; methylation</th>
<th>RE recognition site</th>
<th>Ex of recognition sites</th>
<th>RE cleavage site</th>
<th>ATP required</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Bifunctional enzyme of 3 subunits</td>
<td>Mutually exclusive</td>
<td>Bipartite and asymmetric</td>
<td>EcoAI [GAGNNNNNNNGTCA] EcoKI [AACNNNNNNGTGC]</td>
<td>Non-specific, &gt;1000 bp from recognition site</td>
<td>Yes</td>
</tr>
<tr>
<td>II</td>
<td>Separate endonuclease and methylase</td>
<td>Separate reactions</td>
<td>4-8 bp sequence, often palindromic</td>
<td>EcoRI [G/AATTC] Bal [TGG/CCA]</td>
<td>Same as or close to recognition site</td>
<td>No</td>
</tr>
<tr>
<td>III</td>
<td>Bifunctional enzyme of 2 subunits</td>
<td>Simultaneous</td>
<td>5-7 bp, asymmetric sequence</td>
<td>EcoPI [AGACC] HinfIII [CGAAT]</td>
<td>24-25 bp, downstream of recognition site</td>
<td>Yes</td>
</tr>
</tbody>
</table>

RE- restriction enzyme  
N- A, C, G ou T; R- G ou A
## TP1: Restriction-modification Systems


<table>
<thead>
<tr>
<th>Type</th>
<th>Type I</th>
<th>Type II</th>
<th>Type III</th>
<th>Type IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Features</td>
<td>Oligomeric REase and MTase complex</td>
<td>Separate REase and MTase or combined REase~MTase fusion</td>
<td>Combined REase + MTase complex</td>
<td>Methylation-dependent REase</td>
</tr>
<tr>
<td></td>
<td>Require ATP hydrolysis for restriction</td>
<td>Cleave within or at fixed positions close to recognition site</td>
<td>ATP required for restriction</td>
<td>Cleave at variable distance from recognition site</td>
</tr>
<tr>
<td></td>
<td>Cleave variably, often far from recognition site</td>
<td>Many different subtypes</td>
<td>Cleave at fixed position outside recognition site</td>
<td>Cleave m6A, m5C, hm5C and/or other modified DNA</td>
</tr>
<tr>
<td></td>
<td>‘DEAD-box’ translocating REase</td>
<td></td>
<td>‘DEAD-box’ REase</td>
<td>Many different types</td>
</tr>
<tr>
<td></td>
<td>bipartite DNA recognition domain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Example</td>
<td>e.g. EcoKI</td>
<td>e.g. EcoRI</td>
<td>e.g. EcoP11</td>
<td>No ‘typical’ example</td>
</tr>
<tr>
<td>Genes</td>
<td><em>hSDR, hsdM, hsdS</em></td>
<td>e.g. <em>ecorIR</em>, <em>ecorIM</em></td>
<td>e.g. <em>ecop11M</em>, <em>ecop11R</em></td>
<td>e.g. <em>mcrA</em>, <em>mcrBC</em>, <em>mrr</em></td>
</tr>
<tr>
<td>Subunits</td>
<td>~135, ~62 and ~52 kDa</td>
<td>~31 and ~38 kDa for EcoRI</td>
<td>~106 and ~75 kDa for EcoP11</td>
<td>Unrelated proteins</td>
</tr>
<tr>
<td>Proteins</td>
<td>REase: 2R + 2M + S MTase: 2M + S (± 2R)</td>
<td>Orthodox REase: 2R MTase: M</td>
<td>REase: 1 or 2 R + 2M MTase: 2M (± 2R)</td>
<td>Varies</td>
</tr>
<tr>
<td>REBASE</td>
<td>104 enzymes, 47 genes cloned, 34 genes sequenced, 5140 putatives</td>
<td>3938 enzymes, 633 genes cloned, 597 sequenced, 9632 putatives</td>
<td>21 enzymes, 19 genes cloned &amp; sequenced, 1889 putatives</td>
<td>18 enzymes &amp; genes cloned, 15 sequenced, 4822 putatives</td>
</tr>
</tbody>
</table>

Type I and II are currently divided in 5 and 11 different subclasses, respectively. Few enzymes have been well-characterized, but based on the current avalanche of sequence information many putative genes belonging to all Types and subtypes are being identified and listed on the restriction enzyme website (http://rebase.neb.com). The modification-dependent Type IV enzymes are highly diverse and only a few have been characterized in any detail. In each case, an example is given of one of the best-characterized enzymes within the different Types I, II and III. Note that Type II enzymes range from simple (shown here for EcoRI) to more complex systems (see Table 2 for the diversity of Type II subtypes). REBASE count is as of 16 September 2013 (http://rebase.neb.com/cgi-bin/statlist).
TP1: Restriction-modification Systems

Type I

• Restriction and modification functions in the same multisubunit enzyme (R2M2S)

• The enzymes require Mg, AdoMet, and ATP for activity,

• If the restriction sequence is fully methylated, ATP hydrolysis drives the dissociation of the enzyme from DNA.

• If the restriction site is hemimethylated (only one strand is methylated, e.g., DNA replication) the enzyme methylates the other strand and dissociates.

• If unmethylated, DNA cleavage occurs randomly at significant distances from an asymmetric recognition site (up to 1000bp)
TP1: Restriction-modification Systems

Type I

- In enteric bacteria Type I R-M systems have been demonstrated to be multifunctional protein complexes composed of three subunits encoded by the \textit{hsdS}, \textit{hsdM}, and \textit{hsdR} genes.

The \textit{hsdM} subunit adds chemical markers to the bacteria’s own DNA.

The \textit{hsdS} subunit specificity enzyme recognises these markers. However, if foreign DNA enters the cell, \textit{hsdS} subunit will recognise it as foreign, unmodified DNA.

The \textit{hsdR} subunit – endonuclease cleavage of target sequence.
TP1: Restriction-modification Systems

Ex. EcoK system

EcoKI recognition site
5'-AAC-(N)5-GTGC-3'
5'-AACNNNNNGTGC-3'

Enzymes bind to the target sequence and cleave or methylate according to the methylated state of DNA:
- Cleaves non-methylated DNA
- Methylates hemi-methylated DNA

Typically foreign DNA is unmethylated and will be degraded when introduced in a strain wild type for the three genes
Foreign DNA cloning
- Modification of some *E. coli* strains:
  
  *hsdR* gene is required only for the endonuclease cleavage (*hsdS* and *hsdM* are sufficient for target sequence methylation) - *E. coli* strains mutated for *hsdR* have the phenotype restriction minus, modification plus (r-, m+)

**Restriction modification systems** have been removed from *E. coli* K-12 strains used for cloning

- It can be used to clone foreign DNA that is not methylated at the target sequence
- When foreign DNA is propagated in this genetic background, the target sequence becomes methylated and may then be grown in a strain wildtype for the three genes
TP1: Restriction-modification Systems

Genotypes and phenotypes of *hsd* mutants in the *EcoK* system

*HsdS* - determines *HsdM* e *HsdR* specificity, an *hsdS* minus genotype lead to a *r*⁻*m*⁻ phenotype

*HsdR* - it is required only for nuclease cleavage of target sequence

*HsdM* - it methylates at A⁻⁶ACNNNNNNGTGC ou G⁻⁶ANNNNNNGTT. DNA isolated in *hsdM*⁻ is cleaved in *HsdR* host

The two subunits encoded by *hsdM* and *hsdS*, HsdM and HsdS are both necessary and sufficient for methyltransferase activity. HsdM is important for the positioning of the domains of HsdS

<table>
<thead>
<tr>
<th>GENOTYPE</th>
<th>PHENOTYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsdS</td>
<td>hsdR</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
**Type II**

- Type II R-M systems consist of two separate proteins with independent enzymatic activities, a restriction endonuclease and a DNA methyltransferase.

- In contrast to type I systems, DNA cleavage or methylation occurs within symmetrical recognition sites.

- Most type II enzymes recognize palindromic, duplex DNA sequences, such as GAATTC, whose complementary strand has the same 5′-3′ sequence.

- The restriction endonucleases are homodimeric, require Mg²⁺, and cleave phosphodiester bonds within or immediately adjacent to the recognition sequence to leave a staggered or blunt double-strand cut.

- The methyltransferases are monomeric and require the cofactor AdoMet. Methylation takes place on both strands of the DNA duplex within the recognition sequence.
Methylase of EcoRI (M.EcoRI) catalyzes the transfer of methyl groups from SAM to a specific A in the recognition sequence (*) of DNA.

The modification of adenine (*A) to 6-methyladenine, protects DNA from EcoRI cleavage.
**TP1: Restriction-modification Systems**

**Ex. BamH1**

**Recognition sites:**
- 4 to 6 nucleotides, palindrome (mostly)
- Cleavage inside the recognition site
TP1: Restriction-modification Systems

Type II restriction enzymes generated extremities

**Enzymes**

**HindIII**

Sticky ends/ extremidades coesivas
Protruding 5’P cohesive ends or 3’ recessed ends
5’ projetadas, ou 3’ recuadas

**Smal**

Sticky ends/ extremidades cegas
Blunt ends/ extremidades cegas

**KpnI**

Sticky ends/ extrem. coesivas
Protruding 3’OH cohesive ends or 5’ recessed ends
3’ projetadas, ou 5’ recuadas

http://nptel.ac.in/courses/102103013/module2/lec1/images/4.png
http://www.fastbleep.com/assets/notes/image/7041_1.jpg
## TP1: Restriction-modification Systems

### Restriction enzyme nomenclature

<table>
<thead>
<tr>
<th></th>
<th>EcoRI</th>
<th>HindIII</th>
<th>BamHI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Genus</strong></td>
<td>Escherichia</td>
<td>Haemophilus</td>
<td>Bacillus</td>
</tr>
<tr>
<td><strong>Specie</strong></td>
<td>coli</td>
<td>influenzae</td>
<td>amylo</td>
</tr>
<tr>
<td><strong>Strain</strong></td>
<td>R</td>
<td>D</td>
<td>H</td>
</tr>
<tr>
<td><strong>Isolate</strong></td>
<td>I</td>
<td>III</td>
<td>I</td>
</tr>
<tr>
<td><strong>Recognition site</strong></td>
<td>G^AATTC</td>
<td>A^AGCTT</td>
<td>G^GATGC</td>
</tr>
</tbody>
</table>
EcoRI enzyme name is derived from *Escherichia coli*, strain RY13, 1rst identified

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Microorganism From Which Enzyme Is Isolated</th>
<th>Recognition Sequence</th>
<th>Type of Fragment End Produced</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>BamHI</em></td>
<td><em>Bacillus amyloliquefaciens</em></td>
<td>5′–GGATCC–3′&lt;sup&gt;↓&lt;/sup&gt; 3′–CCTAGG–3′&lt;sup&gt;↑&lt;/sup&gt;</td>
<td>Cohesive</td>
</tr>
<tr>
<td><em>CofI</em></td>
<td><em>Clostridium formicoaceticaenum</em></td>
<td>5′–GGCG–3′&lt;sup&gt;↓&lt;/sup&gt; 3′–GGCG–5′&lt;sup&gt;↑&lt;/sup&gt;</td>
<td>Cohesive</td>
</tr>
<tr>
<td><em>DraI</em></td>
<td><em>Deinococcus radiophilus</em></td>
<td>5′–TTTAAA–3′&lt;sup&gt;↓&lt;/sup&gt; 3′–AAATTT–5′&lt;sup&gt;↑&lt;/sup&gt;</td>
<td>Blunt</td>
</tr>
<tr>
<td><em>EcoRI</em></td>
<td><em>Escherichia coli</em></td>
<td>5′–GAATTC–3′&lt;sup&gt;↓&lt;/sup&gt; 3′–CTTAAG–5′&lt;sup&gt;↑&lt;/sup&gt;</td>
<td>Cohesive</td>
</tr>
<tr>
<td><em>EcoRII</em></td>
<td><em>Escherichia coli</em></td>
<td>5′–CCAGG–3′&lt;sup&gt;↓&lt;/sup&gt; 3′–GGTCC–5′&lt;sup&gt;↑&lt;/sup&gt;</td>
<td>Cohesive</td>
</tr>
<tr>
<td><em>HaeIII</em></td>
<td><em>Haemophilus aegyptius</em></td>
<td>5′–GGCC–3′&lt;sup&gt;↓&lt;/sup&gt; 3′–CCGG–5′&lt;sup&gt;↑&lt;/sup&gt;</td>
<td>Blunt</td>
</tr>
</tbody>
</table>
### Table 18.2 Characteristics of some common type II restriction enzymes used in recombinant DNA technology

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Microorganism From Which Enzyme Is Isolated</th>
<th>Recognition Sequence</th>
<th>Type of Fragment End Produced</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>HindIII</em></td>
<td><em>Haemophilus influenzae</em></td>
<td>5'--AAGCTT--3'</td>
<td>Cohesive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3'--TTCGAA--5'</td>
<td></td>
</tr>
<tr>
<td><em>HpaII</em></td>
<td><em>Haemophilus parainfluenzae</em></td>
<td>5'--CCGG--3'</td>
<td>Cohesive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3'--GGCC--5'</td>
<td></td>
</tr>
<tr>
<td><em>NotI</em></td>
<td><em>Nocardia oitidis-caviarum</em></td>
<td>5'--GCGGGCGCGG--3'</td>
<td>Cohesive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3'--CGGGCCGGCGG--5'</td>
<td></td>
</tr>
<tr>
<td><em>PstI</em></td>
<td><em>Providencia stuartii</em></td>
<td>5'--CTGACG--3'</td>
<td>Cohesive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3'--GACGTC--5'</td>
<td></td>
</tr>
<tr>
<td><em>PvuII</em></td>
<td><em>Proteus vulgaris</em></td>
<td>5'--CAGCTG--3'</td>
<td>Blunt</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3'--GTGCAG--5'</td>
<td></td>
</tr>
<tr>
<td><em>SmaI</em></td>
<td><em>Serratia marcescens</em></td>
<td>5'--CCCGGG--3'</td>
<td>Blunt</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3'--GGGGCC--5'</td>
<td></td>
</tr>
</tbody>
</table>

Note: The first three letters of the abbreviation for each restriction enzyme refer to the bacterial species from which the enzyme was isolated (e.g., Eco refers to *E. coli*). A fourth letter may refer to the strain of bacteria from which the enzyme was isolated (the "R" in *EcoRI* indicates that this enzyme was isolated from the RY13 strain of *E. coli*). Roman numerals that follow the letters allow different enzymes from the same species to be identified. For convenience, molecular geneticists have come up with idiosyncratic pronunciations of the names: *EcoRI* is pronounced “echo-R-one,” *HindIII* is “hin-D-three,” and *HaeIII* is “hay-three.” These common pronunciations obey no formal rules and simply have to be learned.
Type III

• Only a few type III R-M systems are known
• A single bifunctional enzyme catalyzes both the endonuclease and the methyltransferase restriction activities. The enzymes are composed of two nonidentical subunits: the M subunit (encoded by the mod gene) and the R subunit (encoded by the res gene).

• The R subunit must be complexed with the M subunit for restriction activity, because the M subunit provides the sequence specificity for the enzyme. The two enzymatic activities compete for the uninterrupted, asymmetric DNA recognition sequence, which is usually 5-6 bp long.

• Cleavage takes place 25-30 bp away to the 3′ side of the DNA recognition sequence. If one or both of the DNA strands is (are) methylated, no cleavage occurs. The M subunit can act independently as a methyltransferase, requiring AdoMet and methylating only one strand of the duplex recognition sequence at a time, which is sufficient to inhibit the restriction reaction.
## TP1: Restriction-modification Systems

### Other Restriction-Modification systems

#### Modification systems *(site-specific methylases)*

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Recognition Site</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dam</td>
<td>(N6)G*ATC</td>
<td>Dam methylase, Deoxyadenosine methylase</td>
</tr>
<tr>
<td>Dcm</td>
<td>(C5)C<em>CAGG e C</em>CTGG</td>
<td>DNA 5-cytosine methyltransferase</td>
</tr>
</tbody>
</table>

Only methylate DNA sequences

#### Restriction systems *(endonucleases)*

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Recognition Site</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>McrA</td>
<td>C*CGG</td>
<td>Mcr=modified cytosine restriction</td>
</tr>
<tr>
<td>McrB</td>
<td>G*C</td>
<td></td>
</tr>
<tr>
<td>Mrr</td>
<td>C<em>AC and C</em>AG</td>
<td></td>
</tr>
</tbody>
</table>

Only cleave METHYLATED sequences
TP1: Restriction-modification Systems

Restriction system – Dam and Dcm methylation

- The **Dam** (DNA adenine methyltransferase) and **Dcm** (DNA cytosine methyltransferase) modification systems methylate adenines and cytosines located within specific recognition sequences (5’-GATC-3’ for Dam and the second cytosine of 5’-CCA/TGG-3’ for Dcm)

- Some or all of the sites for a restriction endonuclease may be resistant to cleavage when isolated from strains expressing the Dam or Dcm methylases if the methylase recognition site overlaps the endonuclease recognition site. For example, plasmid DNA isolated from *dam+ E. coli* is completely resistant to cleavage by *MboI*, which cleaves at GATC sites.

- Prevent certain restriction enzymes from cutting their target sequence in plasmid DNA. *Eg. XbaI will not cut its recognition sequence (5’-TCTAGA-3’) when the last adenine is methylated.*
**TP1: Restriction-modification Systems**

**Modification system – Dam methylation**

**Dam methylation:** $G^{m6}ATC$  (methyl group at position N6 of adenine: $N6$-methyladenine)

1. example of Type II restriction enzyme **blocked** by Dam methylation

   - *Clal*  $G^{*}ATCGAT$
   - *XbaI*  $TCTAG^{*}ATC$
   - *MboI*  $G^{*}ATC$

2. example of Type II restriction enzyme **not blocked** by Dam methylation

   - *BamHI*  $GG^{*}ATCC$
   - *Pvul*  $CG^{*}ATCG$
   - *Sau3AI*  $G^{*}ATC$

Restriction enzymes have different sensibilities to the various methylation patterns

Ex: *Sau3AI*

**Bold**- type II restriction enzyme recognition sequence
Modification system – mcrA/mcrB/mrr complex

• Direct endonuclease cleavage to DNA targets that are methylated on certain sequences

• There are two mcr (methyl-cytosine restricting) systems in E. coli—mcrA and mcrBC. Another system, mrr (modified adenine recognition and restriction), restricts particular DNA sequences that include methyladenine (m6A) or methylcytosine (m5C) residues.

• The mcrA/mcrB/mrr complex degrades foreign DNA that is not properly methylated, such as methylated DNA obtained from mouse and human cells which contains CpG methylated DNA
Strains mutated in both McrA and McrB systems, as well as in the Mrr system should be considered when cloning DNA from other organisms as plants and mammals usually contain some degree of methylation that can be recognized by these systems.
Ex. 1 – Consequences of restriction on DNA amplified in a Dam$^+$ strain: G*ATC

- **MboI**
  - GATC
  - Doesn’t cleave methylated sequences
  - CTAG
  - G* ATC
  - CTA* G

- **Sau3AI**
  - GATC
  - Cleaves methylated sequences
  - CTAG
  - G* ATC
  - CTA* G

Ex. 2 – Cloning of human DNA, that has *CpG islands*, should be amplified in mcrA$^-$ and mcrBC$^-$

- **mcrA$^-$**
  - McrA- C$^{m5}$CGG

- **mcrBC$^-$**
  - McrBC- G$^{m5}$C
  - G$^{h5}$C
  - G$^{N4}$C
The extensive research on methylation was conducted on bacteria where both A and C can be methylated.

Eukaryotic DNA methylation affects mainly C residues and is specific for CpG sequence.

The protective function of DNA methylation is similar in eukaryotes and prokaryotes. In humans and rodents inserted viral sequences can become methylated in association with silencing of the introduced genes. Thus function of DNA methylation machinery for recognition and/or eliminating of foreign DNA seem to be conserved in evolution.

eukaryotic genome - cytosine methylation plays a role in functional reorganization:

Low 5'-mC----------> high levels of gene expression  
High 5'-mC ----------> low levels of gene expression

Since DNA methylation is reversible and does not directly depend on the sequence context it was described as an epigenetic mechanism of gene regulation.

DNA methylation is observed in most of the organisms at the different stages of evolution. Some species, like Drosophila melanogaster lack DNA methylation.
**Isoesquizomers**

1- Recognize the same sequence and cut on the same sites

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Sequence</th>
<th>Cut Site</th>
<th>Overhang</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>AccIII</td>
<td>TCCGGA</td>
<td>[A G G C C / T]</td>
<td>5' - CCGG</td>
<td><img src="image1.png" alt="Image" /></td>
</tr>
<tr>
<td>BspEI</td>
<td>TCCGGA</td>
<td>[A G G C C / T]</td>
<td>5' - CCGG</td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
</tbody>
</table>

2- Recognize the same sequence but cut in different positions

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Sequence</th>
<th>Cut Site</th>
<th>Overhang</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>XmaI</td>
<td>CCCGGG</td>
<td>[G G G / C C C]</td>
<td>5' - CCGG</td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>Smal</td>
<td>CCCGGG</td>
<td>[C C C / G G G]</td>
<td>blunt</td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
</tbody>
</table>
### Isoesquizomers

3- Recognize the same sequence, different sensitivity to methylation

*Dpn*I only cleaves fully-adenomethylated dam sites G*ATC*

*Dpn*II and *Mbo*I share methylation sensitivity - cleave dam sites which lack adenomethylation and each is blocked by complete dam methylation

*Sau3AI* will cleave all dam sites regardless of adenomethylation

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Sequence</th>
<th>Cut Site</th>
<th>Overhang</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>DpnI</td>
<td>GATC</td>
<td>/G A T C C T A G/</td>
<td>blunt</td>
<td><img src="image" alt="Properties" /></td>
</tr>
<tr>
<td>BfuCl</td>
<td>GATC</td>
<td>/G A T C C T A G/</td>
<td>5’-GATC</td>
<td><img src="image" alt="Properties" /></td>
</tr>
<tr>
<td>DpnII</td>
<td>GATC</td>
<td>/G A T C C T A G/</td>
<td>5’-GATC</td>
<td><img src="image" alt="Properties" /></td>
</tr>
<tr>
<td>MboI</td>
<td>GATC</td>
<td>/G A T C C T A G/</td>
<td>5’-GATC</td>
<td><img src="image" alt="Properties" /></td>
</tr>
<tr>
<td>Sau3AI</td>
<td>GATC</td>
<td>/G A T C C T A G/</td>
<td>5’-GATC</td>
<td><img src="image" alt="Properties" /></td>
</tr>
</tbody>
</table>

*Cleavage with this restriction enzyme may be blocked or impaired when the substrate DNA is methylated by either the dam or dcm or CpG methylase.*
Applications
DNA cloning: bacterial strains

• *E. coli* has several mechanisms for identifying foreign DNA and destroying it - *problem in cloning experiments*, resulting in substantially reduced recovery of desired sequences.

• EcoKI restriction, encoded by the *hsdRMS* genes, attacks DNA that is not protected by adenine methylation at the appropriate recognition site (AAC[N6]GTGC or GCAC[N6]GTT).

• McrA, McrBC, and Mrr, encoded by *mcrA*, *mcrBC*, and *mrr*, are methylation-requiring systems that attack DNA only when it is methylated at specific positions.

The problem can be avoided by the use of strains in which these mechanisms are disabled by mutation.

A strain completely disabled for restriction will be defective at the *hsd*, *mcrA*, *mcrBC*, and *mrr* loci.
Recombinant DNA technology was “born” with the discovery of restriction enzymes.

One of the most important applications of restriction enzyme type II is **DNA cloning**, i.e., the possibility of propagating a particular gene of interest or other DNA fragment using specific DNA molecules (vectors).

The approach used to clone a specific gene, depends to a large degree on:
- the gene
- what is known about it
- objective

Among the several tools needed for cloning, let’s consider:
- restriction enzymes
- vector DNA
- DNA ligase
- host cell
TP1: Restriction-modification Systems

Cleaving and viewing DNA after digestion with type II RE

(a) Linear DNA

Digestion at four sites by HindIII

(b) Circular DNA

Digestion at four sites by HindIII

<table>
<thead>
<tr>
<th>EcoRI alone</th>
<th>BamHI alone</th>
<th>EcoRI + BamHI (double digest)</th>
<th>Size standards</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 kb</td>
<td>9 kb</td>
<td>7 kb</td>
<td>10 kb</td>
</tr>
<tr>
<td>3 kb</td>
<td>4 kb</td>
<td>3 kb</td>
<td>6 kb</td>
</tr>
<tr>
<td>2 kb</td>
<td>2 kb</td>
<td>3 kb</td>
<td>2 kb</td>
</tr>
<tr>
<td>1 kb</td>
<td></td>
<td>1 kb</td>
<td></td>
</tr>
</tbody>
</table>
Plasmid vector important markers

- Plasmids are small circular DNA molecules that are found inside some prokaryotic cells.
- Plasmids as cloning vectors are in vitro constructed and are important tools in genetic engineering.

Antibiotic resistance marker
Bacterial host cell should be sensitive to the antibiotic.
TP1: Restriction-modification Systems

Construction of a recombinant DNA molecule
Isolating and cloning a DNA fragment

**TP1: Restriction-modification Systems**

**Isolating and cloning a DNA fragment**

- **Nucleus**: Contains chromosomes.
- **Chromosomes**: Part of the nucleus.
- **HUMAN CELL**: Contains chromosomes.
- **Double-stranded DNA molecule**: Contains DNA marker.
- **Cleavage**: Represents the cleavage of DNA.
- **DNA fragment**: Represents the fragment of DNA.
- **BACTERIAL CELL**: Contains a bacterial cell with a plasmid.
- **CLONED DNA**: Represents the cloned DNA.
- **Replication in bacterial cells**: Represents the replication process.
- **DNA fragment from any organism**: Represents the DNA fragment from any organism.
- **Cleaved plasmid DNA vector**: Represents the cleaved plasmid DNA vector.
- **Transformation of a bacterium**: Represents the transformation of a bacterium.
- **Growth and cell division**: Represents the growth and cell division process.
- **Clones of plasmid-containing bacterium**: Represents the clones of the plasmid-containing bacterium.
Useful sites/ tools regarding R-M systems

http://rebase.neb.com/rebase/rebase.html

https://www.neb.com/tools-and-resources/interactive-tools/enzyme-finder

http://nc2.neb.com/NEBcutter2/


Type I RE: https://www.youtube.com/watch?v=uyfFLWzUjmM

Type II RE: https://www.youtube.com/watch?v=6U8bGOG90AI