Fundamentos de Biologia Molecular

Curso de Licenciatura em Biologia 2º Ano, 1º Semestre Ano Letivo 2021/2022

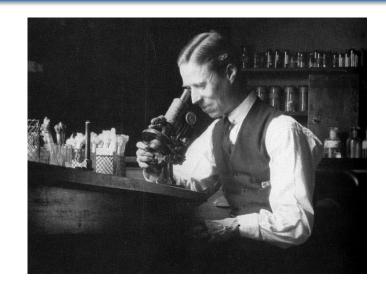
Componente Teórico-Prática



Docente Responsável: Rita Zilhão Docente TPs: Andreia Figueiredo

- 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.

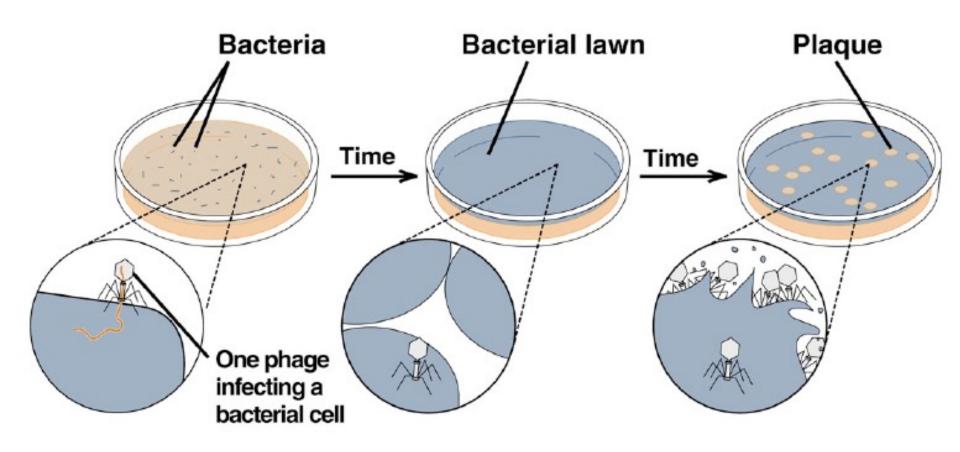


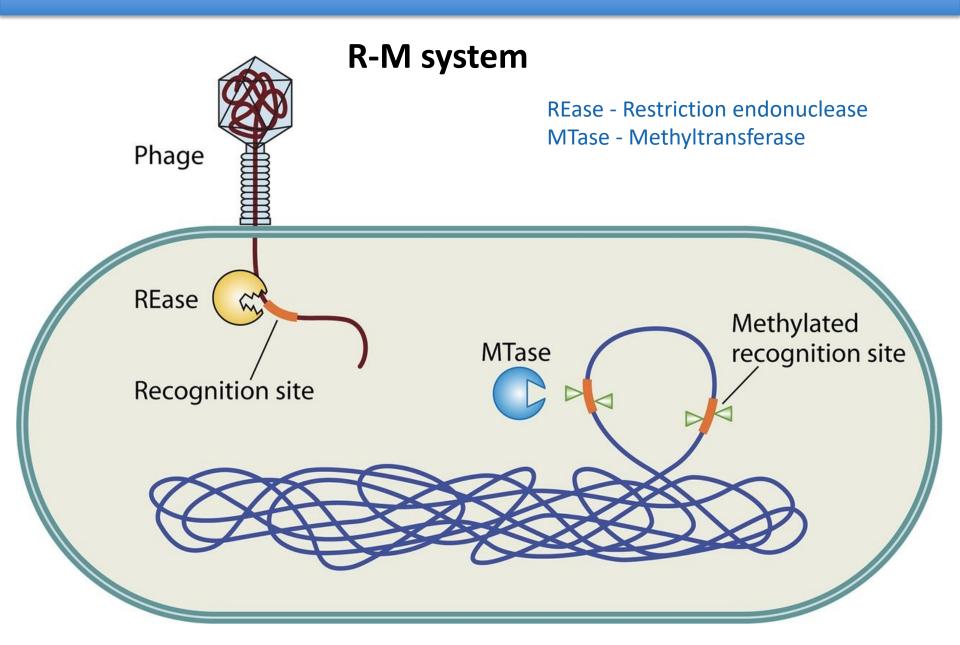
Bacterial colonies

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



Lysis plaques







■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



Hamilton Smith

■In 1972, he purified the first site-specific "**Type II**" restriction enzyme, known as *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 *Hind II* recognized and specifically cleavage.



Daniel Nathans

- ■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.

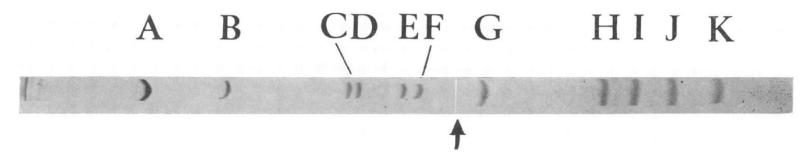


Fig. 3. Radioautographic analysis of SV40 DNA digested with H. influenzae restriction endonuclease. 1 μg of SV40 [14C]-DNA I (3 \times 104 cpm/ μg) was digested (see Fig. 2) for 6 hr in a volume of 55 μ l; 0.0015 unit of enzyme was added at 0 time and at 1, 2, 3, 4, and 5 hr. 20 μ l of sample was electrophoresed for 12.3 hr and the radioautogram was prepared as described in Methods. The origin is at the left. The arrow below the radioautogram indicates a transverse cut made in the gel prior to slicing.

Danna, K., Nathans, D. (1971) Specific cleavage of simian virus 40 DNA by restriction endonuclease of *Hemophilus influenza*. Proc. Natl. Acad. Sci. USA 68: 2913-2917.

The Nobel Prize in Physiology or Medicine 1978



Werner Arber Prize share: 1/3



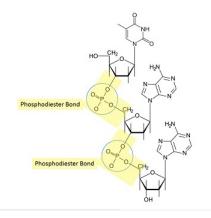
Daniel Nathans Prize share: 1/3



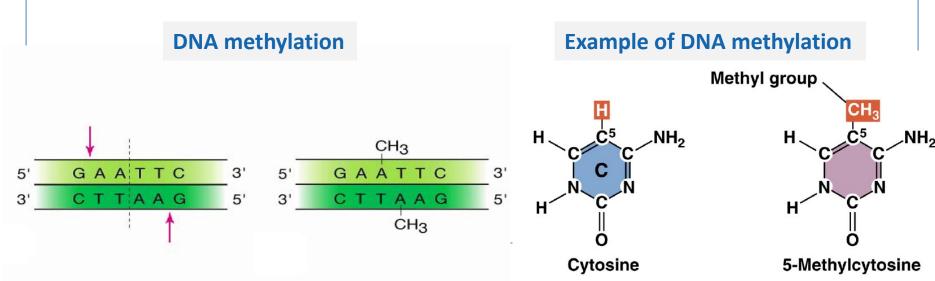
Hamilton O. Smith Prize share: 1/3

"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 (eg 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 *Nci*I: originates 3'P and 5'OH extremities)



- •Restriction enzymes usually cut 4-6 bp long sequences that are frequently palindromes¹.
- •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 (CH₃) groups.

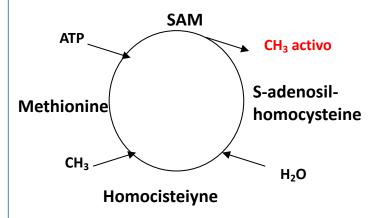


A methyl group is added to the 5-carbon position of cytosine residues in DNA

¹A palindrome is a word, phrase, number, or other sequence of characters which reads the same backward or forward

•DNA methyltransferases (MTases) transfer a methyl group from S-adenosylmethionine to either adenine our cytosine residues

Main donor of methyl groups SAM (S-adenosyl methionine)



S-Adenosyl-L-methionine alternative names: SAM, AdoMet, ademetionine.

Most common DNA methylation

(N4mC)

(N6mA)

•Approximately ¼ of known bacteria present R-M systems and ½ of those present more than one R-M system.

(5mC)

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.

| Туре | Protein structure | Restriction & methylation | RE recognition site | Ex of recognition sites | RE cleavage site | ATP required |
|------|--|---------------------------|------------------------------------|---|--|-----------------|
| I | Bifunctional enzyme of 3 subunits | Mutually exclusive | Bipartite and asymetric | EcoAl GAGNNNNNNNGTCA EcoKl AACNNNNNNGTGC | Non-specific, >1000 bp from recognition site | Yes |
| II | Separate endonuclease and methylase | Separate reactions | 4-8 bp sequence, often palindromic | EcoRI G/AATTC Ball TGG/CCA | Same as or close to recognition site | No |
| III | Bifunctional enzyme of 2 subunits | Simultaneous | 5-7 bp, asymetric sequence | EcoPI AGACC HinfIII CGAAT | 24-25 bp, dowstream of recognition site | Yes |

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, eg 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)

Type I

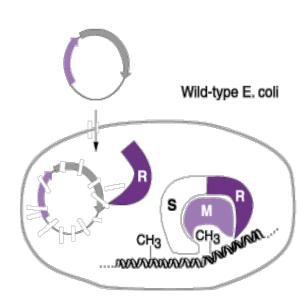
•In enteric bacteria Type I R-M systems have been demonstrated to be multifunctional protein complexes composed of three subunits encoded by the *hsdS*, *hsdM*, and *hsdR* genes



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

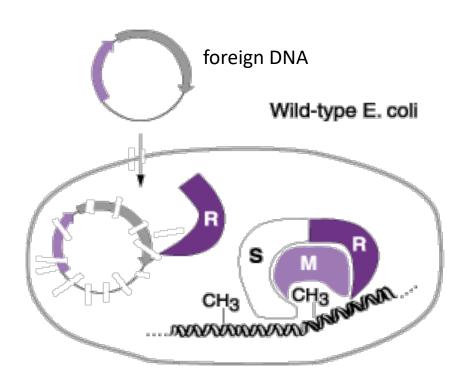
The **hsdR** subunit – endonuclease cleavage of target sequence

The **hsdM** subunit adds chemical markers to the bacteria's own DNA.



Ex. EcoK system – problems when using E.coli K-12 bacteria in the lab

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 (attacks DNA that is *not* protected by adenine methylation)
- Methylates hemi-methylated DNA

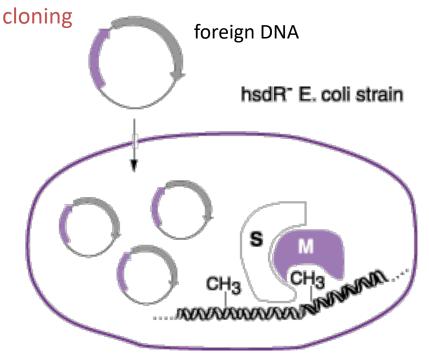
Typically foreign DNA is unmethylated and will be degraded when introduced in a strain wild type for the three genes

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



- •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

Genotypes and phenotypes of hsd mutants in the EcoK system

HsdS - determines **HsdM e HsdR specificity**, an <u>hsdS minus genotype</u> lead to a r - m - phenotype

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

HsdM - it methylates at A^{N6}ACNNNNNNGTGC ou GC^{N6}ANNNNNNGTT. 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

| hsdS | hsdR | PHENOTYPE | |
|------|------|-----------|-------------------------------|
| + | + | + | r ⁺ m ⁺ |
| _ | + | + | r m |
| + | - | + | r⁻m⁺ |
| + | + | - | r ⁻ m ⁻ |

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 Mg, 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.

https://www.youtube.com/watch?v=6U8bGOG9OAI&t=151s

Ex. EcoRI system

EcoRI, as most of all the other type II restriction enzymes, cleaves inside the recognition sequence :

```
...aactGAATTCtcgac...
...ttgaCTTAAGagctg...

Gagctg...

Gagctg...
```

Methylase of *Eco*RI (M.*Eco*RI) catalyzes the transfer of methyl groups from SAM to a specific A in the recognition sequence (*) of DNA

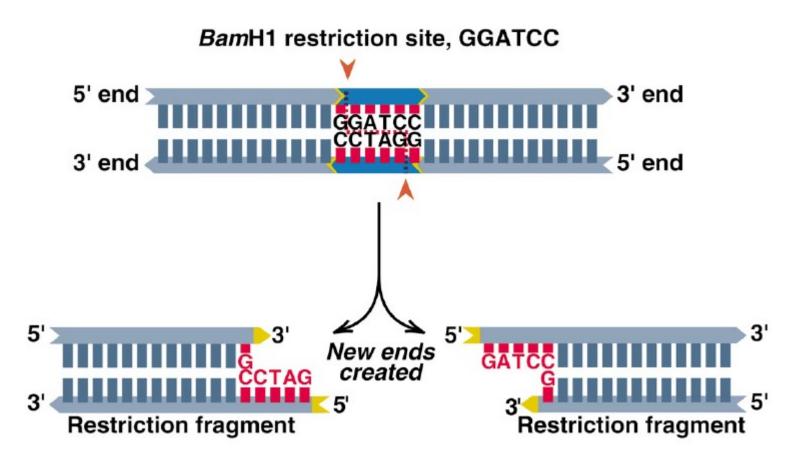
```
...aact G A *A T T C tcgac...
...ttga C T T *A A G agctg...
```

The modification of adenine (*A) to 6-methyladenine, protects DNA from *Eco*RI cleavage

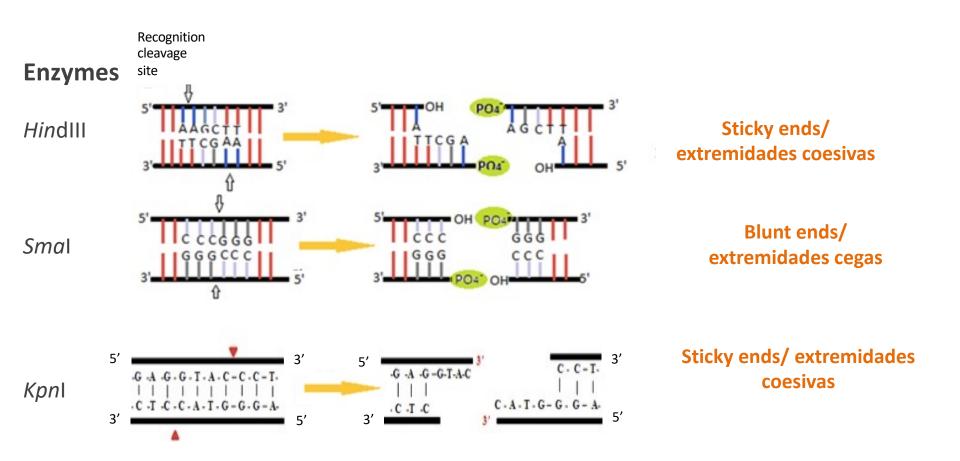
Ex. BamH1

Recognition sites:

- 4 to 6 nucleotides, palindrome (mostly)
- Cleavage inside the recognition site



Type II restriction enzymes generated extremities



Restriction enzyme nomenclature

| | EcoRI | HindIII | BamHI | |
|------------------|-------------|-------------|----------|--|
| Genus | Escherichia | Haemophilus | Bacillus | |
| Specie | coli | influenzae | amylo | |
| Strain | R | D | Н | |
| Isolate | l | III | I | |
| Recognition site | G^AATTC | A^AGCTT | G^GATGC | |

| Table 18.2 Characteristics of some common type II restriction enzymes used in recombinant DNA technology | | | | | |
|--|--|--|----------------------------------|--|--|
| Enzyme | Microorganism From Which Enzyme Is Isolated | Recognition Sequence | Type of Fragment End Produced | | |
| <i>Bam</i> HI | Bacillus amyloliquefaciens | 5'-GGATCC-3' 3'-CCTAGG-3' | Cohesive | | |
| Cofl | Clostridium formicoaceticum | 5'-GCGC-3' 3'-CGCG-5' | Cohesive | | |
| Dral | Deinococcus radiophilus | ↑ ↓ 5′-TTTAAA-3′ 3′-AAATTT-5′ | Blunt | | |
| <i>Eco</i> RI | Escherichia coli | 5'-GAATTC-3' 3'-CTTAAG-5' | Cohesive | | |
| <i>Eco</i> RII | Escherichia coli | 5′-CCAGG-3′ 3′-GGTCC-5′ | Cohesive | | |
| Haelll | Haemophilus aegyptius | 5'-GGCC-3' 3'-CCGG-5' | Blunt | | |

EcoRI enzyme name is derived from *Escherichia coli,* strain RY13, 1rst identified

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.

Other restriction-modification systems

Modification systems (site-specific methylases)

Dam (N6)G*ATC

Dcm (C5)C*CAGG e C*CTGG DNA 5-cytosine methyltransferase

Only methylates DNA sequences

Restriction systems (endonucleases)

McrA C*CGG

McrB G*C

Mrr C*AC and C*AG

Mcr=modified cytosine restriction

Dam methylase, Deoxyadenosine methylase

Only cleave METHYLATED sequences

- Prokaryotic Methylation methylation should be considered when digesting DNA with restriction endonucleases (choice of bacterial strains in the lab)
- Cleavage can be blocked or impaired when a particular base in the recognition site is methylated

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 Mbol, which cleaves at GATC sites.
- •Prevent certain restriction enzymes from cutting their target sequence in plasmid DNA. Eg. Xbal will not cut its recognition sequence (5'-TCTAGA-3') when the last adenine is methylated.

Modification system – Dam methylation

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

1- example of Type II restriction enzymes **blocked** by Dam methylation

Clal G*ATCGAT

Xbal TCTAG*ATC

Mbol G*ATC

2- example of Type II restriction enzymes **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

CLEAVES

Gm6ATC

GATm4C

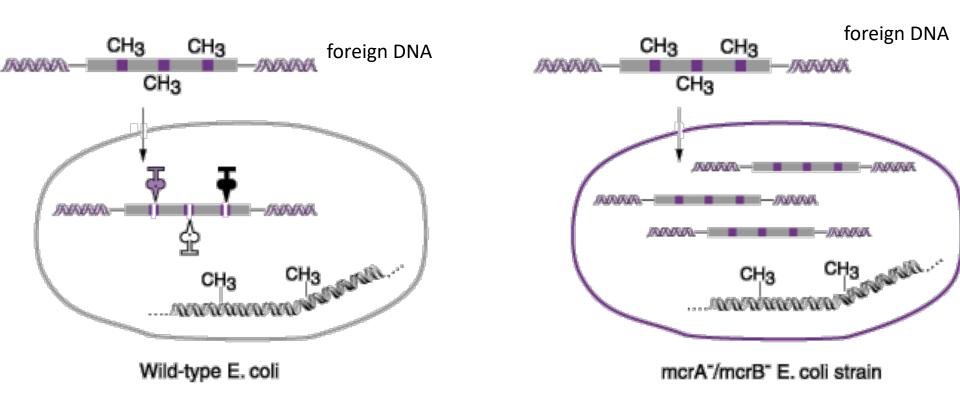
GATm5C

Bold- type II restriciton enzyme recognition sequence

Modification system – mcrA/mcrB/mrr complex

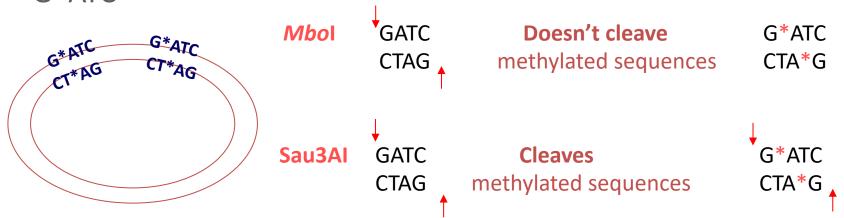
- •Direct endonuclease cleavage to DNA targets that are methylated on certain sequences
- •There are two mcr(methyl-cytosine restricting) systems in *E. co*li—mcrAand mcrBC. Another system, mrr (modified adenine recognition and restriction), restricts particular DNA sequences that include methyladenine (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

Only cleave METHYLATED sequences



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



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 ^{N4} C |

DNA methylation in prokaryotes and eukaryotes

- •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

| Enzyme | Sequence | Cut Site | Overhang | Properties |
|--------|----------|----------------------------|-----------|---------------------------------|
| AccIII | TCCGGA | T/C C G G A A G G C C/T | 5' - CCGG | |
| BspEl | TCCGGA | T/C C G G A A G G C C/T | 5' - CCGG | ⇔ dam CpG RN ₩ 183.1 37° |

2- Recognize the same sequence but cut in different positions

| Enzyme | Sequence | Cut Site | Overhang | Properties |
|--------|----------|--------------------------------|-----------|---------------------|
| Xmal | CCCGGG | C/CCGGG GGCC/C | 5' - CCGG | CpG RX CutSmart 37° |
| Smal | cccgg | C C C / G G G G G G / C C C | blunt | CpG RX CutSmart 25° |

Isoesquizomers

3- Recognize the same sequence, different sensitivity to methylation

DpnI only cleaves fully-adenomethylated dam sites G*ATC

DpnII and **MboI** 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

| Enzyme | Sequence | Cut Site | Overhang | Properties |
|--------|----------|-----------------------|-----------|-------------------------|
| Dpnl | GATC | G A/T C C T/A G | blunt | CpG RX Ves CutSmart 37° |
| BfuCl | GATC | / G A T C C T A G/ | 5' - GATC | CpG CutSmart 37° |
| Dpnll | GATC | / G A T C C T A G/ | 5' - GATC | dam Ril Vis NEBU 37° |
| Mbol | GATC | / G A T C C T A G/ | 5' - GATC | dam CpG RX CutSmart 37° |
| Sau3AI | GATC | / G A T C C T A G/ | 5' - GATC | CpG Ril 137° |

dam CpG dcm

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



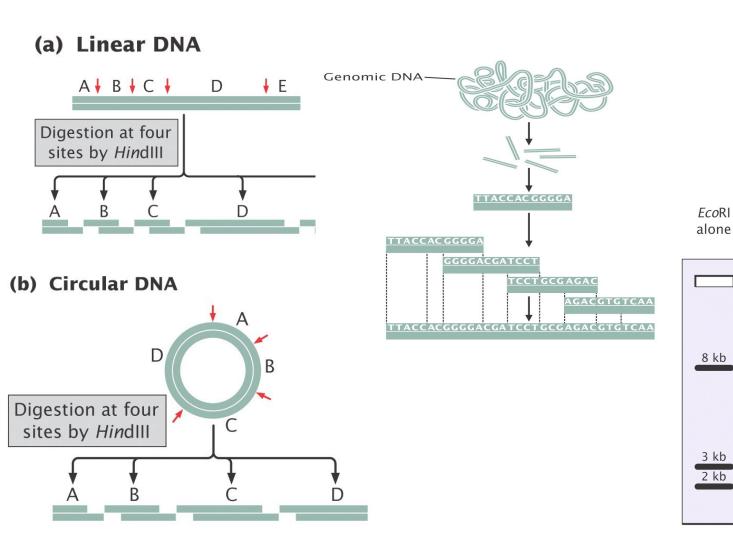
The problem can be avoided by the use of strains in which these mechanisms are disabled by mutation.

A strain completely disabled for restriction will be defective at the *hsd*, *mcrA*, *mcrBC*, and *mrr* loci

Application of restriction enzymes type II: DNA cloning

- 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, lets consider:
 - restriction enzymes
 - vector DNA
 - DNA ligase
 - host cell

Cleaving and viewing DNA after digestion with type II RE



BamH1

alone

9 kb

4 kb

EcoRI +

BamHI

(double digest)

7 kb

3 kb

1 kb

Size standards

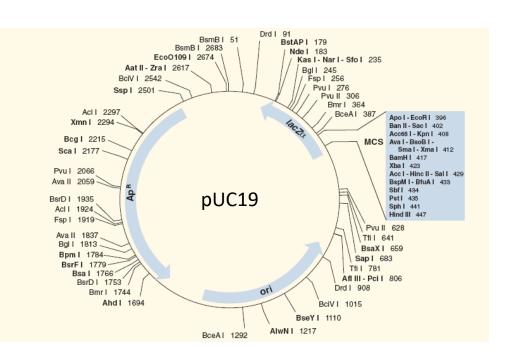
10 kb

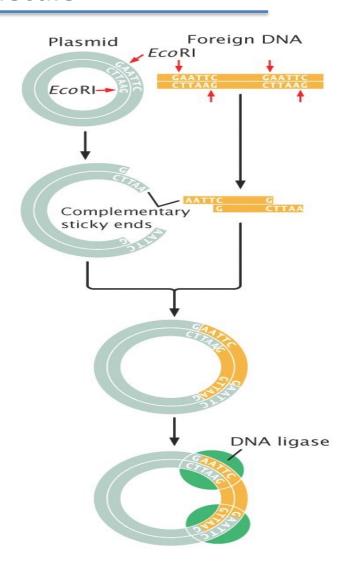
6 kb

3 kb

1 kb

Construction of a recombinant DNA molecule





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/

https://www.ncbi.nlm.nih.gov/genbank/

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

Type II RE: https://www.youtube.com/watch?v=6U8bGOG9OAI

NEB: https://international.neb.com/