

# Introduction to Molecular Genetics



EUROPEAN UNION  
European Structural and Investing Funds  
Operational Programme Research,  
Development and Education



MINISTRY OF EDUCATION,  
YOUTH AND SPORTS

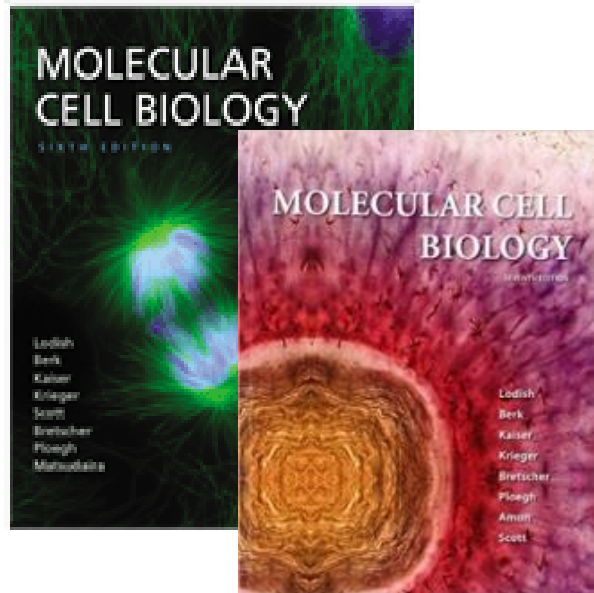


Dílo podléhá licenci Creative Commons 4.0 Česko  
Uveďte původ - Zachovejte licenci

# Molecular genetics and genetic engineering







**Lodish et al.  
Freeman and company**



**Alberts et al.  
Garland Science**

<http://biomikro.vscht.cz/>

password: giruml



Ústav biochemie a mikrobiologie  
Vysoká škola chemicko-technologická v Praze



[Česky / Czech](#)

**HOMEPAGE**

**EDUCATION**

**RESEARCH**

**CONTACTS**

» Home » Education

**Introduction**

**Courses**

## Education at the Department of Biochemistry and Microbiology

Department of Biochemistry and Microbiology offers bachelor courses in core disciplines (biochemistry, microbiology, biology and others). We also participate on courses for the novel bachelor program focused on forensic sciences. Biochemistry courses are offered to students of all faculties of our institute. Our department organizes master programmes General and Applied Biochemistry, Microbiology and Clinical Biochemistry.

[Environmental Microbiology, Food microbiology and Genetic engineering courses](#) are delivered in English for foreign students. For the list of courses in Czech language visit [the Czech version of the website](#).

The department is accredited for doctoral education of biochemists and microbiologists. If you are interested in PhD studies at our department, please contact [group leaders](#) or [the department secretariat](#).





**HOME PAGE**

**EDUCATION**

**RESEARCH**

**CONTACTS**

» [Home](#) » [Education](#)

**Introduction**

**Courses**

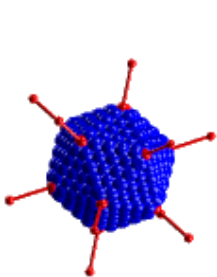
## Courses

---

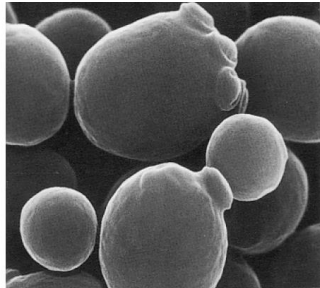
### Courses in English

- [Environmental Microbiology](#)
- [Food microbiology](#)
- [Genetic engineering](#) ([slides](#))
- Psychrotrophic Microorganisms and Cold-Active Enzymes ([slides](#))

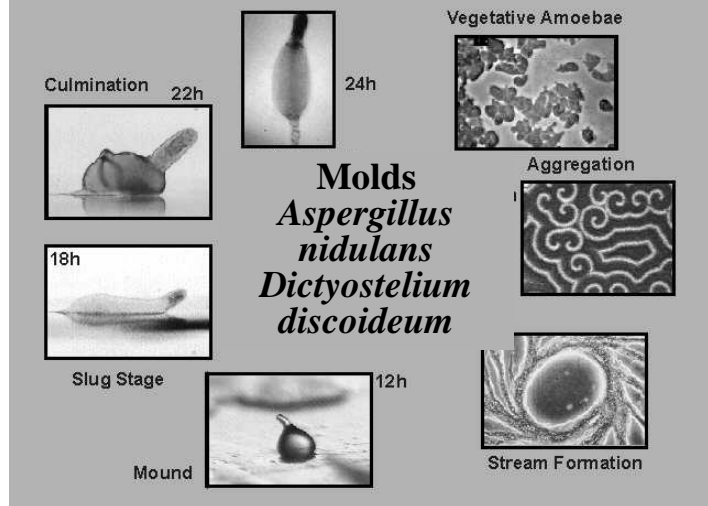
[Other courses \(in Czech\)](#)



*Escherichia coli*



*Saccharomyces cerevisiae*



*Caenorhabditis elegans*



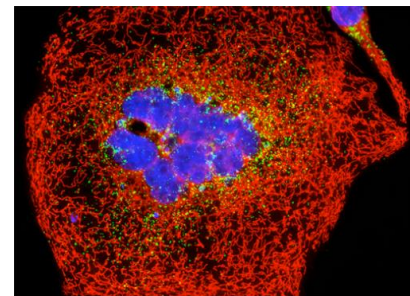
*Drosophila melanogaster*



*Danio rerio*



*Arabidopsis thaliana*



# Model organisms



# Gene

**DNA fragment    ~    protein  
                              RNA  
                              regulatory function**

**Ability to replicate**

**Ability to mutate**

**Genome – whole genetic information of an organism**

**Genome** - the total genetic information of an organism in haploid state

**German botanist Hans Winkler 1920 - GenOme combination of gene and chromosome**

**Genotype** - individual genetic equipment

**(sometimes used to specific studied genes)**

**Phenotype** - the manifestation of the genotype

**Modifications** - reversible changes in phenotype

Diploid cell - homologous (alelomorphic chromosomes)  
Allele - form of the same gene occurring at the same loci in homologous chromosomes but differ in the sequence of bases

Standard x mutated

Allele – in heterozygous or homozygous state

**? Transfer of genetic information ?**

# **Gregor Mendel**

**(1822 – 1884)**

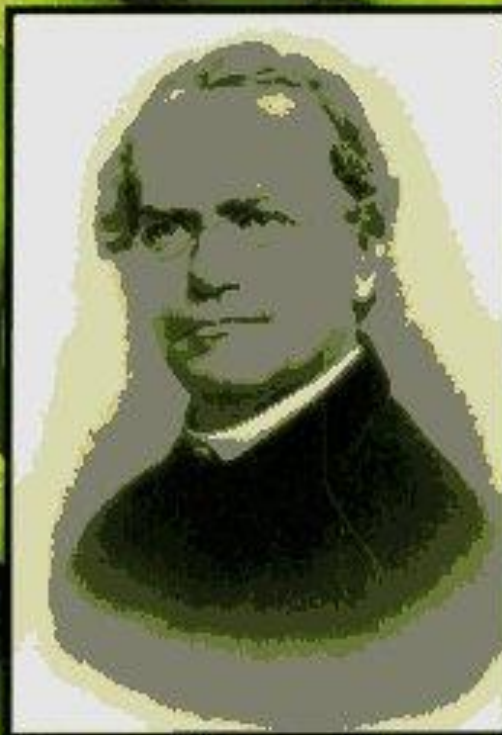
**1866 - Versuche über Pflanzen-Hybriden**  
**- principles of segregation**



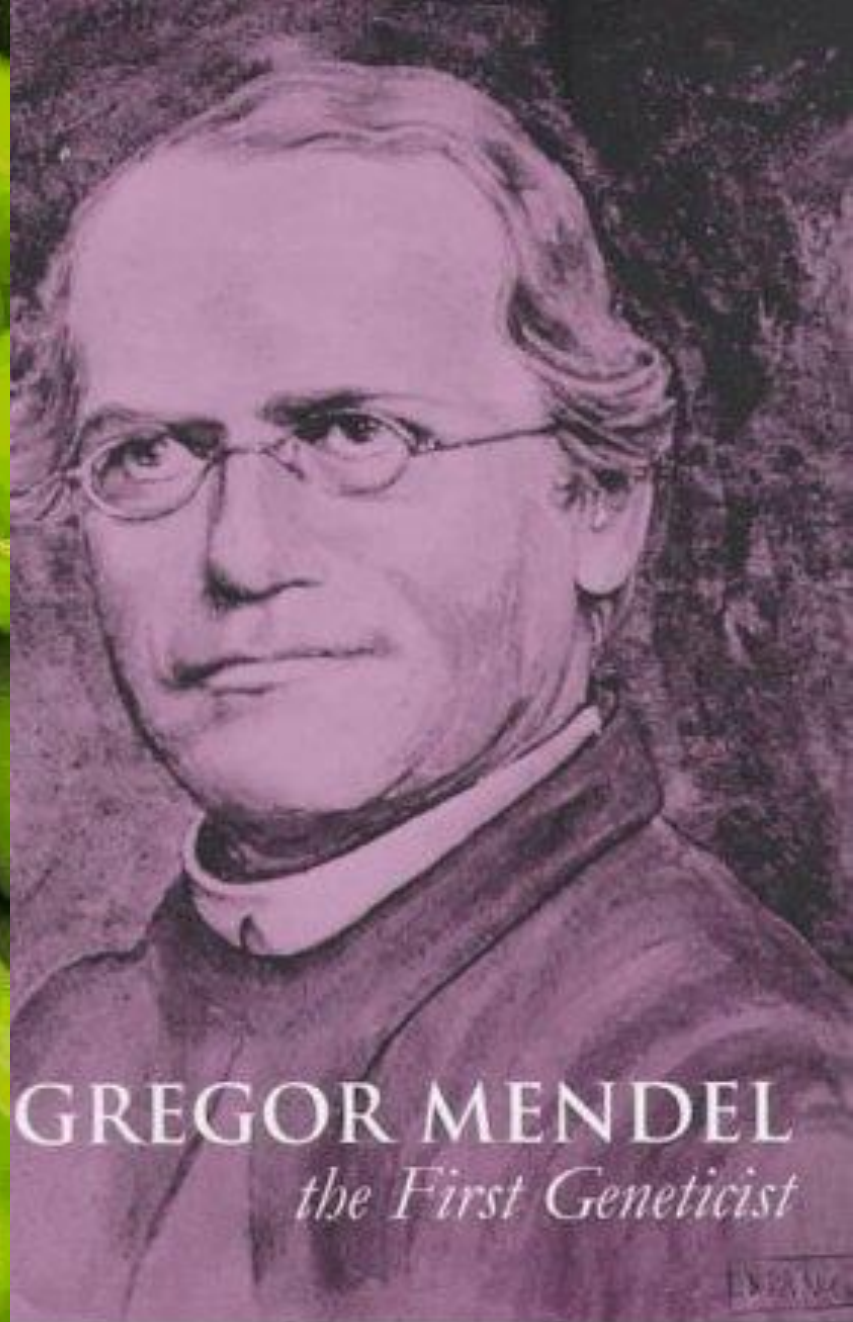
GREAT MINDS OF SCIENCE

# GREGOR MENDEL

FATHER OF  
GENETICS



Roger Klare
















GREGOR MENDEL  
*the First Geneticist*

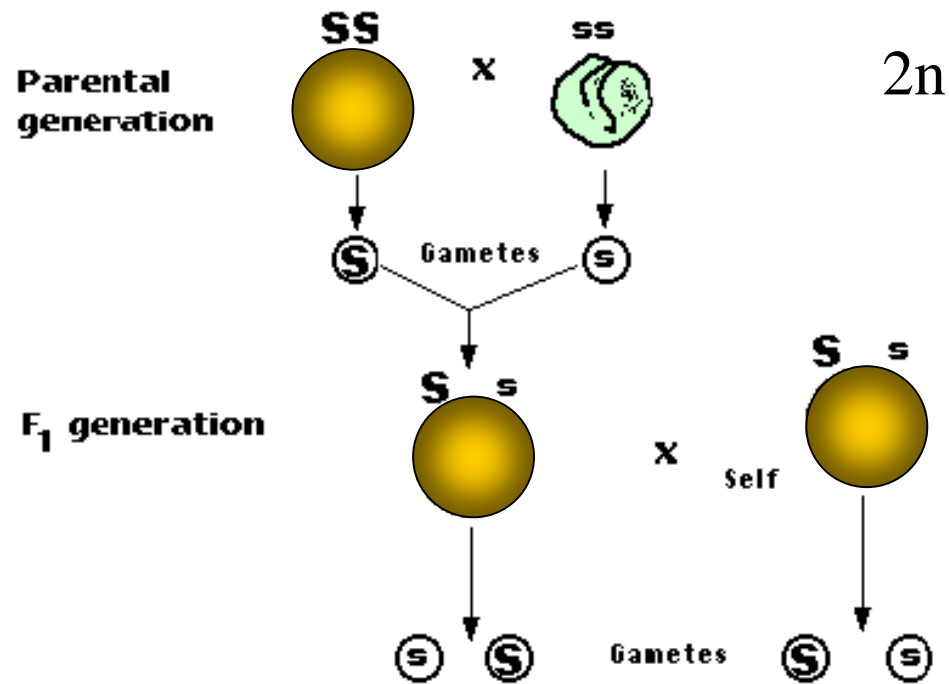
Vítěz



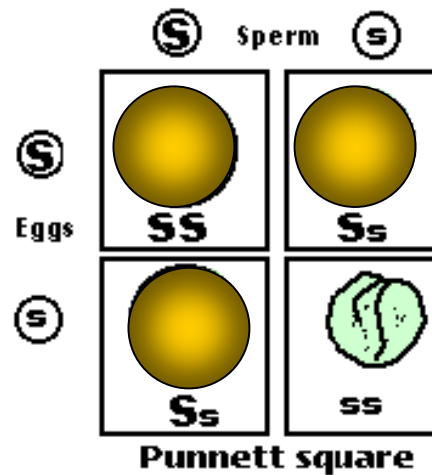
BY SA

Character	Dominant trait	Recessive trait	Character	Dominant trait	Recessive trait
Seed shape	 Spherical	 Wrinkled	Flower position	 Axial	 Terminal
Seed color	 Yellow	 Green		Stem height	 Tall
Flower color	 Purple	 White			
Pod shape	 Inflated	 Constricted			
Pod color	 Green	 Yellow			

**Self-pollination – series of passages – genetically pure strain**



**F<sub>2</sub> generation**



20th century English geneticist  
- Reginald **Punnett**

**Phenotype**

**3 : 1**

# Why is the mutant wrinkled?

Dominant  
allele R



Functional starch  
branching enzyme

RR or Rr



unbranched starch

branched starch

Round

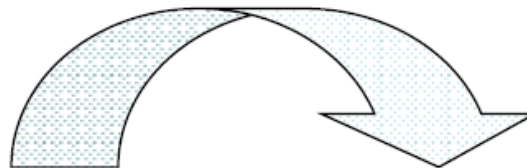


Recessive  
allele r



Inactive starch  
branching enzyme

rr



unbranched starch

unbranched starch

Wrinkled

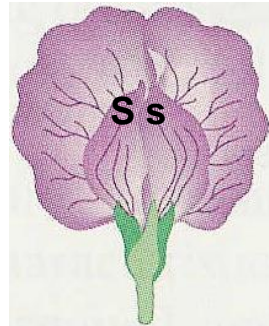




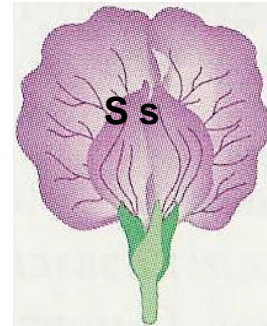
# Crossing hybrid offspring with purple flowers with F<sub>1</sub> generation

Deduced -  
segregation of traits

F<sub>1</sub> generation



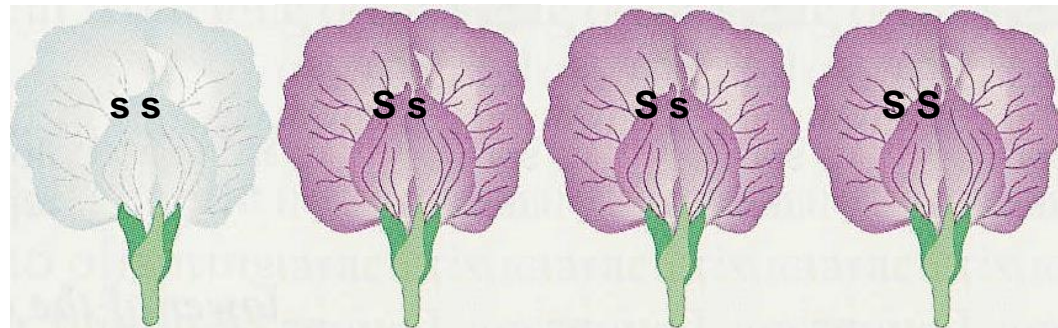
×



Hybrid offspring



F<sub>2</sub> generation  
(2. filial)



$\frac{1}{4}$  white

$\frac{3}{4}$  violet

**Established terms dominant and recessive heritable traits**

**Traits are controlled by certain factors**

**There is a relationship between factors and external expression**

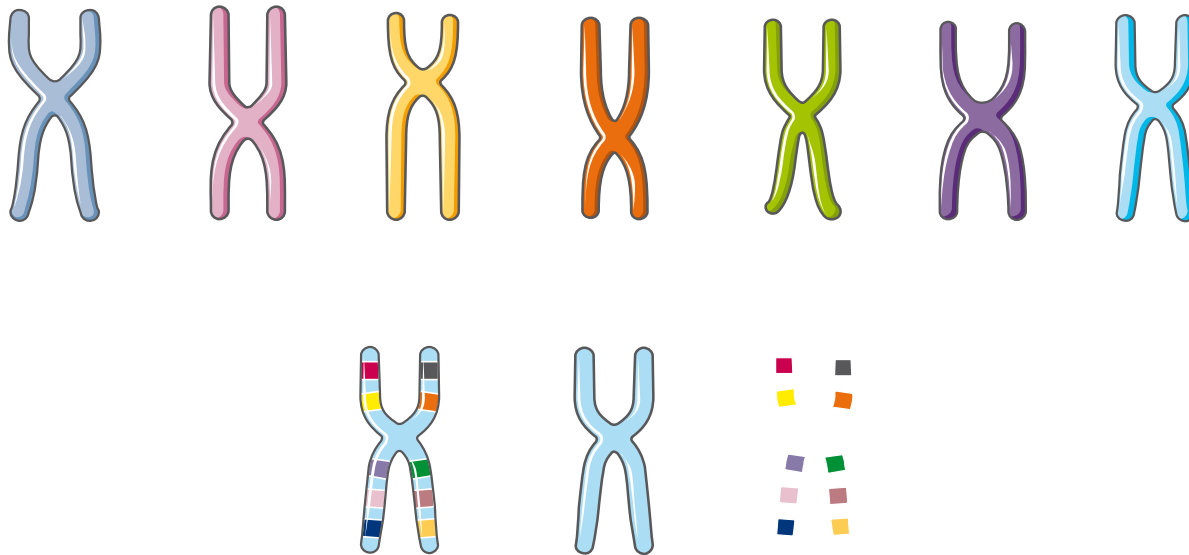
**1. Mendel's law (law of segregation) – Heritable units (pair alleles) in somatic cells are independent and segregate (separate) during meiosis (formation of gametes)**

**2. Mendel's law (law of independent assortment) – separation of any pair of allelic genes occurs independently (if they are on different pair of homologous chromosomes)**

# Diploid cells - two sets of each somatic chromosome

→ two copies of each gene

– genes - in different variants (**alleles**)

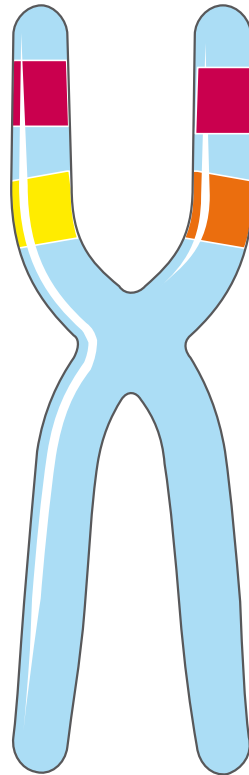


- **Allele** - variant form of a gene with slightly different sequence
- **Alleles** - created by mutations (errors in the copying of DNA)

Individuals can have one or two alleles for any gene

two of the same allele = **homozygous**

two different alleles = **heterozygous**





**Phenotype – the observable characteristics of an organism**

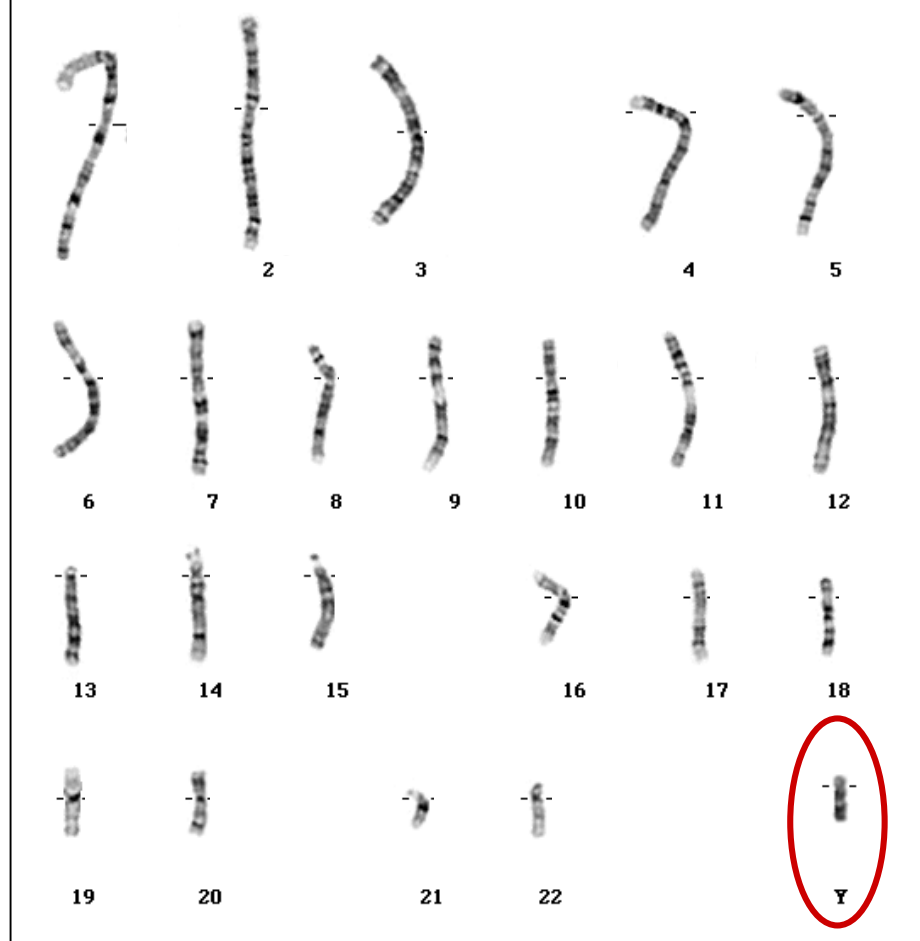
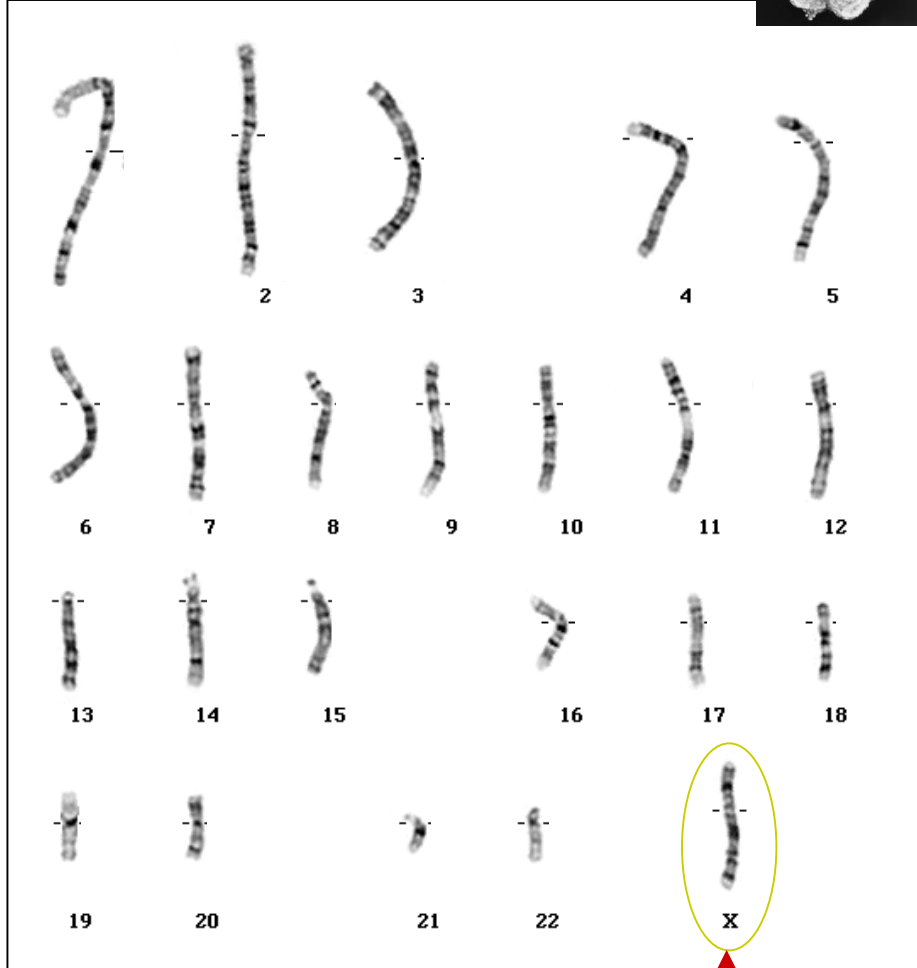
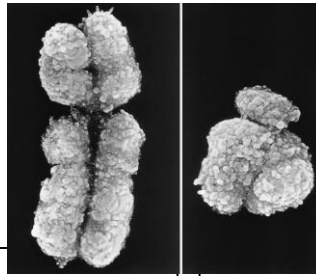
**Alleles: Dominant      Semidominant      Recessive**

**Dominant** alleles - expressed whether there is one copy or two

**Recessive** alleles must be present in two copies to be expressed



# Sperm determines the gender of the descendant



Egg always X

$2n$

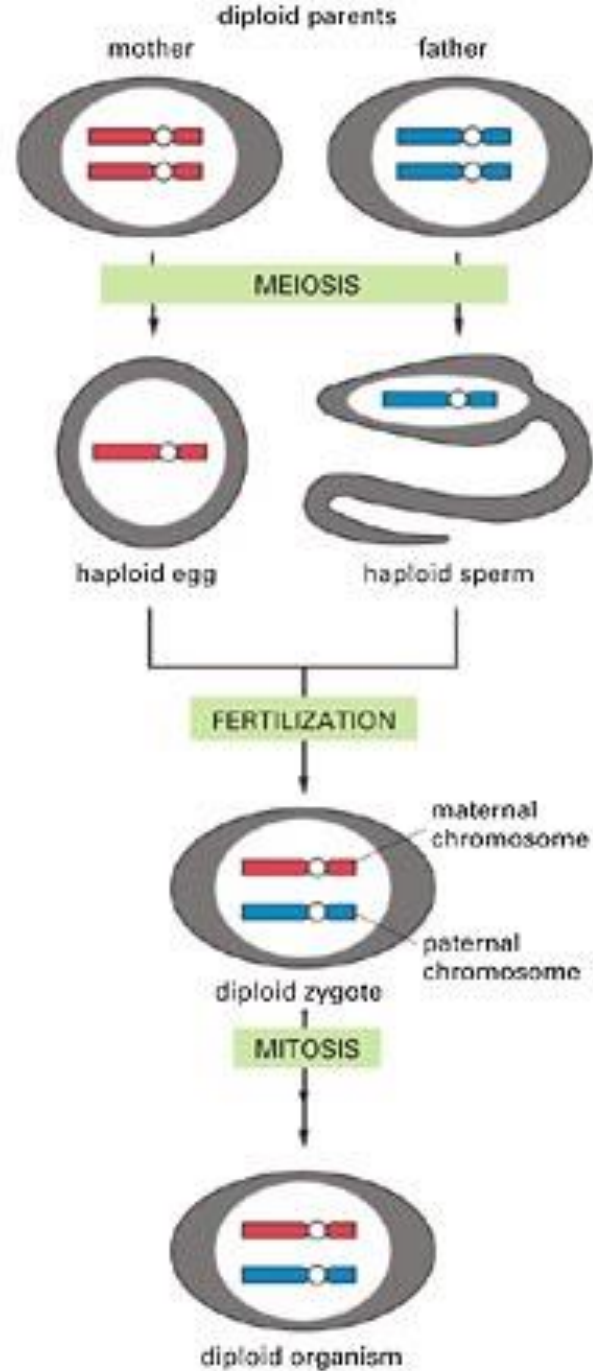
$n$

$2n$

$n$

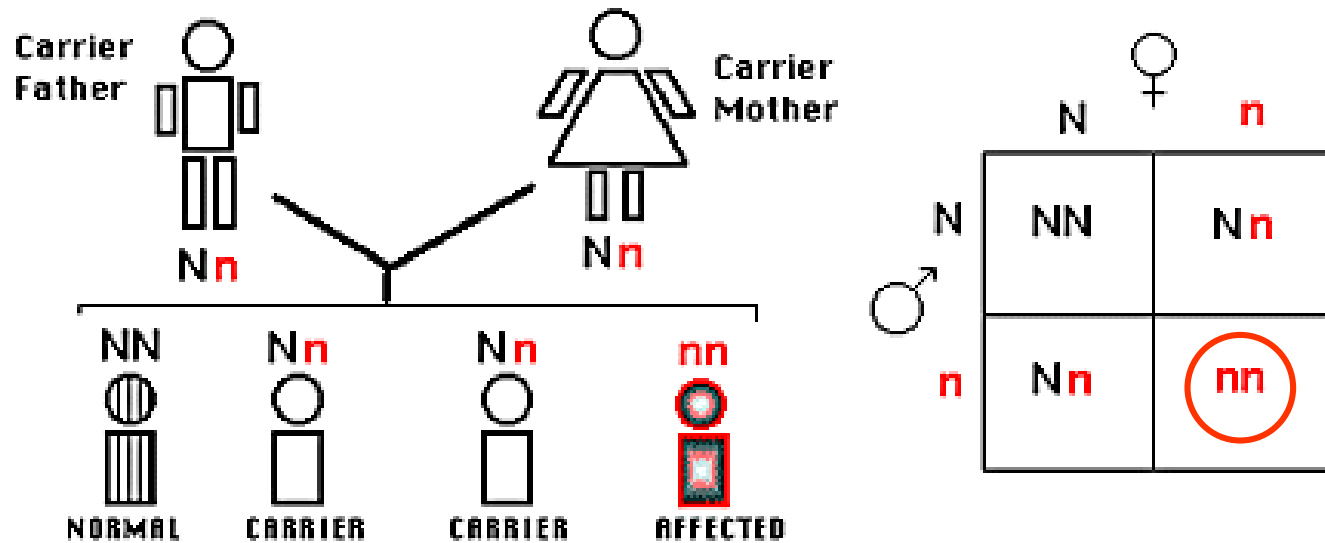


(A)

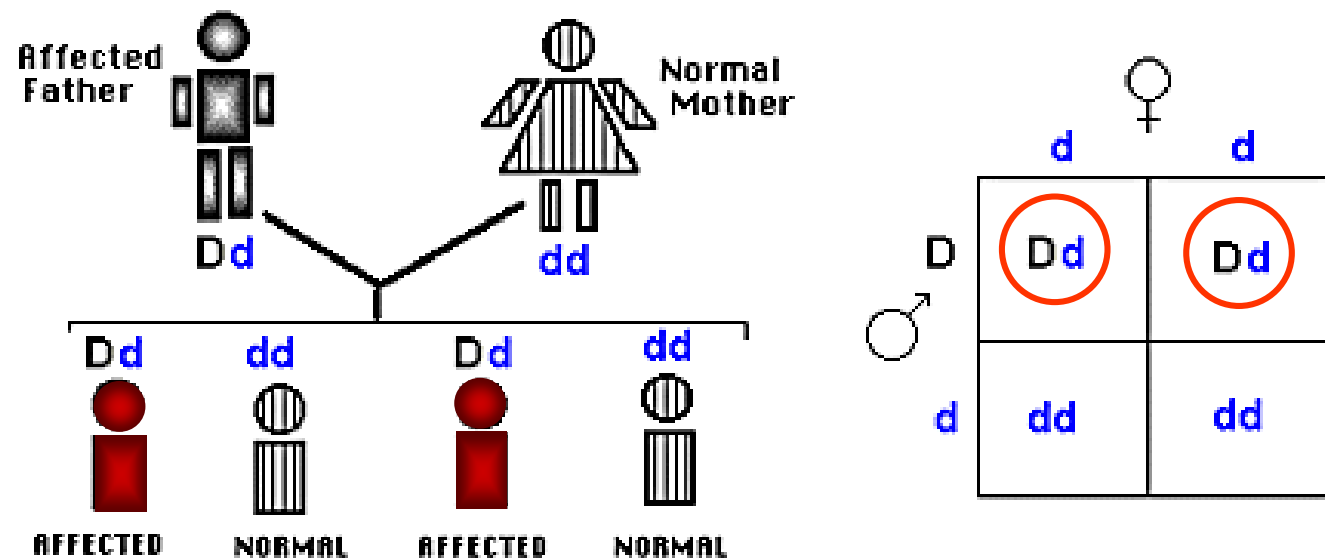


(B)

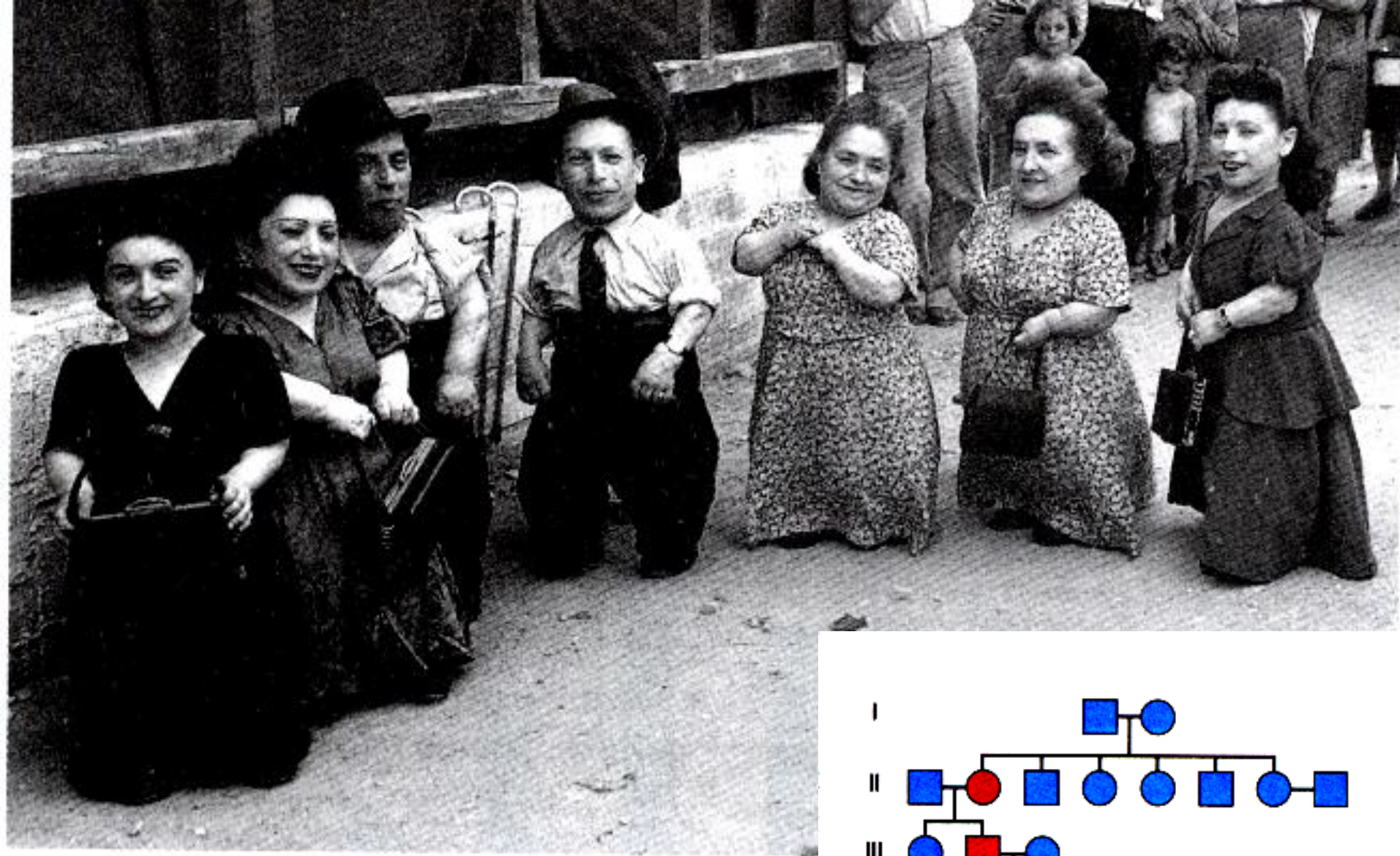
## Recessive inheritance



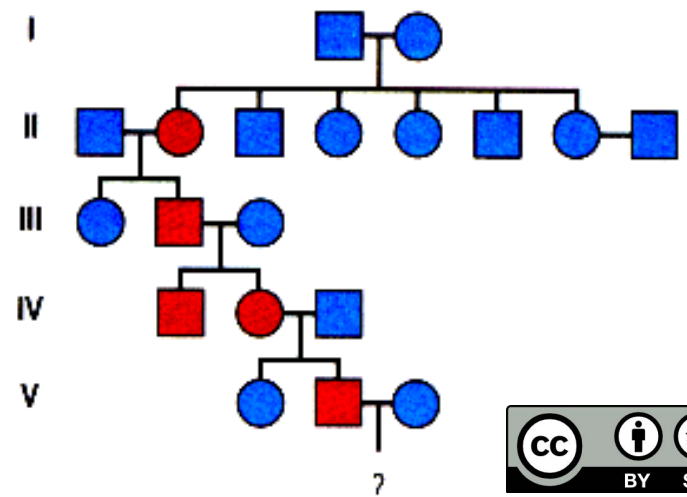
## Dominant inheritance







**Dwarfism – dominant mutation**





## **polydactyly**

**Autosomal dominant mutations of different genes involved in the ontogenesis**

**The Bible mentions Philistine warriors with six fingers and toes**



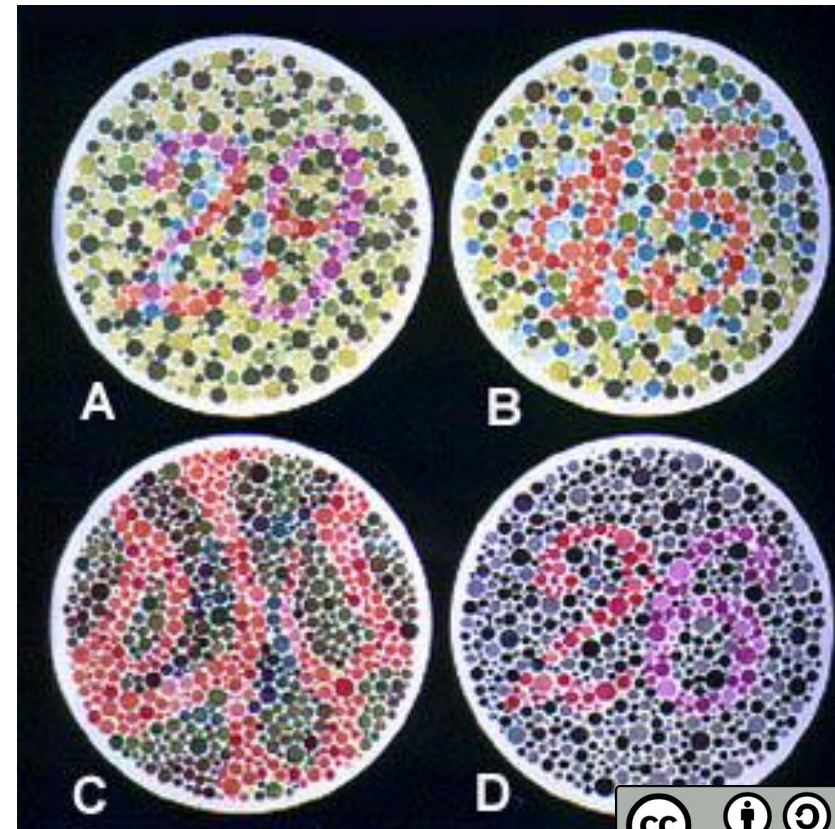
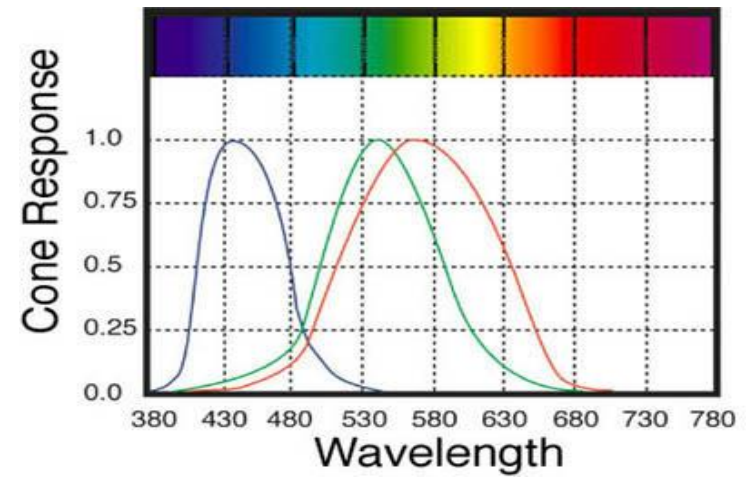
# Colorblindness

## Retina 3 receptors

The gene for the blue receptor - autosomal

The genes for red and green receptor  
- X chromosome (sex-linked)

mutations → most color-blind - men



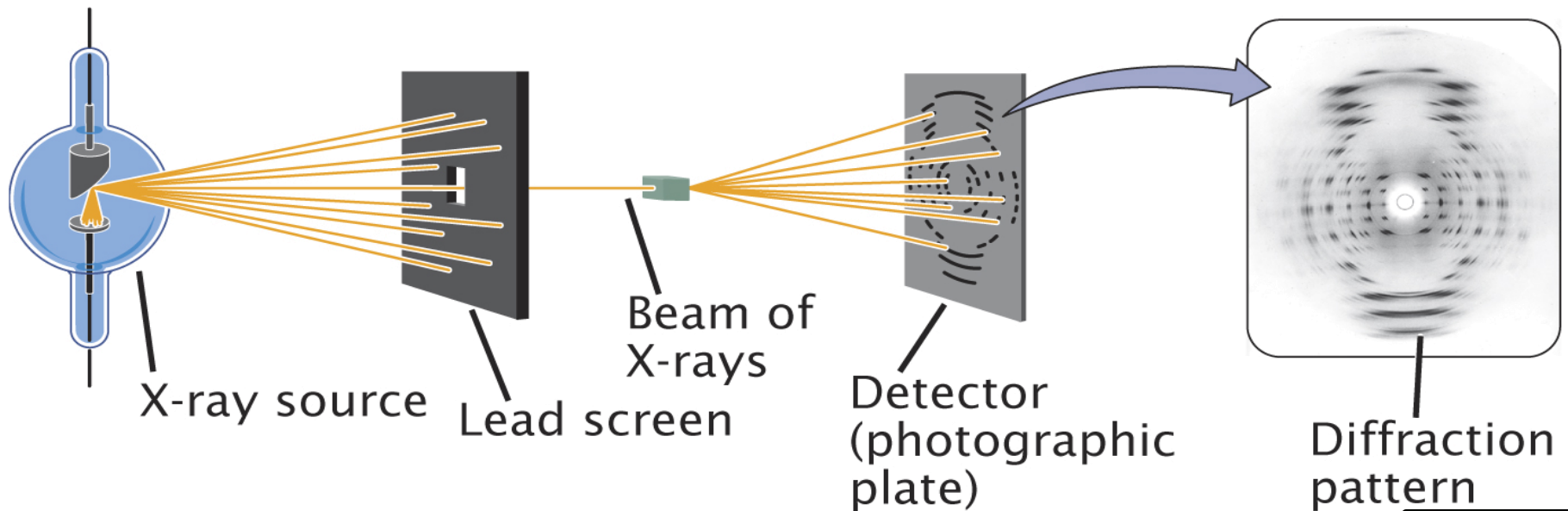
How do genes function?

The discovery of DNA structure

Rosalind Franklin and Maurice Wilkins, 1952

Crystallographic data:

DNA - a long, thin molecule with repeated structural motif (3.4 Å and 34 Å)



Erwin Chargraff, 1950

% A =% T; % C =% G (different species - varying GC content)



Table 3-2 Data Leading to the Formulation of Chargaff’s Rules

Source	Adenine to Guanine	Thymine to Cytosine	Adenine to Thymine	Guanine to Cytosine	Purines to Pyrimidines
Ox	1.29	1.43	1.04	1.00	1.1
Human	1.56	1.75	1.00	1.00	1.0
Hen	1.45	1.29	1.06	0.91	0.99
Salmon	1.43	1.43	1.02	1.02	1.02
Wheat	1.22	1.18	1.00	0.97	0.99
Yeast	1.67	1.92	1.03	1.20	1.0
<i>Hemophilus influenzae</i>	1.74	1.54	1.07	0.91	1.0
<i>E-coli</i> K2	1.05	0.95	1.09	0.99	1.0
Avian tubercle bacillus	0.4	0.4	1.09	1.08	1.1
<i>Serratia marcescens</i>	0.7	0.7	0.95	0.86	0.9
<i>Bacillus schatz</i>	0.7	0.6	1.12	0.89	1.0

SOURCE: After E. Chargaff et al., *J. Biol. Chem.* 177 (1949).

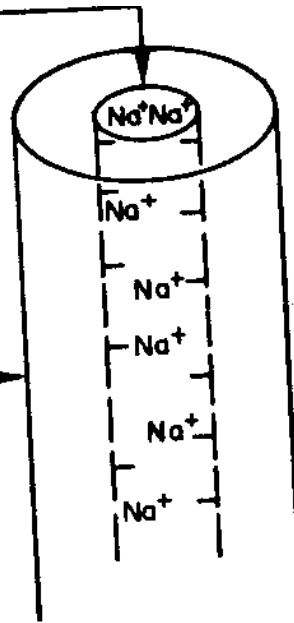


James Watson, Francis Crick, 1953

Model of the DNA structure - the pairing of bases → Replication

Sugar-Phosphate backbones  
on the inside, held together by  
attraction between  $\text{PO}_2^-$  and  $\text{Na}^+$

Bases on outside of  
cylindrical molecule



## Watson and Crick's first model 1951 – 3 chains

### Unpublished manuscript (1951), Crick wrote:

*"There are no atoms which can donate hydrogen bonds except in the basic rings and the water, thus **hydrogen bonding is unlikely to play the dominant part in the structure** that it does in the polypeptide alpha helix. electrostatic forces are so big relative to Van der Waals forces that we may be confident that the  **$\text{Na}^+$  and the  $\text{PO}_2^-$**  will mainly...."*



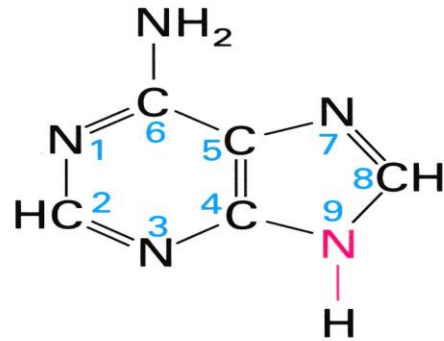
1953



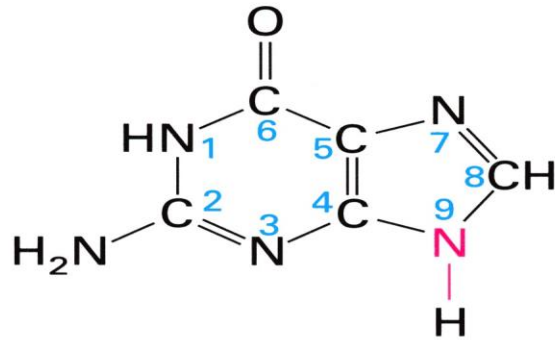
**J.D. Watson**

**F.H.C. Crick**

## PURINES

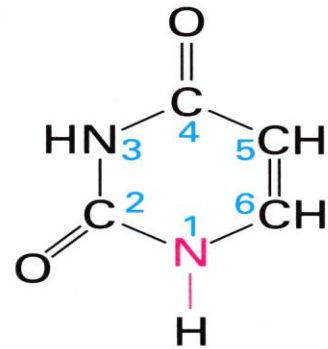


**Adenine (A)**

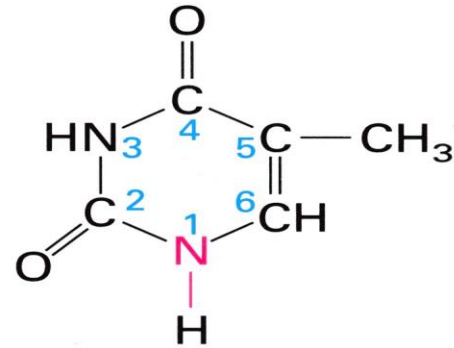


**Guanine (G)**

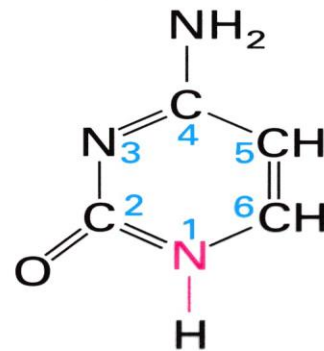
## PYRIMIDINES



**Uracil (U)**

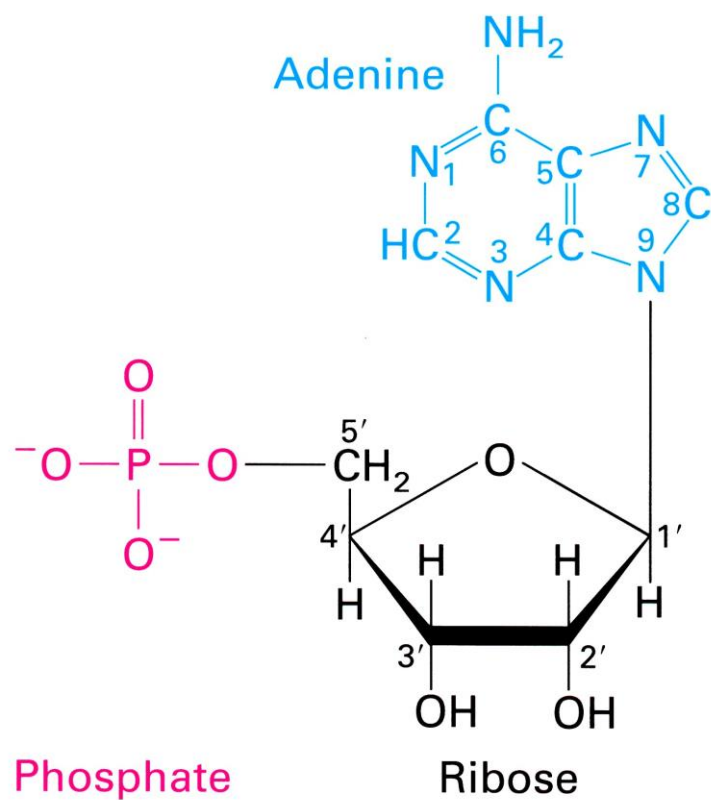


**Thymine (T)**

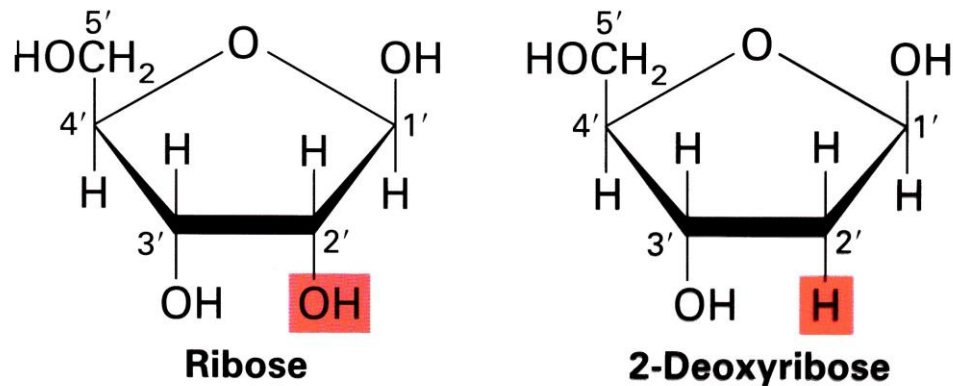


**Cytosine (C)**



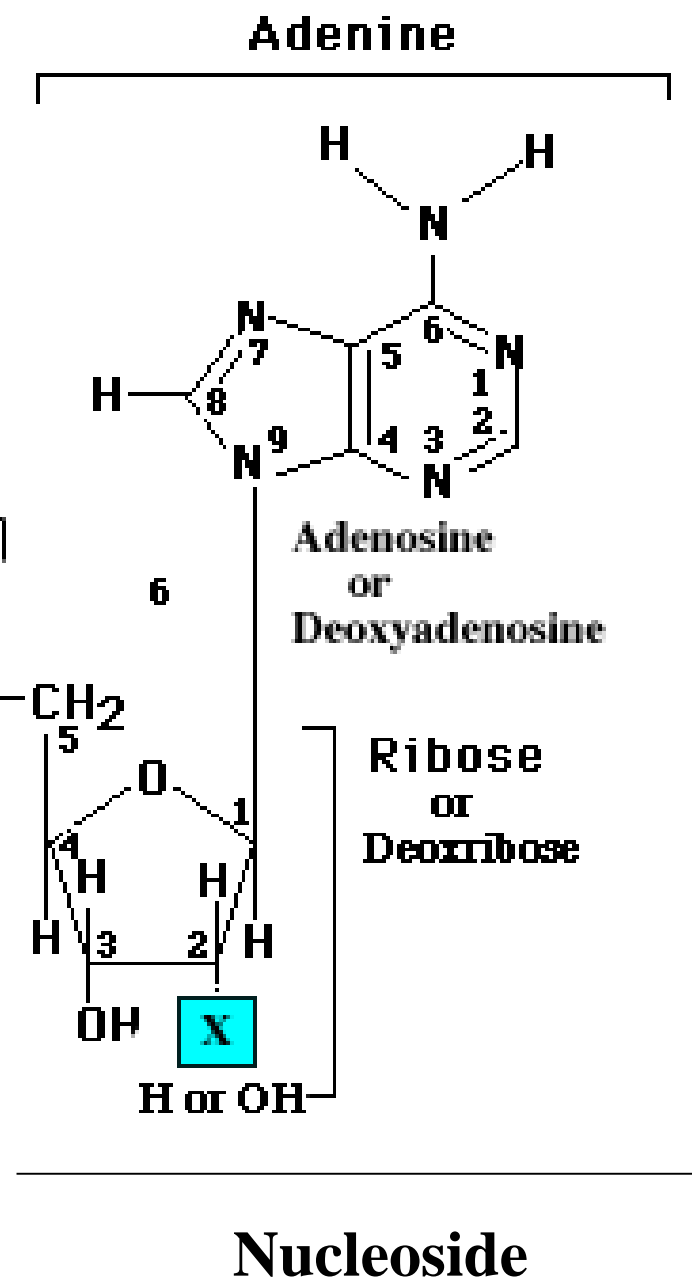
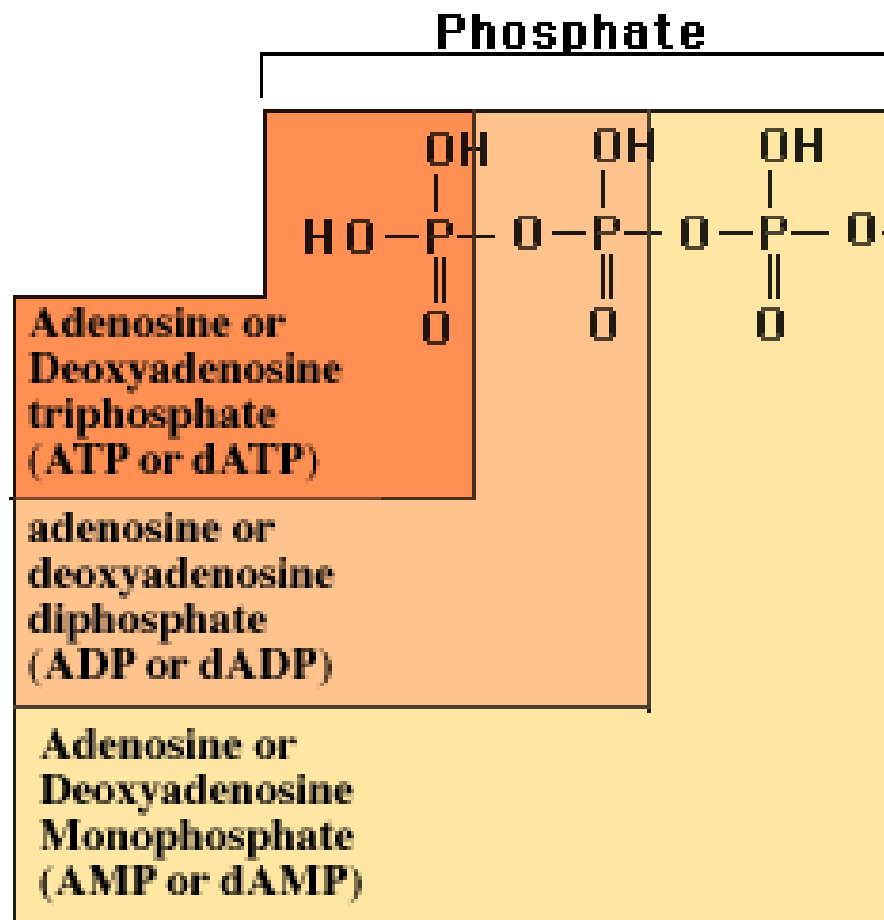


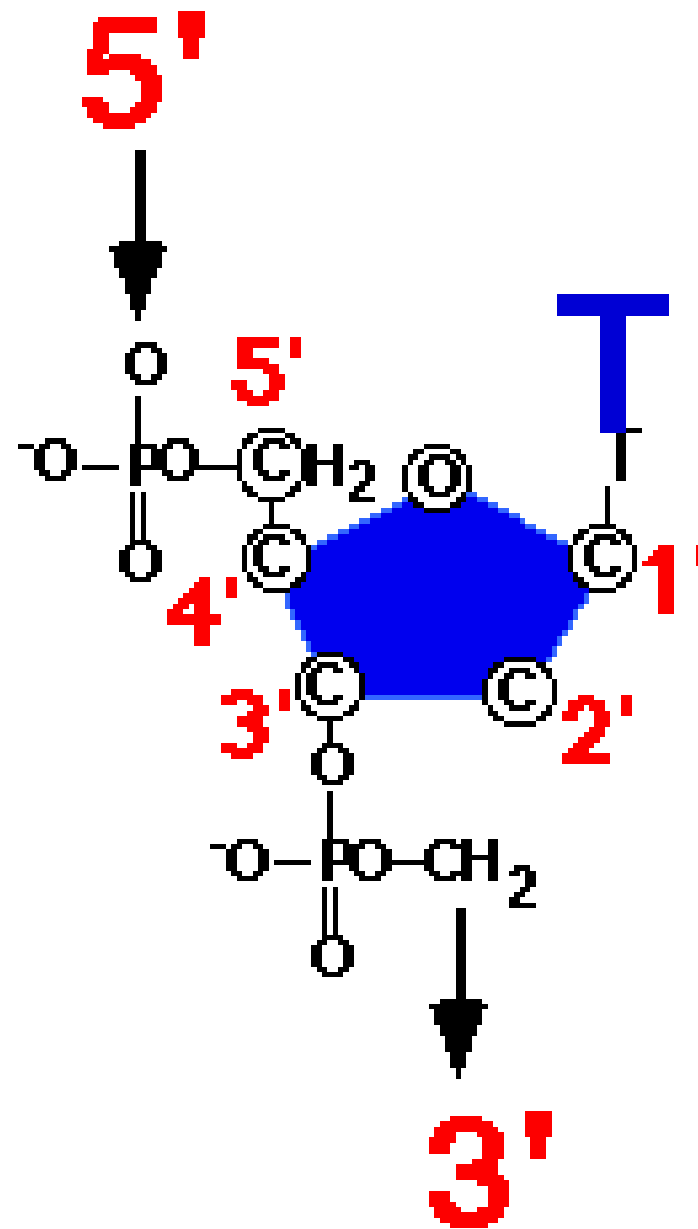
**Adenosine  
5'-monophosphate  
(AMP)**



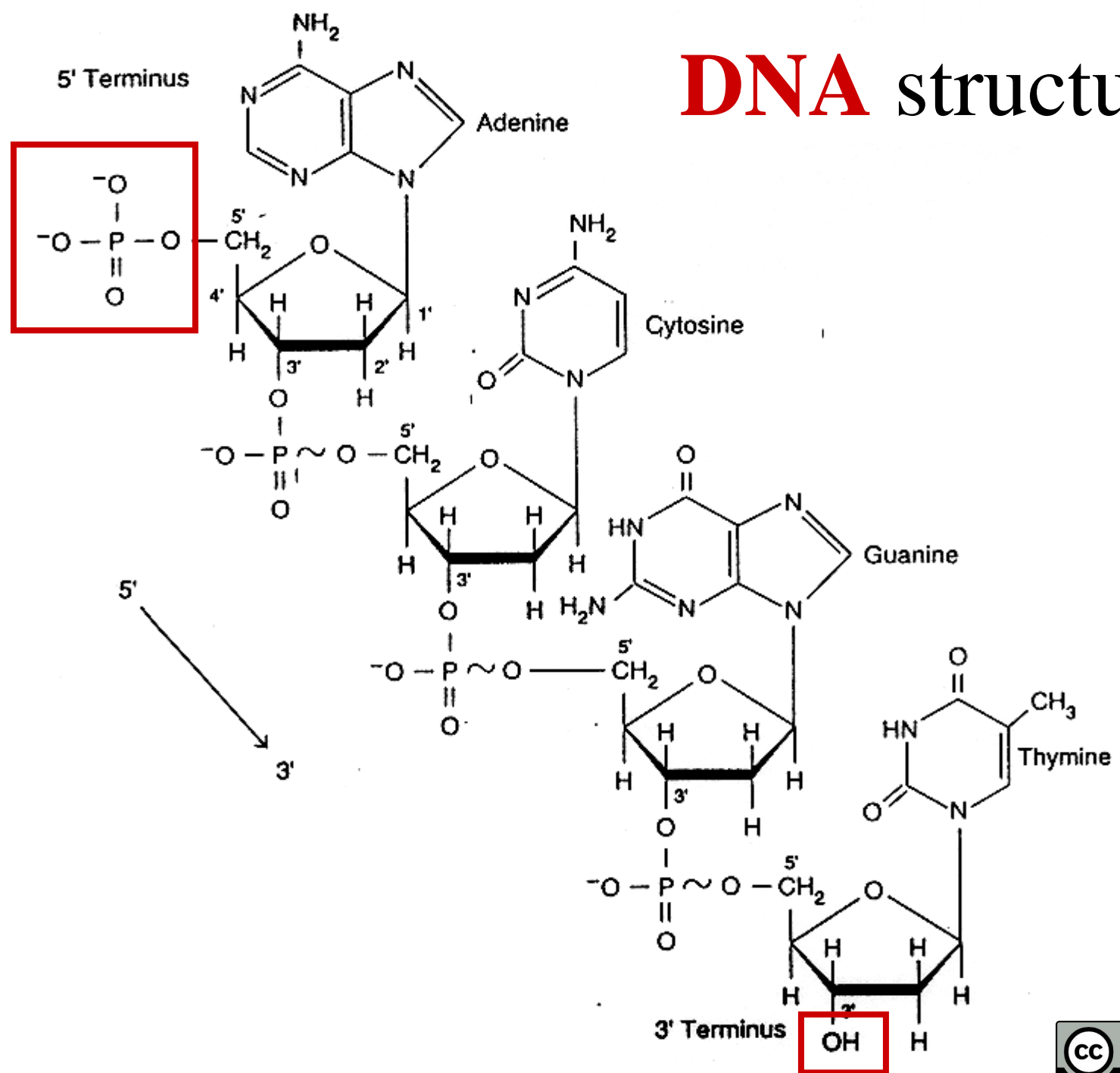
# Nucleotide

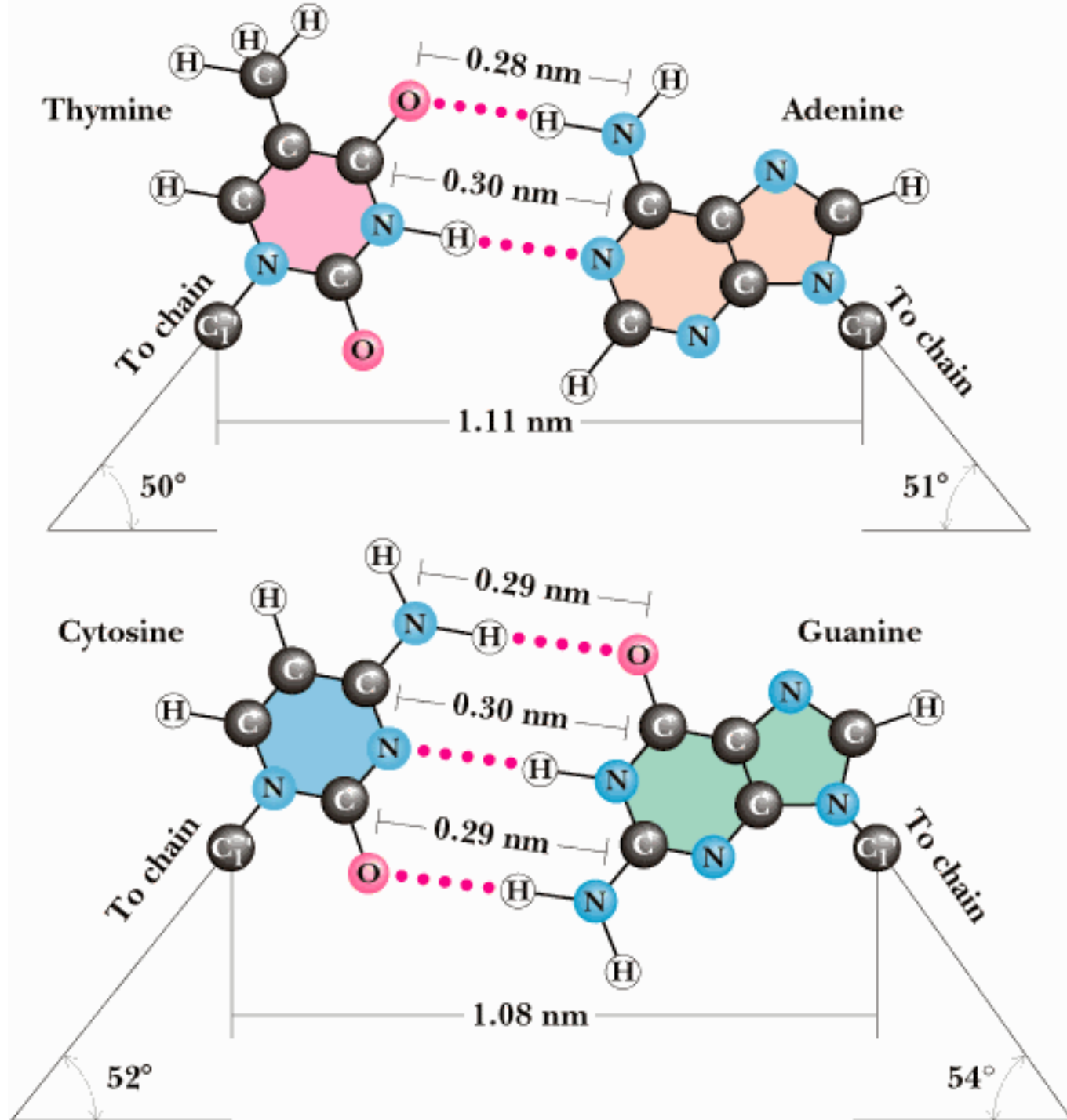
= nucleoside + phosphate





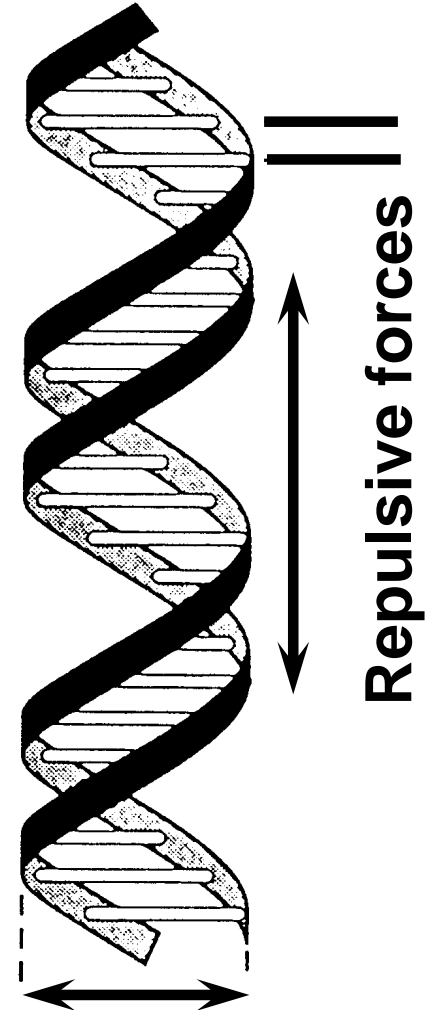
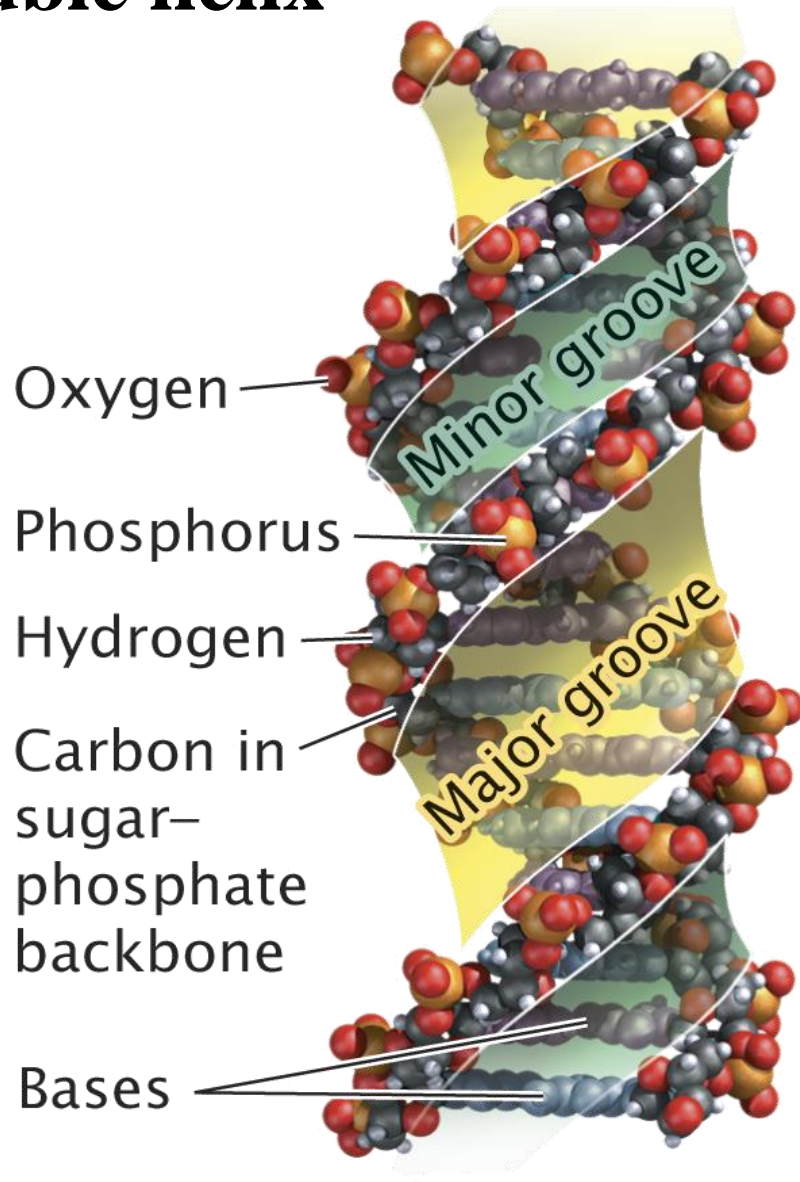
# DNA structure





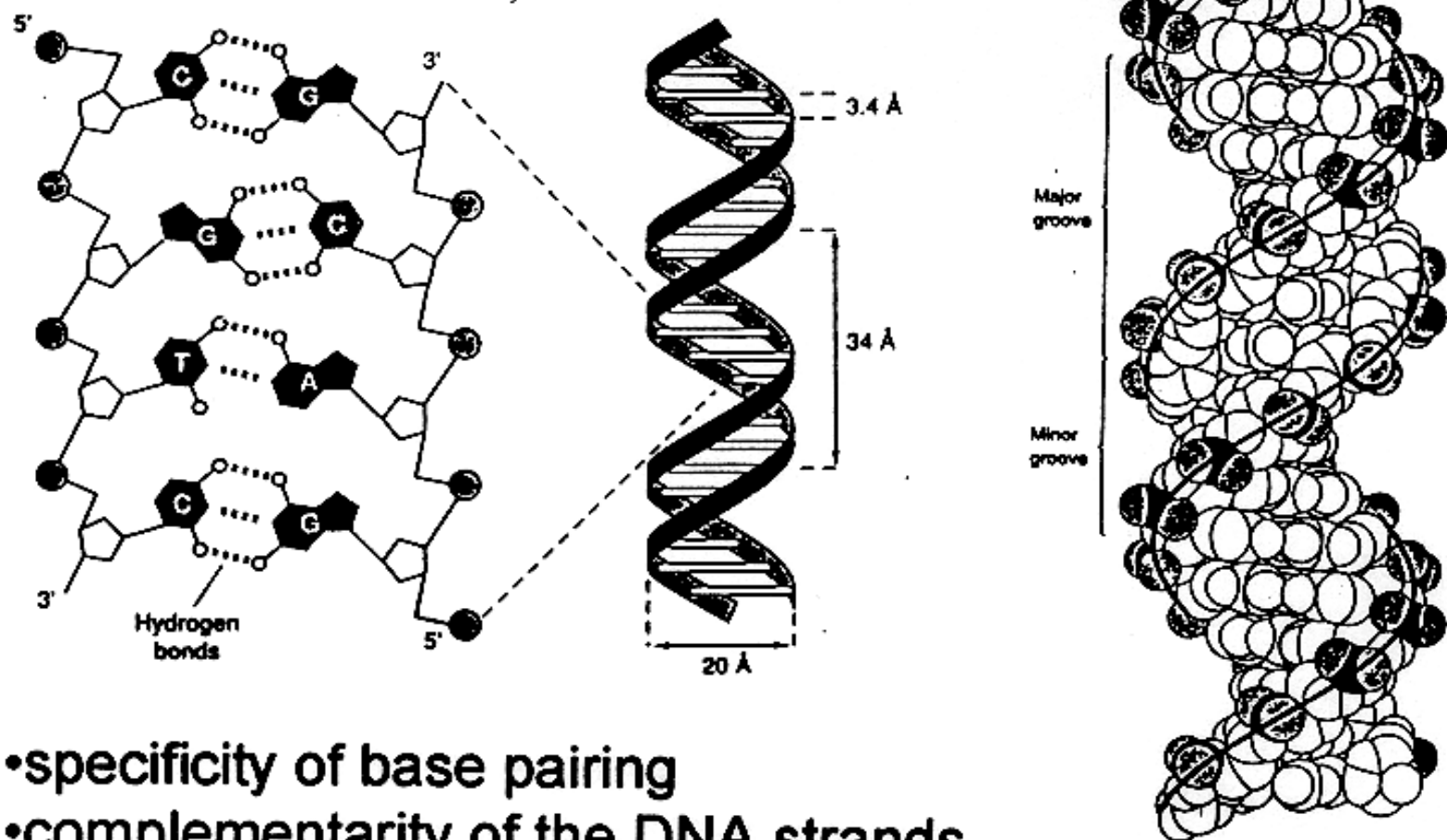
# DNA antiparallel double helix

## Stacking interactions





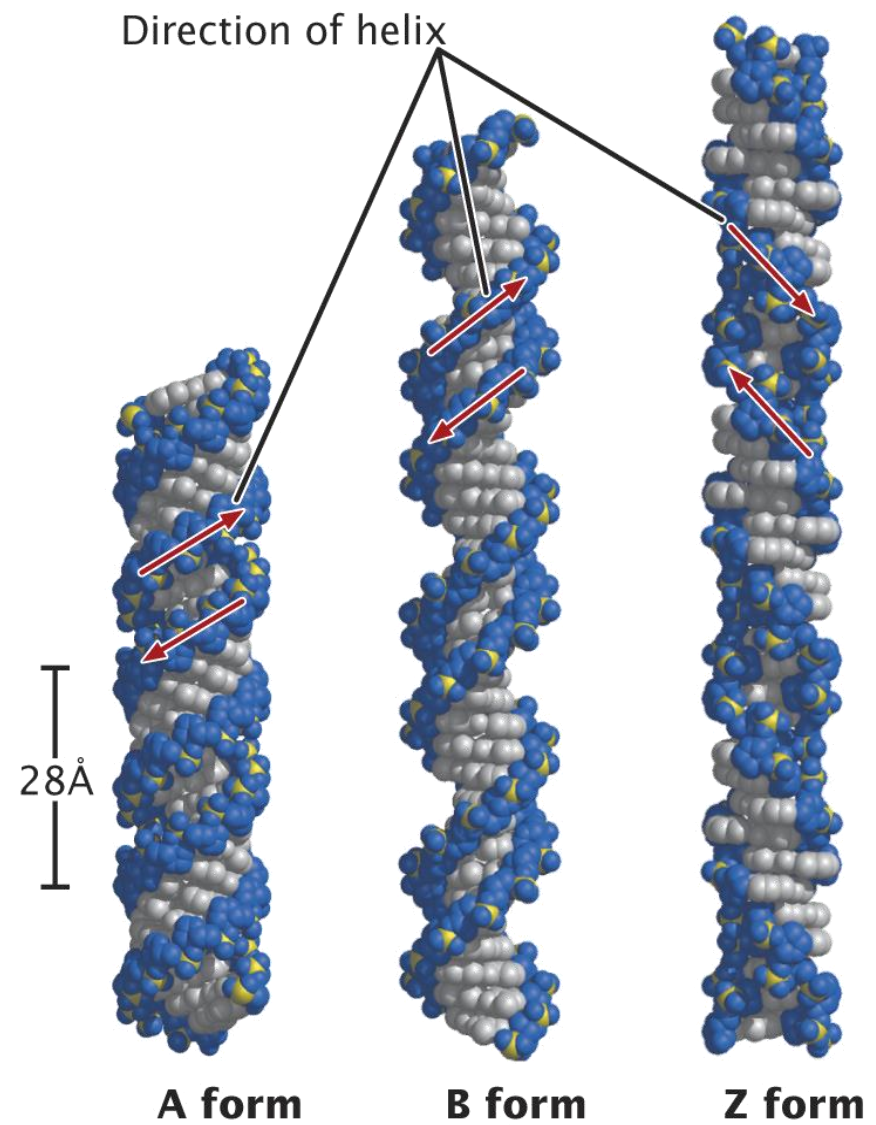
# Double-stranded DNA



- specificity of base pairing
- complementarity of the DNA strands
- B-DNA has 10 base-pairs per turn
- DNA with fewer or more bp per turn is "super"

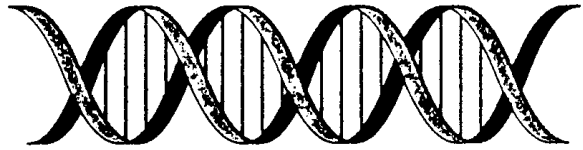
	<b>A</b>	<b>B</b>	<b>Z</b>
<b>bp/turn</b>	<b>11</b>	<b>10,5</b>	<b>12</b>
	<b>RH</b>	<b>RH</b>	<b>LH</b>

**RH – right handed**  
**LH – left handed**

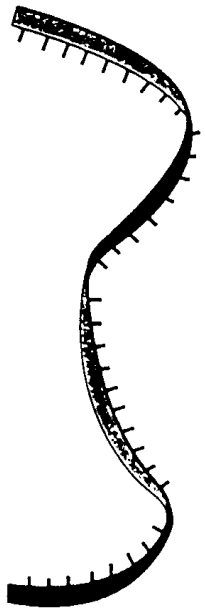
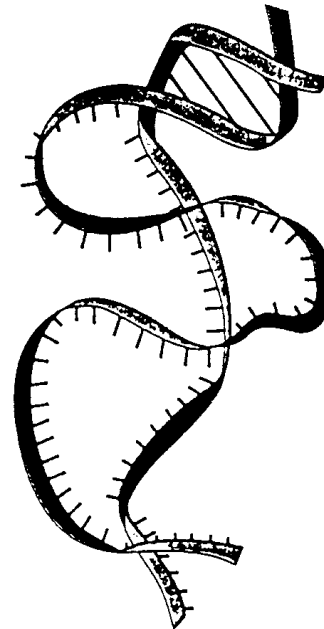
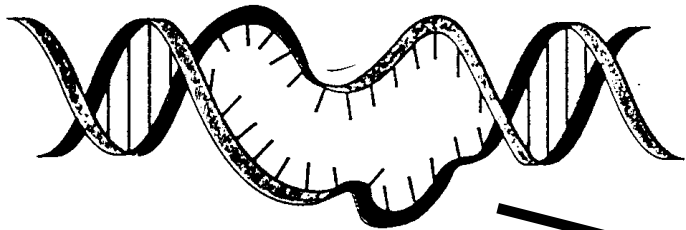


Fig\_10-15 *Genetics, Second Edition* © 2005 W.H. Freeman and Company

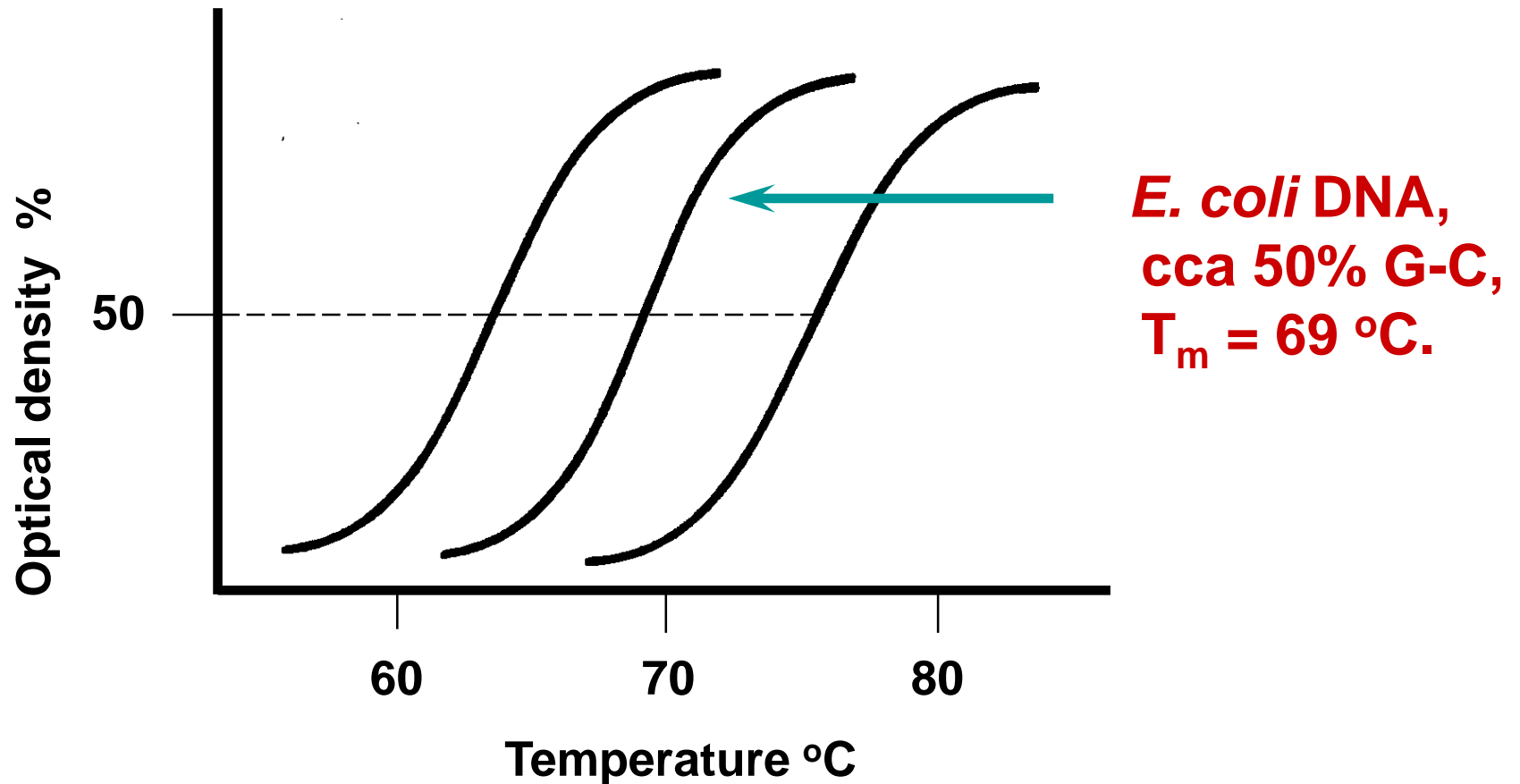
# Extreme pH or high temperature – DNA denaturing



A-T rich regions



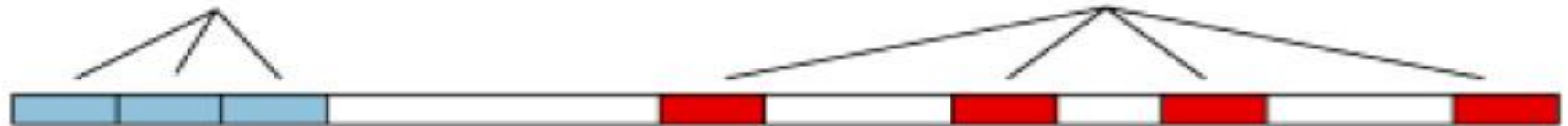
$$T_m \sim \text{G-C}$$



Average G-C content can be determined from the  $T_m$

GATCTNTTTNTTTT consensus

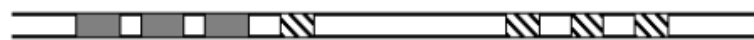
TTATNCANA consensus



Tandem 13-mer

Four 9-mer  
initiation protein  
binding sites

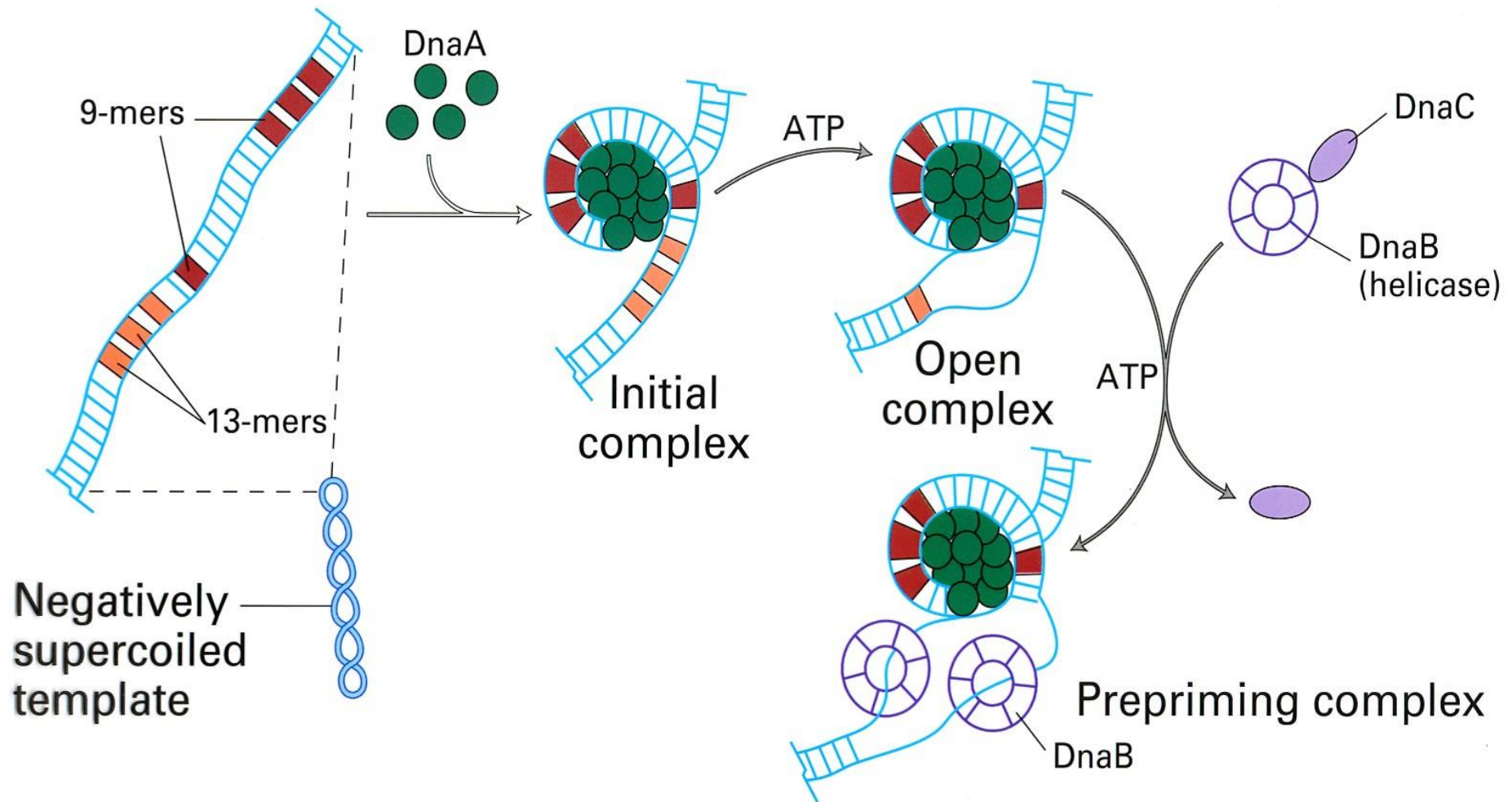
*OriC*  
(245 bp)



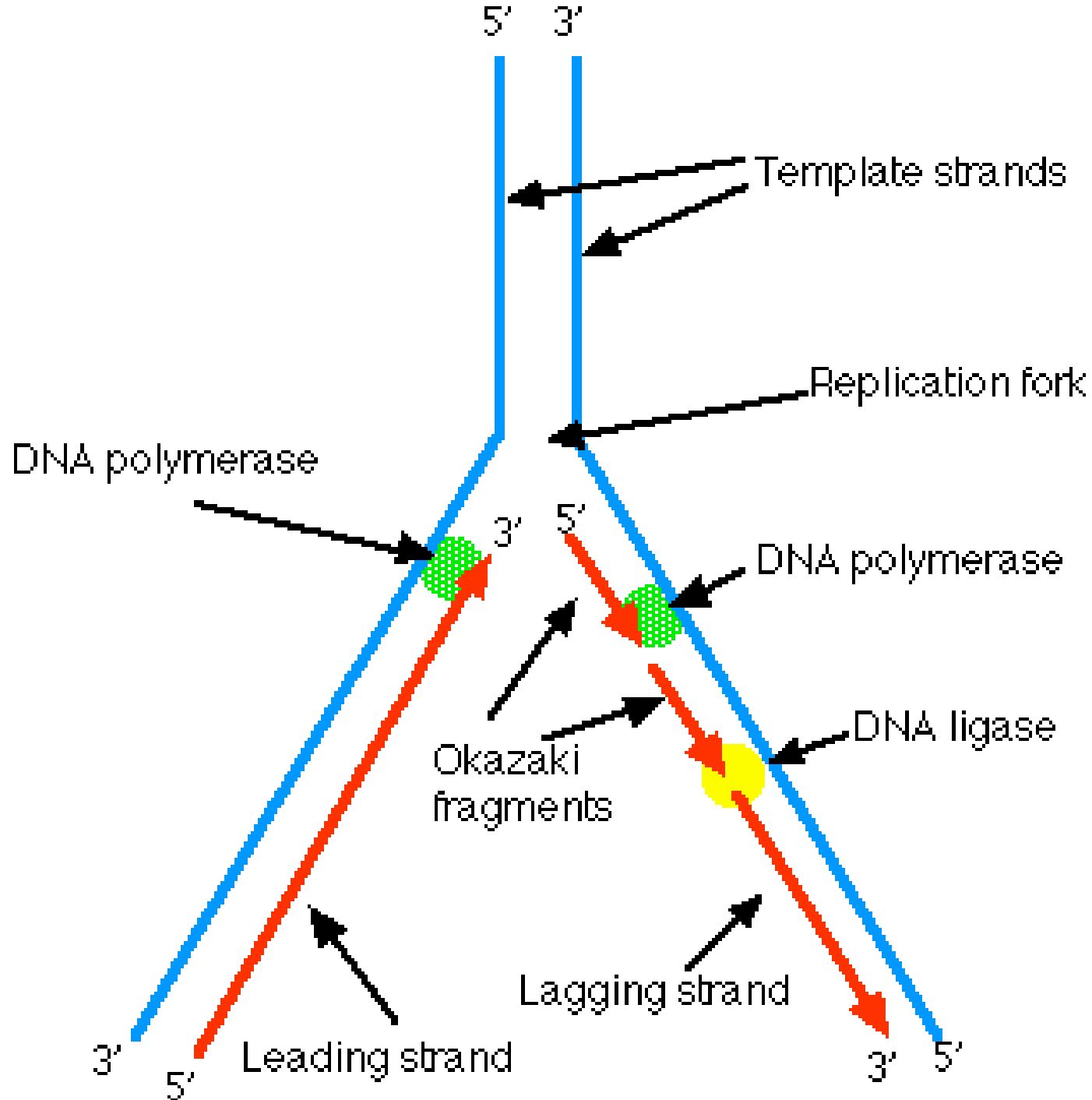
'13-mers'  
GATCTNTTTTATTT  
GATCTNTTNTATT  
GATCTCTTATTAG

'9-mers'  
TGTGGATAA  
TTATACACA  
TTTGGATAA  
TTATCCACA

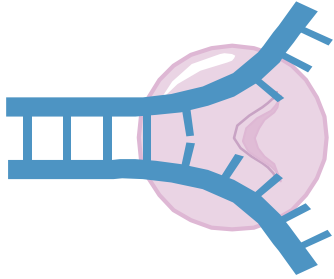




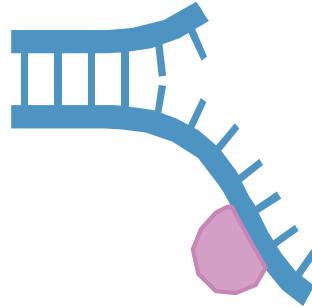




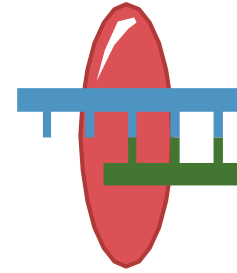
# Enzymes in DNA replication



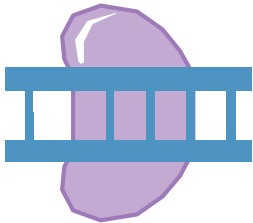
**Helicase unwinds  
parental double helix**



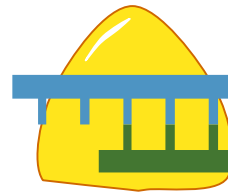
**Binding proteins  
stabilise separate  
strands**



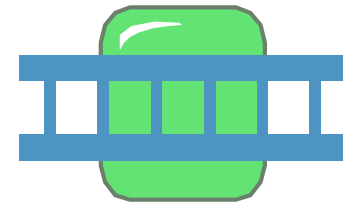
**Primase adds  
short primer  
to template strand**



**DNA polymerase  
binds nucleotides  
to form new strands**



**DNA polymerase I  
(Exonuclease) removes  
RNA primer and inserts  
the correct bases**

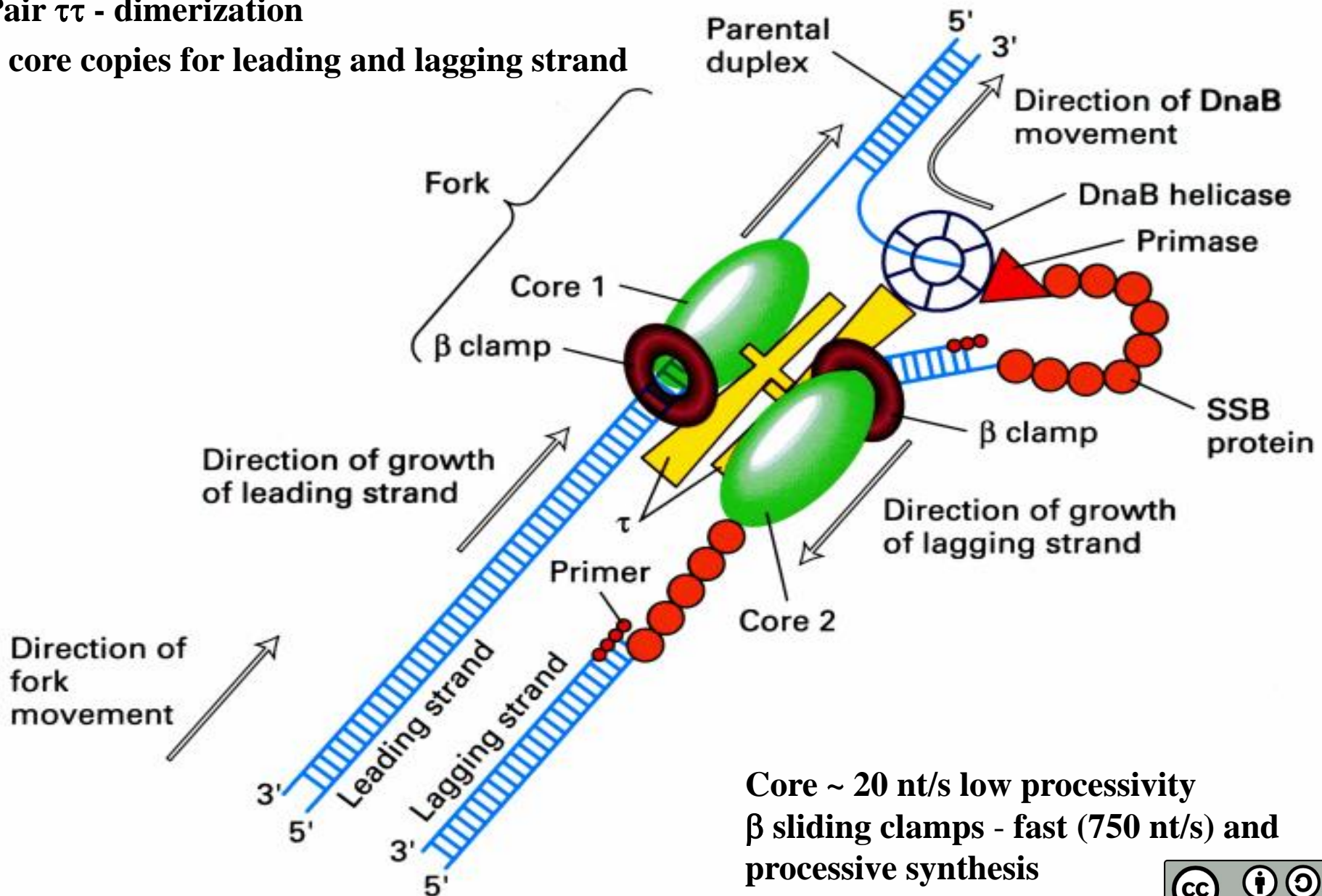


**Ligase joins Okazaki  
fragments and seals  
other nicks in sugar-  
phosphate backbone**

# REPLISOME COMPLEX - 2 POLYMERASES

Pair  $\tau\tau$  - dimerization

2 core copies for leading and lagging strand



Core ~ 20 nt/s low processivity  
 $\beta$  sliding clamps - fast (750 nt/s) and processive synthesis

# ***E. coli* DNA polymerases**

**3' → 5' exonuclease**

**5' → 3' exonuclease**

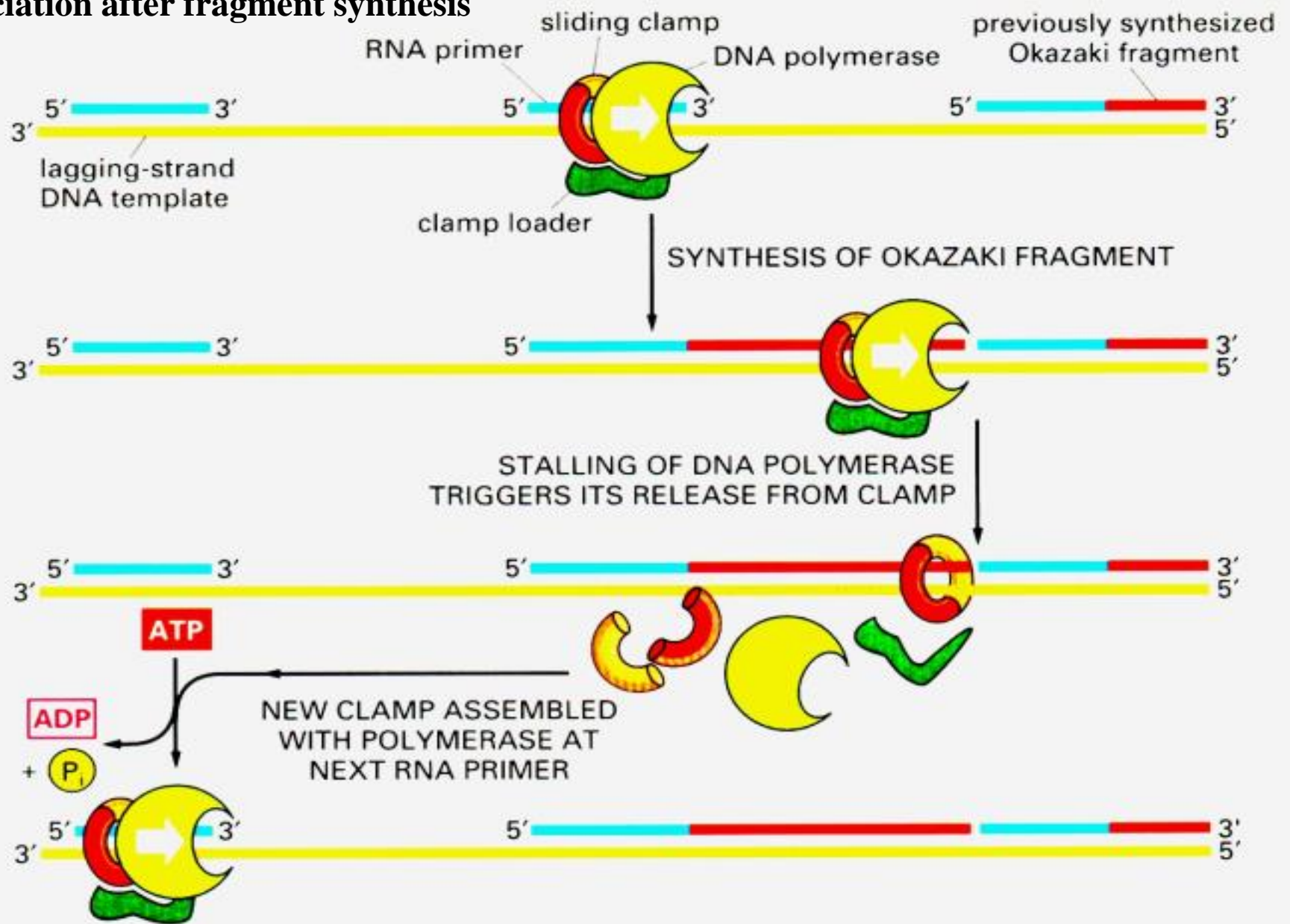
**5' → 3' polymerization**

**I      Repair**

**II     Repair**

**III    Replication + repair**

# Dissociation after fragment synthesis

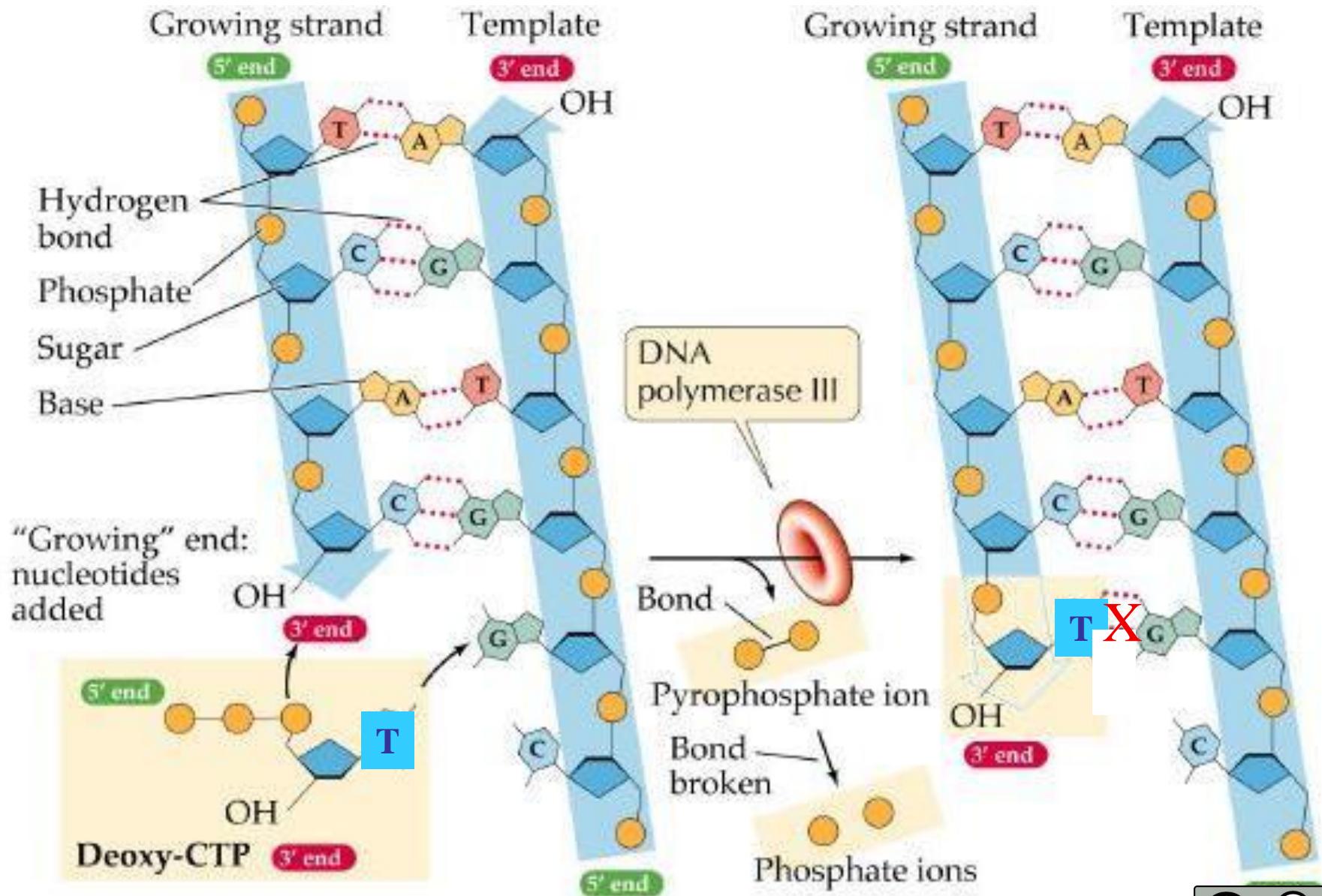


[Clamp – video1](#)

[Clamp - vi](#)

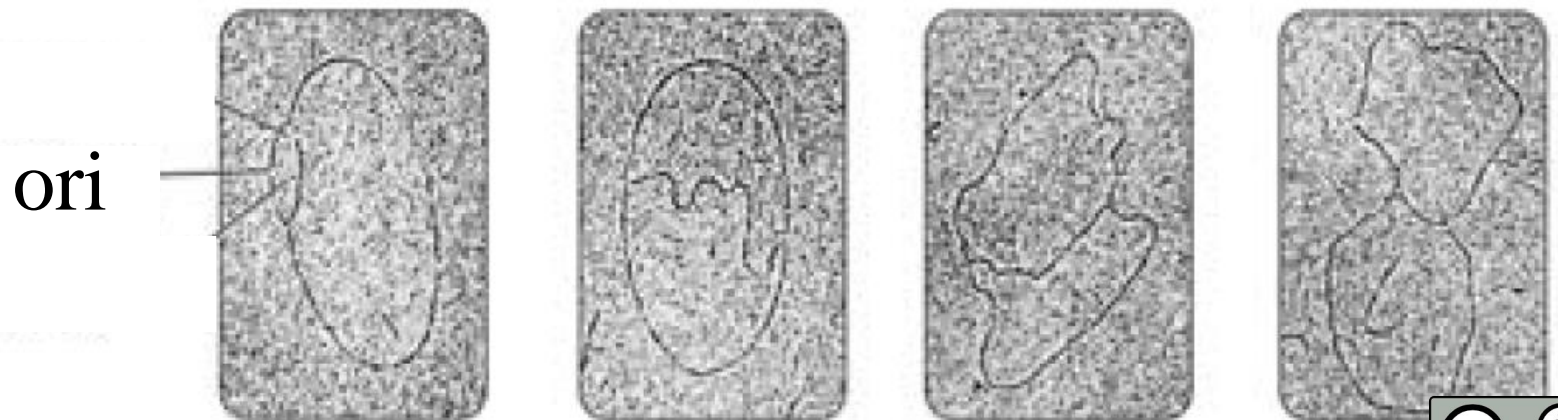
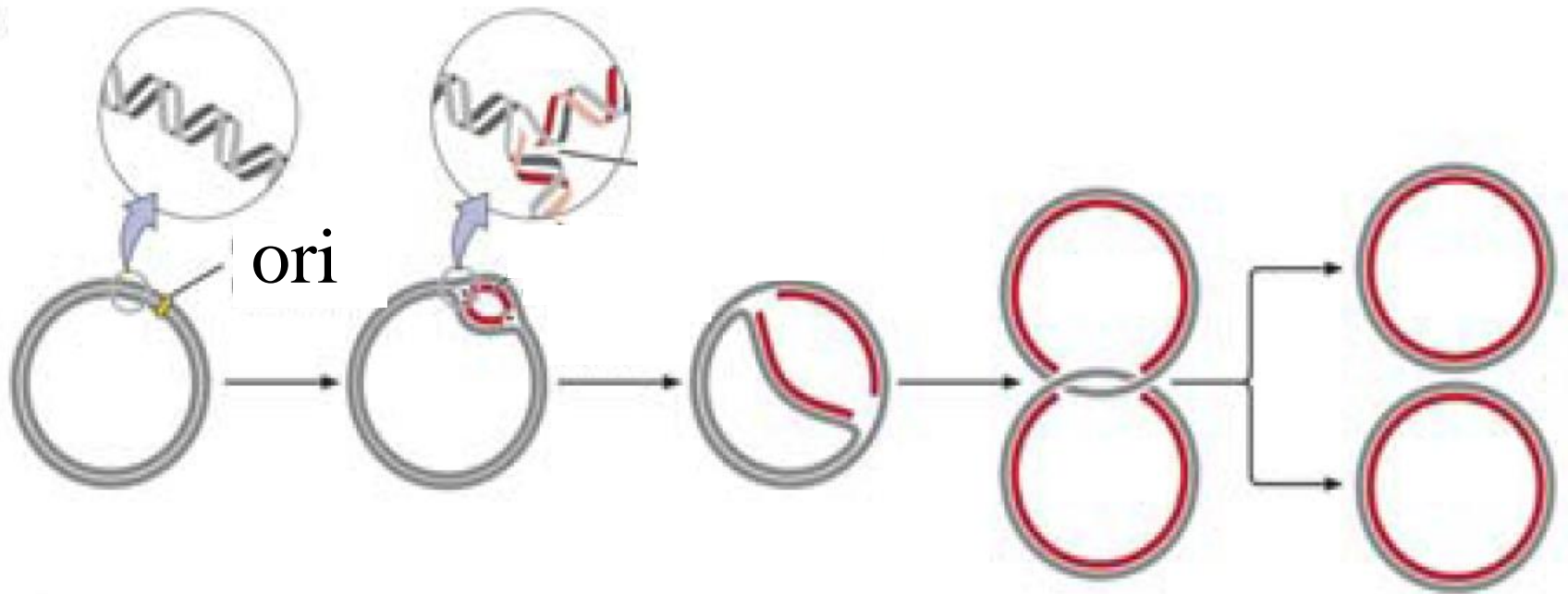


# Proofreading





# Theta replication $\Theta$



# Rolling circle

**Most bacteriophages e.g.  $\lambda$**

**DNA molecule – multiple length of genome**

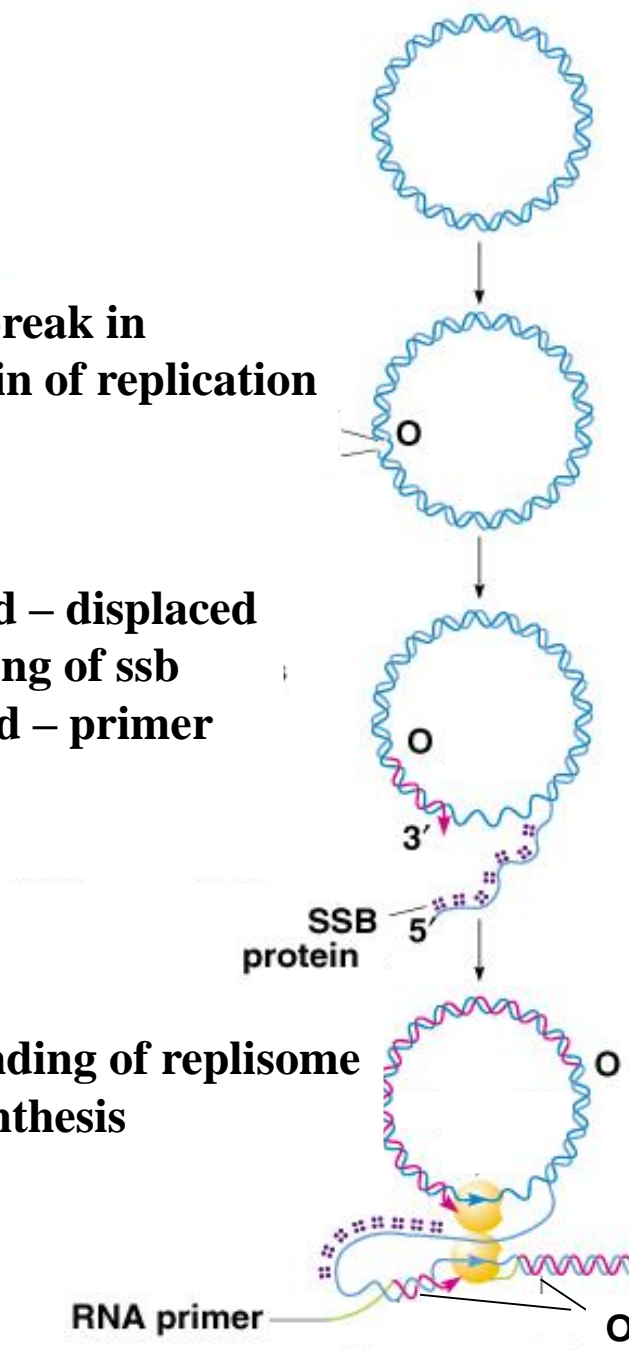
**During assembly – cleavage of DNA to individual chromosomes**

**SS break in origin of replication**

**5' end – displaced  
binding of ssb  
3' end – primer**

**Binding of replisome  
Synthesis**

**RNA primer** **Okazaki fragments**



# Eukaryotes

## Complex chromatine structure

**2 m DNA in nucelus Ø 6 µm**

~

**40 km of thin thread in tennis ball**

# Histone octamer – nucleosome core

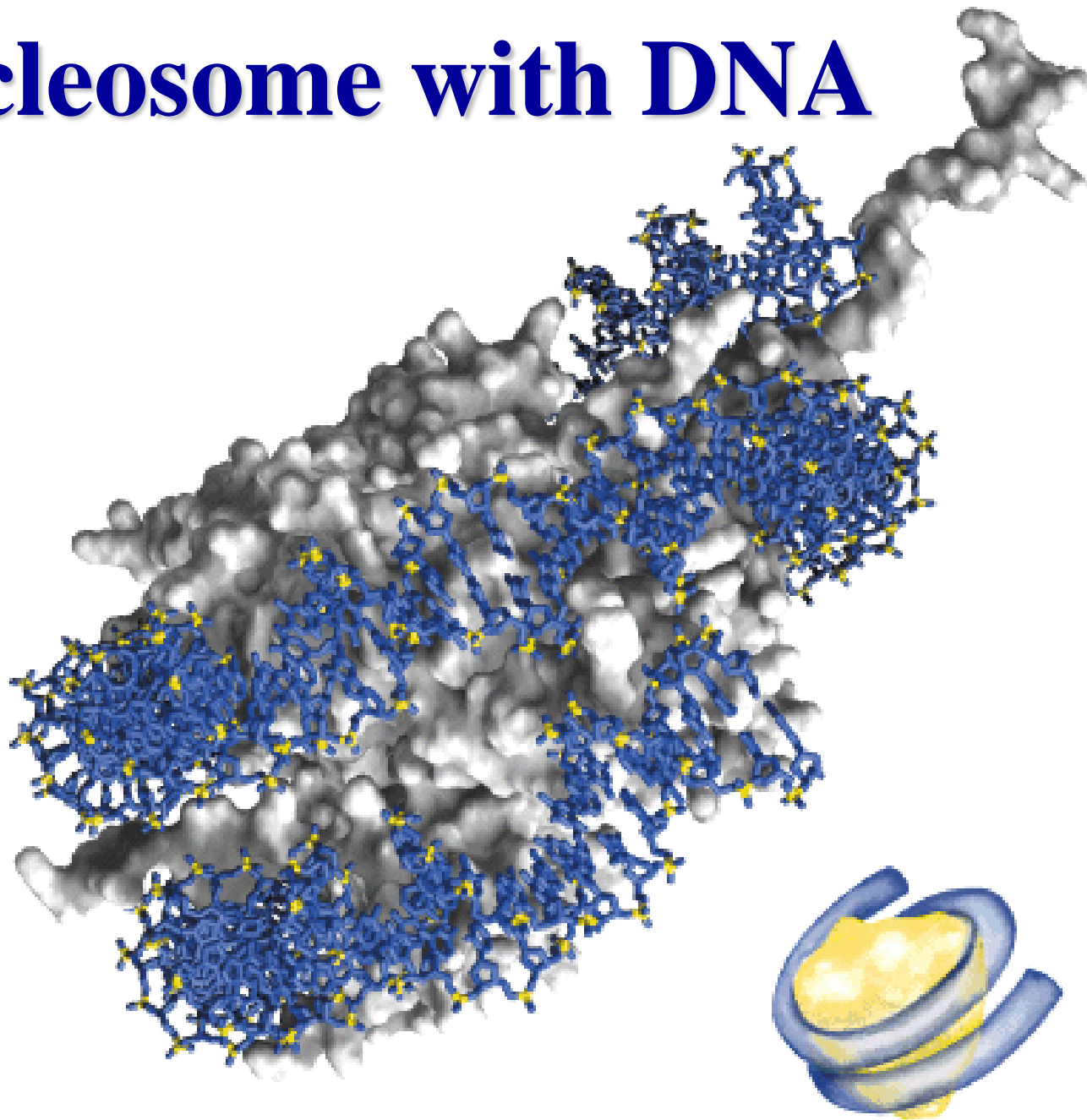
- 2 x H2A, H2B, H3, H4
- H3+H4  $\rightarrow$  H3(2)H4(2) tetramer
- two H2A-H2B heterodimers 1 above and 1 under tetramer

# Histone types

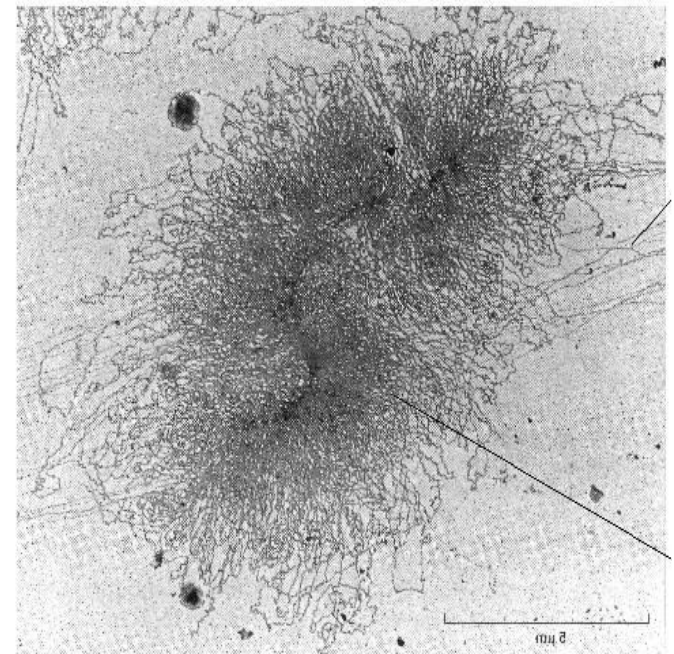
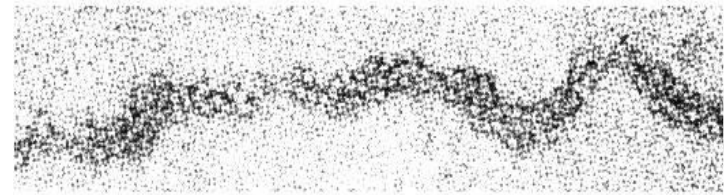
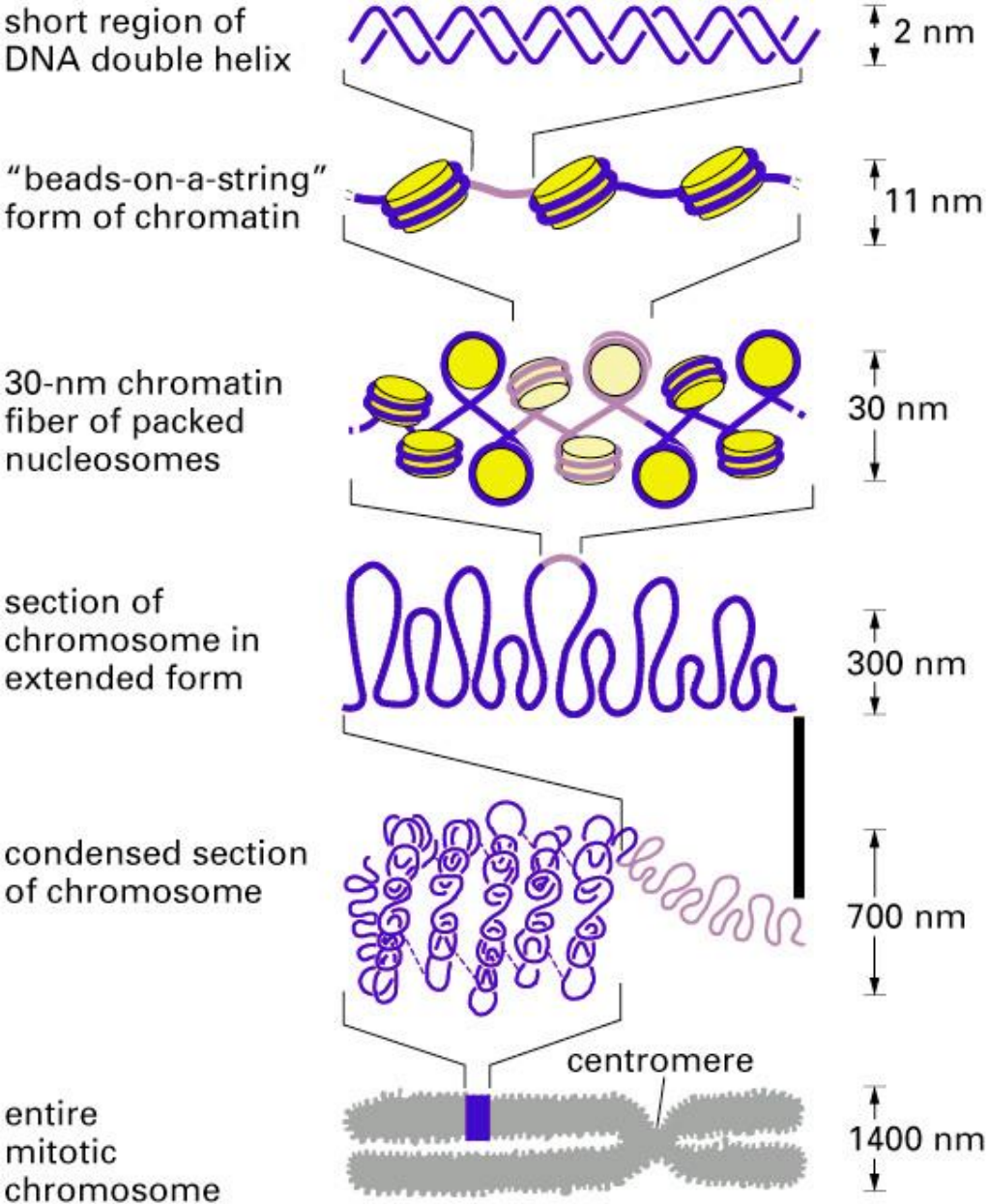
Histone	Molecular mass	Number of amino acids	Content of basic amino acids (%)	
			Lys	Arg
H1*	21,130	223	29.5	1.3
H2A*	13,960	129	10.9	9.3
H2B*	13,774	125	16.0	6.4
H3	15,273	135	9.6	13.3
H4	11,236	102	10.8	13.7

**Histones: Lys and Arg rich**

# Nucleosome with DNA

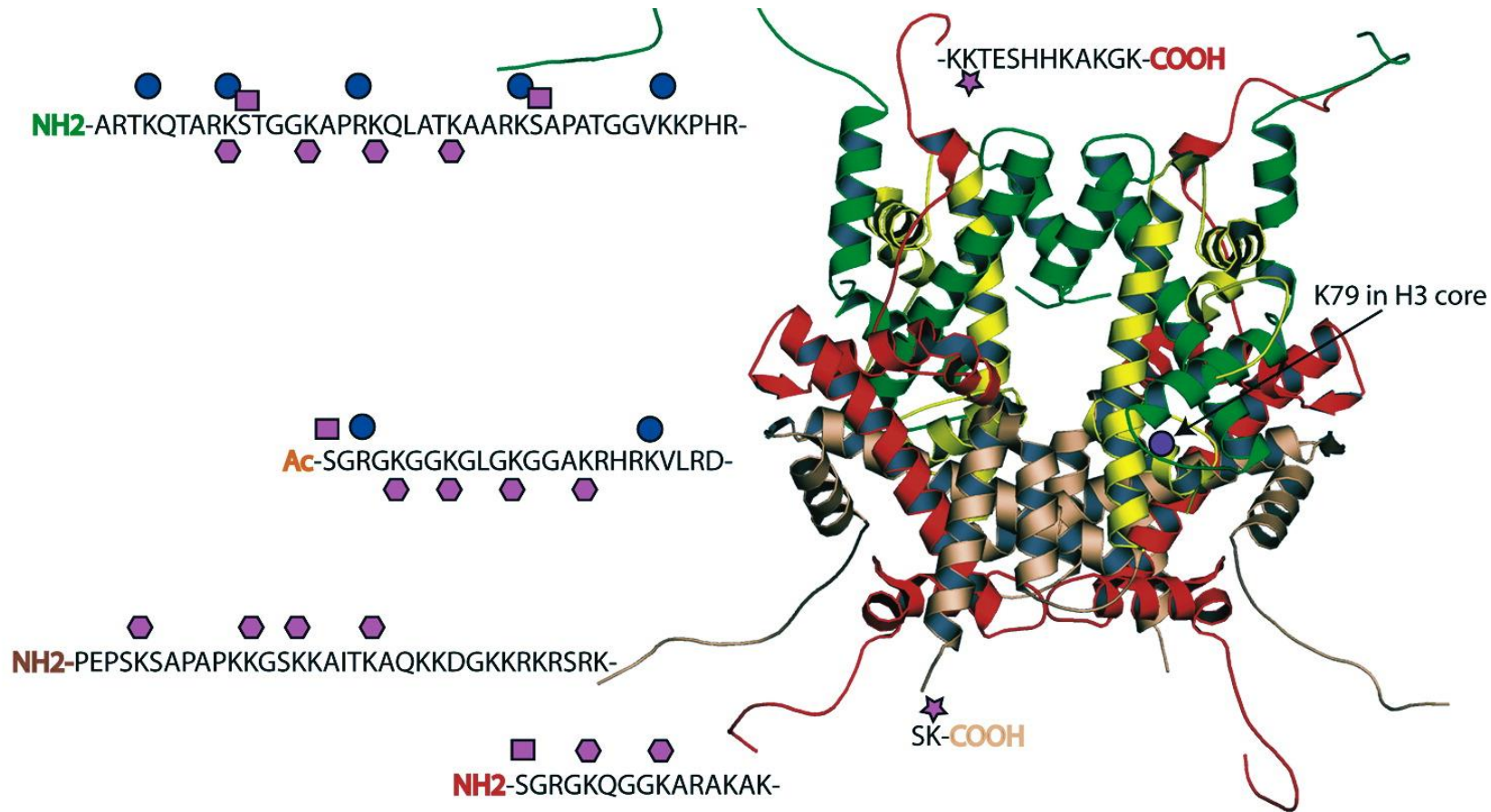




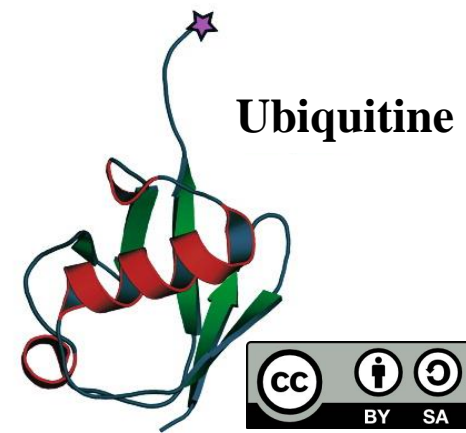
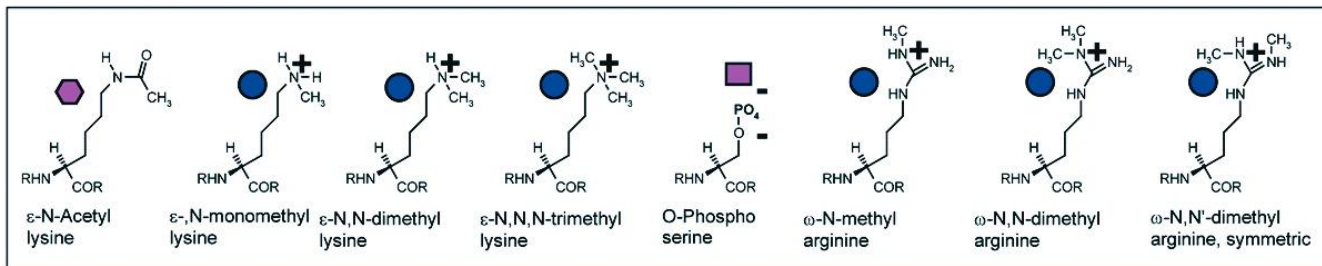


**NET RESULT: EACH DNA MOLECULE HAS BEEN PACKAGED INTO A MITOTIC CHROMOSOME THAT IS 10,000-FOLD SHORTER THAN ITS EXTENDED LENGTH**

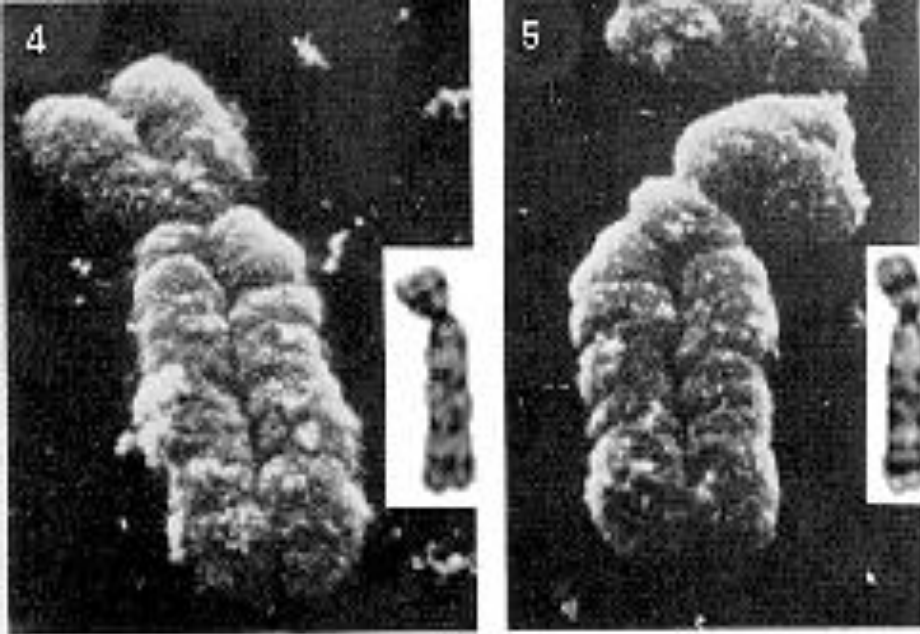
# Histone code – enhances capacity of genetic information



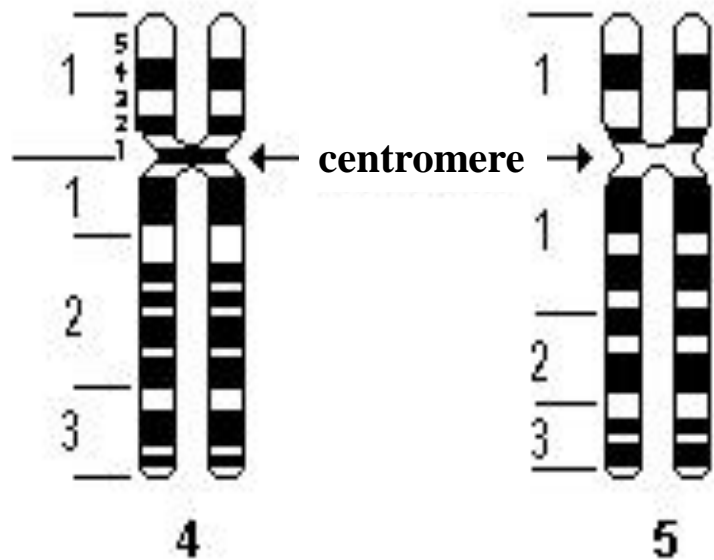
B







# Mitotic chromosome

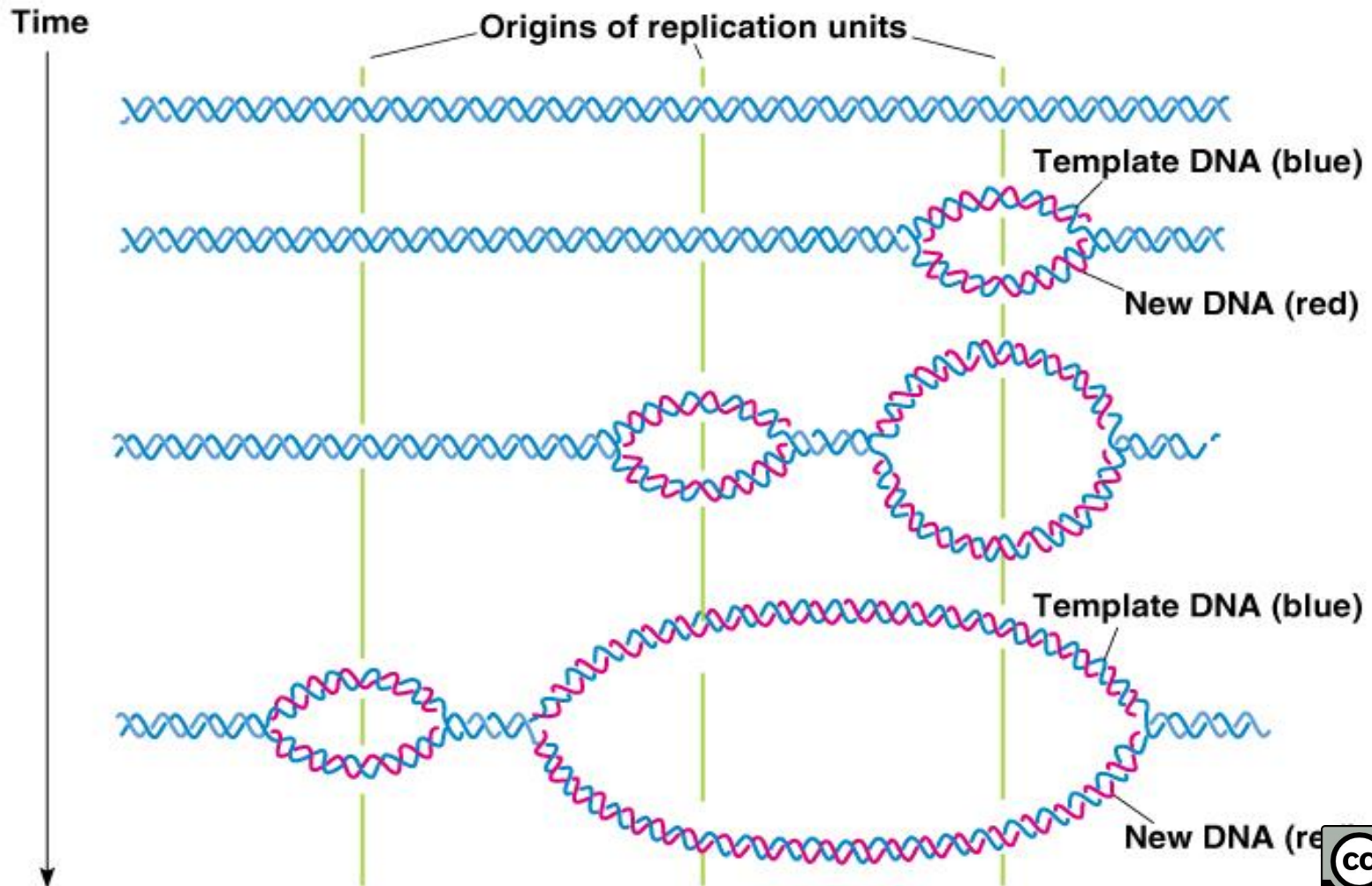


# Eukaryotic chromosomes - linear DNA double helixes

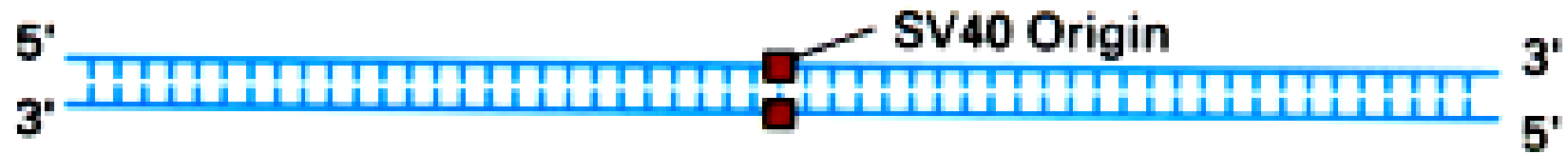
$\sim 10^8$  bp

Replication 2 kb/min.

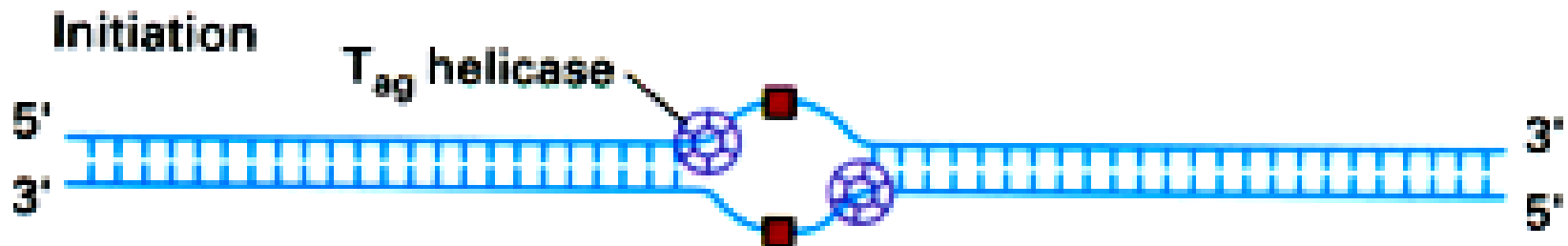
---> more origins



## SV40 DNA



1 Binding of  $T_{ag}$  at origin



2 Unwinding of duplex and binding of RPA to unwound template strands  
RPA replication protein A

Formation of prepriming complex



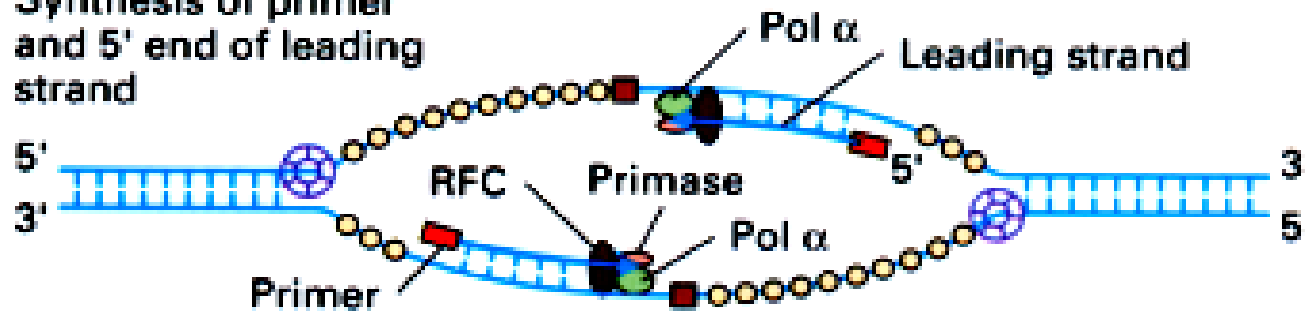


3

Binding of primase–Pol  $\alpha$   
 Primer synthesis by primase  
 Binding of RFC  
 DNA-chain synthesis by Pol  $\alpha$   
 stimulated by RFC

replication factor C

Synthesis of primer  
 and 5' end of leading  
 strand

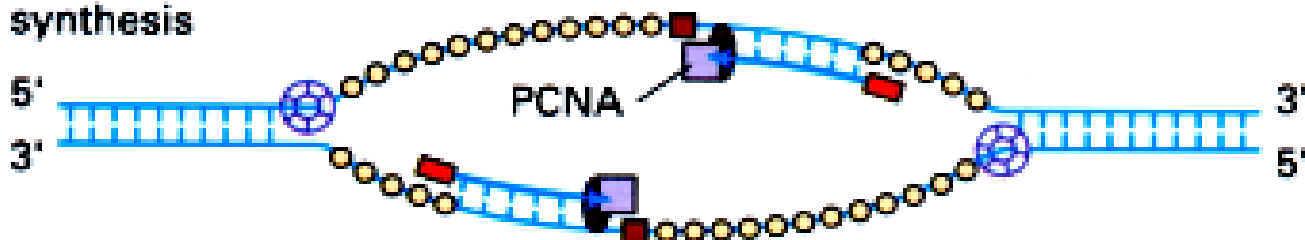


proliferating cell nuclear antigen

4

Binding of PCNA  
 displaces primase–Pol  $\alpha$

Interruption of  
 leading-strand  
 synthesis





Interruption of  
leading-strand  
synthesis

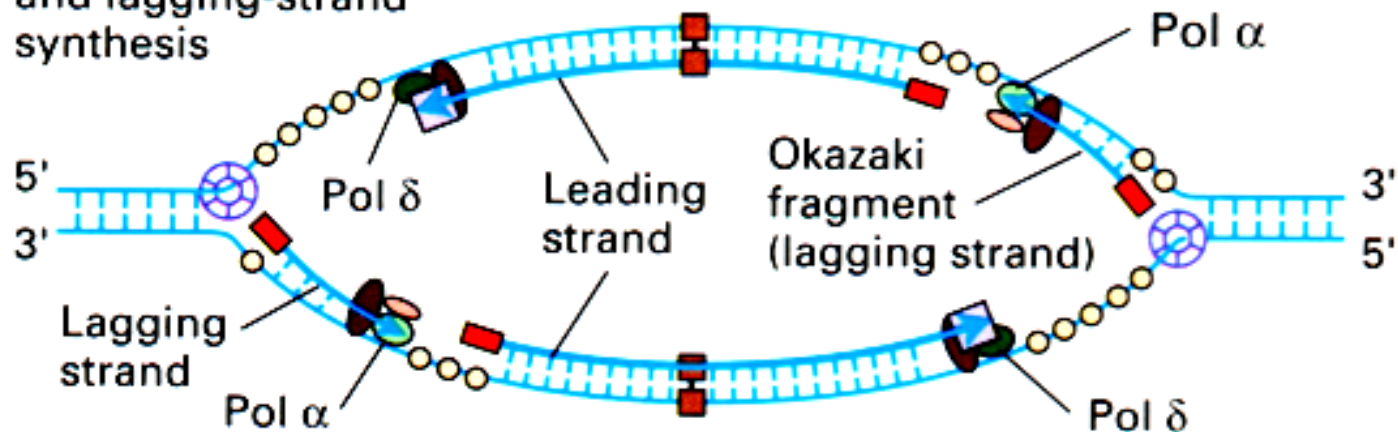


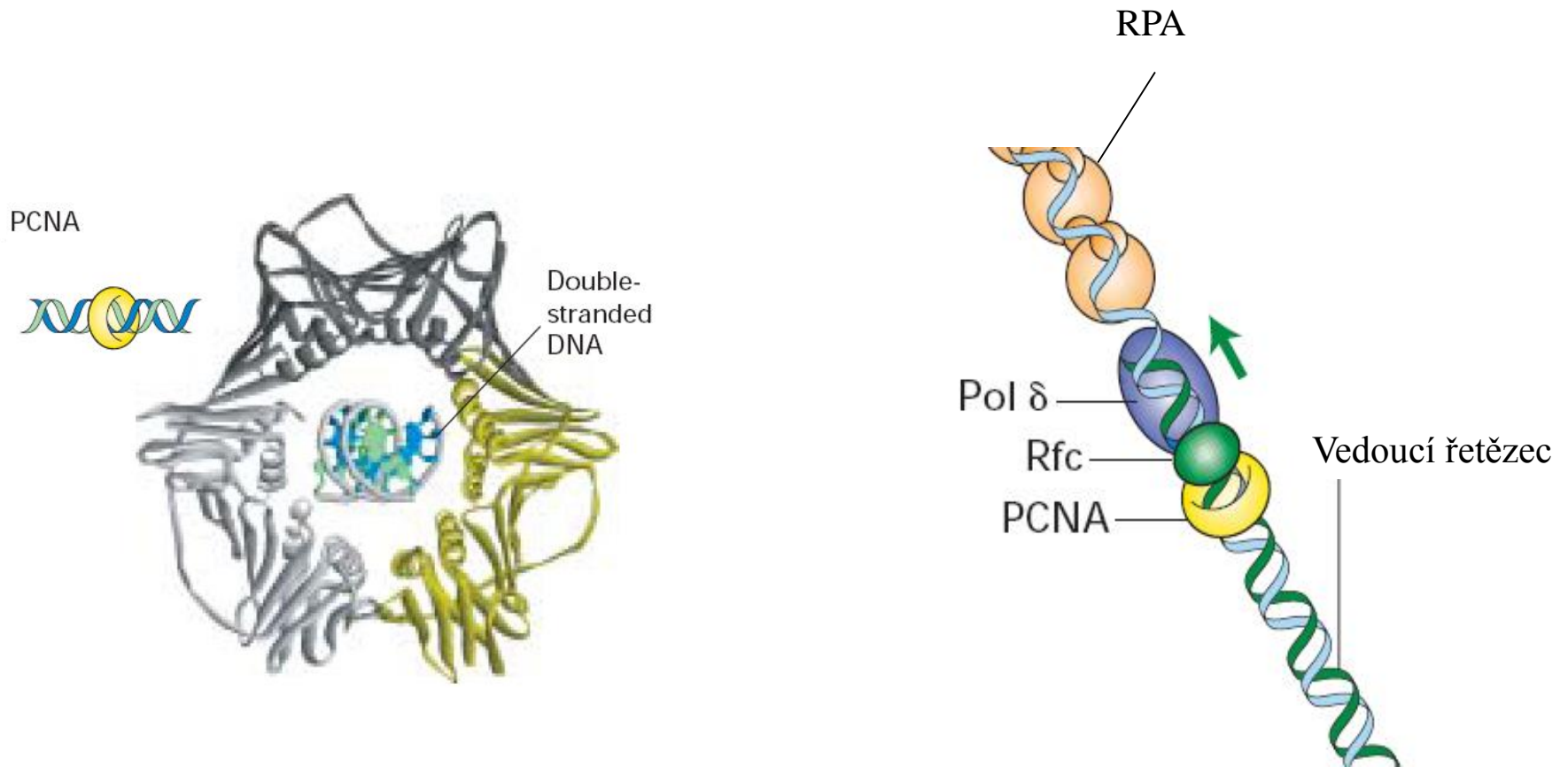
Binding of Pol  $\delta$  to PCNA-RFC  
complex and synthesis of  
leading strands

5

Binding of primase-Pol  $\alpha$  and  
RFC to lagging-strand template  
and synthesis of lagging  
strands discontinuously

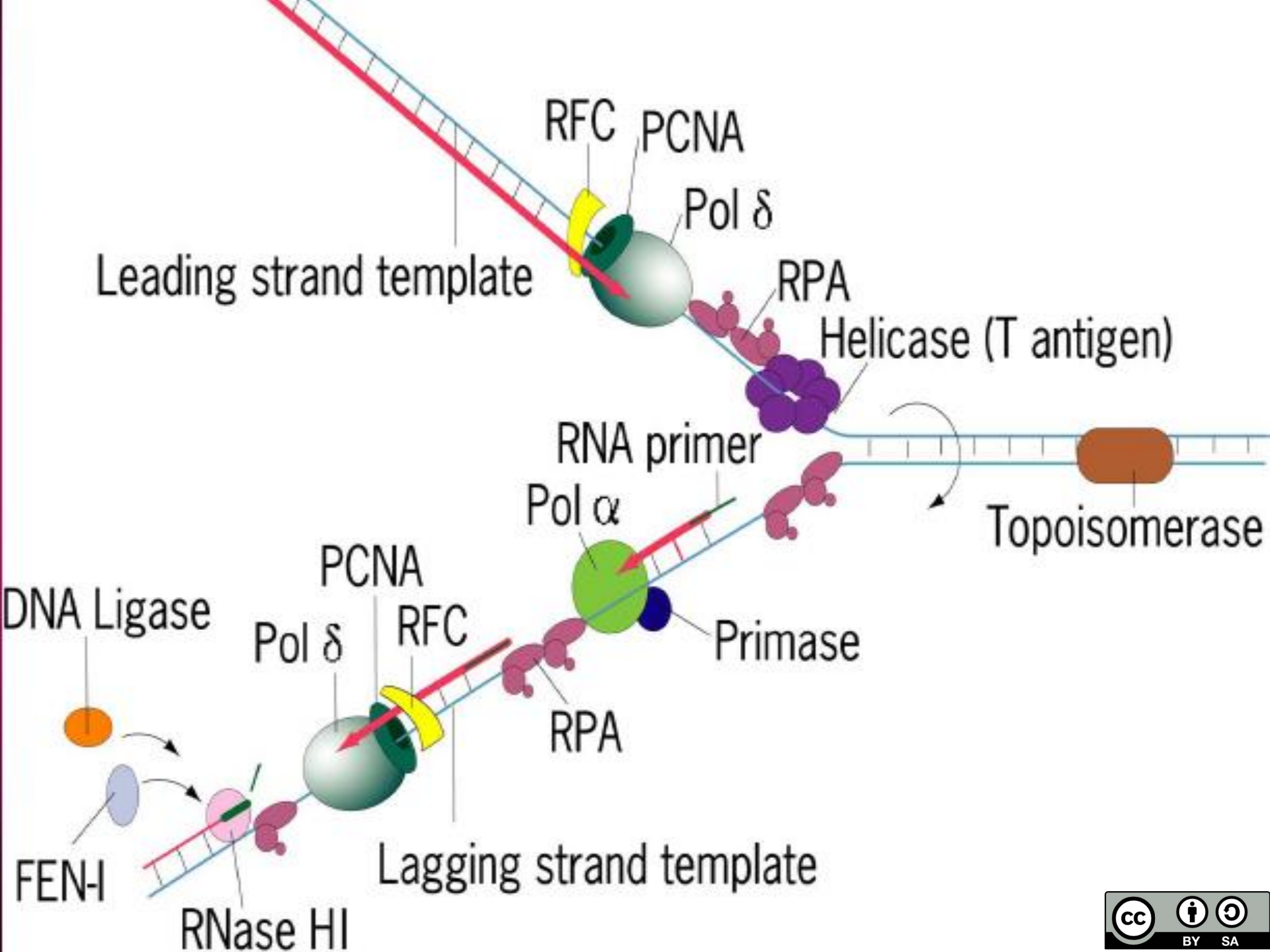
Continued leading-  
and lagging-strand  
synthesis





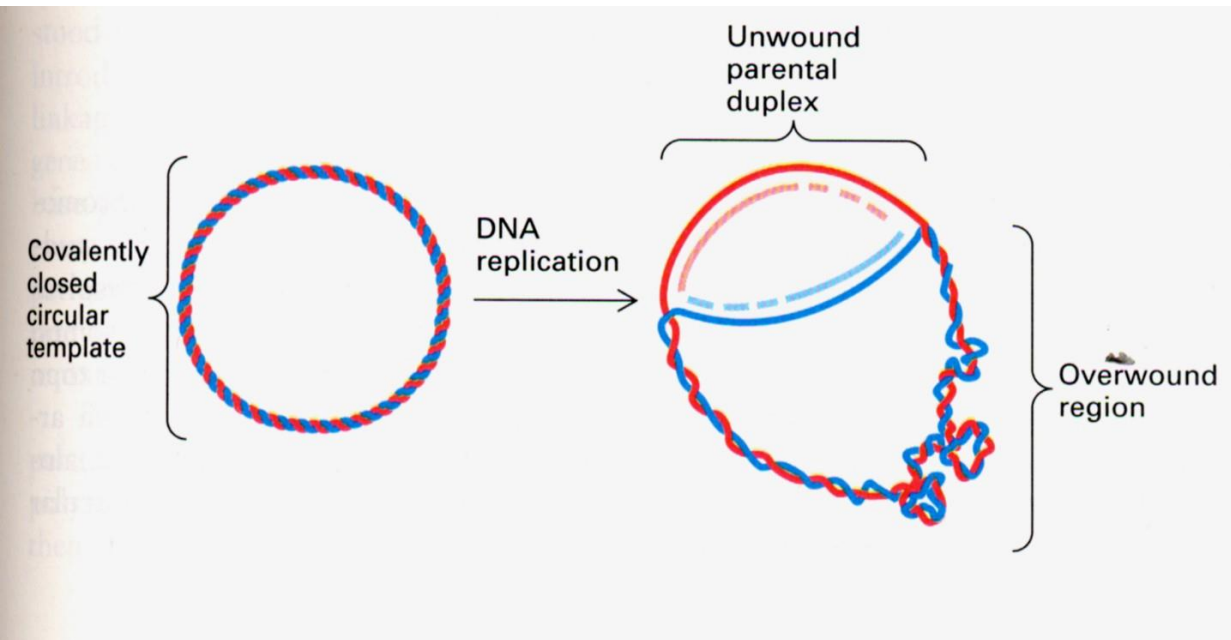
## PCNA – homotrimer

around DNA – prevents dissociation of the complex  
PCNA-RFC-Pol $\delta$  from the template

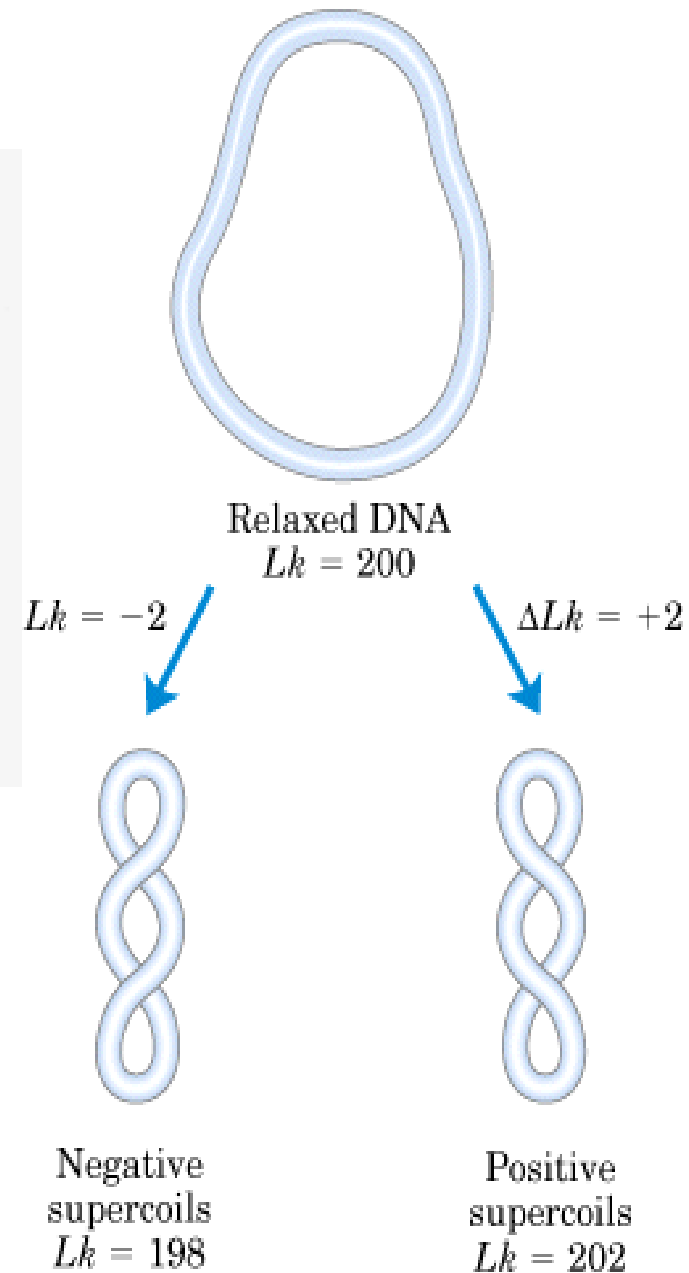


# positive supercoil – overwound DNA

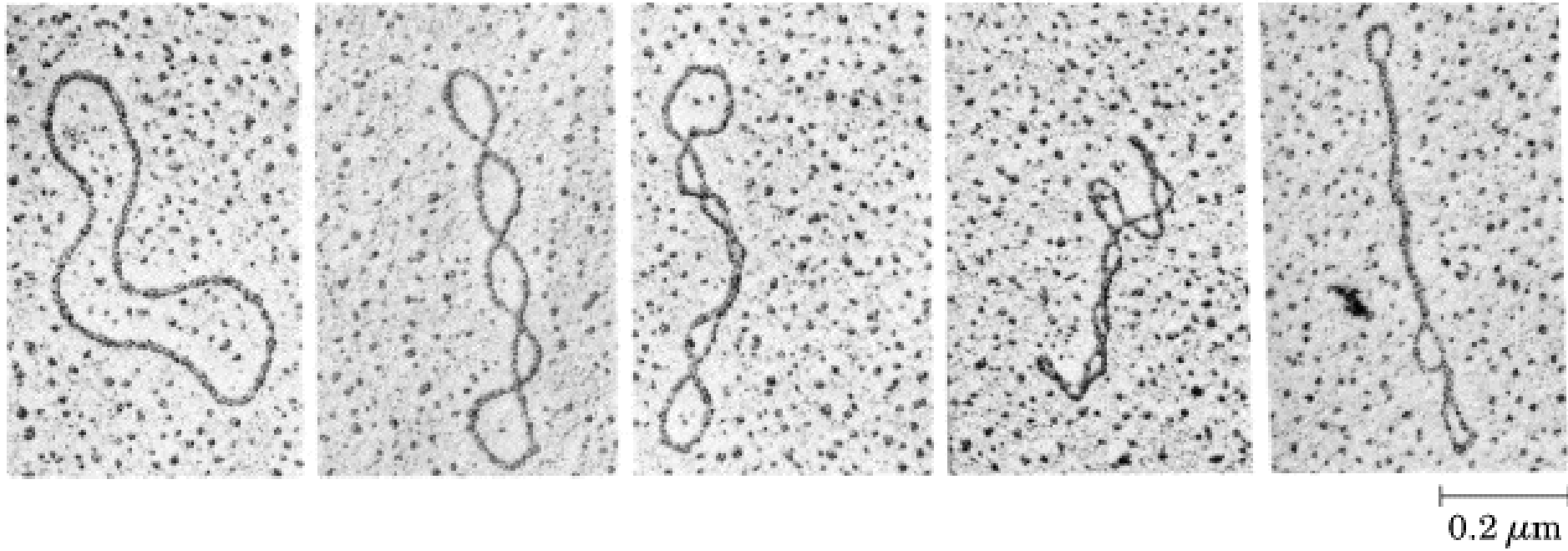
## - Product of DNA replication



negative supercoil from underwound DNA  
- present in nucleosomes



## Supercoil DNA



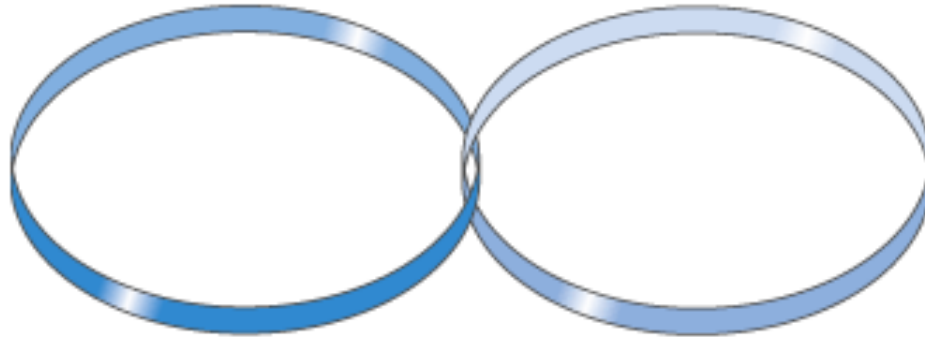
## Over or under-wound thread- supercoil



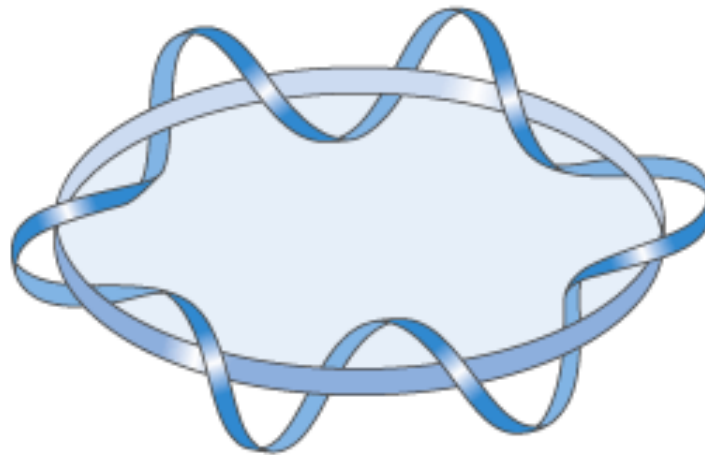


**L** - linking number

**#** multiplicity of threading of DNA chain around another chain



$$Lk = 1$$



$$Lk = 6$$

## Superhelix in front of the fork

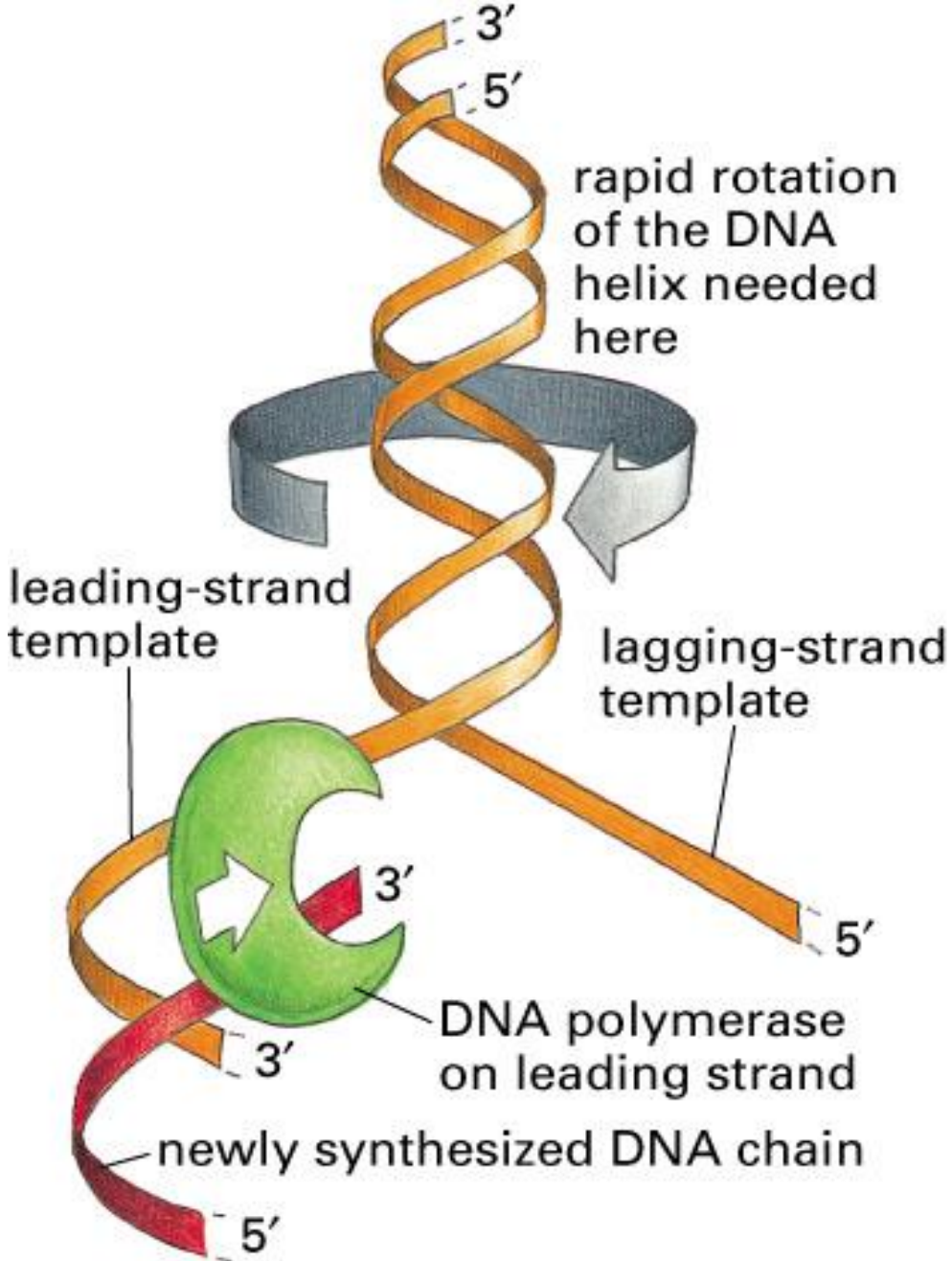
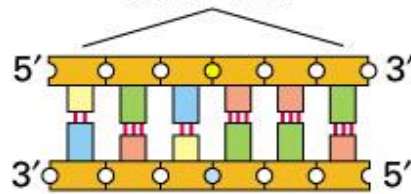
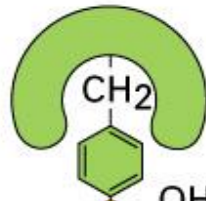
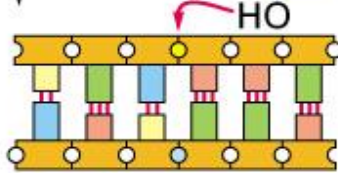


Figure 5-24. Molecular Biology of the Cell, 4th Edition.

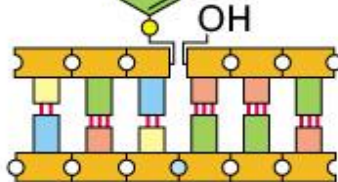
one end of the DNA  
double helix cannot  
rotate relative to the  
other end



type I DNA  
topoisomerase  
with tyrosine at  
the active site



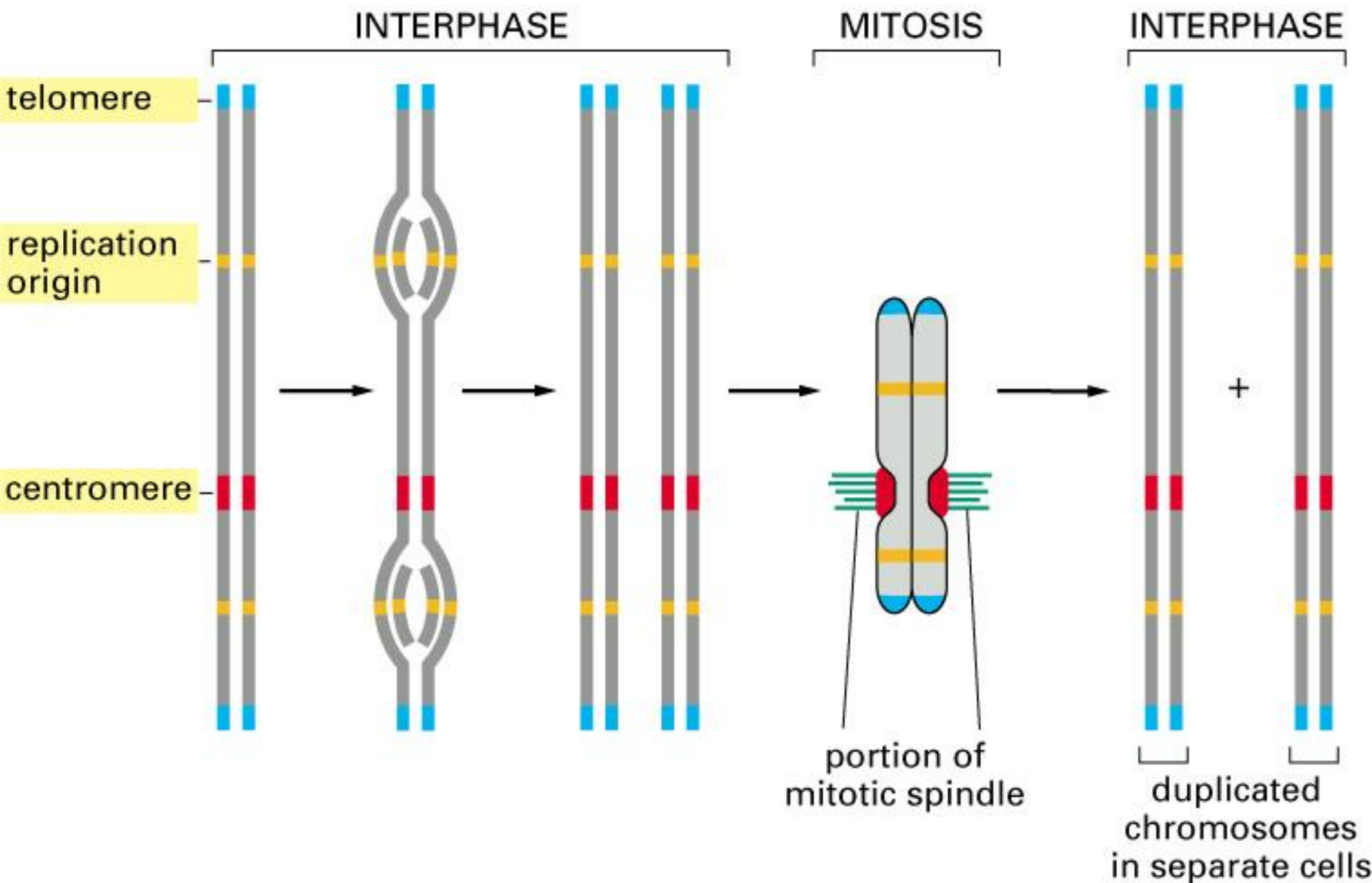
DNA topoisomerase covalently  
attaches to a DNA phosphate,  
thereby breaking a phosphodiester  
linkage in one DNA strand



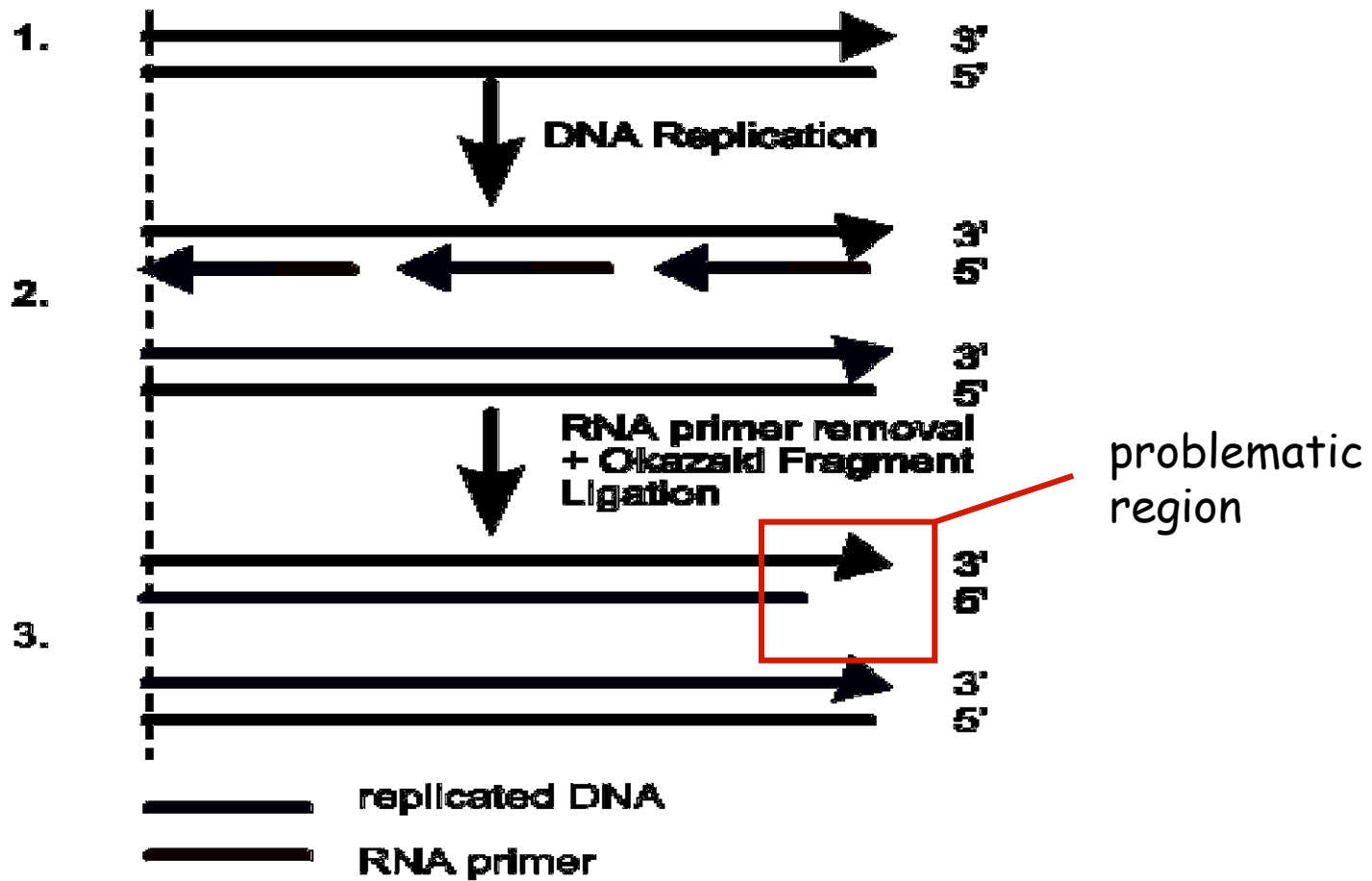
the two ends of the DNA double  
helix can now rotate relative to  
each other, relieving accumulated  
strain



Figure 5–25 part 1 of 2. Molecular Biology of the Cell, 4th Edition.



## Problem - ends of replicated chromosomes



# T4 or $\lambda$ - *cos* sites - circularization

## single-stranded RNA

tobacco mosaic virus  
bacteriophage R17  
poliovirus



Either + or - strand

## single-stranded DNA

parvovirus



## single-stranded circular DNA

M13

$\phi$ 174 bacteriophages



## double-stranded circular DNA

SV40  
polyoma viruses



## double-stranded RNA

reovirus



## double-stranded DNA

T4 bacteriophage  
herpes viruses



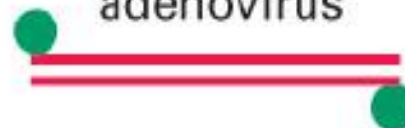
## double-stranded DNA with each end covalently sealed

poxvirus



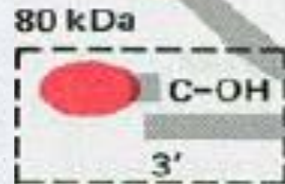
## double-stranded DNA with covalently linked terminal protein

adenovirus



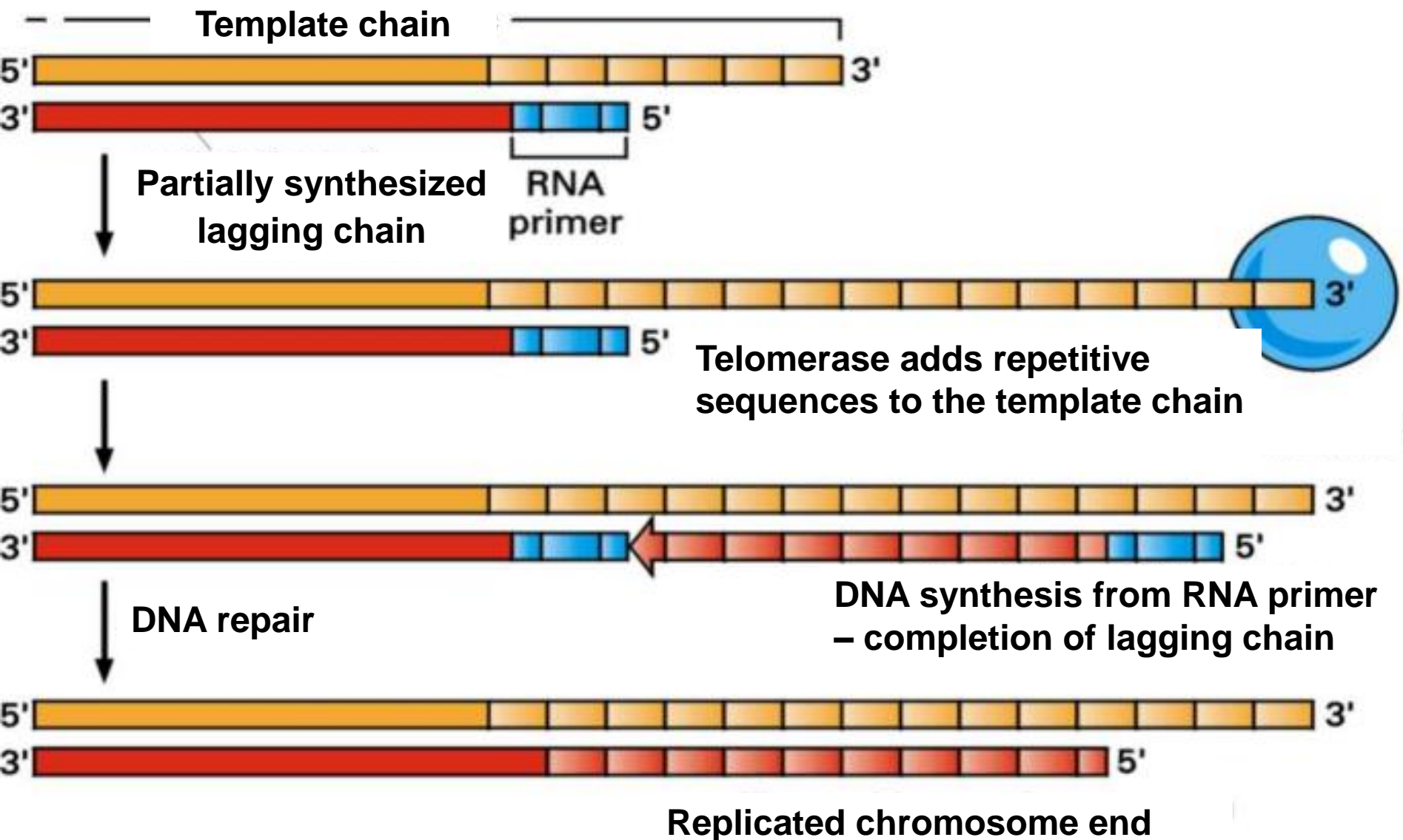


# Covalent binding to DNA



# Telomerase





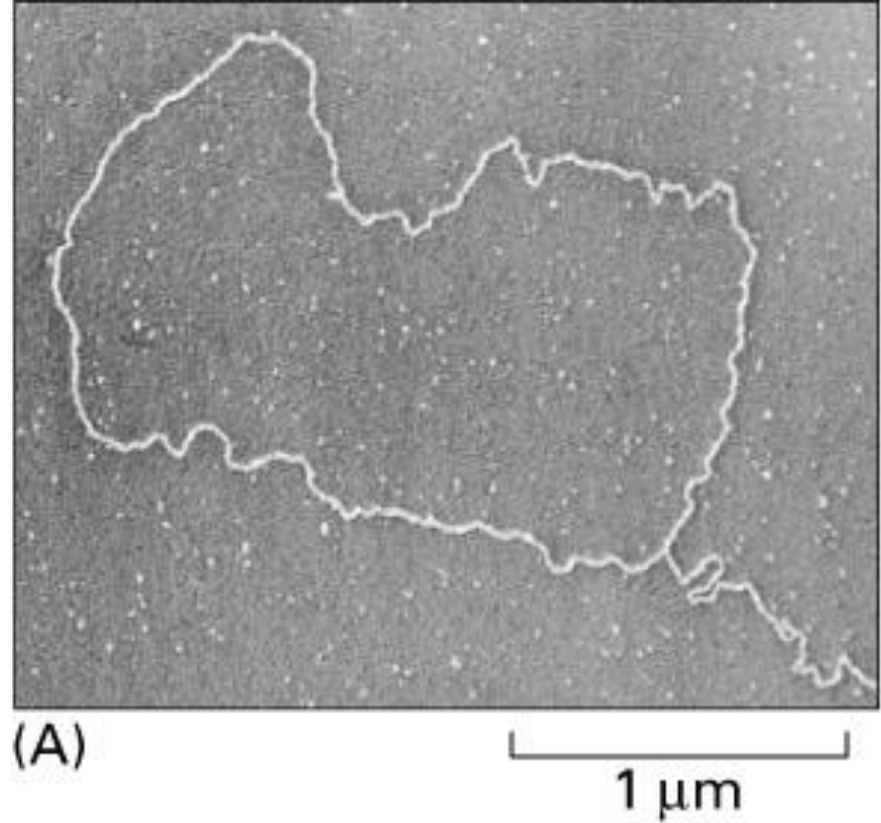




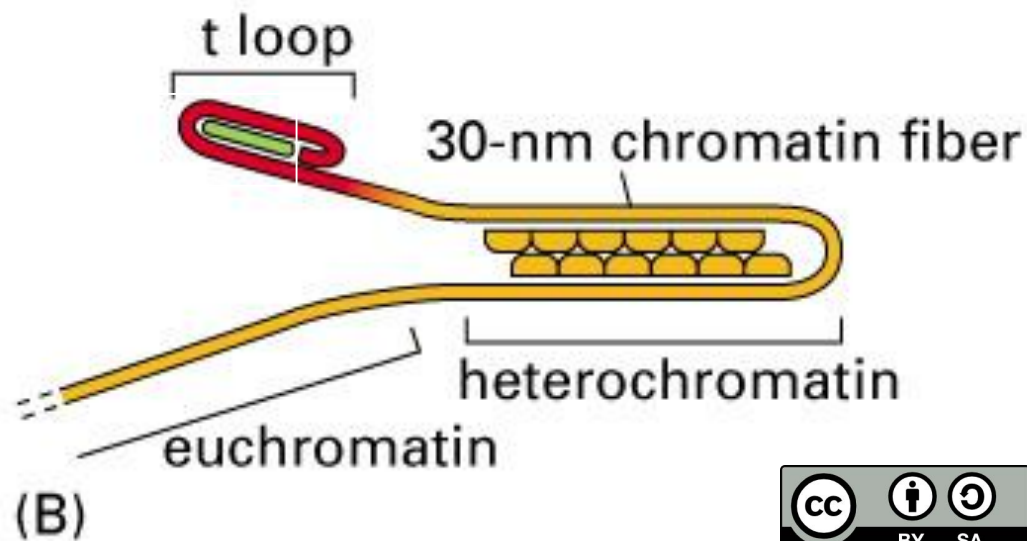


**8 and 9 years**  
**Premature aging – telomerase defect**

T-loop

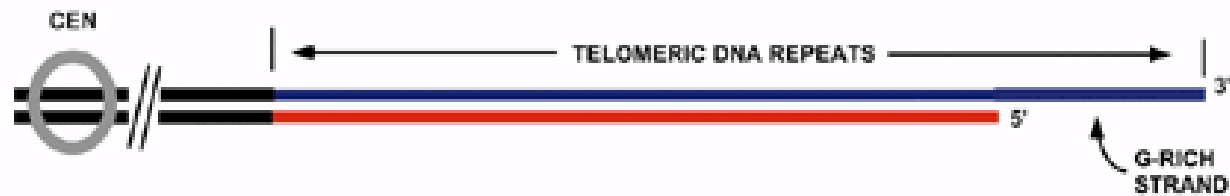


**Green - telomere-specific proteins**

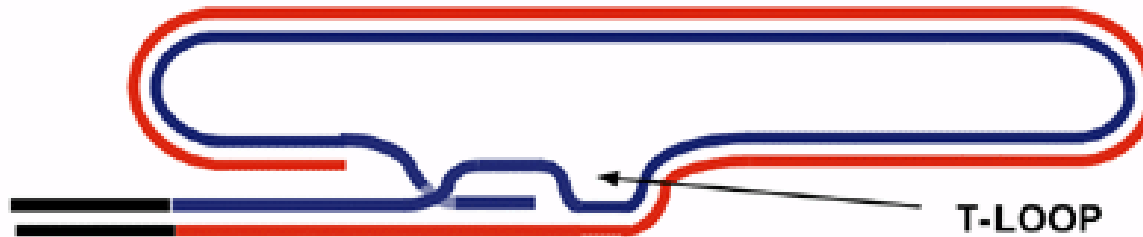




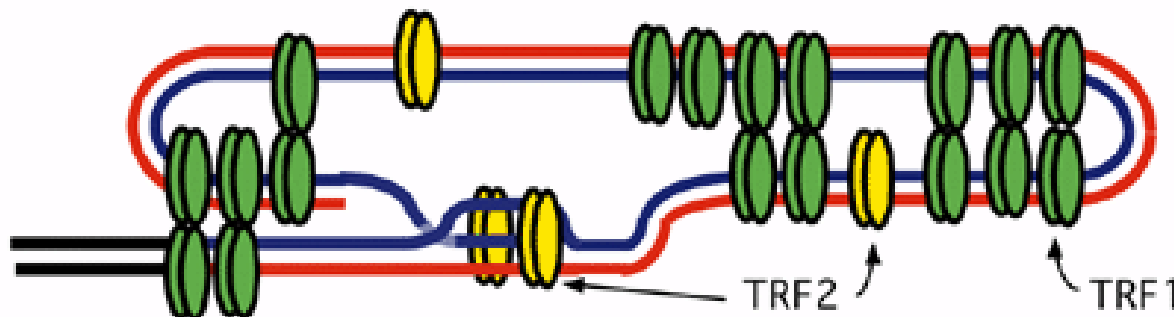
# Telomere ends -- T-loop



**Telomere DNA**

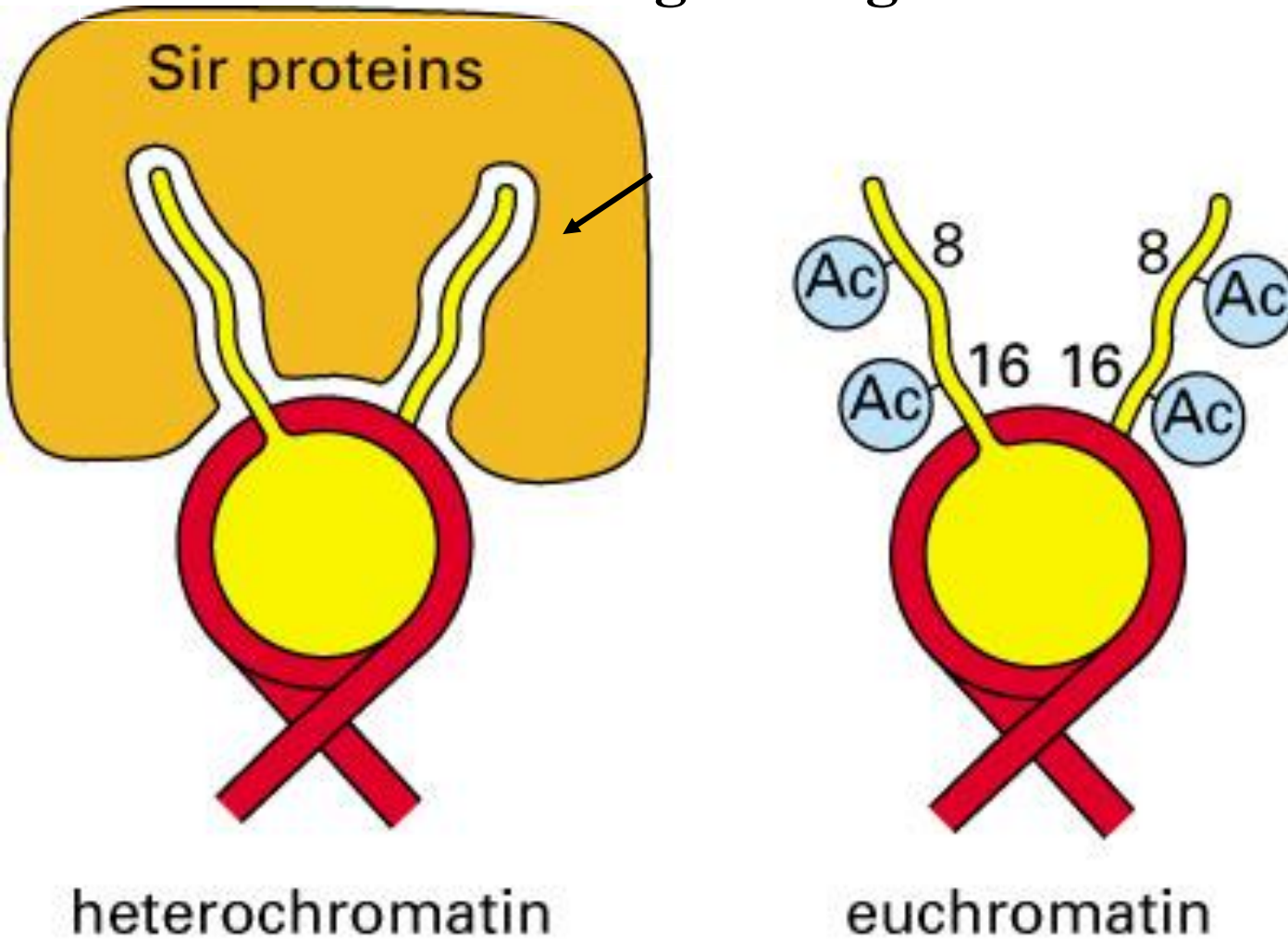


**T- loop**



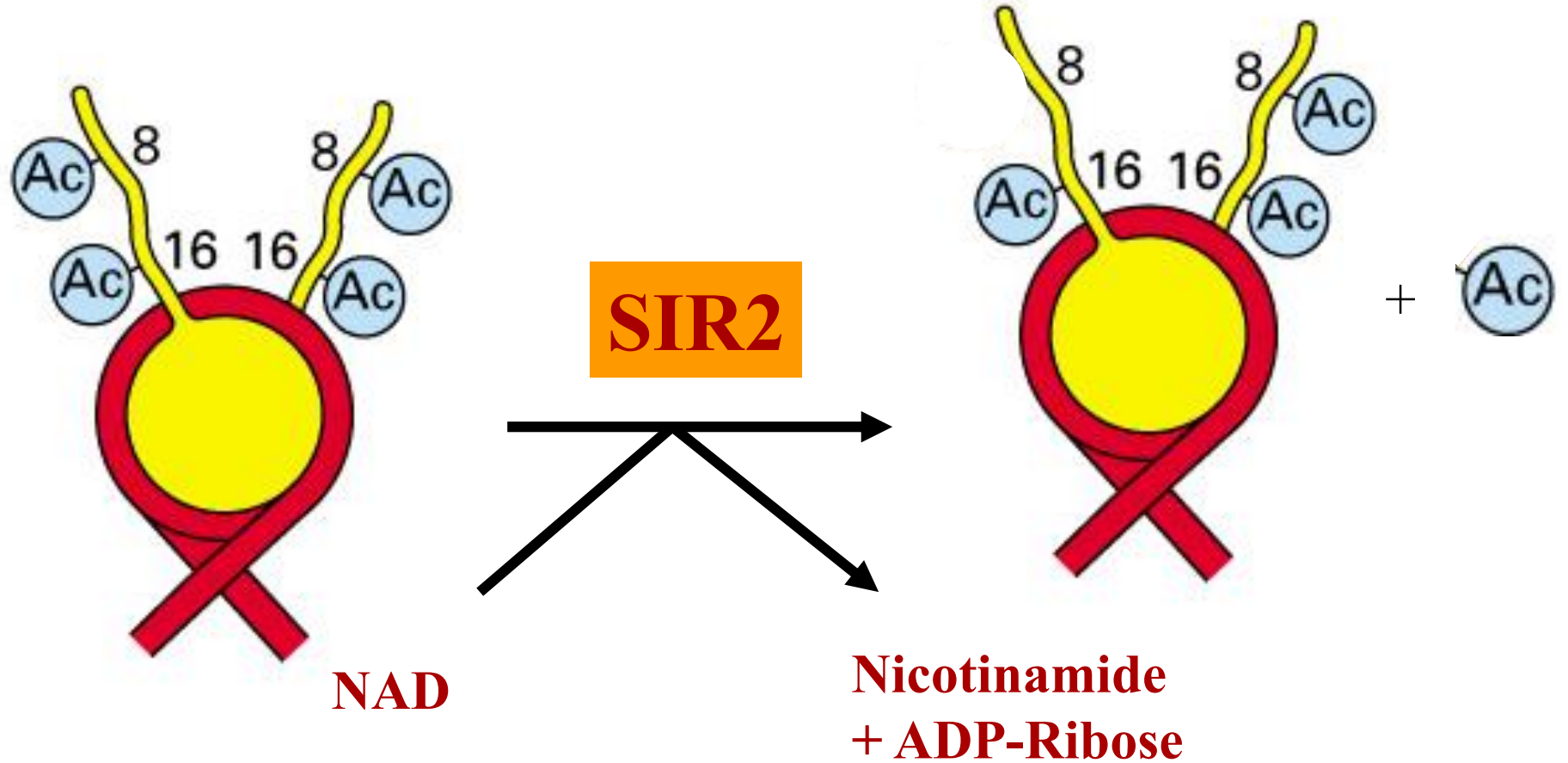
**TRF1 and TRF2  
proteins in  
T-loop**

# Histone modification – important for binding of factors organising heterochromatin



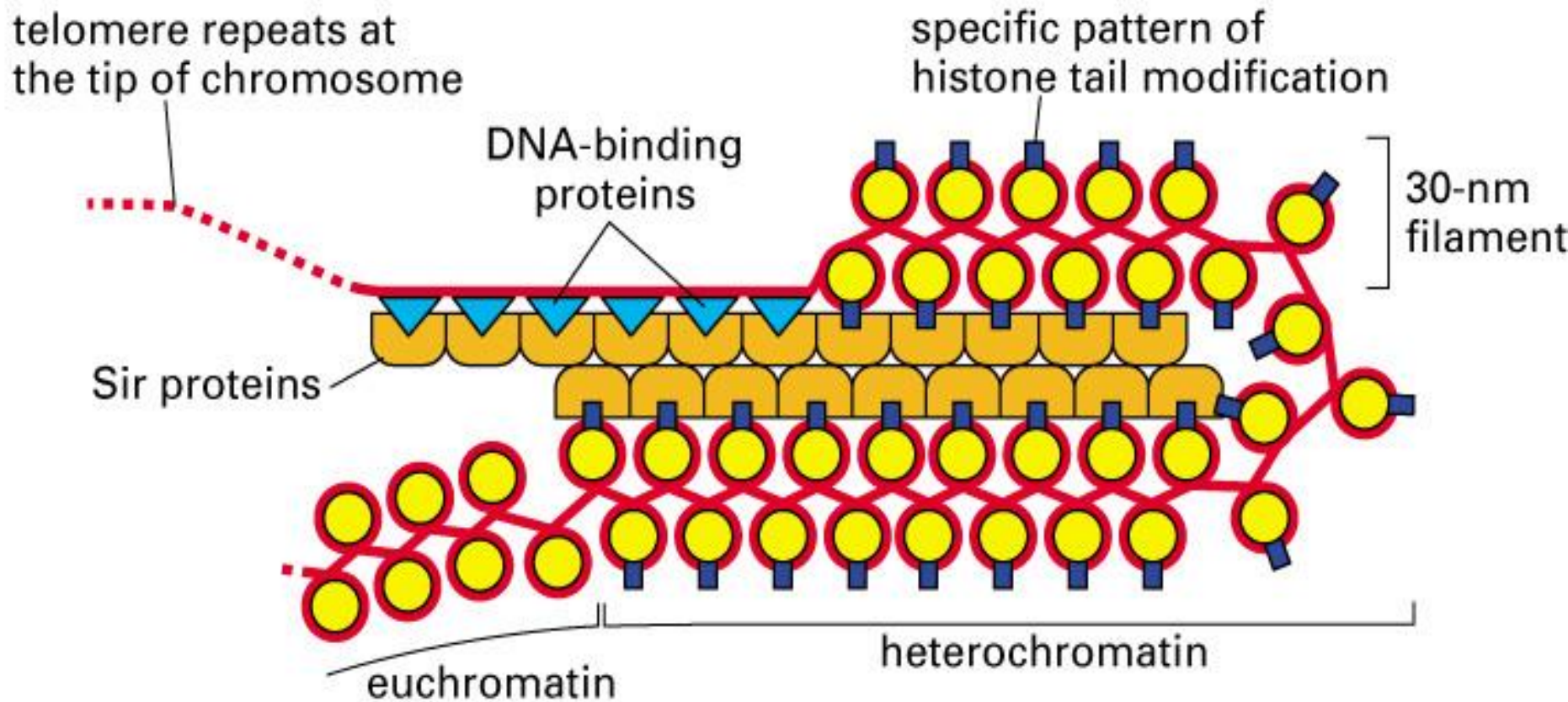
(A)

# Sir2 NAD-dependent histone deacetylase

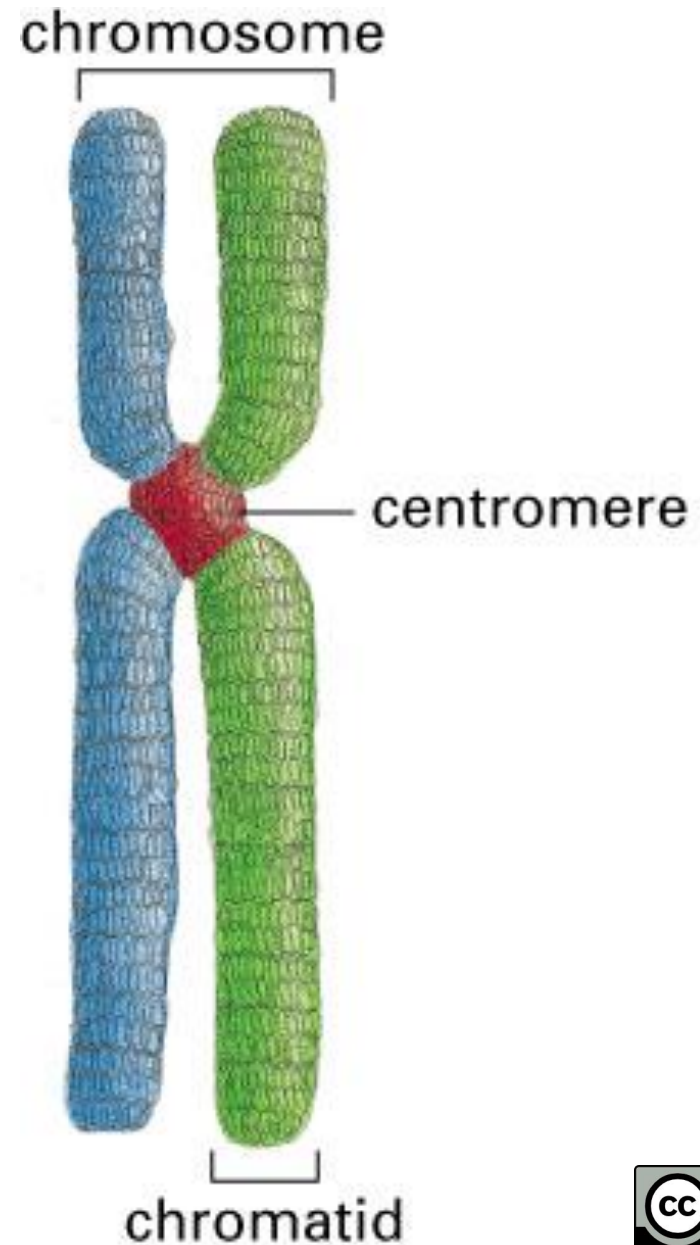


NAD consumed in reaction

# Heterochromatin proteins - condense 30nm thread in compact structure



# Centromere – chromatids segregation during mitosis

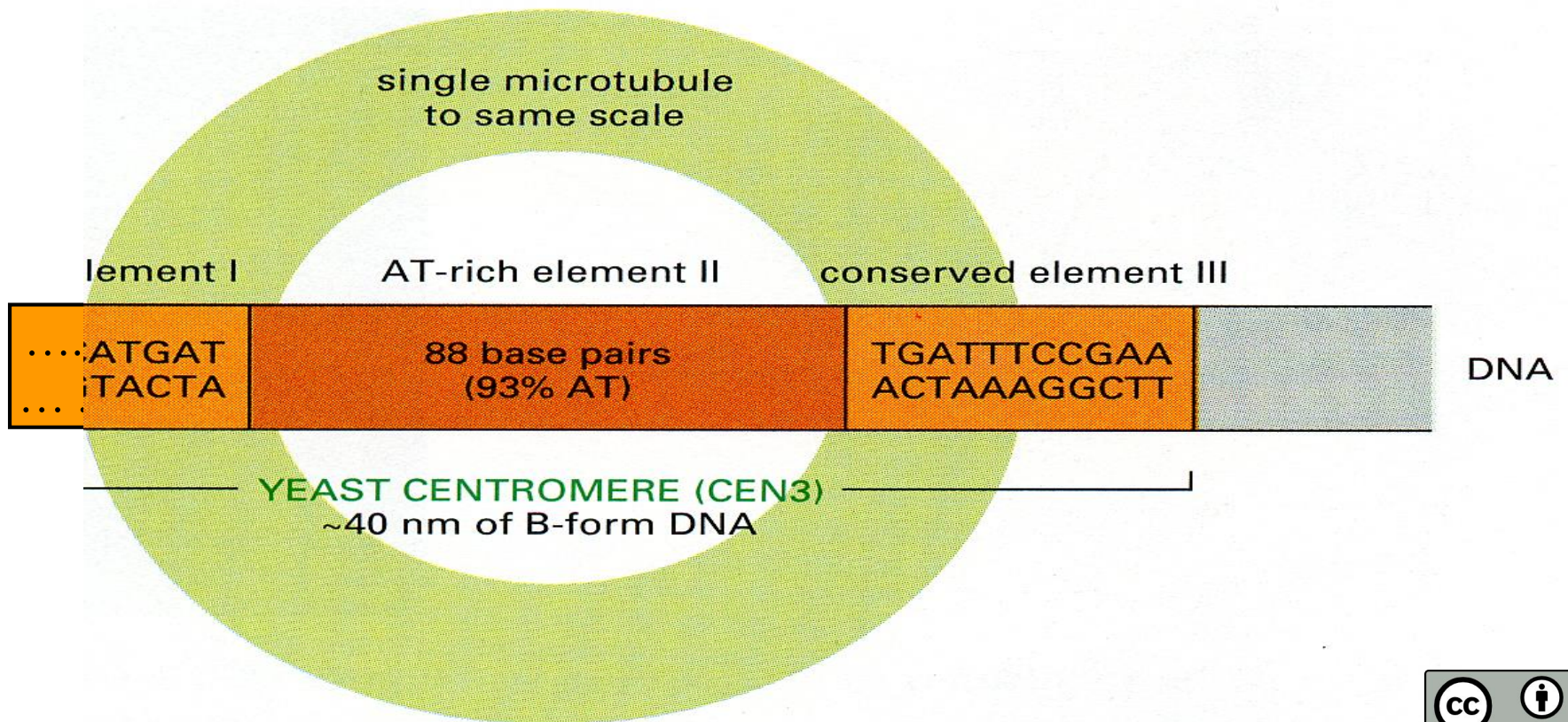




# Centromere

**Yeast**            **GTCACGTG**            **78 – 86 bp**    **CCGAAA**

**Drosophila** **GTCACATAG**    **264 bp**            **CCGAAA**





# Human centromere

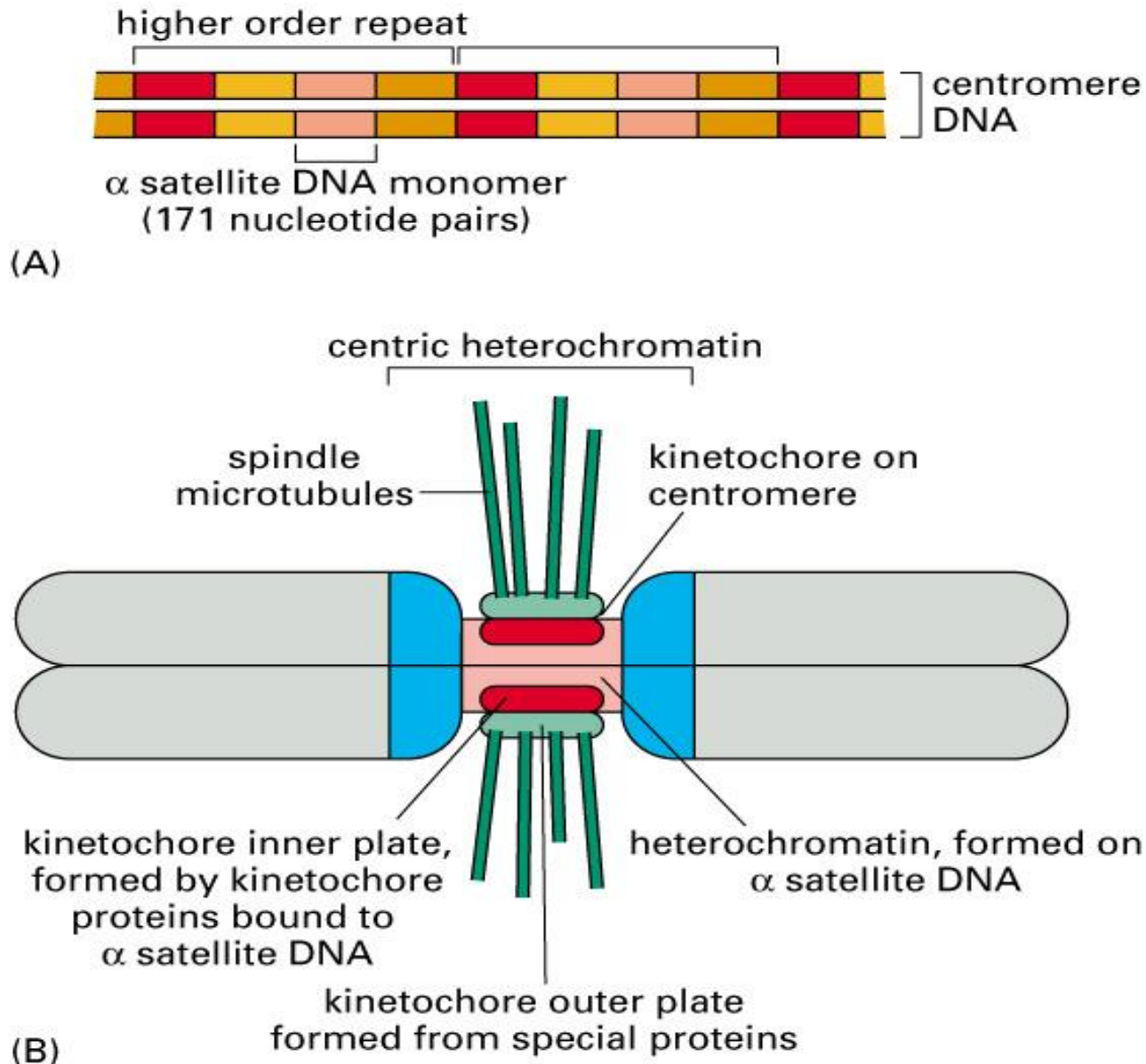


Figure 4–50. Molecular Biology of the Cell, 4th Edition.

# Heterochromatine of centromere - special nucleosomes

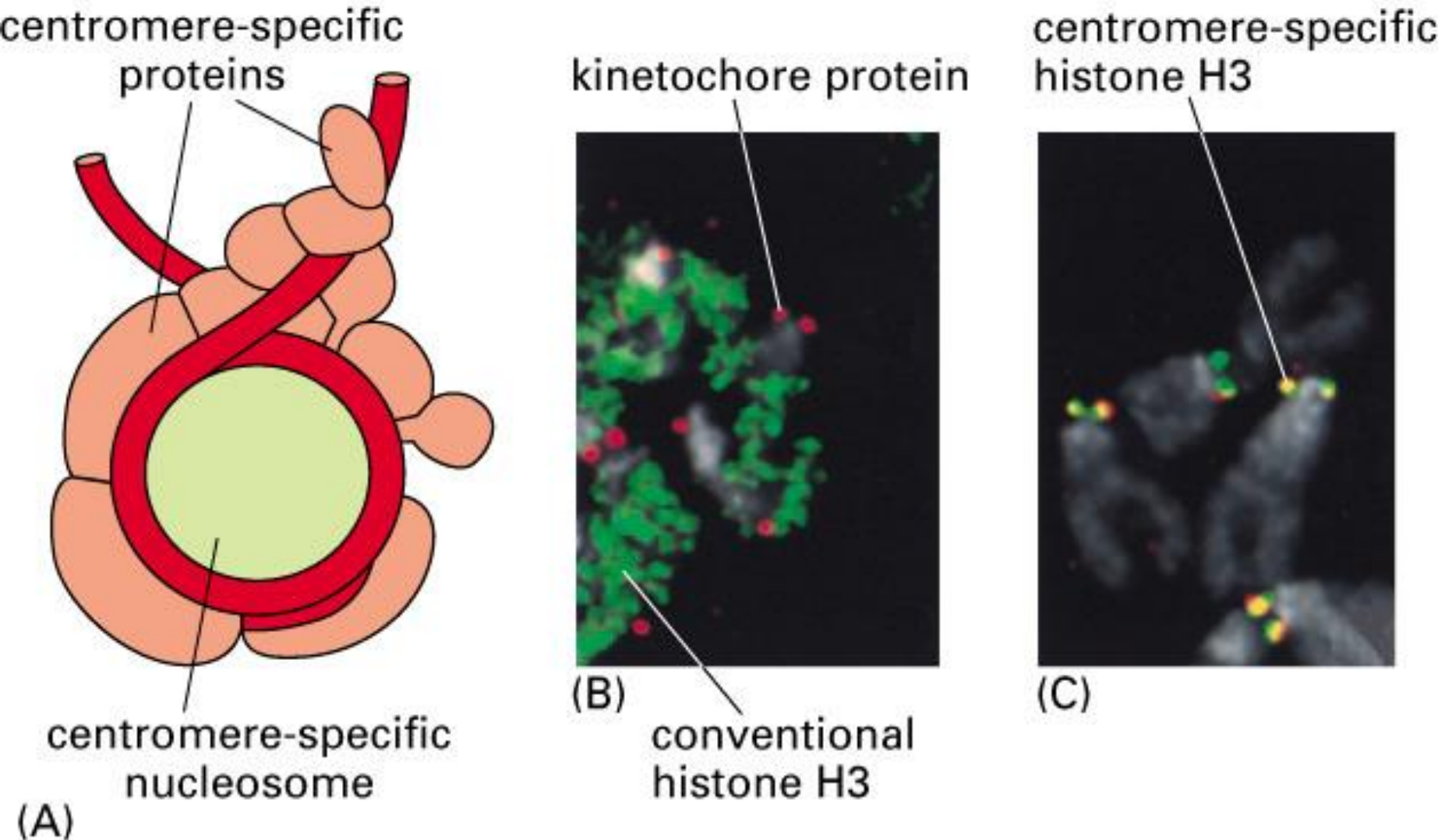


Figure 4-49. Molecular Biology of the Cell, 4th Edition.

Uveřejněné materiály jsou určeny studentům Vysoké školy chemicko-technologické v Praze

jako studijní materiál. Některá textová i obrazová data v nich obsažená jsou převzata

z veřejných zdrojů. V případě nedostatečných citací nebylo cílem autora/ů záměrně poškodit event. autora/y původního díla.

S eventuálními výhradami se prosím obraťte na autora/y konkrétního výukového materiálu, aby

bylo možné zjednat nápravu.

The published materials are intended for students of the University of Chemistry and Technology, Prague as a study material. Some text and image data contained therein are taken from public sources. In the case of insufficient quotations, the author's intention was not to intentionally infringe the possible author(s) rights to the original work.

If you have any reservations, please contact the author(s) of the specific teaching material in order to remedy the situation.

# Introduction to Molecular Genetics II



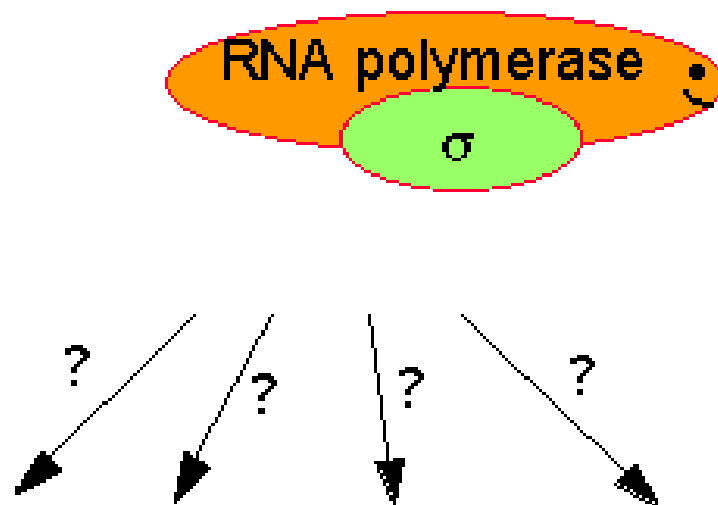
EUROPEAN UNION  
European Structural and Investing Funds  
Operational Programme Research,  
Development and Education



MINISTRY OF EDUCATION,  
YOUTH AND SPORTS

# TRANSCRIPTION

How does RNA polymerase know where to start transcription?



DNA

**DNA** —→ **RNA** —→ **Protein**

## **Regulation**

Transcription

Post-transcriptional regulation

RNA transport

Degradation

Translation

Transcription  
Constitutive  
Induction x Repression

Activity regulation  
Feedback inhibition – quick response

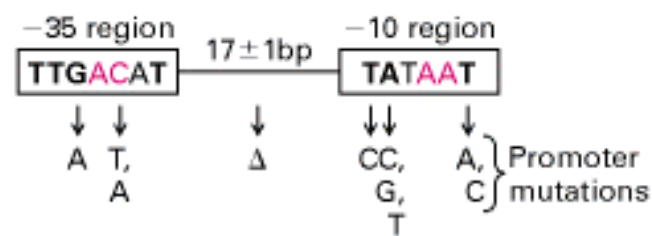




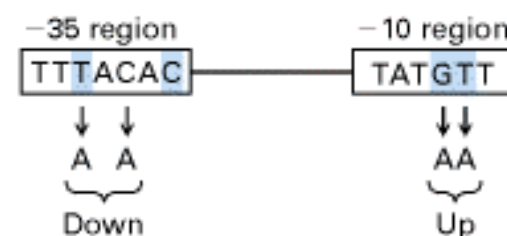
## Strong *E. coli* promoters

tyr tRNA	TCTCAACGTAACAC	TTTACAG	GCGGCG	•	CGTCATTTGA	TATGAT	GC	•	GCCCCG	GCTTCCCGATAAGGG
rrn D1	GATCAAAAAAATAC	TTGTG	CAAAAAA	•	TTGGGATCCC	TATAAT	GCGCCTCCG	TTGAGACGACAACG		
rrn X1	ATGCATTTTTTCCGC	TTGTCT	TCTTGA	•	GCCGACTCCC	TATAAT	GCGCCTCCAT	CGACACGCGCGGAT		
rrn (DXE) <sub>2</sub>	CCTGAAATTCAGGG	TTGACT	CTGAAA	•	GAGGAAAGCG	TAATAT	AC	•	GCCACCT	CGCGACAGTGAGC
rrn E1	CTGCAATTTTTCTA	TTGCGG	CCTGCG	•	GAGAACTCCC	TATAAT	GCGCCTCCAT	CGACACGCGCGGAT		
rrn A1	TTTTAAATTTTCTC	TTGTCA	GGCCGG	•	AATAACTCCC	TATAAT	GCGCCACCA	CTGACACGGAACAA		
rrn A2	GCAAAAATAAAATGC	TTGACT	CTGTAG	•	CGGGAAGGCG	TATTAT	GC	•	ACACCC	CGCGCCGCTGAGAA
λ P <sub>R</sub>	TAACACCGTGCGTG	TTGACT	ATTTTA	•	CCTCTGGCGGT	GATAAT	GG	•	TTGCAT	TGTACTAAGGAGGT
λ P <sub>L</sub>	TATCTCTGGCGGTG	TTGACAT	AAATA	•	CCACTGGCGGT	GATACT	GA	•	GCACAT	CAGCAGGACGCAC
T7 A3	GTGAAACAAAACGG	TTGACA	ACATGA	•	AGTAAACACGG	TACGAT	GT	•	ACCACAT	GAAACGACAGTGA
T7 A1	TATCAAAAAGAGTA	TTGACTT	TAAAGT	•	CTAACCTATAGG	TACTTA	•	CAGCCAT	CGAGAGGGACACG	
T7 A2	ACGAAAAACAGGTA	TTGACA	ACATGA	AGTAACATGCAG	TAAGAT	AC	•	AAATCG	CTAGGTAACACTAG	
fd VIII	GATACAAATCTCCG	TTGTACT	TTTGTT	•	TCGCGCTTGG	TATAAT	CG	•	CTGGGG	GTCAAAGATGAGTG
		-35				-10			+1	

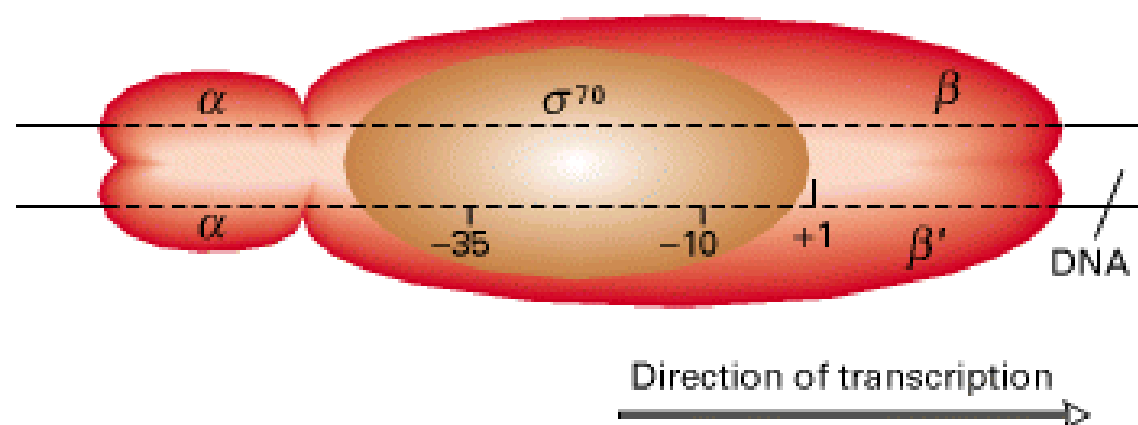
## Consensus sequences of $\sigma^{70}$ promoters



## (c) *Lac* promoter sequence

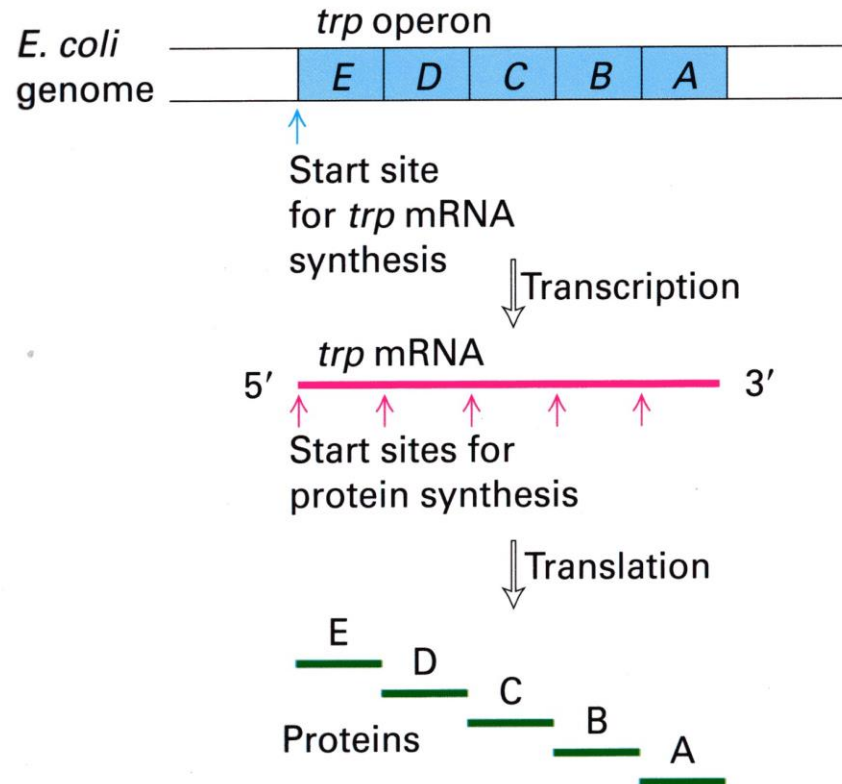


← Upstream, downstream →



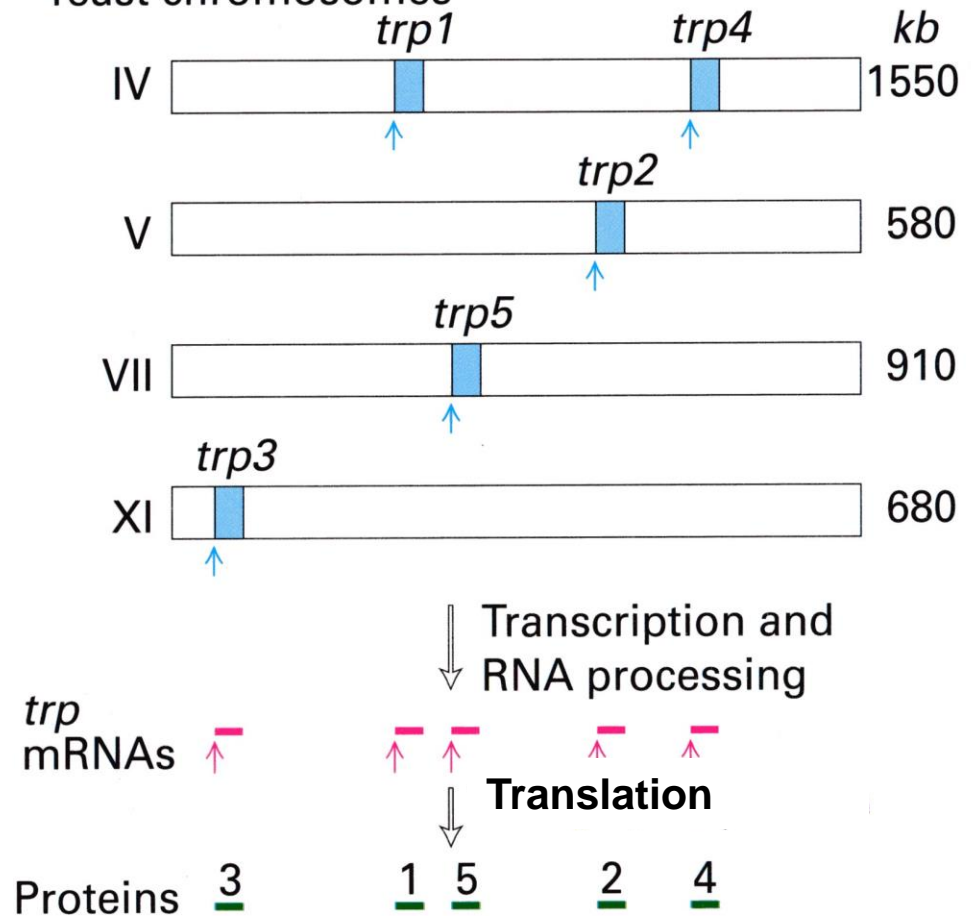
# Structural genes of metabolic chain

## Prokaryotes

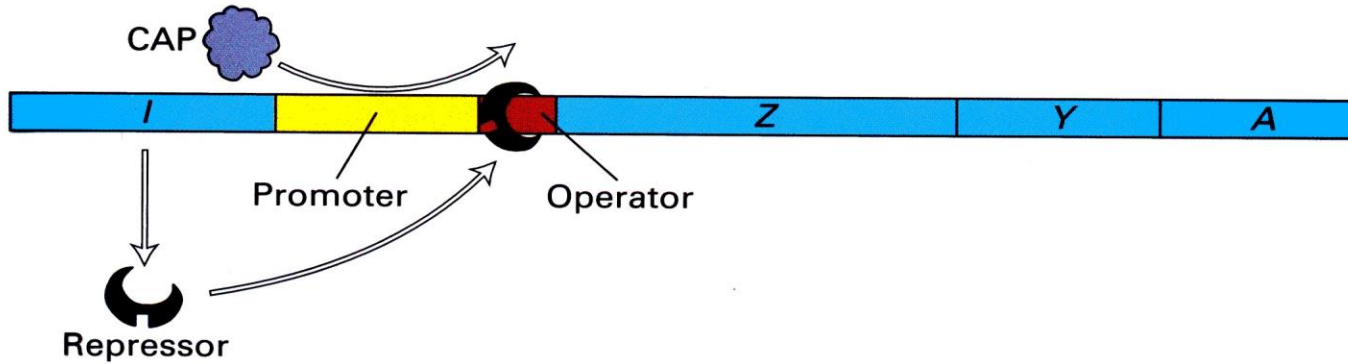


## Eukaryotes

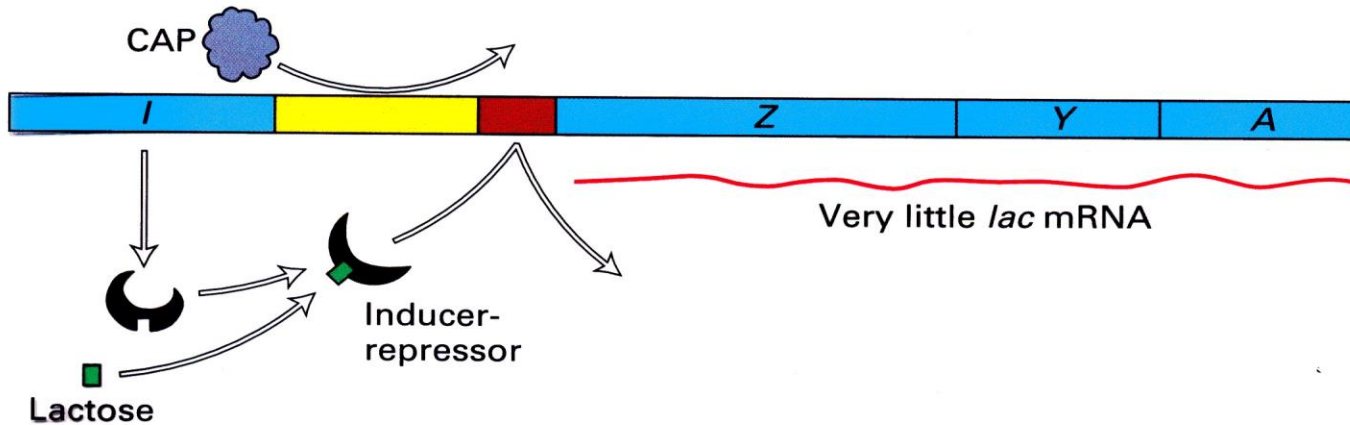
Yeast chromosomes



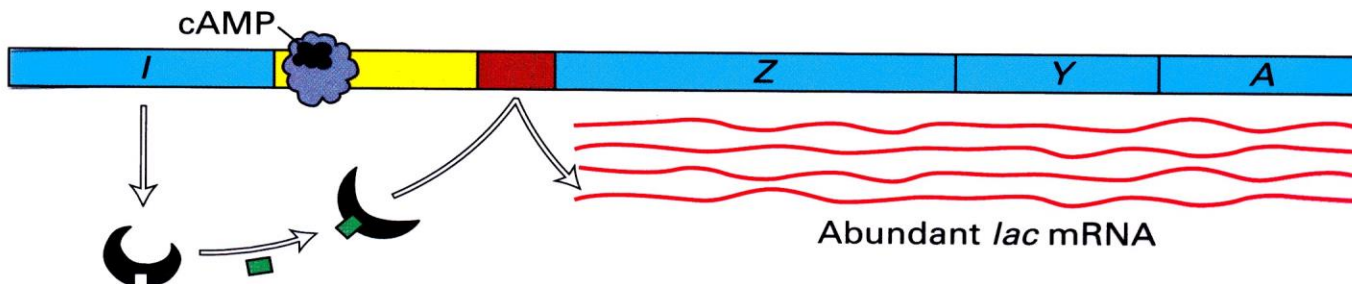
(a) Glucose present (cAMP low); no lactose



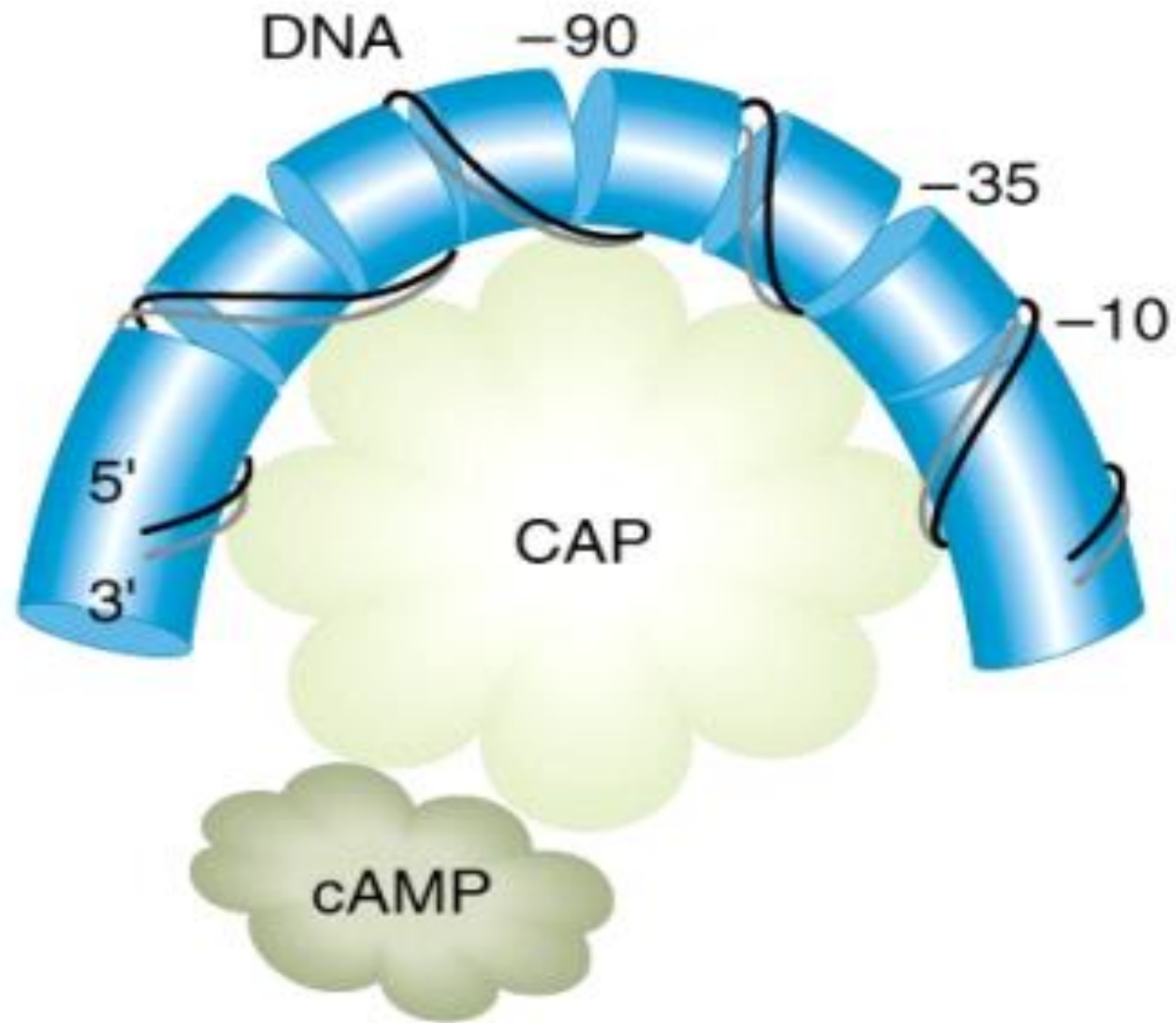
(b) Glucose present (cAMP low); lactose present



(c) No glucose present (cAMP high); lactose present

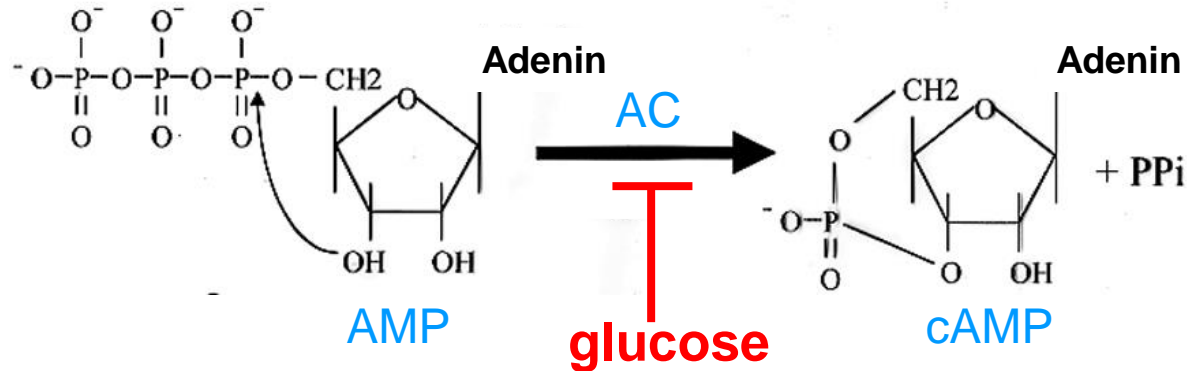


# CAP - bends DNA



# Adenylate cyclase and CAP – glucose repression of Lac

Adenylate cyclase (AC) - enzyme synthesizing cyclic AMP (cAMP) from ATP

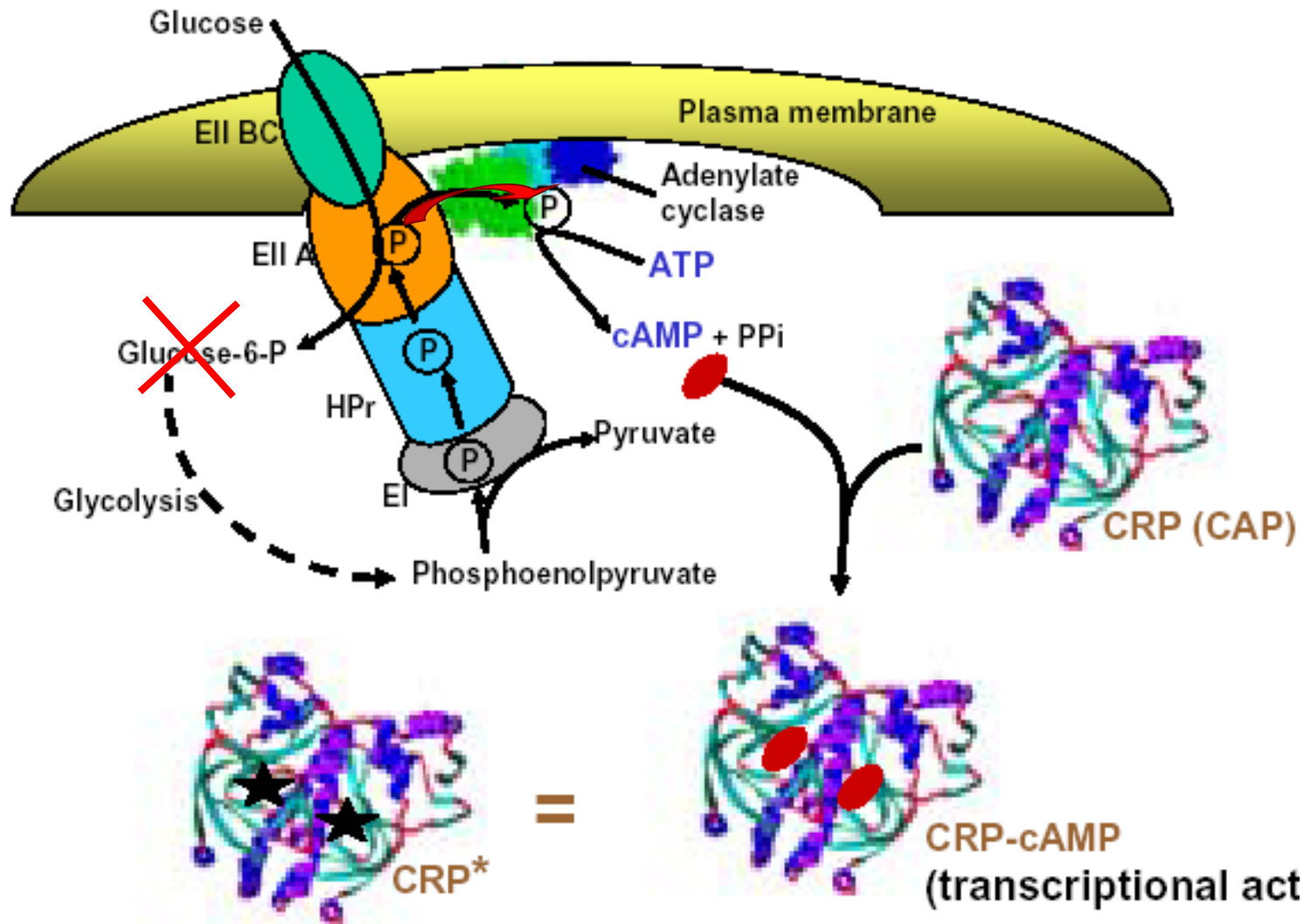


**High concentration of glucose**

- ⇒ inhibition of adenylate cyclase (indirectly by catabolic product)
- ⇒ low levels of cAMP

**absence of glucose ⇒ adenylate cyclase is not repressed**

⇒ high levels of cAMP

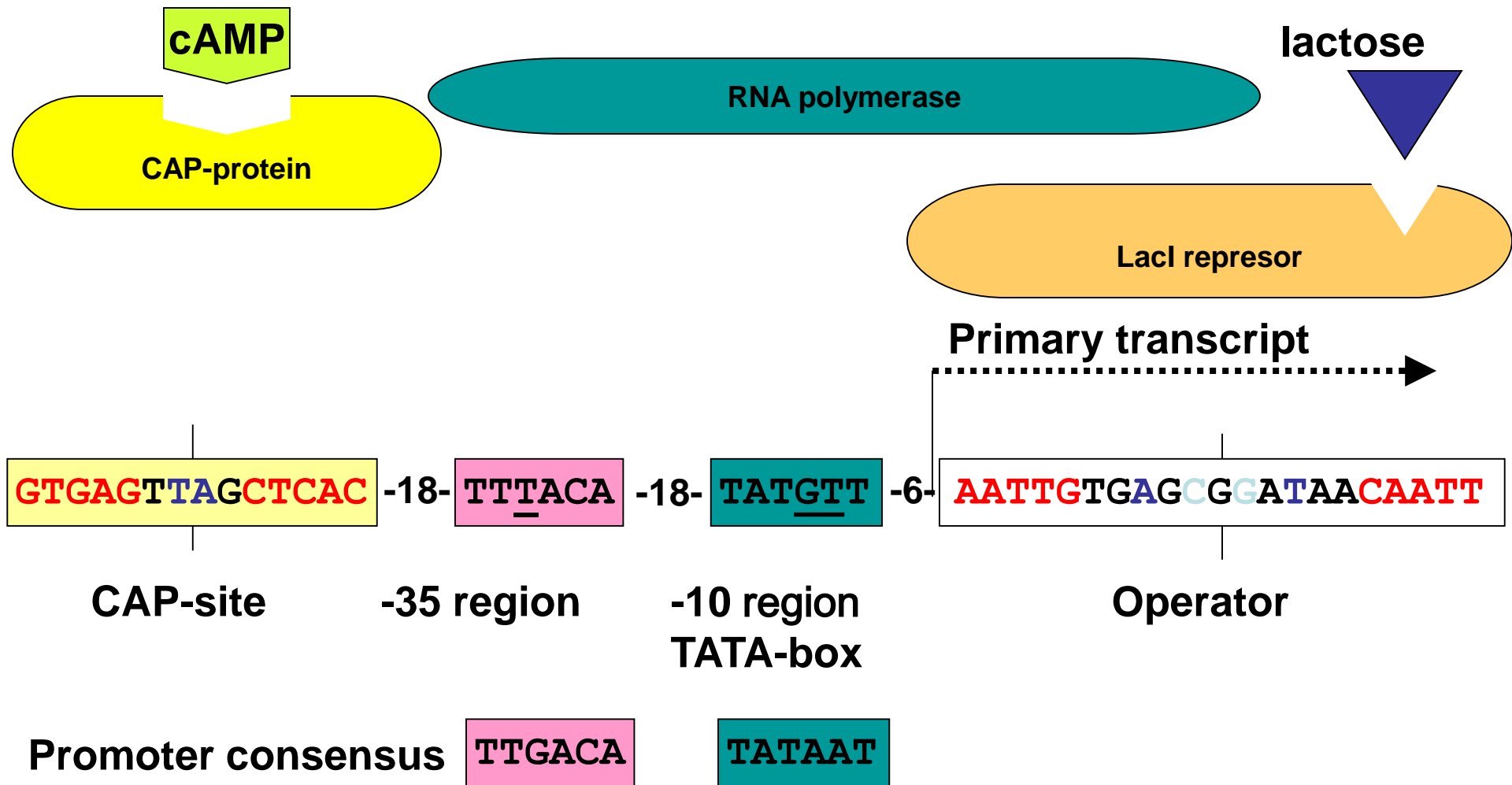


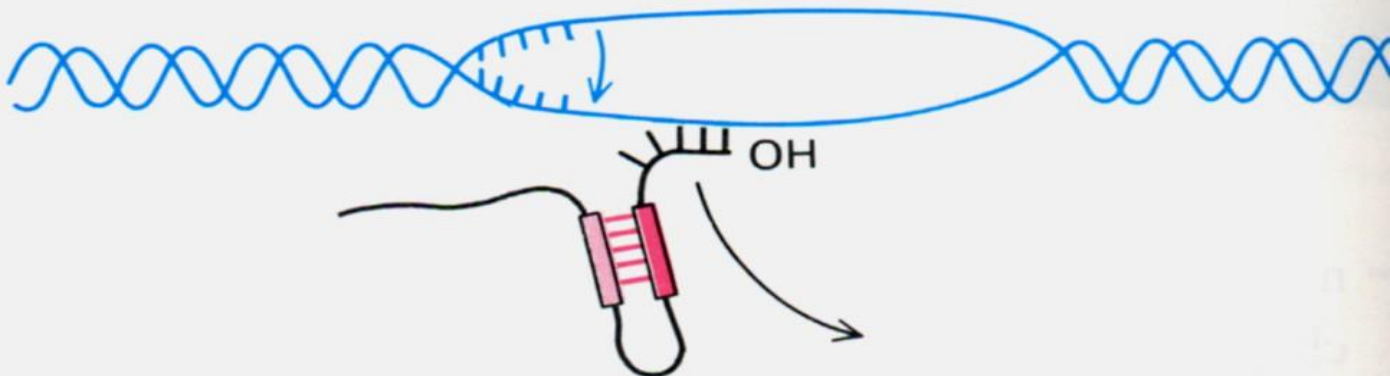
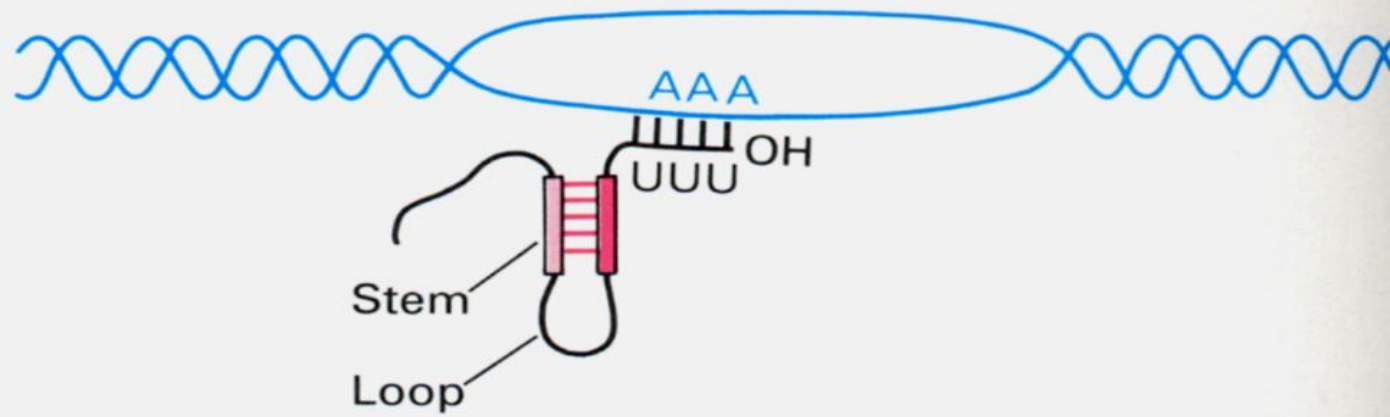
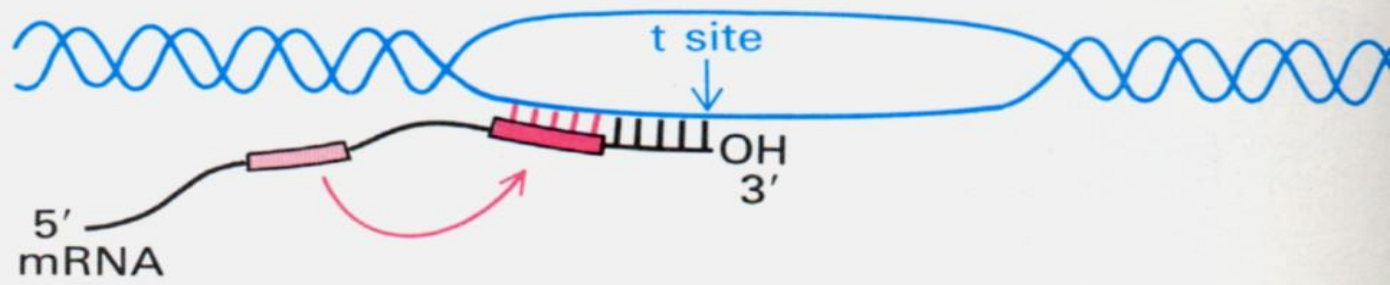
The absence of extracellular glucose

- Glucose transporter transferring phosphate from phosphoenolpyruvate to adenylate cyclase



# Lac promoter



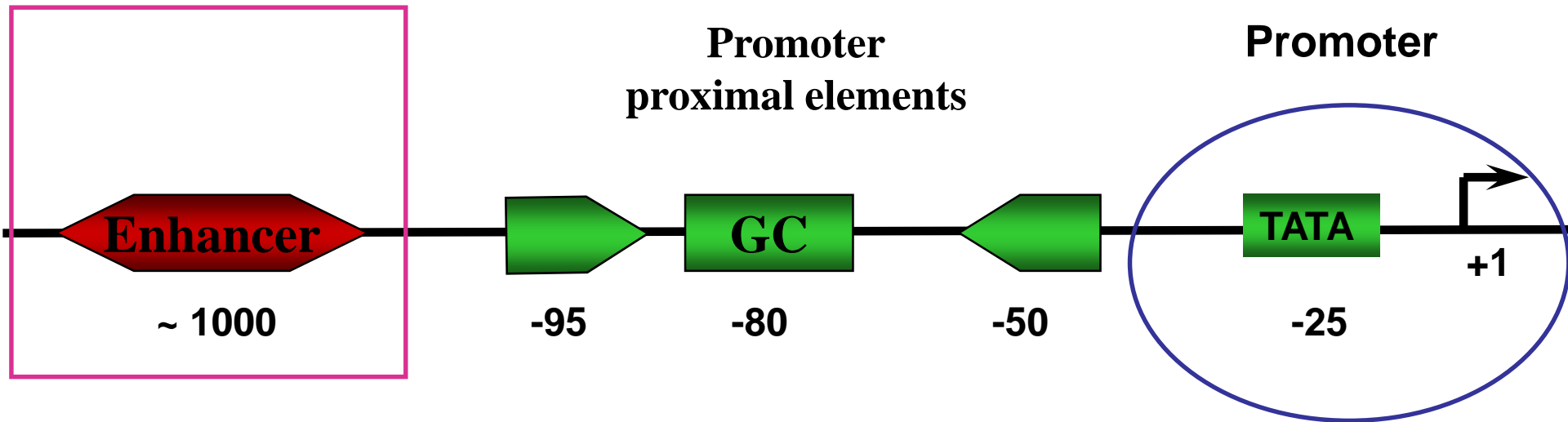


# Eukaryotic transcription

## Control of gene expression

- decision – „right“ genes, time, cells
- cis regulatory elements
  - targets for binding of regulatory proteins

# Eukaryotic genes

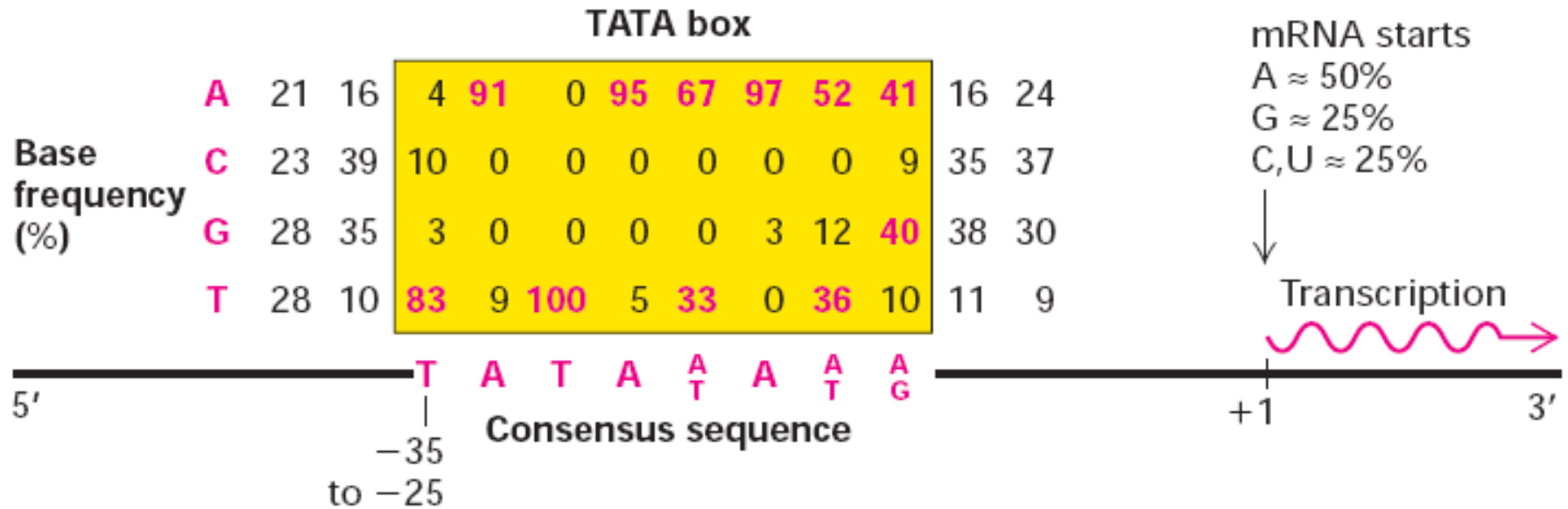


**Promotor** - sequence recognized by RNA polymerase

TATA box (TATAAAA) (Hogness box)      TATA  $\begin{smallmatrix} A & A \\ T & T \end{smallmatrix}$

# TATA box

- Most frequent regulatable promoter
- 25 - 30 bp upstream (-25 až -30) from start



# Initiators

Some eukaryotic genes – instead of TATA box

Not conservative – 5' YY<sup>A+1</sup>NT/AYYY

(Y – pyrimidine)



## **Alternative promoters**

Neither TATA box nor initiator

**CG island** – GC rich regions

**Housekeeping genes** (constitutive genes) for central metabolic pathways (e.g. TCA cycle)

# Human RNA polymerases

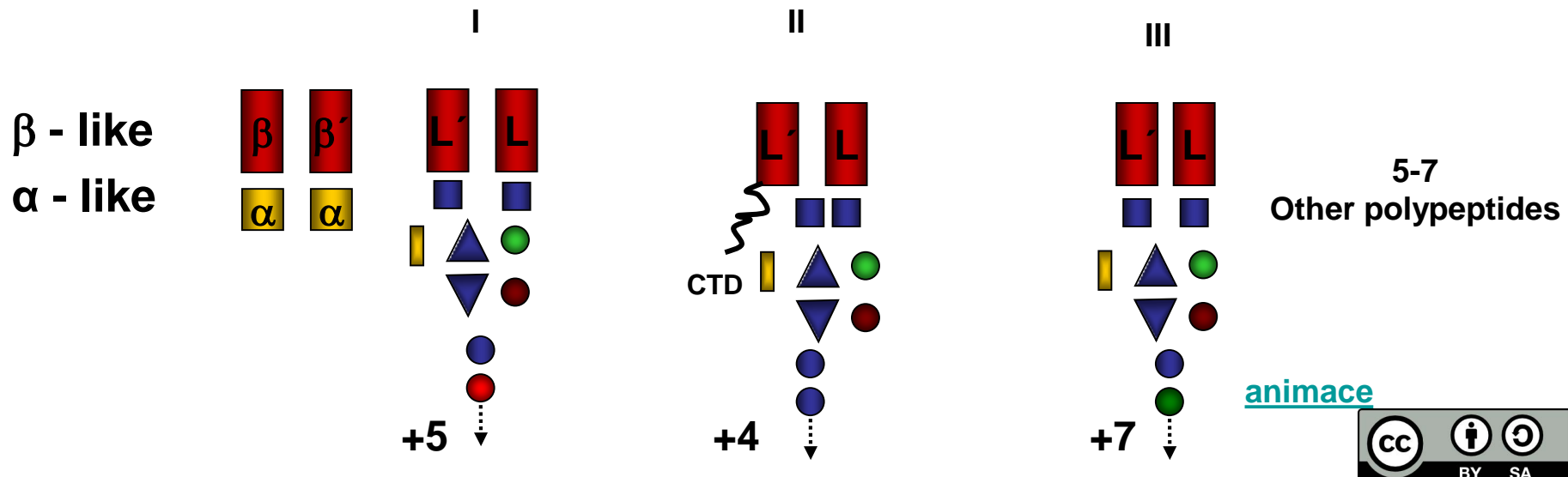
<u>Polymerase</u>	<u>Localization</u>	<u>Product</u>
RNA polymerase I	nucleolus	18S, 28S, 5.8S rRNA
RNA polymerase II	nucleoplasm	hnRNA/mRNA, U1, U2, U4, U5 snRNA
RNA polymerase III	nucleoplasm	tRNA, 5S RNA, U6 snRNA, 7SL RNA
mitochondrial RNA polymerase	mitochondria	all mitochondrial RNAs

---

2 large subunits – L' 190 – 220 kDa  
 L 140 – 150 kDa

70-80% identity of all three polymerases

Some regions – homologous with *E. coli*  $\alpha$  and  $\beta$   
 Other subunits – also homologous  
 (5 subunits common to all three pol.)



# CTD

Heptapeptide repetition Tyr-Ser-Pro-Thr-Ser-Pro-Ser

Only Pol II

- in yeast ~ 26x
- in mammals ~ 52x
- Ser and Tyr – phosphorylation
- phosphorylation enhances processivity
- *in vitro* phosphorylated after transcription initiation

# Initiation of transcription - participation of TFs

## General transcription factors

Human Genome  $\approx$  2,000 transcription factors

## Proteins regulating synthesis of eukaryotic mRNA

binding + activation domain

Classification by type of binding domain

**RNA polymerase II binds to the promoter via TFIID**

**Transcription factors bind promoter elements and interact with proteins on the promoter**

- Stabilization (or inhibition) of pre-initiation complex formation**



**Beginning of assembly of transcription system**

**TBP (TATA box binding protein) - binds a TATA box**

**Factors supporting TAF (TBP-Associated Factors - 13 proteins)**

**- Bind TBP - form the TFIID complex (many subunits)**

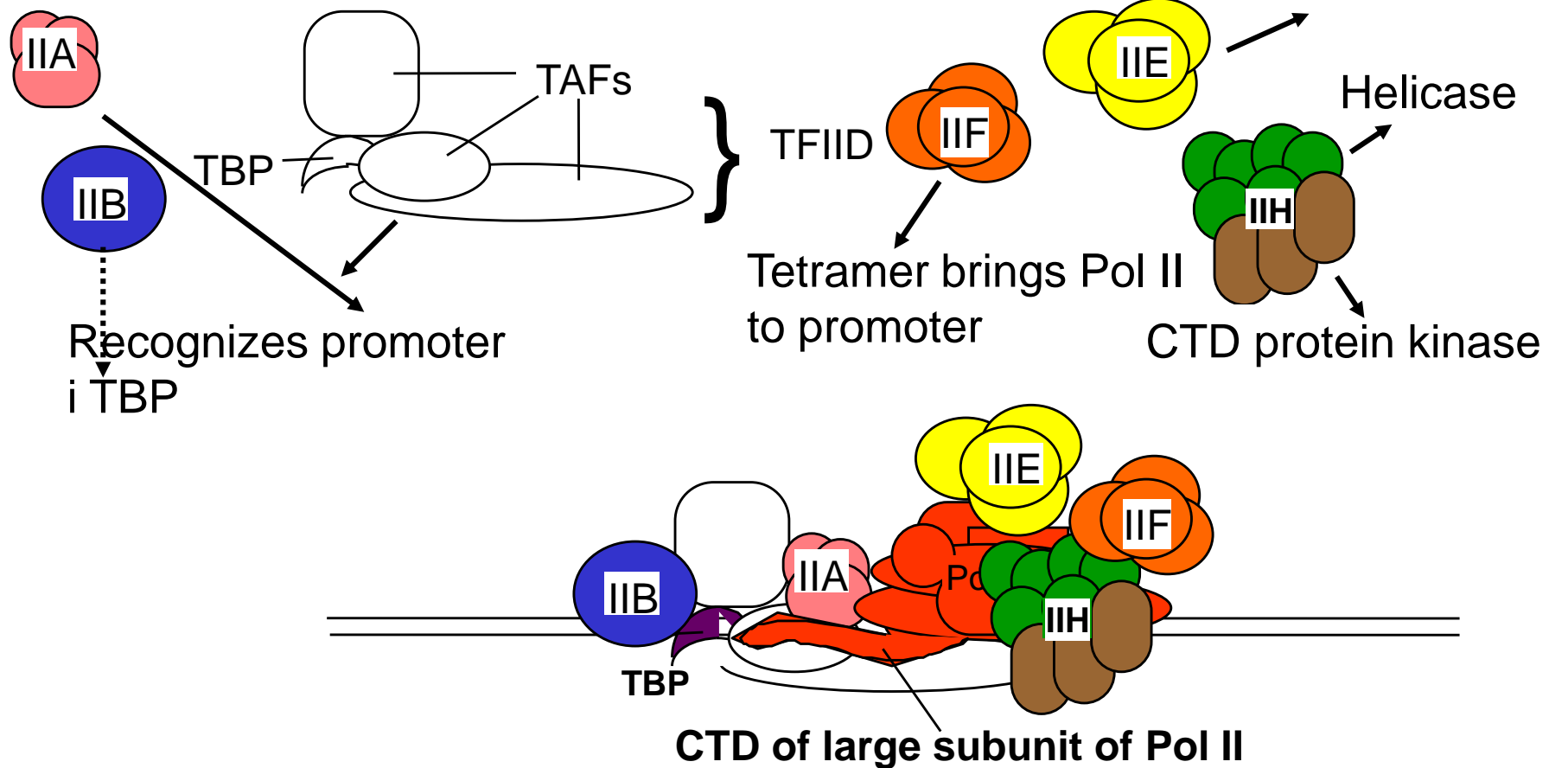
**TFIIA, TFIIB, TFIIIE, TFIIF, TFIIH combine with TFIID**

[TATA binding protein](#)

**TFIIH – phosphorylation of RNA polymerase II**

# Assembly of preinitiation complex of RNA polymerase II

(TFIIA – essential *in vivo*)



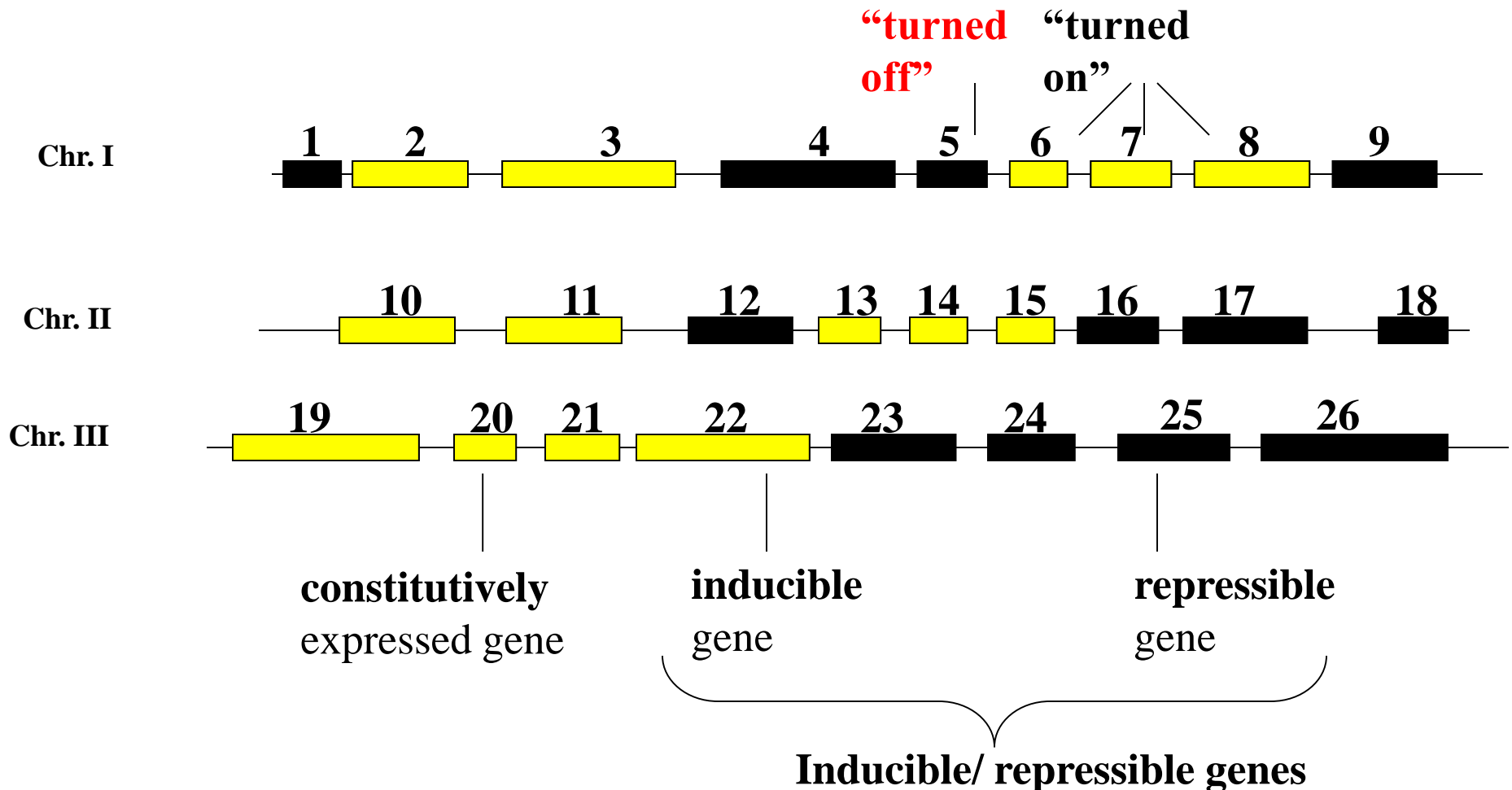
(HNF1 hepatocyte nuclear factor 1, C/EBP interacts with CCAAT)

**GTFs (general transcription factors)**  
**TBP – TATA box binding protein**  
**TAF – TBP-associated factors**

**initiation of other than TATA promoters**  
**binding on the initiator and**  
**downstream (30 bp) promoter**  
**of genes without TATA**

# Gene regulation

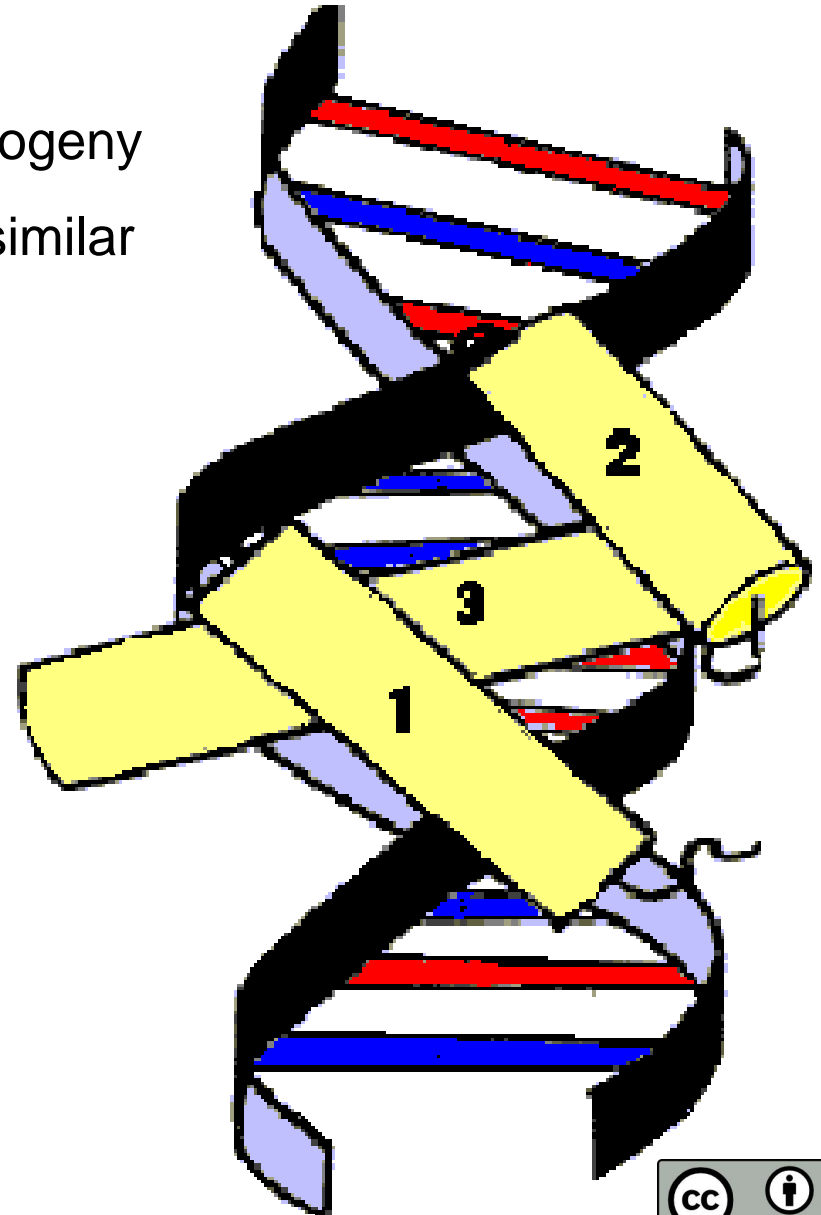
Conditions 2



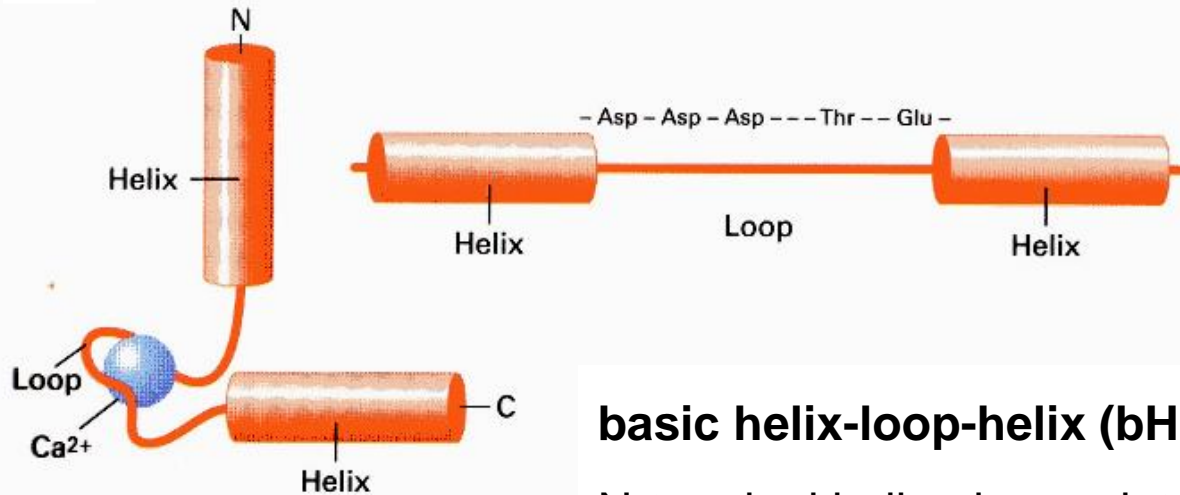
# Helix-turn-helix

3 roviny helixů - vazba žlábků DNA

Many eukaryotic TFs involved in the ontogeny  
conservative 60-AA DNA-binding motif similar  
homeodomain helix-turn-helix bacterial  
repressors



# Helix-loop-helix



## basic helix-loop-helix (bHLH)

N-terminal helix - base - interaction with DNA

C-terminal region - hydrophobic AA -

intervals characteristic of a class A amphipathic helix.

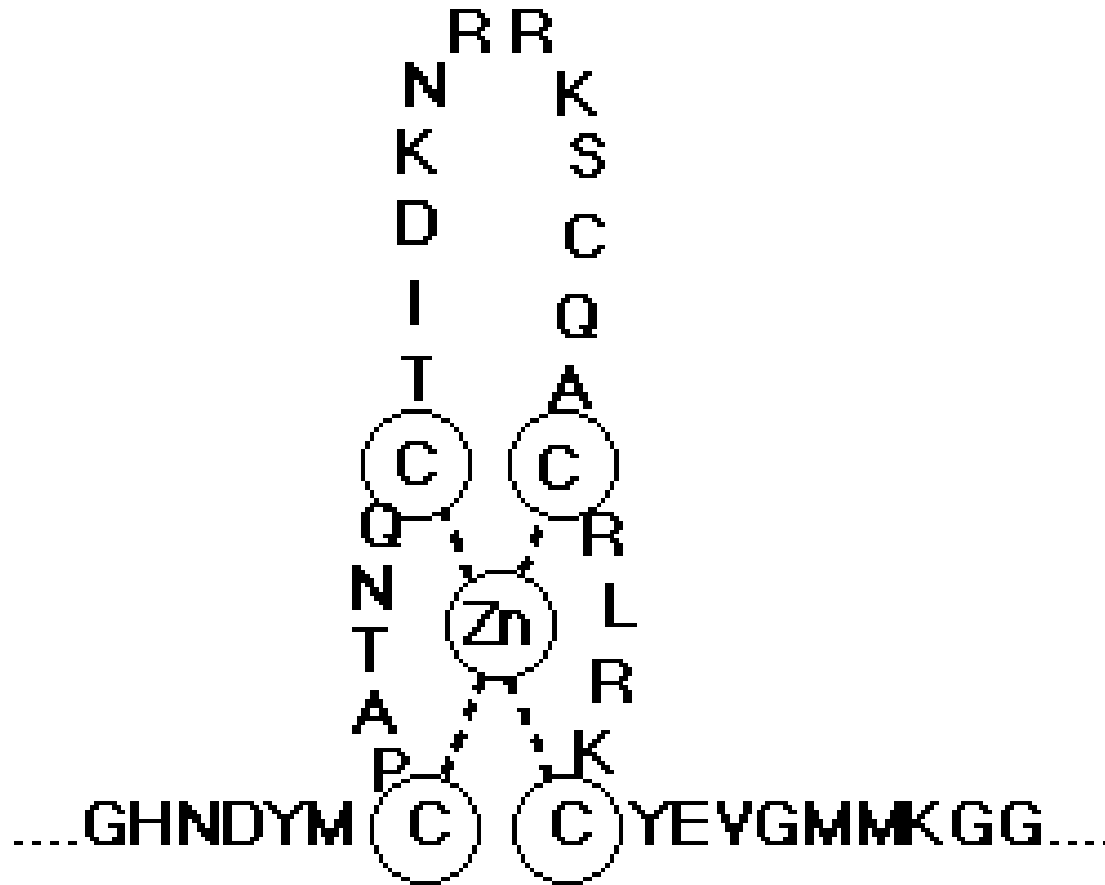
Various bHLH proteins - formation of heterodimers

# Zinc fingers

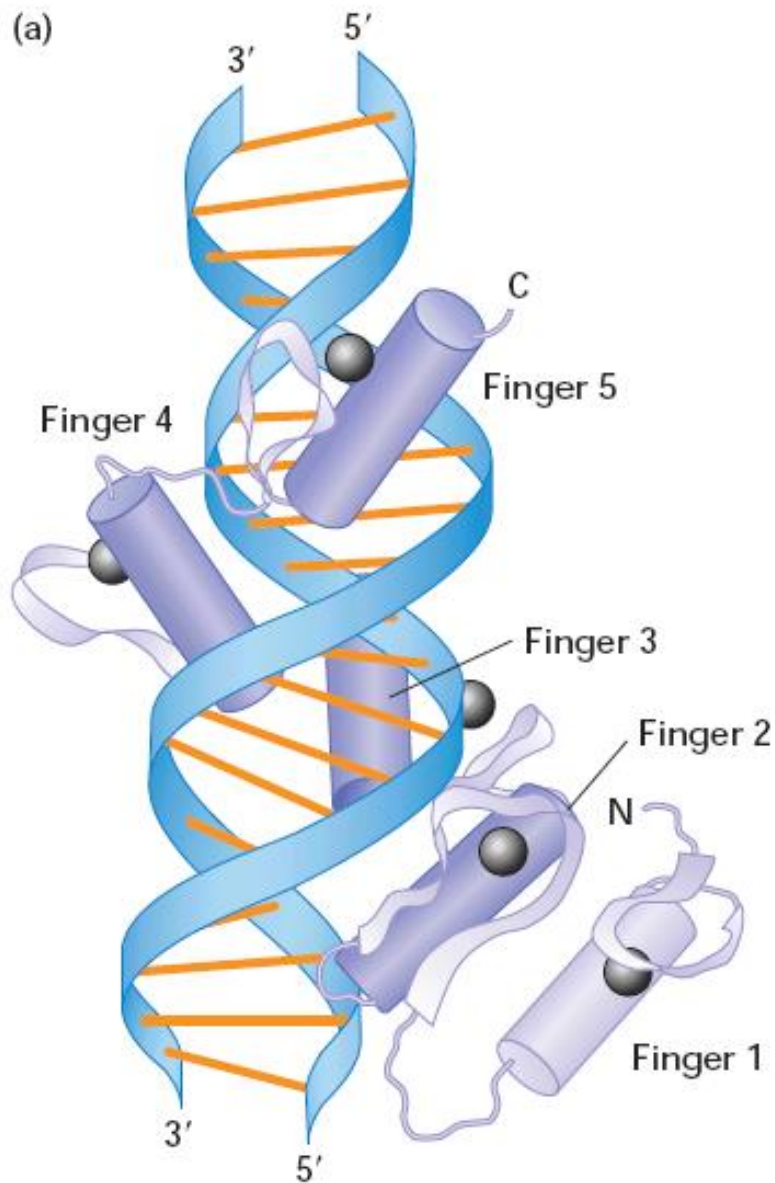
cysteine and histidine residues ( $C_2H_2$ —most common)

- Bind  $Zn^{2+}$

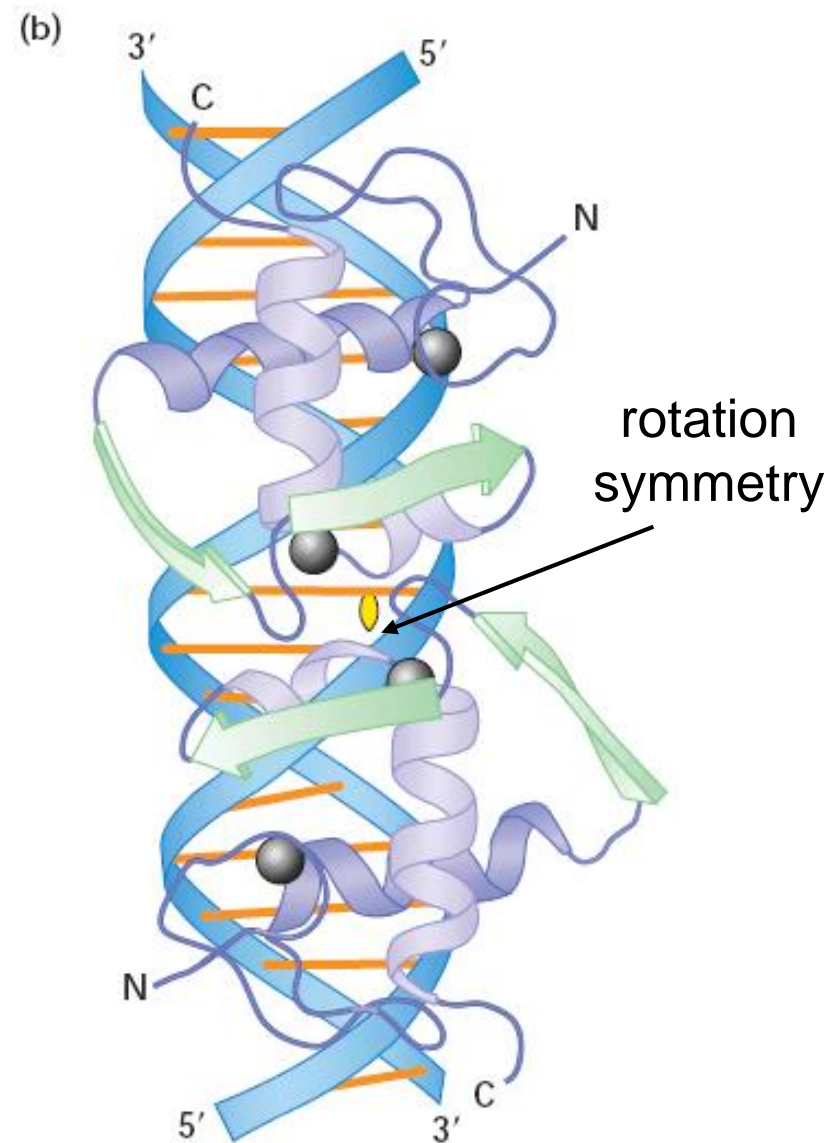
Finger fits in DNA grooves



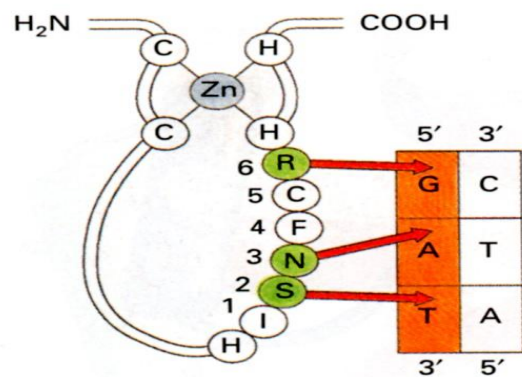




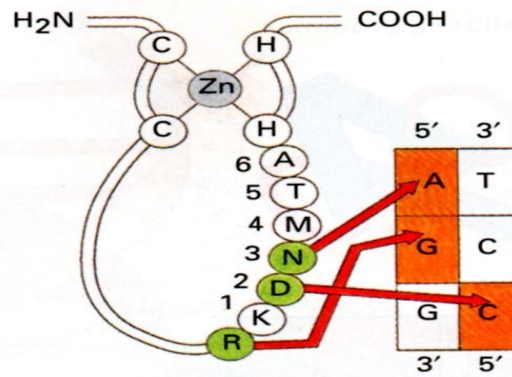
$C_2H_2$  – common - mammals, plants



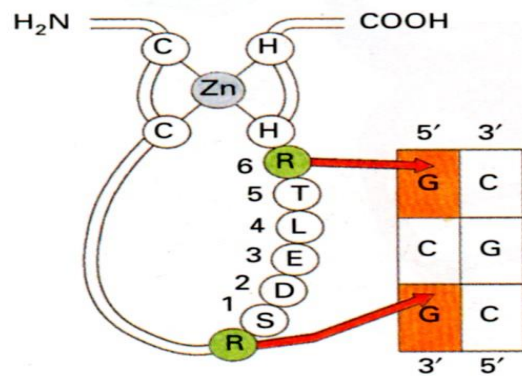
$C_4$  homodimer  $\approx$  50 humanTFs,  
for hormone receptors



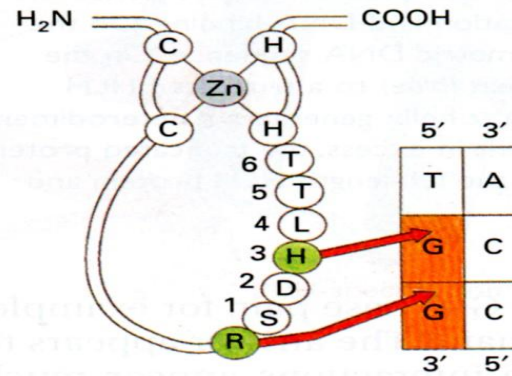
TTK Finger 1



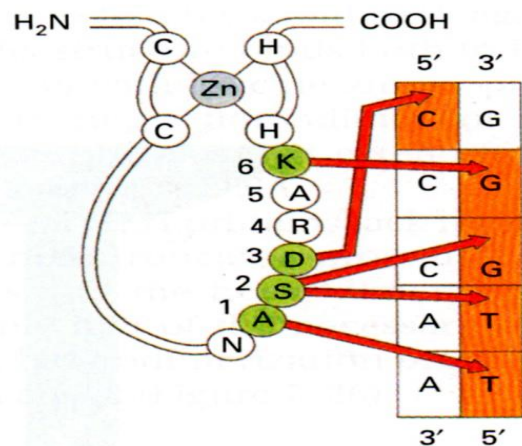
TTK Finger 2



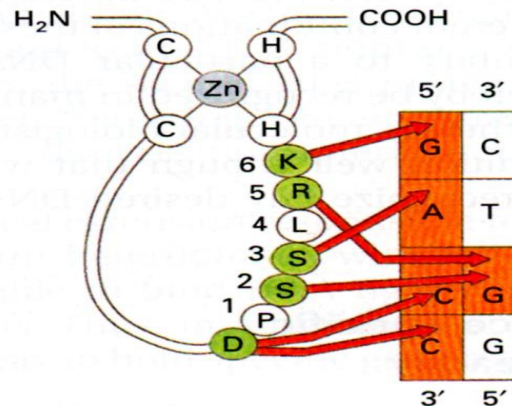
Zif Finger 1



Zif Finger 2



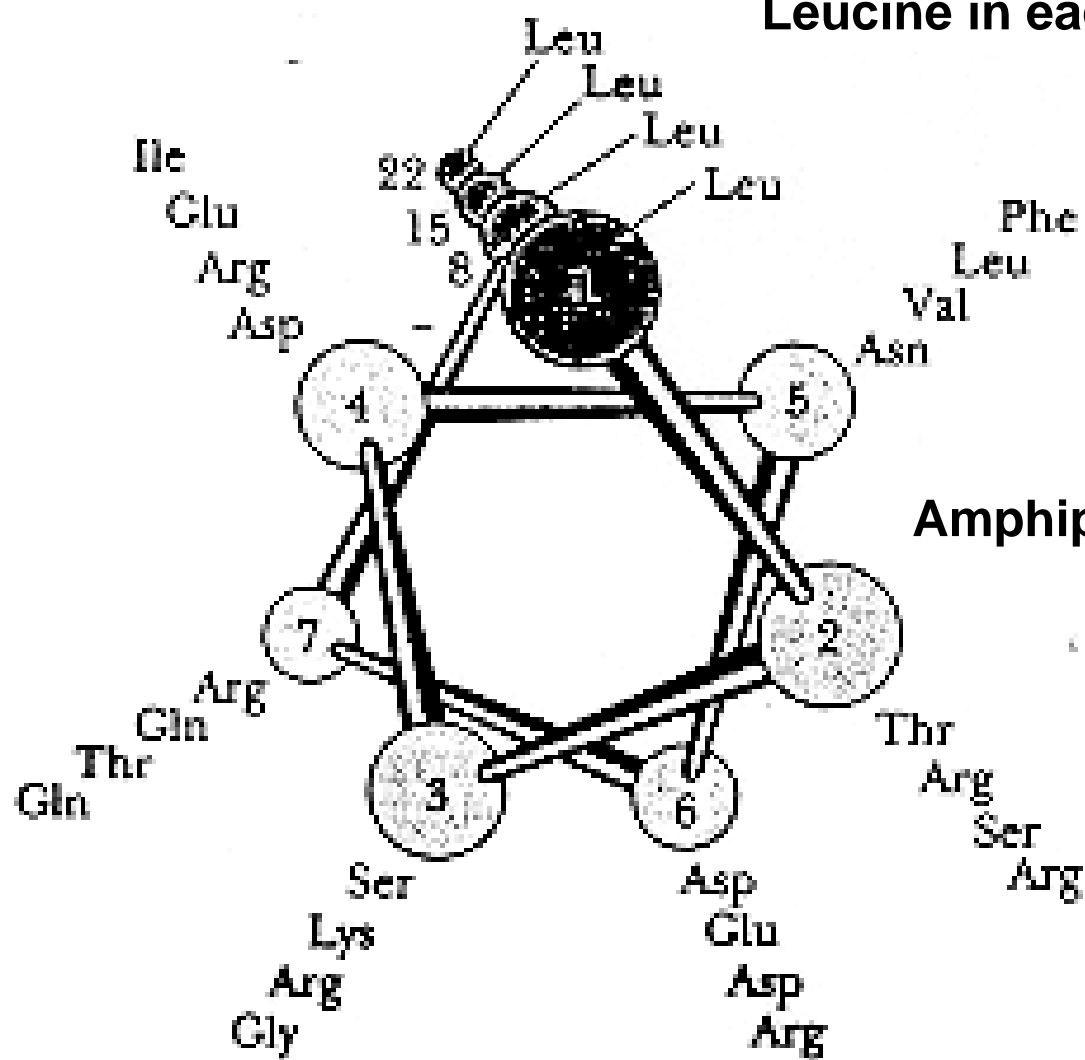
GLI Finger 4



GLI Finger 5

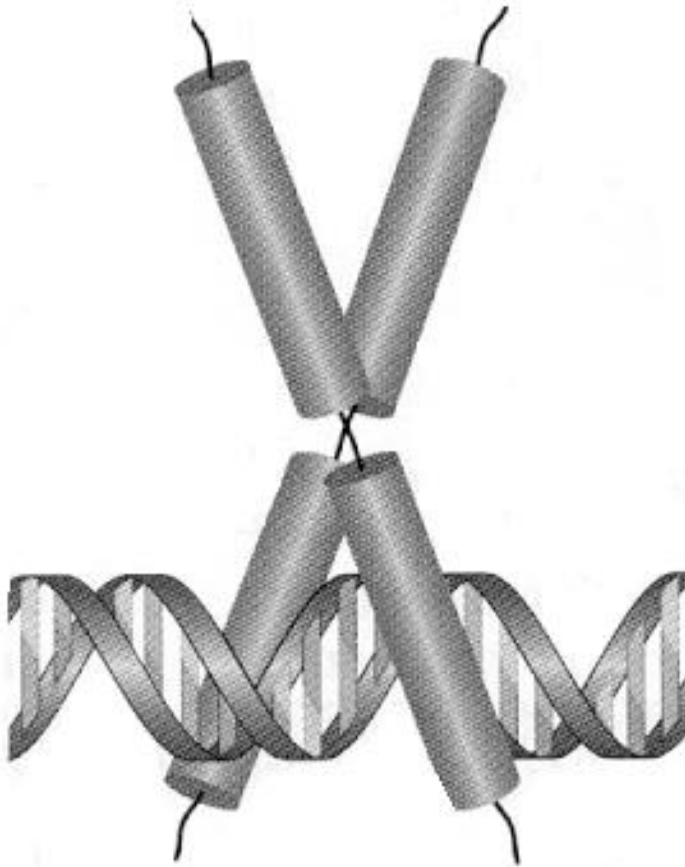
# Leucine zipper

Leucine in each seventh position



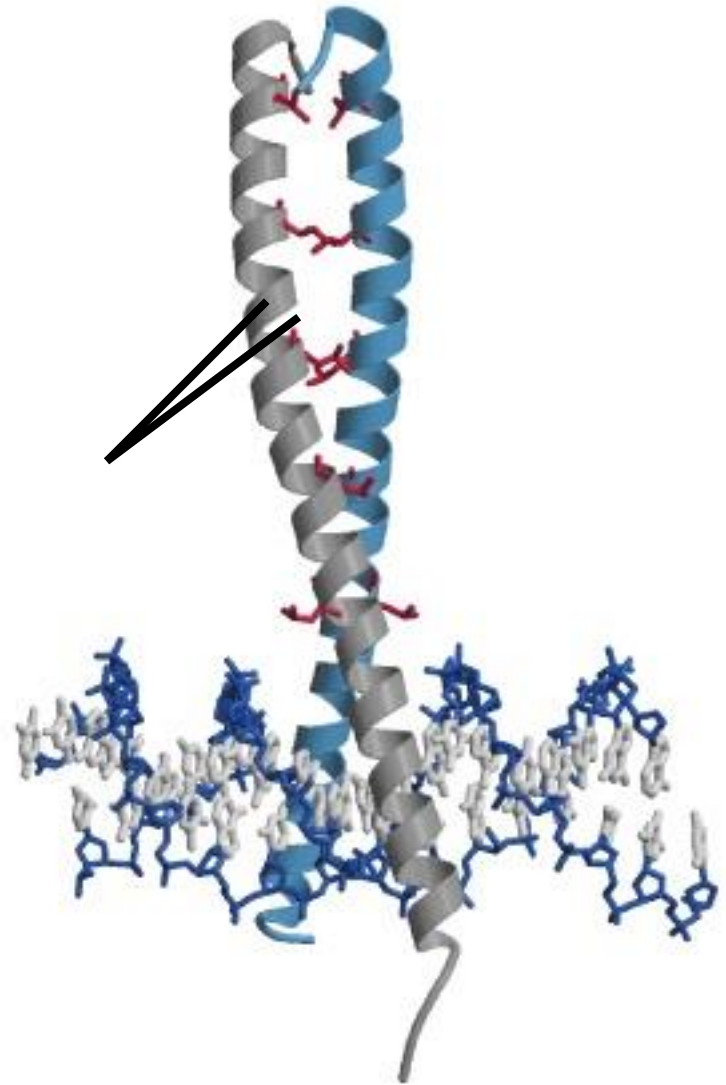
Amphipathic alpha helix

# Leucine zipper



dimers - Leu residues - each turn in  $\alpha$ -helix  
 $\alpha$ -helical region - leucine zipper  
Adjacent to zipper - Y shaped area - DNA binding

Leucines



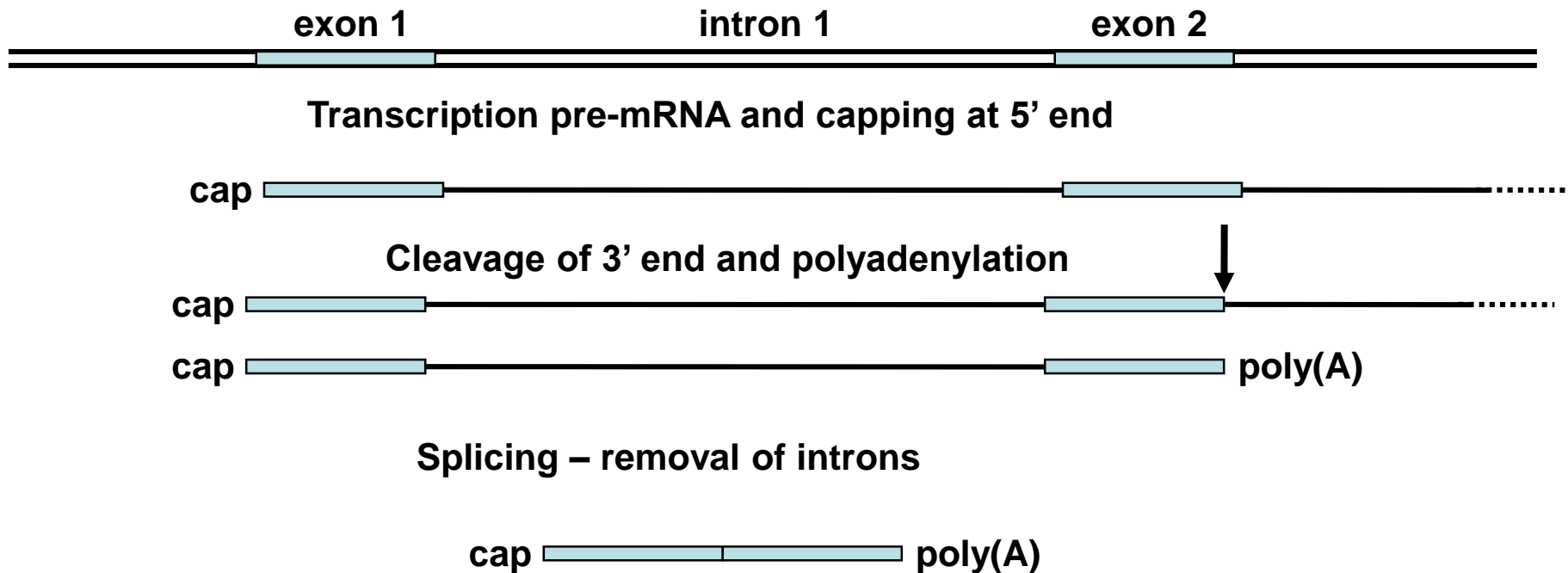
# Posttranscription modifications of RNA

**mRNA (hnRNA - precursor mRNA)**

**cap - capping (cotranscriptional)**

**cleavage and polyadenylation (- 3' end)**

**splicing (before the export from nucleus)**

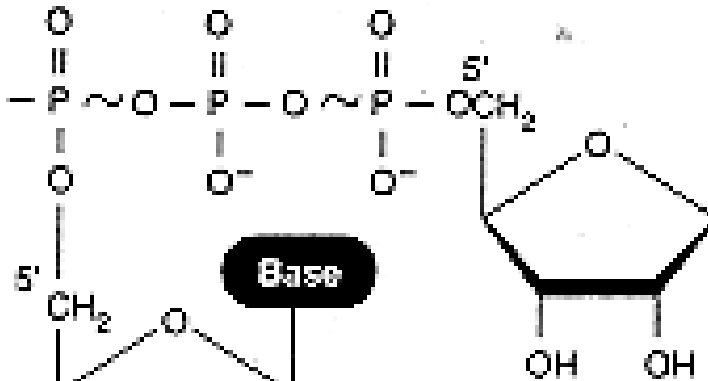
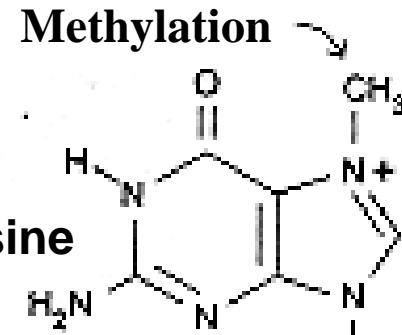


# Cap

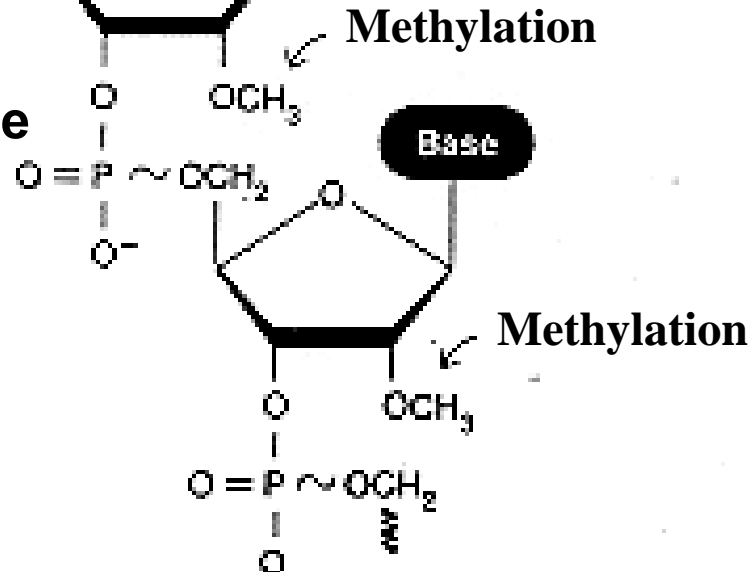
Protection against nucleases,  
Required for translation initiation

Triphosphate bond

7-methylguanosine



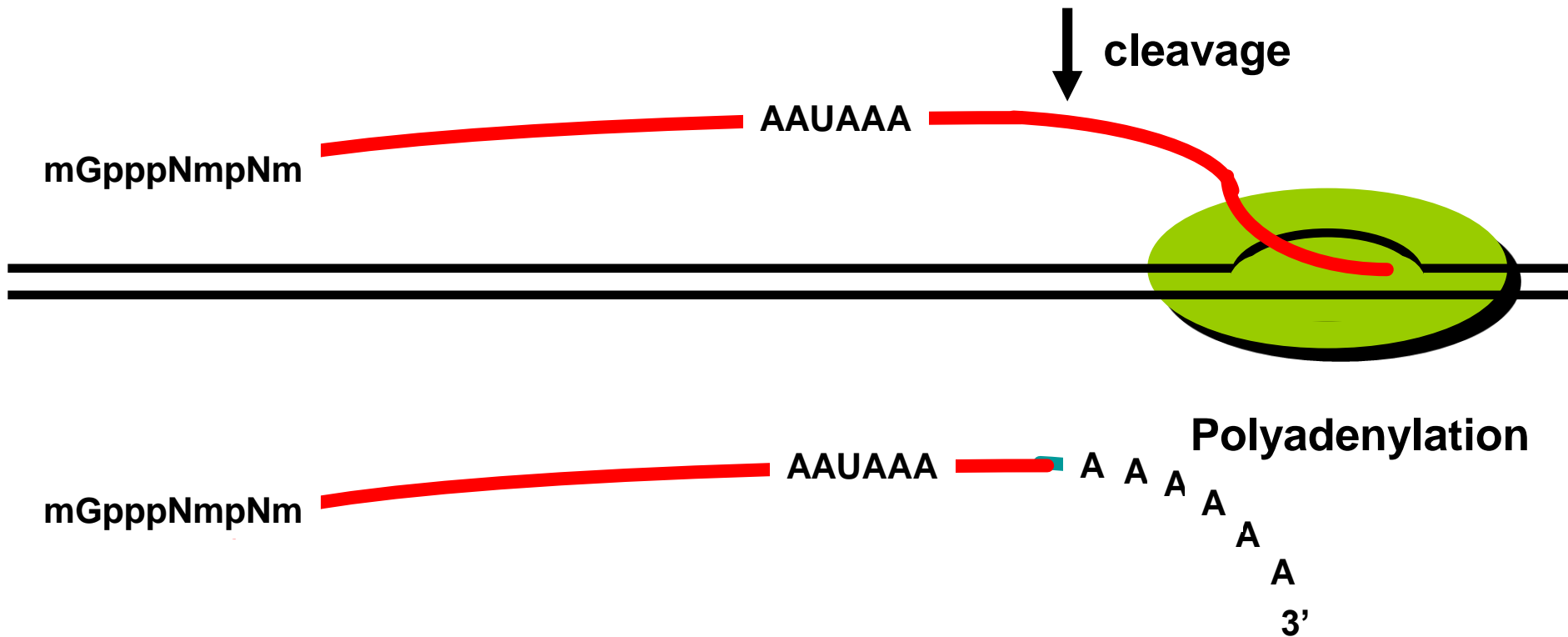
Methylation of 2' ribose

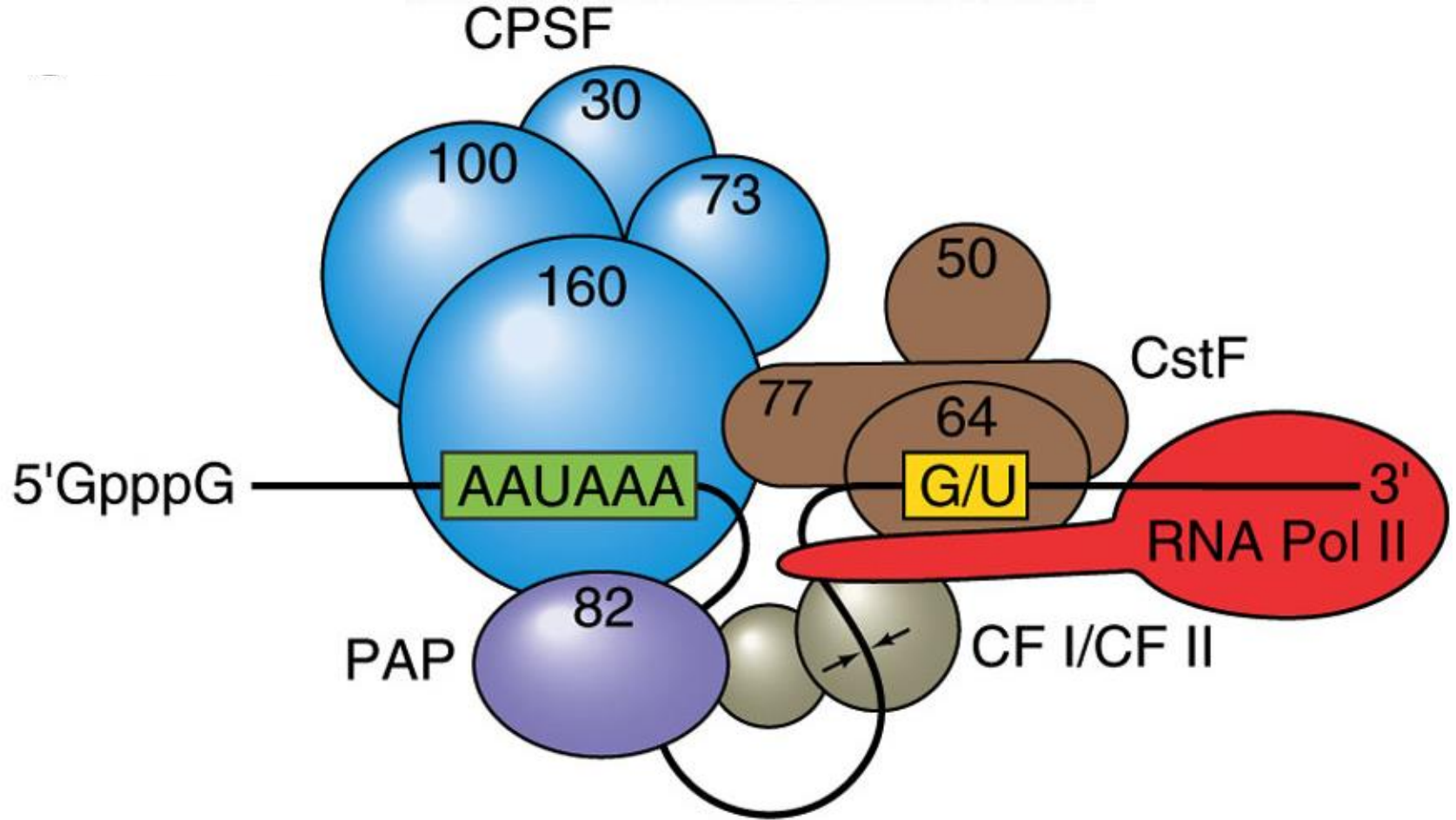




10-30 nucleotides downstream from AAUAAA  
consensus sequence

- polyadenylation - **poly(A) polymerase**
- ~ 200 adenylates / binding of specific proteins
- stability



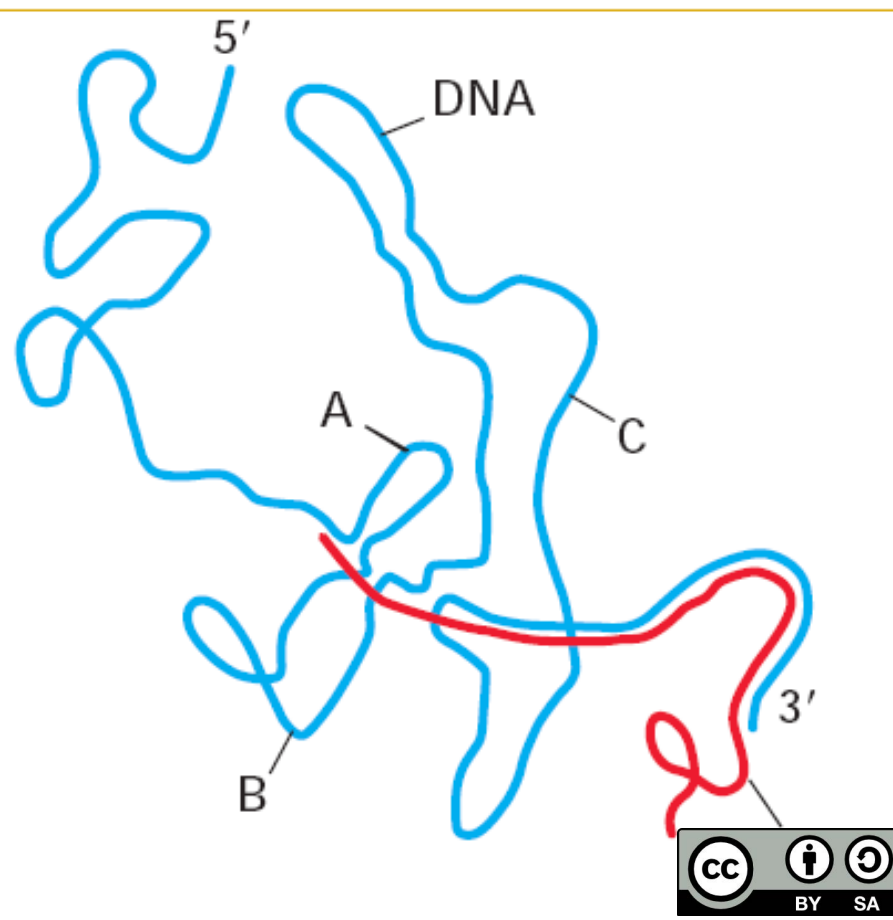
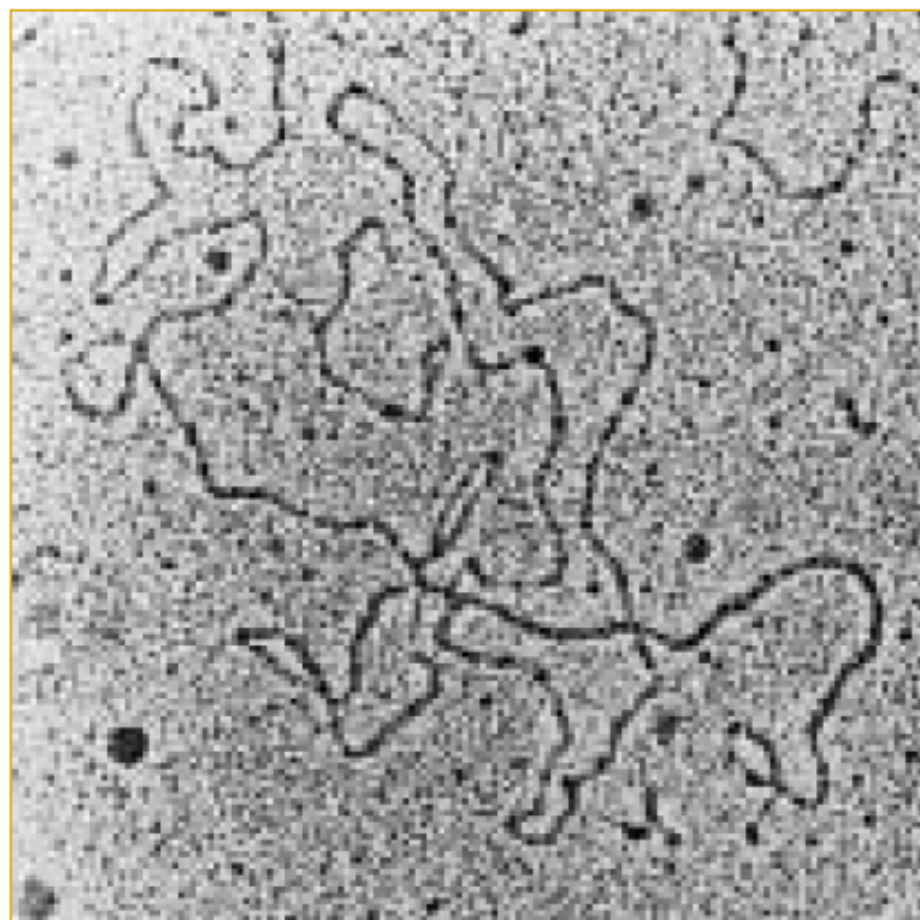
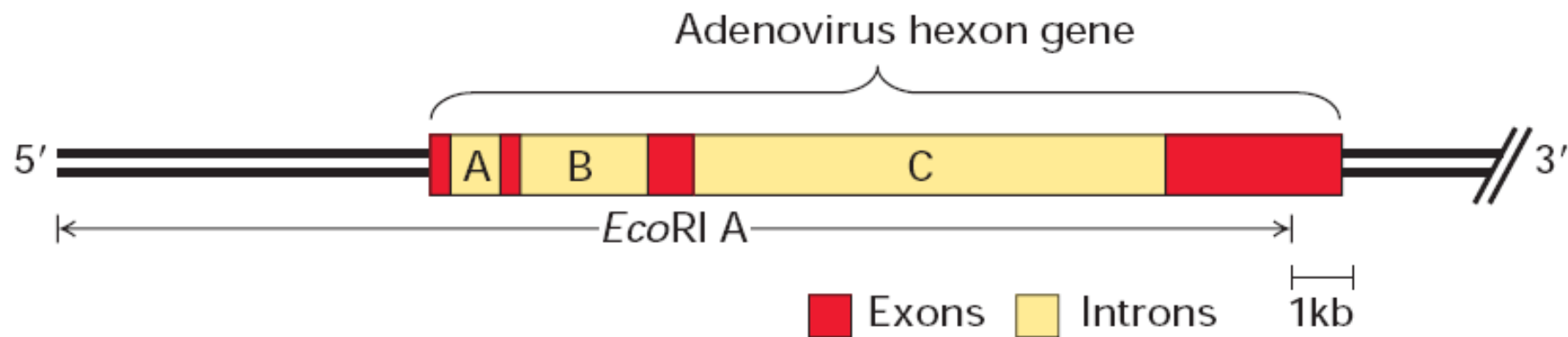


**Cleavage and Polyadenylation Specificity Factor (CPSF)**

**Cleavage stimulating factor (CstF)**

**Cleavage factors I and II (CF I, CF II)**

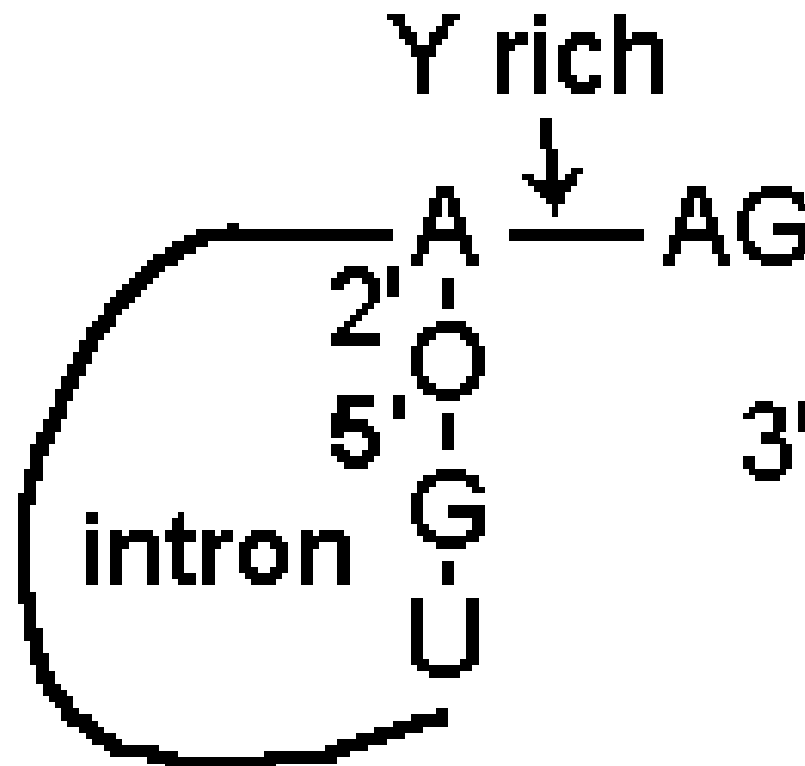
**PolyA polymerase (PAP)**

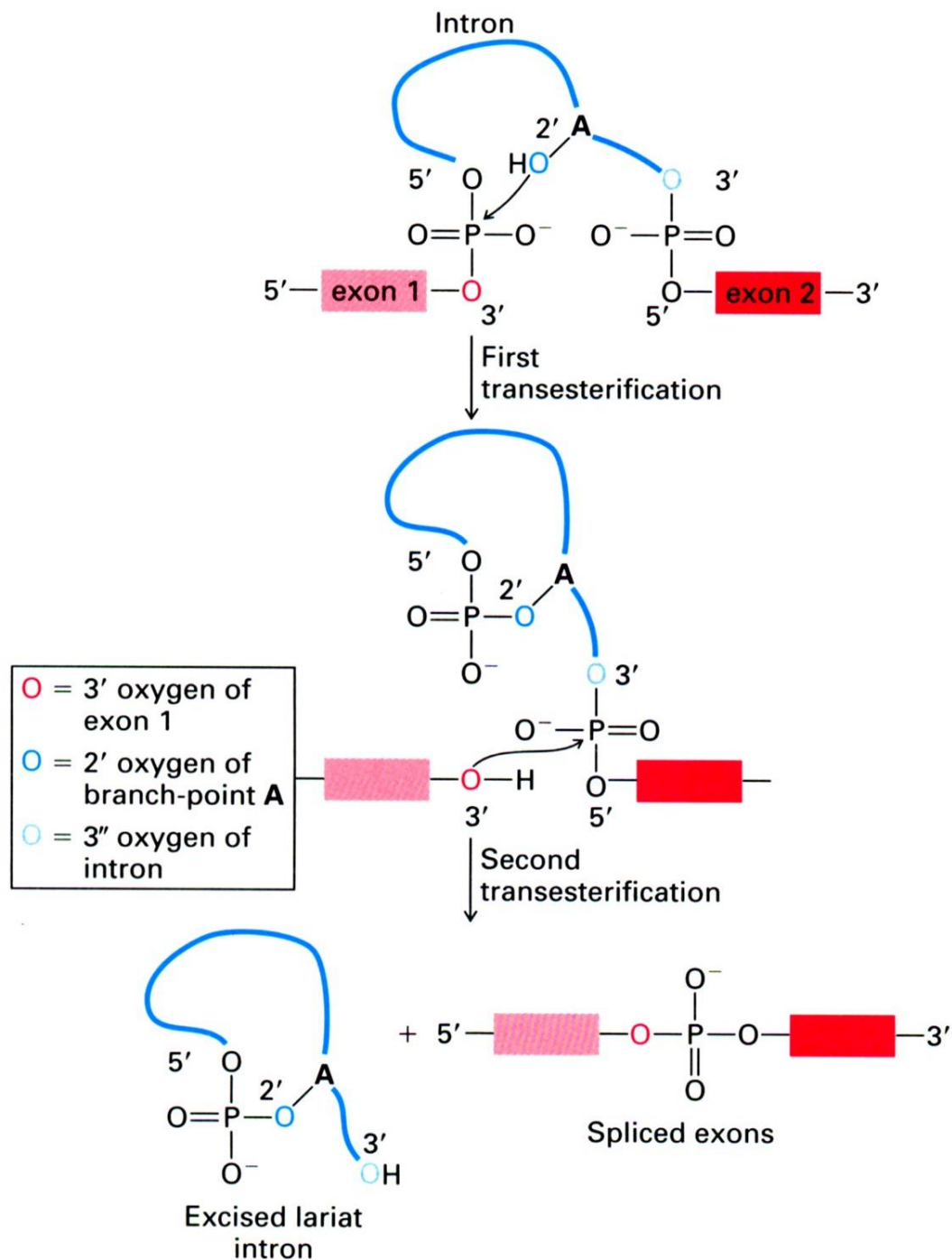




Exon 1

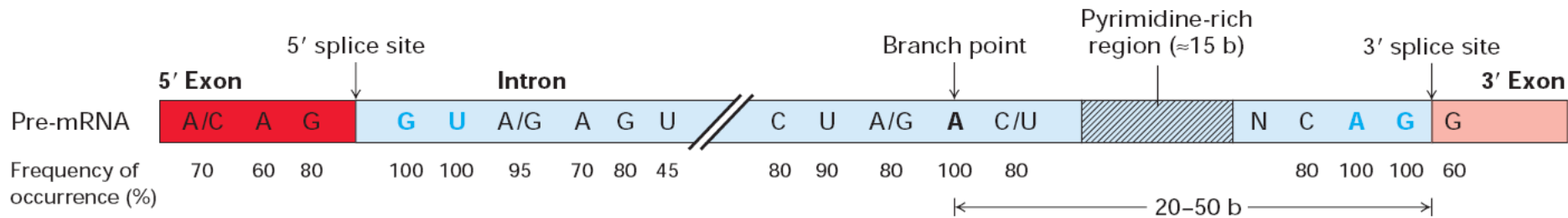
Exon2







# Frequency of bases in splice sites

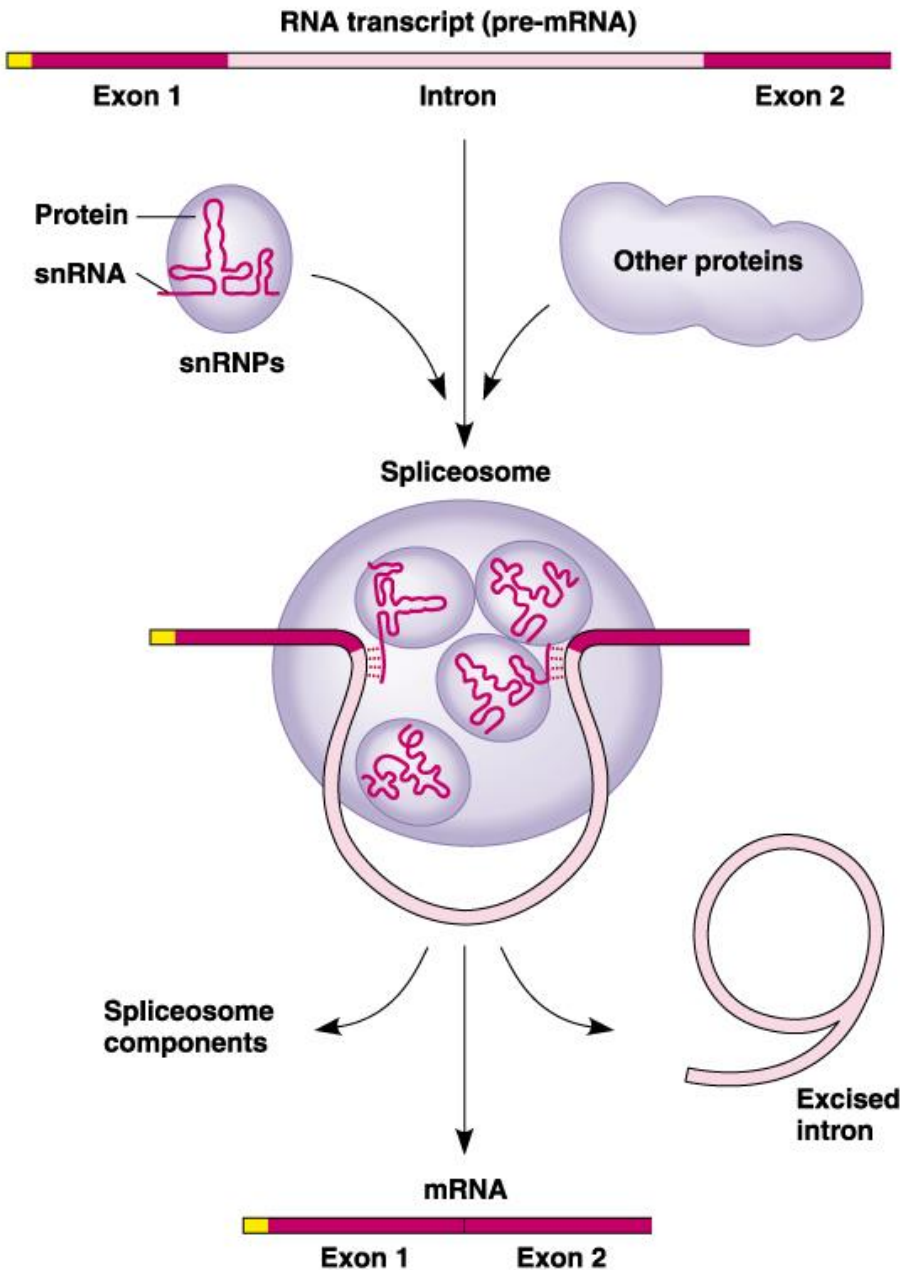


	exon				intron							
%A	30	40	64	9	0	0	62	68	9	17	39	24
%U	20	7	13	12	0	100	6	12	5	63	22	26
%C	30	43	12	6	0	0	2	9	2	12	21	29
%G	19	9	12	73	100	0	29	12	84	9	18	20
			A	G	<u>G</u>	<u>U</u>	A	A	G	U		

## Acceptor 3' sequence

	intron														exon		
%A	15	10	10	15	6	15	11	19	12	3	10	25	4	100	0	22	17
%U	51	44	50	53	60	49	49	45	45	57	58	29	31	0	0	8	37
%C	19	25	31	21	24	30	33	28	36	36	28	22	65	0	0	18	22
%G	15	21	10	10	10	6	7	9	7	7	5	24	1	0	100	52	25
	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	A	G	G	

Polypyrimidine tract (Y = U or C; N = any nucleotide)



## Spliceosome

Recognition of specific sites of intron

Intron cleavage

Joining of exons

snRNPs

Small nuclear ribonucleoproteins

RNA



Spliceosome

Splicing regulated by ribonucleoproteins

## **Spliceosome**

- snRNA (small nuclear RNA)
- snRNPs “snurps” (sn ribonucleoproteins)

Five subunits - U1, U2, U4, U5, U6

U1 recognition of 5'- site

U2 recognition of branching site

U4 Interacts with U6 – complementary complex

U5 3'- linker binds U4-U6 complex

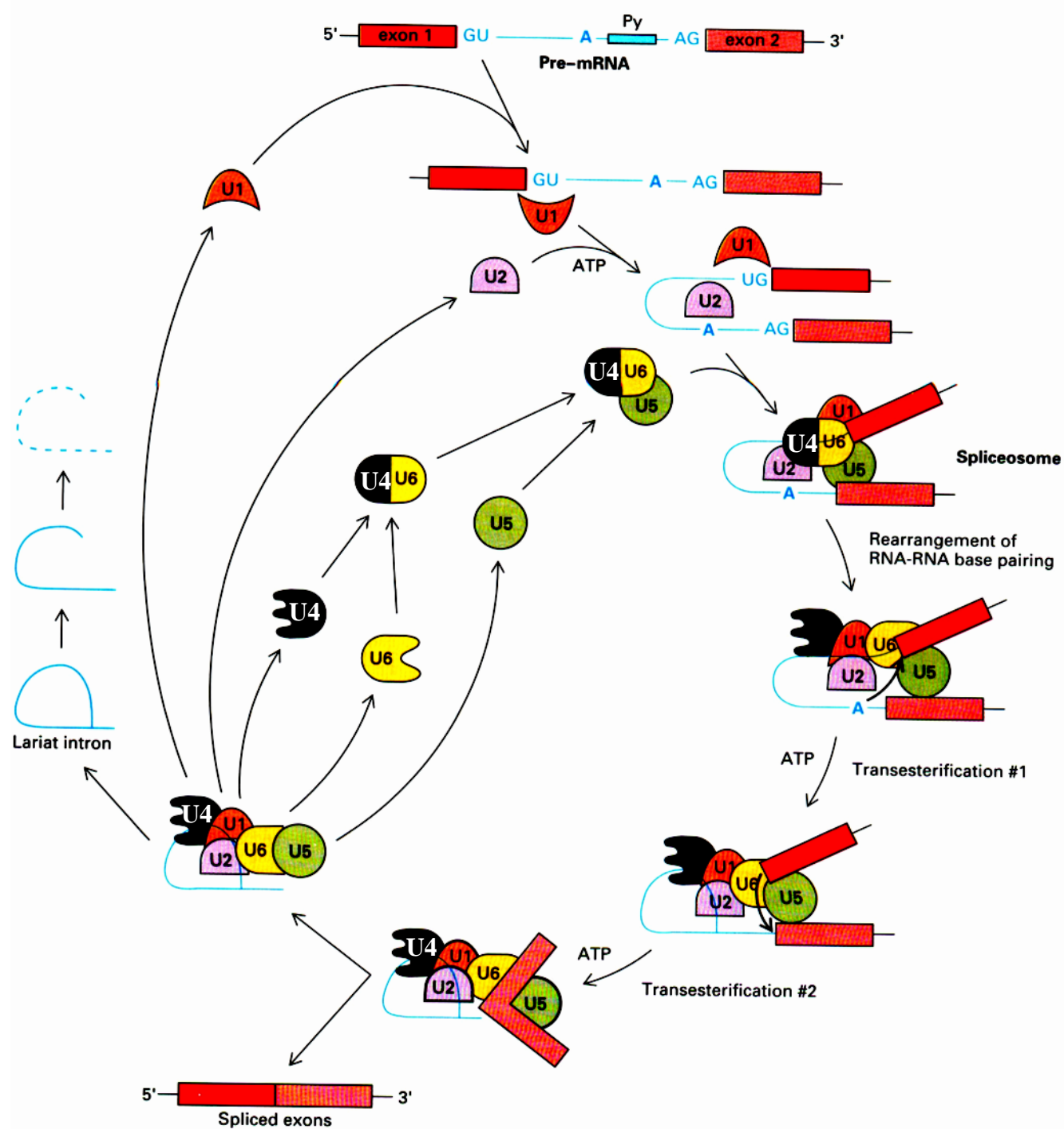
U6 complex with U4 creates transesterase of spliceosome

RNA



Spliceosom





# Alternative splicing

- Various protein forms from one gene

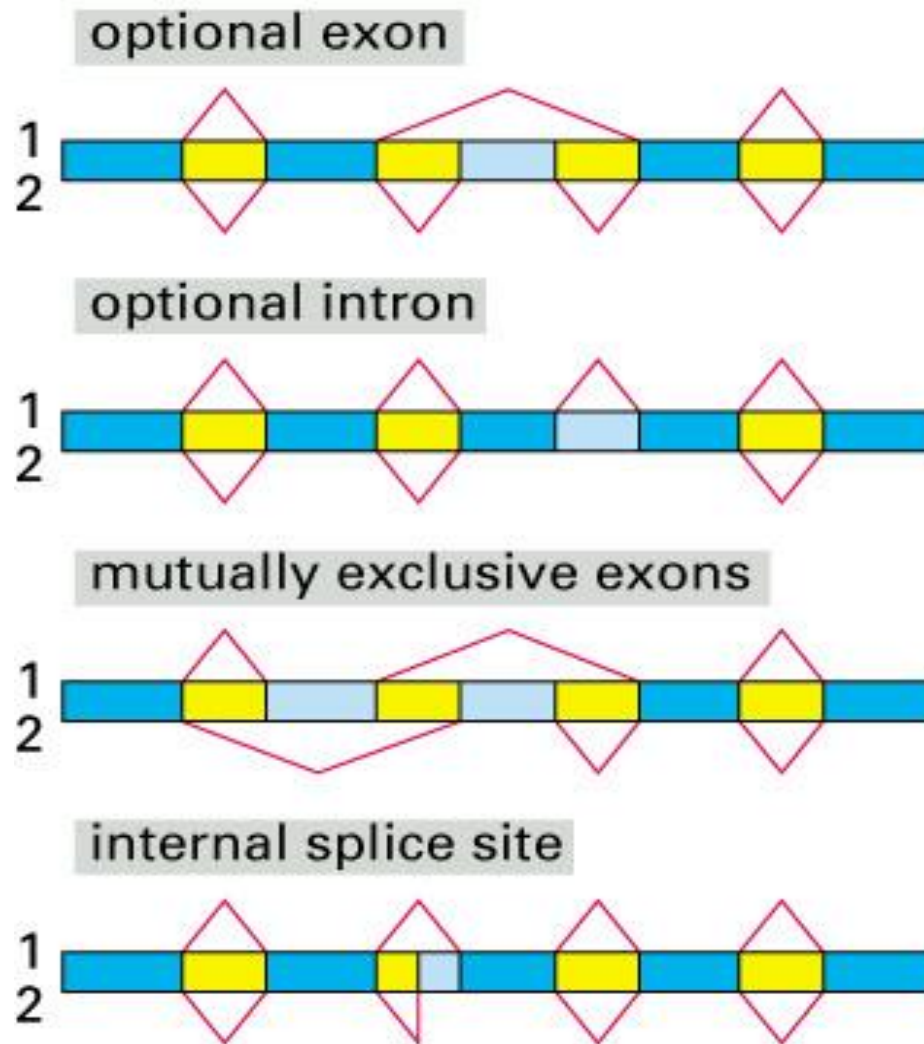


Figure 7-88. Molecular Biology of the Cell, 4th Edition.



# Alternative Pre-mRNA splicing – large diveristy

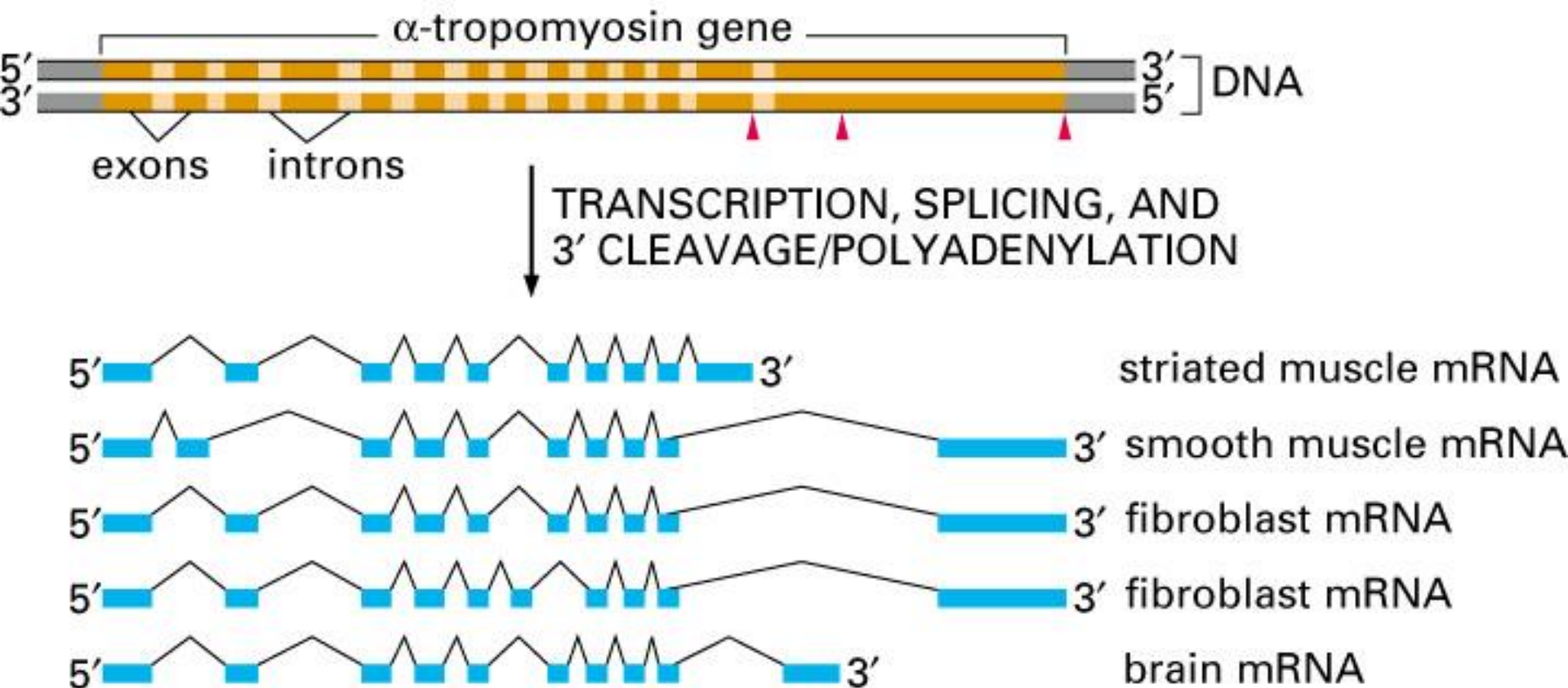


Figure 6-27. Molecular Biology of the Cell, 4th Edition.

## $\beta$ -thalassemia

- Autosomal recessive disease

Different types of mutations - deletions or point mutations

about 100,000 children a year ~ 10,000 in India

Not synthesized  $\beta$ -globin

- Dependence on transfusions and iron intake

Avg. Life expectancy well-treated patients, about 25 years



Mild untreated  $\beta$  thalassemia

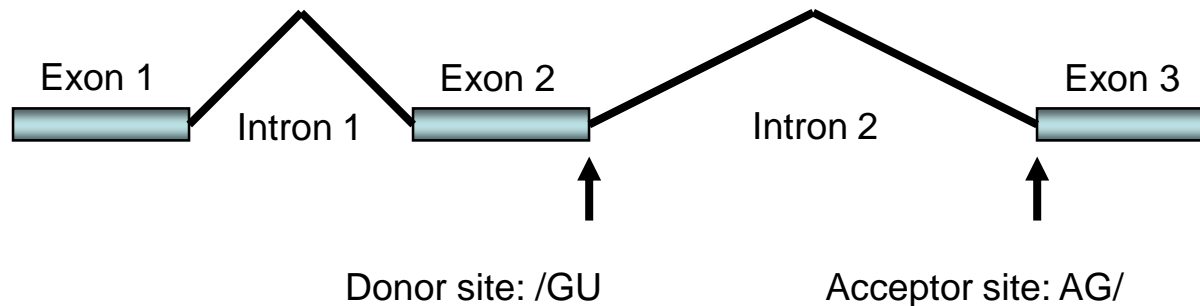


Thalassemia – chronic anemia (hematopoiesis in spleen and liver - enlarged)

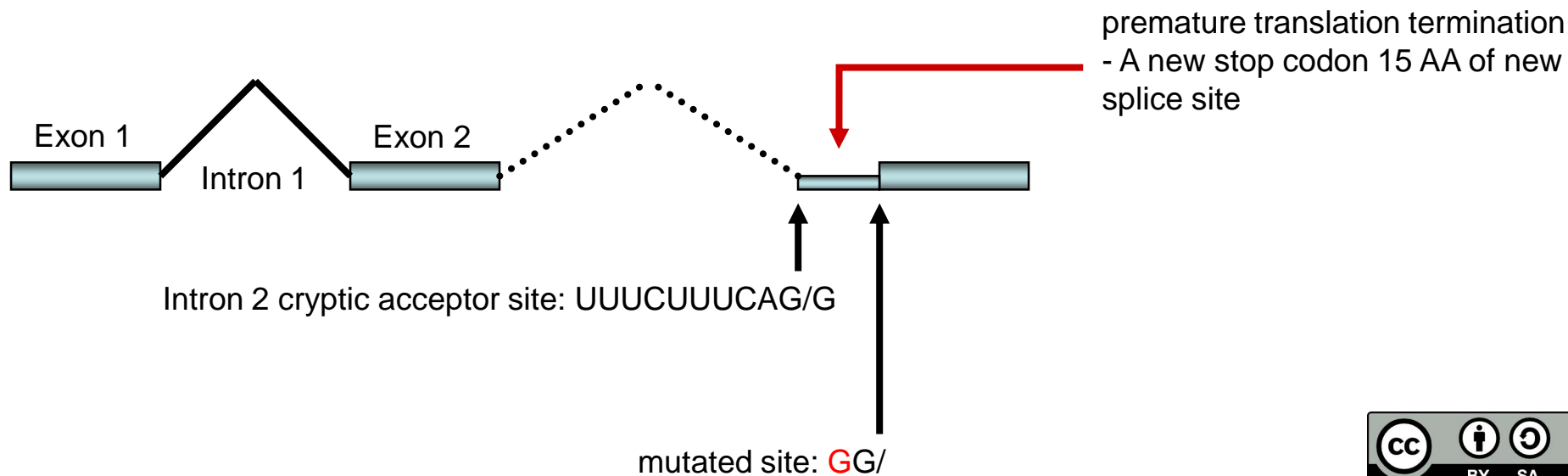
# Mutations disrupting splicing the gene for hemoglobin

$\beta^0$ -thalassemia - not synthesized  $\beta$  - chain  
 $\beta^+$  -thalassemia - weak synthesis of  $\beta$  -chain

Normal splicing:



**Intron 2  $\beta^0$  mutation of acceptor site - not functioning; use of cryptic sites in intron 2**  
**- Longer mRNA**

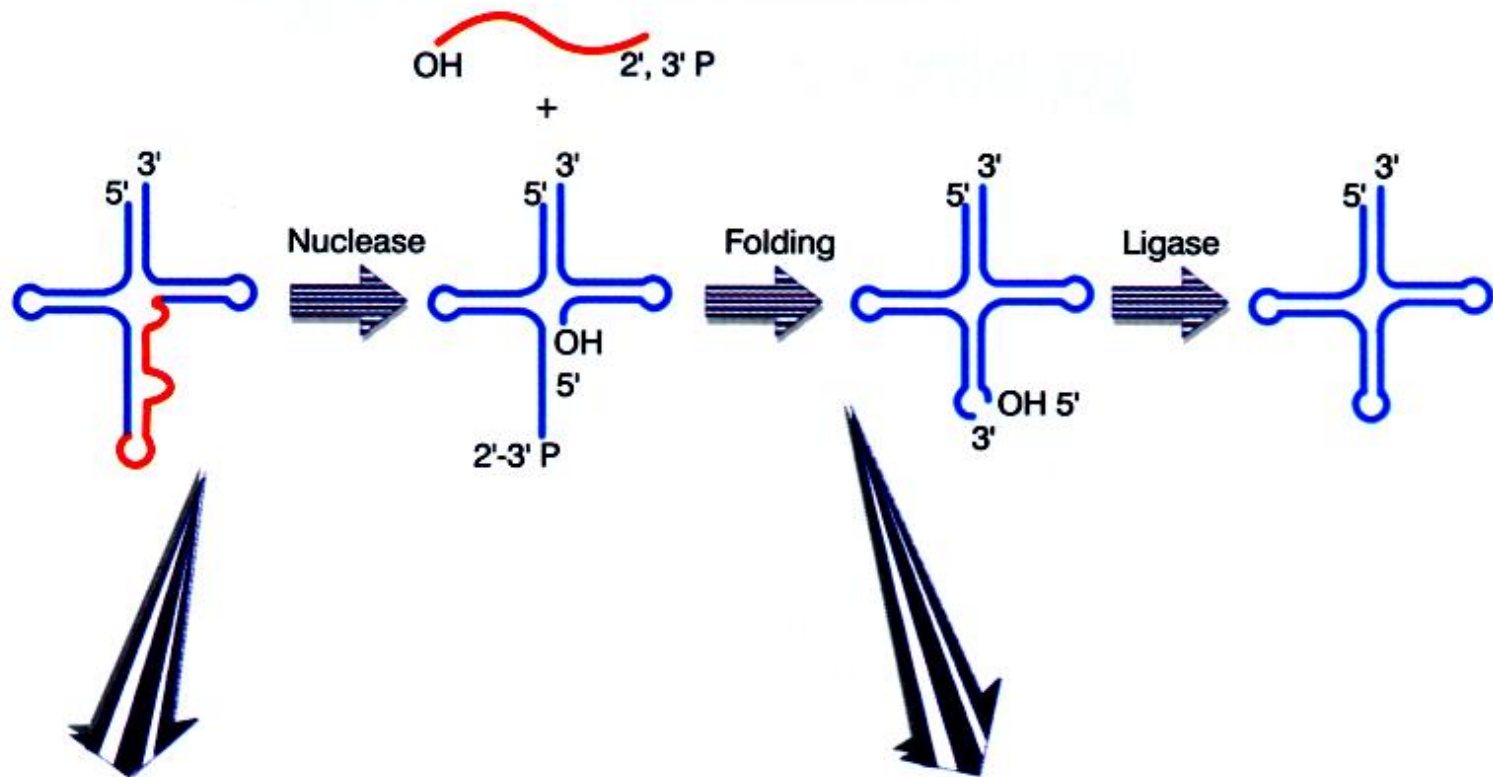


# Genetic code

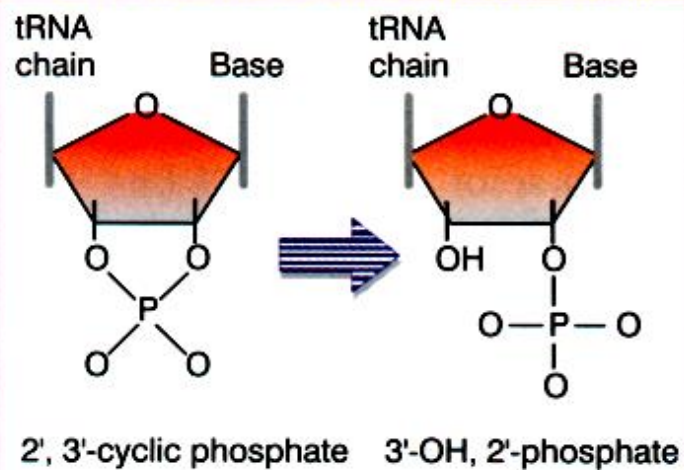
UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
UUC		UCC		UAC		UGC	
UUA	Leu	UCA		UAA	Stop	UGA	Stop
UUG		UCG		UAG		UGG	Trp
CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
CUC		CCC		CAC		CGC	
CUA		CCA		CAA	Gln	CGA	
CUG		CCG		CAG		CGG	
AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
AUC		ACC		AAC		AGC	
AUA		ACA		AAA	Lys	AGA	Arg
AUG	Met	ACG		AAG		AGG	
GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
GUC		GCC		GAC		GGC	
GUA		GCA		GAA	Glu	GGA	
GUG		GCG		GAG		GGG	



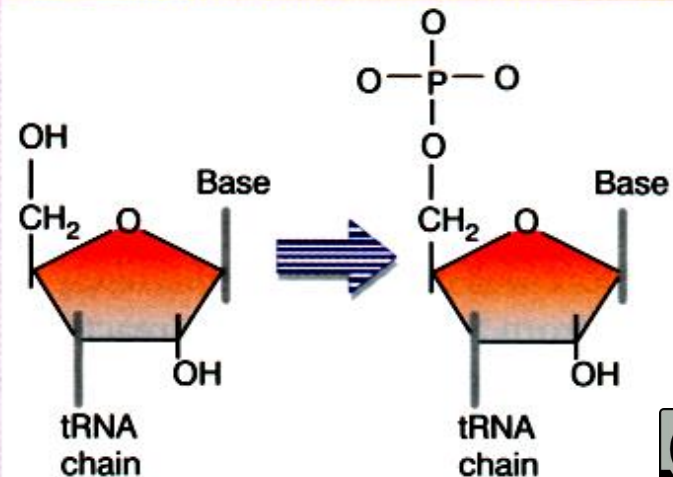




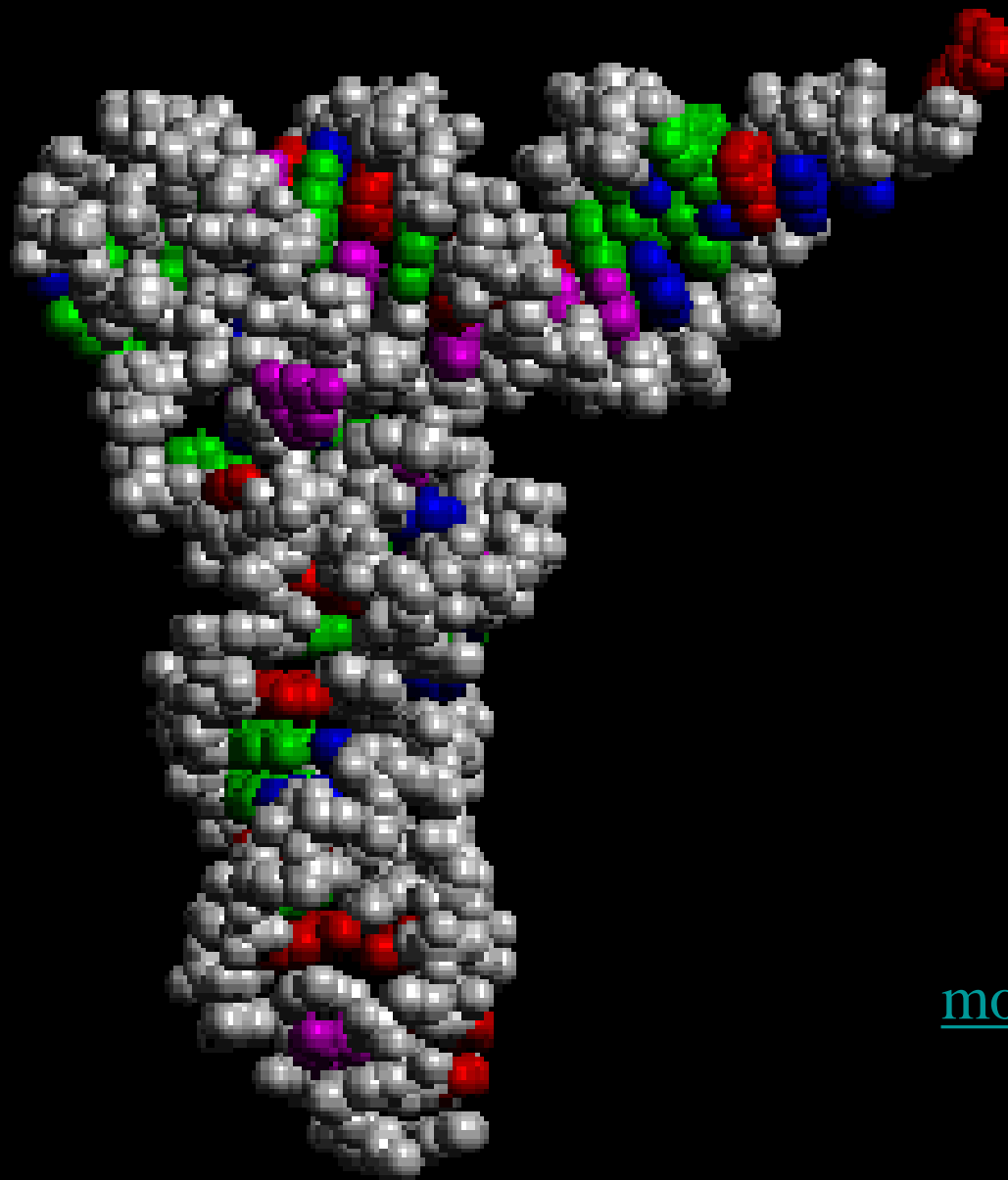
#### Phosphodiesterase opens phosphate ring



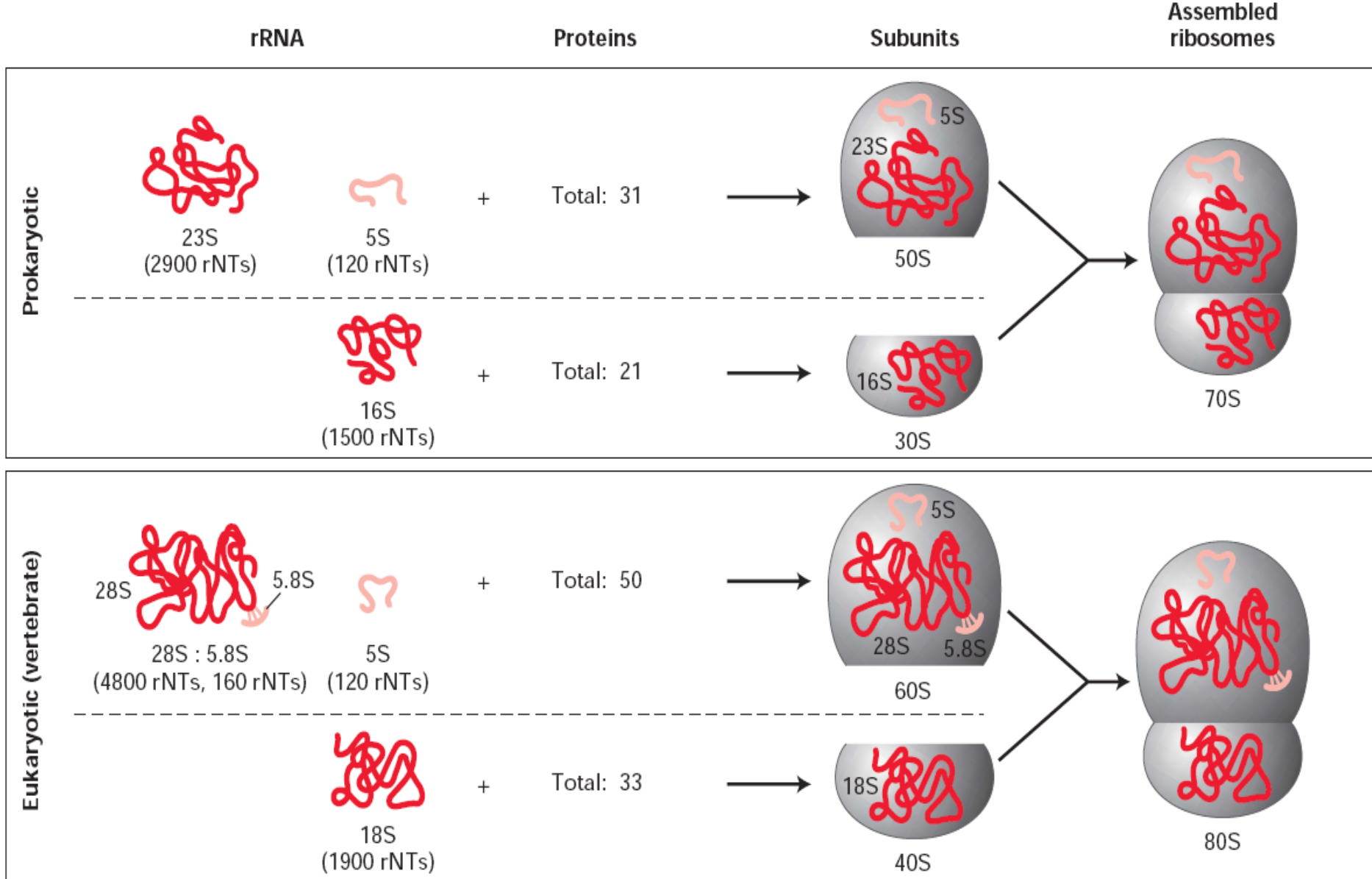
#### Kinase phosphorylates 5'-OH terminus



# tRNA

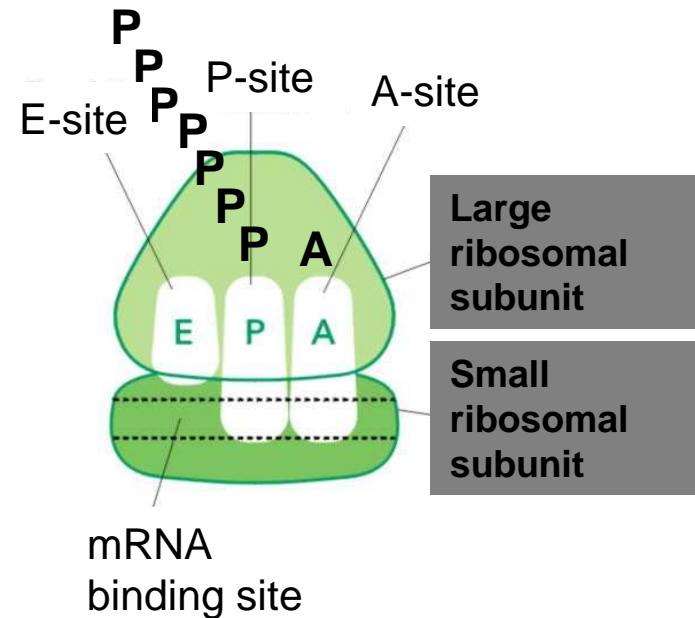


mov



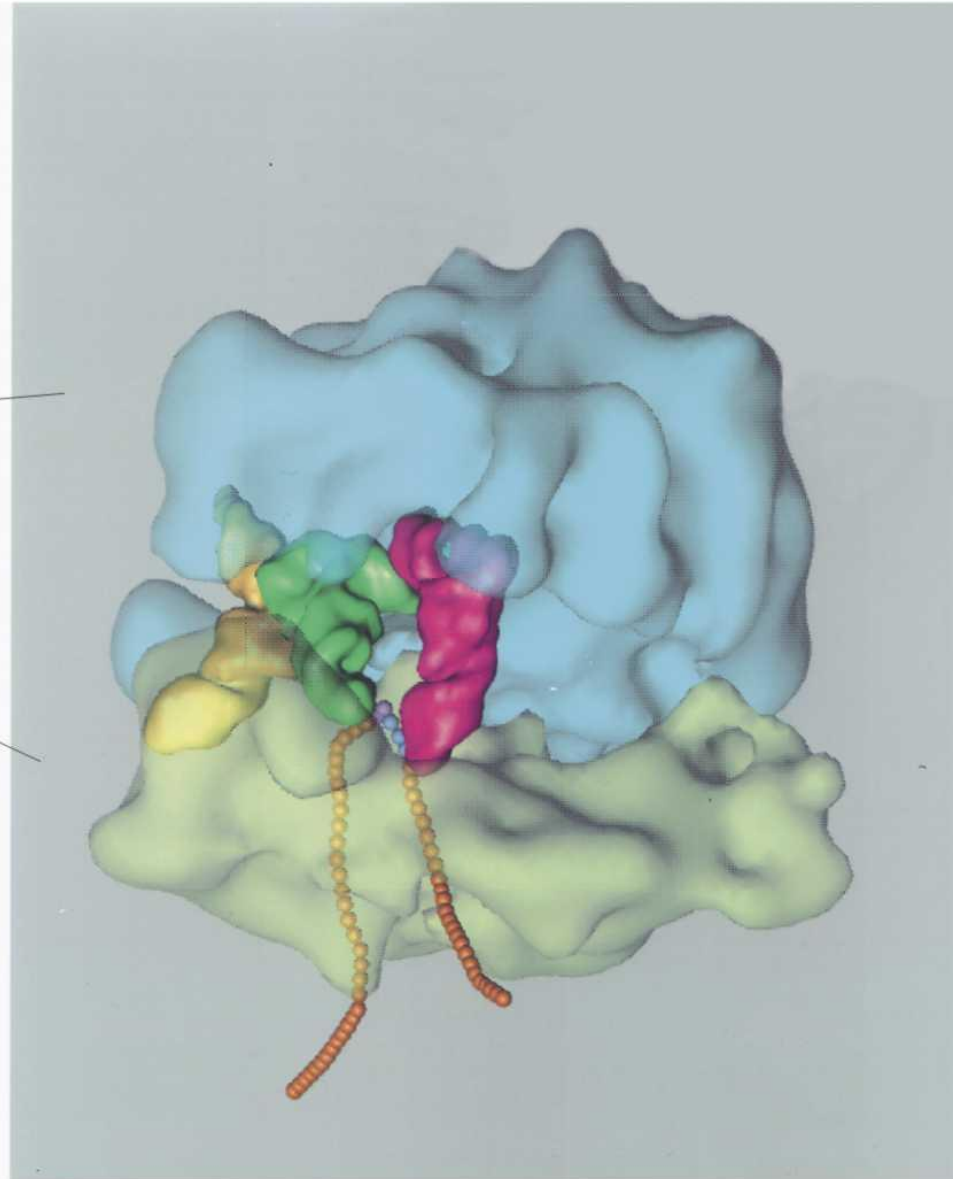
rRNAs – in different species – sequence variability – similar secondary structures  
 rRNA ~ 60% of molar weight of ribosomes

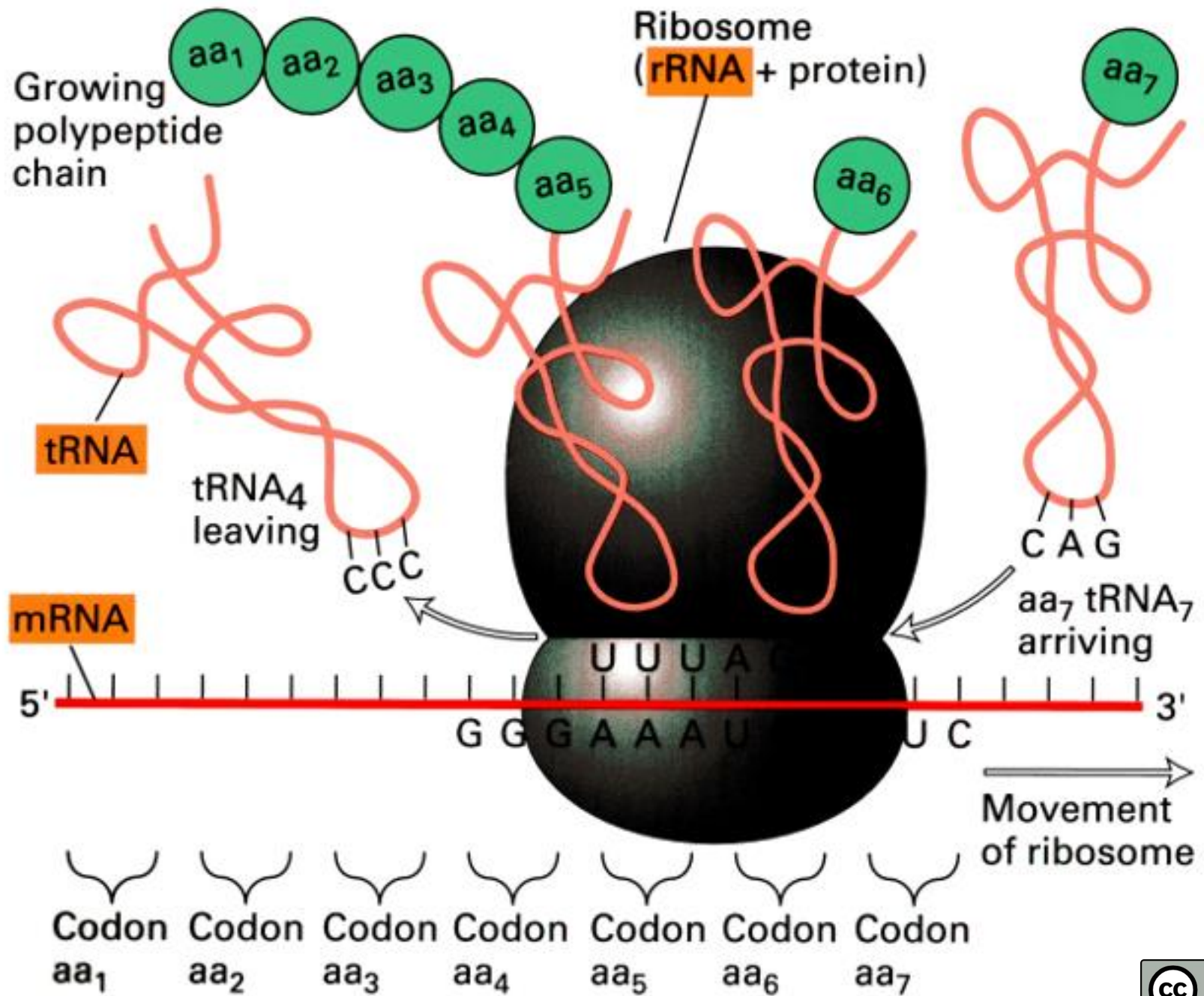
# Ribosome structure



**P-site**  
peptidyl tRNA site

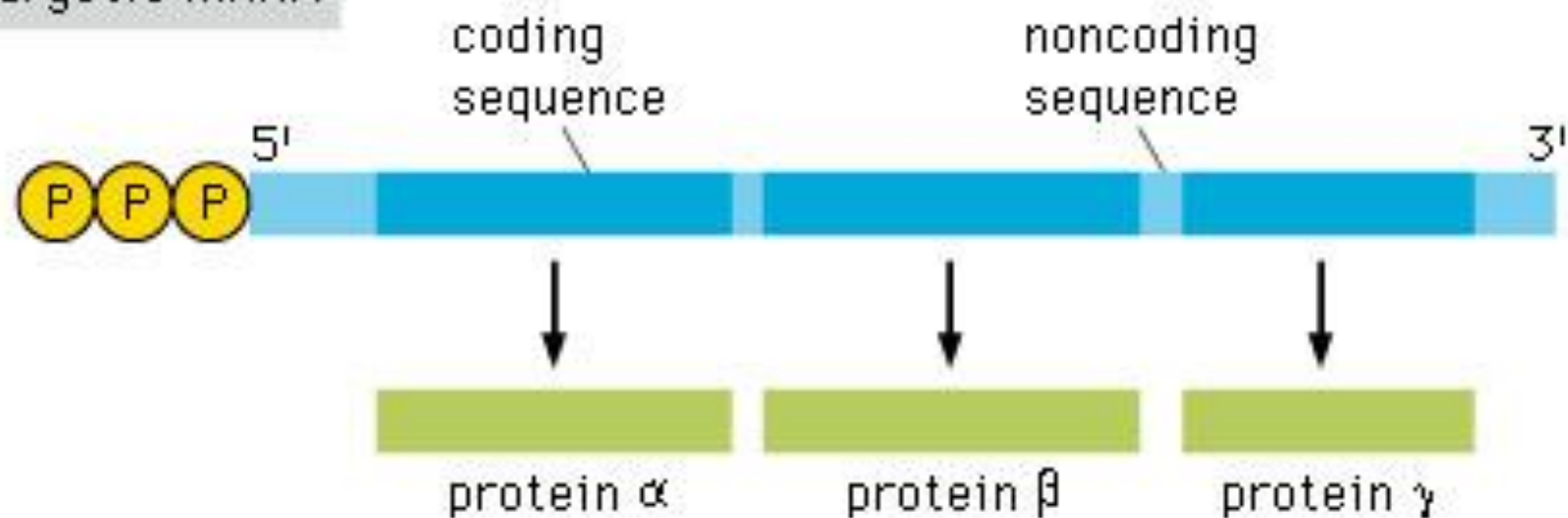
**A-site**  
aminoacyl tRNA site



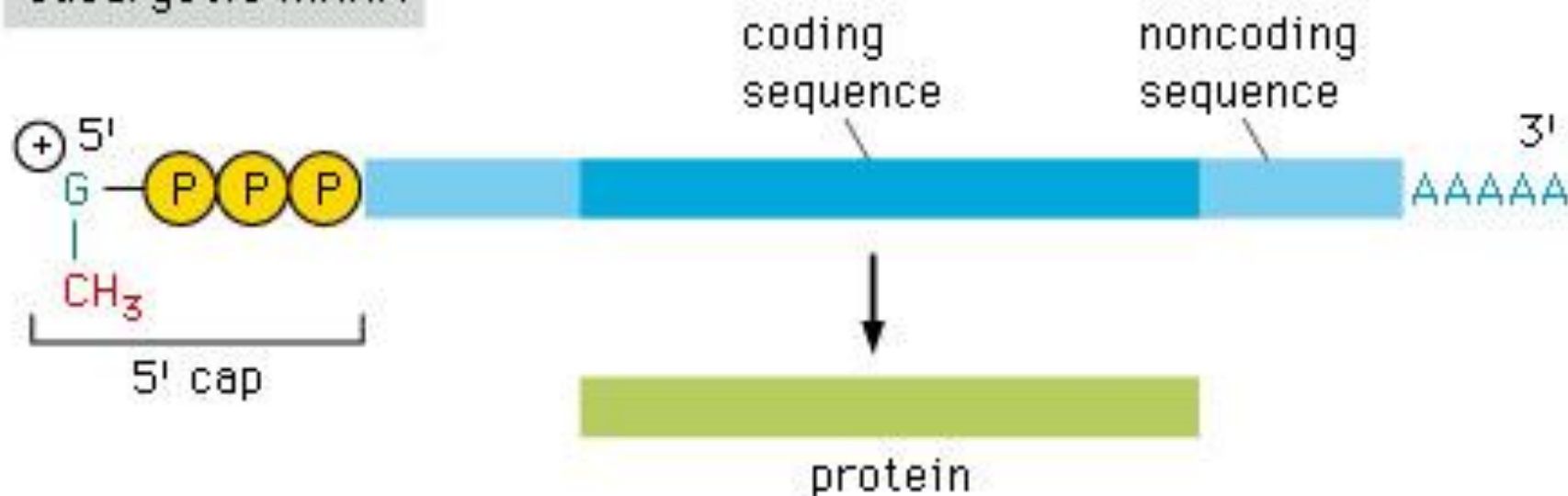




## procaryotic mRNA



## eucaryotic mRNA

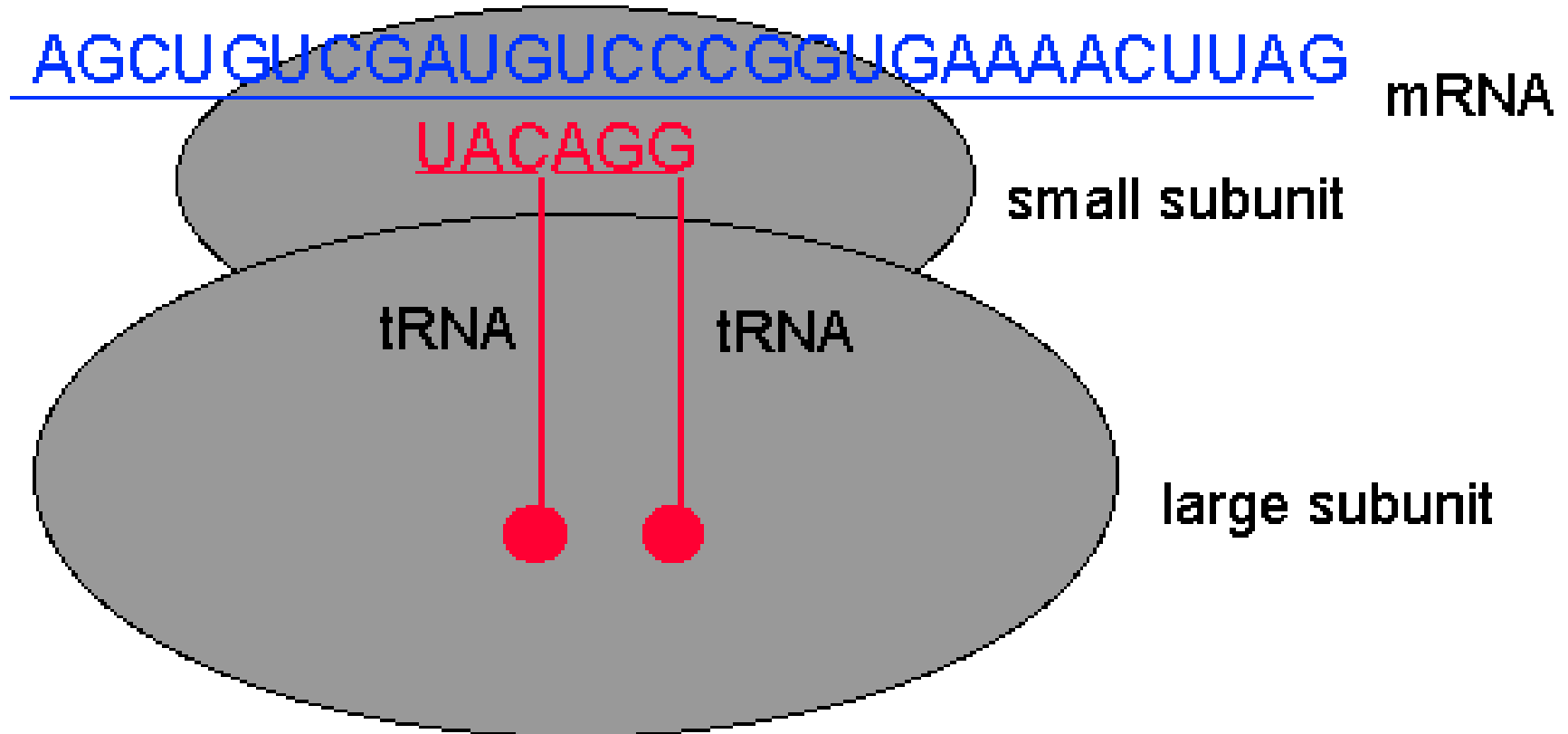




# Functions of ribosomal subunits

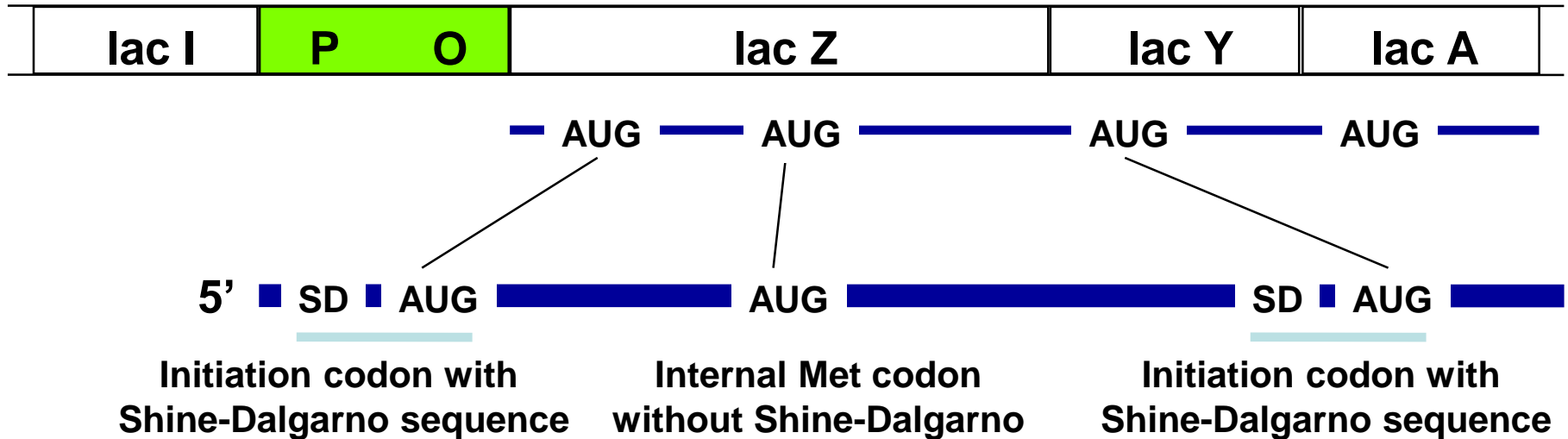
**Small subunit – codon-anticodon interaction**

**Large subunit – peptide bond formation**



# Initiation in prokaryotes

- also in internal AUG codons in mRNA
- lac operon in *E. coli* transcribed as polycistronic mRNA with multiple AUG codons



## Initiation in eukaryotes only at the first AUG codon

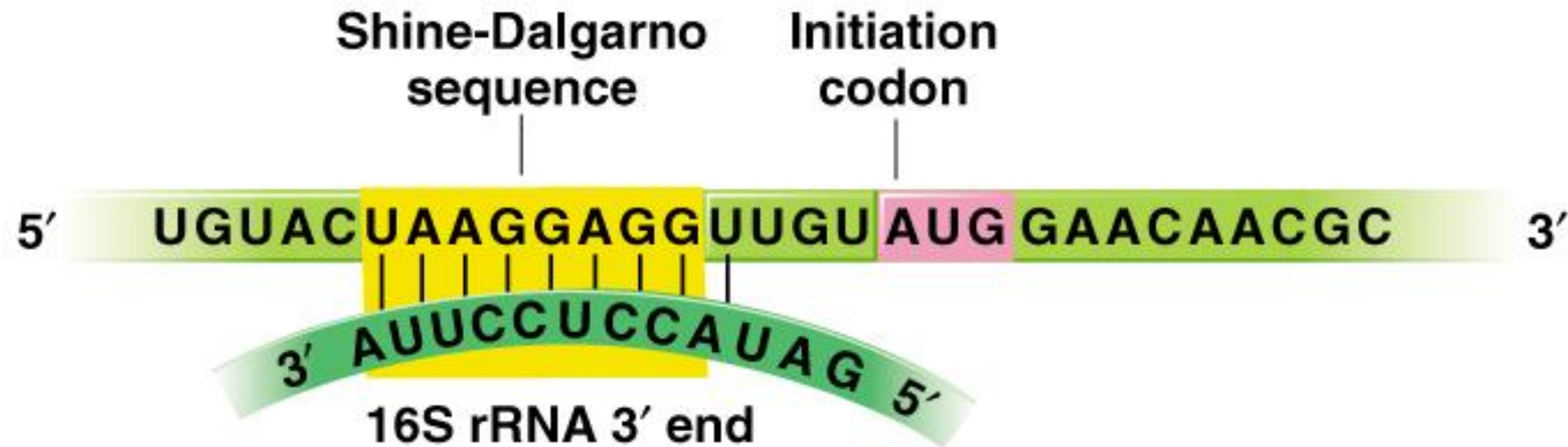


# Prokaryotic ribosome binding site in mRNA

## a) Sequence at 3' end of 16S rRNA

3' **AUUCCUCCAUAAG** 5'

## b) Example of mRNA leader and 16S rRNA pairing

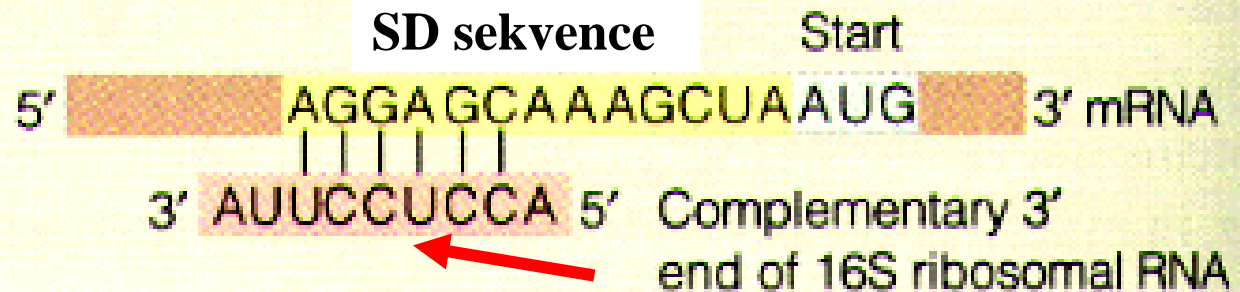


# mRNA Shine-Dalgarno

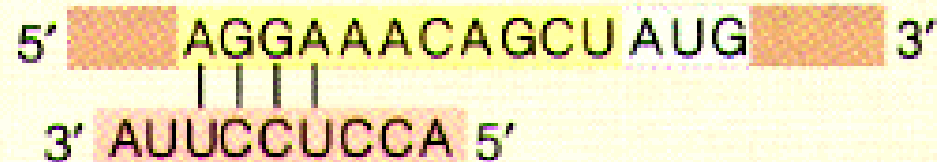
Message for

Shine-Dalgarno Sequence

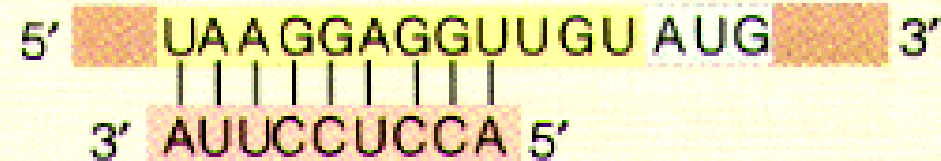
Ribosomal protein L10

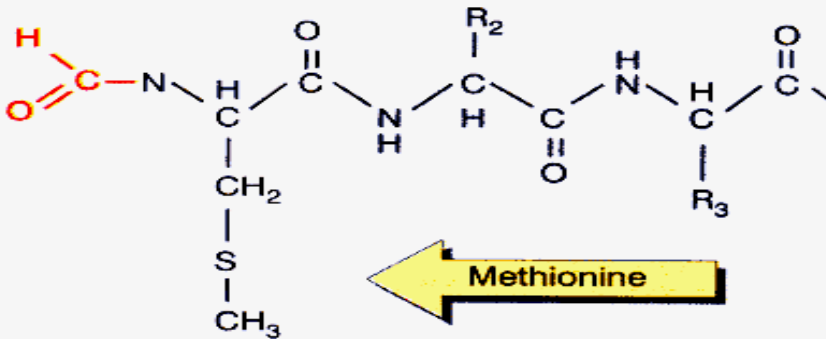
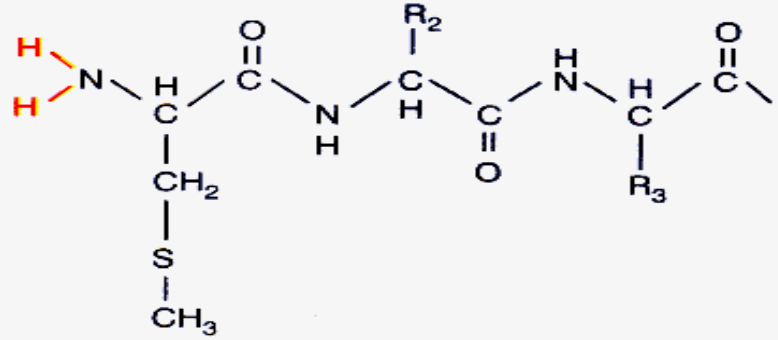
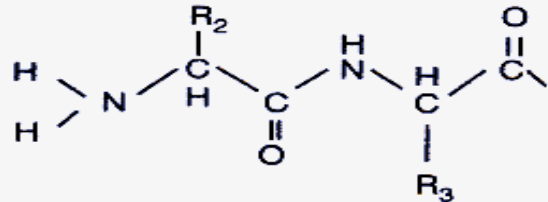


*E. coli lac z*

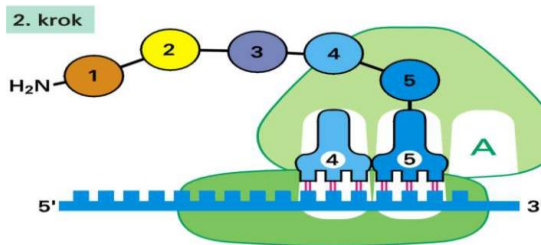
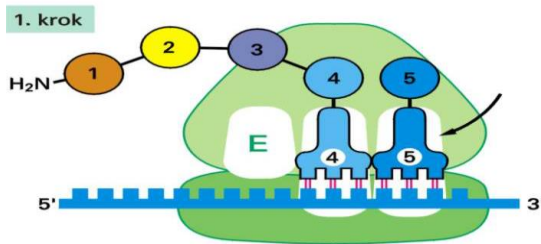
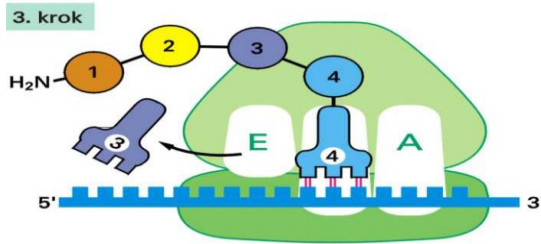
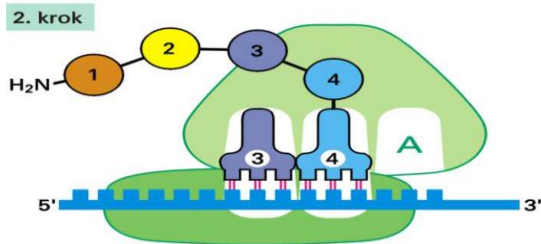
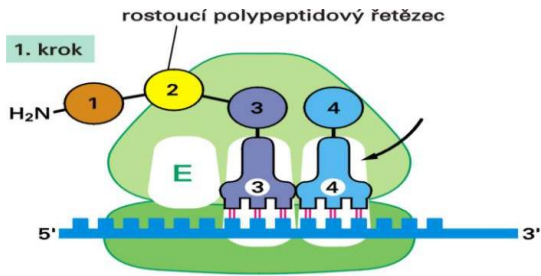


$\lambda$  phage Cro



Terminus	Structure
Formyl group	 <p>Chemical structure showing a peptide chain with a formyl group (H-C=O) attached to the N-terminus. The chain includes a methionine residue (R<sub>3</sub>) and an unknown residue (R<sub>2</sub>). A yellow arrow labeled "Methionine" points to the methionine residue.</p>
Deformylase ↓	
Amino group	 <p>Chemical structure showing the peptide chain after the formyl group has been removed. The N-terminus now has two hydrogen atoms (H) attached to the nitrogen.</p>
Aminopeptidase ↓	
R <sub>2</sub> amino acid	 <p>Chemical structure showing the released R<sub>2</sub> amino acid, which is an alpha-amino acid with two hydrogen atoms on the nitrogen and a carboxyl group.</p>

# Elongation



[mov](#)

[mov2](#)



Uveřejněné materiály jsou určeny studentům Vysoké školy chemicko-technologické v Praze

jako studijní materiál. Některá textová i obrazová data v nich obsažená jsou převzata

z veřejných zdrojů. V případě nedostatečných citací nebylo cílem autora/ů záměrně poškodit event. autora/y původního díla.

S eventuálními výhradami se prosím obraťte na autora/y konkrétního výukového materiálu, aby

bylo možné zjednat nápravu.

The published materials are intended for students of the University of Chemistry and Technology, Prague as a study material. Some text and image data contained therein are taken from public sources. In the case of insufficient quotations, the author's intention was not to intentionally infringe the possible author(s) rights to the original work.

If you have any reservations, please contact the author(s) of the specific teaching material in order to remedy the situation.

# Sequence analysis



EUROPEAN UNION  
European Structural and Investing Funds  
Operational Programme Research,  
Development and Education



MINISTRY OF EDUCATION,  
YOUTH AND SPORTS

# DNA isolation

## **Alkaline lysis**

lysozyme

chromosomal DNA, proteins and SDS precipitate during neutralization of basic solution by potassium phosphate

Removal by phenol/chloroform extraction

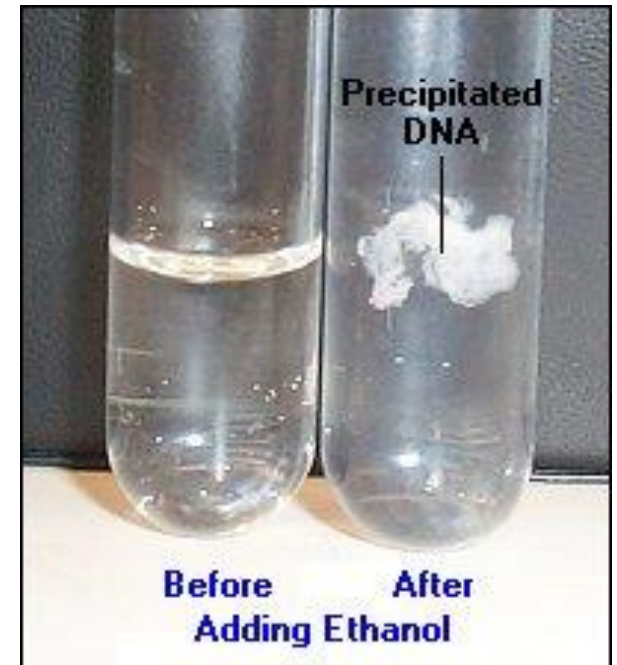
Plasmid DNA - ethanol or isopropanol precipitation

## **Boiling method**

lysozyme, Triton-X100, boiling.

Chromosomal DNA - attached to the membrane – pelleted together with denatured proteins

Plasmid DNA - ethanol or isopropanol precipitation





Centrifugation  
Vacuum



# **Centrifugation purification of plasmids**

**Plasmids – small, covalently closed, circular dsDNA**

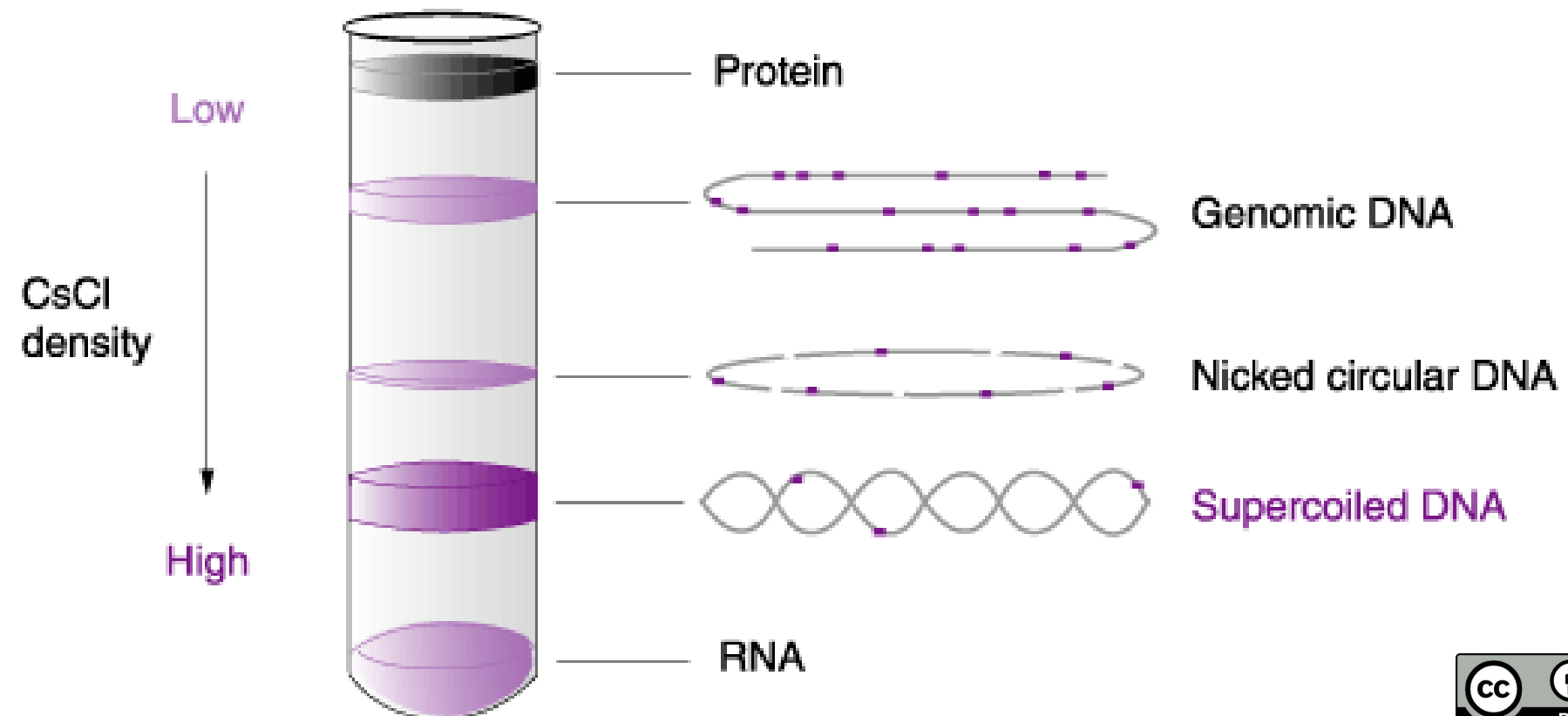
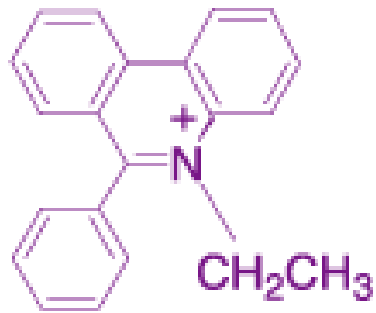
**Chromosomal DNA – large – broken into linear fragments**

**Results in different:**

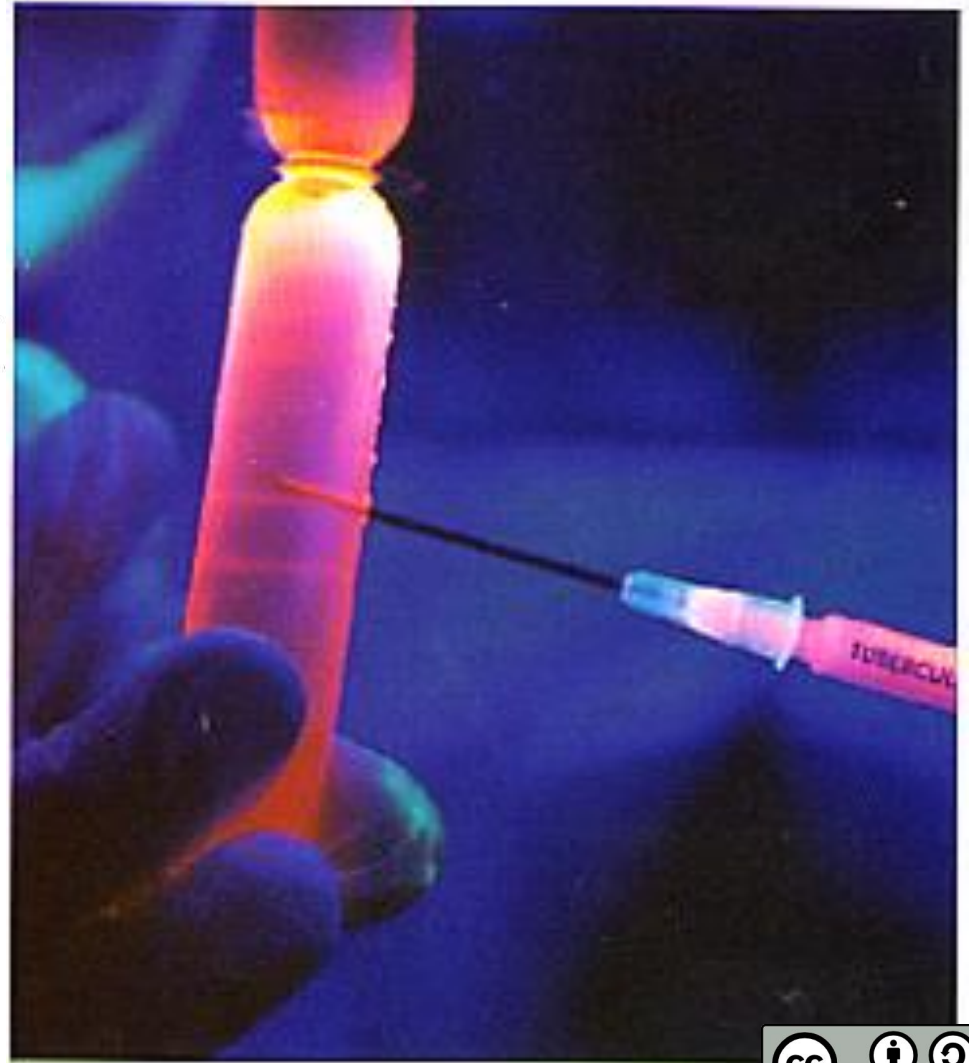
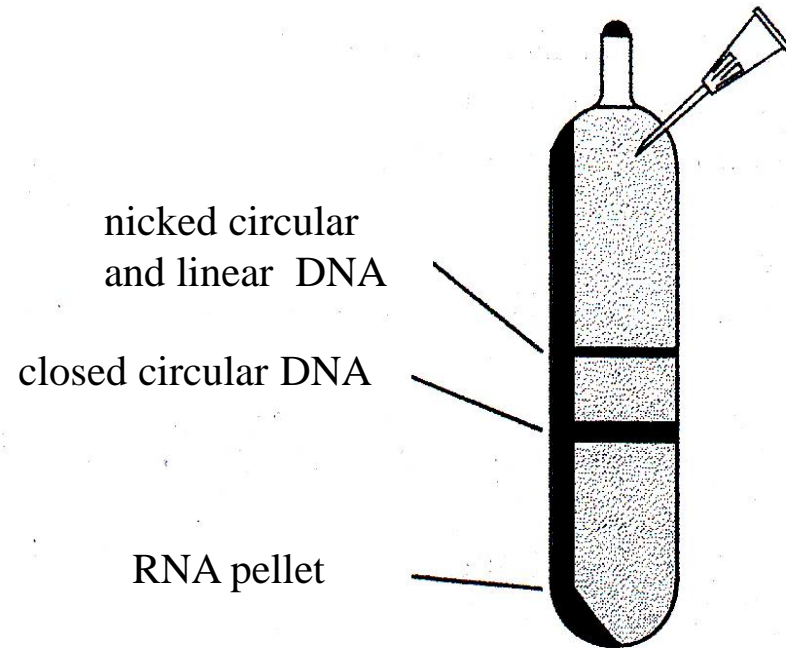
- binding of ethidium bromide,**
- different densities (CsCl - bands)**



# Ethidium Bromide



hypodermic needle



# Labeling of nucleic acids

- radioisotopes
- fluorescence markers
- biotin

Radionuclide half-life		Type/energy (MeV)	Spec. activity (Ci/mmol)
$^{32}\text{P}$	14.3 d	$\beta/1,71$	400-6000
$^{35}\text{S}$	87.4 d	$\beta/0,167$	400-1500
$^{125}\text{I}$	59.6 d	$\gamma/0,035$ $\beta/0,035$	1000-2000
$^3\text{H}$	12.4 y	$\beta/0,018$	25-100
$^{14}\text{C}$	5,730 y	$\beta/0,156$	

**Phosphate incorporation -  $[\gamma\text{-}^{32}\text{P}]\text{ATP}$  - polynucleotide kinase T4**

**Fill in recessive ends – DNA polymerase**

**Nucleotide incorporation – terminal transferase**

**Exchange reaction - excess of ADP**

**transfer of phosphate from the 5' end to ADP**

**re-phosphorylation by radioactive phosphate**

**from  $[\gamma\text{-}^{32}\text{P}]\text{ATP}$**

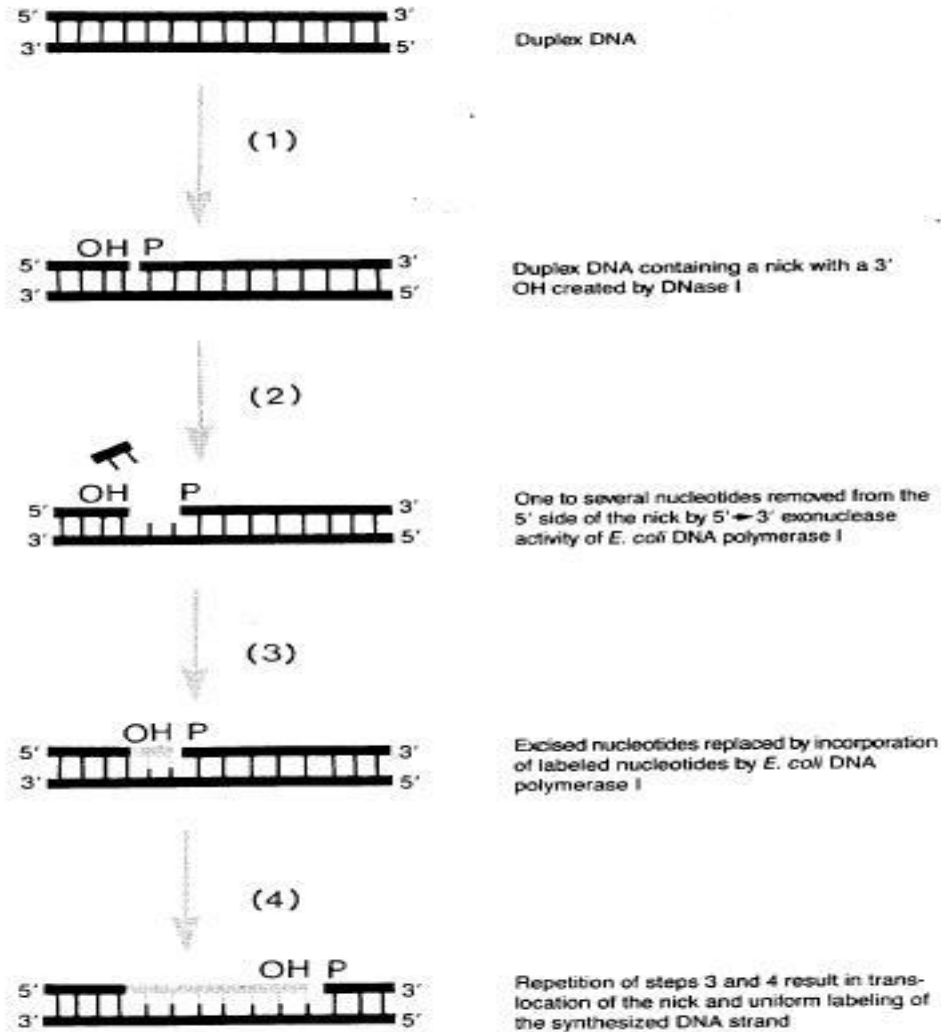
# **End labeling of nucleic acids**

3' and 5' ends ss and ds DNA, RNA and oligonucleotides

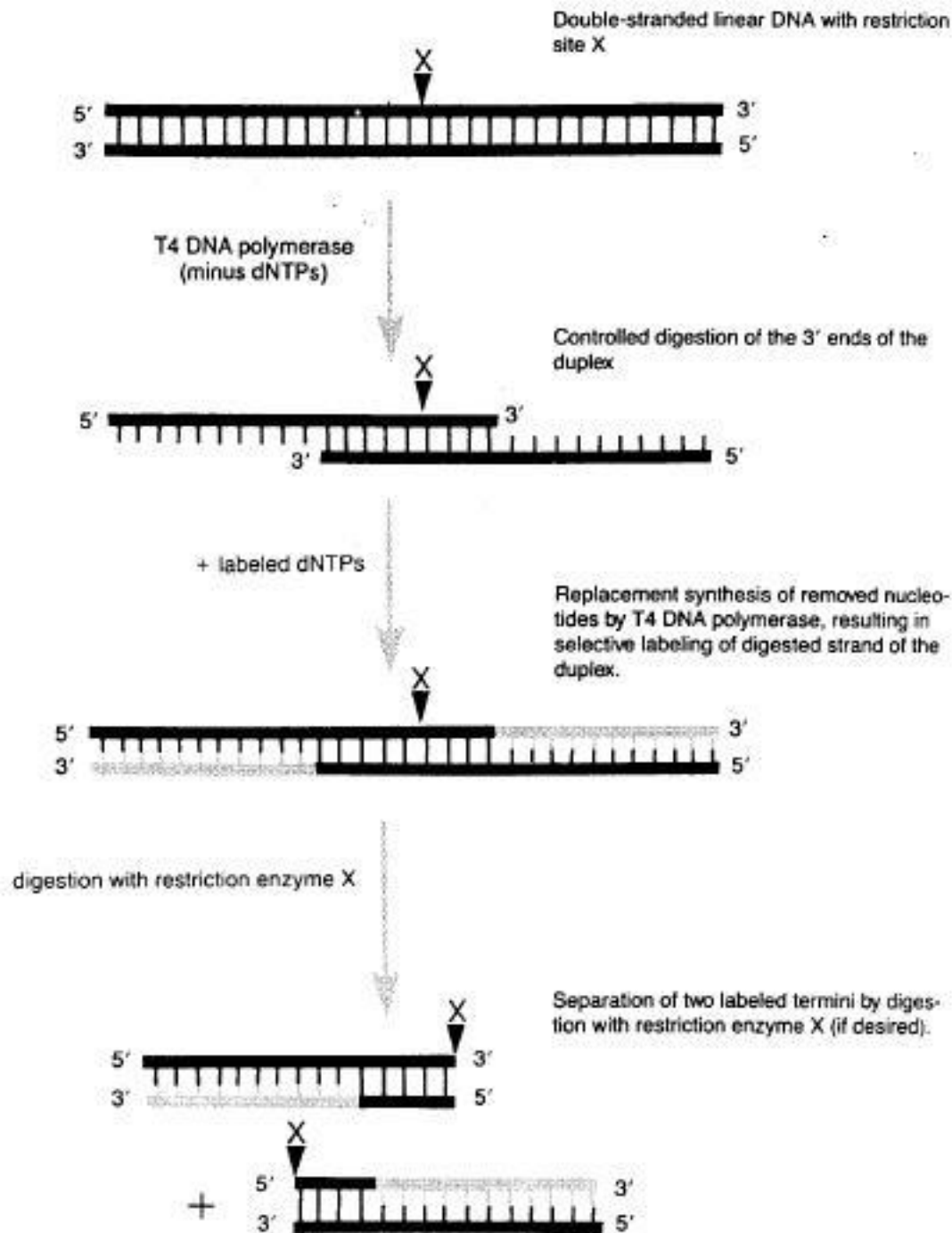
# Nick translation

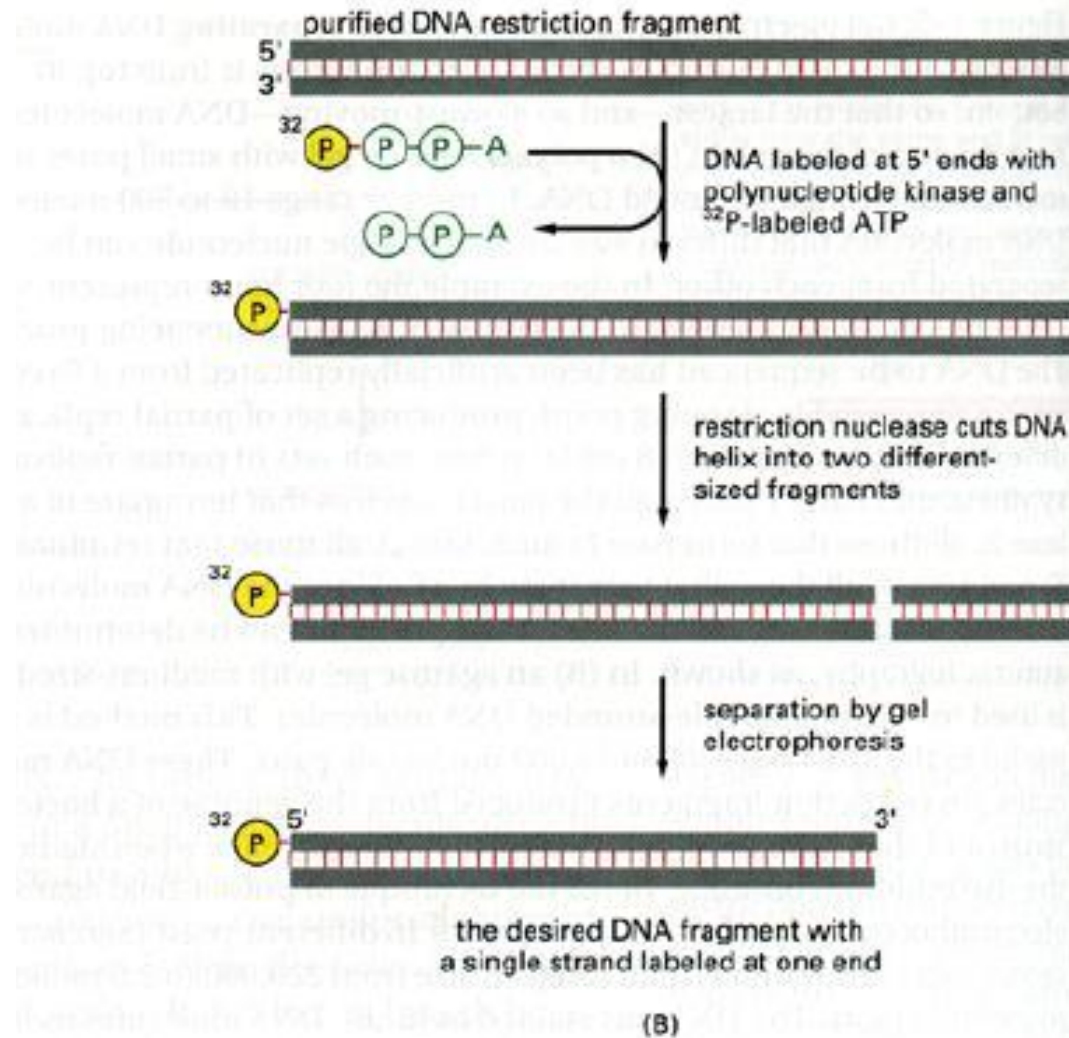
Nonspecific breaks by endonuclease,

Break extension by 5'-3' exonuclease activity of DNA pol. I *E. coli*  
synthesis by 5'-3' polymerase activity DNA pol. I *E. coli*





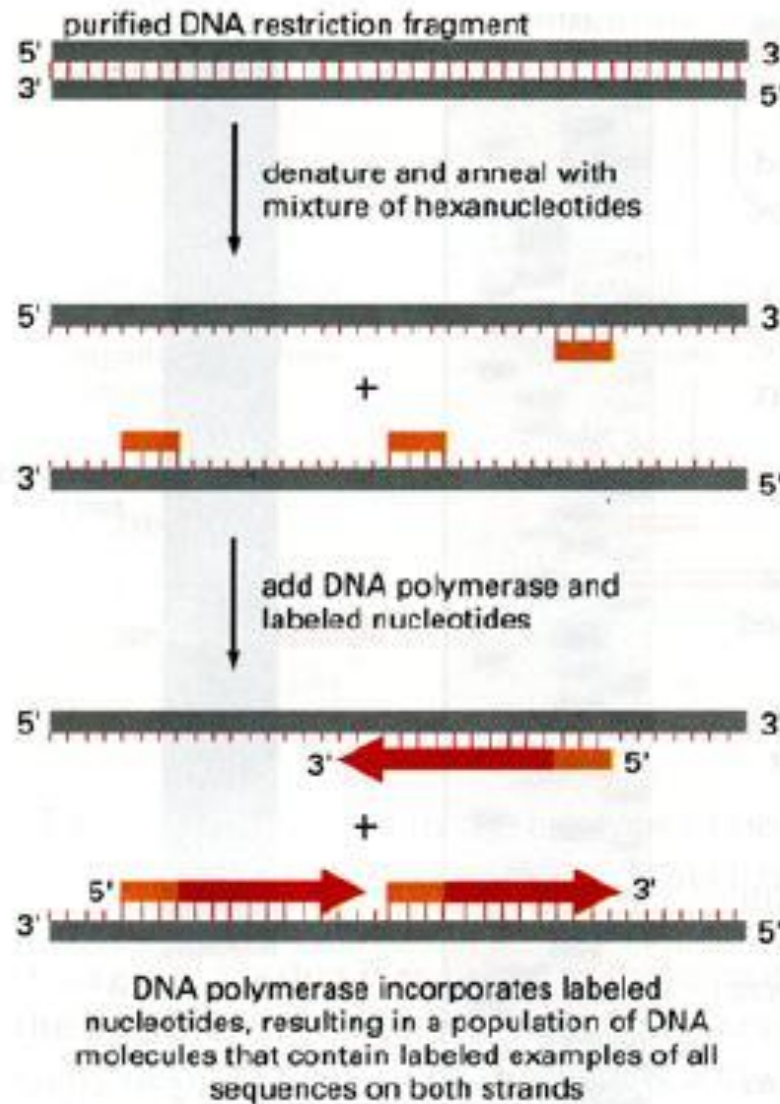




5' ends

alkaline phosphatase - free hydroxyl

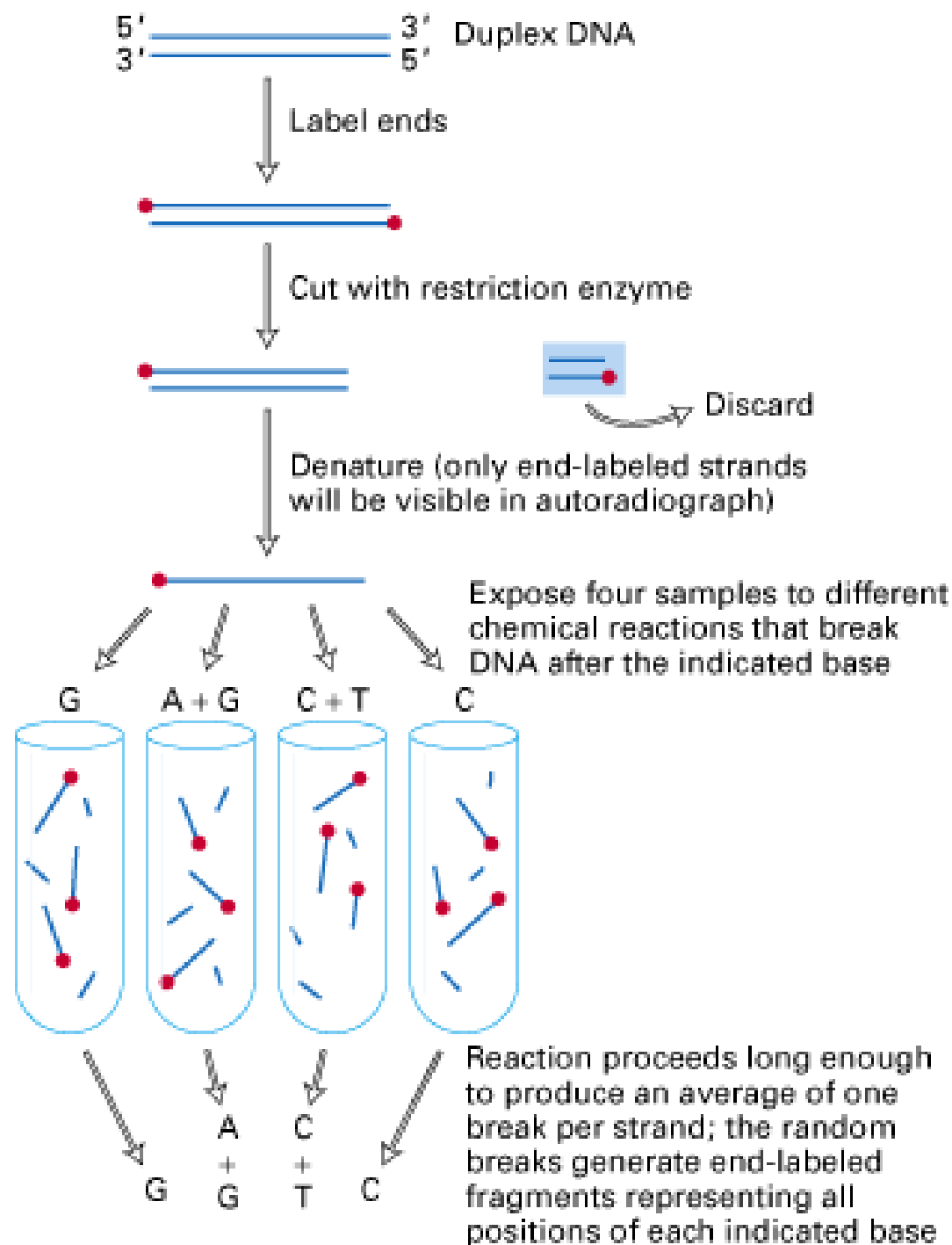
phosphate incorporation - [ $\gamma$ -<sup>32</sup>P]ATP - polynucleotide kinase T4



(A)

# **DNA sequencing**

## **Maxam-Gilbert method - chemical**

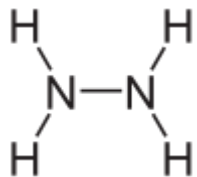


**G:** DMS methylates N<sup>7</sup> G – opens between C<sup>8</sup> a N<sup>9</sup>.  
Piperidine removes modified G from saccharide moiety.

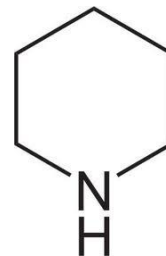
**G + A:** formic acid protonates purine ring,  
weakens glycoside bonds in A and G.  
Purines – removed by piperidine.

**T + C:** Hydrazine breaks rings in T and C.  
Fragments - removed by piperidine.

**C:** hydrazine, in presence of NaCl reacts only with C.  
Modified C – is removed by piperidine.

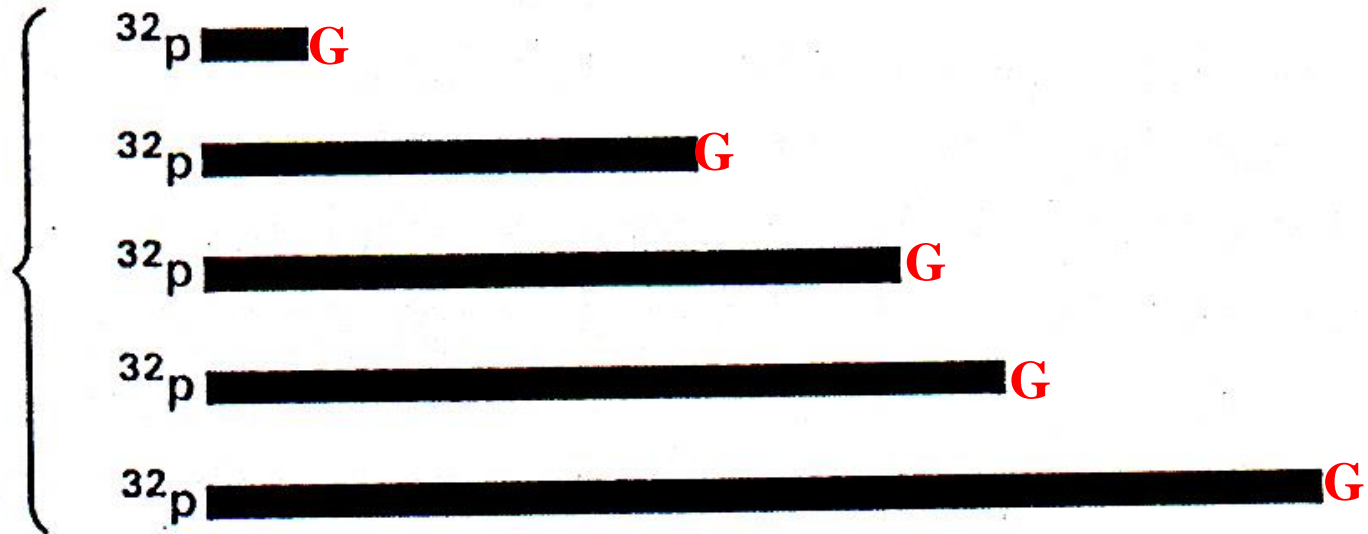


hydrazine



piperidine

Labeled fragments





# **DNA sequencing**

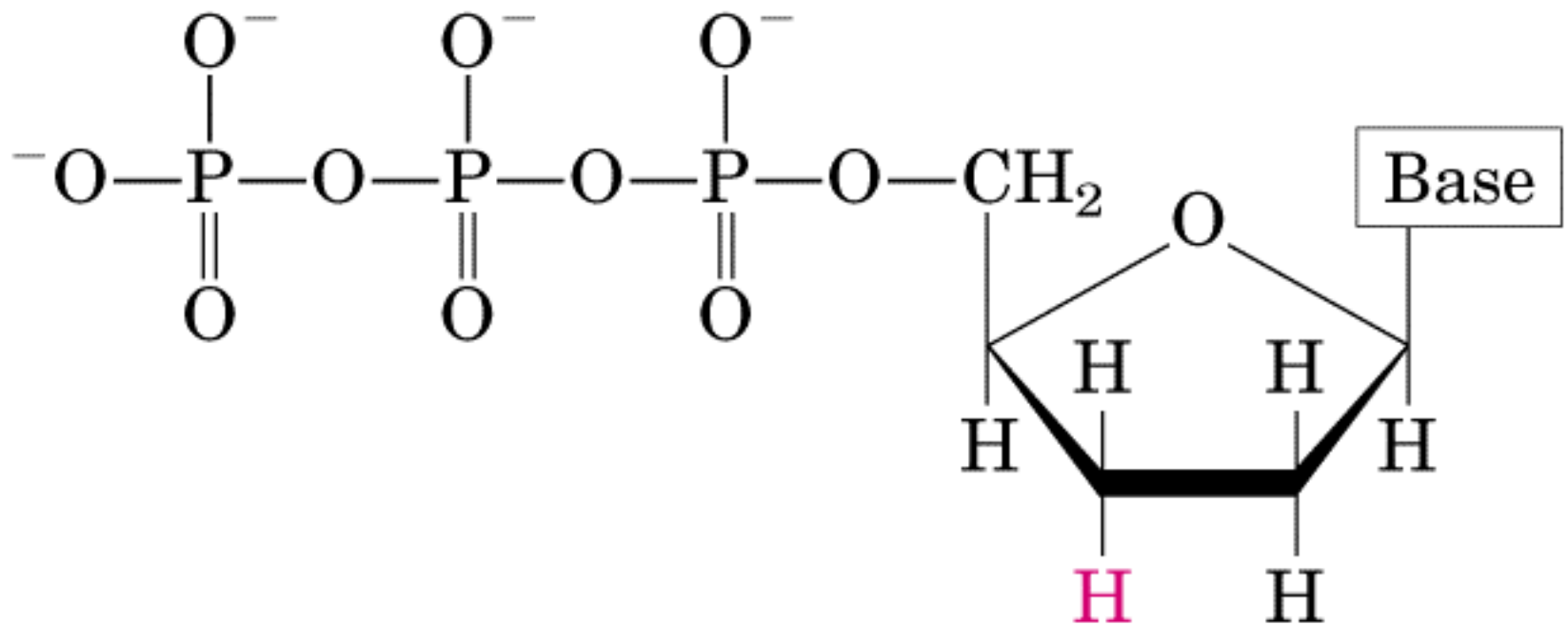
**Sanger method**

**DNA polymerase**

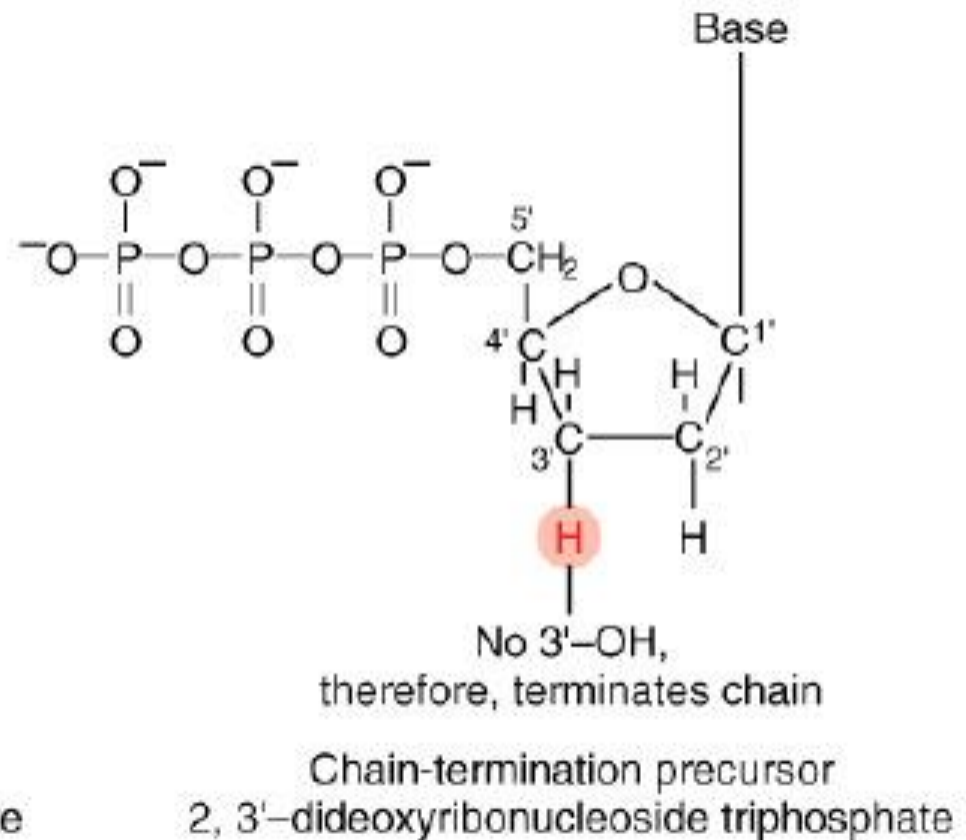
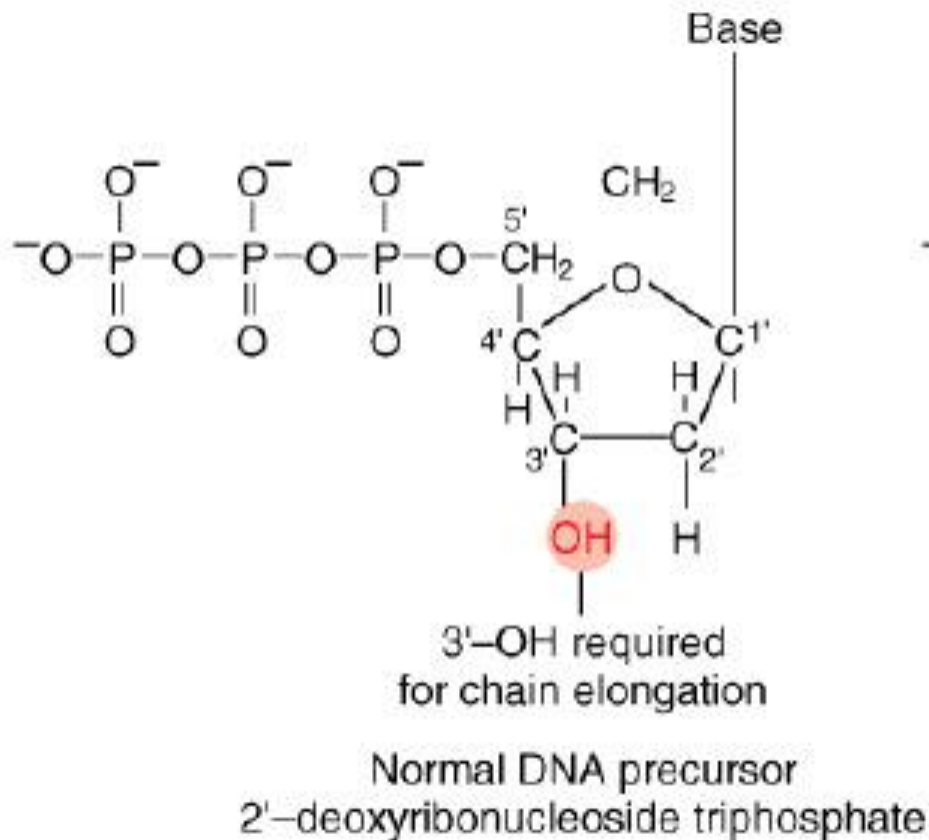
**dNTP**

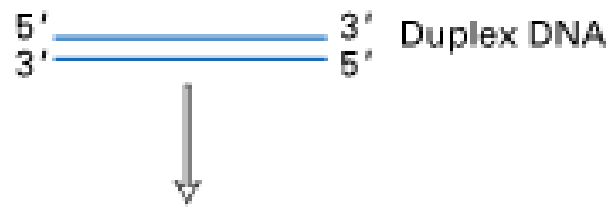
**ddNTP**

**primer**

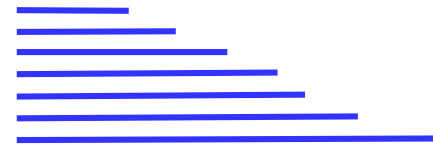
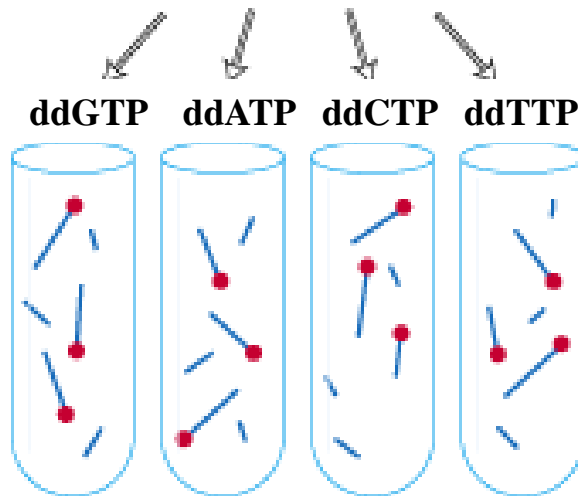


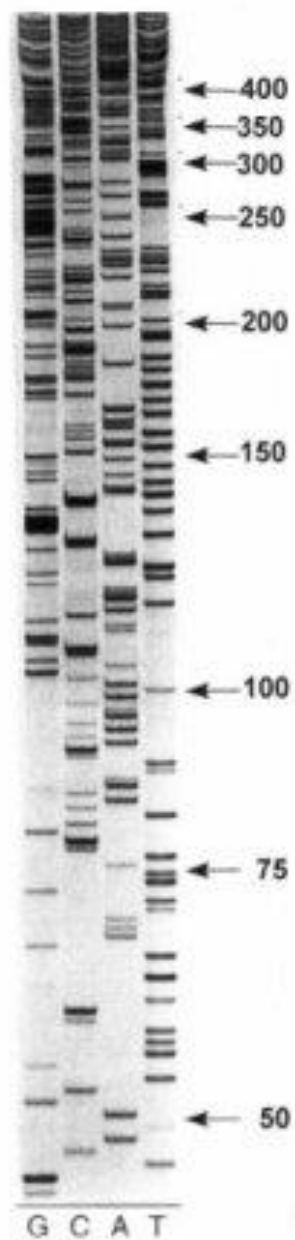
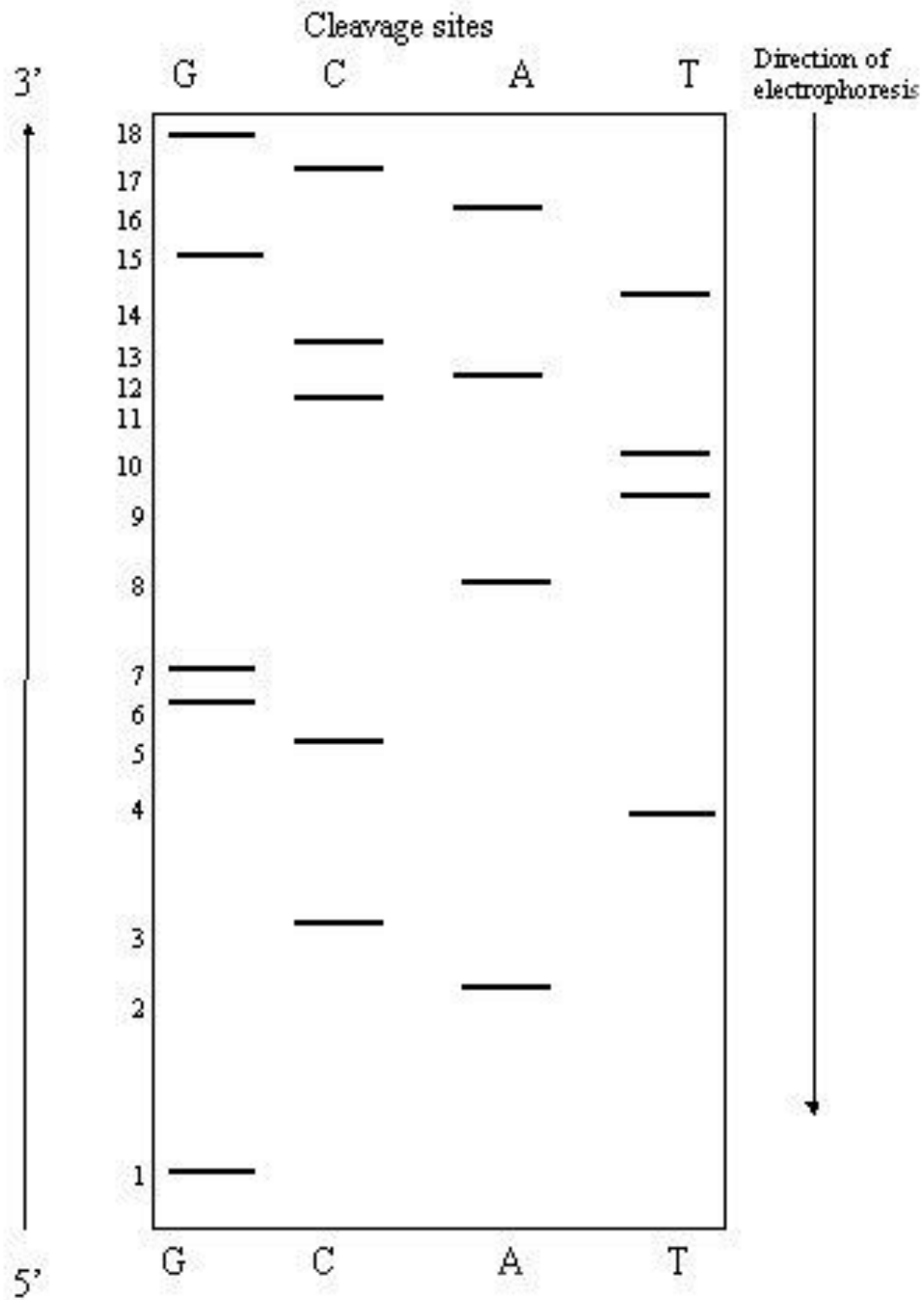
**ddNTP analogue**





**DNA polymerase**  
**All dNTPs + labeled dNTP**  
**primer**

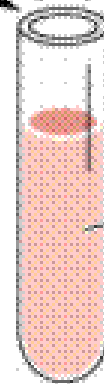




5' TTAGACCCGATAAGCCCGCA 3'

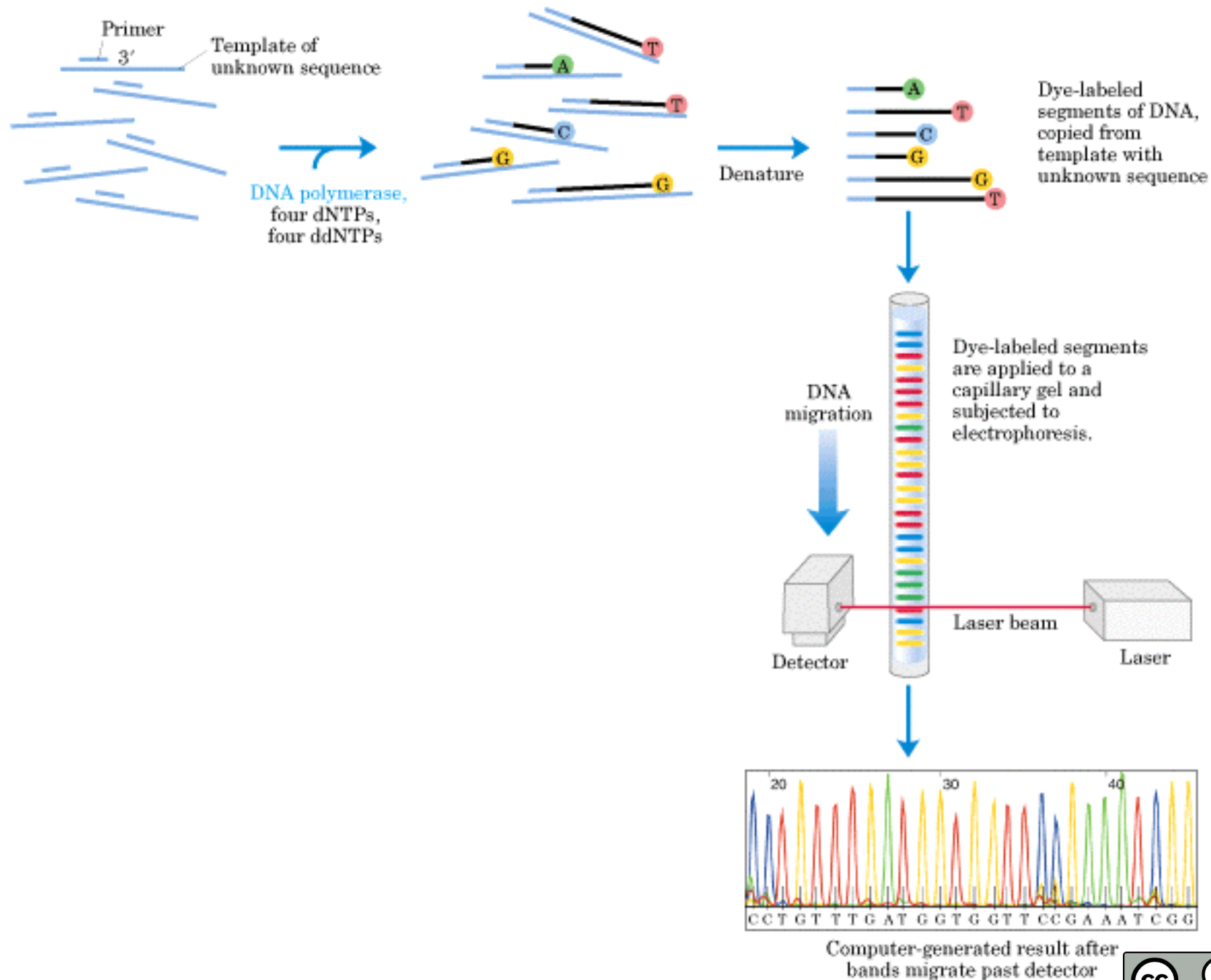
ddCTP ddGTP ddTTP ddATP

DNA sequence  
to be determined

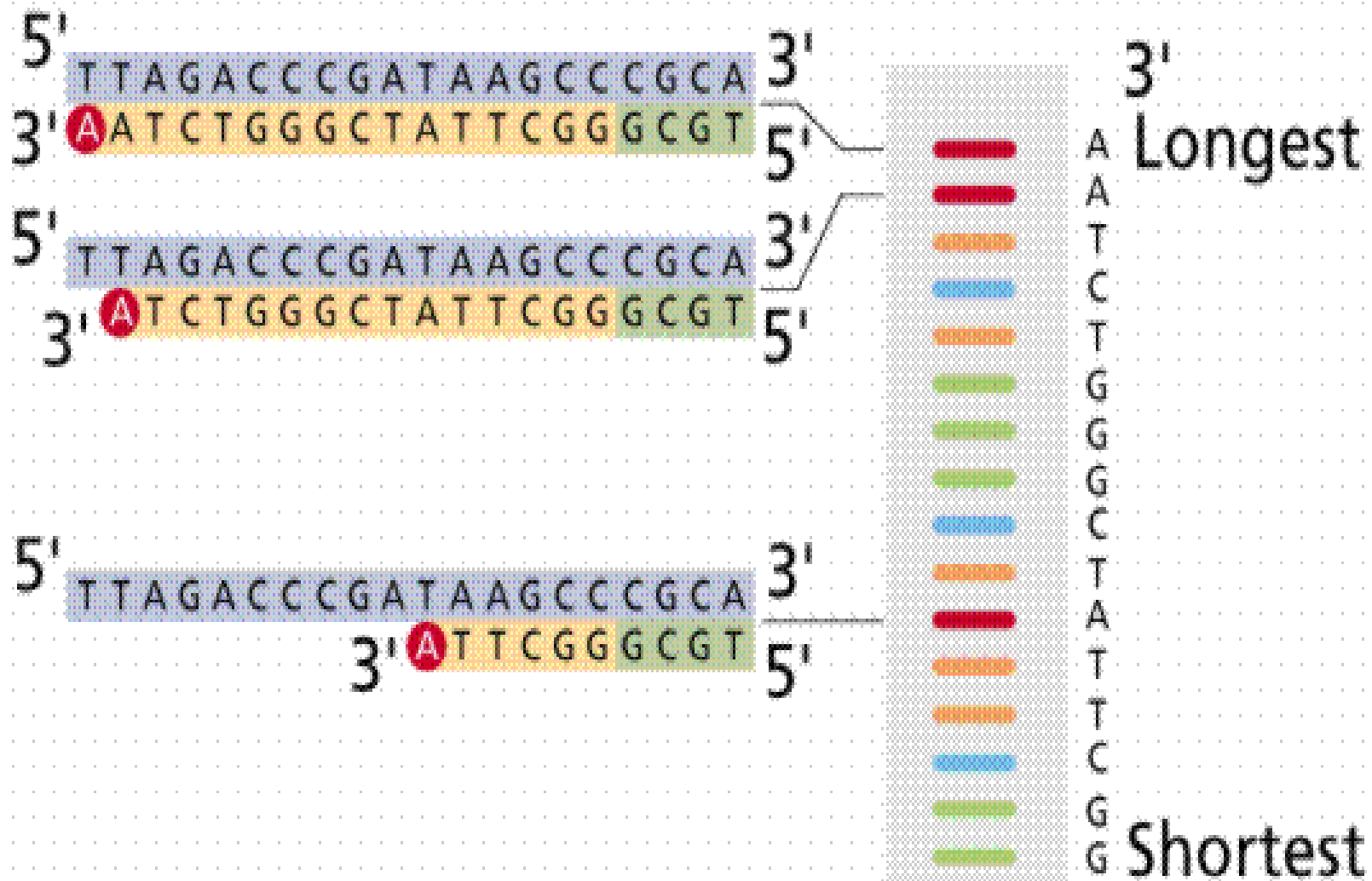


Analysis of  
reaction mixture  
for A shown  
below

5' TTAGACCCGATAAGCCCGCA 3'  
3' GCGT 5' Primer







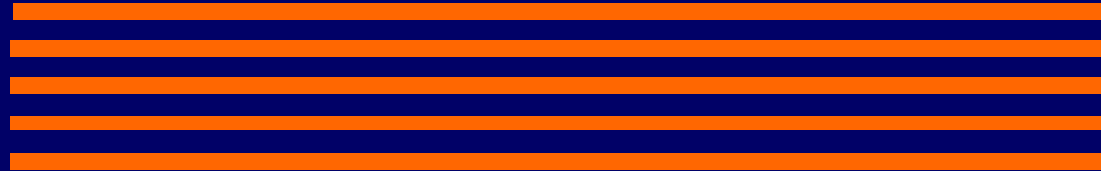
# Subclone Construction

```
GATCGTCTAGAGATCTC
GAGACTCTCTGAGAGTTC
GTGGGAAACTGTGTGA
TTTGACTAGCCCAAGT
TACGTTGTAGAGATGT
ATGATGCACCTGACCC
GAGTTTCACCTCTCAGC
GACTCACTCCAGCTCA
GAGGCCACCGGCGCT
GTGCAGTCCAGCCAC
GATTATTACCATTTTA
ATCCTTAGAGATTGACA
```

**BAC DNA**



**Prepare Multiple Copies**



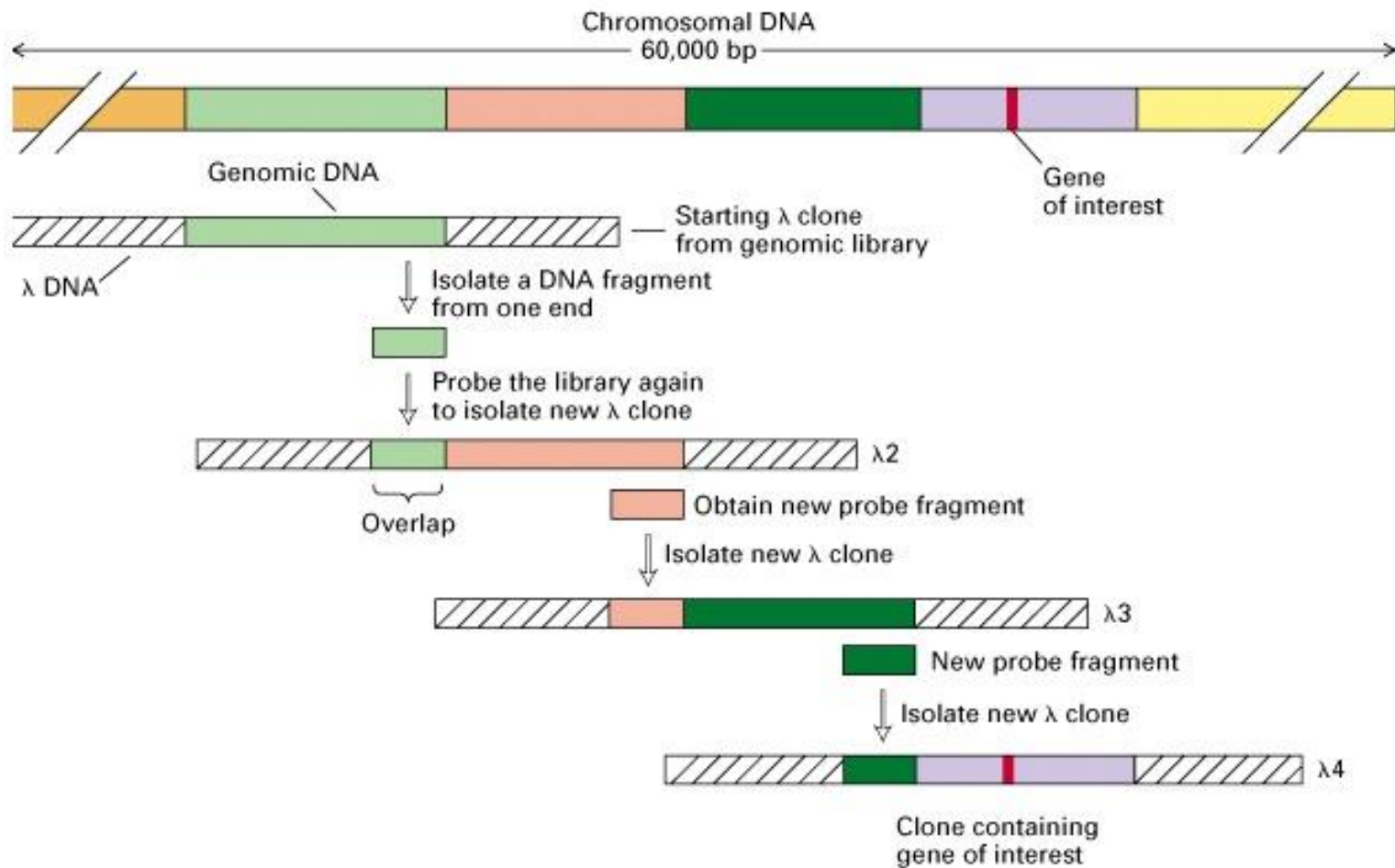
**Randomly Fragment**



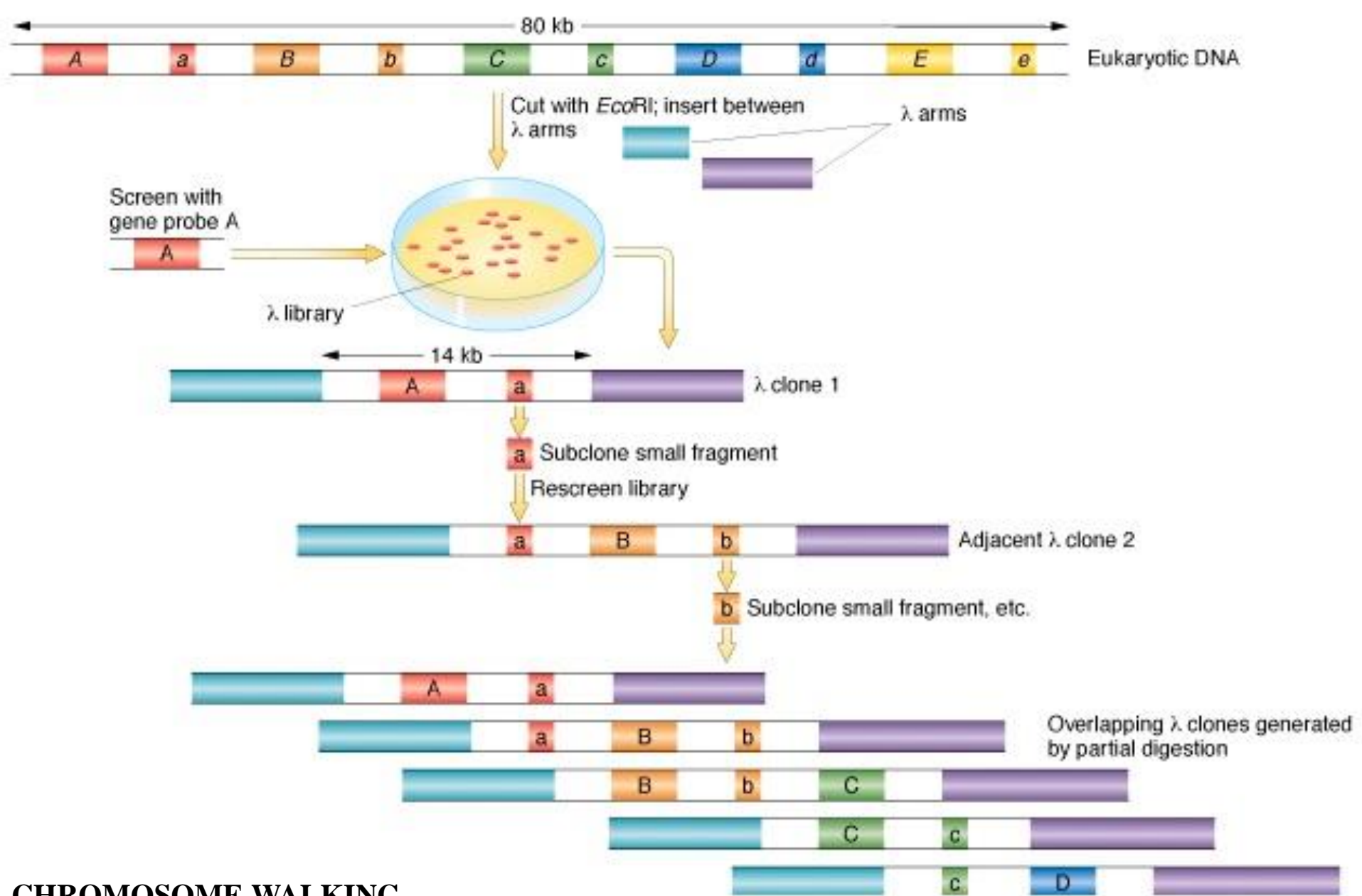
**Subclone Fragments**



GA	GA	GA	GA	GATCGTCTAGAGATCTC
GA	GA	GA	GA	GAGACTCTCTGAGAGTTC
GT	GT	GT	GT	GTGGGAAACTGTGTGA
TT	TT	TT	TT	TTTGACTAGCCCAAGT
TA	TA	TA	TA	TACGTTGTAGAGATGT
AT	AT	AT	AT	ATGATGCACCTGACCC
GA	GA	GA	GA	GAGTTTCACCTCTCAGC
GA	GA	GA	GA	GACTCACTCCAGCTCA
GA	GA	GA	GA	GAGGCCACCGGCGCT
GT	GT	GT	GT	GTGCAGTCCAGCCAC
GA	GA	GA	GA	GATTATTACCATTTTA
AT	AT	AT	AT	ATCCTTAGAGATTGACA

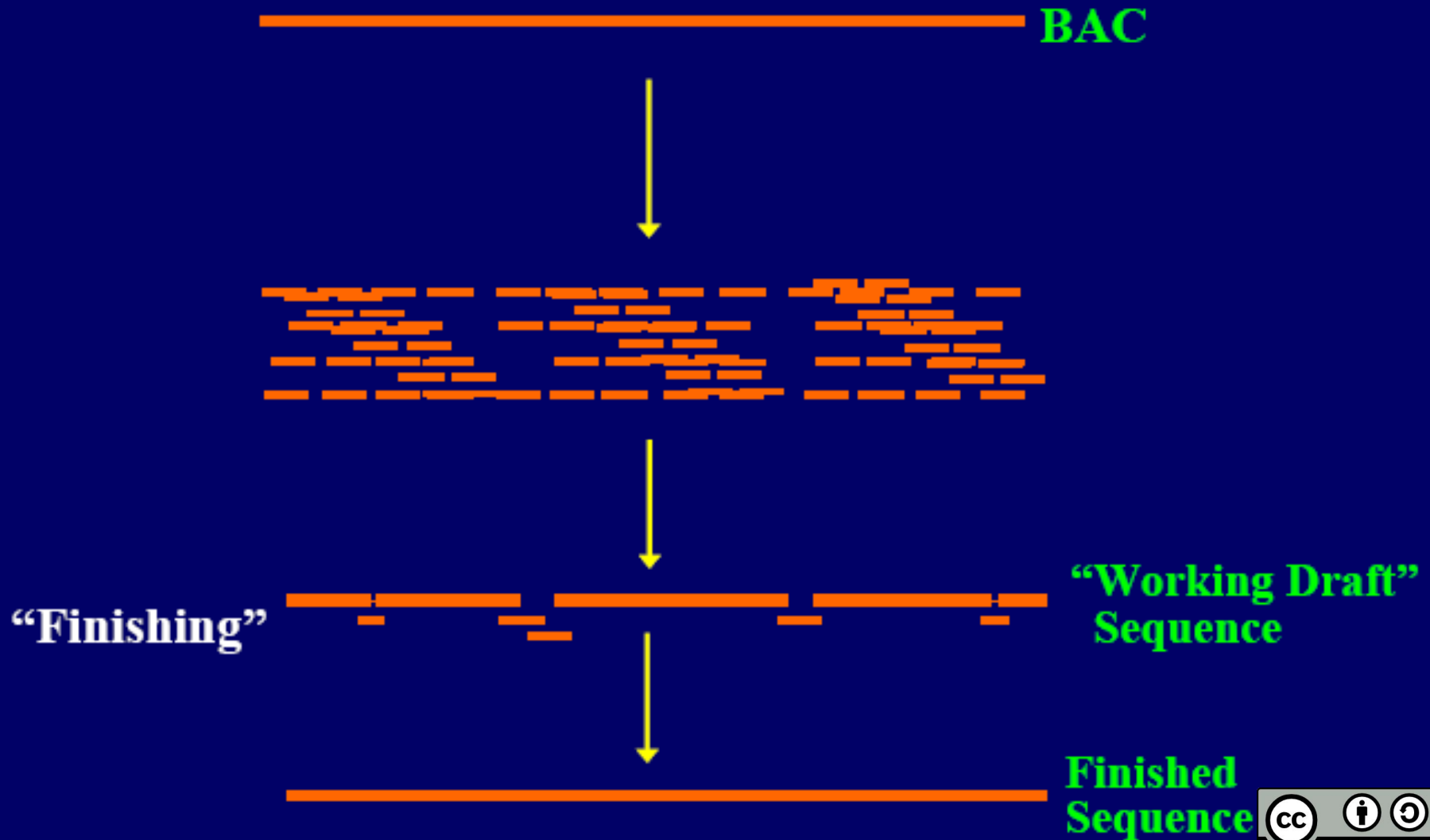


**CHROMOSOME WALKING** "crawl along the chromosome" - positional mapping begins with the known fragment that is bound to the analyzed unknown fragment / gene. Sequential hybridization, fragmentation and re-cloning - reconstruction of the sequence of the chromosome segments



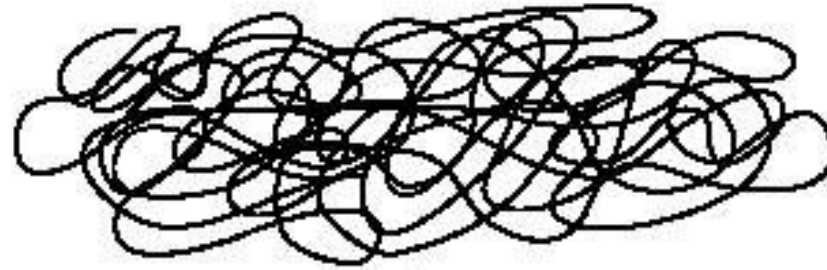
## CHROMOSOME WALKING

# Shotgun Sequencing Strategy



# Hierarchical shotgun sequencing

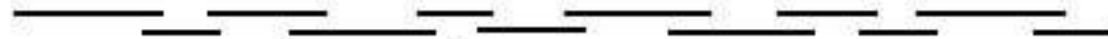
Genomic DNA



BAC library

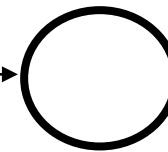
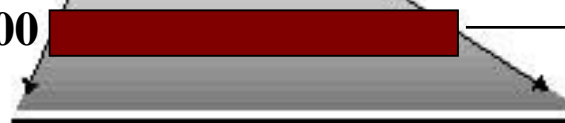


Organized mapped large clone contigs



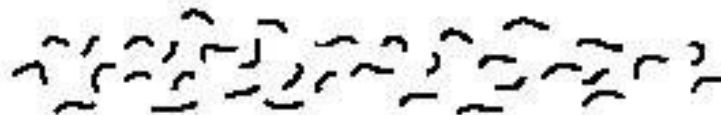
BAC to be sequenced

1500-2000



Lambda library – fragments sequenced from lambda primer

Shotgun clones

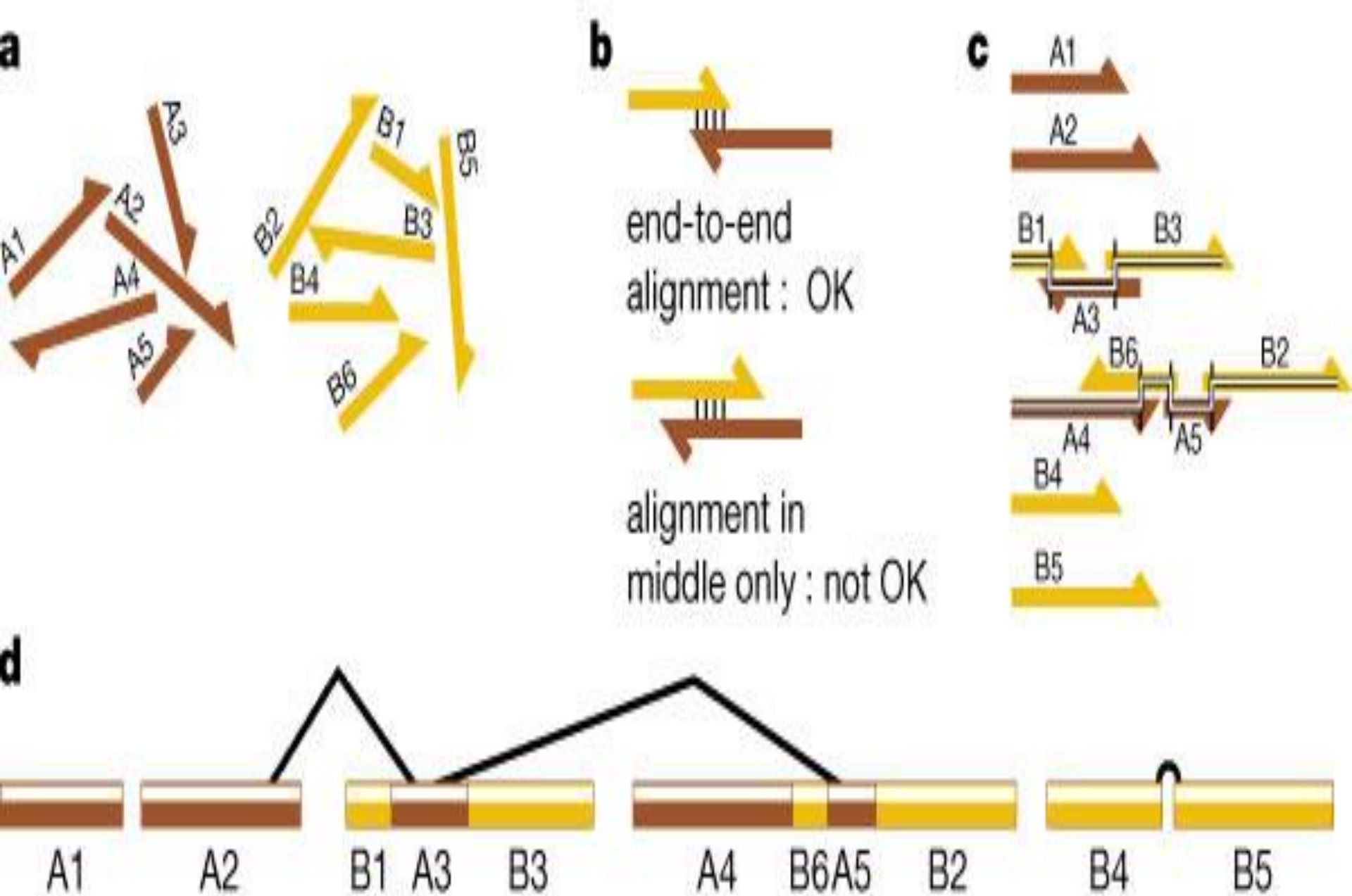


Shotgun sequence

...ACCGTAAATGGGCTGATCATGCTTAA  
TGATCATGCTTAAACCCTGTGCATCCTACTG...

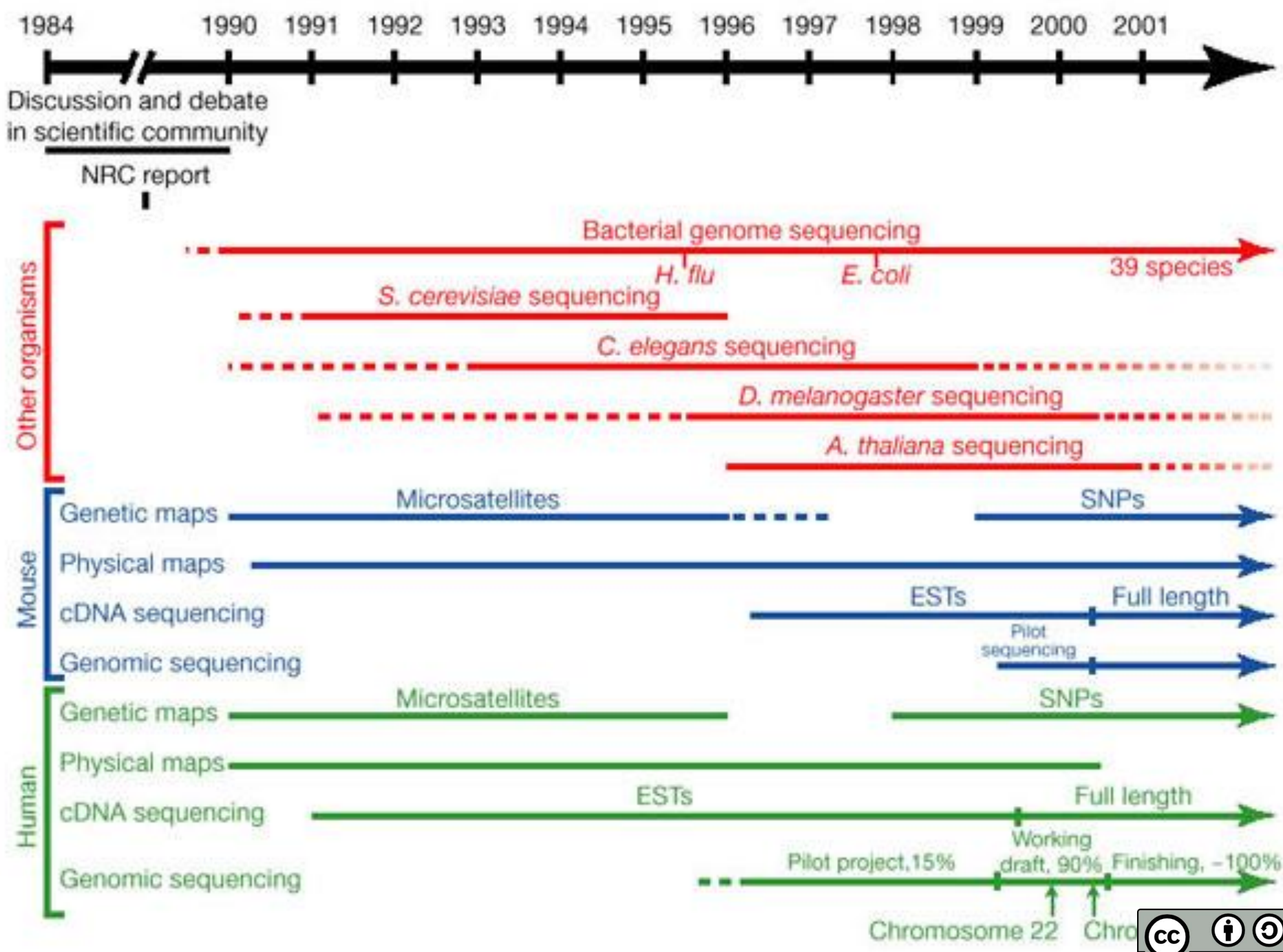
Assembly

...ACCGTAAATGGGCTGATCATGCTTAAACCCTGTGCATCCTACTG...



The key steps in assembling individual sequenced clones into the draft genome sequence. A1–A5 represent initial contigs derived from shotgun sequencing of clone A, and B1–B6 are from clone B.





Common Name	Genus and Species	Diploid Chromosome #
Buffalo	<i>Bison bison</i>	60
Cat	<i>Felis catus</i>	38
Cattle	<i>Bos taurus, B. indicus</i>	60
Dog	<i>Canis familiaris</i>	78
Donkey	<i>E. asinus</i>	62
Goat	<i>Capra hircus</i>	60
Horse	<i>Equus caballus</i>	64
Human	<i>Homo sapiens</i>	46
Pig	<i>Sus scrofa</i>	38
Sheep	<i>Ovis aries</i>	54

# Project **HUGO** – human genome project

1990 – aim: all  $3.10^9$  bp - year 2005

Estimated price  $3.10^9$  \$

8 teams (USA a UK)

1998 – Craig Venter joined with PerkinElmer in Connecticut

## **Celera Genomics**

announced, that human genome will be sequenced in 3 years  
for (300 mil. \$)

2 preconditions –      effective method  
                                 fast, automatic sequencers



1998 – PerkinElmer – fully automatic 96-well sequencer  
24 h a day ~ 10 sequences  
Celera – 230 sequencers, 100 millions nts a day  
(in 1995 ~ 1 million nts by one worker)



## Genetic Code of Human Life Is Cracked by Scientists



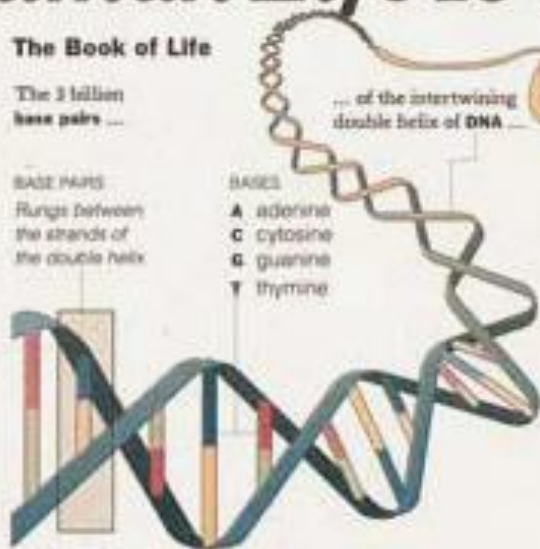
become part  
that Congress was entitled to the last  
word because Miranda's presump-  
tion that a confession was not volun-

### The Book of Life

The 3 billion  
base pairs ...

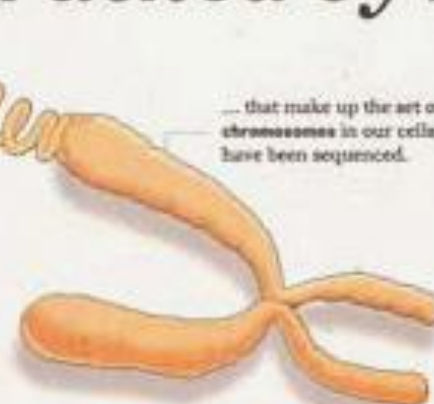
**BASE PAIRS**  
Rungs between  
the strands of  
the double helix

**BASES**  
A adenine  
C cytosine  
G guanine  
T thymine



... of the intertwining  
double helix of DNA

... that make up the set of  
chromosomes in our cells,  
have been sequenced.



By ordering the base units, scientists hope to  
locate the genes and determine their functions.

The New York Times

### A SHARED SUCCESS

2 Rivals' Announcements  
Marks New Medical  
Era, Risks and All

By NICHOLAS WADE

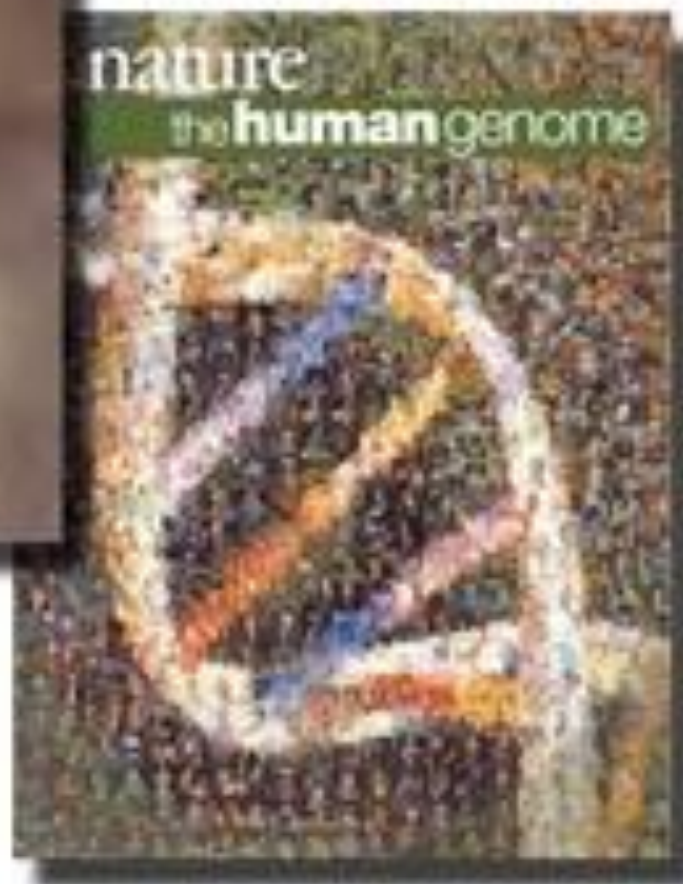
WASHINGTON, June 26 — In  
an achievement that represents a  
milestone of human self-knowledge,  
two rival groups of scientists said  
today that they had deciphered the  
genetic code, the set of instruc-  
tions that defines the human organism.

Francis Collins head of HUGO project  
Craig Venter president of Celera Genomics



**16.2.2001 – Celera  
95% of genome**

**Man ~ 30 000 genes**

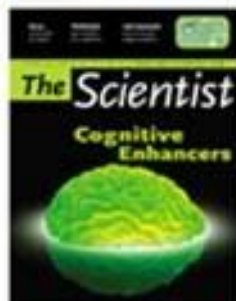


**16.2.2001  
HUGO consortium  
92% genome**



(Celera – approach to the consortium data)





## Lab Consumer

Volume 16 | Issue 20 | 53 | Oct. 14, 2002

[Previous](#) | [Next](#)

## To Dream the Not-So-Impossible Genomics Dream

Solexa's TotalGenotyping technology promises fast, economical whole-genome sequencing | [By Aileen Constans](#)

Photo: Courtesy of Solexa



☛ Solexa's single-molecule detection apparatus

Nick McCooke, CEO of

**Solexa**, has a bold goal: to analyze, in one day, the whole genome of an individual for one thousand dollars. The

Cambridge, UK-based company ([www.solexa.com](http://www.solexa.com)) is developing TotalGenotyping™, a method based on the Single Molecule Array™ technology invented by Solexa founders and Cambridge University academics **Shankar Balasubramanian**, **David**

Search

Go

[Advanced Search](#)

**BioMed  
Scientist  
JOBS**

Job of the Week

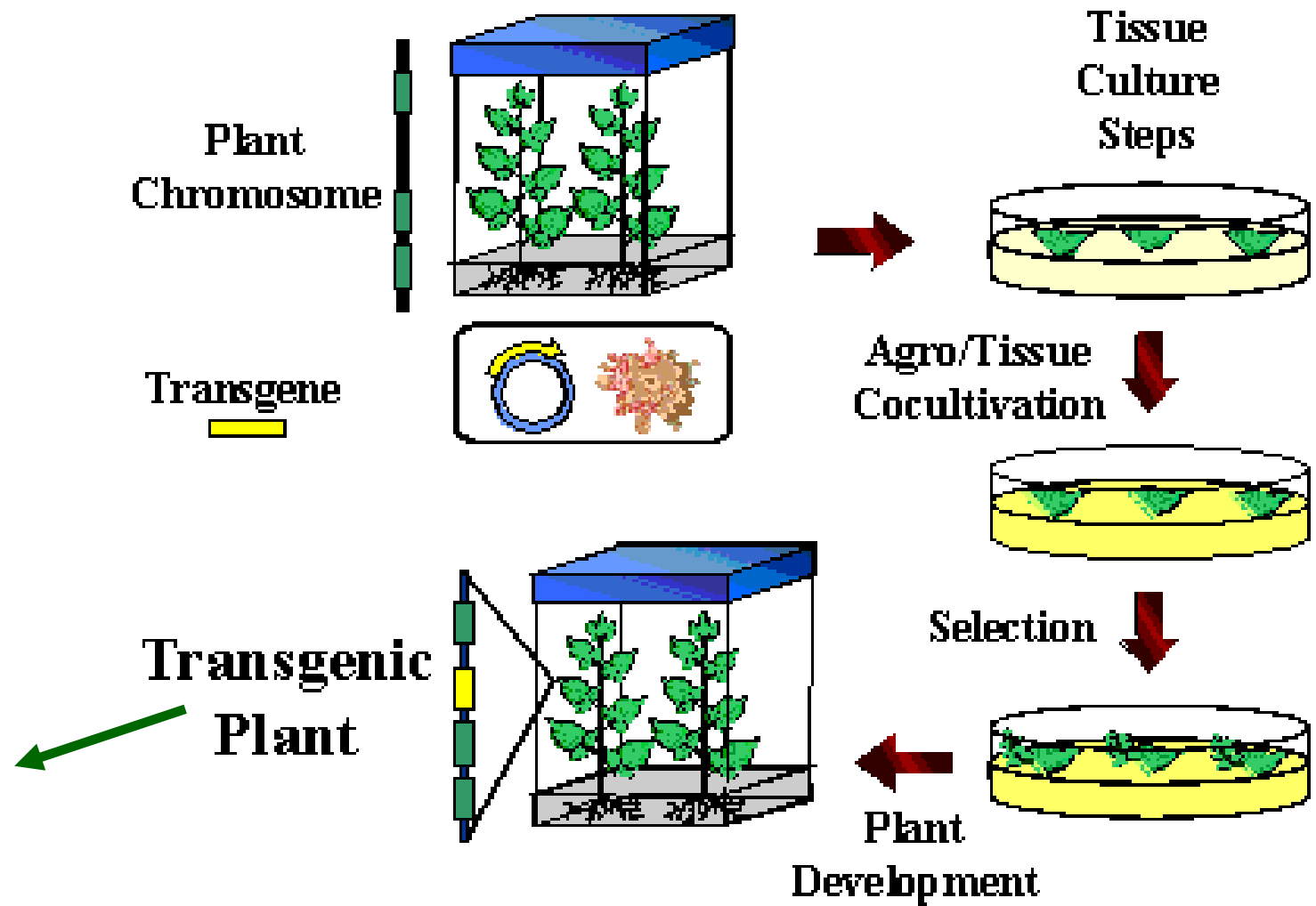
**BioMed Central**



# Farmacogenomics

- The use of DNA sequence information to measure and predict the reaction of individuals to drugs.
- Personalized drugs
- Faster clinical trials
  - Selected trial populations
- Less drug side effects
  - Toxicogenomics

# GMO – Plants the Lab Steps

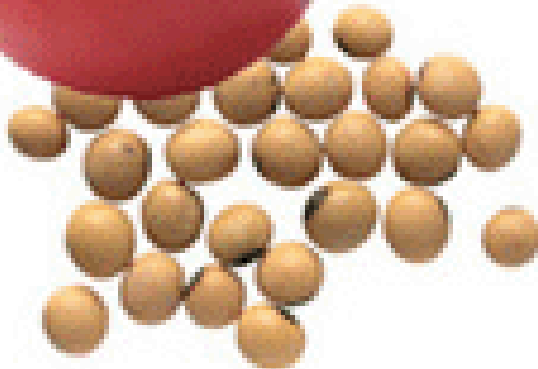


# Why are transgenic plants important?

We can develop organisms that express a “novel” trait not normally found in the species



Extended shelf-life tomato (Flavr-Savr)



Herbicide resistant soybean (Roundup Ready)

# Transgenic plants – commercially available

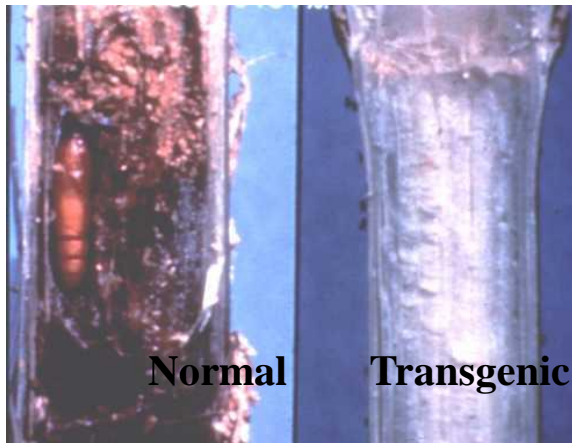


Source: USDA

## *Insect resistant cotton*

– Bt (*B. thuringiensis*) toxin kills the cotton boll worm

- *transgene = Bt protein*



*Insect resistant corn* – Bt toxin kills the European corn borer

- *transgene = Bt protein*



Source: Monsanto

## ***Herbicide resistant crops***

**Now:** soybean, corn, canola (safflower)

**Coming:** sugar beet, lettuce, strawberry  
alfalfa, potato, wheat (?)

- *transgene = modified EPSP synthase or phosphinothricin-N-acetyltransferase*



***Virus resistance*** - papaya resistant to  
papaya ringspot virus

- *transgene = virus coat protein*



Source: Chr. Hansen

*Biotech chymosine; the enzyme used to curdle milk products*

- *transgene = genetically engineered enzyme*



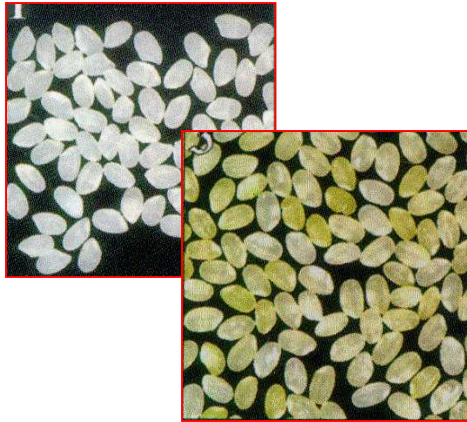
Source: Rent Mother Nature

*bST; bovine somatotropine (Growth hormone )  
- to increase milk production*

- *transgene = genetically engineered enzyme*



# Next Generation of Ag Biotech Products



***Golden Rice*** – increased Vitamine A content  
(but not without controversy)

*transgene = three pathway enzymes*



***Sunflower*** – white mold resistance

*transgene = oxalate oxidase from wheat*





***Turfgrass*** – herbicide resistance;  
slower growing (= reduced mowing)

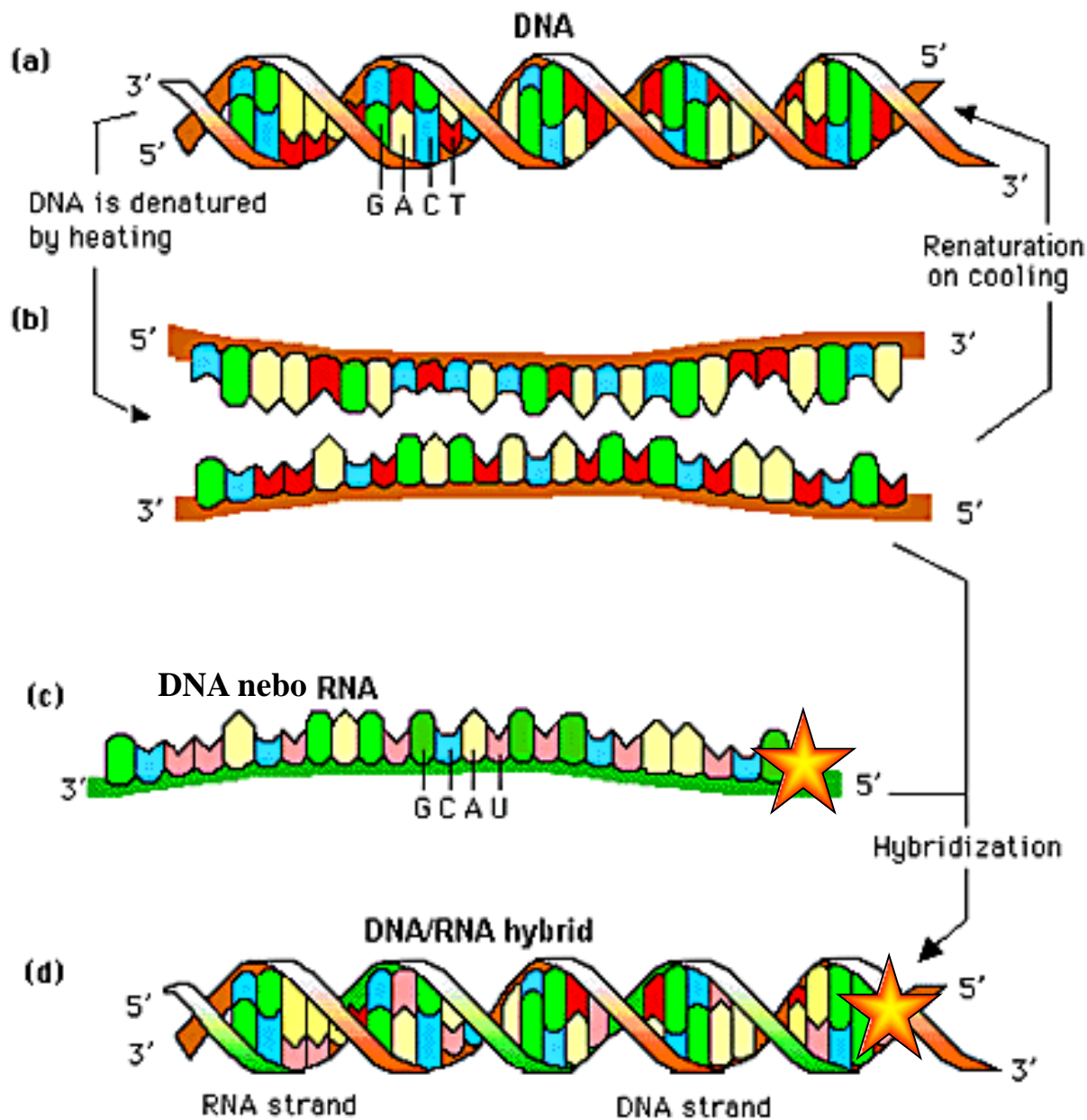


***Bio Steel*** – spider silk expressed in goats; used to  
make soft-body bullet proof vests (Nexia)

# Edible Vaccines

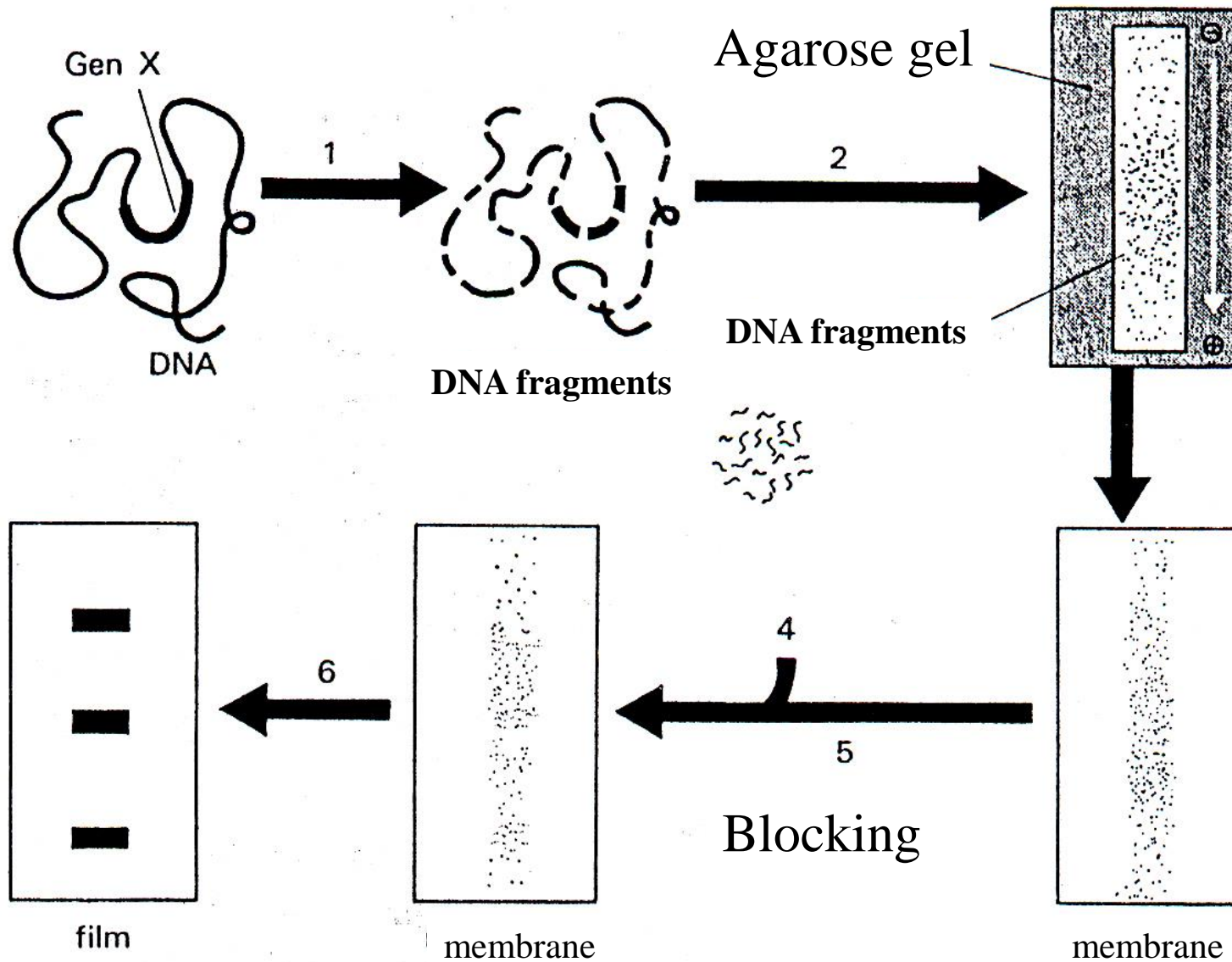
## *Transgenic Plants Serving Human Health Needs*

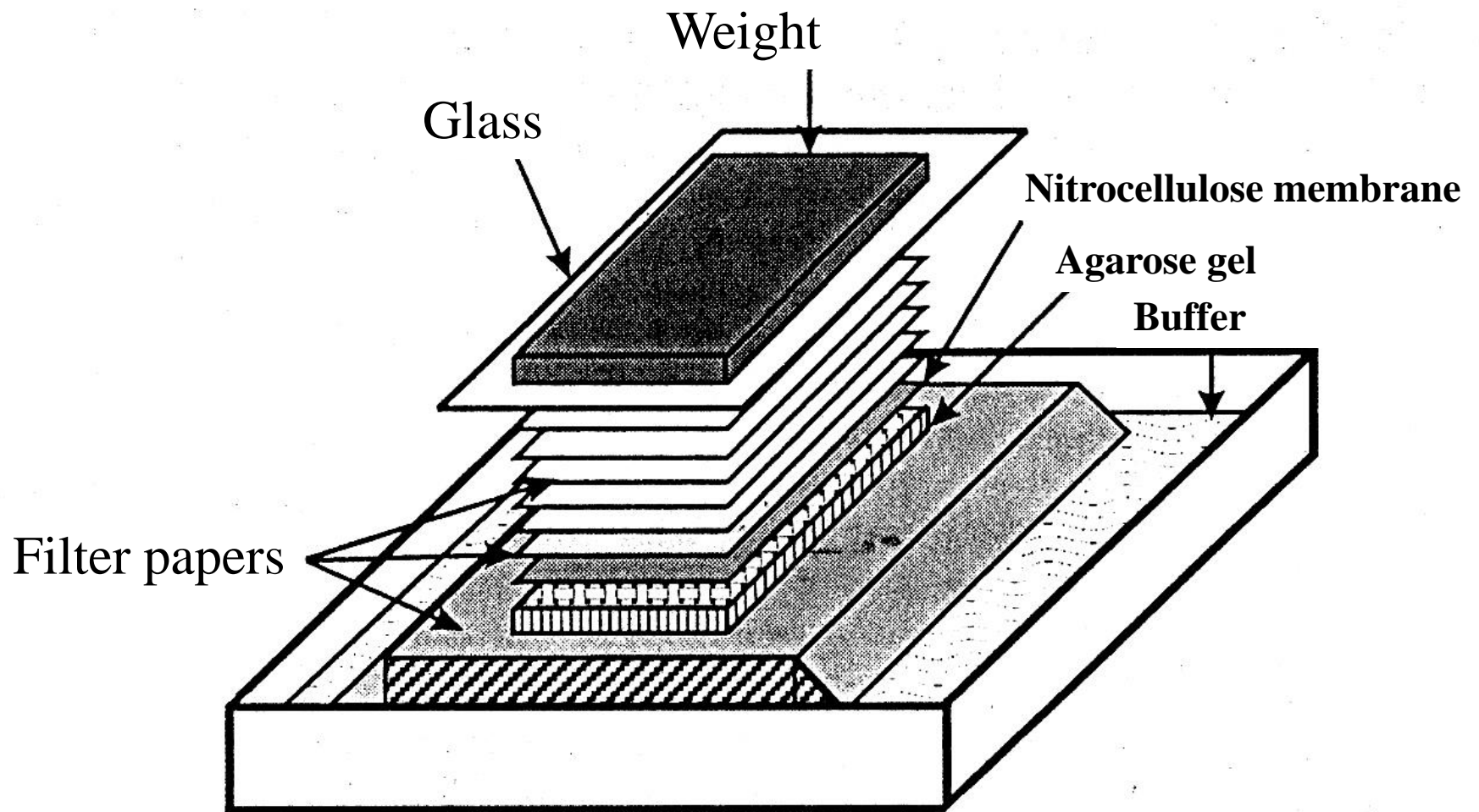
- Works like any vaccine
- A transgenic plant with a pathogen protein gene is developed
- Potato, banana, and tomato are targets
- Humans eat the plant
- The body produces antibodies against pathogen protein
- Humans are “*immunized*” against the pathogen
- Examples:
  - ✓ Diarrhea
  - ✓ Hepatitis B
  - ✓ Measles



## Nucleic Acid Hybridization

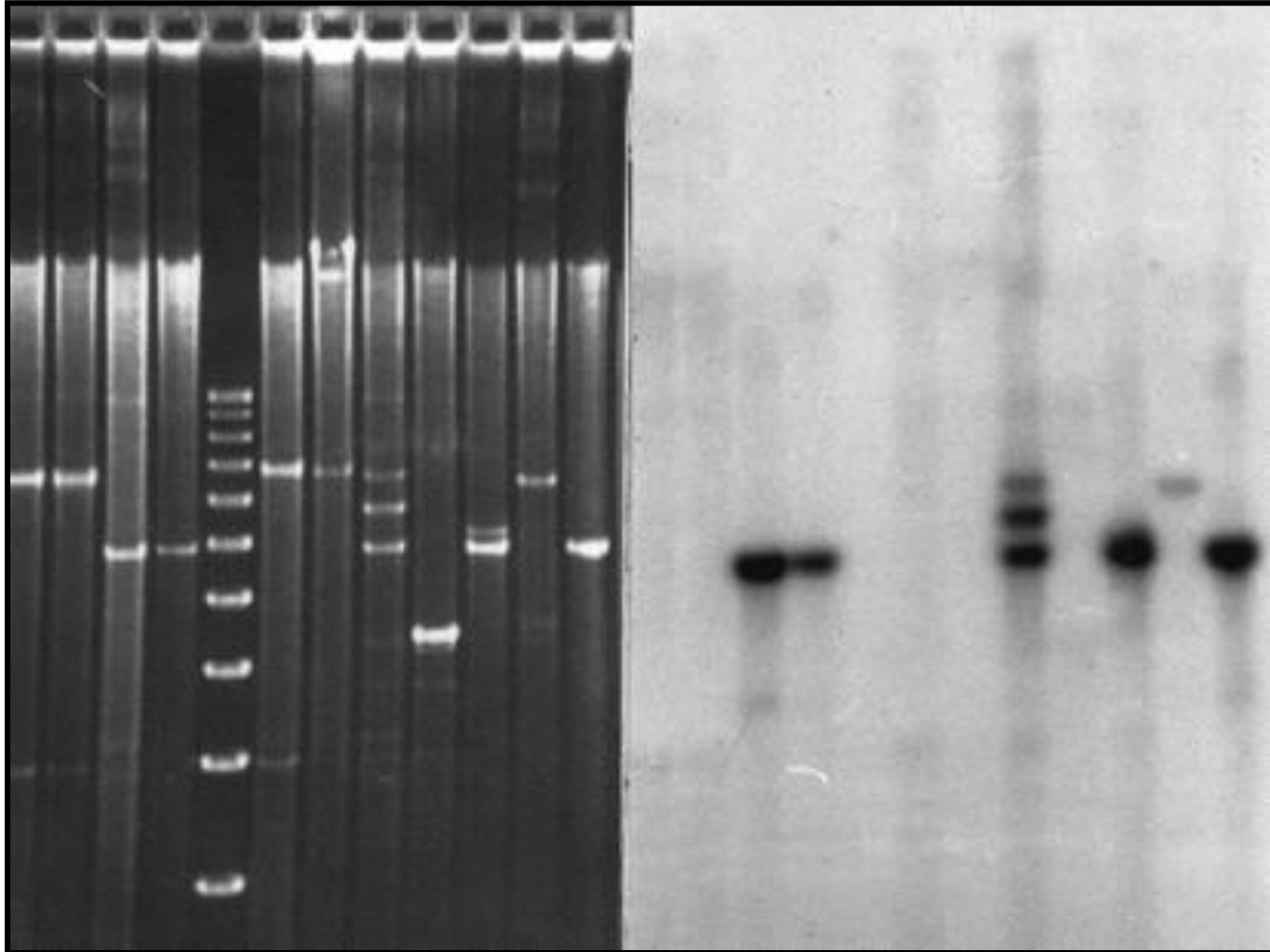
# Southern blot (DNA)







# Southern Blot

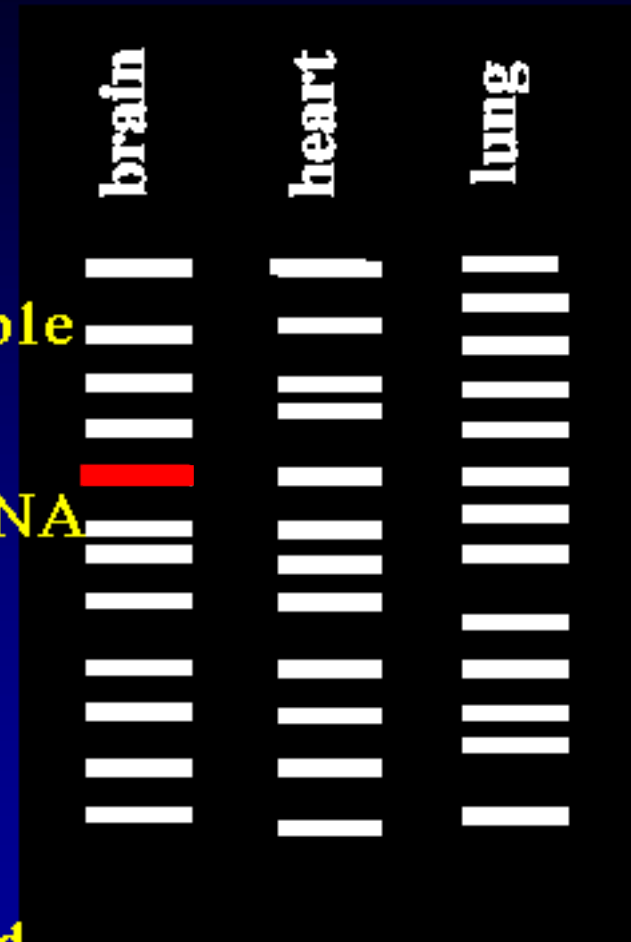


DNA stained with  
ethidium bromide

Autoradiograph of  
hybridized membrane

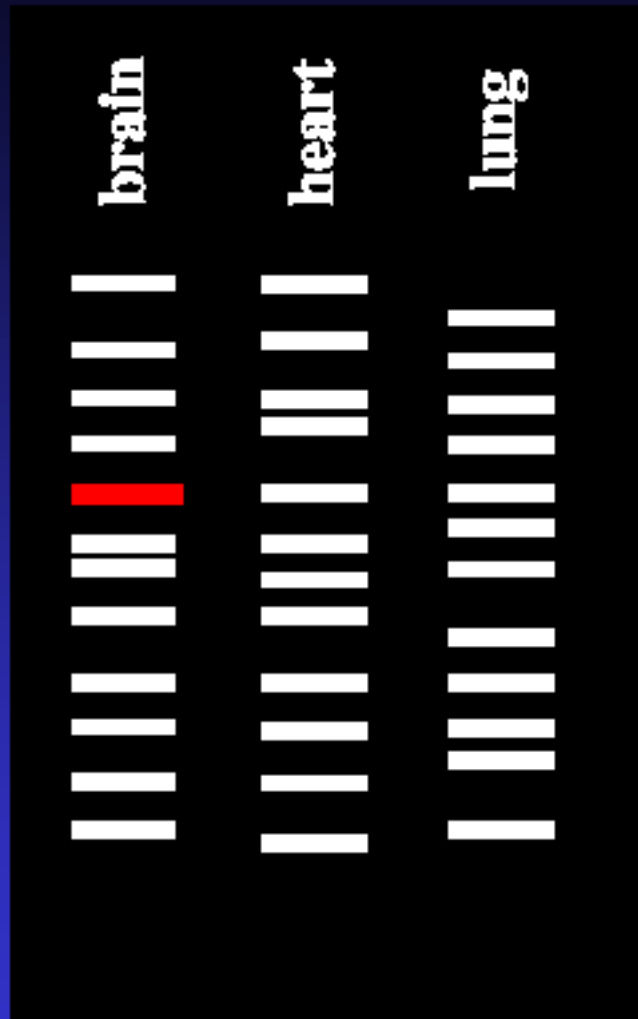
# Analysis of RNA - Northern Blotting

- Isolate tissue RNA
- Electrophorese the RNA
- Transfer the RNA to a stable support - membrane
- Label a piece of cloned DNA
- Allow the labelled DNA to base pair with the RNA on the membrane
- Visualise the labelled band

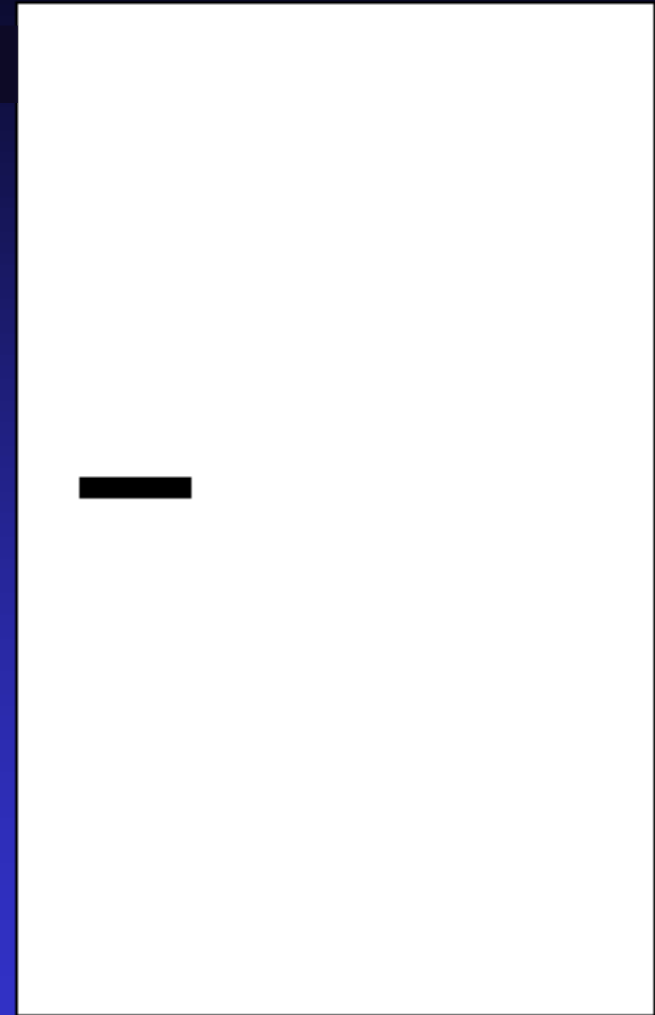




# Autoradiography - visualisation of RNA transcript

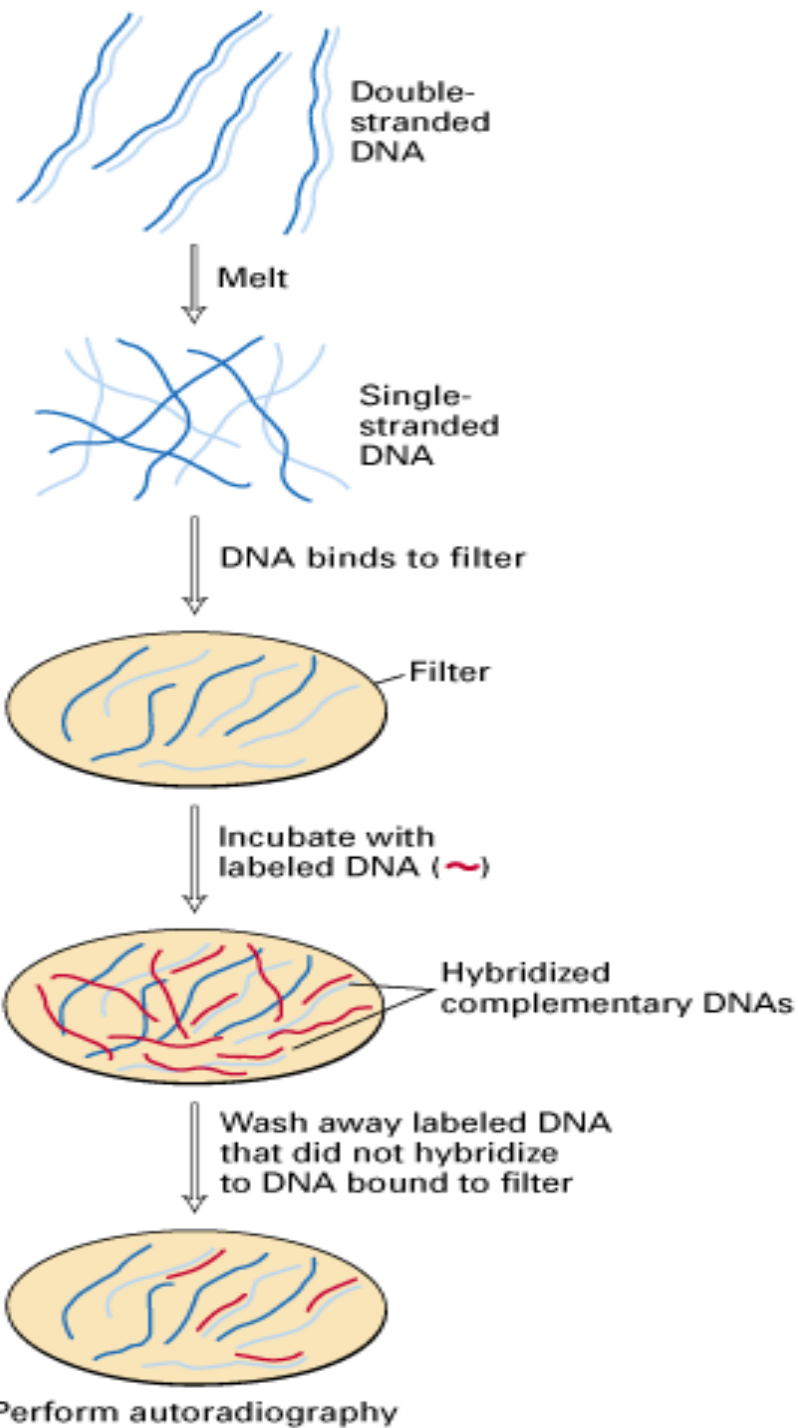


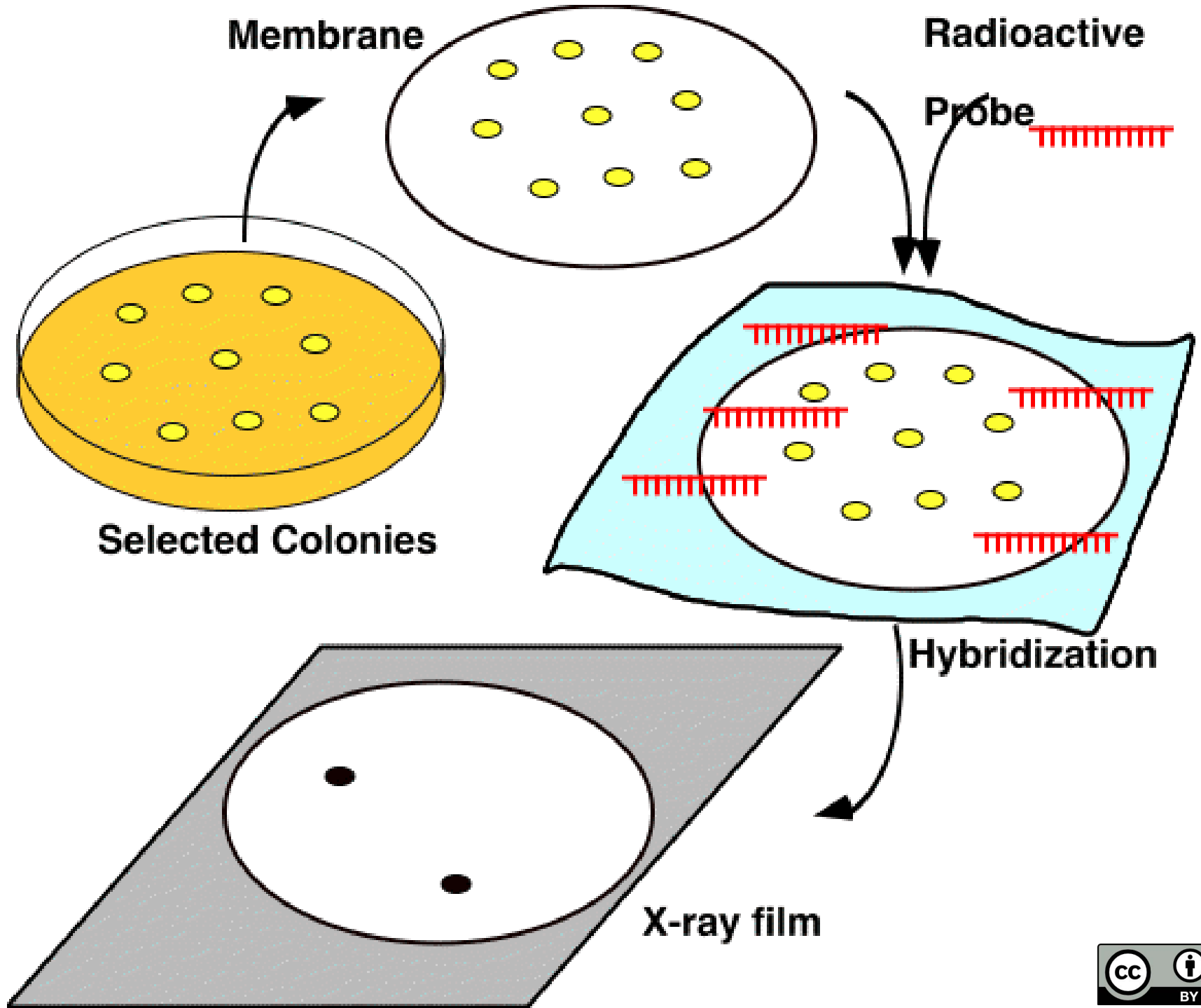
Gel



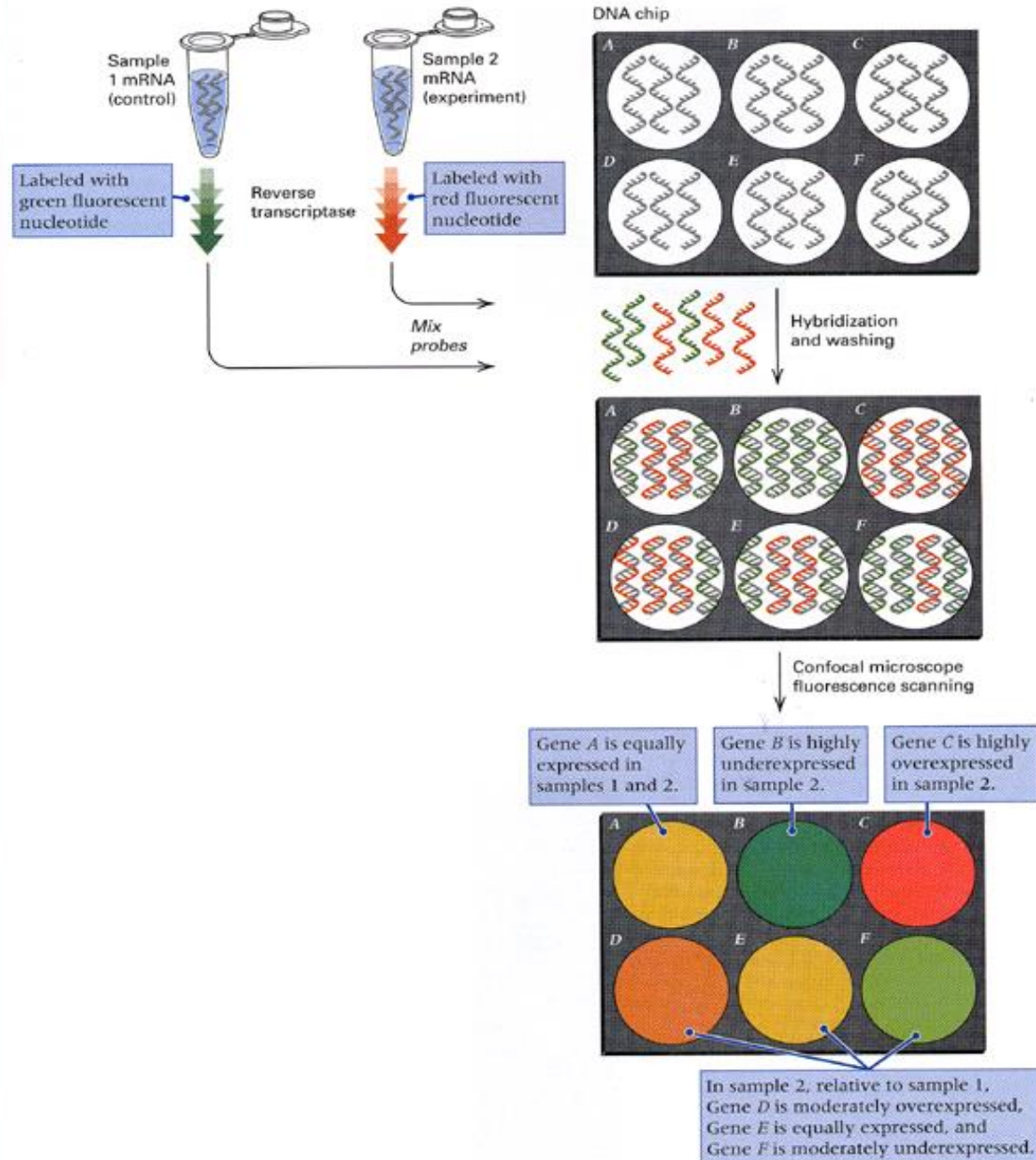
X-ray film image

<b><u>Blot</u></b>	<b><u>Analyte</u></b>	<b><u>Probe and detection</u></b>
<b>Southern</b>	<b>DNA</b>	<b>labeled DNA probe</b>
<b>Northern</b>	<b>RNA</b>	<b>labeled DNA probe</b>
<b>Western</b>	<b>protein</b>	<b>specific antibody</b>
<b>Far Western</b>	<b>protein</b>	<b>specific antibody</b>
<b>South Western</b>	<b>protein</b>	<b>labeled DNA probe</b>

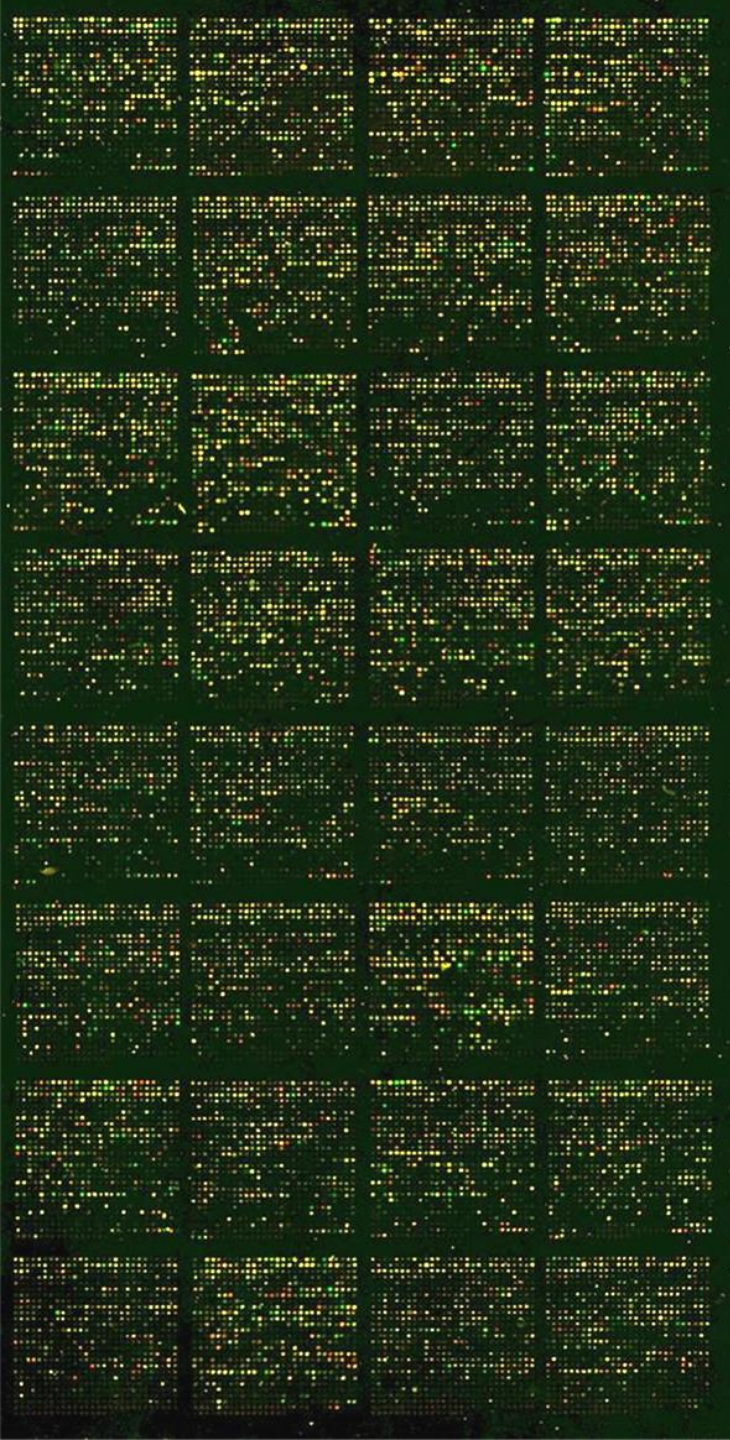




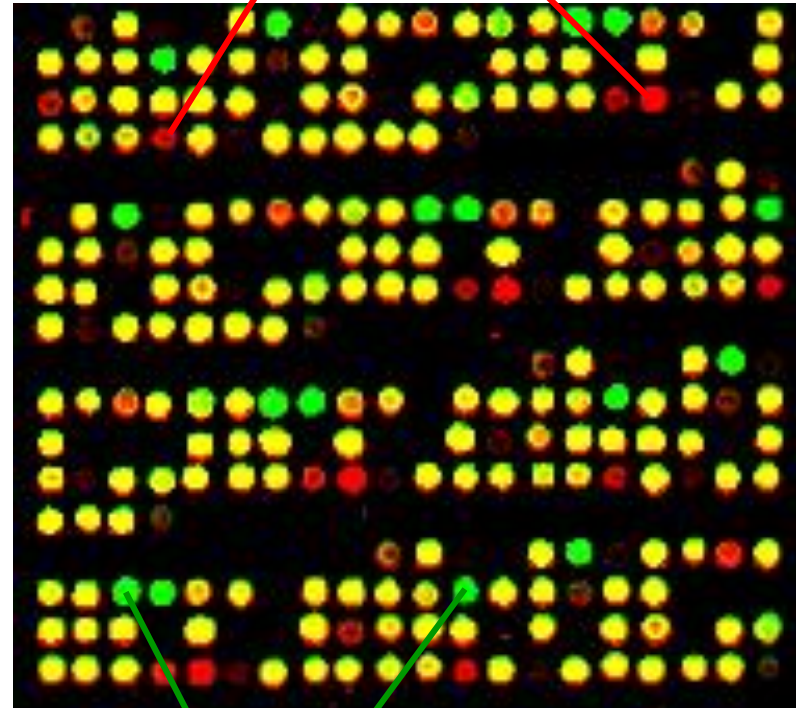
# Hybridizing to Microarrays





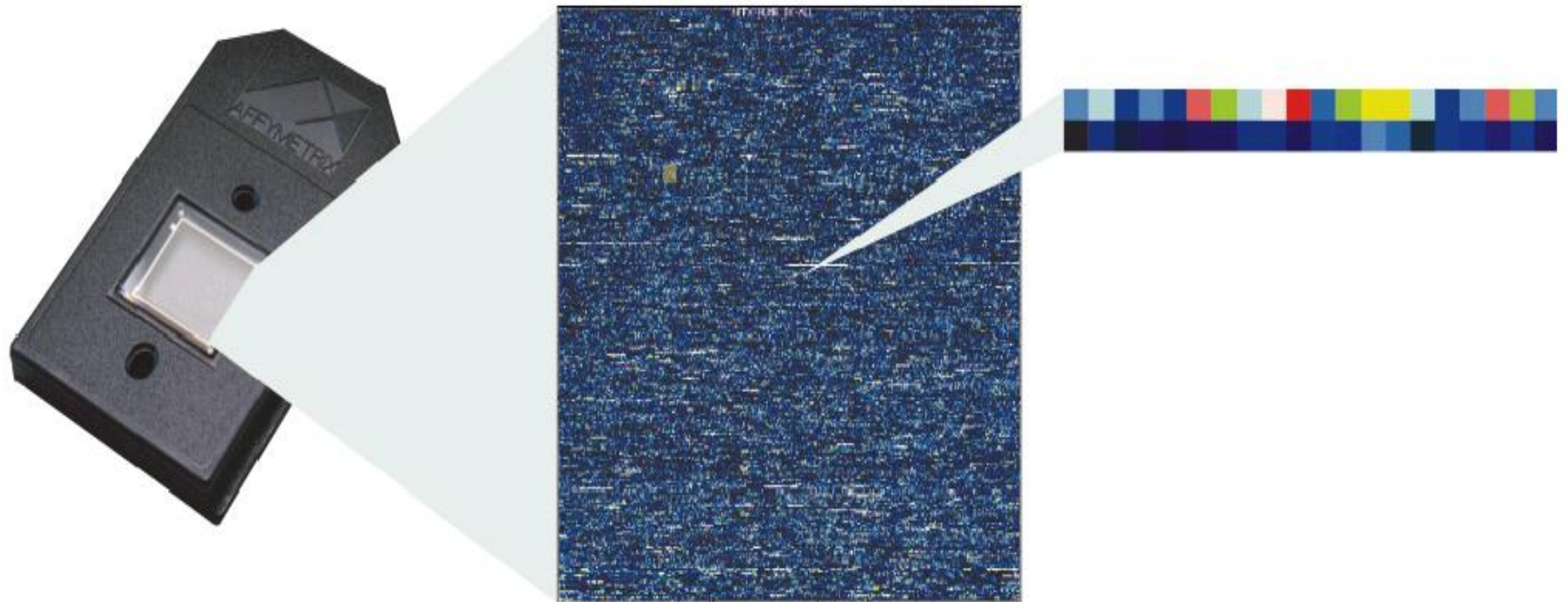


increased expression



decreased expression

# Affymetrix microarray

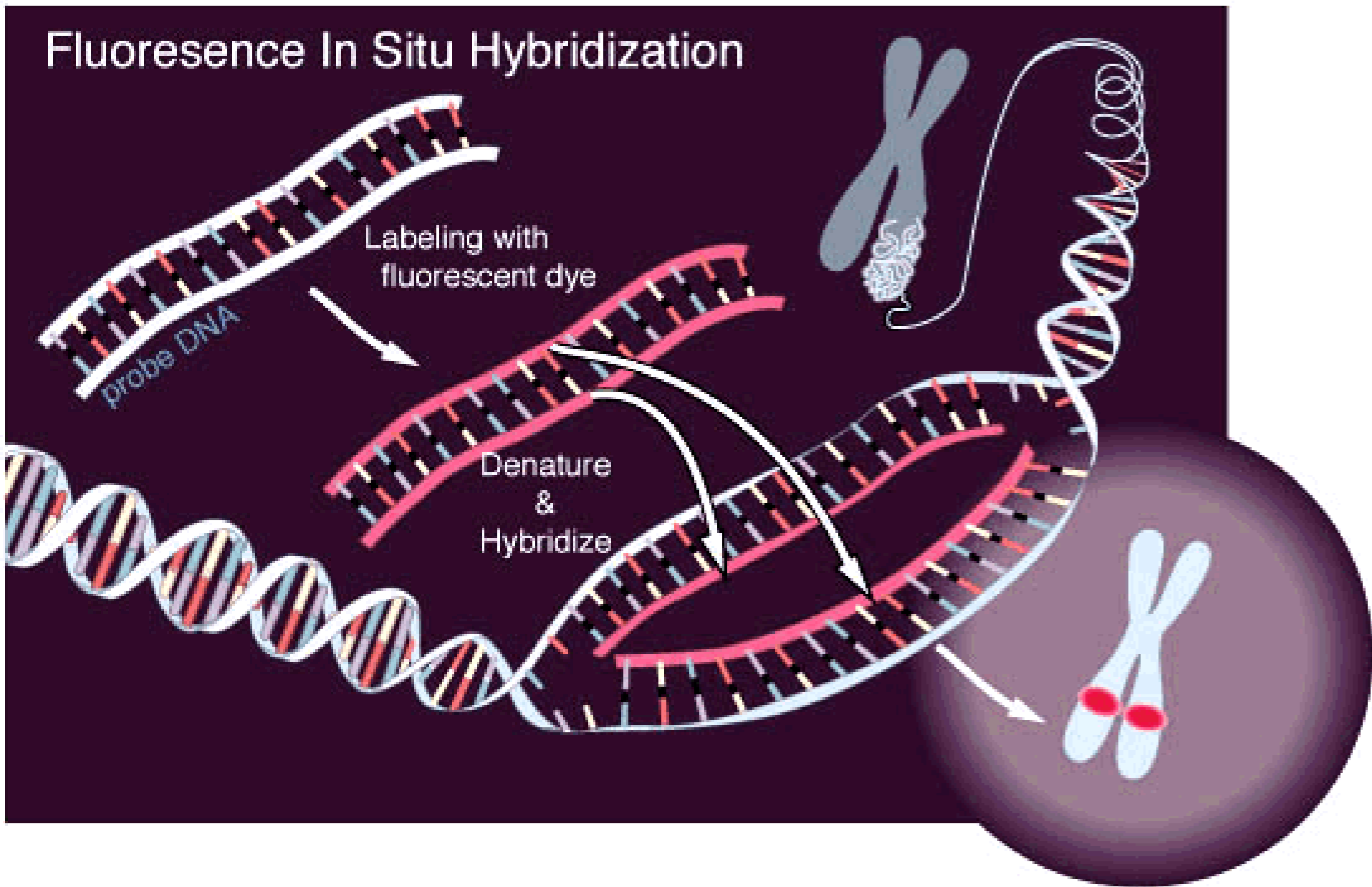




# Typical Microarray Workflow

- a) isolate RNA
- b) reverse transcription → cDNA (oligo dT primers)
- c) amplify/label cDNA → material for microarray hybridization
- d) hybridize labeled cDNA to microarray
- e) wash step (removes unbound material)
- f) scan (fluorescent microarray scanner)
- g) analyze data with software program

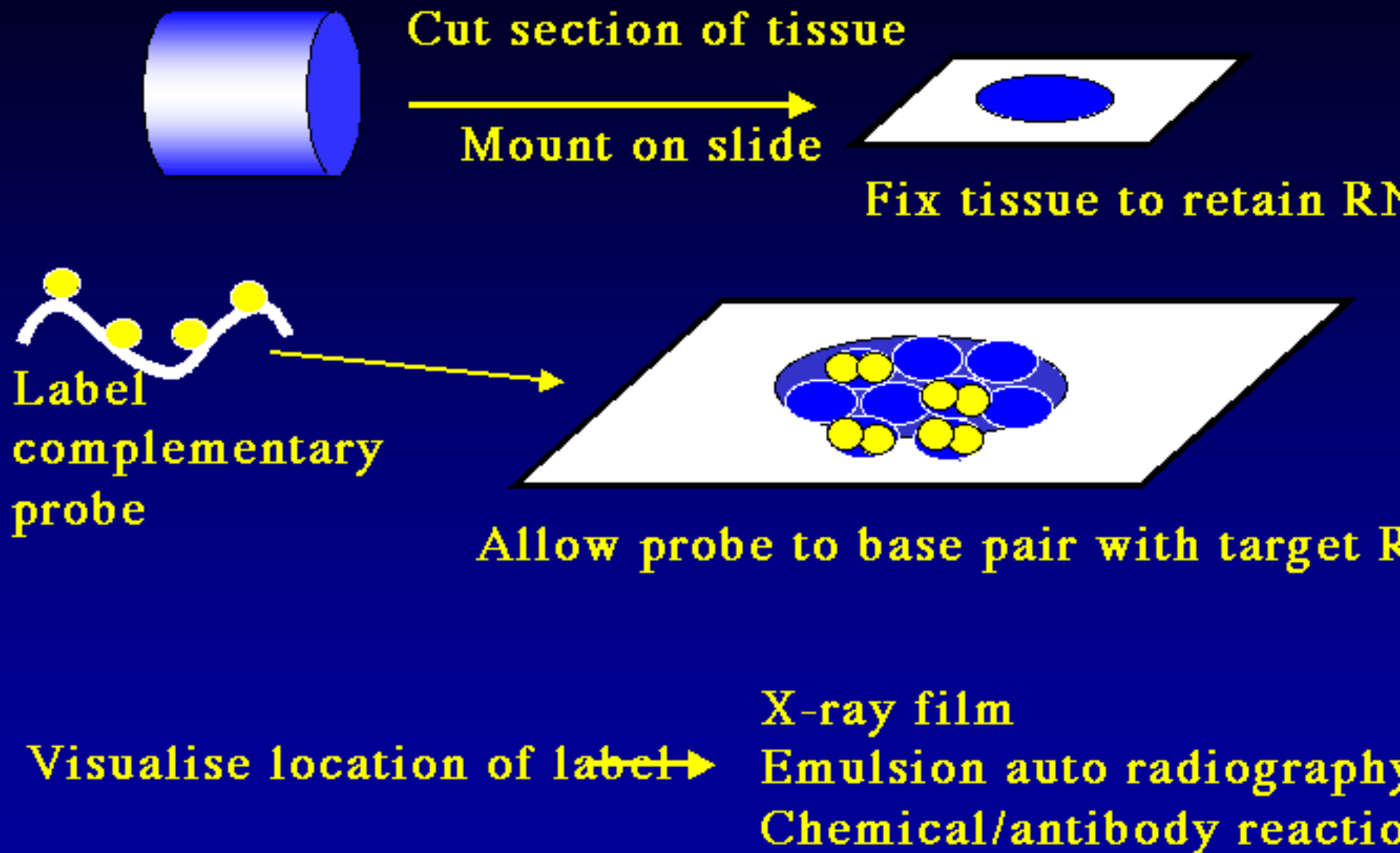
# Fluorescence In Situ Hybridization



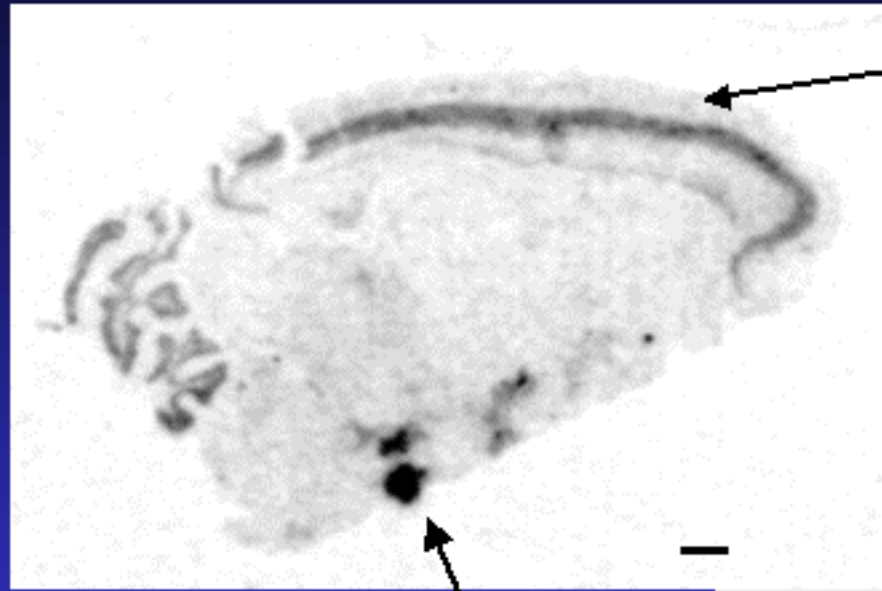


Here the site of  
hybridization is labeled  
green.

# In situ hybridisation - method



# In situ hybridisation for two related genes in the rat brain

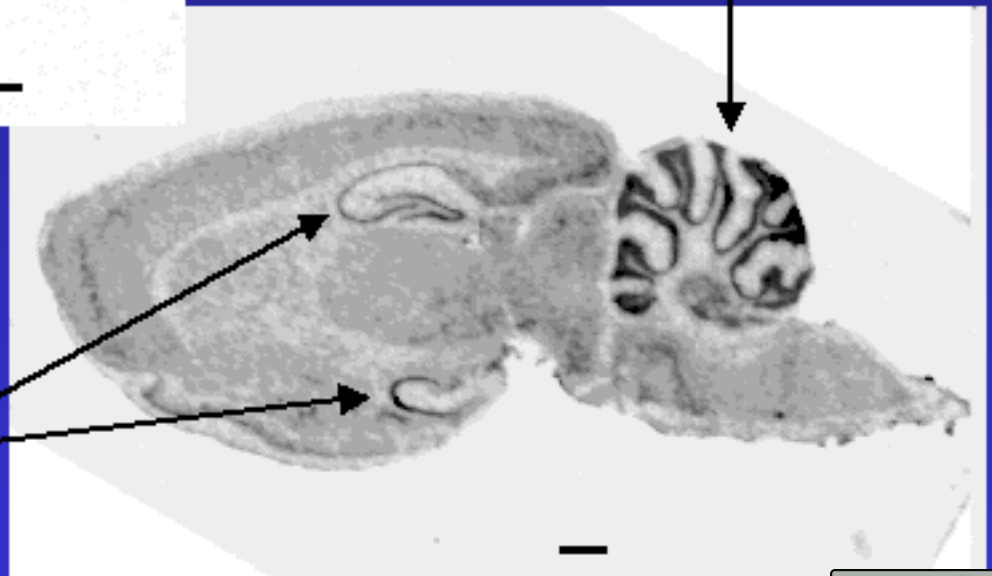


Cortical layers IV & V

cerebellum

hypothalamus

hippocampus



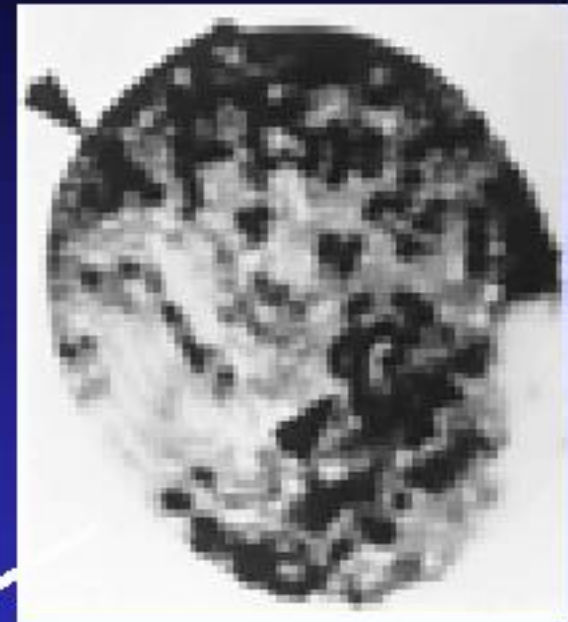
# Changes in mRNA expression levels during development



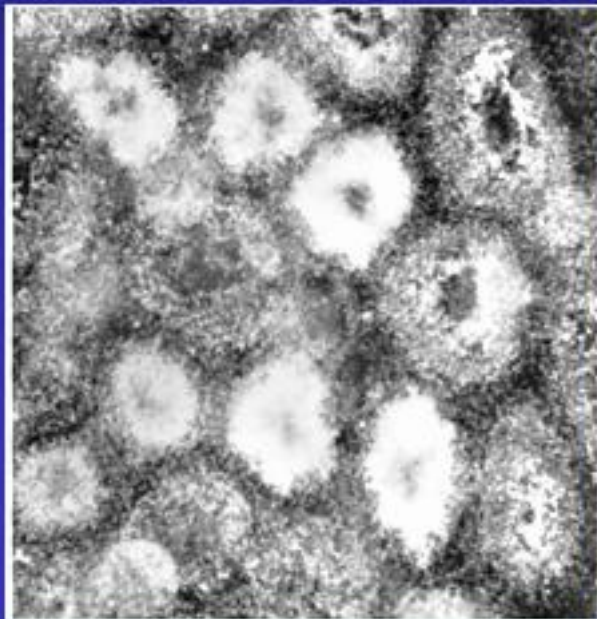
Day 10



Day 30



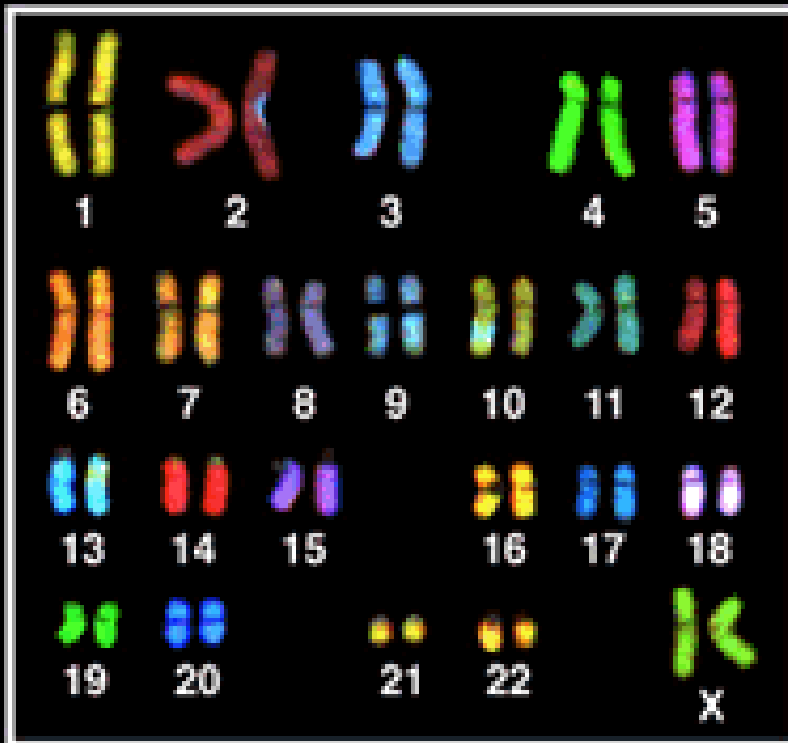
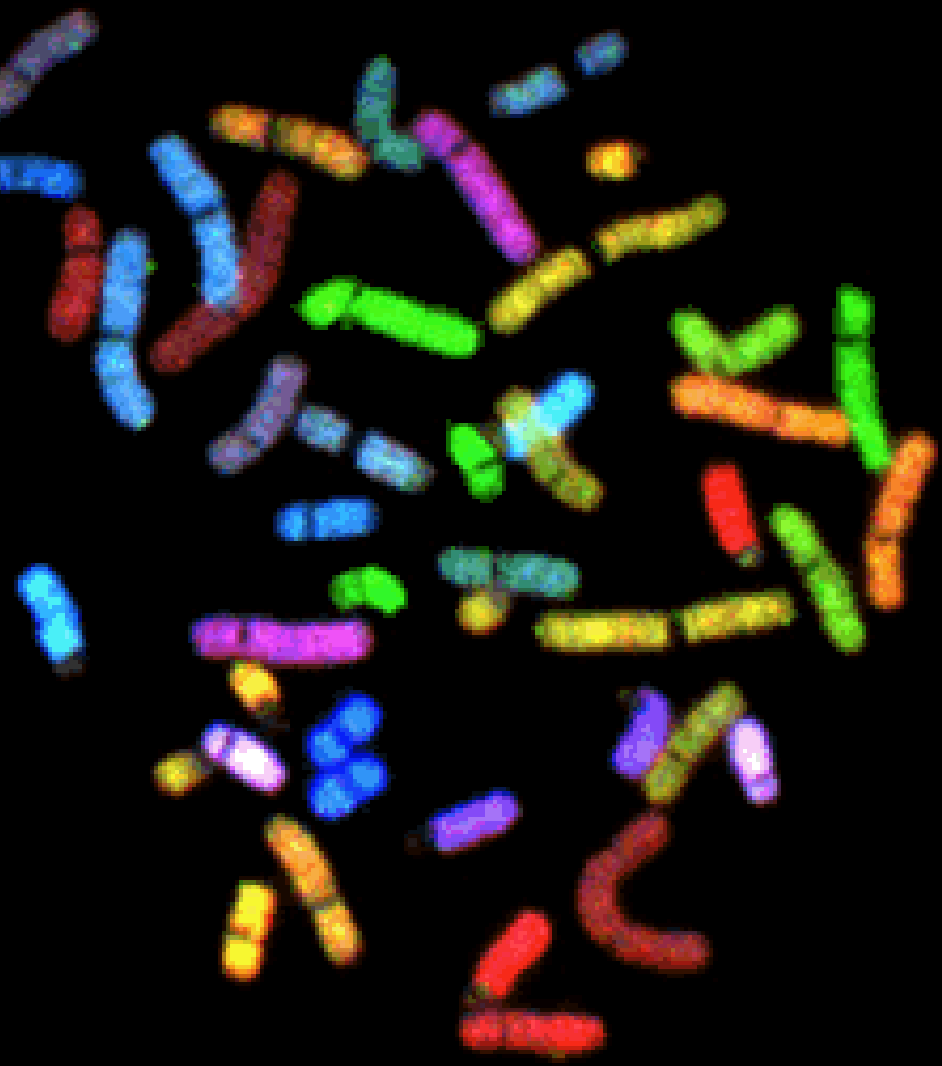
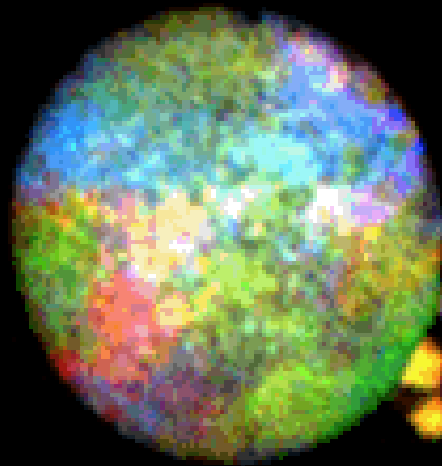
Day 55



Dark field microscope image of PK $\beta$  in 55 day rat testis

# Spectral Karyotyping – SKY

## Detection of translocations





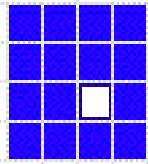
This photograph shows a complete set of chromosomes from an acute promyelocytic leukemia (APL) patient. A new technique called chromosome painting allows visual distinction between chromosomes and can be used to show the chromosome translocations that frequently occur in human cancers. In the case of APL, chromosome 13 is lost, there is a translocation between chromosomes 7 and 15, translocation between chromosomes 11, 15, 17, and between chromosomes 9 and 18. (Look for chromosomes painted with more than one color.) With thanks to Thomas Ried, National Human Genome Research Institute, NIH, for supplying the picture.



# Human genetic variation



Two randomly chosen people differ on average at 0.09% of their base pairs.



APBiotech - AstraZeneca - Aventis - Bayer - Bristol-Myers Squibb - F.Hoffman-La Roche - Glaxo Wellcome

**THE SNP CONSORTIUM LTD**

IBM - Motorola - Novartis - Pfizer - Searle - SmithKline Beecham - Wellcome Trust

## **Human genome sequence**

**1 in 1,000 nucleotides (letters) in the genome differ among individuals**

**That variability is key to understanding variability in disease susceptibility and outcomes**

**SNP = single nucleotide polymorphism**

**>1.4 million SNPs identified**

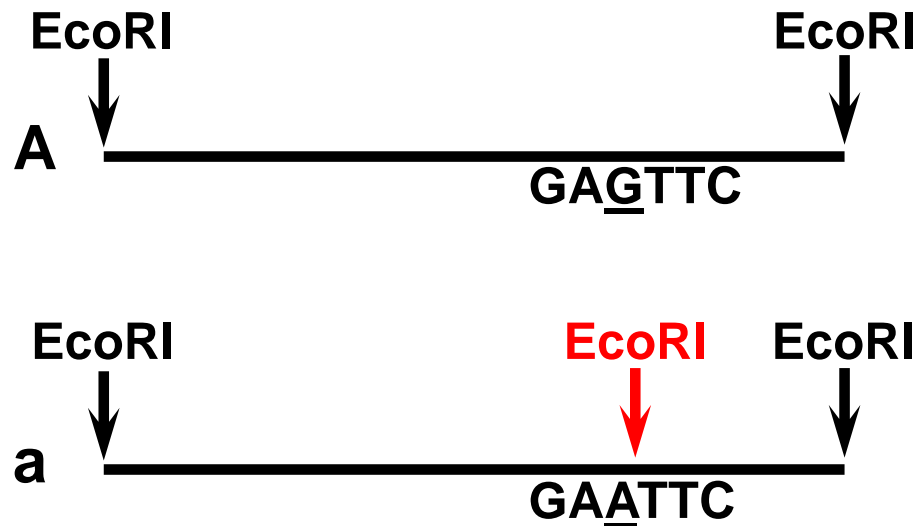
# Restriction fragment length polymorphisms (RFLPs)

- polymorphism

variation of nucleotide sequence caused by point mutation, deletion or insertion...

- RFLP

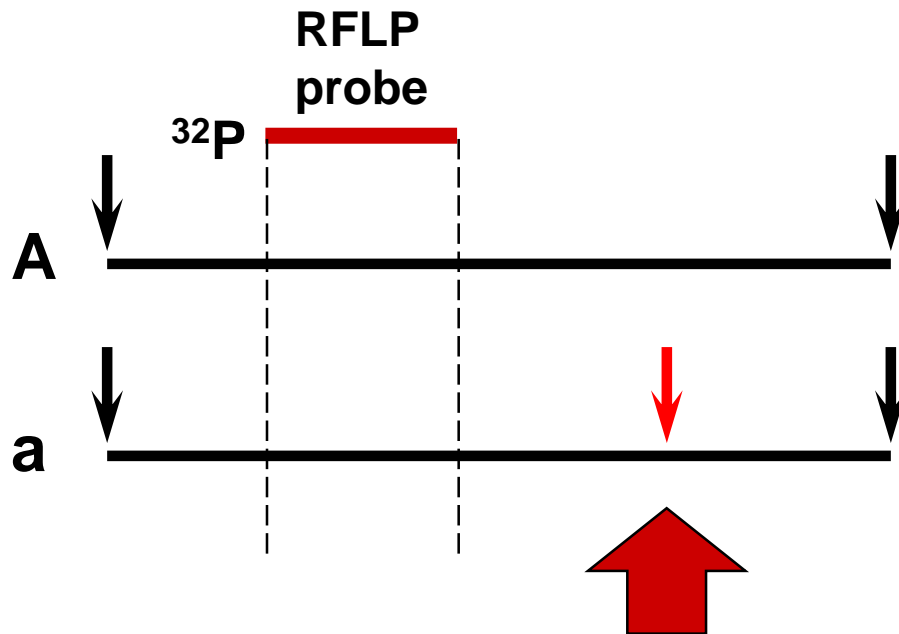
Southern blot analysis



Two alleles 'A' and 'a'

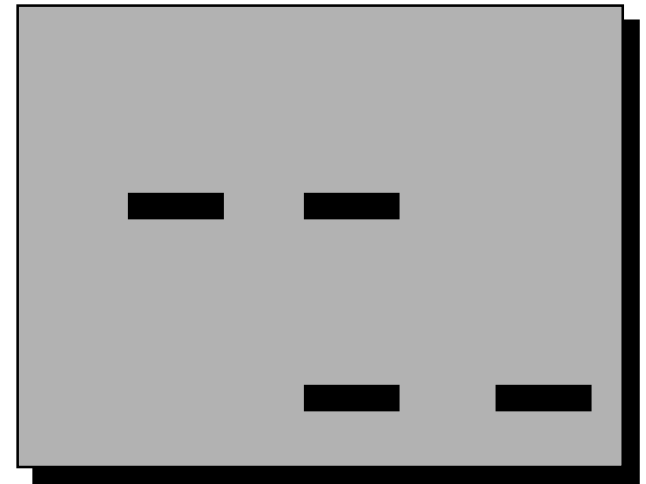
# Restriction fragment length polymorphisms (RFLPs)

- RFLP  
size - Southern blot analysis
- polymorphism - absence or presence single cut site:

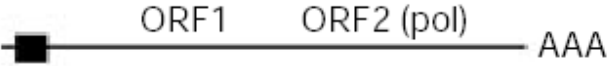

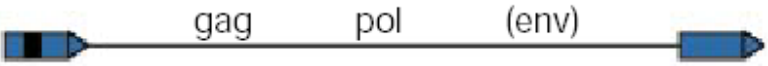
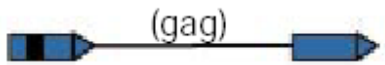
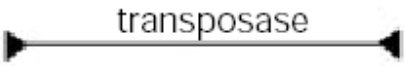



Target site of the enzyme  
new / eliminated

“A” a “a” - alely  
AA Aa aa

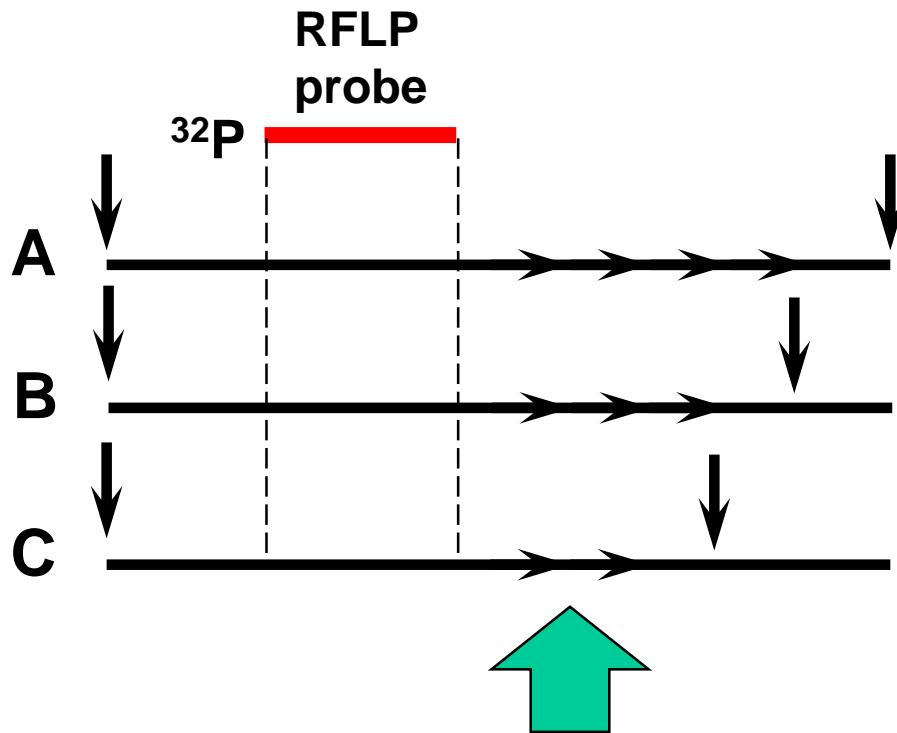


Southern hybridization  
 $^{32}\text{P}$ - probe

		Length	Copy number	Fraction of genome
Autonomous		6–8 kb	850,000	21%
Non-autonomous		100–300 bp	1,500,000	13%
Autonomous		6–11 kb	450,000	8%
Non-autonomous		1.5–3 kb		
Autonomous		2–3 kb	300,000	3%
Non-autonomous		80–3,000 bp		

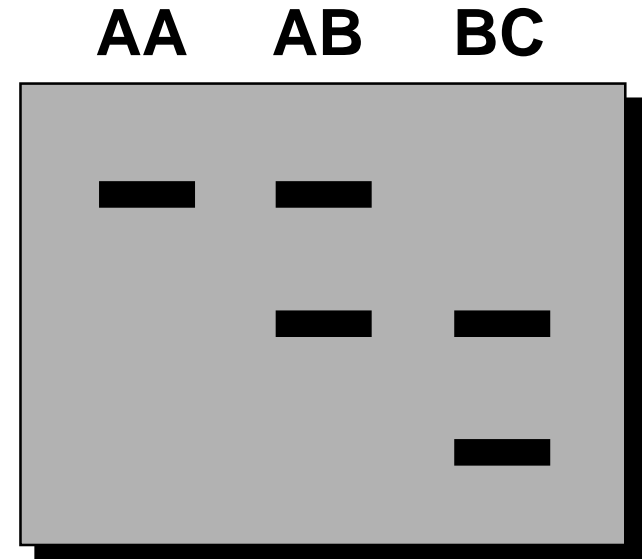
# Polymorphism - variable amount of repetitive sequences

## Variable number of tandem repeats (VNTR)



Target site of the enzyme  
new / eliminated

'A', 'B', a 'C' are alleles



Southern hybridization  
 $^{32}\text{P}$ - probe



# **RFLP – detection of sickle cell anemia**

**$\beta$ -hemoglobinopathy**

- High percentage of genetic diseases**
- Millions affected in all continents**

**sickle cell anemia**

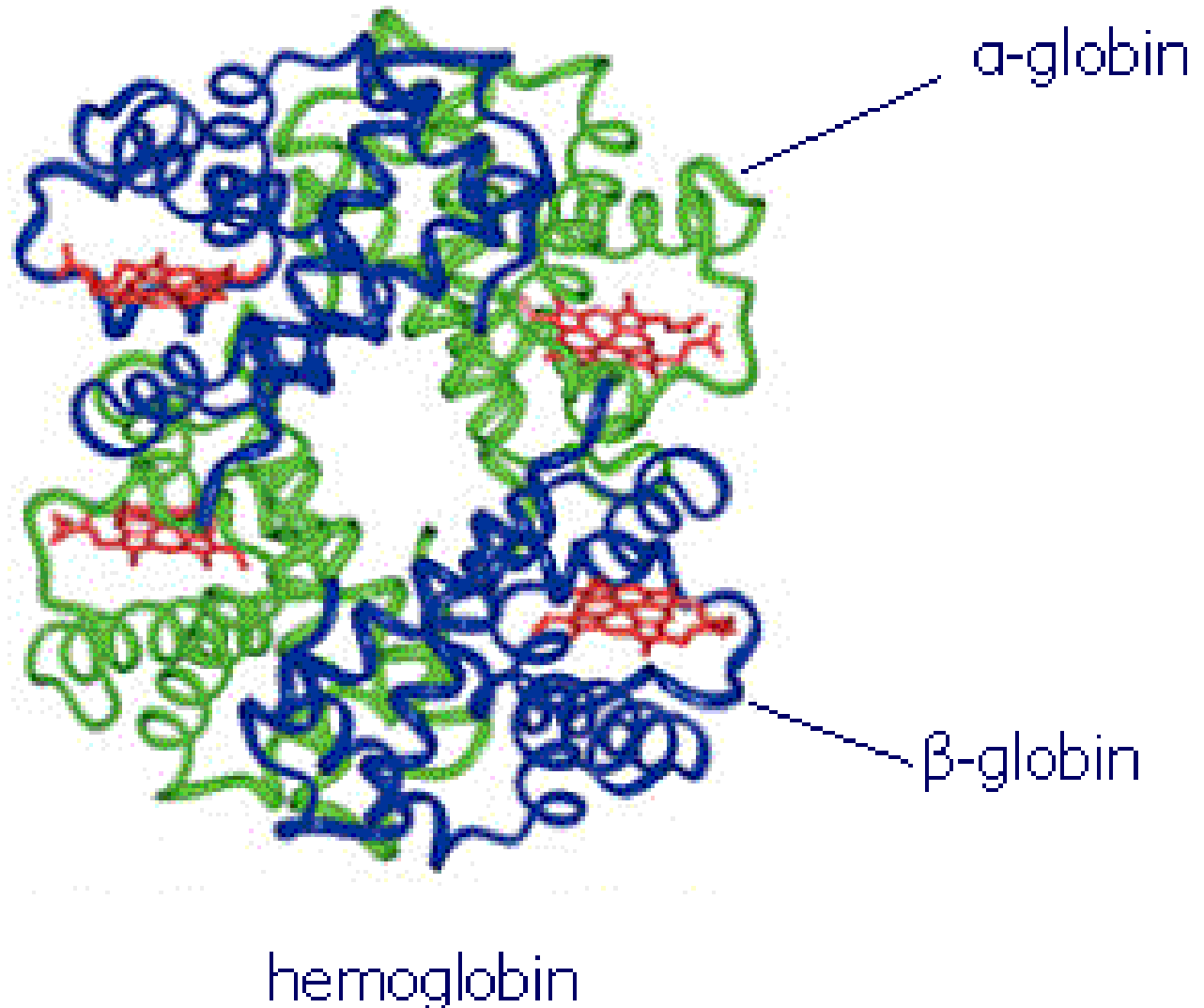
**~1/15 of US black people - gen SCD (sickle cell disease)**

**Whole life treatment ~ \$150,000/year for each SCD patient**

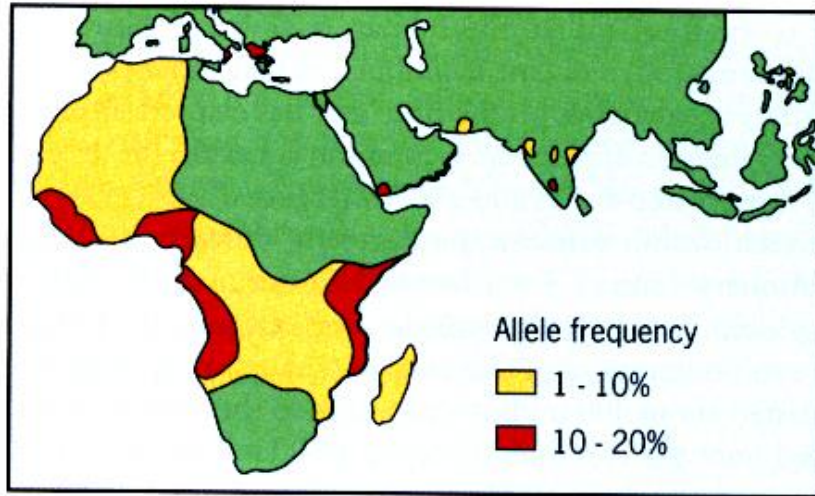
**Hemoglobin** - adults  
- fetus

**HbA** - 2  $\alpha$  and 2  $\beta$  chains

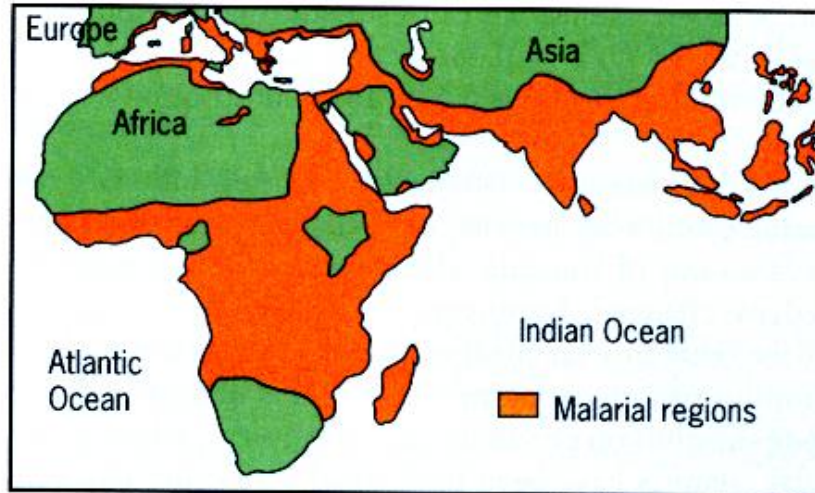
**HbF** - 2  $\alpha$  and 2  $\gamma$  (higher affinity to O<sub>2</sub>)



# Heterozygotes HbS – cca 40% HbS – resistant to malaria



(a) Distribution of sickle-cell anemia allele ( $Hb^s$ ).



(b) Distribution of falciparum malaria.

**Lower flexibility and passage through capillaries**

***Plasmodium falciparum* – parasite in erythrocytes – increased pH by 0.4 enhanced adhesion of erythrocytes (that should be degraded in spleen)**

**Sickle erythrocytes**

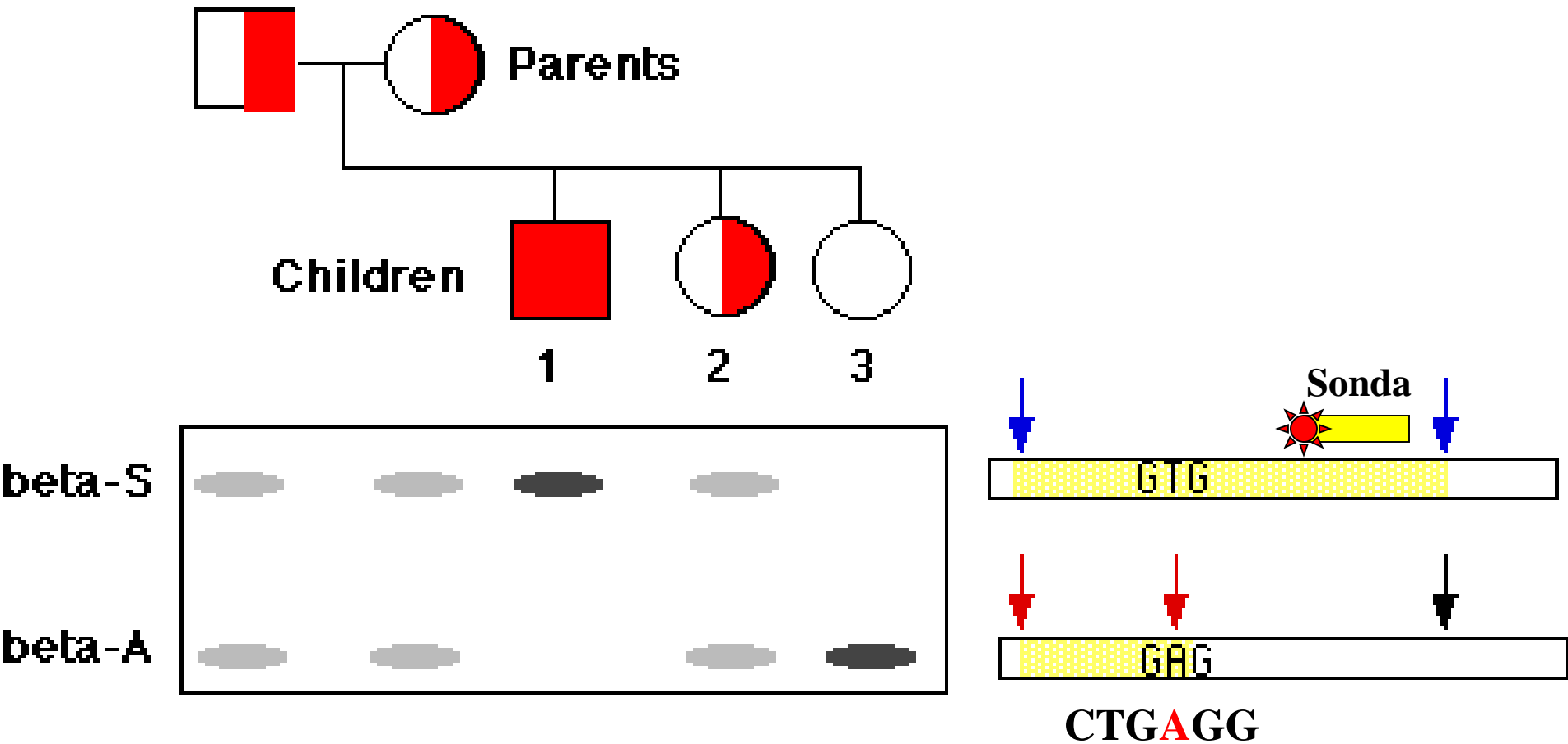
- Lower concentration of  $K^+$  (higher membrane permeability)**
- Limitation with  $K^+$  for *Plasmodium***

	...	A	C	T	C	C	T	G	T	G	G	A	G	...	beta <sup>S</sup> gene
Mutation			Thr			Pro			Val			Glu			beta <sup>S</sup> chain
			Thr			Pro			Glu			Glu			beta <sup>A</sup> chain
	...	A	C	T	C	C	T	G	A	G	G	A	G	...	beta <sup>A</sup> gene
Codone #		4			5			6			7				

## sickle cell anemia

### -Genetic disorder

- homozygous point mutation – both genes for HbS
- both genes encode Val instead of Glu in β-hemoglobine



GAG codon (for Glu) → GTG for Val ⇒  
 Elimination of target sequence CTGAGG for restriction endonuclease **MstII**

# Interspersed sequences - intergene regions

**Satellites** tandem highly repetitive

- length of repeats up to several thousands bps
- clusters at centrosomes and telomeres, frequently also on Y chromosome

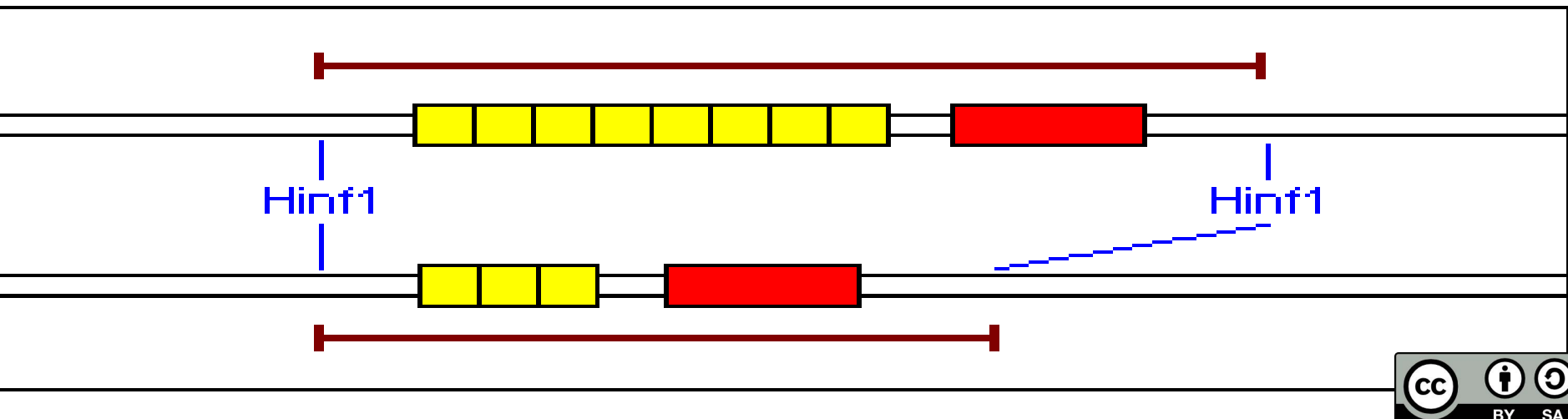
**Minisatellites** (14 to 500 bps, usually 15 bps) repetitions

Average length ~ 0.5 to 30 kb

In mammals, fungi and plants

**Microsatellites** moderately repetitive (1-13 bp) repetitions in mammals, insects, plants. Human genome ~ 30 000 microsatellite loci

Copy number in population variable ~ 10 to 100





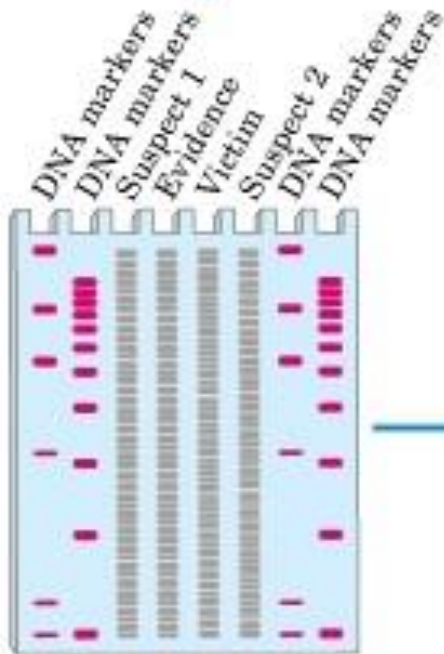
Chromosomal DNA  
(e.g., Suspect 1)

Cleave with restriction  
endonucleases.



DNA fragments

Separate fragments by agarose  
gel electrophoresis (unlabeled).



Denature DNA, and  
transfer to nylon  
membrane.



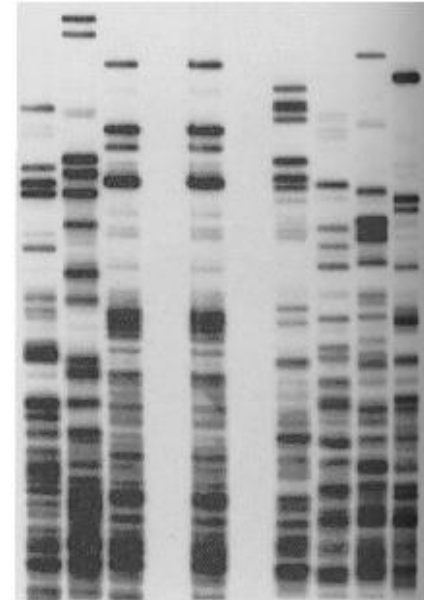
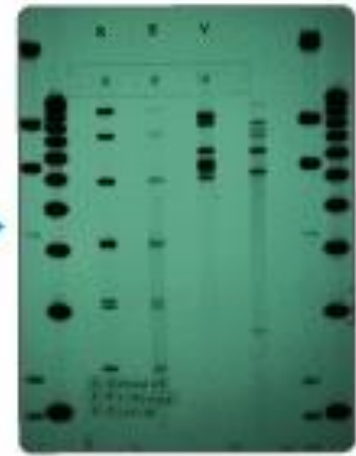
Incubate  
with probe,  
then wash.



Radiolabeled  
DNA probe



Expose x-ray  
film to  
membrane.



1 2 3 Bloodstain 4 5 6 7  
Suspects



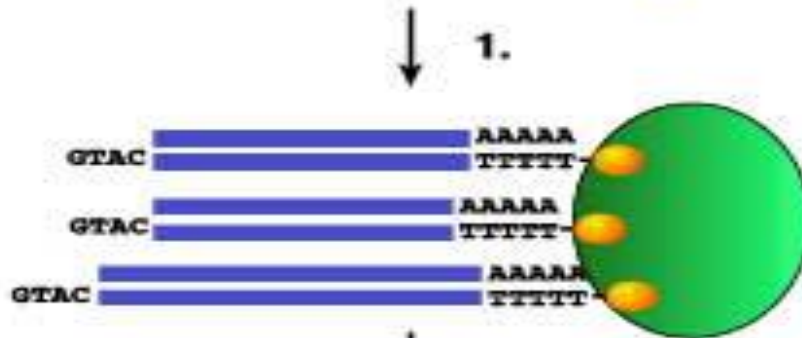
# SAGE

## serial analysis of gene expression

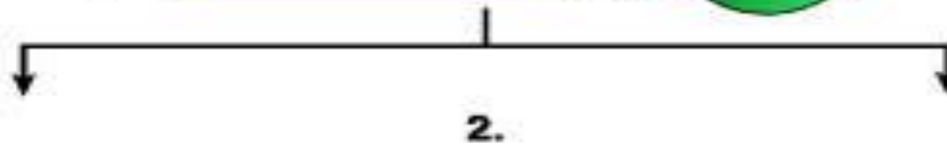
Transformation of mRNA  
to dsDNA



Digest *Nla*III →  
overhanging chain



2 aliquots



Joining of linkers



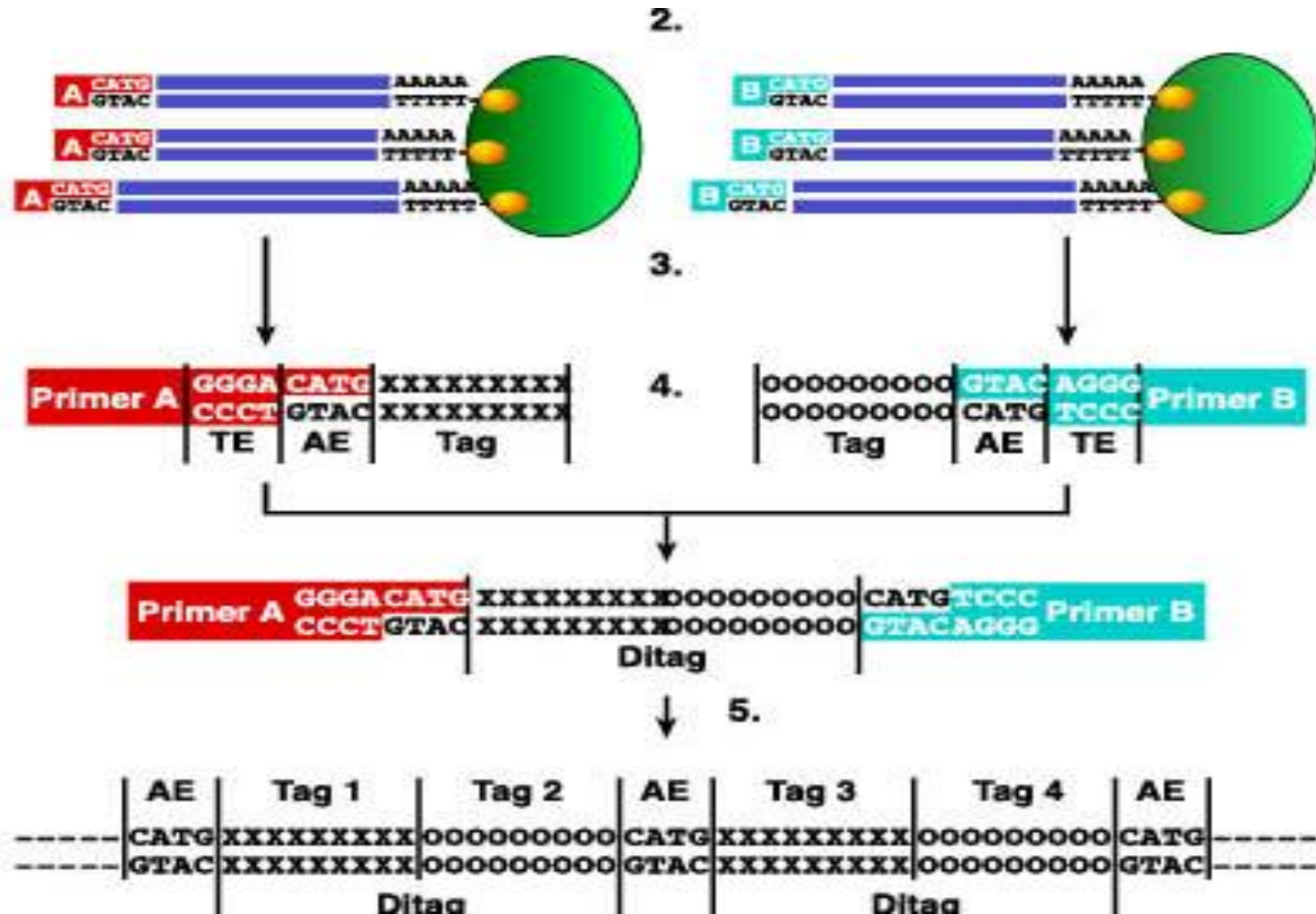
# SAGE

Linkers have  
PCR & Tagging  
Endonuclease

Cut with TE  
BsmF1

Mix both aliquots  
Blunt-end ligate  
to make "Ditag"

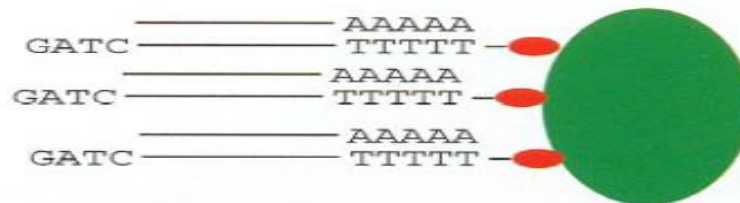
Concatenate  
& Sequence



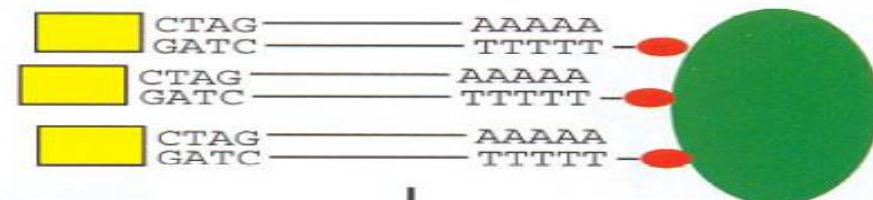
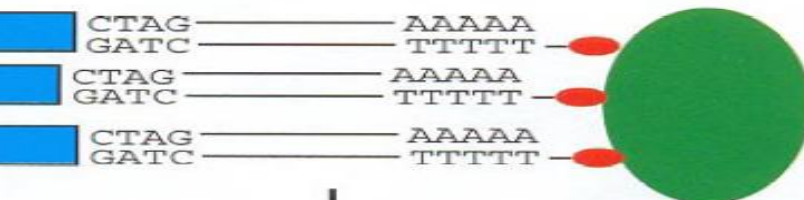


↓

Cleave with anchoring enzyme (AE) and bind to streptavidin beads



↙ Divide in half and ligate to linkers A and B

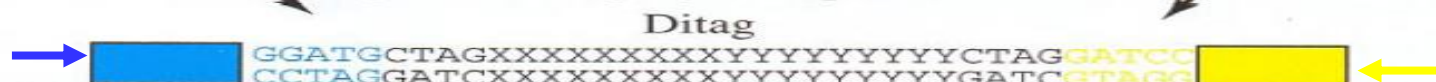


↓

Cleave with tagging enzyme (TE)



↘ Ligate and amplify with primers A and B



↓

Cleave with anchoring enzyme  
Isolate ditags  
Concatenate and clone

CTAGXXXXXXXXXXXXYYYYYYYYYYCTACXXXXXXXXXXXXYYYYYYYYYYCTACXXXXXXXXXXXXYYYYYY  
GATCXXXXXXXXXXXXYYYYYYYYYYGATCXXXXXXXXXXXXYYYYYYYYYYGATCXXXXXXXXXXXXYYYYYY

Ditag

Ditag

Ditag

# Serial Analysis of Gene Expression

>GS0017BA.SCF

```
nnnnnnXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXAAGGATAGCAGGCGCATTTACAACATGCGC
CTGGTCCAGCTTGCCTGCATCATGTACGTAAAAAACCTATCATTGCCATGGAG
ATTAACGGCCACTGCTTGCATGCCCGCCGTGGGGCTTGAGAGACAATGCAGAA
AAAAATACTATTACGCATGTTCAAGAATAAATGTCGGCCGCCATGCCCTCAAGAG
ACACCATCTTGCAGTCATGCCAGCAACGCGAACCACAGAGAGCATGAAGCGCACC
GTATAGTCAGCATAATGCTCCTAGACTGCCTCCGGTTACATGTGGGTTCCTTC
CTATGATGGTACATGCAGCAGAAGTAAGTTGTTAGAGACATGTAAGGCGTGTGGG
GGCTGTCTACATGCAGCTCCTCCGCCTCTTTGTCTCATGGAGACCCTTCGCCAGA
GCGTTTACATGTATGTTGTGAATAGAGAGGTGGACATGCCGAGATCGATAGGAAA
TTACCTCAnnnnATCAAGGGGGCCTGGTTTCATGACCGTGGAATAACTTCTTCT
TGATGCGGTGGAGACTTTTGTGAGCAAAAATGTCCTATTAAGGTCTTGTAXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXnnnnn
```

# Expression Profile in *Giardia* Trophozoites

<u>Tag</u>	<u>Freq.</u>	<u>Contigs</u>	<u>ORFs</u>	<u>Function</u>
CATGACTGCAAGAT	2.3658%	1	1	<b>ornithine carbamoyltransferase</b>
CATGGAGGAGGACG	1.2603%	2	2	<b>alpha-2-tubulin</b>
CATGTGAGTGTGAG	1.2293%	1	0	putative giardin chain
CATGTAAACGCTCT	0.9861%	1	0	putative 40S ribosomal protein S7
CATGAGACAAAGAG	0.9021%	1	1	putative dynein light chain
CATGACGACGAAGG	0.7650%	1	1	<b>giardin gamma chain</b>
CATGGTACAAGTAA	0.7341%	1	0	<b>fructose-1,6-bisphosphate aldolase</b>
CATGCTCAACAGCA	0.6810%	1	1	<b>glyceraldehyde 3-phosphate dehydrogenase</b>
CATGAAGCGCACCG	0.6191%	2	2	<b>elongation factor-1 alpha</b>
CATGCAGCTCCTCC	0.4466%	1	1	putative 14-3-3 protein
CATGGAACGCCTTT	0.4068%	1	0	putative 40S ribosomal protein S21
CATGCAGAGATCAA	0.3449%	1	0	putative 40S ribosomal protein S28
CATGGCTGCAAACC	0.3051%	1	0	<b>carbamate kinase</b>
CATGGCGGCCGACA	0.2432%	1	1	putative dynein heavy chain
CATGGGGAGACTTG	0.1459%	1	1	putative lysyl-tRNA synthetase



# How many tags need to be sequenced?

For yeast (Velculescu *et al.*, 1997 Cell 88:243-251)...

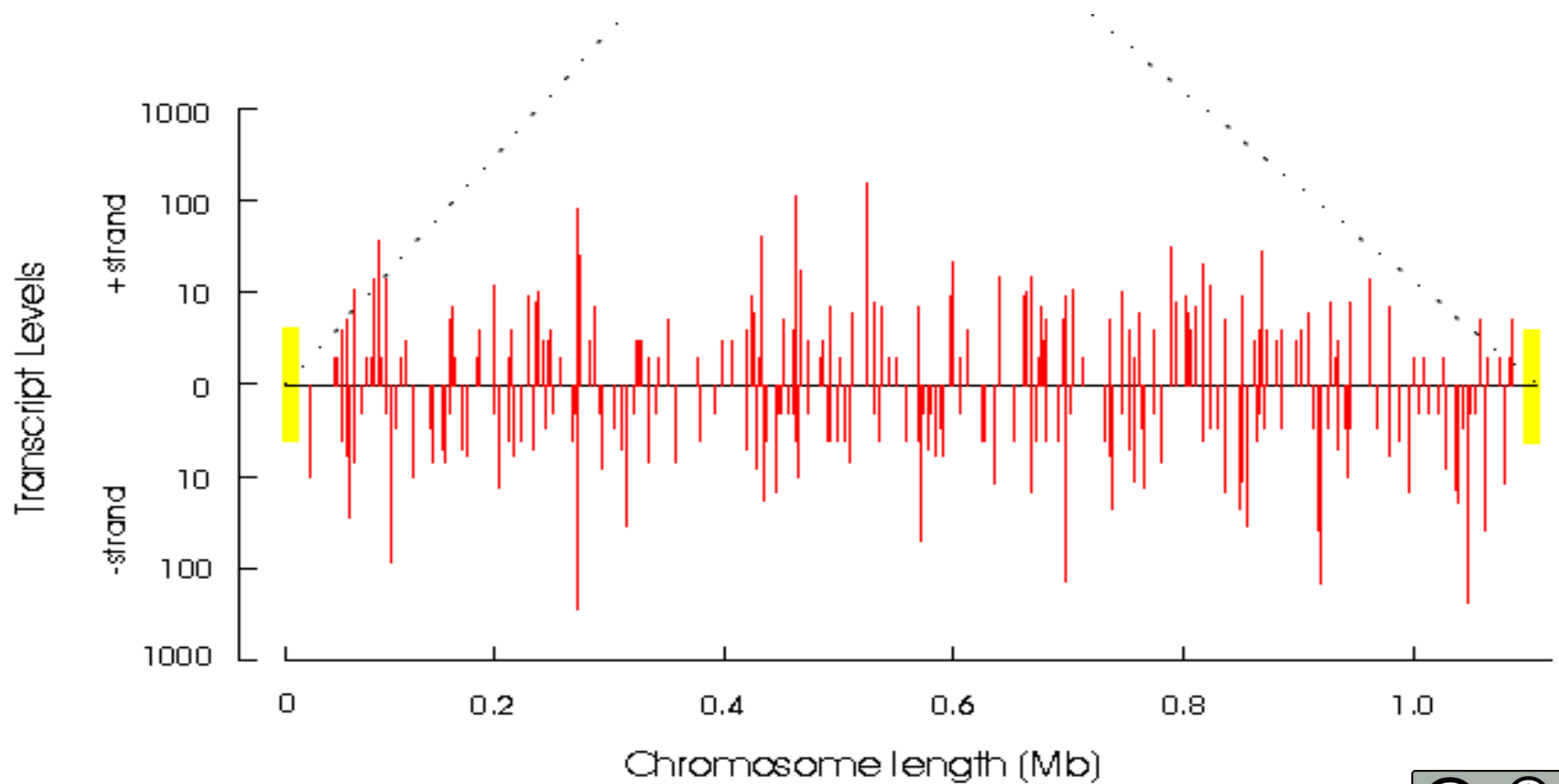
- assume 15,000 mRNA molecules / cell
- if 20,000 tags were sequenced...

...represents 1.3 coverage for single copy mRNA

...72% chance of detecting single copy message  
(based on Monte Carlo modeling method)

...95% of predicted genes expected to be detectable  
based on presence *Nla*III site

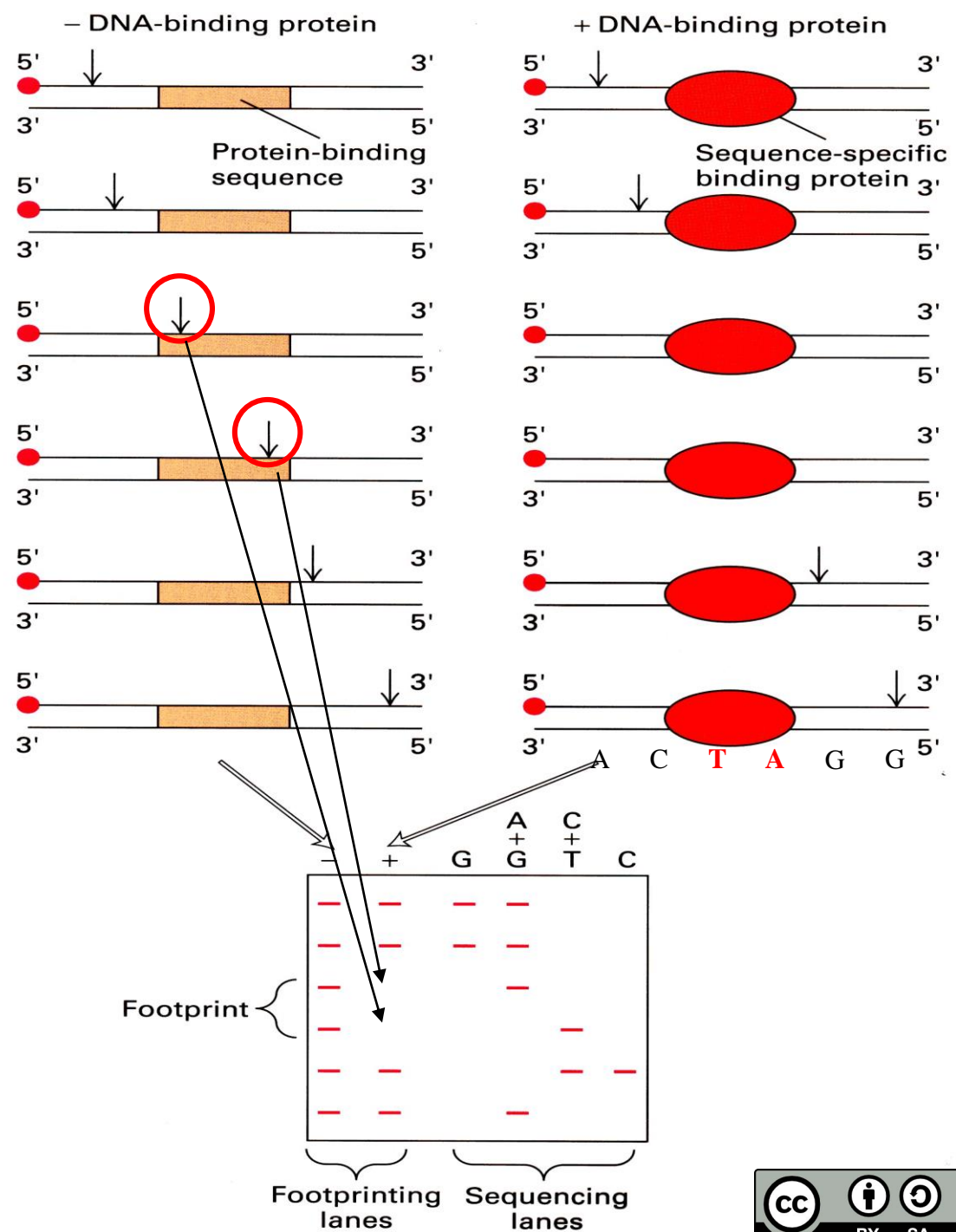
# SAGE of Yeast Chromosome

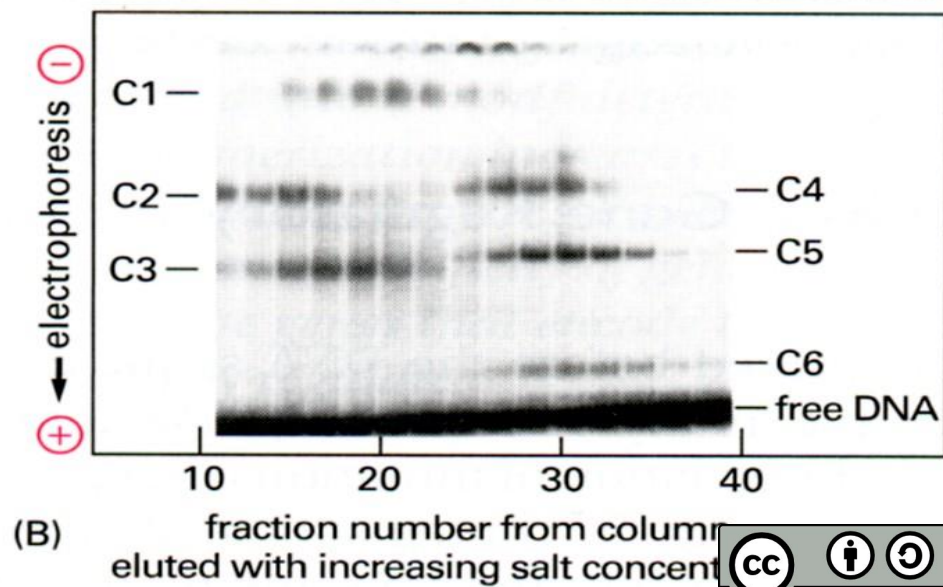
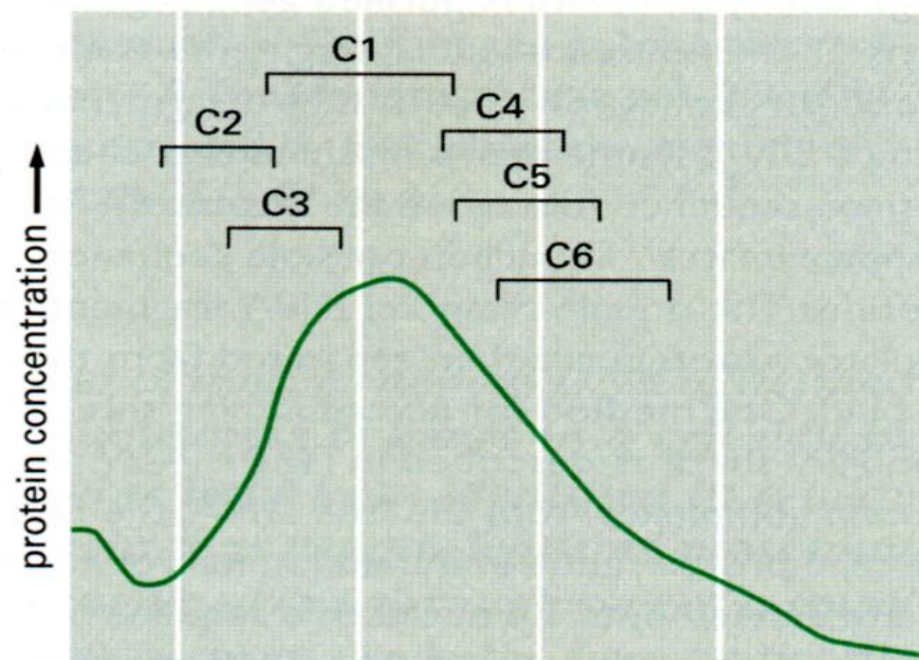
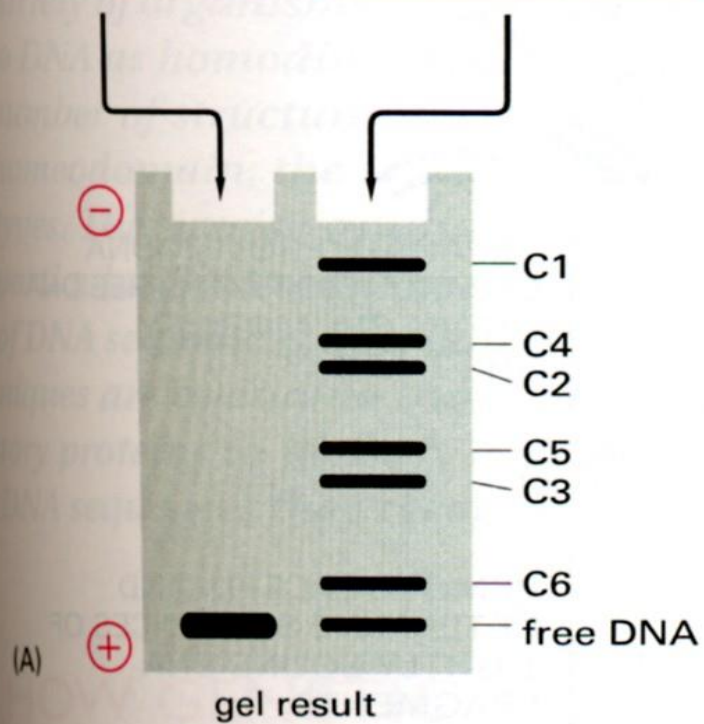
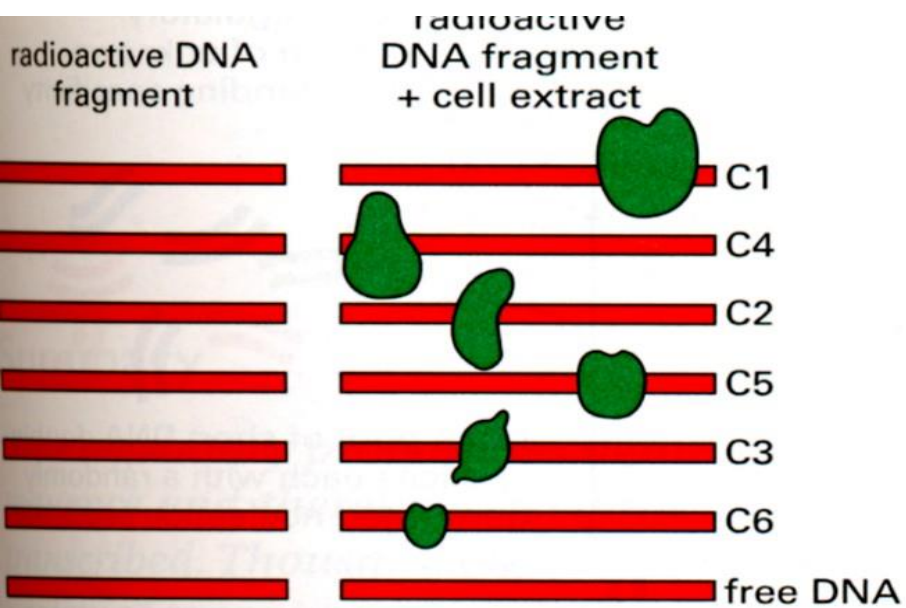






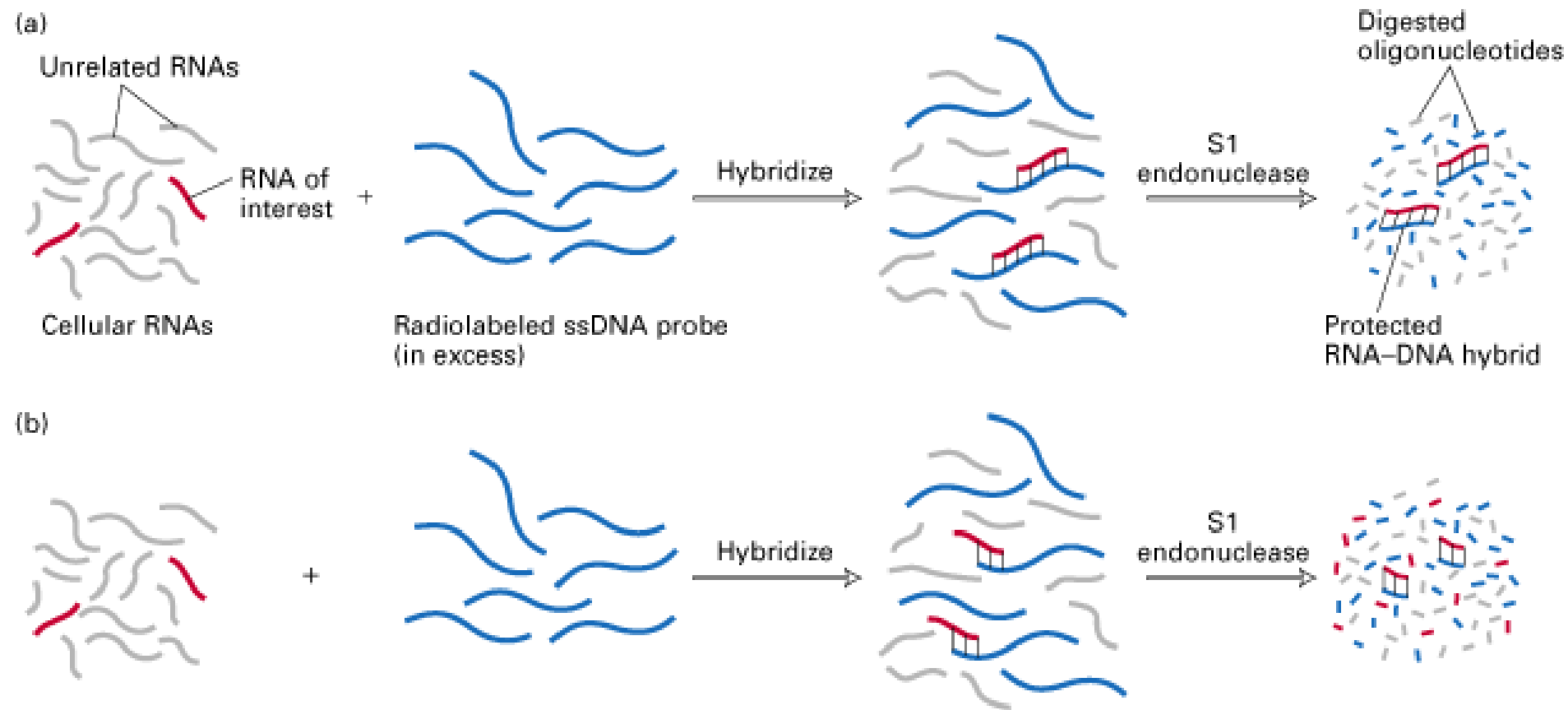
**Gel shift assay (EMSA electrophoretic mobility SA)**  
complex **DNA-protein**



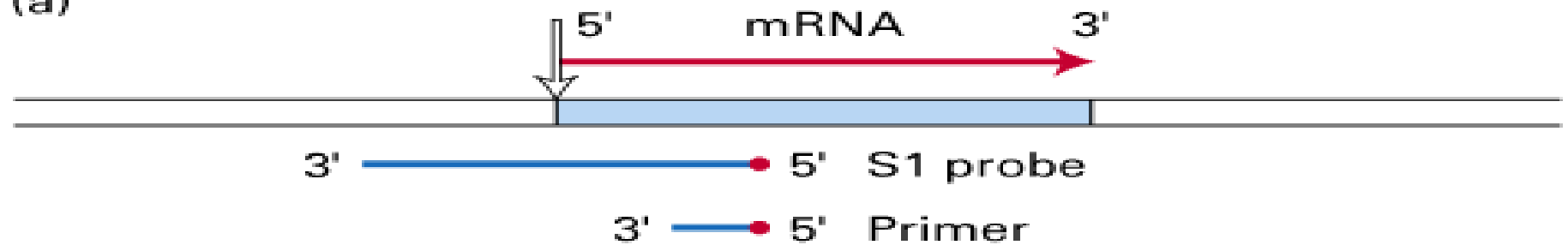


# Binding of RNA to DNA

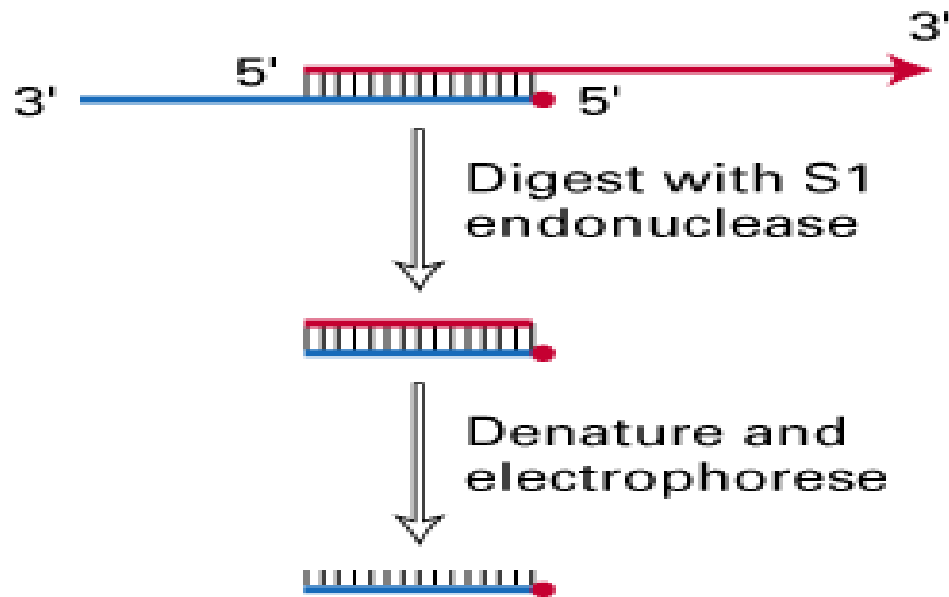
## RNA protection assay



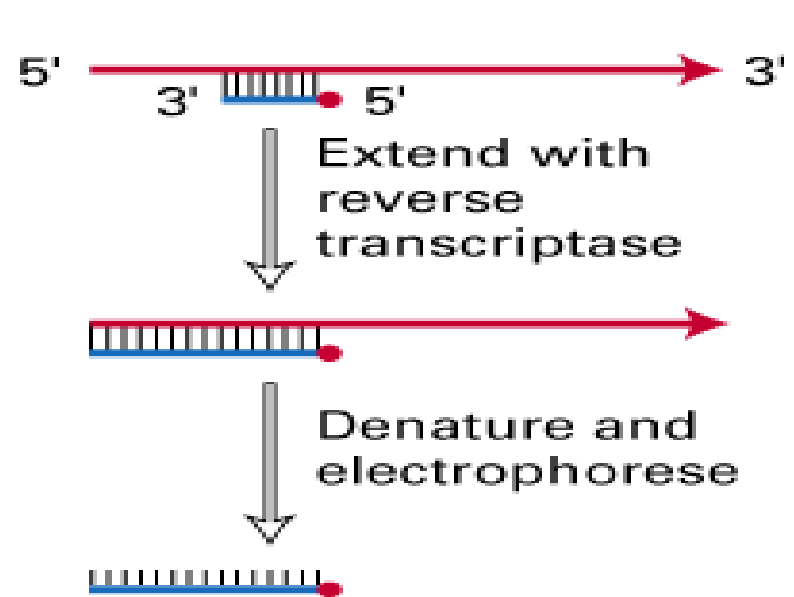
(a)



(b)



(c)





**Reporter genes** – mapping of regulatory sequences *in vivo*

# Subtraction hybridization

RNA from  
experimental tissue

RNA from  
control tissue

Reverse transcription

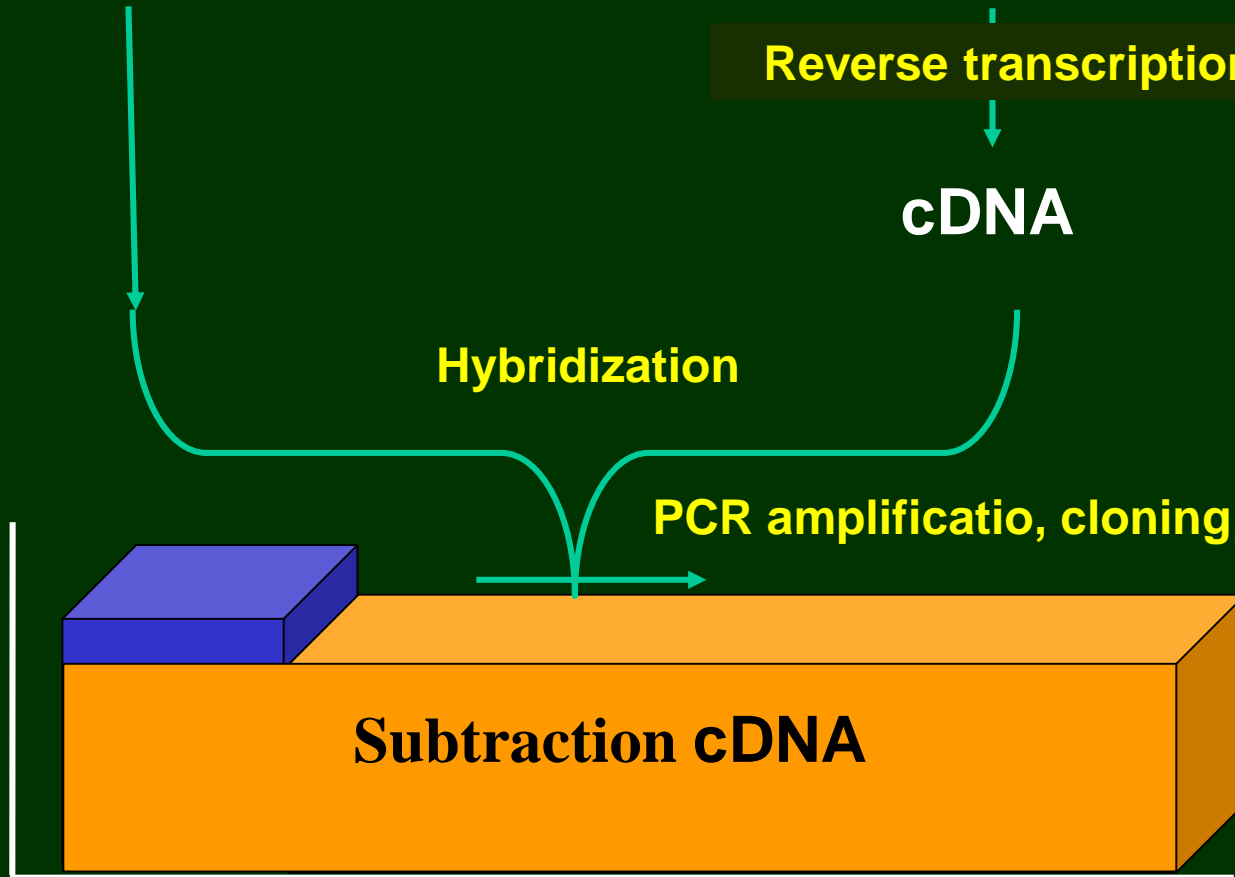
cDNA

Hybridization

PCR amplification, cloning, sequencing

Amt

Subtraction cDNA





mRNA from target cells

mRNA from subtractor cells



TTTTTTTTTT  
AAAAAAAAAAAA

Elute

AAAAAAAAAAAA  
AAAAAAAAAAAA  
AAAAAAAAAAAA

AAAAAAAA  
TTTTTTTTTT

Solid-phase cDNA synthesis  
denaturation

TTTTTTTTTT  
TTTTTTTTTT  
TTTTTTTTTT  
TTTTTTTTTT  
TTTTTTTTTT  
TTTTTTTTTT

Subtracted mRNA

AAAAAAAAAAAA

Common mRNA

AAAAAAAAAAAA  
TTTTTTTTTT

TTTTTTTTTT

AAAAAAAAAAAA  
TTTTTTTTTT

Uveřejněné materiály jsou určeny studentům Vysoké školy chemicko-technologické v Praze

jako studijní materiál. Některá textová i obrazová data v nich obsažená jsou převzata

z veřejných zdrojů. V případě nedostatečných citací nebylo cílem autora/ů záměrně poškodit event. autora/y původního díla.

S eventuálními výhradami se prosím obračejte na autora/y konkrétního výukového materiálu, aby

bylo možné zjednat nápravu.

The published materials are intended for students of the University of Chemistry and Technology, Prague as a study material. Some text and image data contained therein are taken from public sources. In the case of insufficient quotations, the author's intention was not to intentionally infringe the possible author(s) rights to the original work.

If you have any reservations, please contact the author(s) of the specific teaching material in order to remedy the situation.

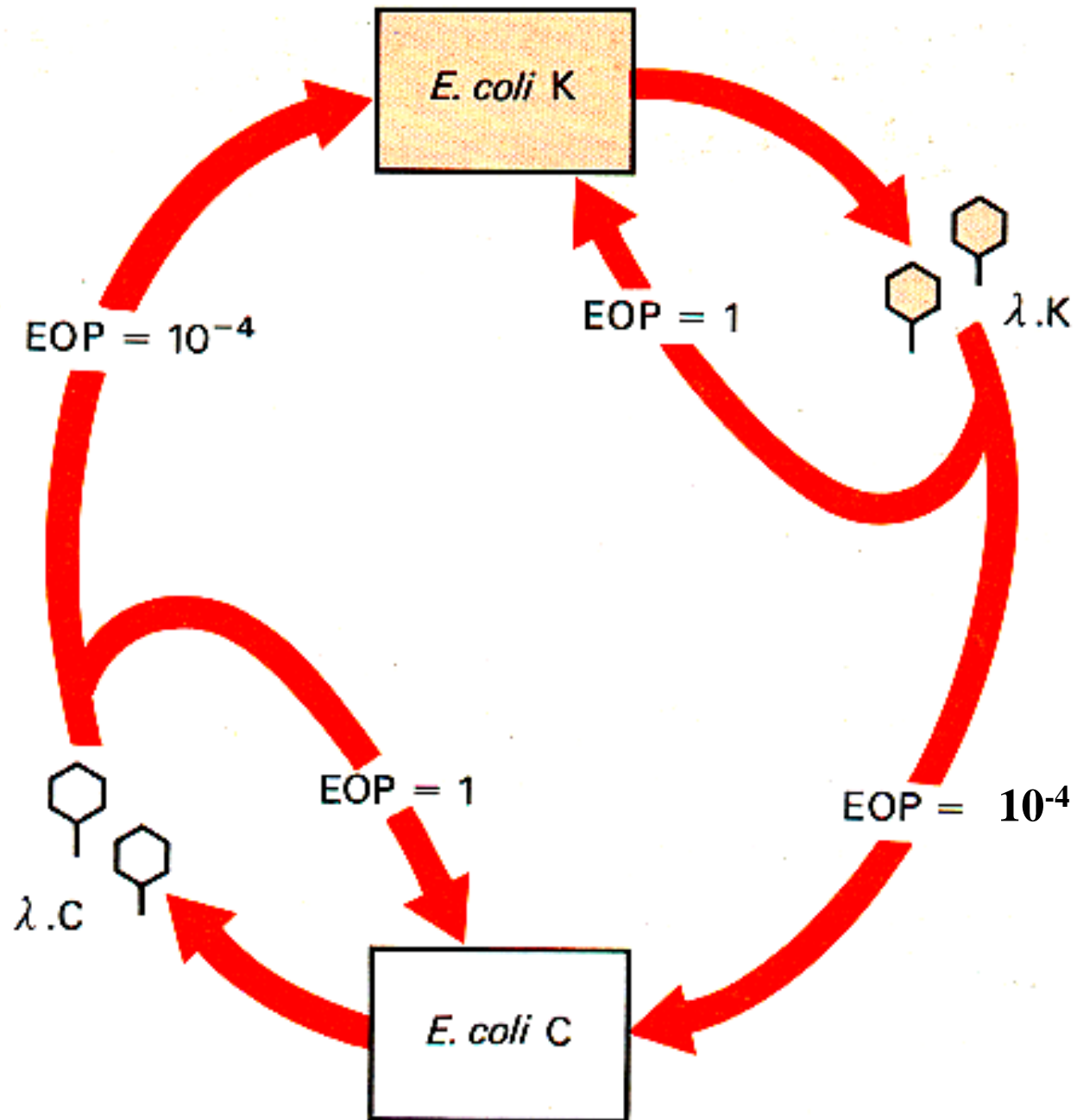
# Restriction endonucleases



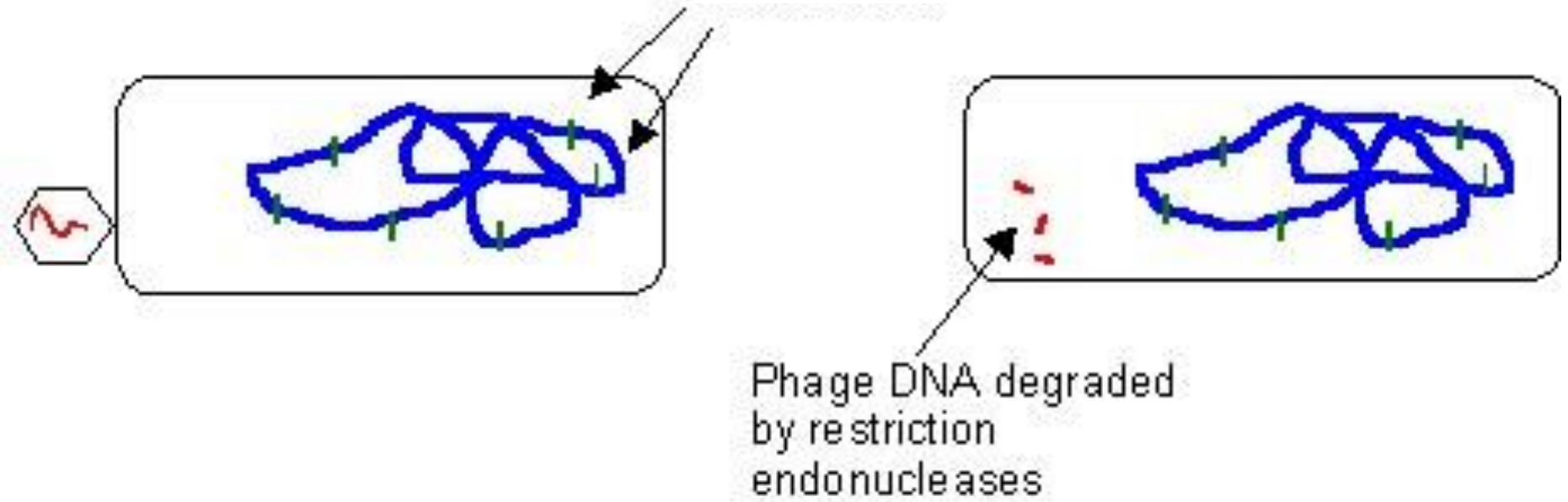
EUROPEAN UNION  
European Structural and Investing Funds  
Operational Programme Research,  
Development and Education



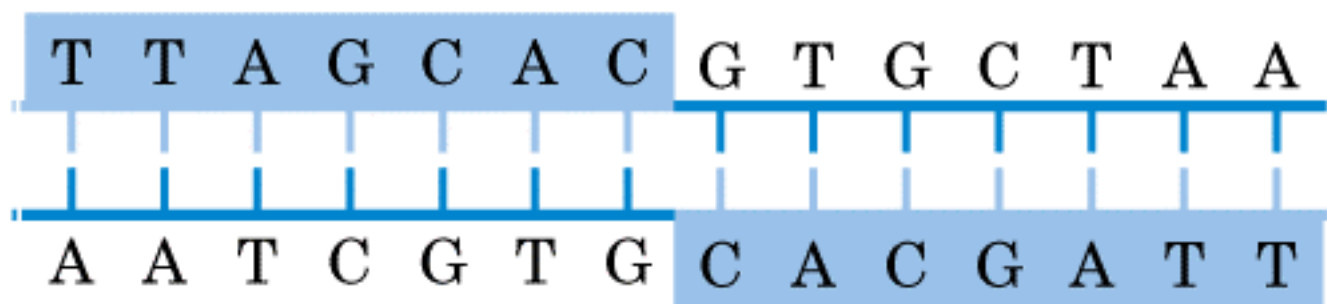
MINISTRY OF EDUCATION,  
YOUTH AND SPORTS



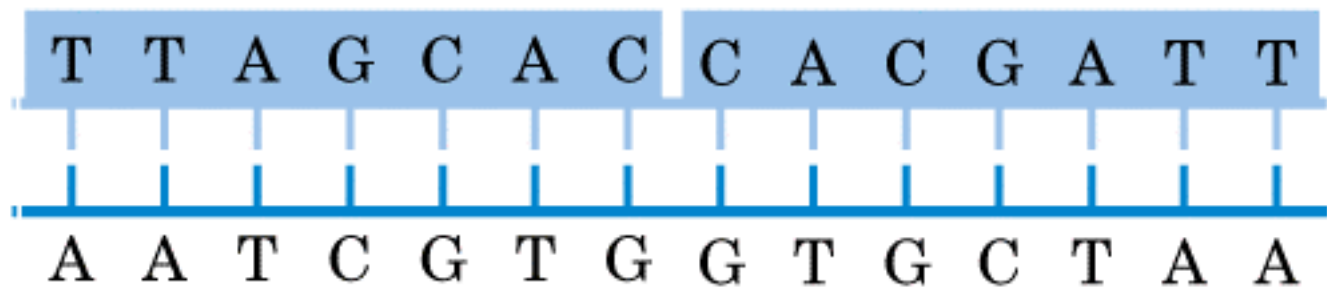
Chromosomal DNA  
modification

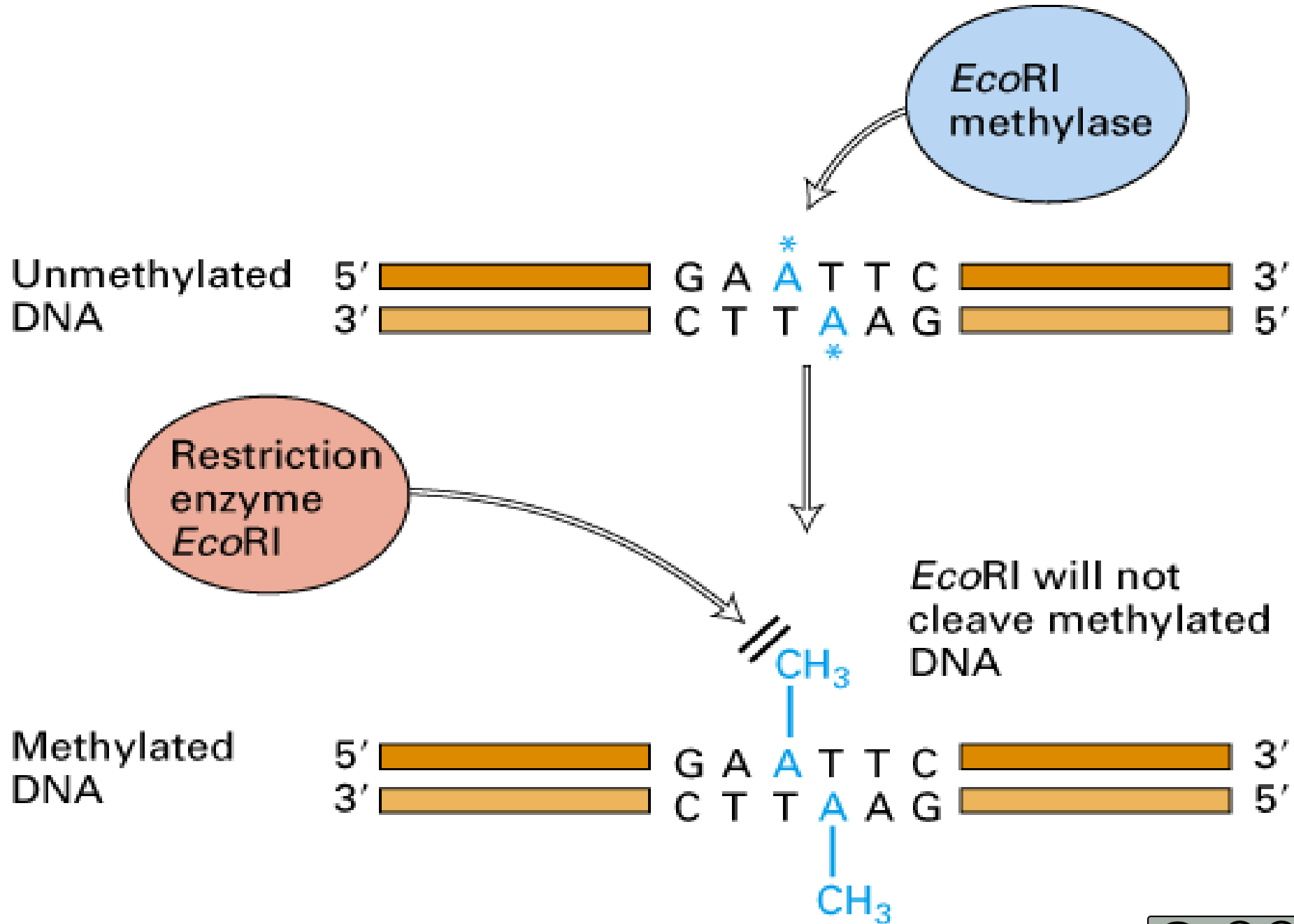


## Palindrome



## Mirror repeat







Restriction  
enzyme  
*EcoRI*

Unmethylated DNA

5' [orange bar] G A A T T C [orange bar] 3'

3' [orange bar] C T T A A G [orange bar] 5'

Cleavage

Sticky ends

5' [orange bar] G [bracket] A A T T [bracket] C [orange bar] 3'

3' [orange bar] C [bracket] T T A A [bracket] G [orange bar] 5'

<b>Name</b>	<b>Source Microorganism</b>	<b>Recognition Sequence</b>
<b>Bam HI</b>	<b>Bacillus amyloliquefaciens</b>	<b>G ↓ GATCC</b>
<b>Eco RI</b>	<b>Eschericia coli RY13</b>	<b>G ↓ AATTC</b>
<b>Hind III</b>	<b>Haemophilus influenzae Rd</b>	<b>A ↓ AGCTT</b>
<b>Not I</b>	<b>Nocardia otitidis-caviarum</b>	<b>GC ↓ GGCCGC</b>
<b>Pst I</b>	<b>Providencia stuartii</b>	<b>CTGCA ↓ G</b>
<b>Sma I</b>	<b>Serratia marcescens</b>	<b>CCC ↓ GGG</b>

# Three groups of RE


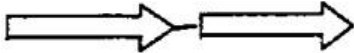

I Binding - specific recognition sequence breaks in an undefined site outside

III - recognize specific sequences (not necessarily symmetric)  
cleaves outside this region

I and III - modification and ATP-dependent restriction activity

II palindromic sequence

Separated restriction and methylation activity

	Type I	Type II	Type III
	<ul style="list-style-type: none"> <li>Hetero-oligomeric enzymes</li> <li>Require ATP hydrolysis for restriction</li> <li>Cut DNA at sites remote from the recognition sequence</li> <li>DEAD-box proteins</li> </ul>	<ul style="list-style-type: none"> <li>ENase and MTase separate enzymes</li> <li>Cut DNA within, or close to, their recognition sequence</li> </ul>	<ul style="list-style-type: none"> <li>Hetero-oligomeric ENase</li> <li>ATP required for restriction</li> <li>Cut DNA close to recognition sequence</li> <li>DEAD-box proteins</li> </ul>
	e.g. <i>EcoKI</i>	e.g. <i>EcoRI</i>	e.g. <i>StyLTI</i>
Genes	<i>hsdR</i> <i>hsdM</i> <i>hsdS</i>	<i>ecorIR</i> <i>ecorIM</i>	<i>mod</i> <i>res</i>
			
Subunits	HsdR    HsdM    HsdS	Res    Mod	Mod    Res
	<hr/>	<hr/>	<hr/>
	MTase	ENase    MTase	MTase
	<hr/>		<hr/>
Activities	ENase, ATPase and MTase		ENase, ATPase and MTase

**1 U - 1  $\mu$ g DNA**

**1 h at optimal conditions**

Work with RE

Izoschizomers

# Star activity

High glycerol concentration

high enzyme to DNA ratio ( $> 100 \text{ U}/\mu\text{g DNA}$ ).

low ionic strength.

High pH.

Presence of organic solvents

(DMSO, ethanol, ethylenglycol)

Exchange of  $\text{Mg}^{2+}$  with other ions  $\text{Mn}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Co}^{2+}$ .

(e.g. *EcoRI*, *BamHI*, *HindIII*, *KpnI* nebo *SalI*)

**EcoR I**  
**#R0101S**

10,000 units

**\$50 (USA)**

for high (5X) concentration, order #R0101T (10,000 units) or #R0101M (50,000 units)

5' ...

**G<sup>^</sup>A A T T C**

... 3'

3' ...

**C T T A A<sup>^</sup>G**

... 5'

**Source:** An *E. coli* strain that carries the cloned EcoR I gene from *E. coli* RY 13 (R. N. Yoshimori)

**Reaction Buffer:** (Supplied with enzyme) NEBuffer EcoR I

50 mM NaCl, 100 mM Tris-HCl, 10 mM MgCl<sub>2</sub>, 0.025% Triton X-100 (pH 7.5 @ 25°C). Incubate at 37°C.

**Ligation and Recutting:** After 100-fold overdigestion with EcoR I, > 95% of the DNA fragments can be ligated and recut.

**Concentration:** 20,000 and 100,000 units/ml. Assayed on lambda DNA.

**Storage Conditions:** 300 mM NaCl, 10 mM KPO<sub>4</sub> (pH 7.5), 0.1 mM EDTA, 10 mM 2-mercaptoethanol, 0.15% Triton X-100, 200 µg/ml BSA, and 50% glycerol. Store at -20°C.

**Diluent Compatibility:** Diluent C

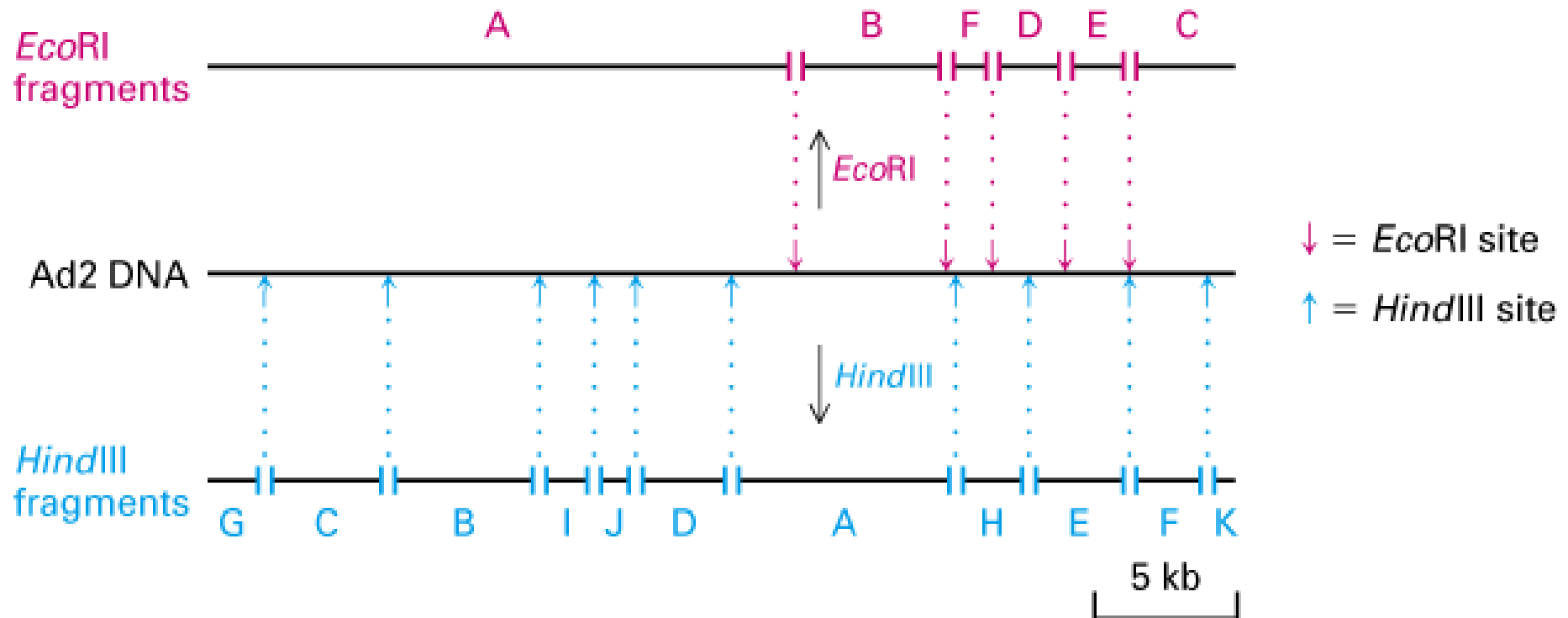
**Heat Inactivation:** 65°C for 20 minutes.

**Note:** Cleavage of mammalian genomic DNA is impaired by overlapping CpG methylation. For performing double digests with EcoR I, click [here](#).

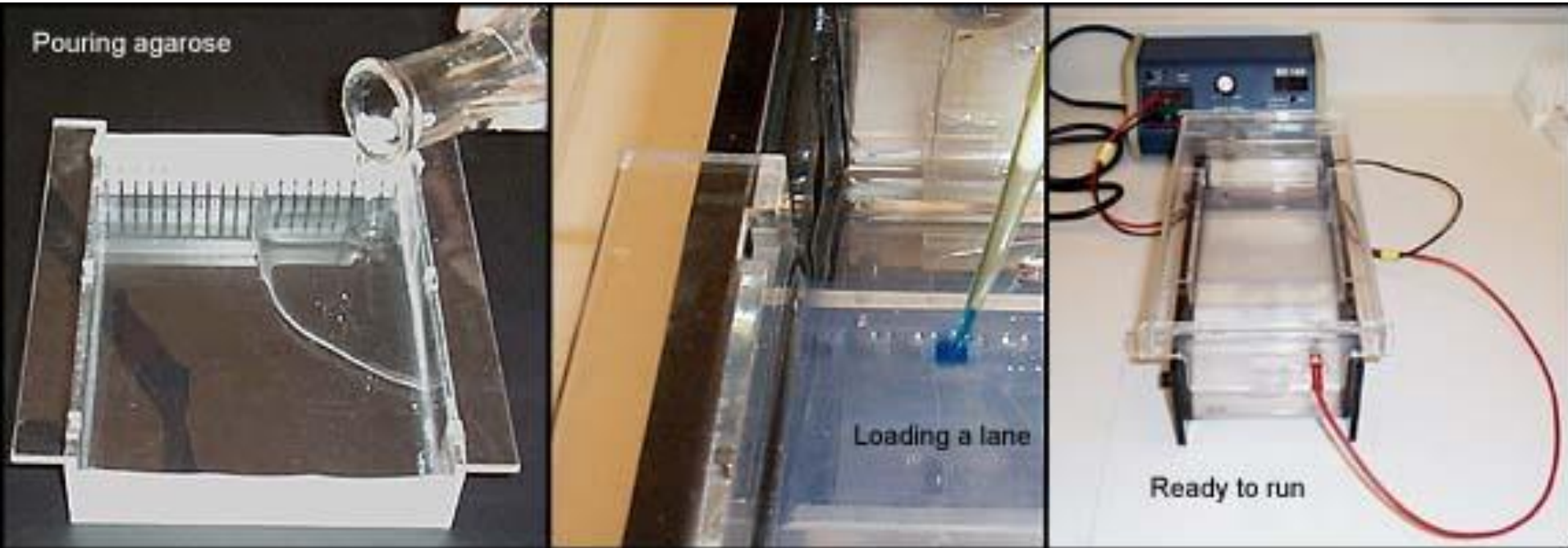
Conditions of low ionic strength, high enzyme concentration, glycerol concentration > 50% or pH > 8.0 may result in [star activity](#).



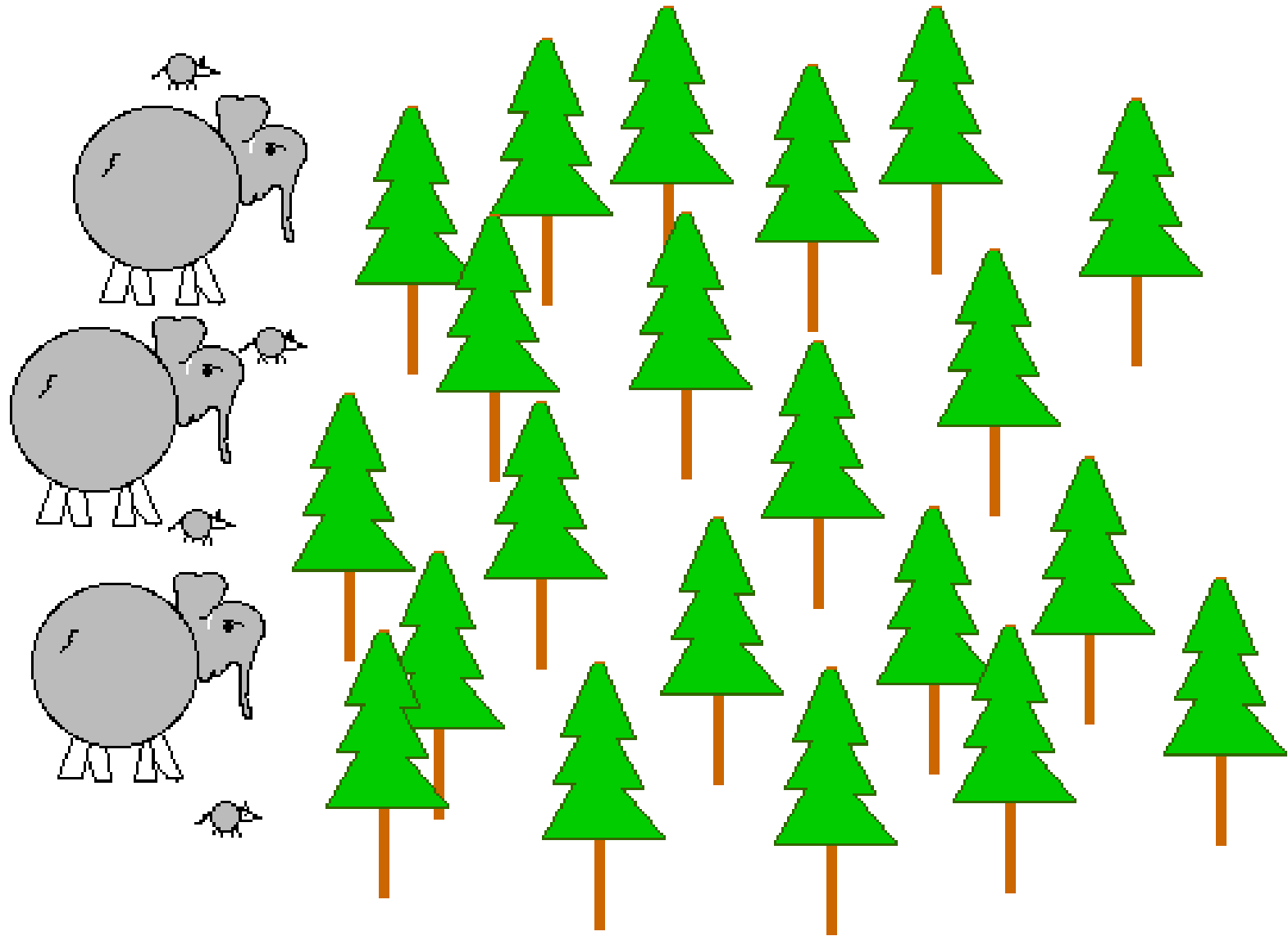


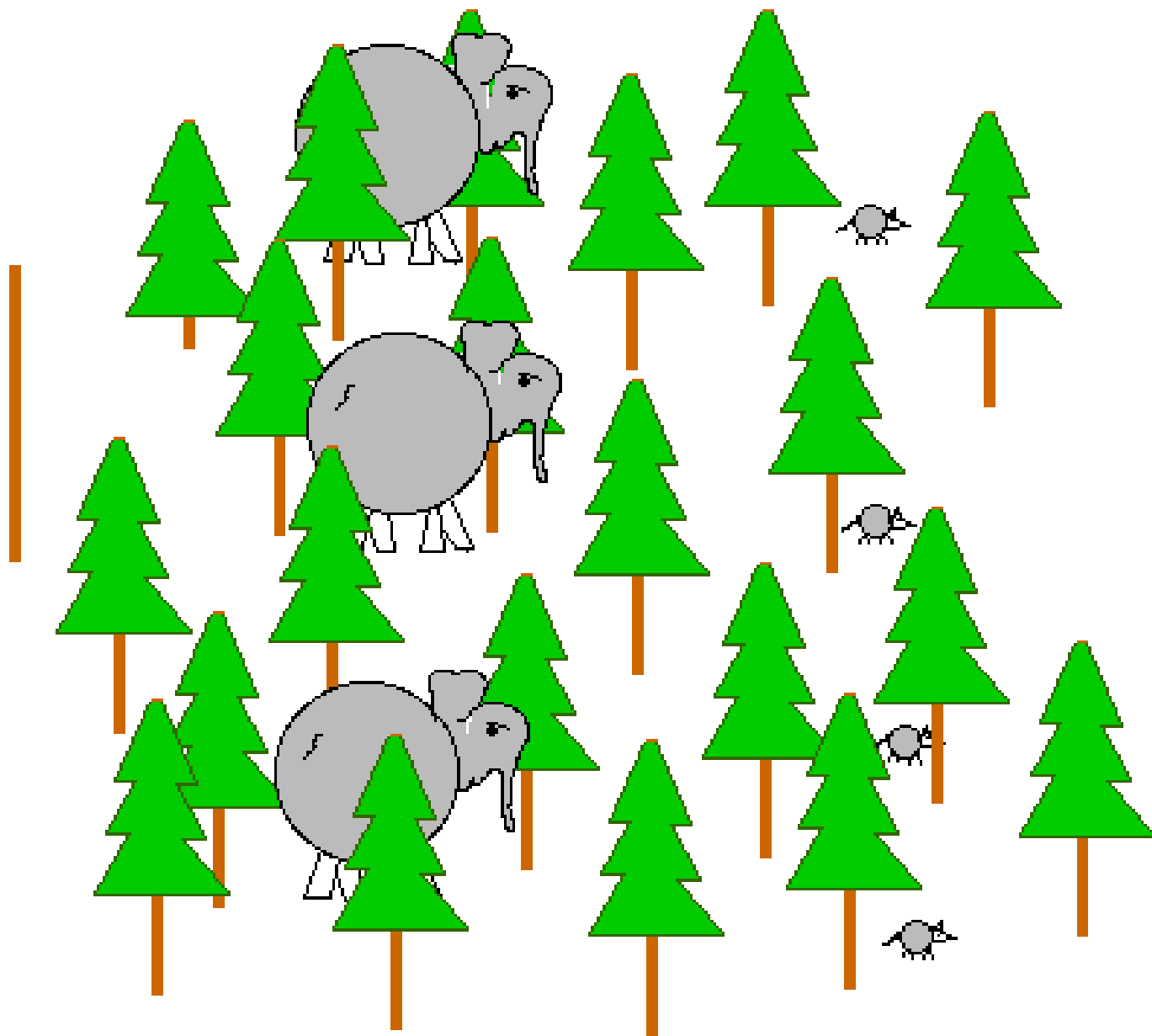


# Analysis of restriction fragments

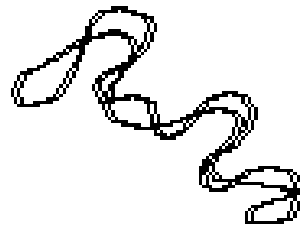


**"Mice run through the forest faster than elephants"**

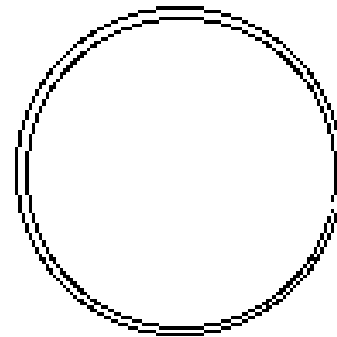




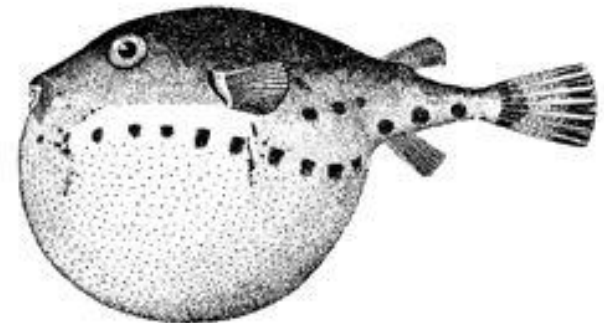
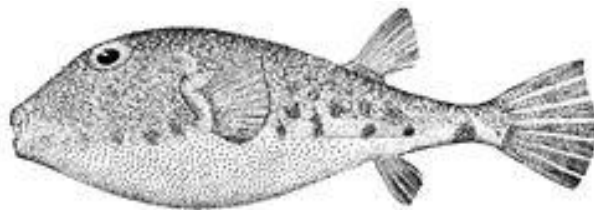
Conformation changes influence the speed of migration

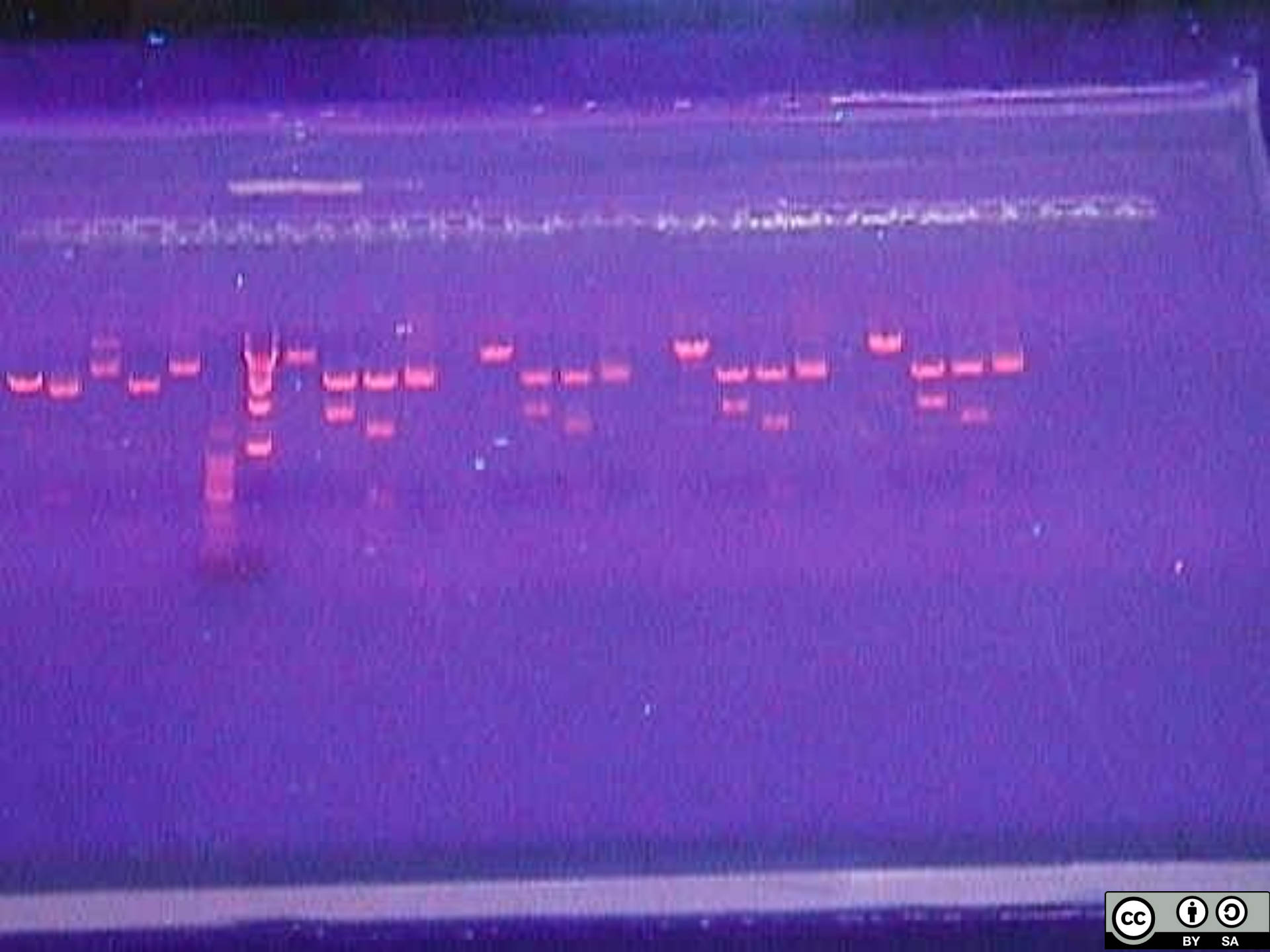


**Supercoiled circle**



**Nicked circle**





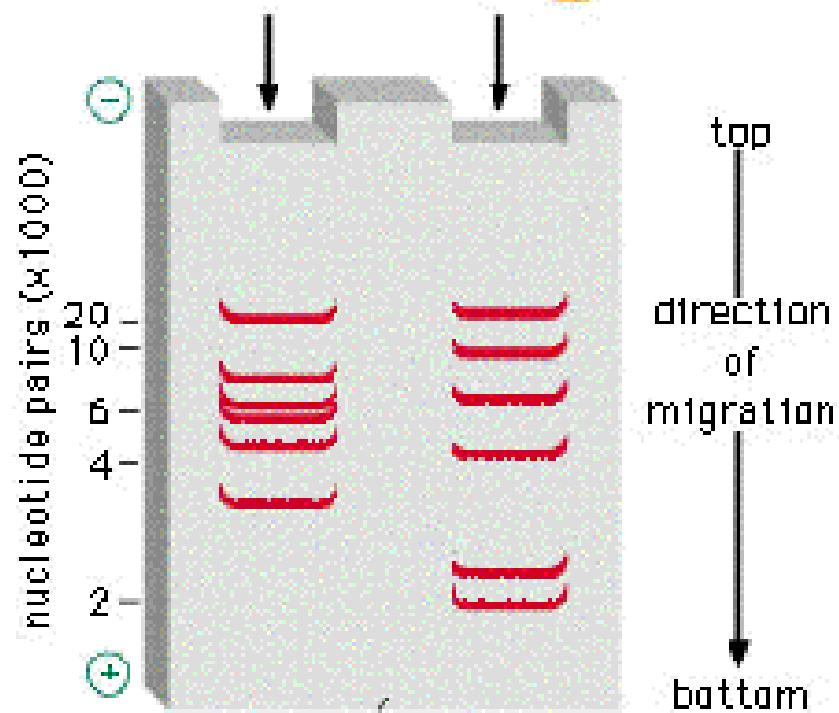
double-stranded  
lambda DNA



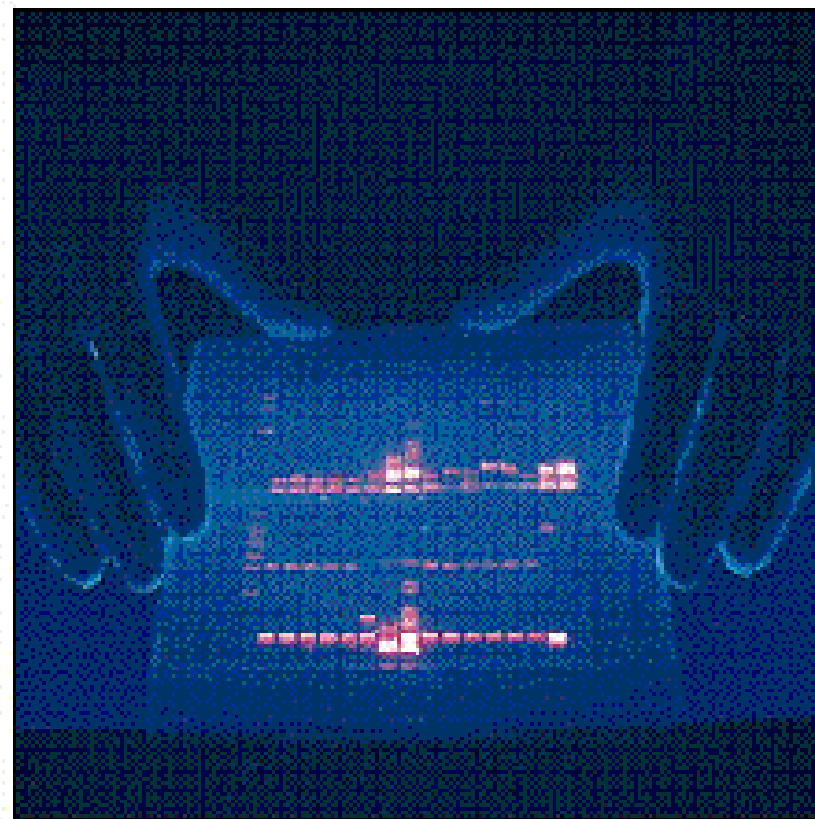
cut with  
Eco RI



cut with  
Hind III

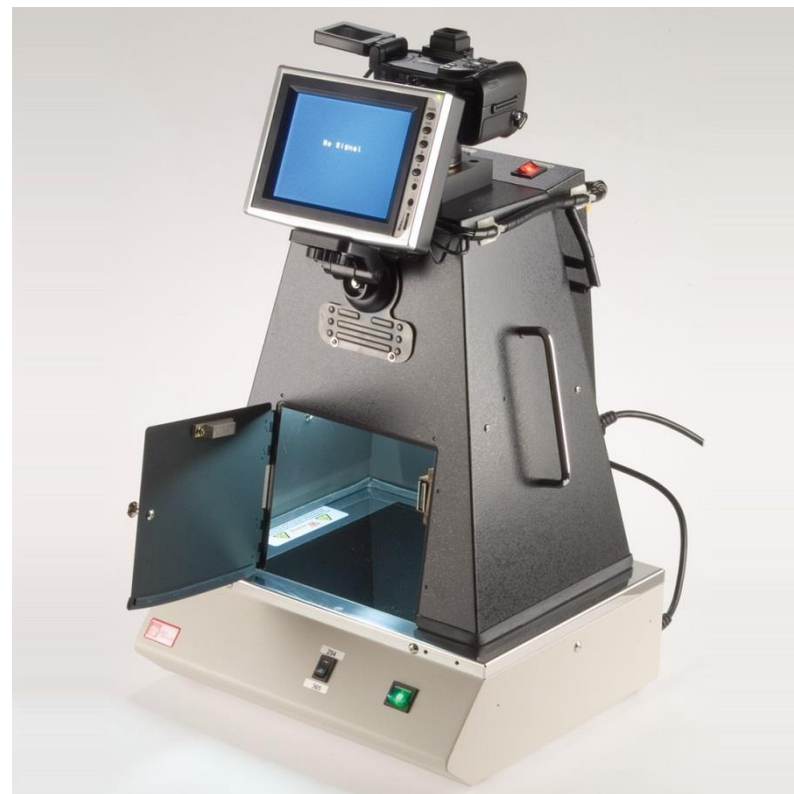


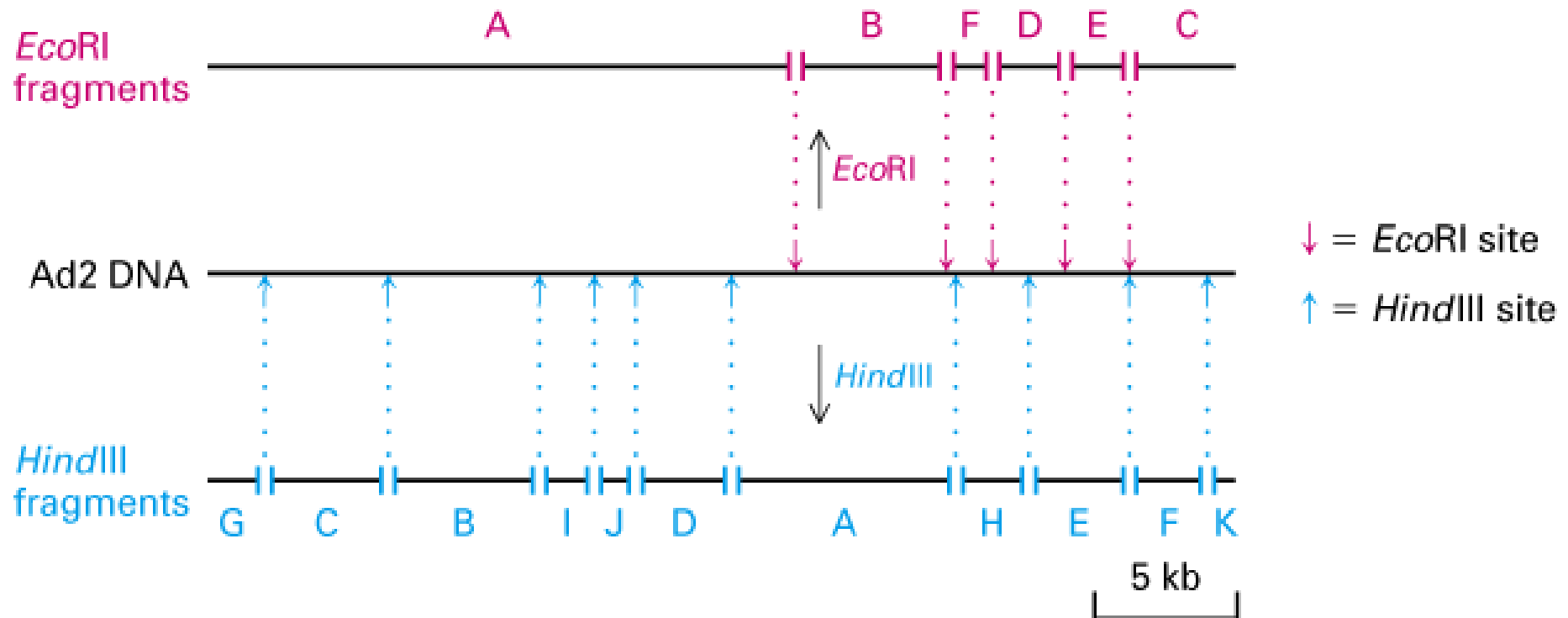
(A) slab of agarose gel



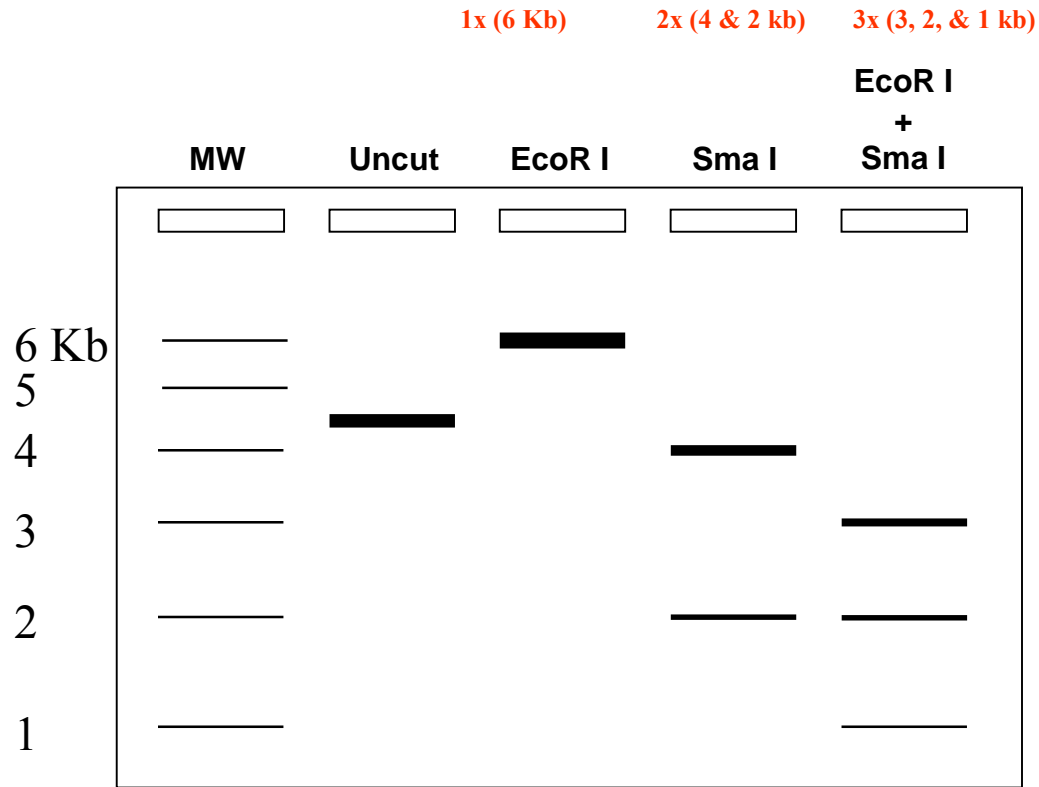
(B)





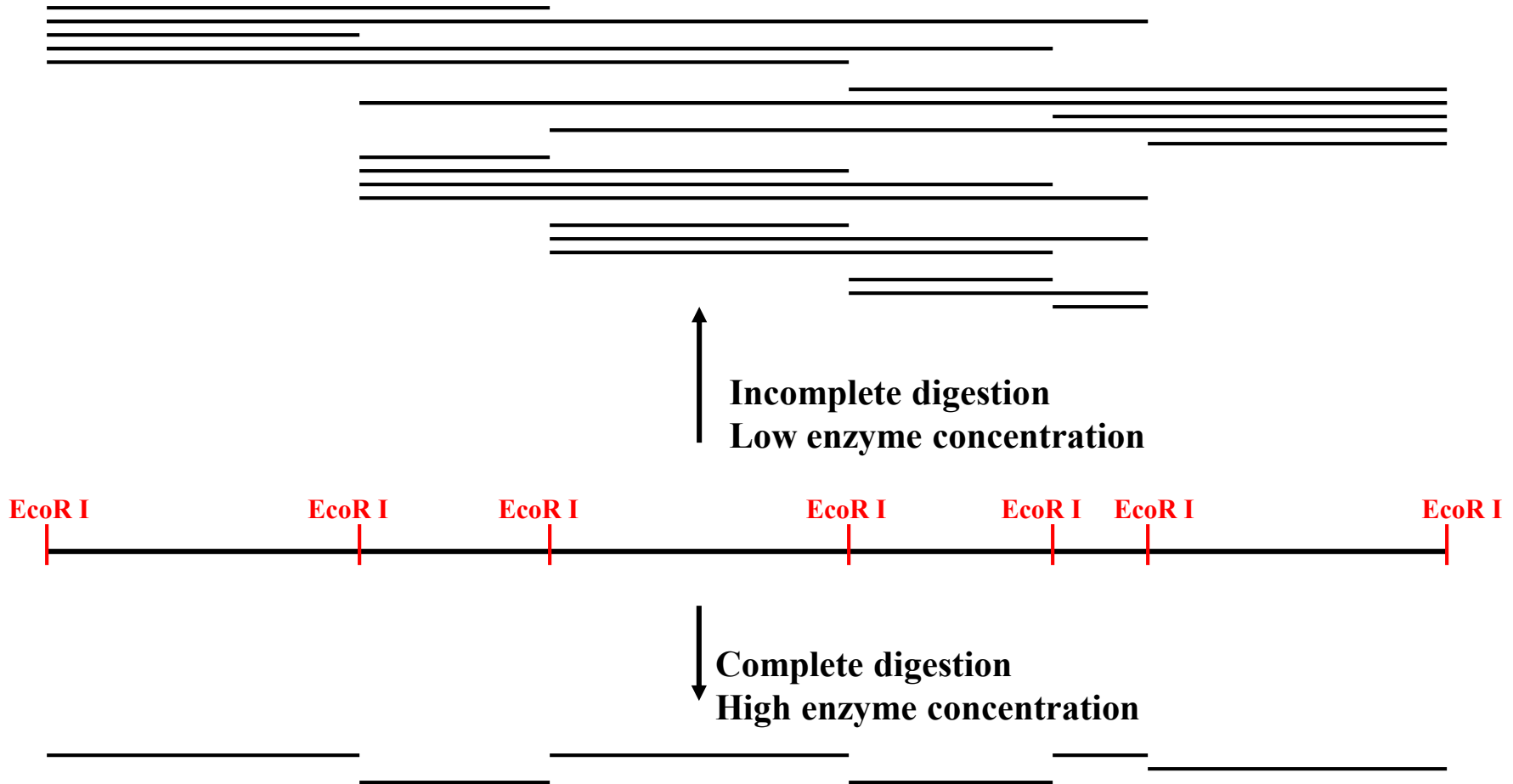


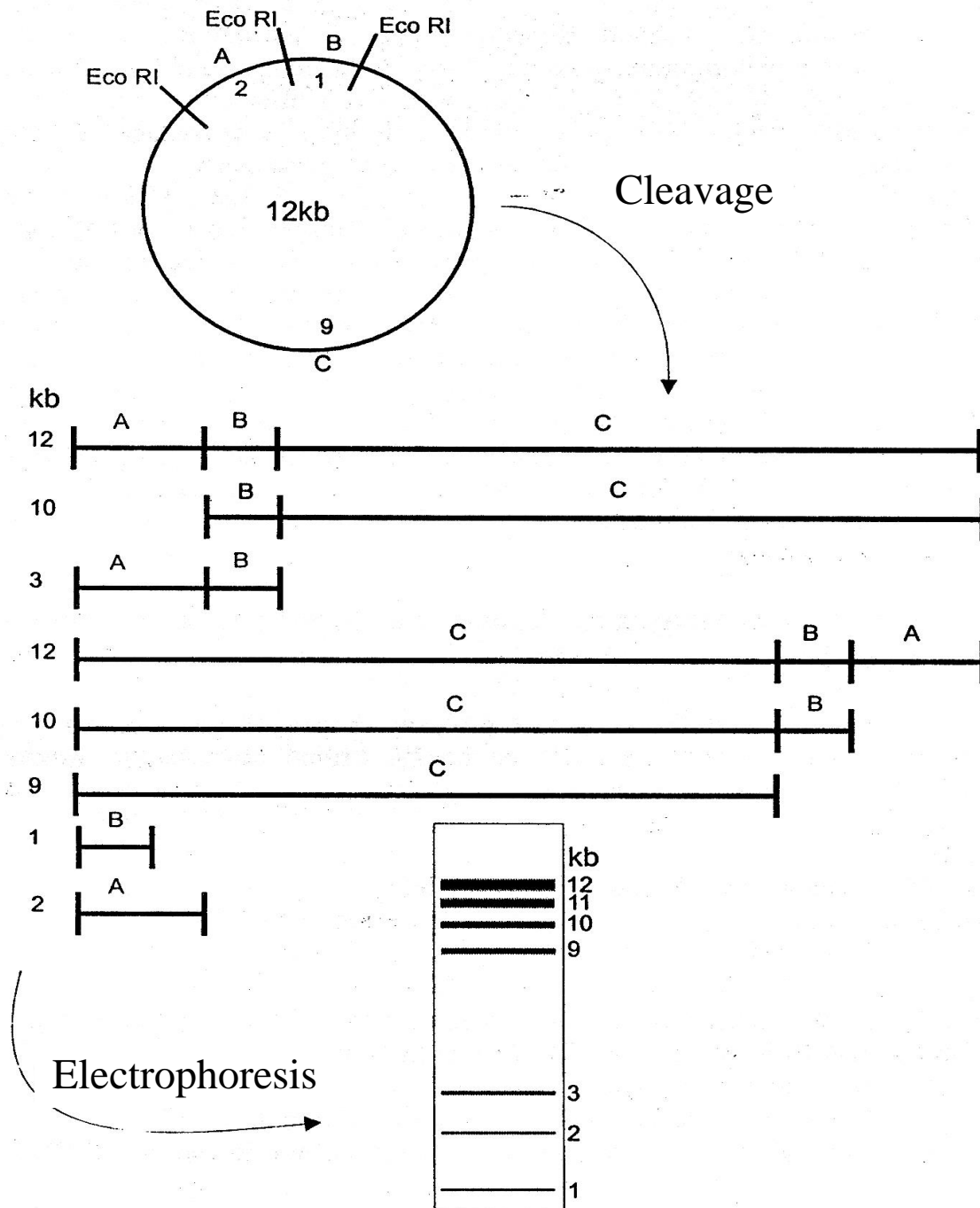
# Restriction mapping



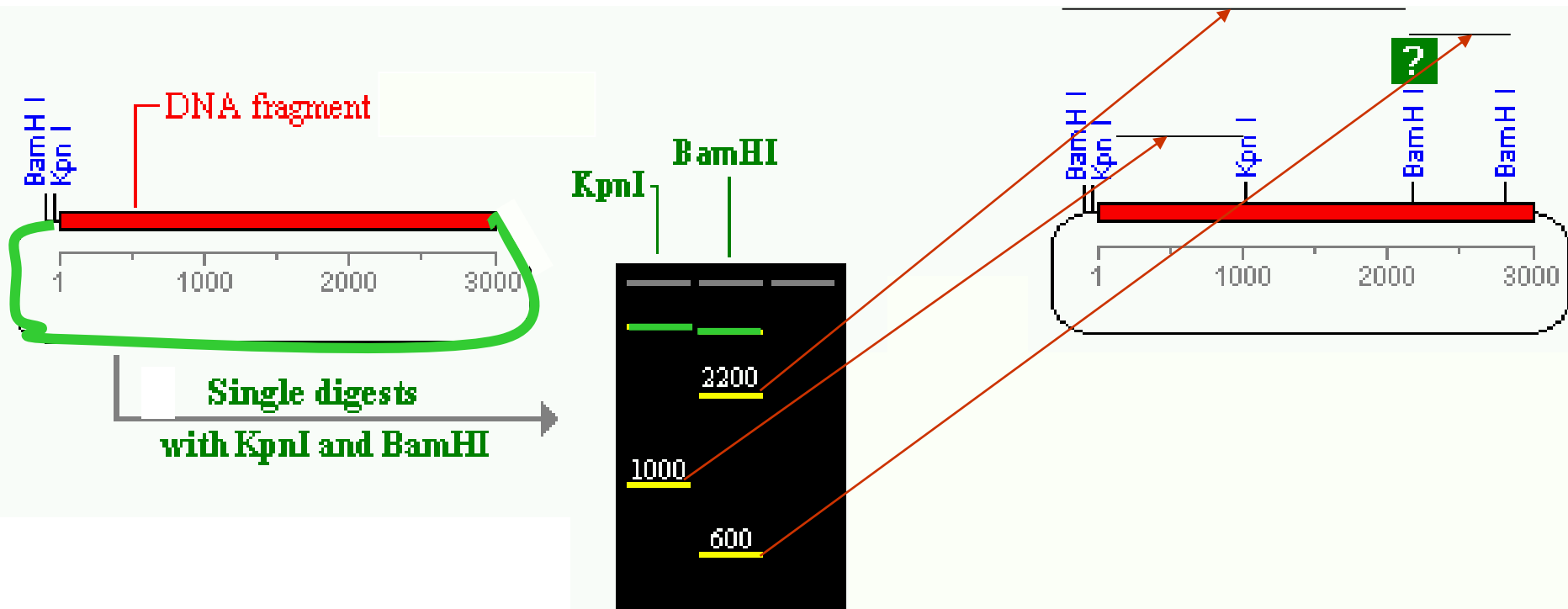
- 1) Plasmid size
- 2) Number and size of restriction fragments
- 3) Is the sum of the fragments equal to the length of the original plasmid?  
(What is their stoichiometry?)

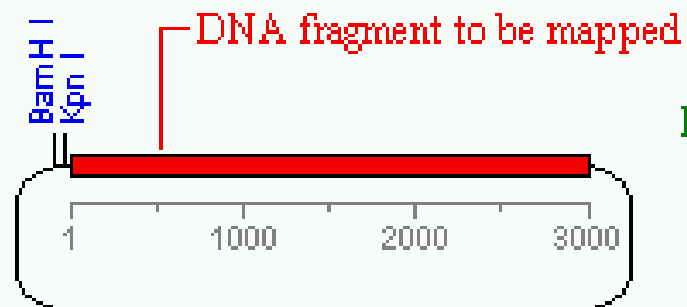
# Partial digestion





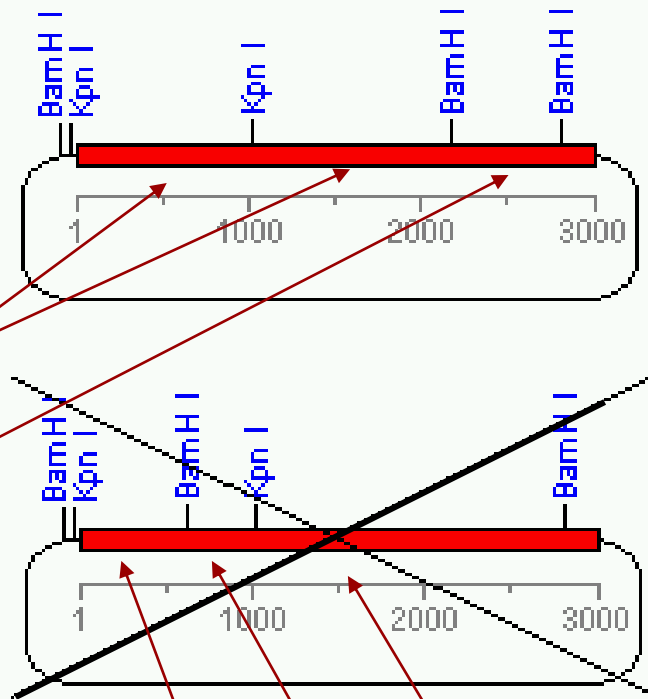
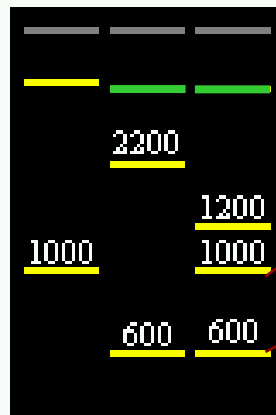
# Orientation of fragment in the vector





Single and double digests  
with KpnI and BamHI

KpnI BamHI KpnI + BamHI



300, 600, 1200



# LINEAR DNA SEPARATION

Agarose concentration

(% w/v)

0,3

0,6

0,7

0,9

1,2

1,5

2,0

Molecular weight

(kb)

5 - 25

1 - 20

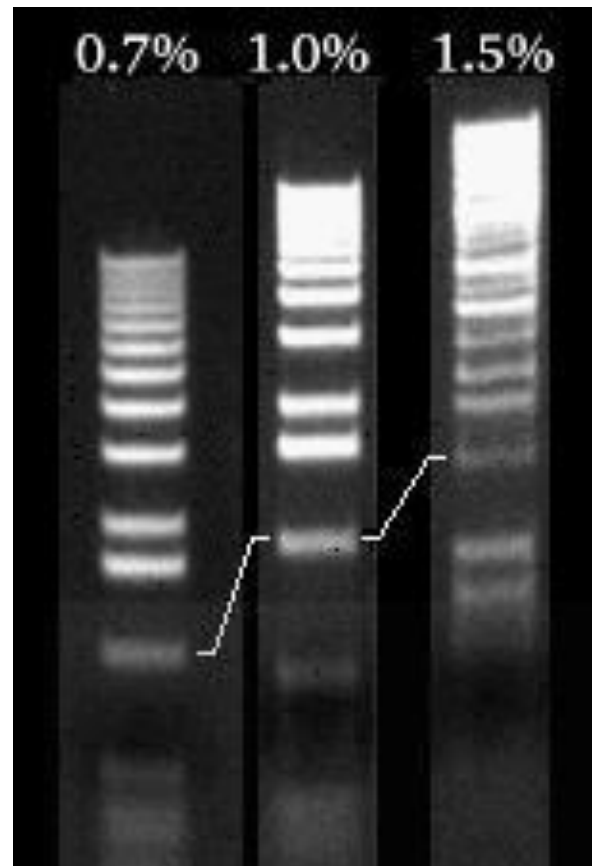
0,8 - 10

0,5 - 7

0,4 - 6

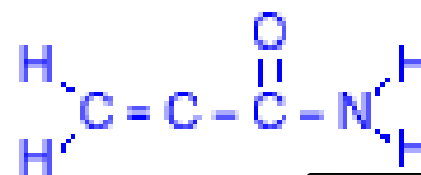
0,2 - 3

0,1 - 2



Acrylamide	Fragment size	Migration of bromophenol blue
------------	---------------	----------------------------------

(%)	(bp)	(bp)
3,5	100 - 1000	100
5,0	100 - 500	65
8,0	60 - 400	45
12	50 - 200	20
20	5 - 100	



Acrylamide

100

1,000

10,000

100,000

1,000,000

Yeast Chromosome PFG Marker

Lambda Ladder PFG Marker

MidRange PFG Markers I & II

Low Range PFG Marker

Lambda DNA-Mono Cuts

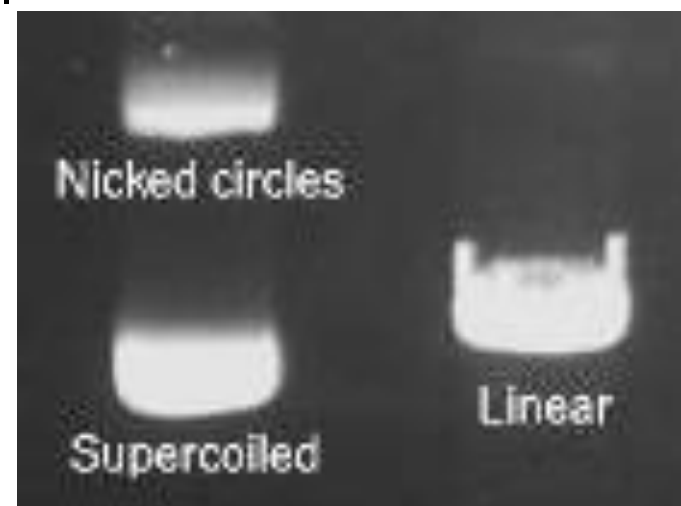
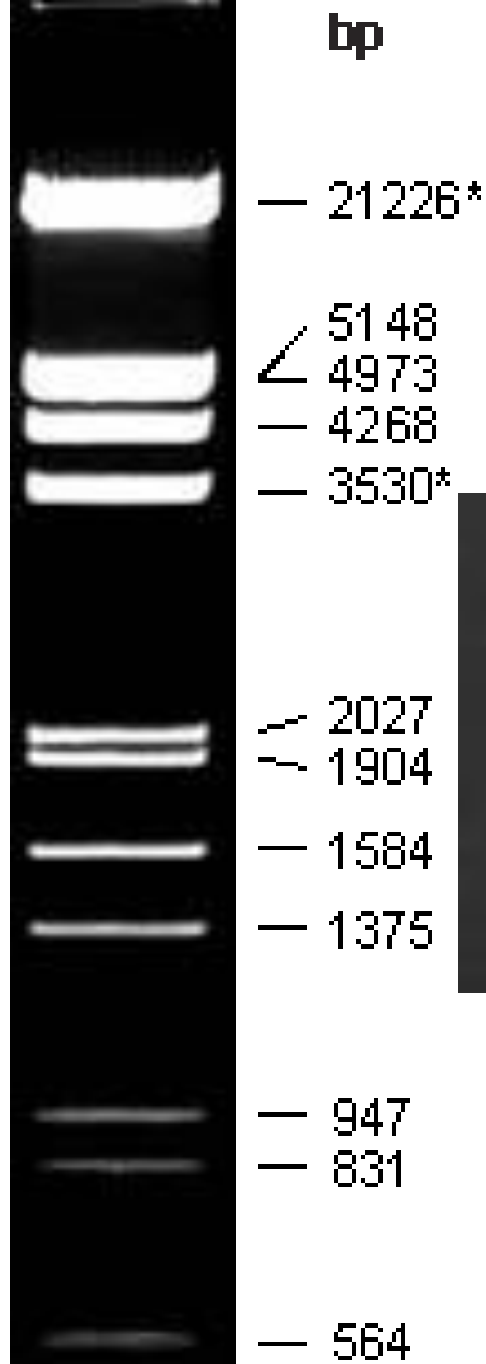
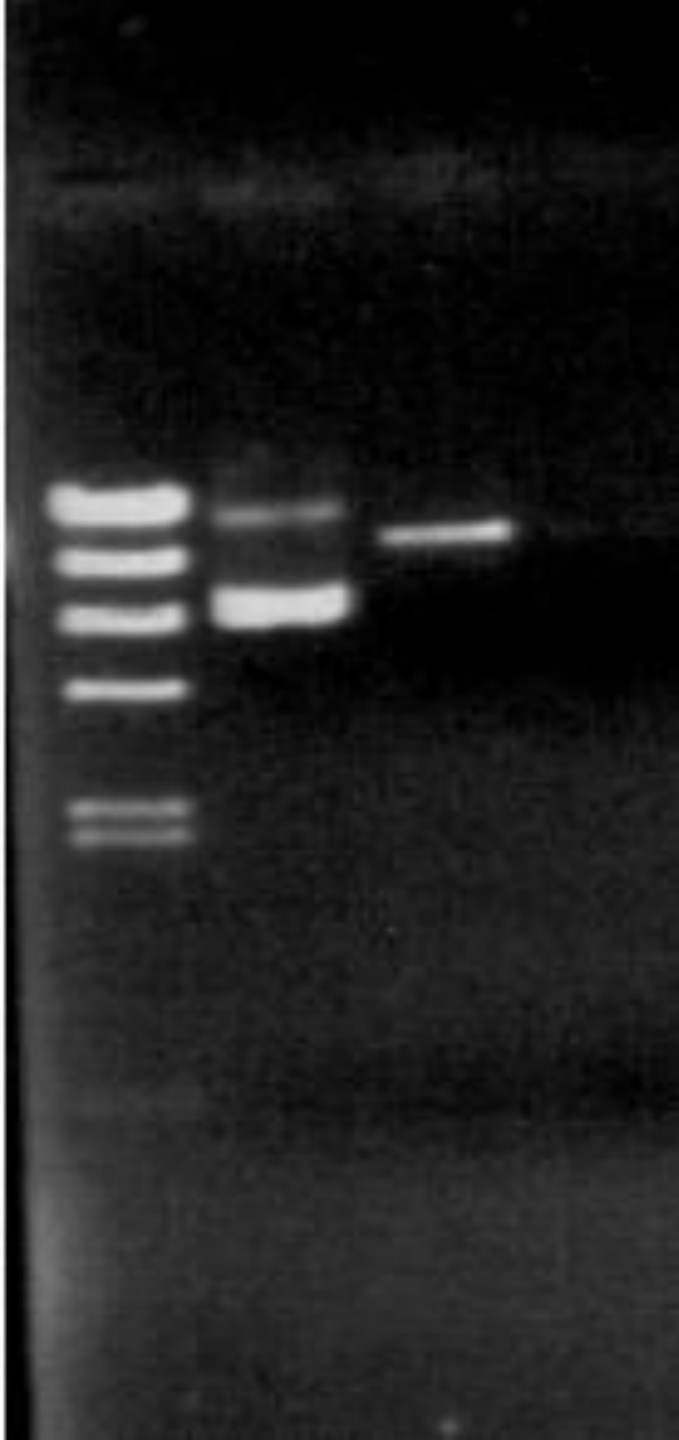
Lambda DNA-*Hind* III digest

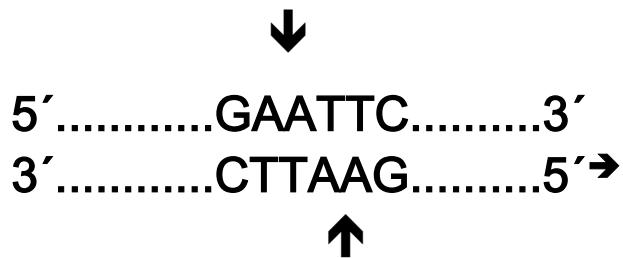
Lambda DNA-*Bst*E II digest

pBR322 DNA-*Bst*N I digest

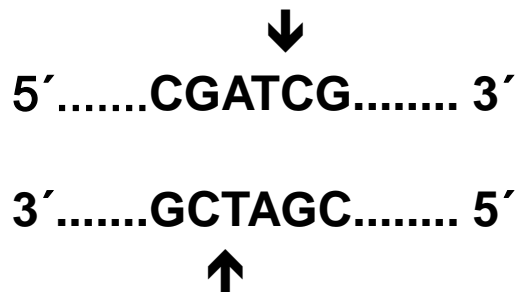
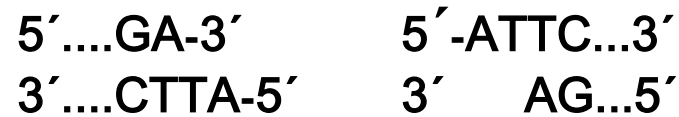
$\phi$ X DNA-*Hae* III digest

pBR322 DNA-*Msp* I digest

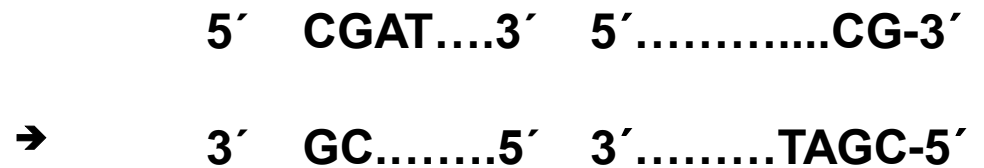




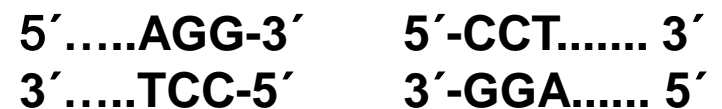
5' overhang



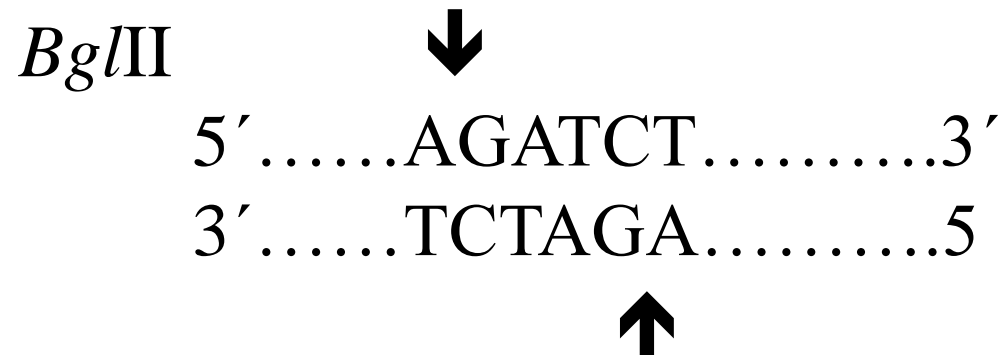
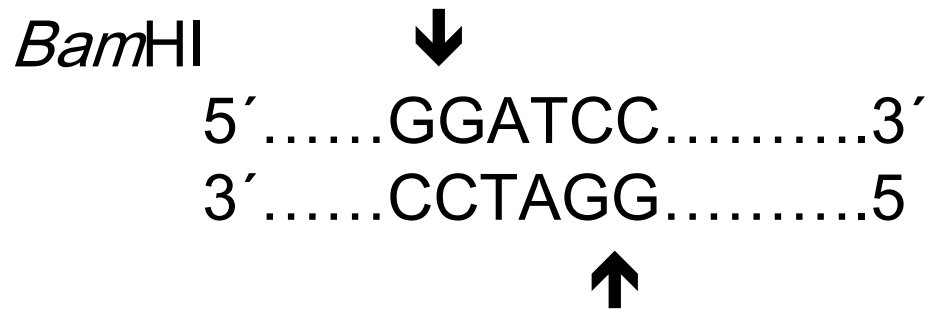
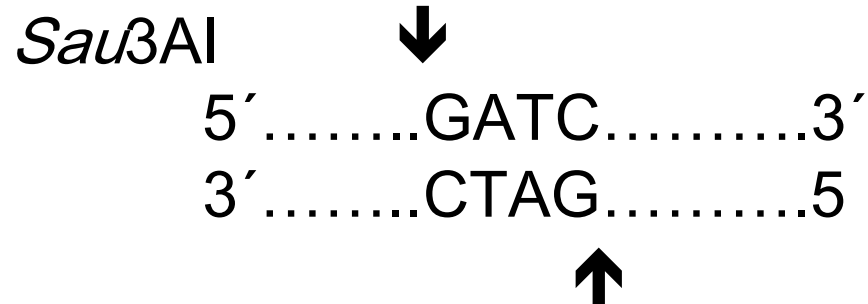
3' overhang



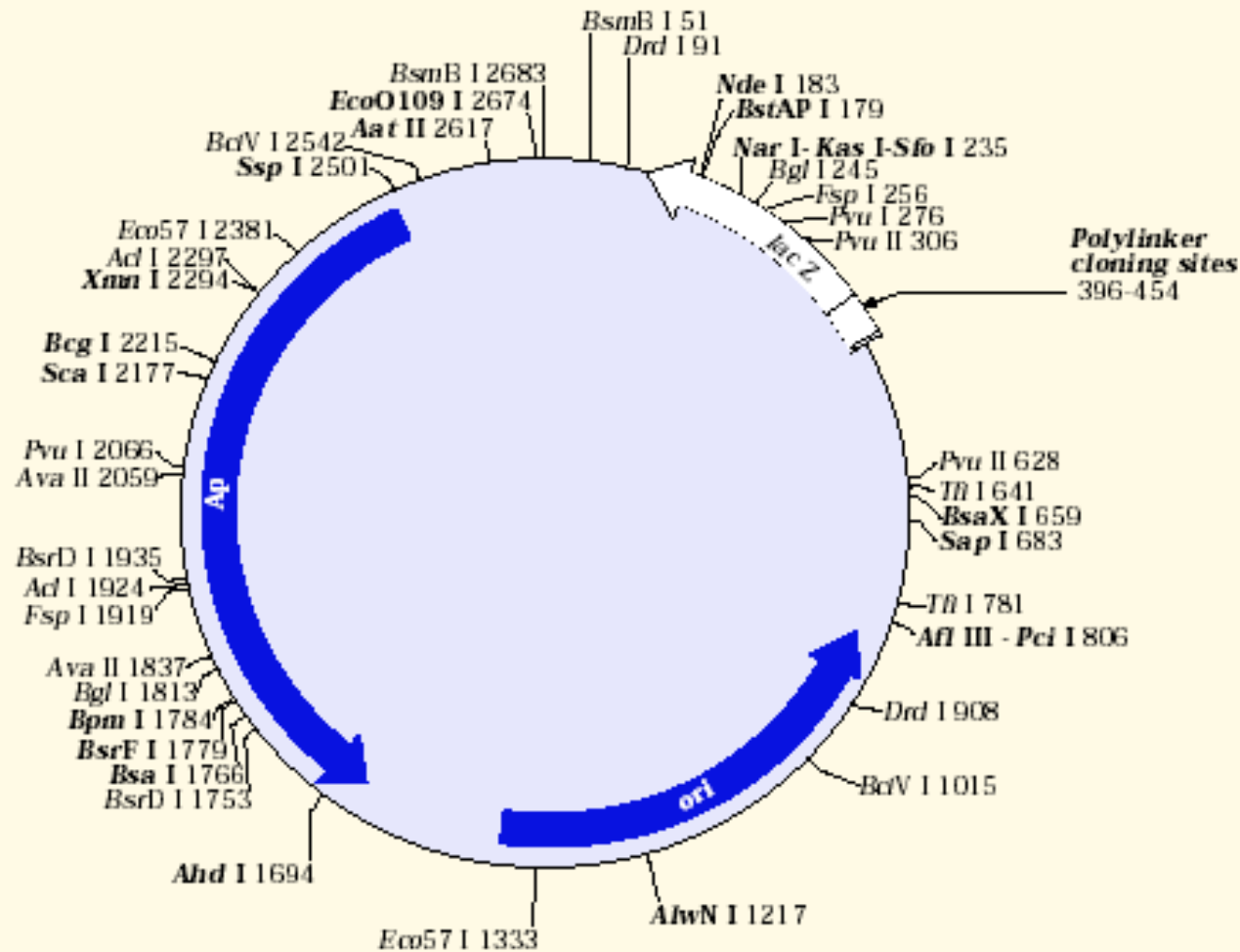
blunt end



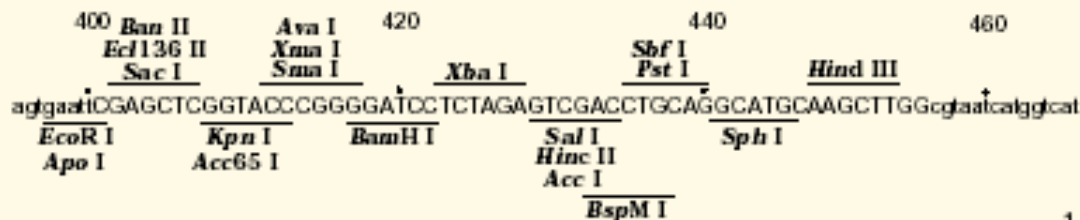
# The same cohesive ends generated by different RE



# pUC 19

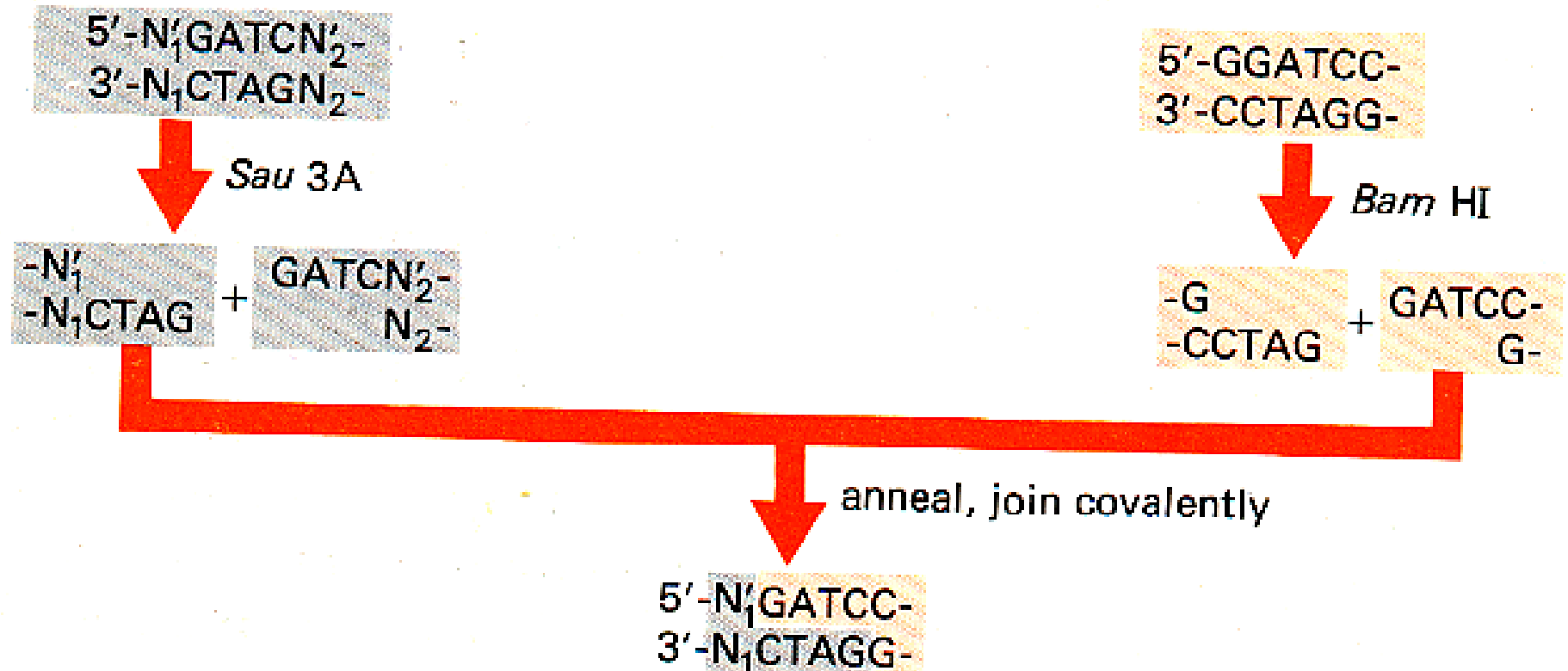
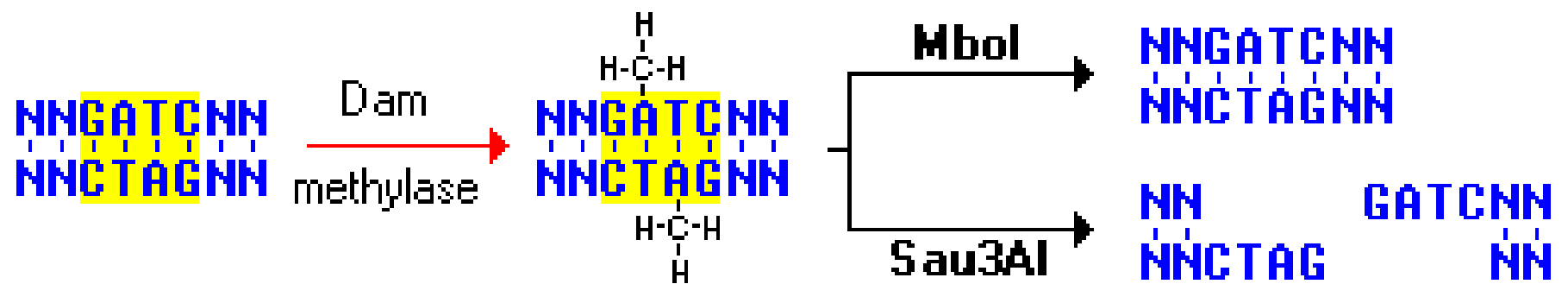


## pUC19 Polylinker

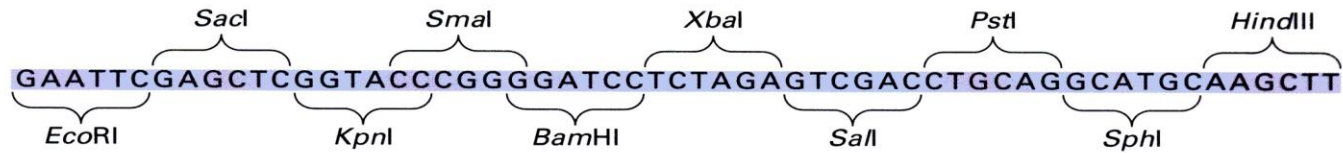


LacZ' Ala Leu Ser Asn Ser Ser Pro Val Arg Pro Asp Glu Leu Thr Ser Arg Cys Ala His Leu Ser Pro Thr Ile Met Thr

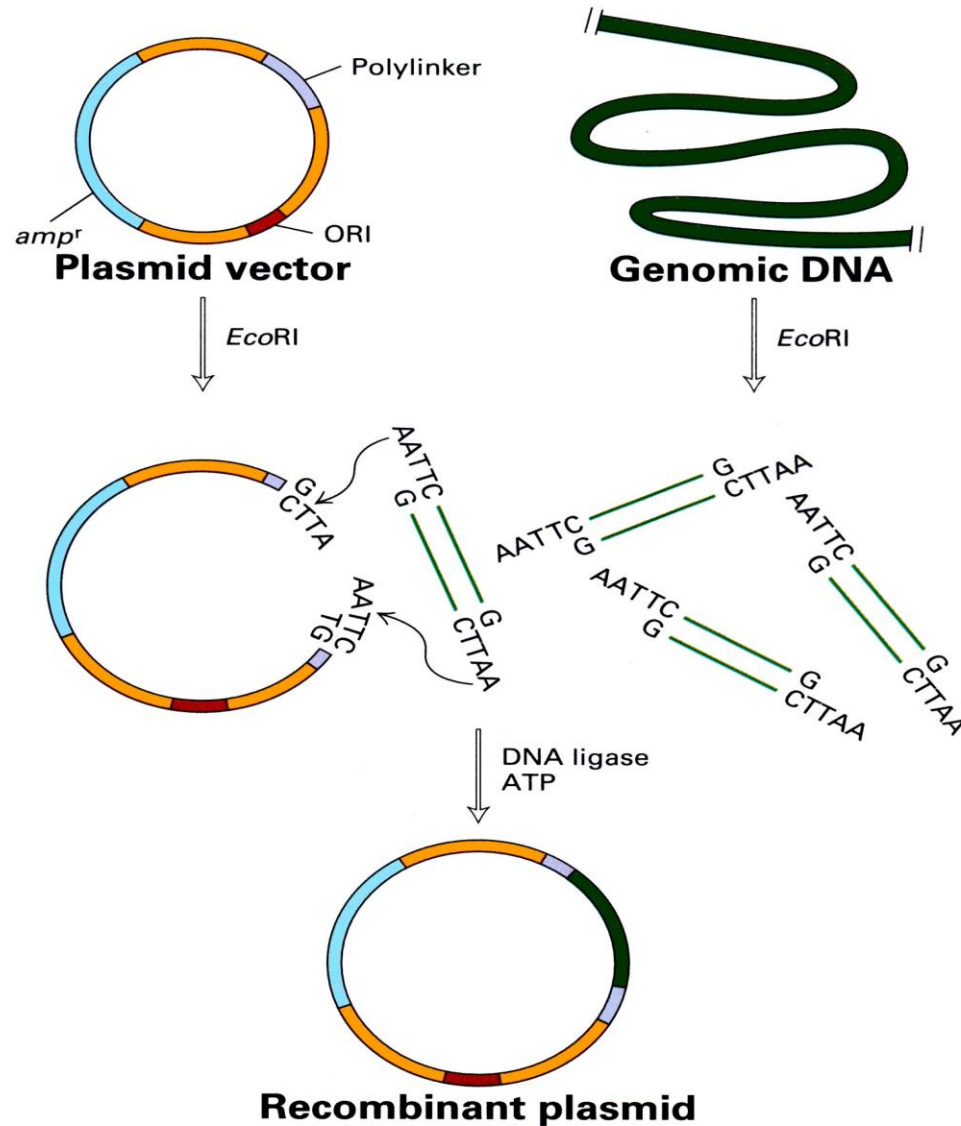


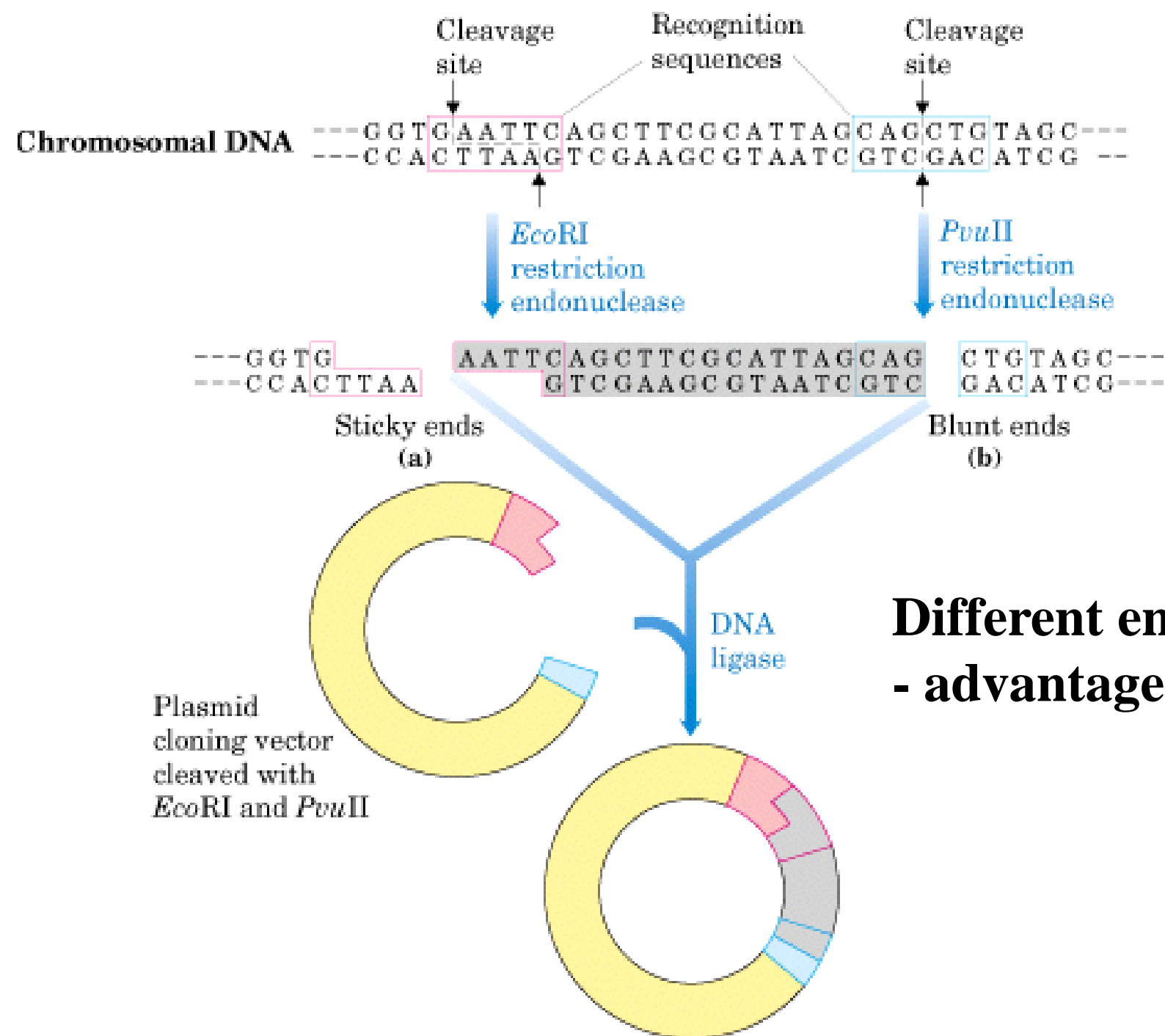


(a) Sequence of polylinker



(b) Insertion of *EcoRI* restriction fragments

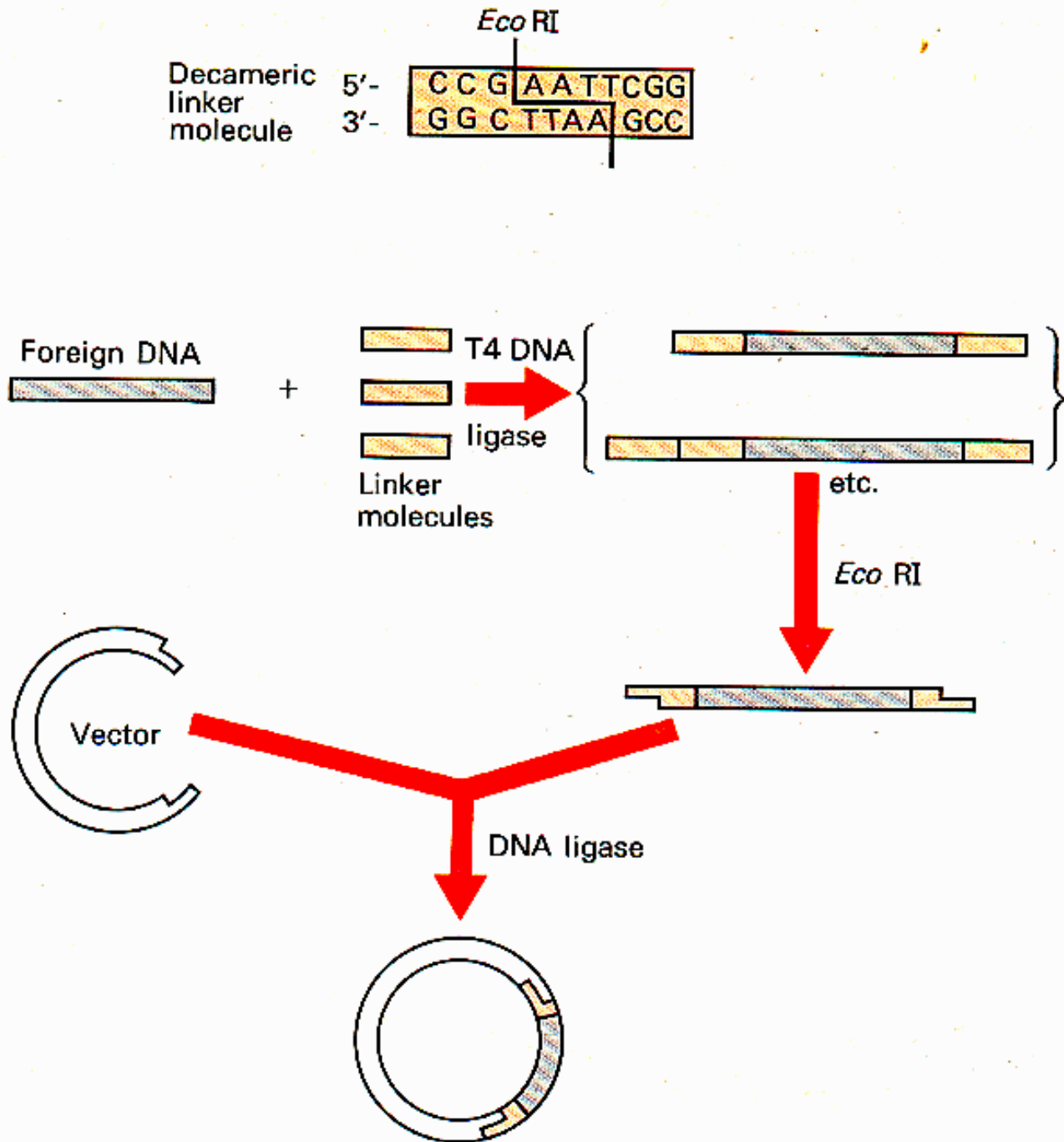




**End modification**

**Adapters, linkers, controlled degradation, synthesis  
of overhang**

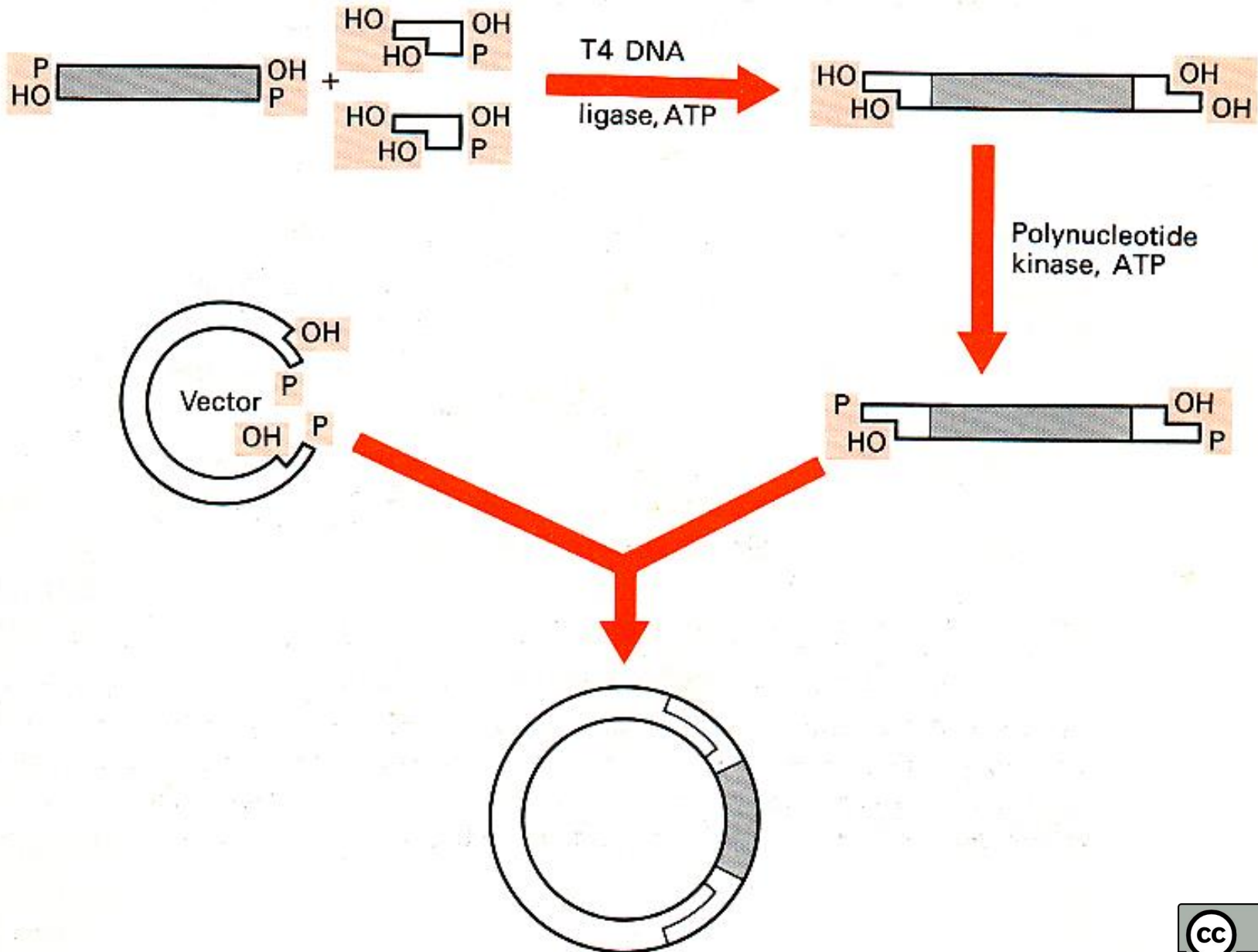
# Linkers



Adaptor molecule

5'-HO-GATCCCCGGG-OH  
3'-HO-GGGCCC-P

# Adapters

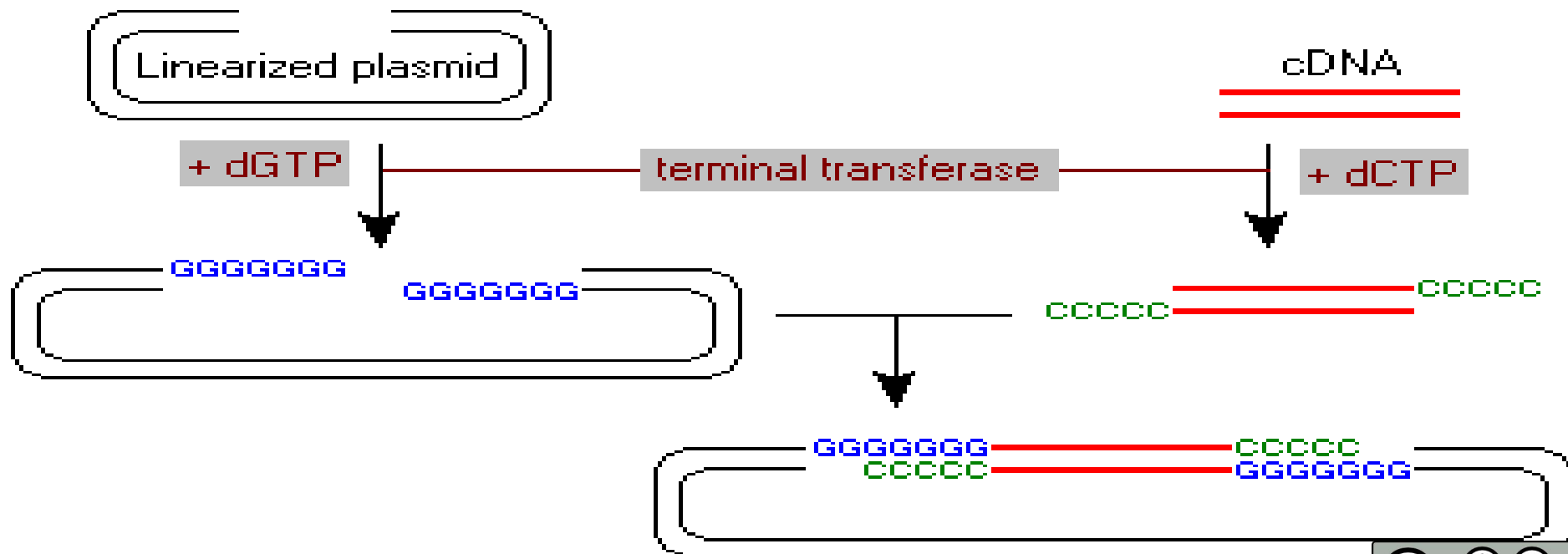


# Overhang synthesis (only 3')

5' G-A-T-C-A-C-T-G-C-A  
A-C-G-T-C-T-A-G-T-G 5'

Terminal transferase  
+ dTTP

T-T-T-T-T-T-T-T-A-C-G-T-C-T-A-G-T-G  
G-A-T-C-A-C-T-G-C-A-T-T-T-T-T-T-T-T





# Correction of the protruding ends to blunt ends using:

T4 DNA polymerase

ss exonucleases

5'  $\rightarrow$  3' and 3'  $\rightarrow$  5'; removal of the 5' and 3' overhang ends and hairpins

- Bal 31 nuclease (*Alteromonas espejiana*)

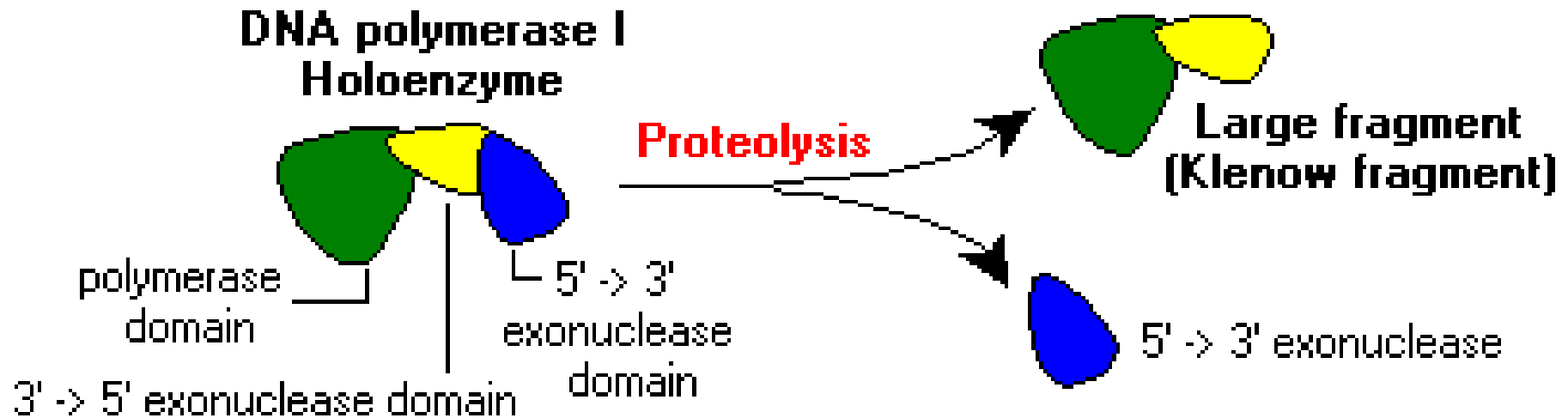
- S1 nuclease (*Aspergillus oryzae*)

Mung bean nuclease (Mung beans)

Exonuclease VII (*E. coli*)

Klenow fragment of DNA polymerase I

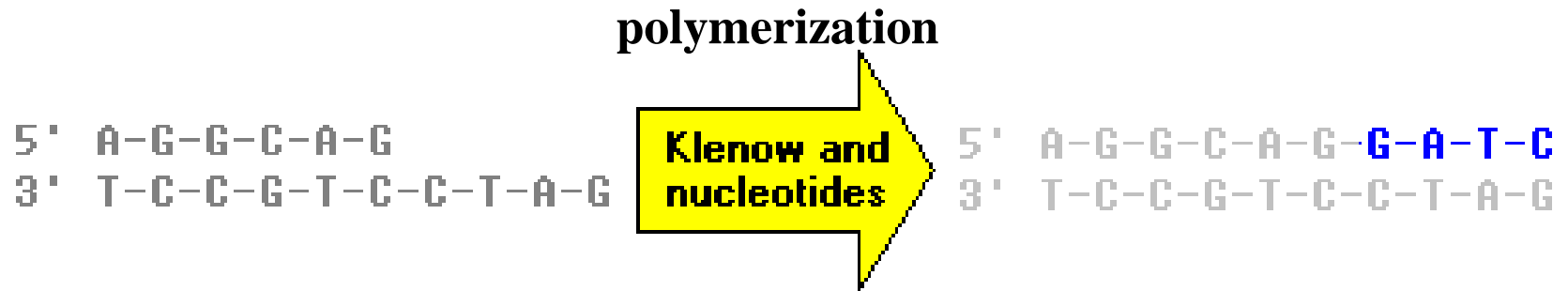
# Synthesis and degradation of ends - Klenow fragment



## Klenow fragment of DNA polymerase I *E. coli*

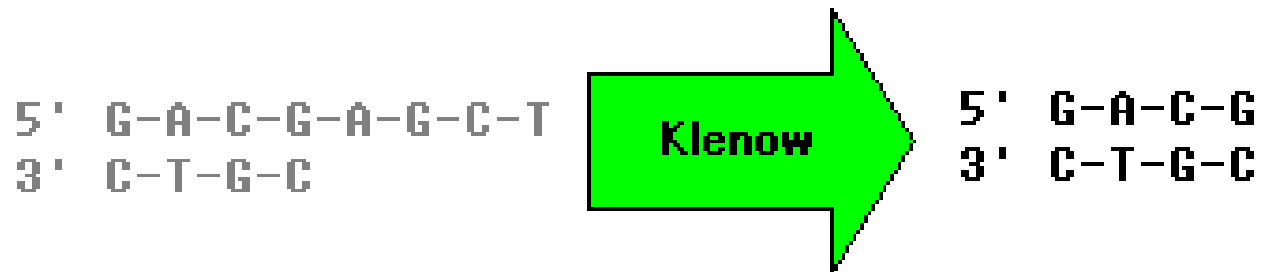
- C-terminal
- subtilisin
- no 5' → 3' exonuclease
- Recombinant

# Filling of recessive end



## Degradation of protruding end by Klenow fragment of DNA polymerase I

$3' \rightarrow 5'$  exonuclease activity

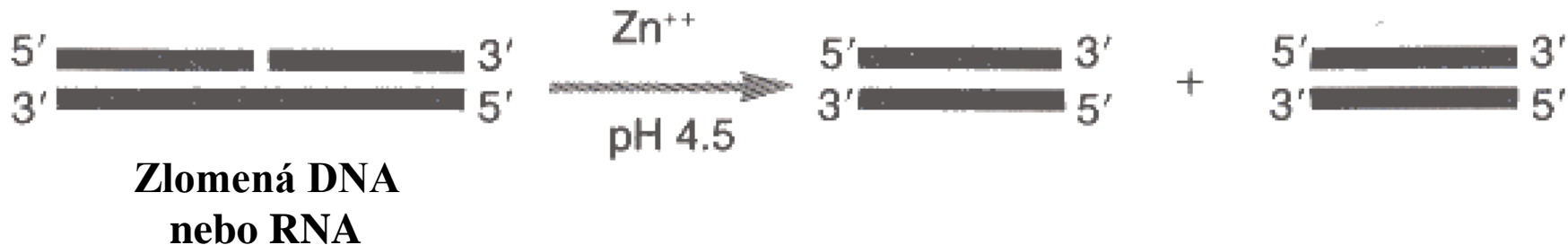


# S1 nuclease from *Aspergillus orizae*

I. Activity on single stranded DNA or RNA: five times more active on DNA than on RNA



II. Activity at a nick or gap in duplex DNA or RNA:

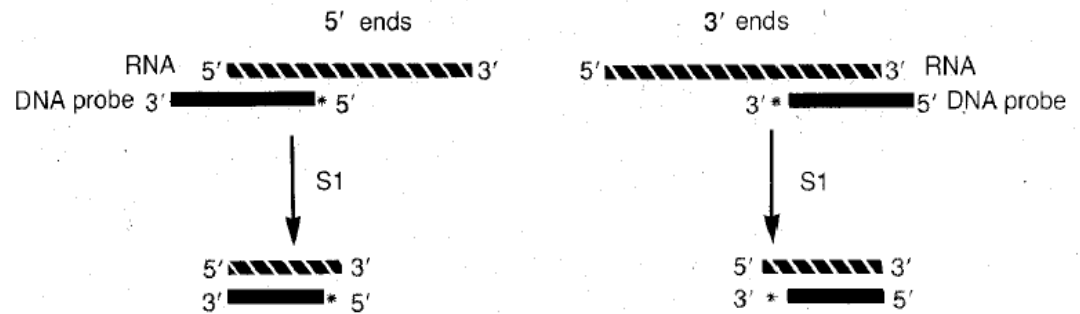


Stable in urea, SDS, formamide

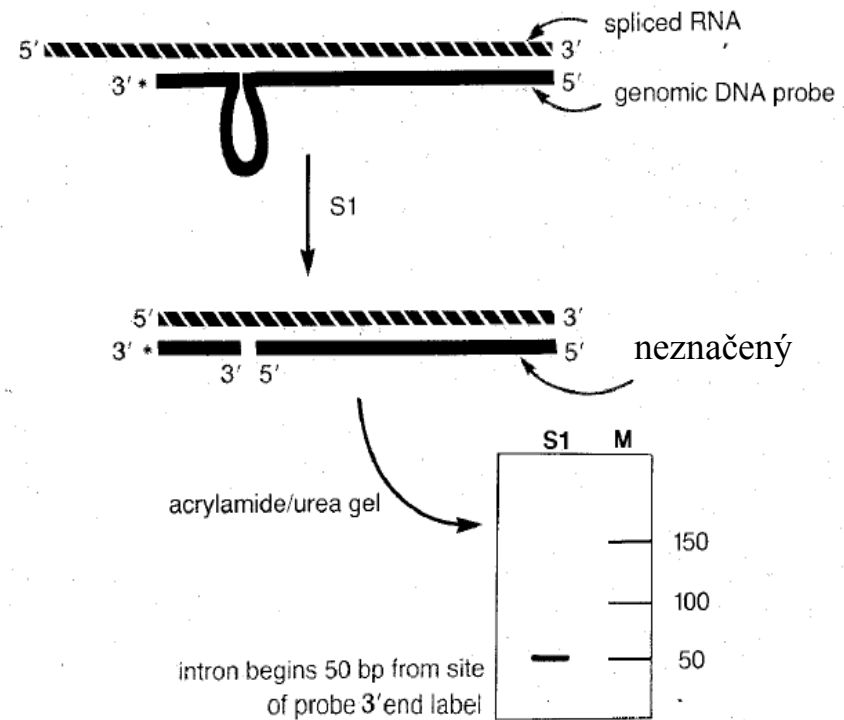
Mung bean is similar, but not so stable

# Applications of S1 nuclease

## Produce blunt ends



## Mapping the locations of introns



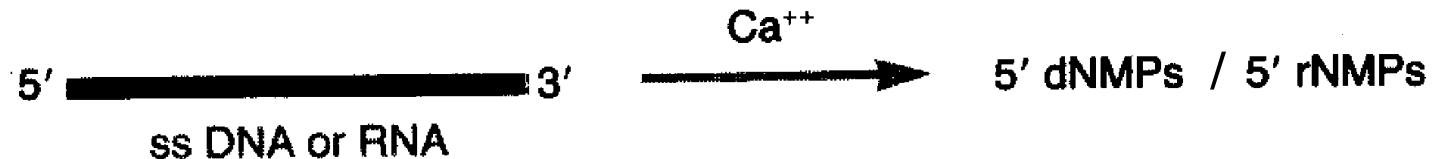
## **ad S1 nuclease application**

- A primary substrate is ssDNA,**

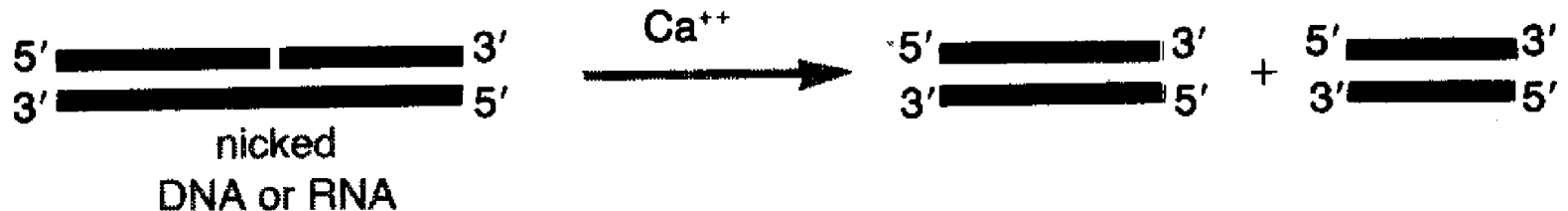
**But generates nicks in dsDNA, dsRNA  
or DNA-RNA hybrid**

# Bal 31 nuclease from *Alteromonas espejiana*

## I. Activity on single stranded DNA or RNA:



## II. Activity at a nick or gap in duplex DNA or RNA:



## III. Activity at the ends of duplex DNA fragments:





# ds nucleases

## **DNase I** – human endonuclease

ss and ds DNA, chromatin and RNA:DNA hybrids

(mutations associated with systemic lupus erythematosus; as a therapy of cystic fibrosis)

- **ds exonuclease  $3' \rightarrow 5'$**

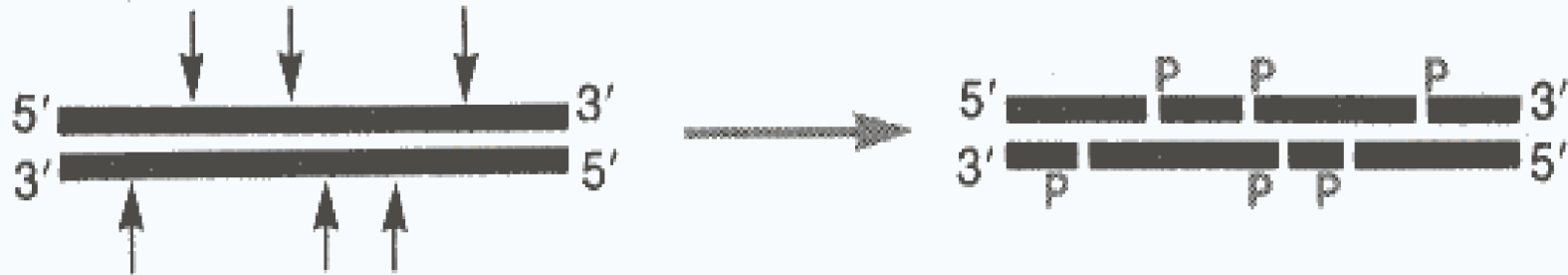
exonuclease III (*E. coli* K-12, from 3' end)

- **ds exonuclease  $5' \rightarrow 3'$**

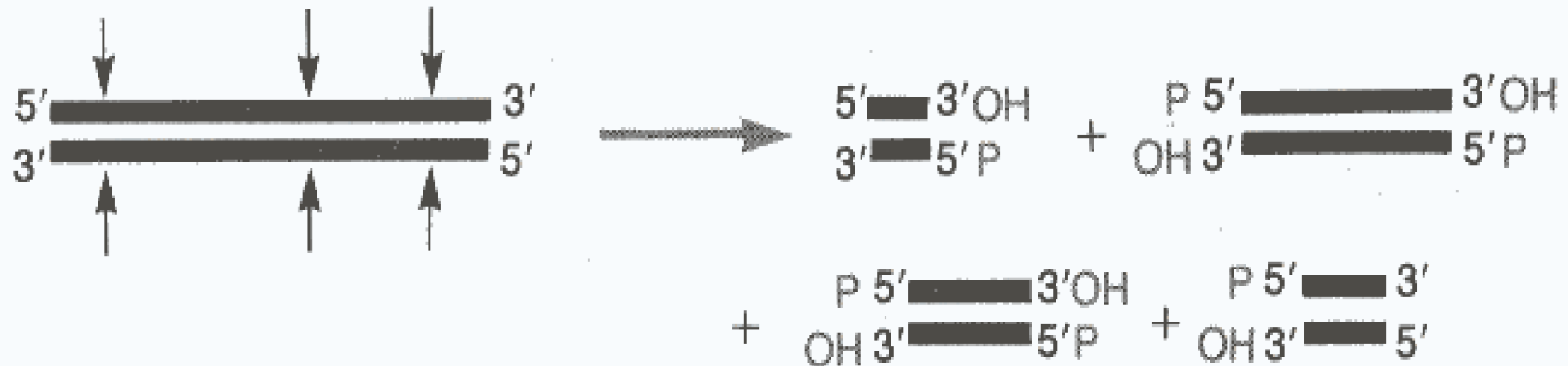
lambda exonuclease ( $5' \Rightarrow 3'$  exodeoxyribonuclease, no activity in nicks)

# DNase I

## I. Activity in the presence of $Mg^{++}$ :



## II. Activity in the presence of $Mn^{++}$ :



# Controlled degradation of 3' ends

**Original fragment**

GAGG *AGCGCT* AATAC  
CTCC *TCGCGA* TTATG

**Cleaved *AfeI***

GAGG *AGC*      *GCT* AATAC  
CTCC *TCG*      *CGA* TTATG

**Partially degraded**

**presence of d**A**TP and**

**Klenow fragment**

GAGG **A**      *GCT* AATAC  
CTCC *TCG*      **A** TTATG

# Partial filling of the ends

**Original fragment**

GAGGCTCGAGAATAC  
CTCCGAGCTCTTATG

**cleaved *Xho*I**

GAGGC      TCGAGAATAC  
CTCCGAGCT      CTTATG

**Partially filled with dCTP  
and dTTP and Klenow  
fragment**

GAGGCTC      TCGAGAATAC  
CTCCGAGCT      CTCCTTATG

# ds exonuclease

**exonuclease III** (*E. coli* K-12, from 3' end)

- ds exonuclease  $3' \rightarrow 5'$

# lambda exonuclease

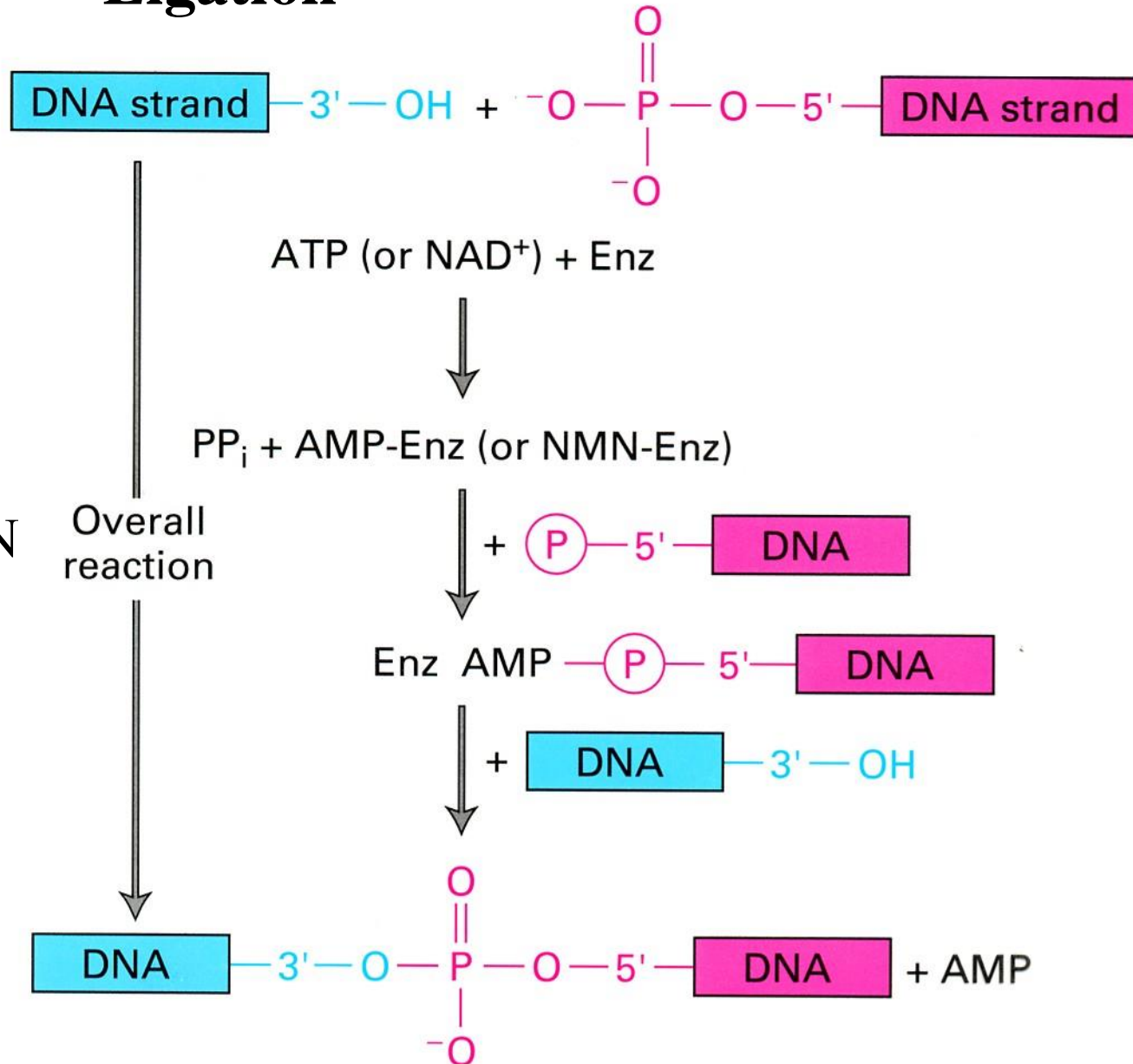
- ds exonuclease  $5' \rightarrow 3'$ , no activity in nicks

# Ligation

## Cofactors

Ligase *E. coli*  
 $\text{ATP} \rightarrow \text{AMP} + \text{PP}_i$

Ligase T7 phage  
 $\text{NAD}^+ \rightarrow \text{AMP} + \text{NMN}$



## DNA ligase T4

**DNA ligase *E. coli*** – less effective

not for blunt ends

higher specificity for complementary ends joining

~ 16°C

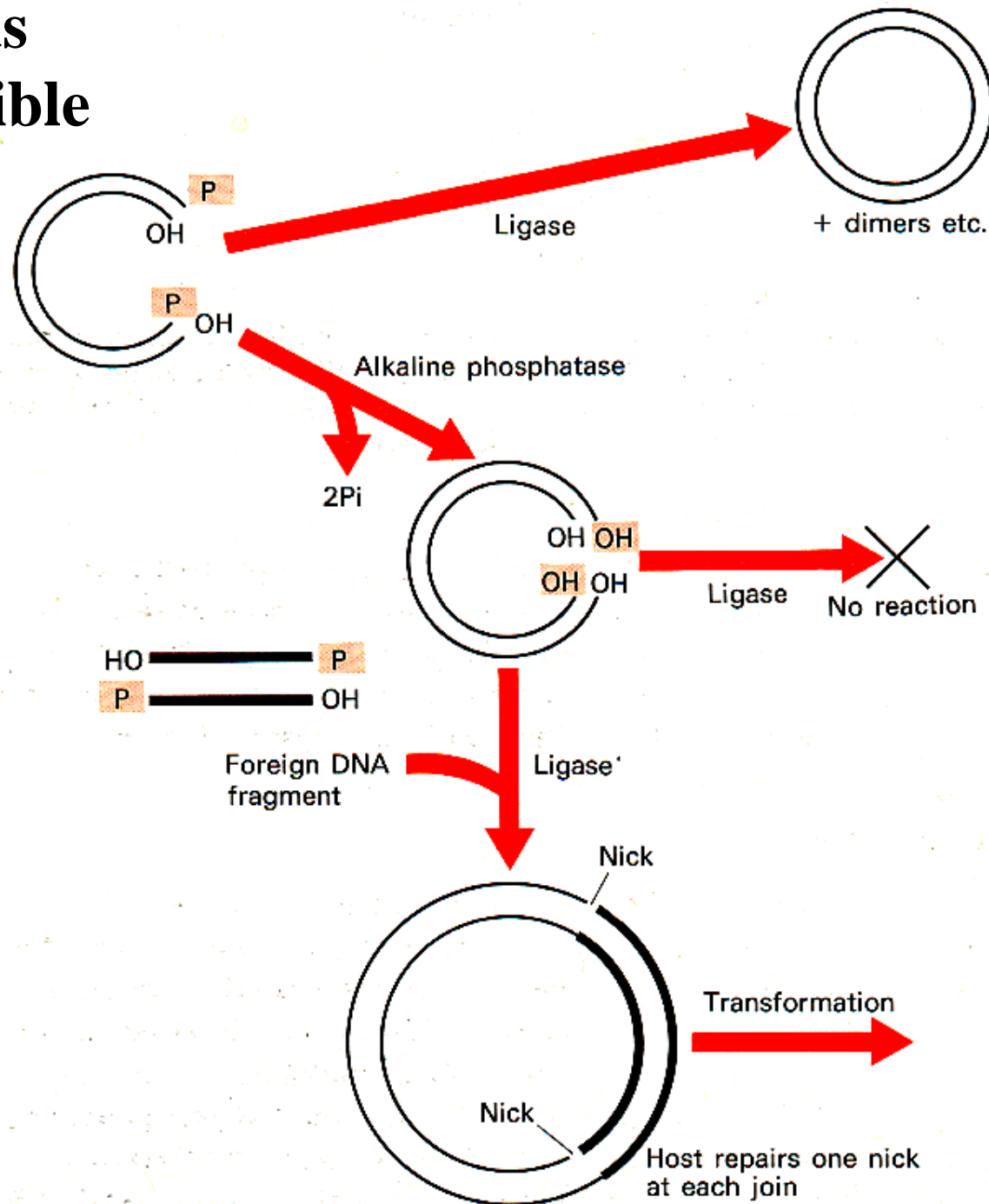
molar excess of insert

high concentration (or PEG 8 000)



# Blunt ends + compatible

- ligation

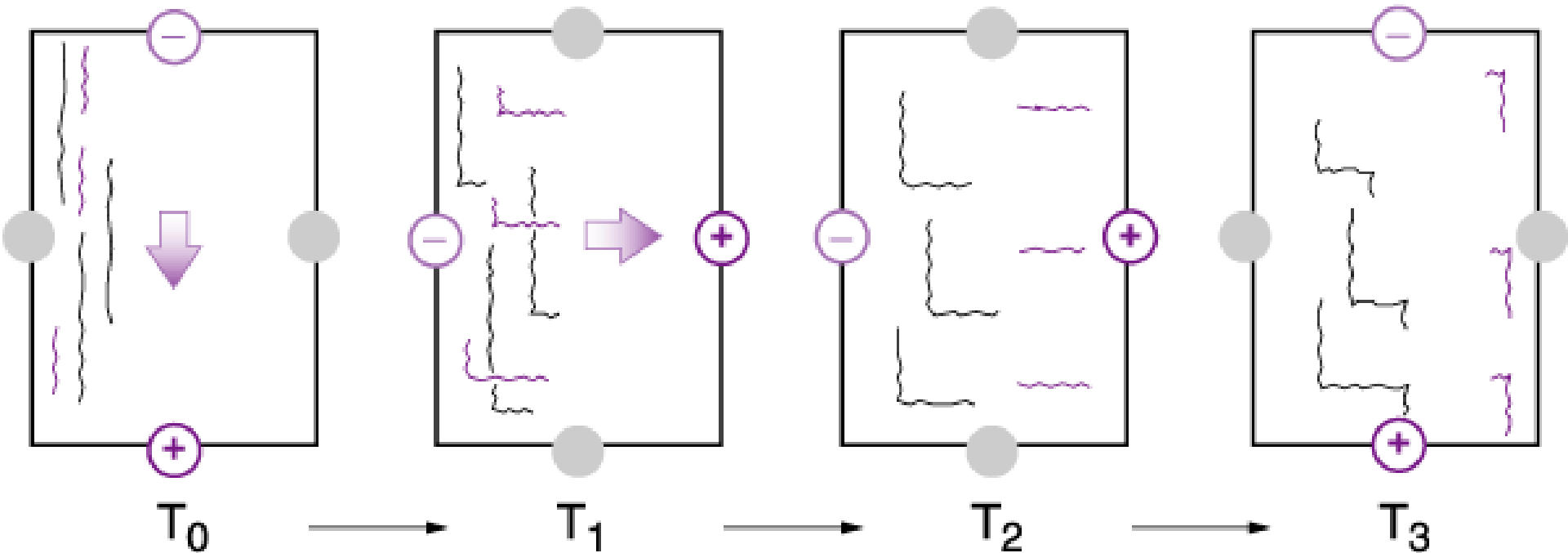


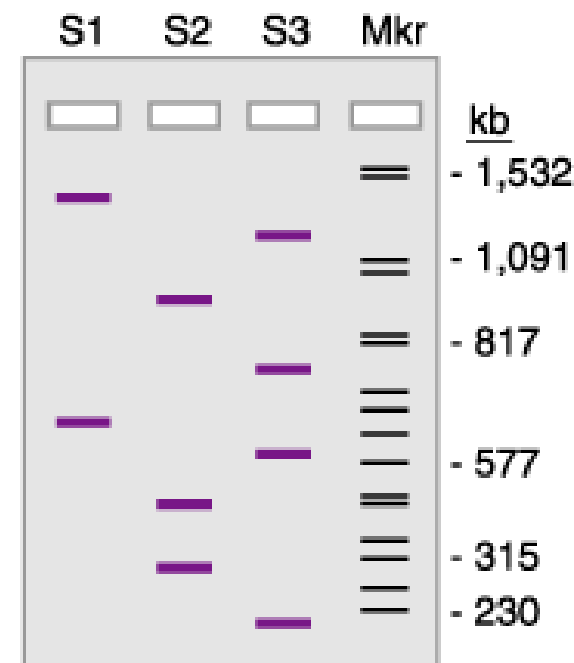
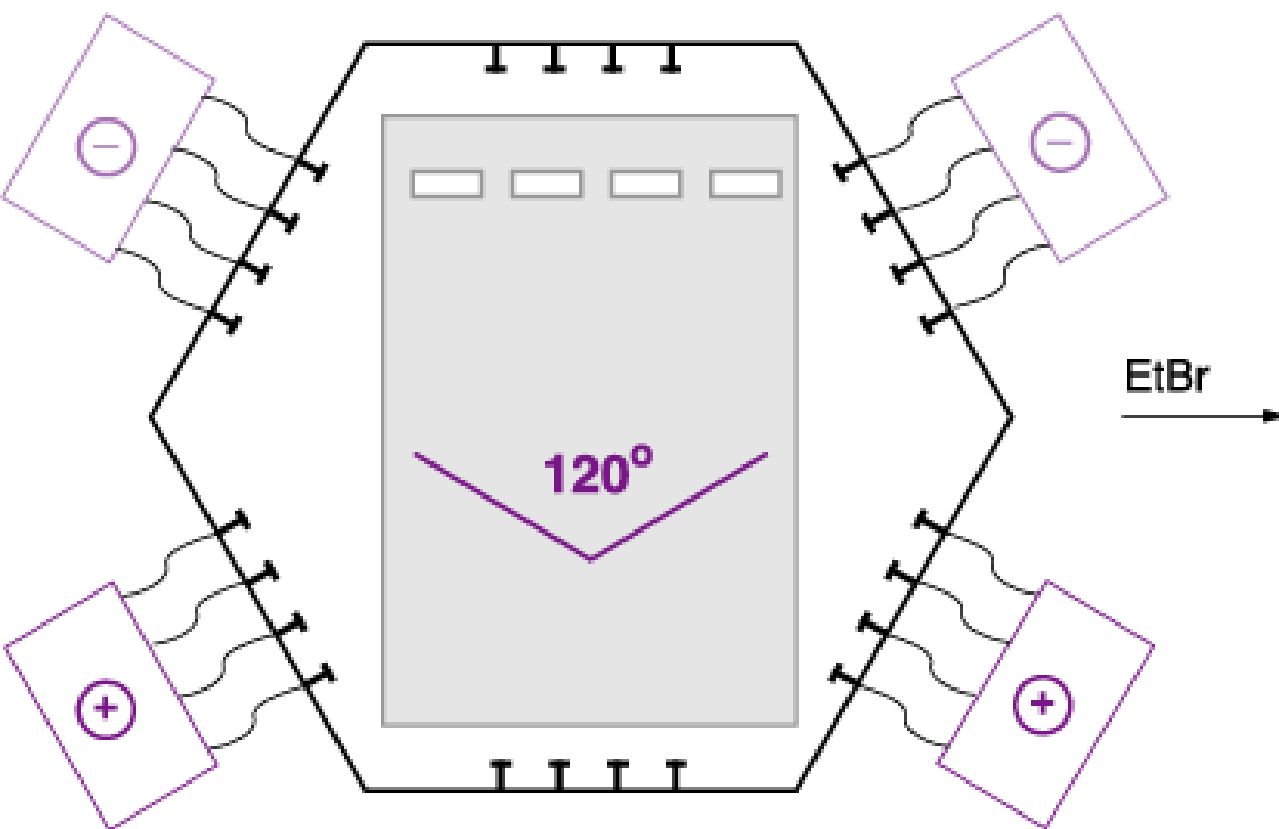
# Other

- **Single-stranded nucleases**
  - Bal 31 nuclease
  - S1 nuclease, Mung bean nuclease
  - Exonuclease VII
- **nonspecific double stranded**
  - DNase I
- **Double-stranded 3'→5'**
  - Exonuclease III
- **Double stranded 5'→3'**
  - lambda exonuclease

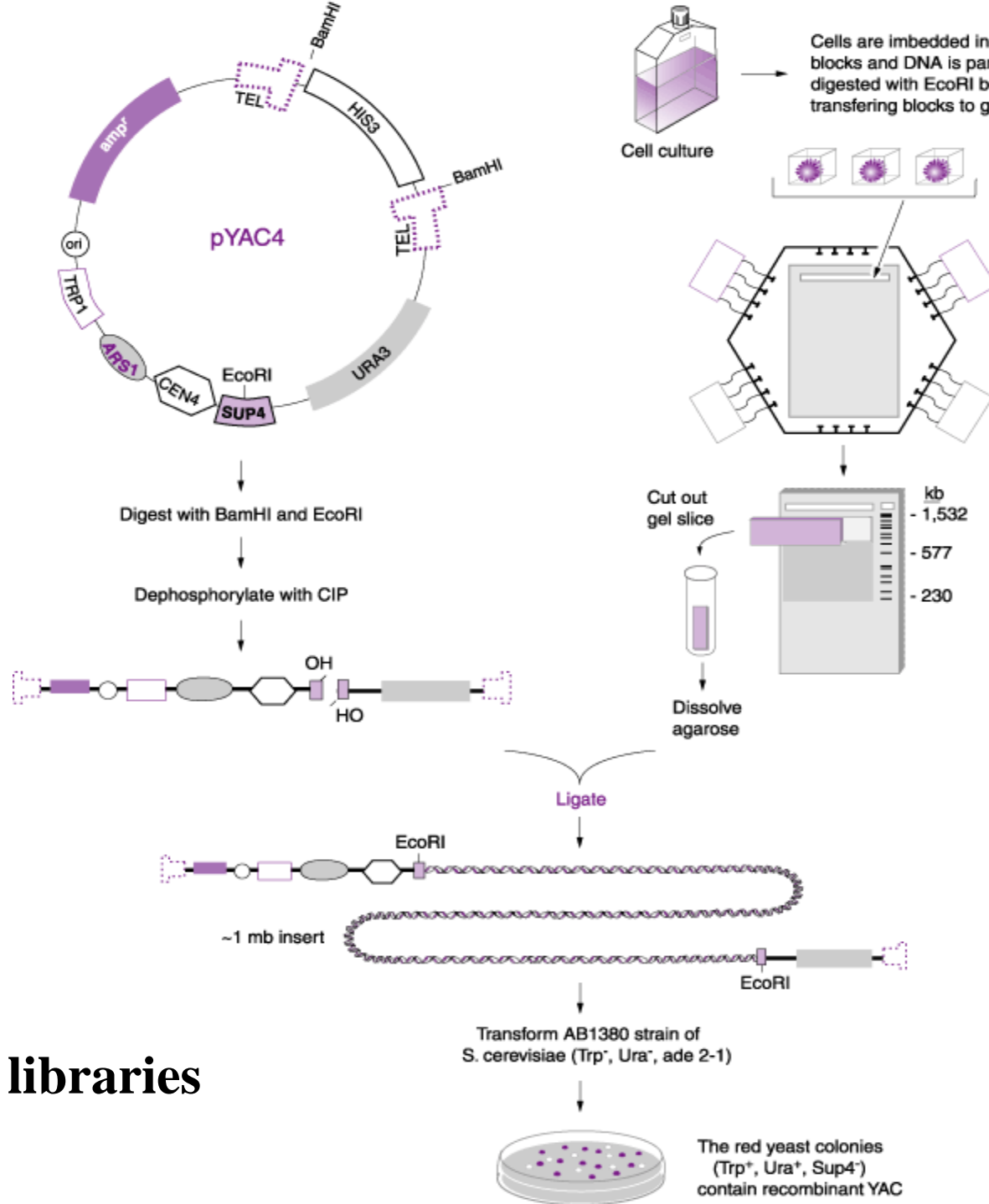
## Pulse-field gel electrophoresis (PFGE)

Large DNA molecules cross the agarose in a „snake-like“ manner. The shorter molecules re-orient faster upon switching the field polarisation



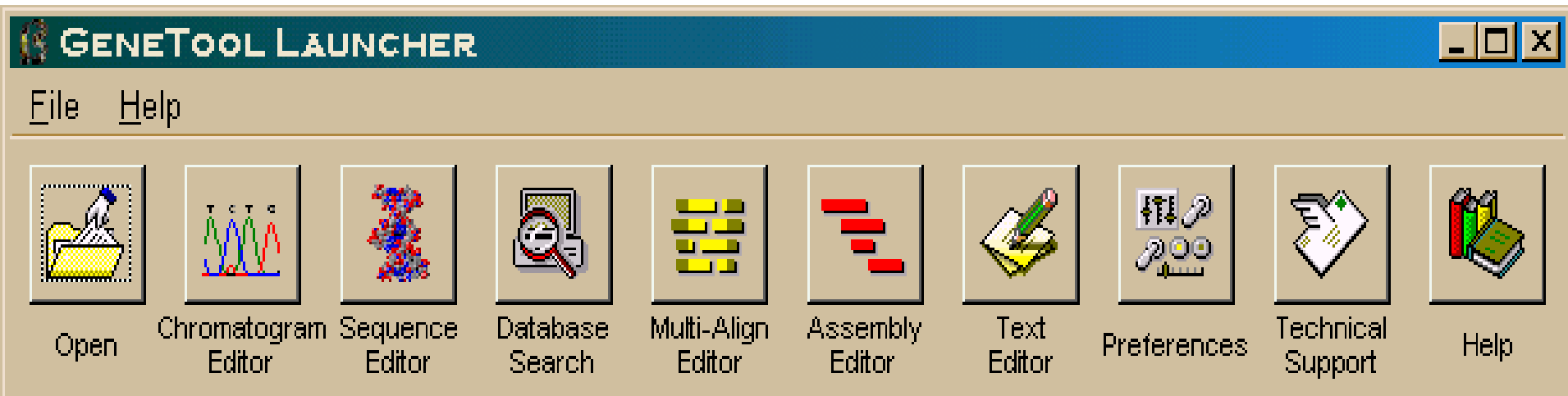


Electric field alternates  $120^\circ$  every 90 seconds for 18 to 24 hours at  $14^\circ\text{C}$



## YAC - libraries

# GENETOOLS-TOOLBAR



# Database Search Results

 DATA BROWSER: UNNAMED-1

File Edit View Transfer Help

Entrez search on "DNA"

1. (AW583558) ia01b10.y1 Human Pancreatic Islets Homo sapiens cDNA 5' similar to gb:X70508 INSUL
2. (AW583365) ia01e05.x1 Human Pancreatic Islets Homo sapiens cDNA 3' similar to gb:X70508 INSUL
3. (AW583585) ia01e05.y1 Human Pancreatic Islets Homo sapiens cDNA 5' similar to gb:X70508 INSUL
4. (AW583366) ia01e06.x1 Human Pancreatic Islets Homo sapiens cDNA 3' similar to gb:X70508 INSUL
5. (AW583377) ia01f11.x1 Human Pancreatic Islets Homo sapiens cDNA 3' similar to gb:X70508 INSUL
6. (AW583598) ia01f11.y1 Human Pancreatic Islets Homo sapiens cDNA 5' similar to gb:X70508 INSUL

## Summary

Accession: AW583558      Sequence Length: 447



Title: ia01b10.y1 Human Pancreatic Islets Homo sapiens cDNA 5' similar to gb:X70508 INSUL...

Keywords:

1	GCTGCATCAG AAGAGGCCAT CAAGCAGATC ACTGTCCTTC TGCCATGGCC CTGTGGATGC
61	GCCTCCTGCC CCTGCTGGCG CTGCTGGCCC TCTGGGGACC TGACCCAGCC GCAGCCTTTG
121	TGAACCAACA CCTGTGCGGC TCACACCTGG TGGAAGCTCT CTACCTAGTG TGCGGGGAAC
181	GAGGCTTCTT CTACACACCC AAGACCCGCC GGGAGGCAGA GGACCTGCAG GTGGGGCAGG
241	TGGAGCTGGG CGGGGGCCCT GGTGCATGCA GCCTGCAGCC CTTGGCCCTG GAGGGGTCCC
301	TGCAGAAAGCG TGGCATTGTG GAACAATGCT GTACCAGCAT CTGCTCCCTC TACCAGCTGG
361	TCCTGCTGTC CTCTGCTGTC CTCTGCTGTC CTCTGCTGTC CTCTGCTGTC CTCTGCTGTC



BY SA

# Text editor

TEXT EDITOR: DENGUE4.TXT

File Edit Transfer Help

```
1 agttgttagt ctgtgtggac cgacaaggac agttccaaat cggaagcttg cttaacacag
  61 ttctaacagt ttgtttgaat agagagcaga tctctggaaa aatgaaccaa cgaaaaaagg
 121 tggtagacc acctttcaat atgctgaaac gcgagagaaa ccgcgtatca acccctcaag
 181 ggttggtgaa gagattctca accggacttt tttctgggaa aggaccctta cggatggtgc
 241 tagcattcat cacgtttttg cgagtccttt ccacccacc aacagcaggg attctgaaga
 301 gatggggaca gttgaagaaa aataaggcca tcaagatact gattggattc aggaaggaga
 361 taggcgcgat gctgaacatc ctgaacggga gaaaaagggtc aacgataaca ttgctgtgct
 421 tgattcccac cgtaatggcg ttttccttgt caacaagaga tggcgaaccc ctcatgatag
 481 tggcaaaaca tgaaaggggg agacctctct tgtttaagac aacagagggg atcaacaaat
 541 gcactctcat tgccatggac ttgggtgaaa tgtgtgagga cactgtcacg tataaatgcc
 601 ccctactggt caataccgaa cctgaagaca ttgattgctg gtgcaacctc acgtctacat
 661 gggtcatgta tgggacatgc acccagagcg gagaacggag acgagagaag cgctcagtag
 721 ctttaacacc acattcagga atgggattgg aaacaagagc tgagacatgg atgtcatcgg
 781 aaggggcttg gaagcatgct cagagagtag agagctggat actcagaaac ccaagattcg
 841 cgctcttggc aggatttatg gcttatatga ttgggcaaac aggaatccag cgaactgtct
 901 tctttgtcct aatgatgctg gtcgccccat cctacggaat gcgatgcgta ggagtaggaa
 961 acagagactt tgtggaagga gtctcaggtg gagcatgggt cgacctagtg ctagaacatg
1021 gaggatgcgt cacaaccatg gcccaaggaa aaccaacctt ggattttgaa ctgactaaga
1081 caacagccaa ggaagtggct ctgttaagaa cctattgcat tgaagcctca atatcaaaca
1141 taactacggc aacaagatgt ccaacgcaag gagagcctta tctgaaagag gaacaggacc
1201 aacagtacat ttgccggaga gatgtggtag acagaggggtg gggcaatggc tgtggcttgt
1261 ttggaaaagg aggagtgtg acatgtgcga agttttcatg ttcggggaag ataacaggca
1321 atttgytccg aattgagaac cttgaataca cagtgyttgt aacagtcac aatggagaca
1381 cccatgcagt aggaaatgac acatccaatc atggagttac agccatgata actcctaggt
```



BY SA



# Restriction analysis

<http://bioweb.pasteur.fr/seqanal/interfaces/restrict.html>

```
# Restrict of , from 1 to 150
# Minimum cuts per enzyme: 1
# Maximum cuts per enzyme: 2000000000
# Minimum length of recognition site: 4
# Blunt ends allowed    # Sticky ends allowed
# DNA is linear        # No ambiguities allowed
# Number of hits: 24
# Base
```

Number	Enzyme	Site	5'	3'	[5'	3']
8	NgoAIV	GCCGGC	8	12		
8	NaeI	GCCGGC	10	10		
9	MspI	CCGG	9	11		
15	Acil	CCGC	12	14		
18	MnII	CCTC	28	27		
36	HphI	GGTGA	24	23		
45	Tsp509I	AATT	44	48		
59	Tsp509I	AATT	58	62		
69	Acil	CCGC	66	68		
78	BsmI	GAATGC	84	82		
81	FspI	TGCGCA	83	83		
82	Hin6I	GCGC	82	84		
82	HhaI	GCGC	84	82		
114	ThaI	CGCG	115	115		
119	Acil	CCGC	119	121		
119	HgaI	GACGC	104	109		
123	Ecil	GGCGGA	108	106		
131	BbvI	GCAGC	143	147		
131	BpmI	CTGGAG	111	109		

# WebCutter

Webcutter – freely available on-line

analysis of restriction targets in nucleotide sequence

<http://www.firstmarket.com/cutter/cut2.html>

NgoAIV

Esp1396I

MroNI

AccB7I

aatggaagccggcggcacctcgctaacggattcaccactccaagaattggagccaatcaattcttgcggagaact base pairs

ttaccttcggccgccgtggagcgattgcctaagtggtagggttctaacctcggttagttaagaacgcctcttga 1276 to 1350

NgoMI

PfIMI

NaeI

Van91I

AviII EcoT14I

Mva1269I StyI Esp1396I

FspI

Eco130I

GsuI

gtgaatgcgcaaaccaacccttggcagaacatatccatcgcggtccgccatctccagcagccgcacgcggcgcatc base pairs

cacttacgcggttggttggaaccgtctgtataggttagcgcgaggcggttagaggctgctcggcgtgcgccgcgtag 1351 to 1425

BsmI

ErhI PfiMI

BpmI

BsaMI

BssT1I Van91I

Acc16I

AccB7I

# Layout Sequence

**SEQUENCE ANALYSIS: pBR322.BTI**

File Edit View Transfer Help

Layout Cloning vector pBR322, complete genome.

BamHI

351 TCATGGCGAC CACACCCGTC CTGTGGATCC TCTACGCCGG ACGCATCGTG  
AGTACCGCTG GTGTGGGCAG GACACCTAGG AGATGCGGCC TGCCTAGCAC

R H H R R H R C G C W R L Y R R

401 GCCGGCATCA CCGGCGCCAC AGGTGCGGTT GCTGGCGCCT ATATCGCCGA  
CGGCCGTAGT GGCCGCGGTG TCCACGCCAA CGACCGCGGA TATAGCGGCT

H H R W G R S G S P L R A H E R L

451 CATCACCGAT GGGGAAGATC GGGCTCGCCA CTTGCGGCTC ATGAGCGCTT  
GTAGTGGCTA CCCCTTCTAG CCCGAGCGGT GAAGCCCGAG TACTCGCGAA

F R R G Y G G R P R G R G T V G R

501 GTTTCGGCGT GGGTATGGTG GCAGGCCCCG TGGCCGGGGG ACTGTTGGGC  
CAAAGCCGCA CCCATACCAC CGTCCGGGGC ACCGGCCCCC TGACAACCCG

H L L A C T I P C G G G A Q R P

551 GCCATCTCCT TGCATGCACC ATTCTTGGC GCGGCGGTGC TCAACGGCCT  
CGGTAGAGGA ACGTACGTGG TAAGGAACGC CGCCGCCACG AGTTGCCGGA

Q P T T G L L P N A G V A \* G R A

601 CAACCTACTA CTGGGCTGCT TCCTAATGCA GGAGTCGCAT AAGGGAGAGC  
GTTGGATGAT GACCCGACGA AGGATTACGT CCTCAGCGTA TTCCCTCTCG

SalI

S T D A L E S L Q P S Q L L P V G

Selection: 1 ●1 to 4361 ●2

☒ Double Stranded  
☒ Uppercase  
☐ Bold  
☐ Underline  
☐ Italic

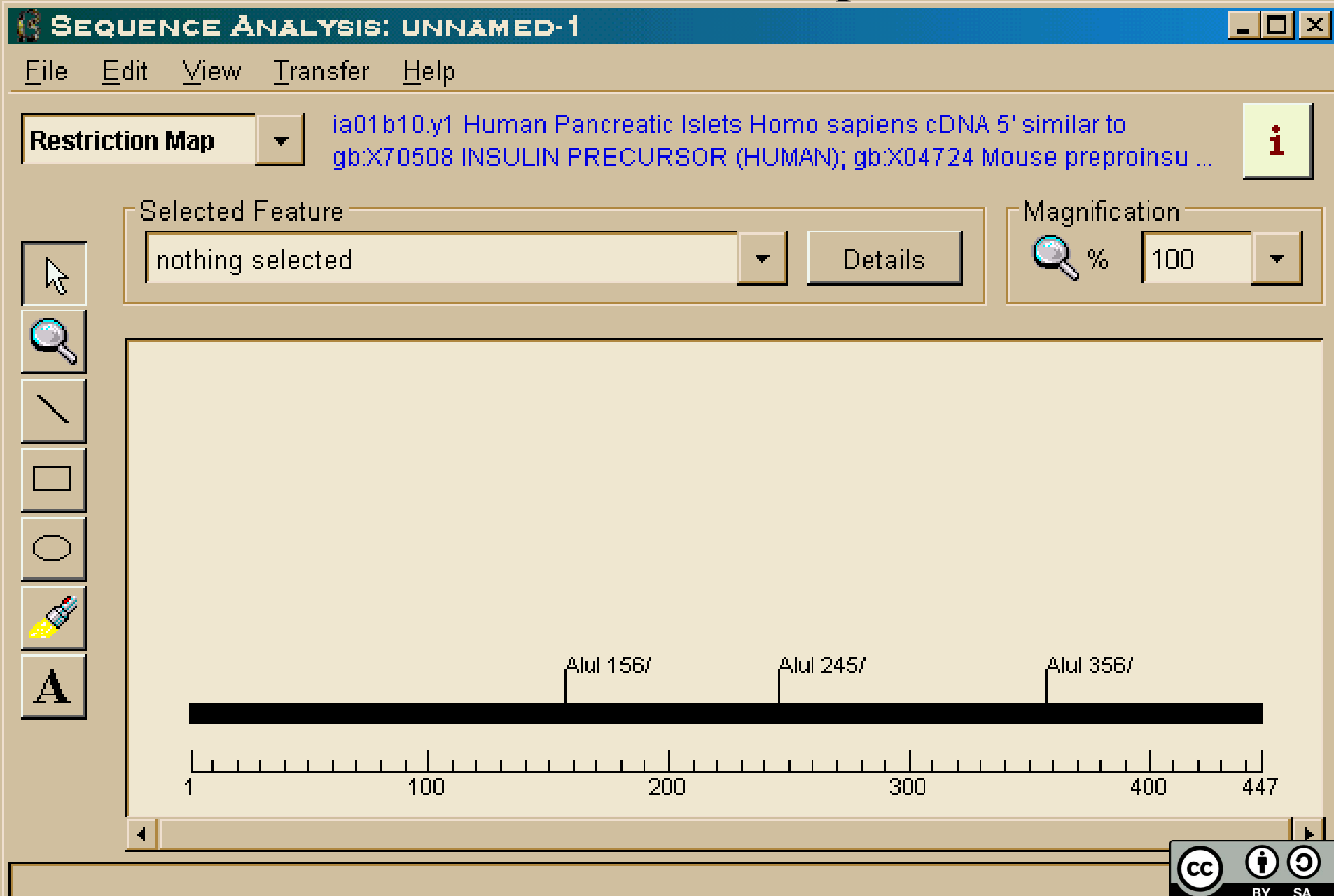
Grouping: 10  
Ruler: Left  
Translation: 1 Frame

Show Restriction Sites...

Clear

\*Showing Restriction Sites, Double Stranded, and Translation

# Restriction Map



```

1   GAATTCATACCAGATCACCGAAAACTGTCCTCCAAATGTGTCCCCCTCACACTCCCAAAT
61  TCGCGGGCTTCTGCTCTTAGACCACTCTACCCCTATTCCCCACACTCACCGGAGCCAAAGC
121 CGCGGCCCTTCCGTTCCTTTGCTTTTGAAAGACCCCCACCCGTAGGTGGCAAGCTAGCTTA
181 AGTAACGCCACTTTGCAAGGCATGGAAAAATACATAACTGAGAATAGGAAAAGTTCAGATC
241 AAGGTCAGGAACAAAGAAACAGCTGAATACCAAACAGGATATCTGTGGTAAGCGGTTCTCT
301 GCCCCGGCTCAGGGCCAAGAACAGATGAGACAGCTGAGTGATGGGCCAAACAGGATATCT
361 GTGGTAAGCAGTTCTTGCCCCGGCTCGGGGGCCAAGAACAGATGGTCCCCAGATGCGGTCC
421 AGCCCTCAGCAGTTCTTAGTGAATCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAAA
481 ATGACCCTGTACCTTATTTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTTCGCGCGC
541 TTCCGCTCTCCGAGCTCAATAAAAAGAGCCCAACCCCTCACTCGGCGCGCCAGTCTTCC
601 GATAGACTGCGTCGCCCCGGGTACCCGTATTCCCAATAAAGCCTCTTGCTGTTTGCATCCG
661 AATCGTGGTCTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTGATTGACTACCCACGACGG
721 GGGTCTTTCATTTGGGGGCTCGTCCGGGATTTGGAGACCCCTGCCAGGGACCACCGACC

```

Create Map

Clear DNA

Get Demo DNA

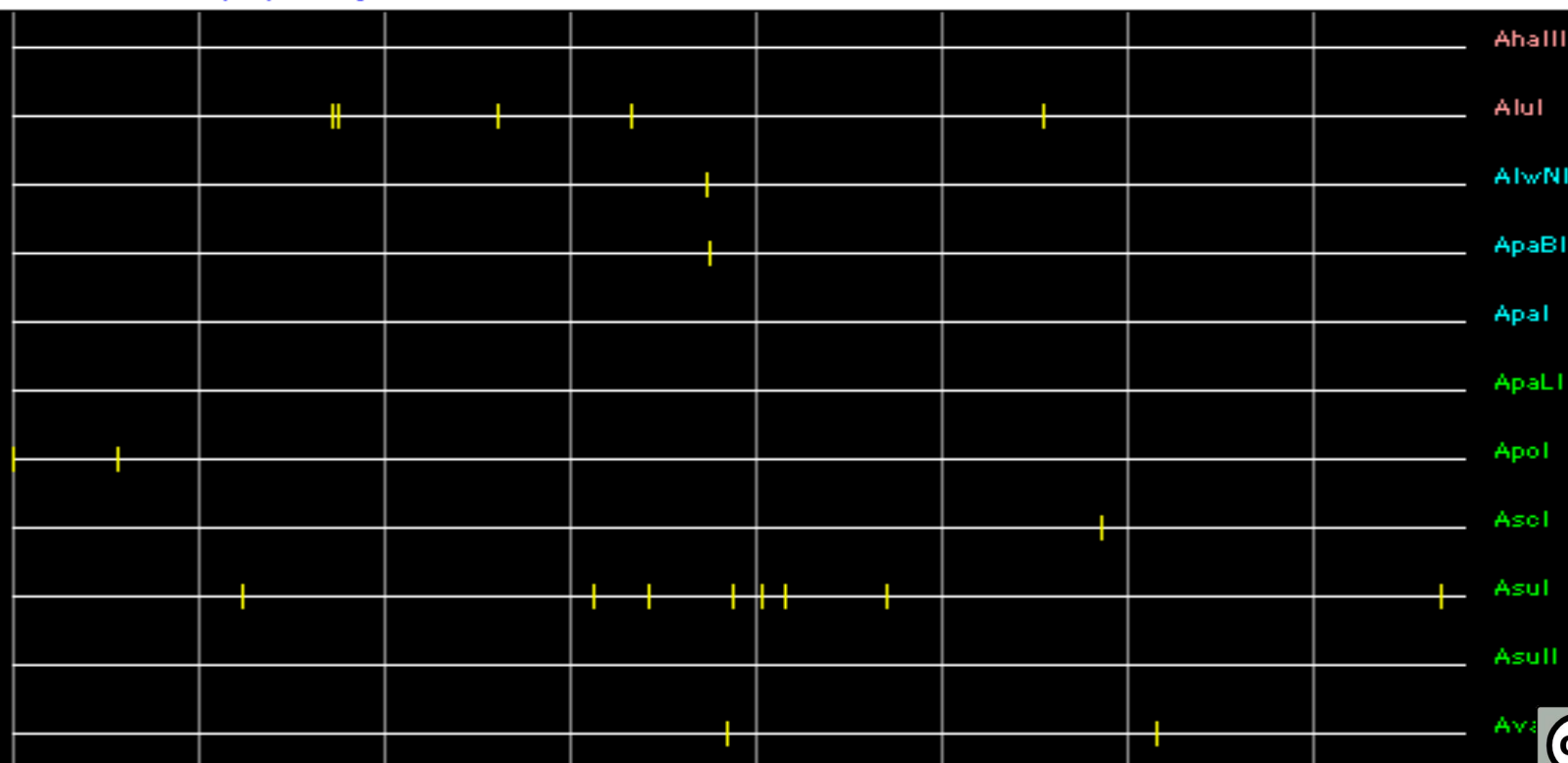
DNA is 780 base pairs in length

All restriction enzymes ▼

Graphic Display

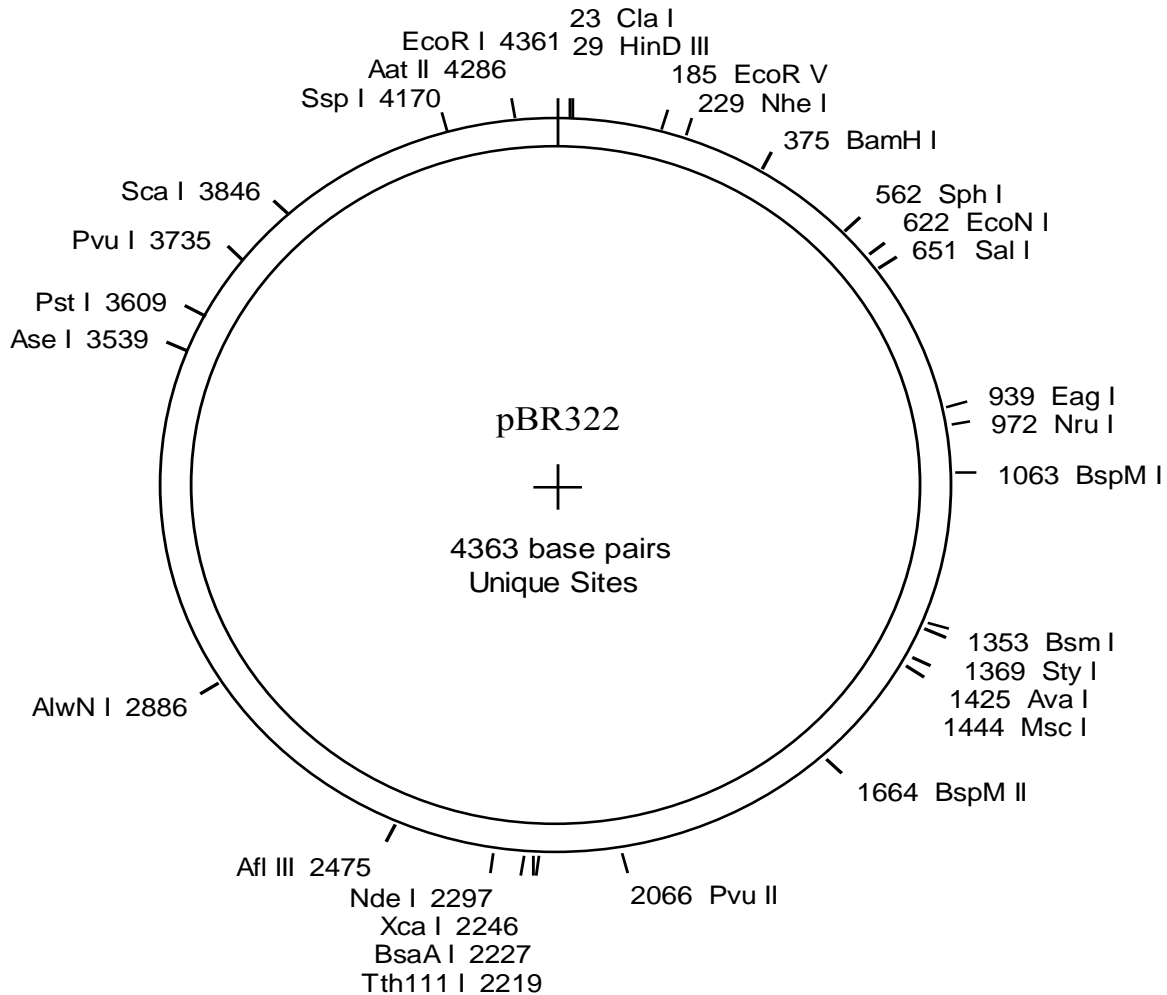
Text Display

Clicked within a 723 bp *ApoI* fragment on base 735



BY SA

# DNA Strider is simple, but still elegant



# Compute Statistics

NUCLEIC ACID STATISTICS: ia01b10.y1 HUMAN PANCREATIC ISLETS HO...

Print...

Close

## Sequence Data

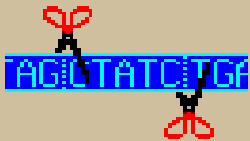
Sequence: ia01b10.y1 Human Pancreatic Islets Homo sapiens cDNA 5' similar to gb:X70508 INSULIN PRECURSOR (HUMAN); gb:X04724 Mouse preproinsulin gene II (MOUSE);, mRNA sequence.  
Accession AW583558

Total Number of Bases	447
Total Number of Degenerates	0
Percent GC Density	63.3%
Percent AT Density	36.7%
Sense Strand Mol Weight as DNA	137.88 kilodaltons
Antisense Strand Mol Weight as DNA	138.42 kilodaltons
Weight of DNA Duplex	276.3 kilodaltons
Sense Strand Mol Weight as RNA	143.92 kilodaltons
Antisense Strand Mol Weight as RNA	144.38 kilodaltons
Weight of RNA Duplex	288.3 kilodaltons
UV Absorbance as ssDNA	OD(260) 1.0 ~ 40 ug/mL = 290.12 nM
UV Absorbance as dsDNA	OD(260) 1.0 ~ 50 ug/mL = 180.96 nM
Melting Temp. as dsDNA ([K+] = 50 mM)	84.49 C

## Nucleotide Frequencies

# Gel Simulation INPUT

**GENETOOL: GEL SIMULATION** [X]



**Enzyme Selection**

Enzyme Library:

Common Enzymes [v]

Enzymes:

- ☒ AluI
- ☐ BamHI
- ☐ BglI
- ☐ BglII
- ☐ EcoRI

[All] [None]

**Cut Type**

- ☒ All
- ☐ Blunt End
- ☐ 3' Overhangs
- ☐ 5' Overhangs

**Enzyme Cut Frequency**

- ☐ Unique Cuts
- ☐ 2 Cuts or less
- ☐ 3 Cuts or less
- ☐ 4 Cuts or less
- ☐ 5 Cuts or less
- ☒ Unlimited

**Experiment Parameters**


Voltage: [100] V

Concentration: [1] %

Time: [0] : [30]  
(HH) : (MM)

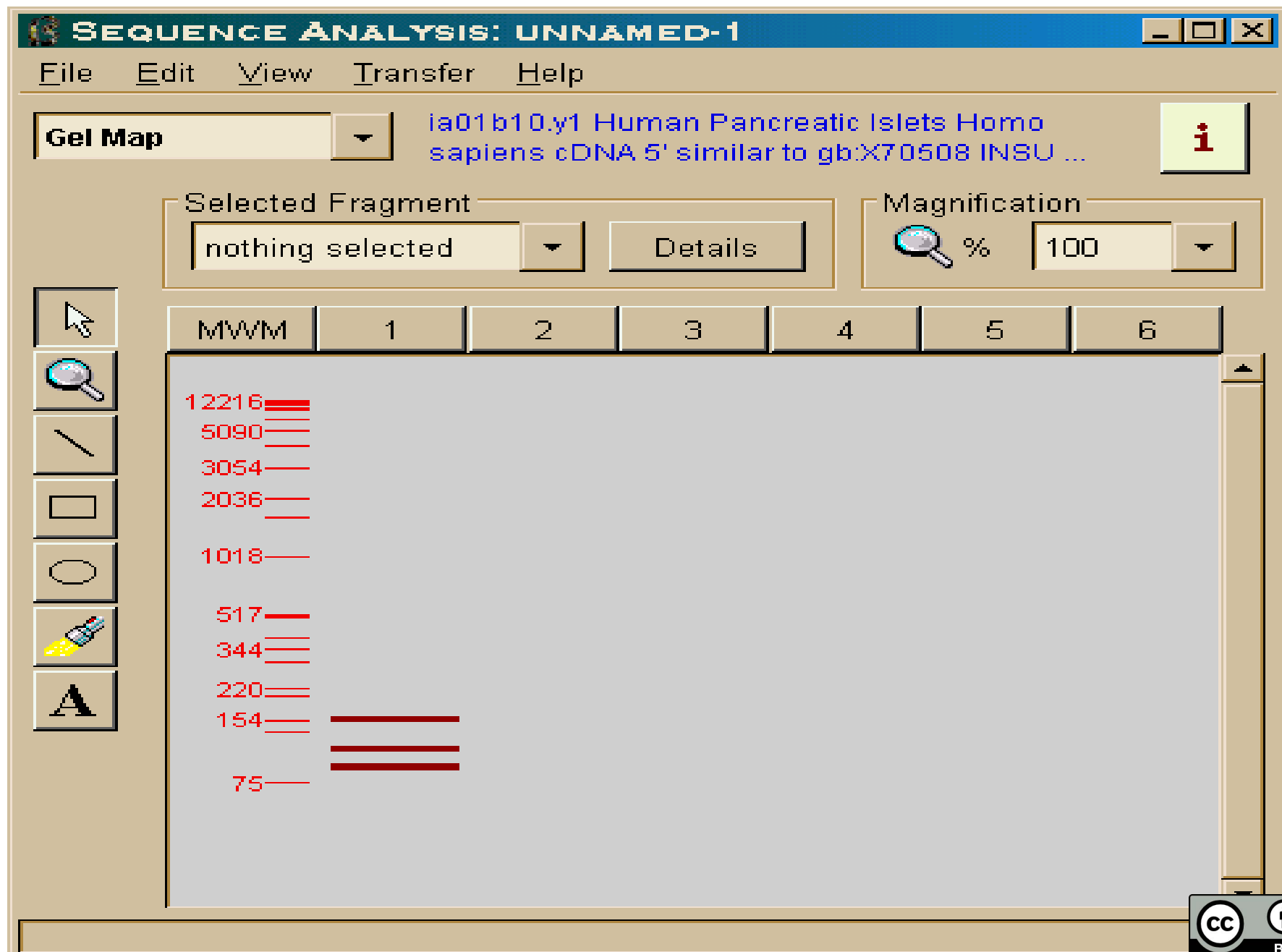
Stop searching after finding [500] cut sites.

[OK] [Cancel] [Options...]

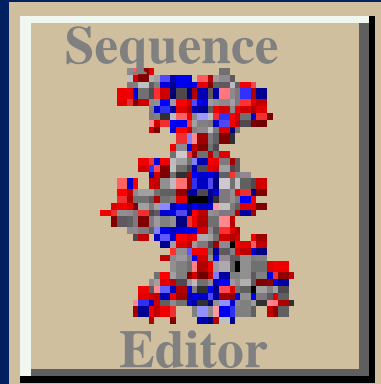




# Gel Simulation OUTPUT



## ANALYSIS TOOLS



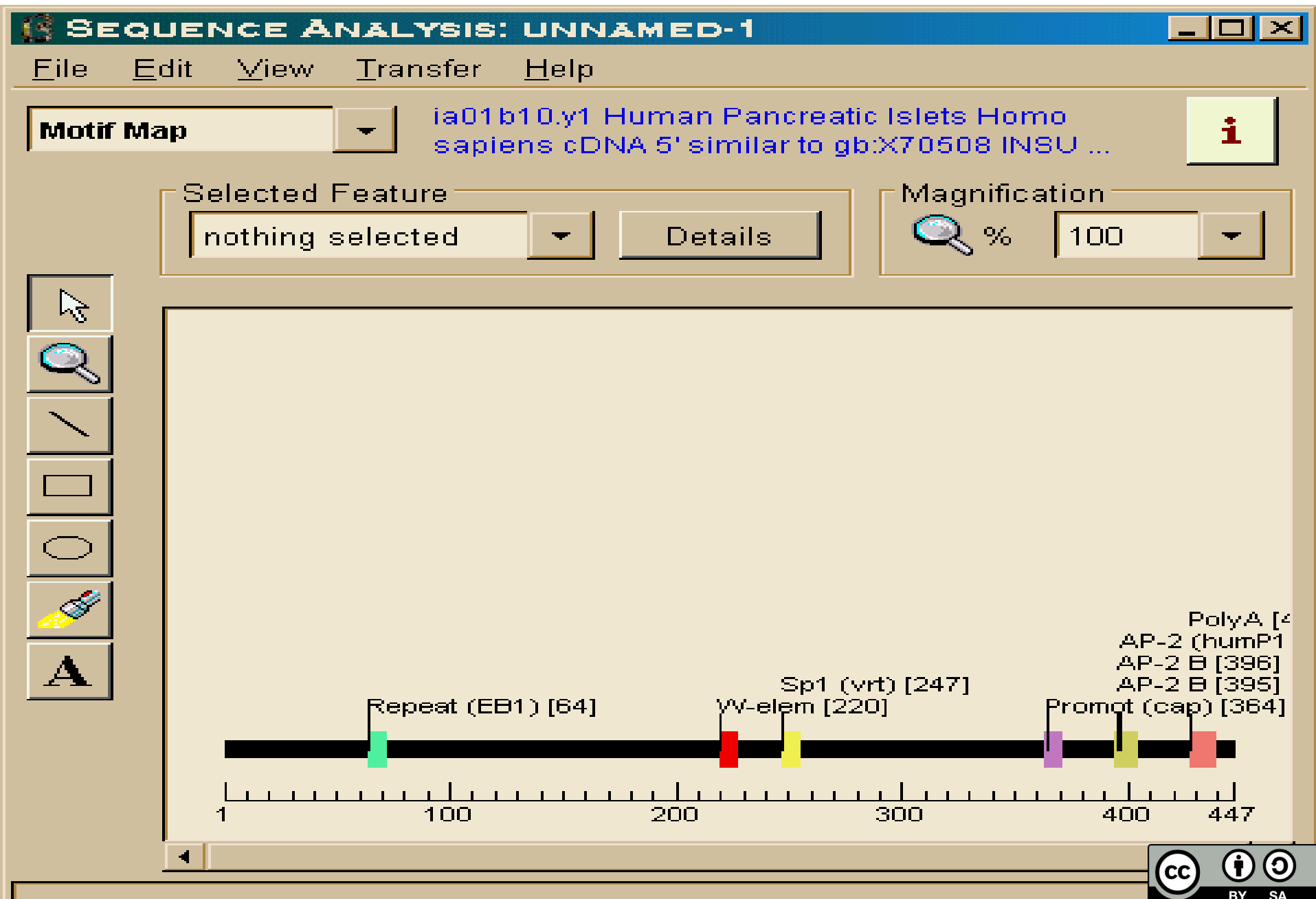
### ➤ ANALYZE MENU

- Translate
- Simulate Gel
- Compute Statistics
- Find: Motifs, ORFs, Restriction Sites, Exons, Repeats and Vectors, PCR Primers
- FAST Alignment
- Remote BLAST

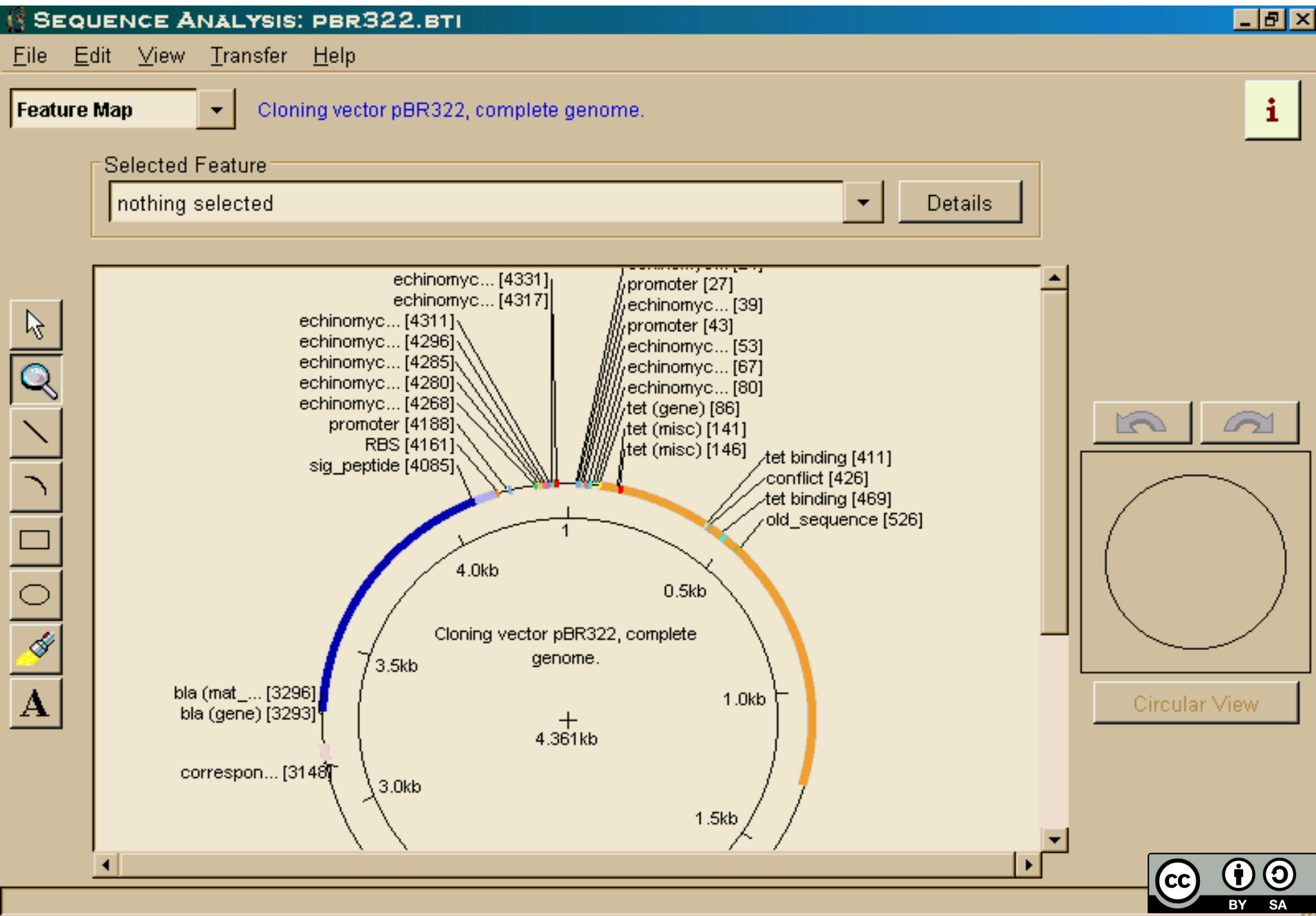
### ➤ PLOT MENU

- GC Density
- AT Density
- CpG Islands
- Fickett's Method

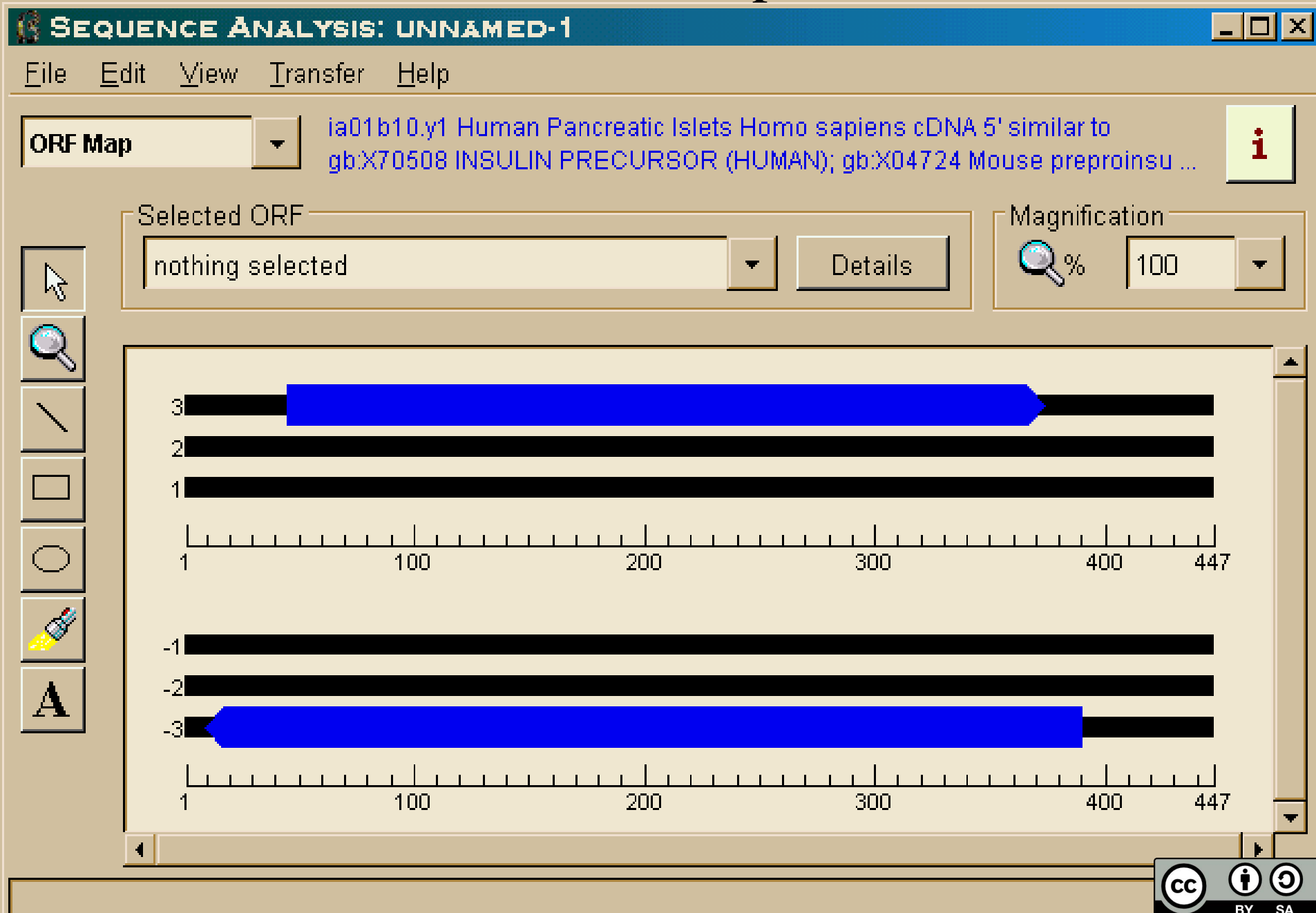
# Find Motifs



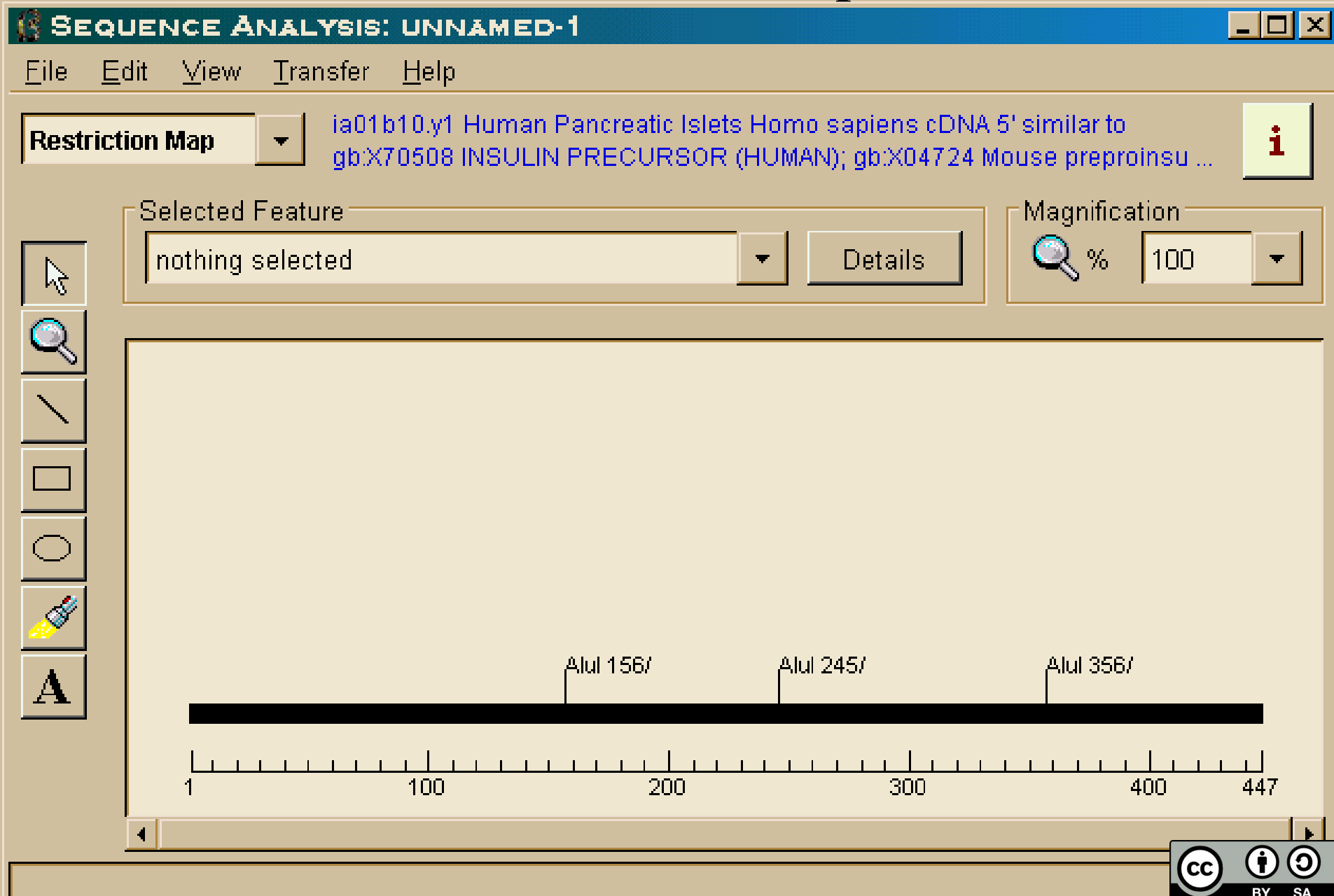
# Feature Map



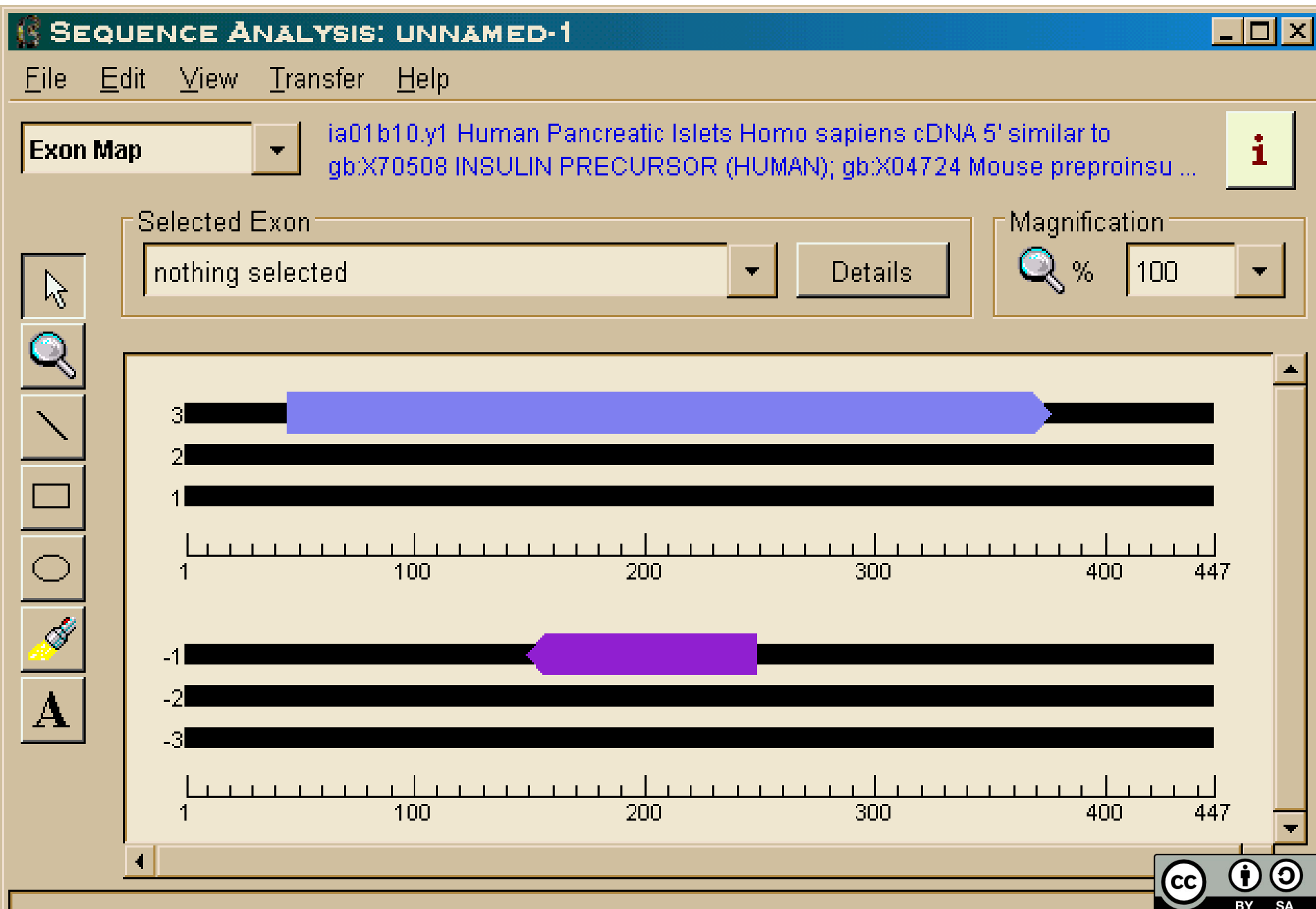
# ORF Map



# Restriction Map



# Exon Map



# Summary:

## What can you do with GeneTool?

- Design primers
- Find restriction sites
- Help identify exons
- Map and simulate gels
- Assemble contigs
- Align sequences
- Find repeats, motifs and ORFs



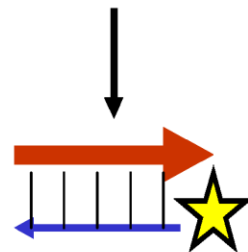




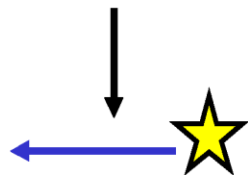
# Primer Extension : Another method to determine DNA sequence encoding the 5' end of mRNA



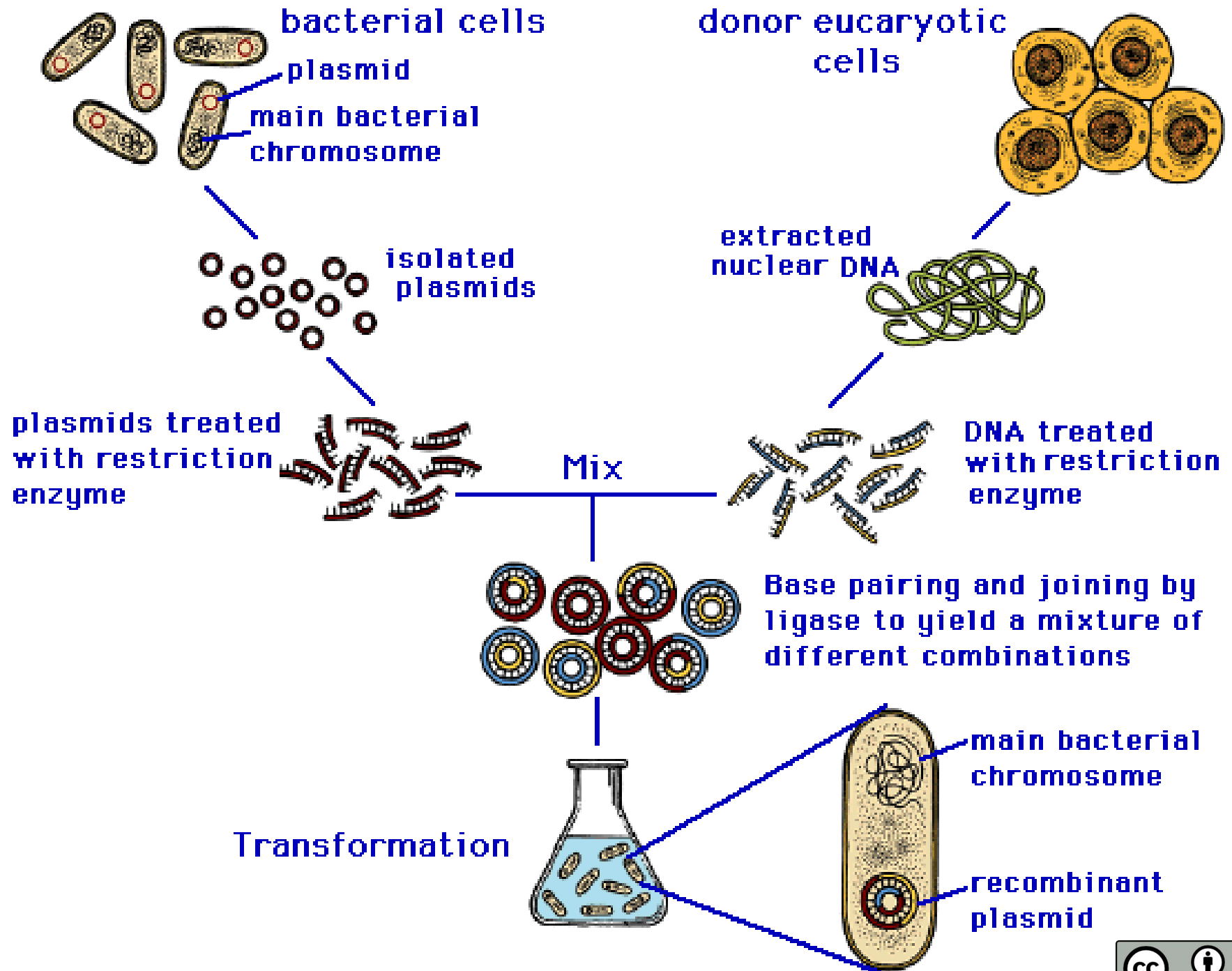
Anneal a primer, complementary to RNA, labeled on 5' end:



Reverse transcriptase + dNTPs to extend primer to 5' end of RNA



Denaturing gel electrophoresis





Organismus	Genom velikost (Mb)	Typ I	Typ II	Typ III	M-DRS*
<i>Aeropyrum pernix</i>	1·67		7		
<i>Aquifex aeolicus</i>	1·55				1
<i>Archaeoglobus fulgidus</i>	2·18	1	2	1	
<i>Bacillus subtilis</i>	4·21	†	2		1
<i>Borrelia burgdorferi</i>	1·44		2		
<i>Campylobacter jejuni</i>	1·64	1	4		1
<i>Chlamydia muridarum</i>	1·07				
<i>Chlamydia trachomatis</i>	1·05				
<i>Chlamydia pneumoniae</i> AR39	1·23				
<i>Deinococcus radiodurans</i>	2·65		4		3
<i>Escherichia coli</i> K-12	4·60	1‡			3
<i>Haemophilus influenzae</i>	1·83	2	3	1	
<i>Helicobacter pylori</i> 26695	1·66	3	14	2	
<i>Helicobacter pylori</i> J99	1·64	3	16	2	
<i>Methanobacterium thermoautotrophicum</i>	1·75	1	1		3
<i>Methanococcus jannaschii</i>	1·66	3	8		
<i>Mycobacterium tuberculosis</i>	4·40	1	1		
<i>Mycoplasma genitalium</i>	0·58	1			
<i>Mycoplasma pneumoniae</i>	0·81	1	1		
<i>Neisseria meningitidis</i> serotype A	2·18	3	7	2	
<i>Neisseria meningitidis</i> serotype B	2·27	1	4	1	
<i>Pyrococcus abyssi</i>	1·77	1	4		
<i>Pyrococcus horikoshii</i>	1·74		3		
<i>Rickettsia prowazekii</i>	1·10				
<i>Synechocystis</i> species	3·57		1		1
<i>Thermatoga maritima</i>	1·80		1		
<i>Treponema pallidum</i>	1·16				
<i>Ureaplasma urealyticum</i>	0·71	1	1		

\* Putative methylation-dependent restriction systems.

† Some strains of *B. subtilis* do have a type I R-M system.

‡ Many strains of *E. coli* have a chromosomally encoded type I R-M system; to date alleles conferring 11 different specificities have been identified (Barcus *et al.*, 1995).

Uveřejněné materiály jsou určeny studentům Vysoké školy chemicko-technologické v Praze

jako studijní materiál. Některá textová i obrazová data v nich obsažená jsou převzata

z veřejných zdrojů. V případě nedostatečných citací nebylo cílem autora/ů záměrně poškodit event. autora/y původního díla.

S eventuálními výhradami se prosím obraťte na autora/y konkrétního výukového materiálu, aby

bylo možné zjednat nápravu.

The published materials are intended for students of the University of Chemistry and Technology, Prague as a study material. Some text and image data contained therein are taken from public sources. In the case of insufficient quotations, the author's intention was not to intentionally infringe the possible author(s) rights to the original work.

If you have any reservations, please contact the author(s) of the specific teaching material in order to remedy the situation.

# Gene expression



EUROPEAN UNION  
European Structural and Investing Funds  
Operational Programme Research,  
Development and Education



MINISTRY OF EDUCATION,  
YOUTH AND SPORTS

# **Gene Expression**

## **prokaryotic x eukaryotic systems**

**Bacteria:**  
**high expression,**  
**growth speed,**  
**cost effective medium**

# **Which vector?**

## **Compatibility with the host cells**

### **Combinations:**

- strong promoter**
  - binding site for ribosome**
  - termination sequence**
- specific sequences – for isolation, solubilisation, detection...**

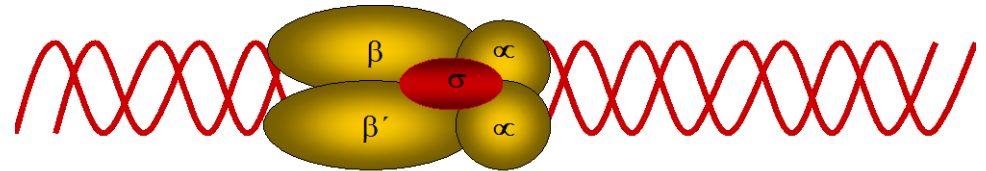


## Organism

Characteristic	<i>E. coli</i>	<i>S. cerevisiae</i>	<i>P. pastoris</i>	Insect	Mammalian
High growth rate	E <sup>a</sup>	VG	VG	P-F	P-F
Availability of genetic systems	E	G	F	F-G	F-G
Expression levels	E	VG	E	G-E	P-G
Low-cost media available	E	E	E	P	P
Protein folding	F	F-G	F-G	VG-E	E
Simple glycosylation	No	Yes	Yes	Yes	Yes
Complex glycosylation	No	No	No	Yes <sup>b</sup>	Yes
Low levels of proteolytic degradation	F-G	G	G	VG	VG
Excretion or secretion	P normally VG in special cases	VG	VG	VG	E
Safety	VG	E	VG	E	G

# Strong bacteriophage promoters.

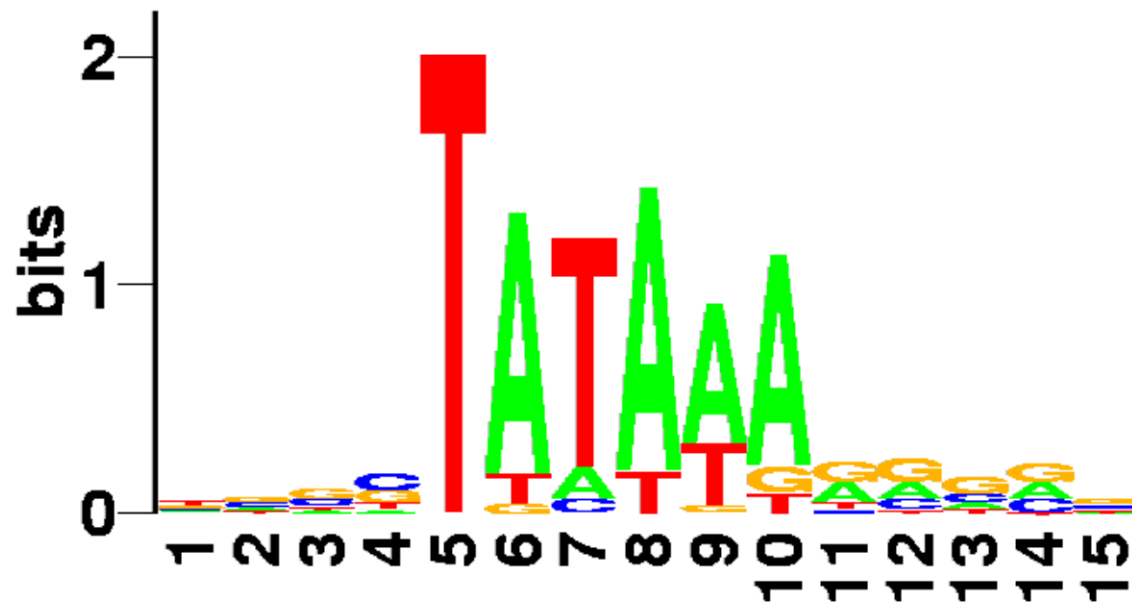
From bacteriophages T3, T7 and SP6



## Cell promoters

*trp* promoter

*lac* promoter



-10 "TATA" box pro 60 lidských promotorů

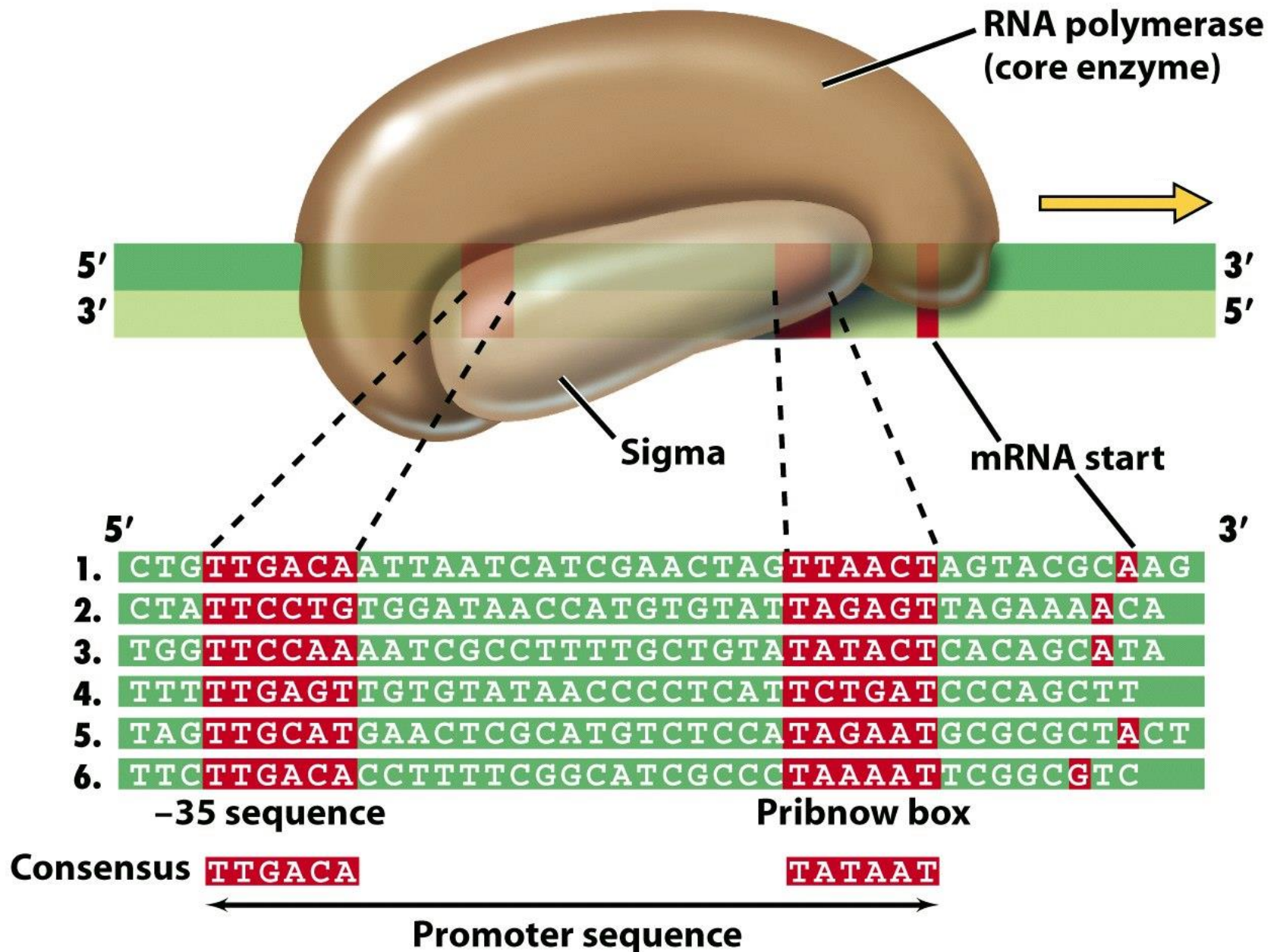


Figure 7-30 Brock Biology of Microorganisms 11/e  
© 2006 Pearson Prentice Hall, Inc.

# Host-vector systems

## Bacteria

Gram-negative

*E. coli*: known physiology and genetics, fast growth, high cell density, high yield

Troublesome secretion, proteolytic degradation, formation of inclusion bodies

# *E. coli*

Solubilization (inclusion bodies)

Secretion

Benefits

- Protection against proteolysis

- Support of correct folding (e.g. S-S bridges)

- Less contaminants (endotoxins)

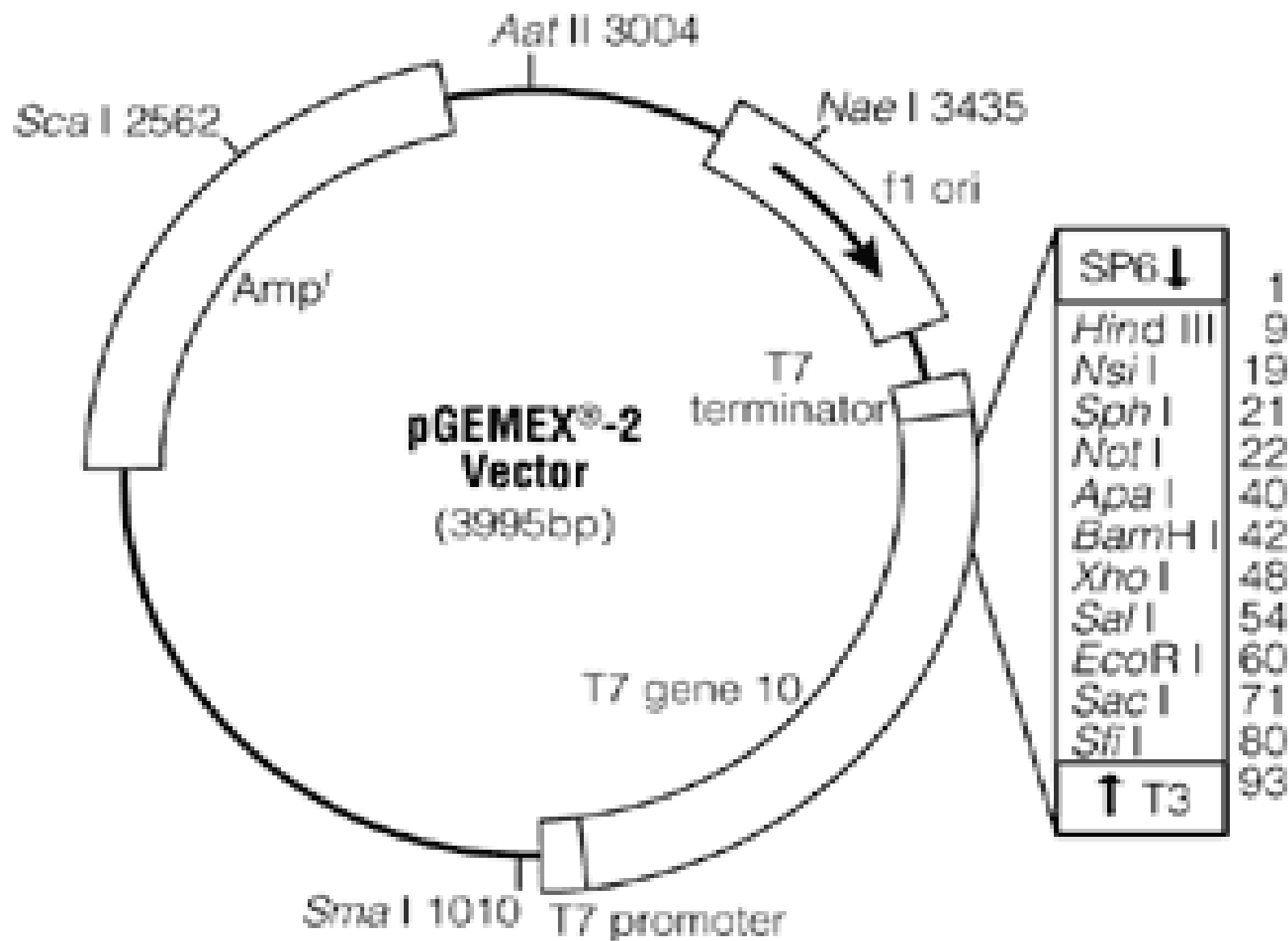
- Continual system with immobilized cells possible

# Gram-positive bacteria

## *Bacillus subtilis*

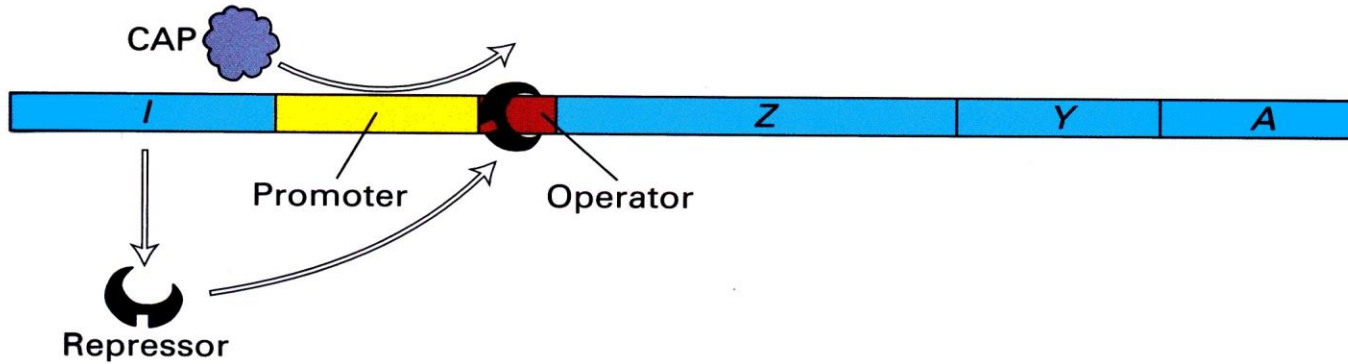
- Benefits: fast growth, lack of outer membrane, efficient secretion
- Drawbacks: proteases, smaller choice of vectors and promoters

*Streptococcus cremoris* and *Streptomyces* sp.

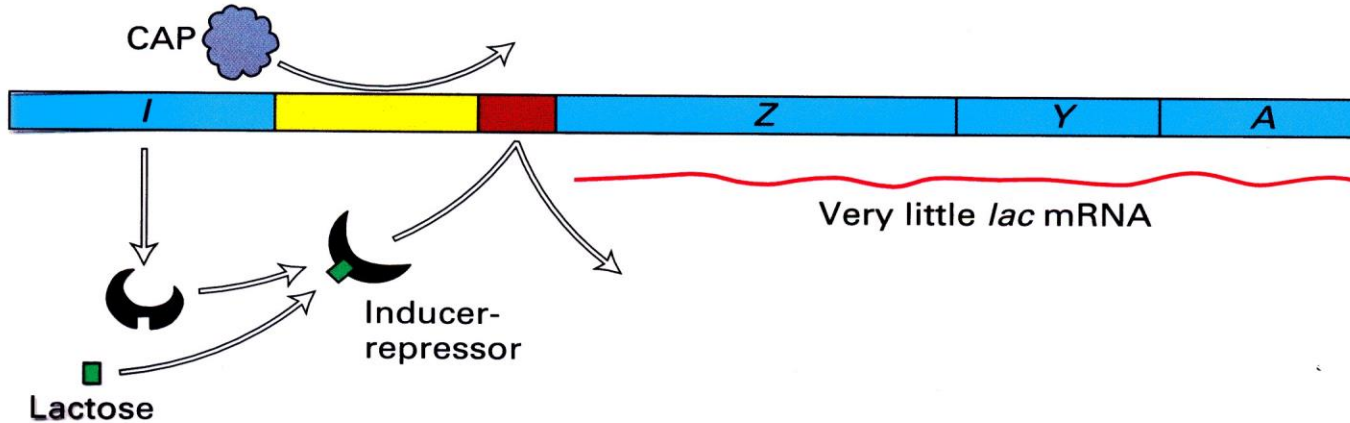


0011VB003\_0A

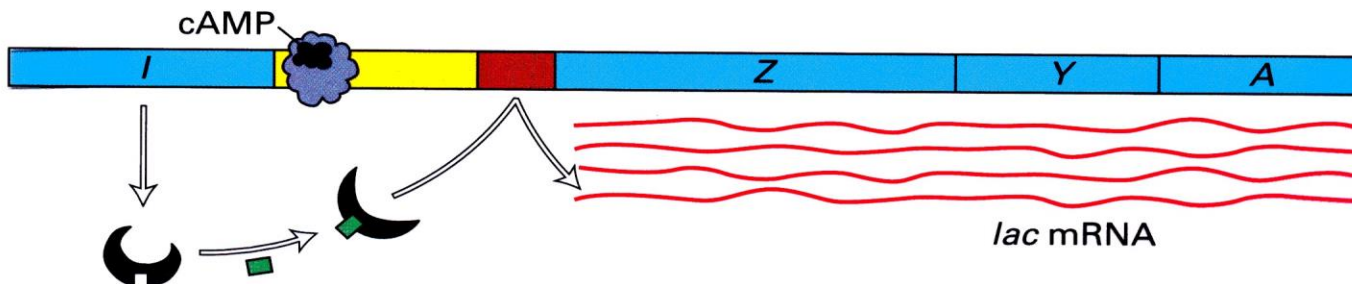
(a) Glucose present (cAMP low); no lactose



(b) Glucose present (cAMP low); lactose present



(c) No glucose present (cAMP high); lactose present





# Lac promoter - complex regulation

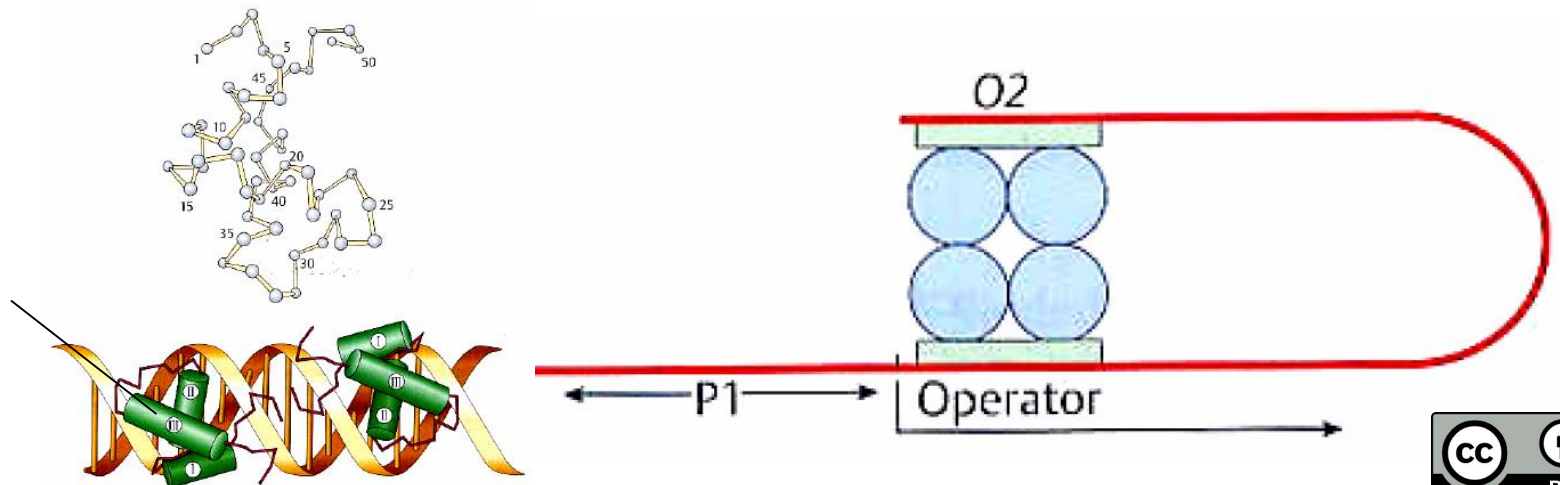
**3 promoters – P1, P2, P3**

**3 operators – O1, O2, O3 – palindrom sequences  
(O1 – main regulator)**

**Repressor LacI – homotetramer – subunits MW 38 kDa**

**Two dimers – binding to O1 and O2 or O3**

**→ loop – transcription block**



# CAP

**Homodimer – both monomers – ATP binding**

**Activation of transcription CAP regulone**

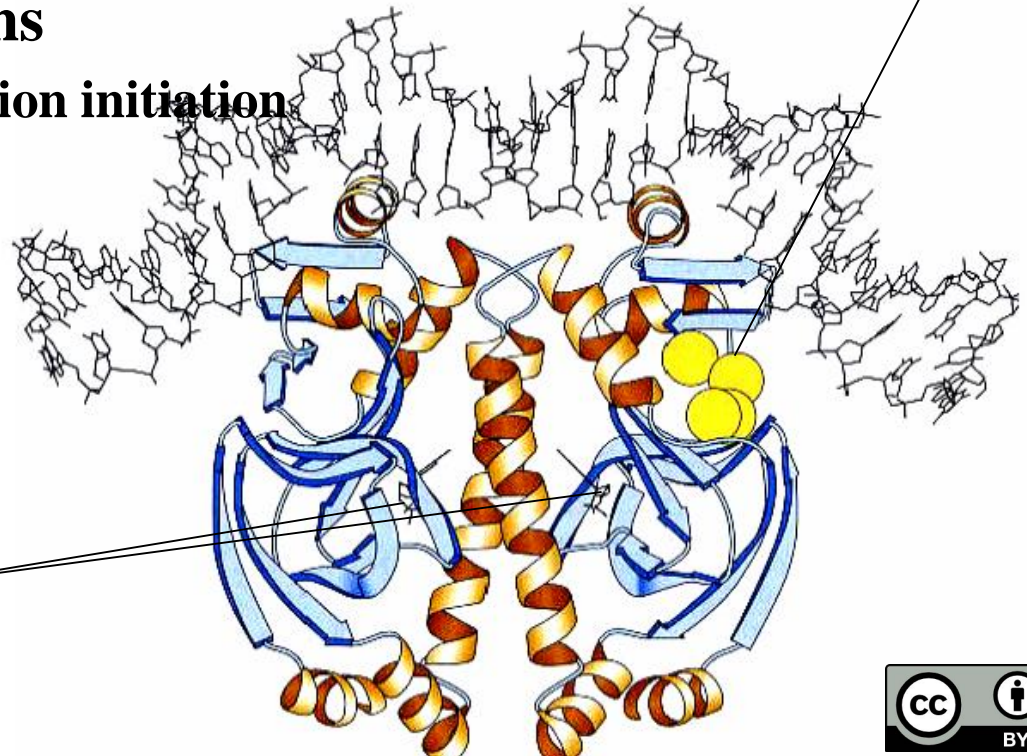
- sensitive to catabolites
- lac, gal, ara, mal

**Binding to various regions**

**– various degree of transcription initiation**

**AA in contact with  
RNA polymerase**

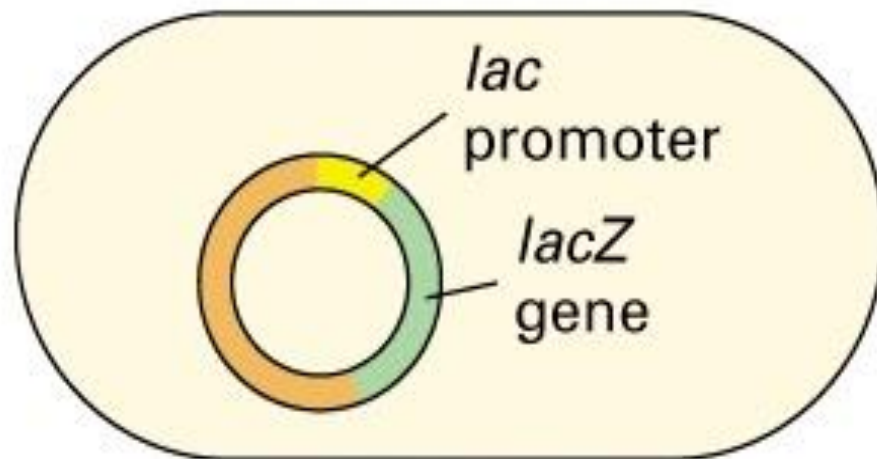
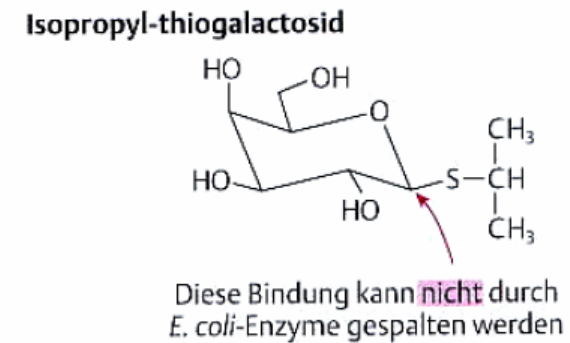
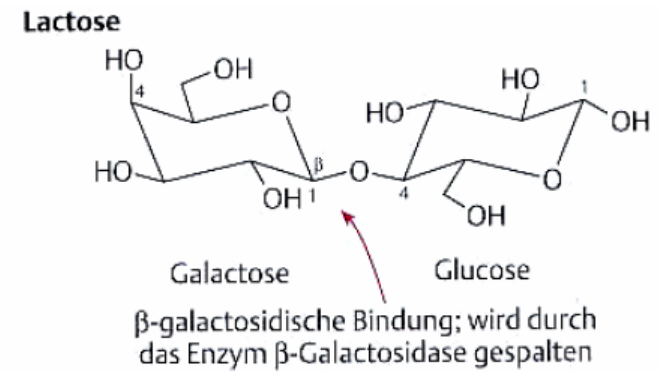
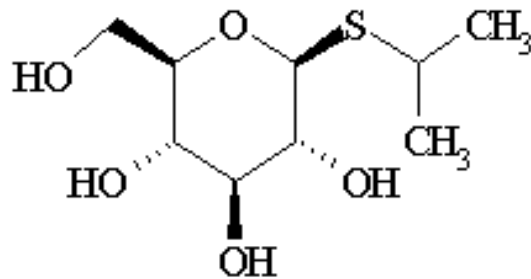
**cAMP**



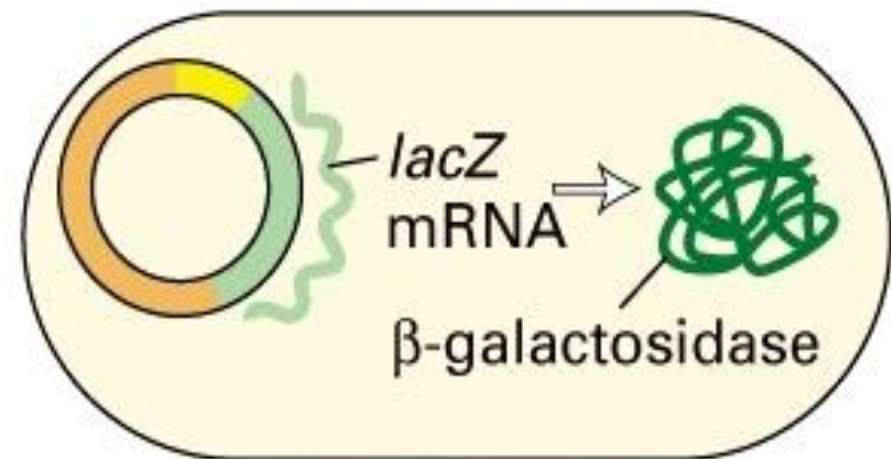
# Expression system

## Induction of *lac* promoter with IPTG

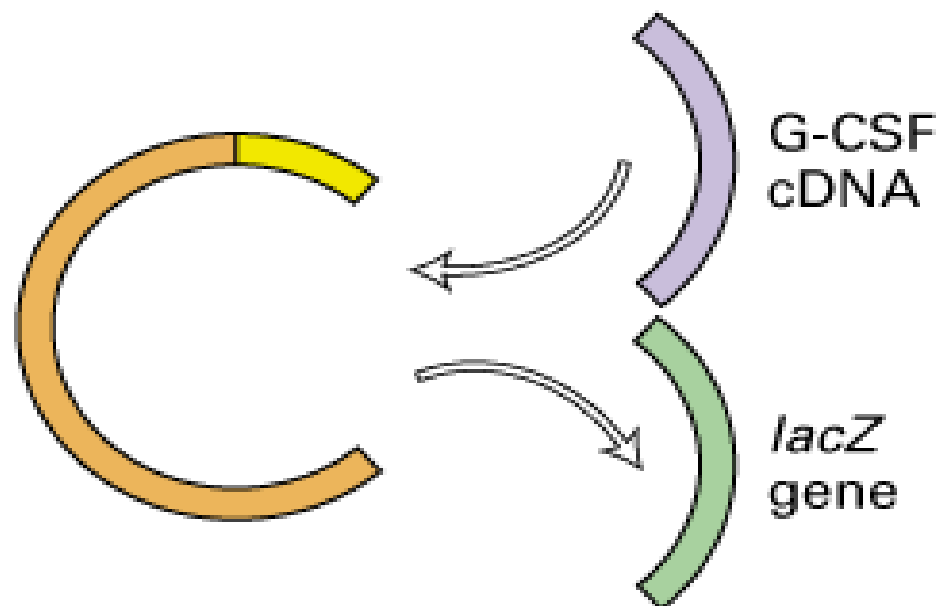
### isopropyl- $\beta$ -D-thiogalaktopyranoside



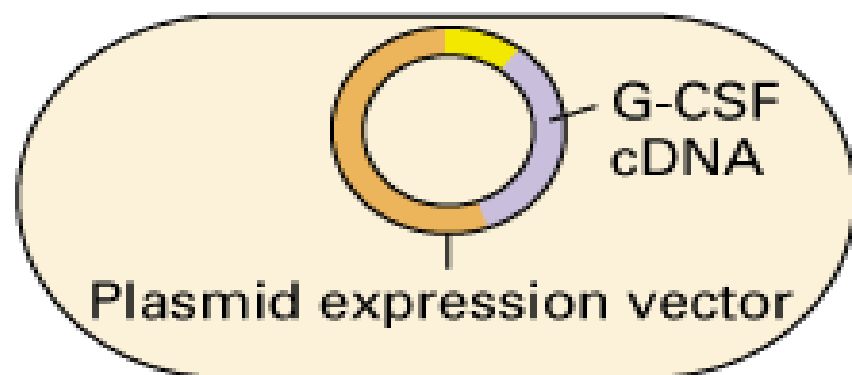
- IPTG



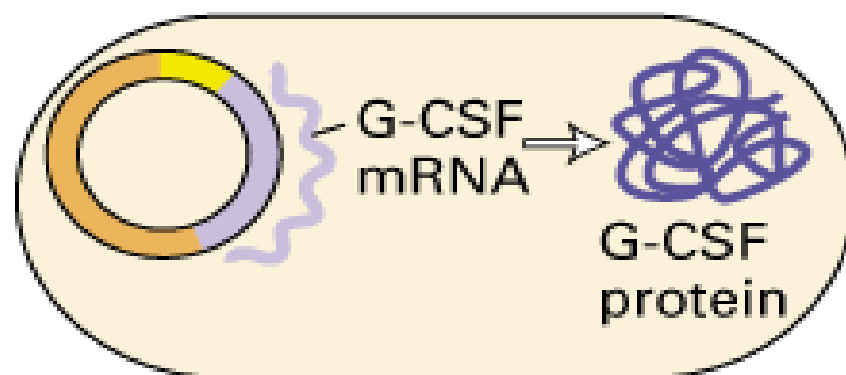
+ IPTG



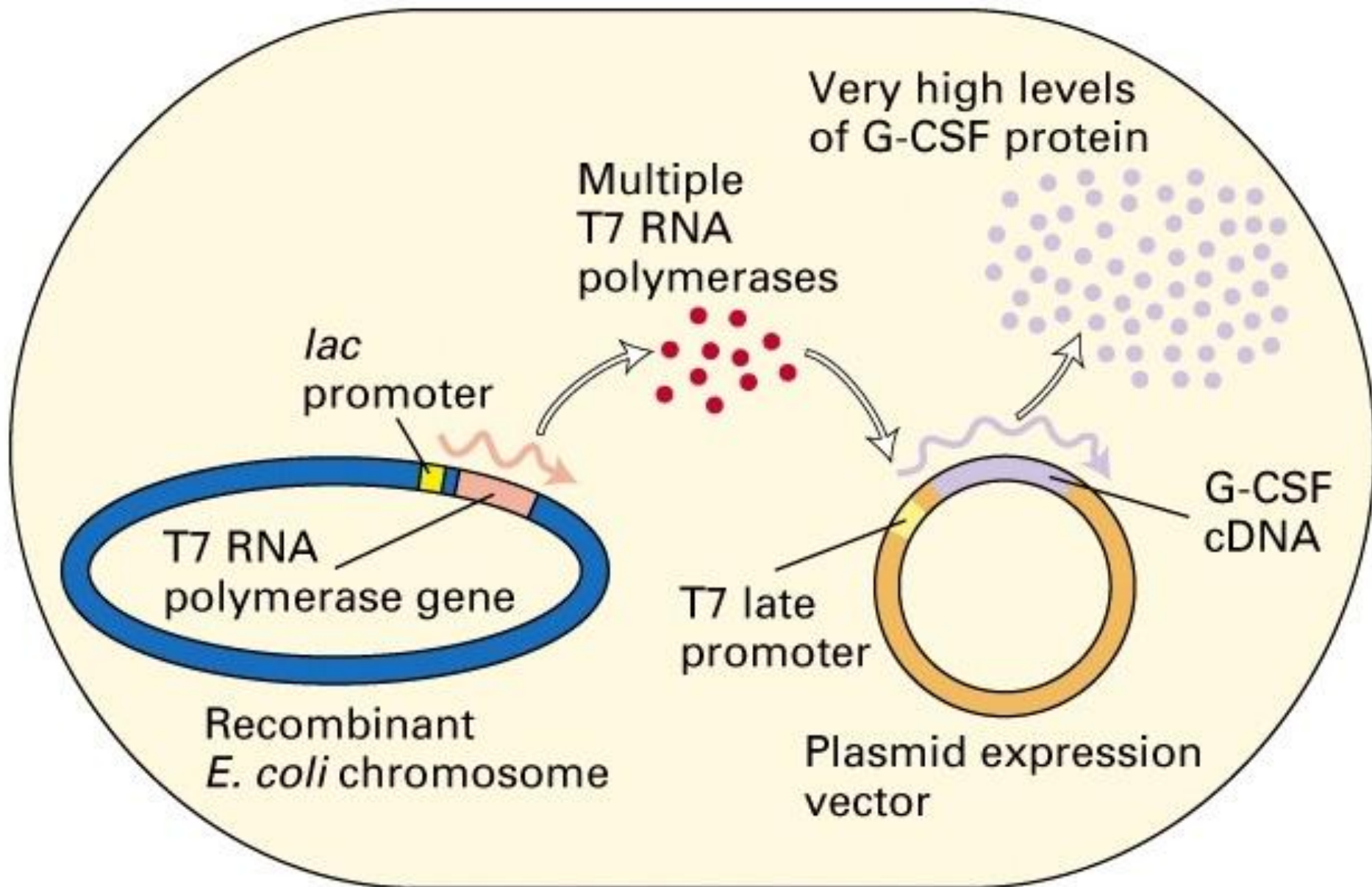
Transform  
*E. coli*



– IPTG



+ IPTG

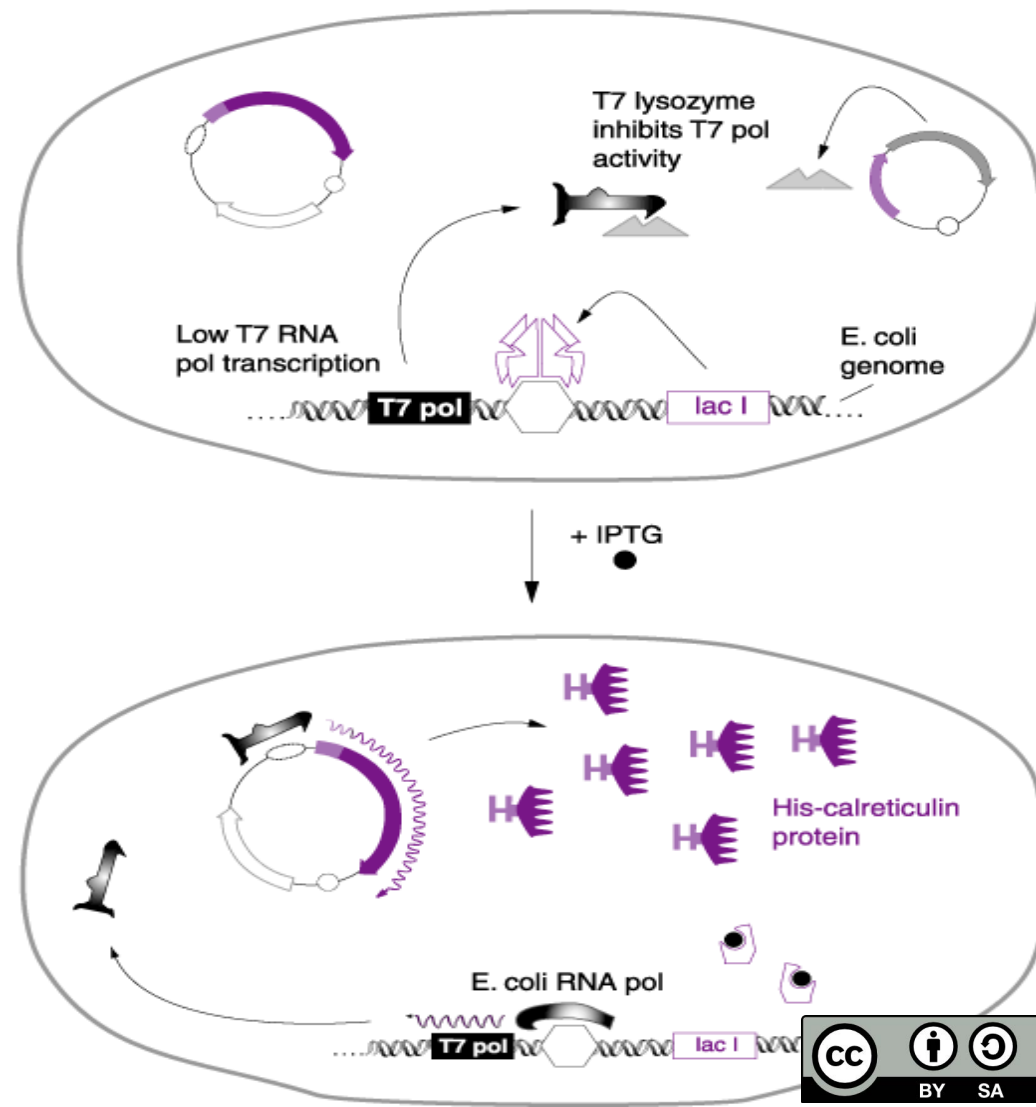
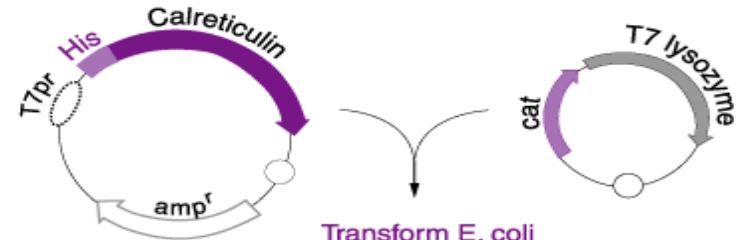
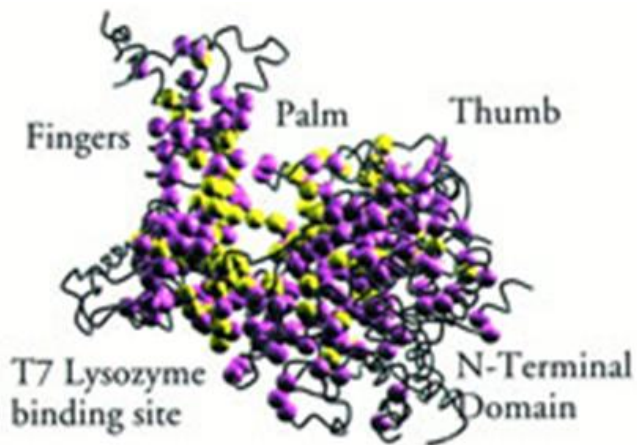




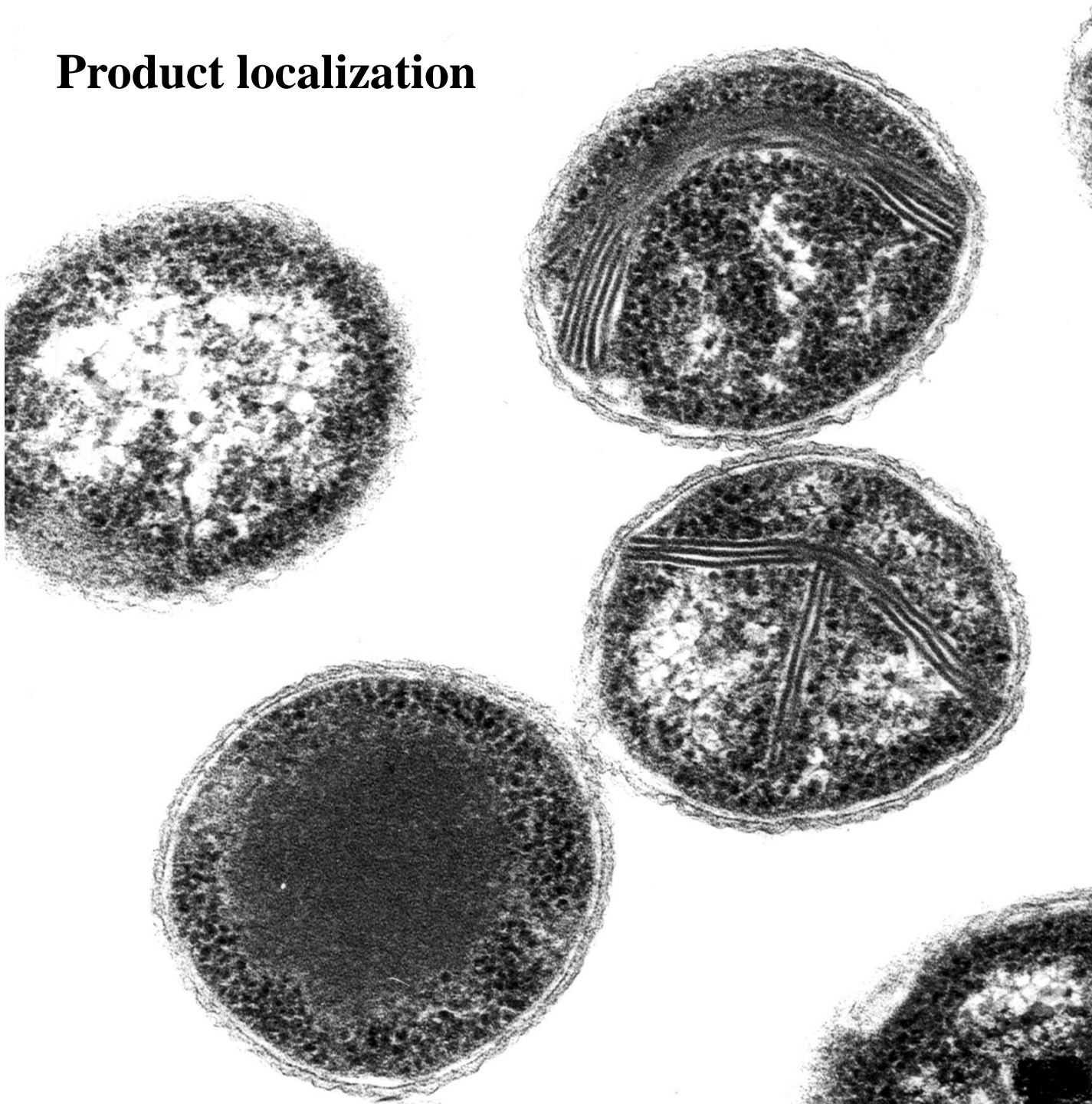
# Regulation of expression in *E. coli*

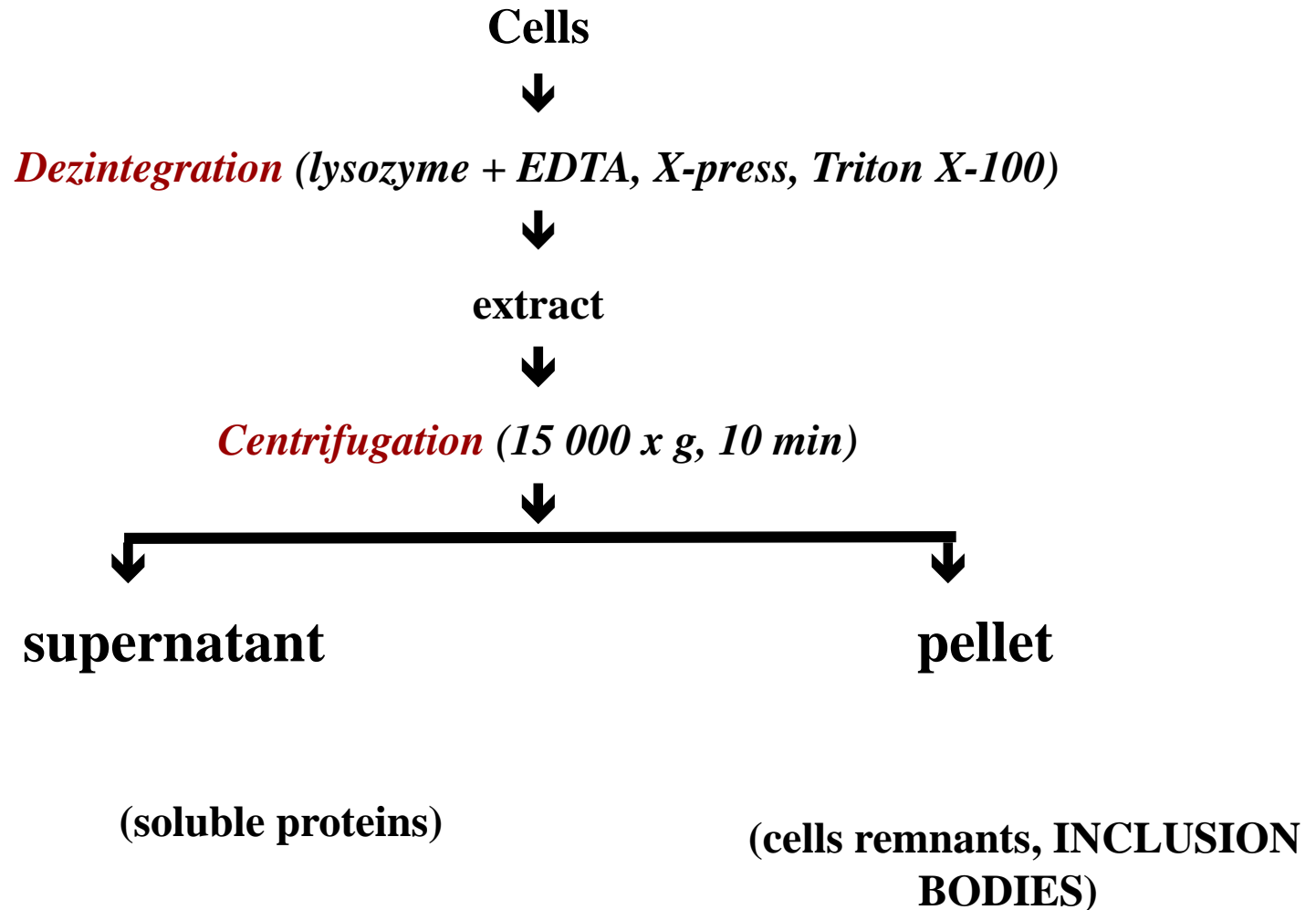
Lac repressor – repression of transcription of T7 polymerase in the absence of IPTG (imperfect)

T7 lysozyme – inhibition of T7 polymerase in cells with pLys plasmide (constitutive production of T7 lysozyme)



# Product localization







## ***Dissolution***

***8 M urea + DTT***

***6 – 8 M guanidinium.HCl + DTT***

***10 – 20% acetic acid***



**denatured protein**

**in monomeric reduced form**



## ***Purification***

***ionex chromatography, gel filtration, affinity chromatography***



**purified, denatured protein**



***getting off the denaturation agents via dilution or dialysis***

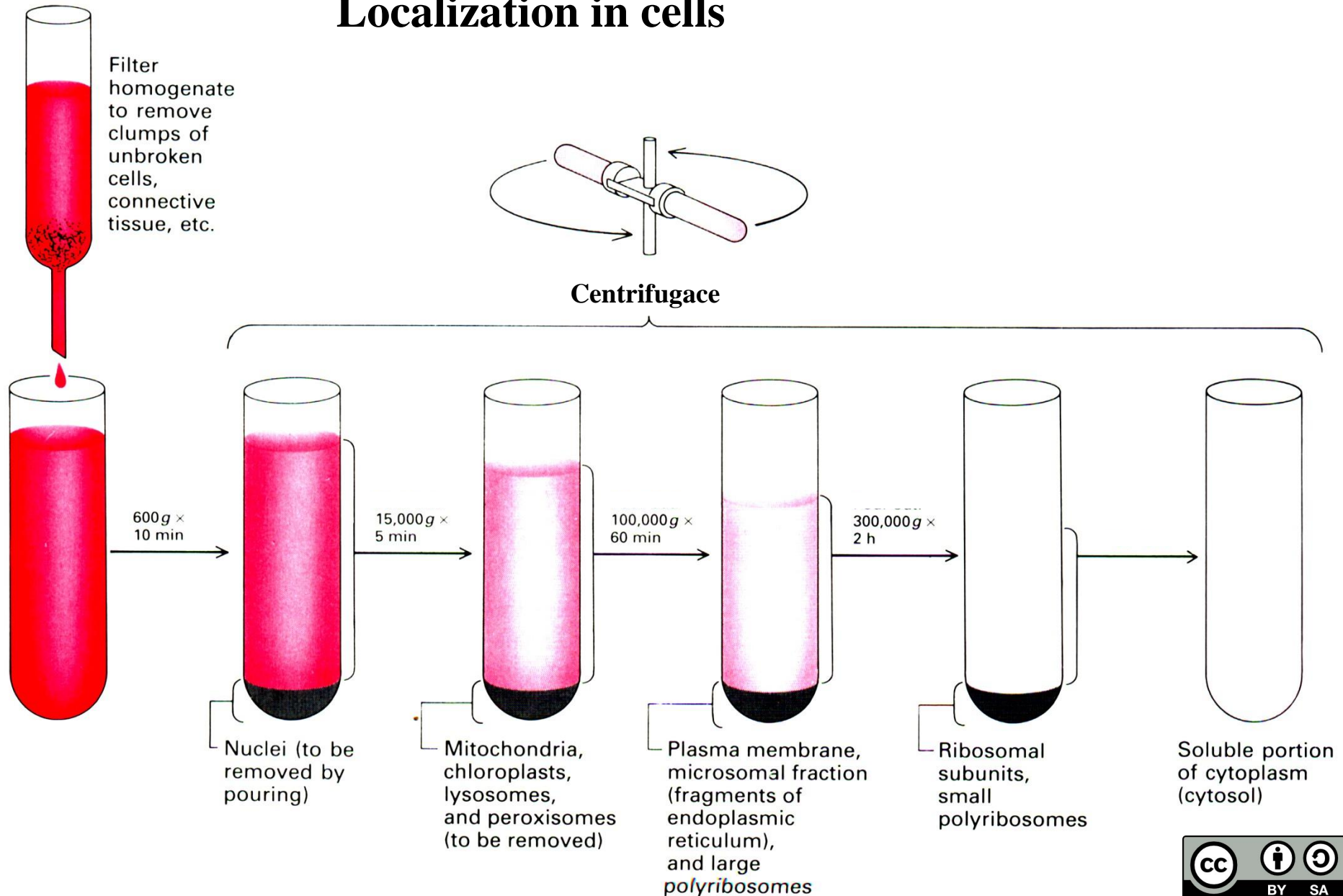


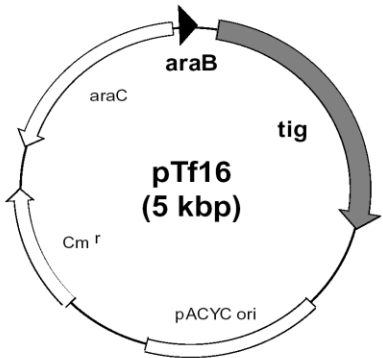
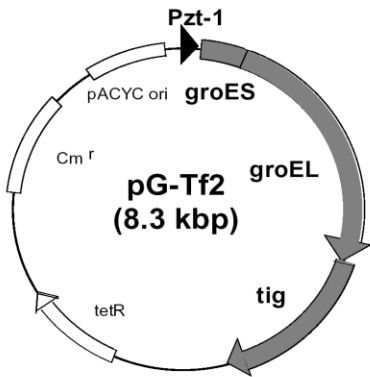
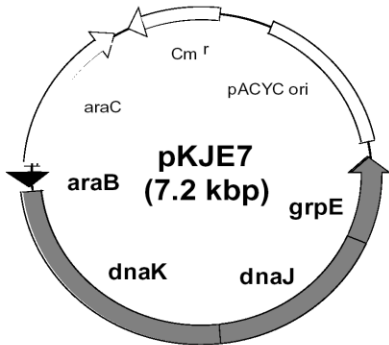
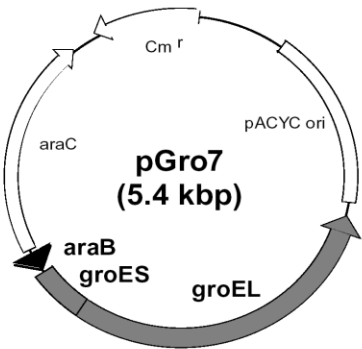
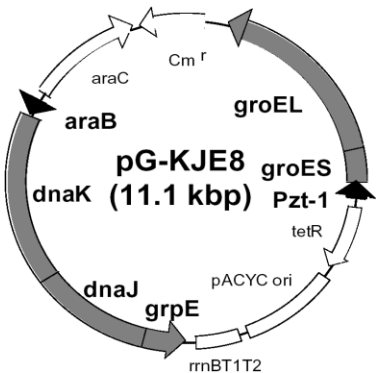
## ***Renaturation***

# Localization of product

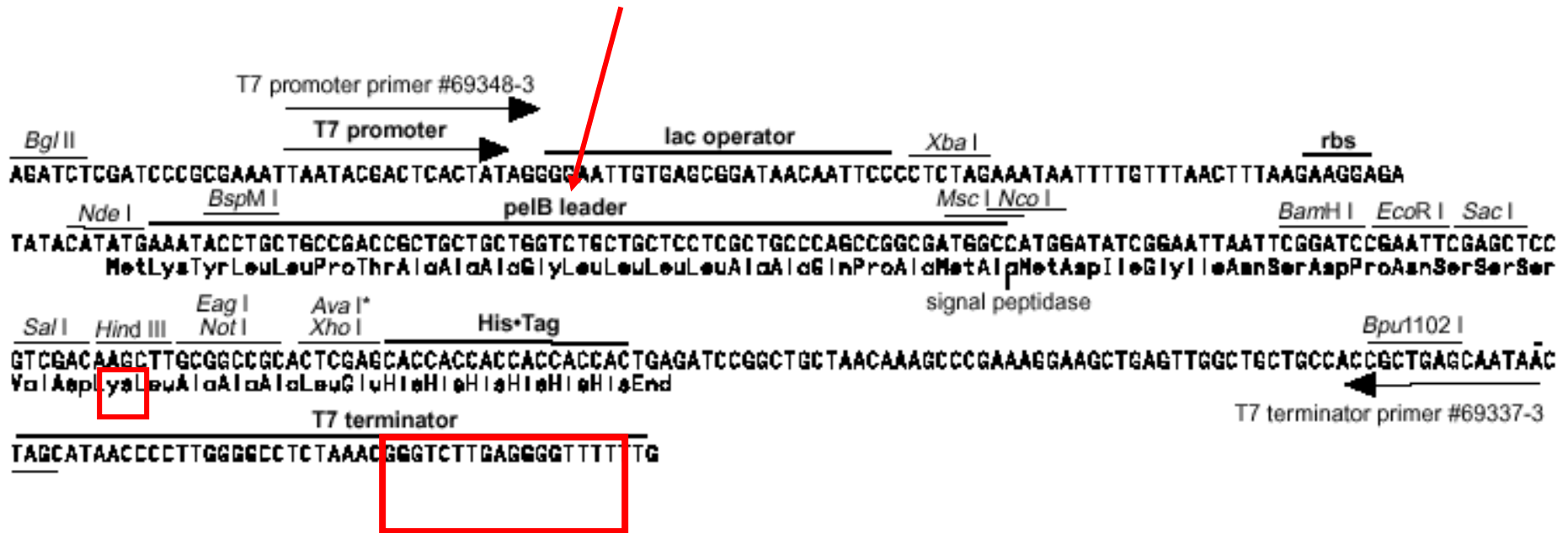
## Fractionation of cells / product isolation

### Localization in cells



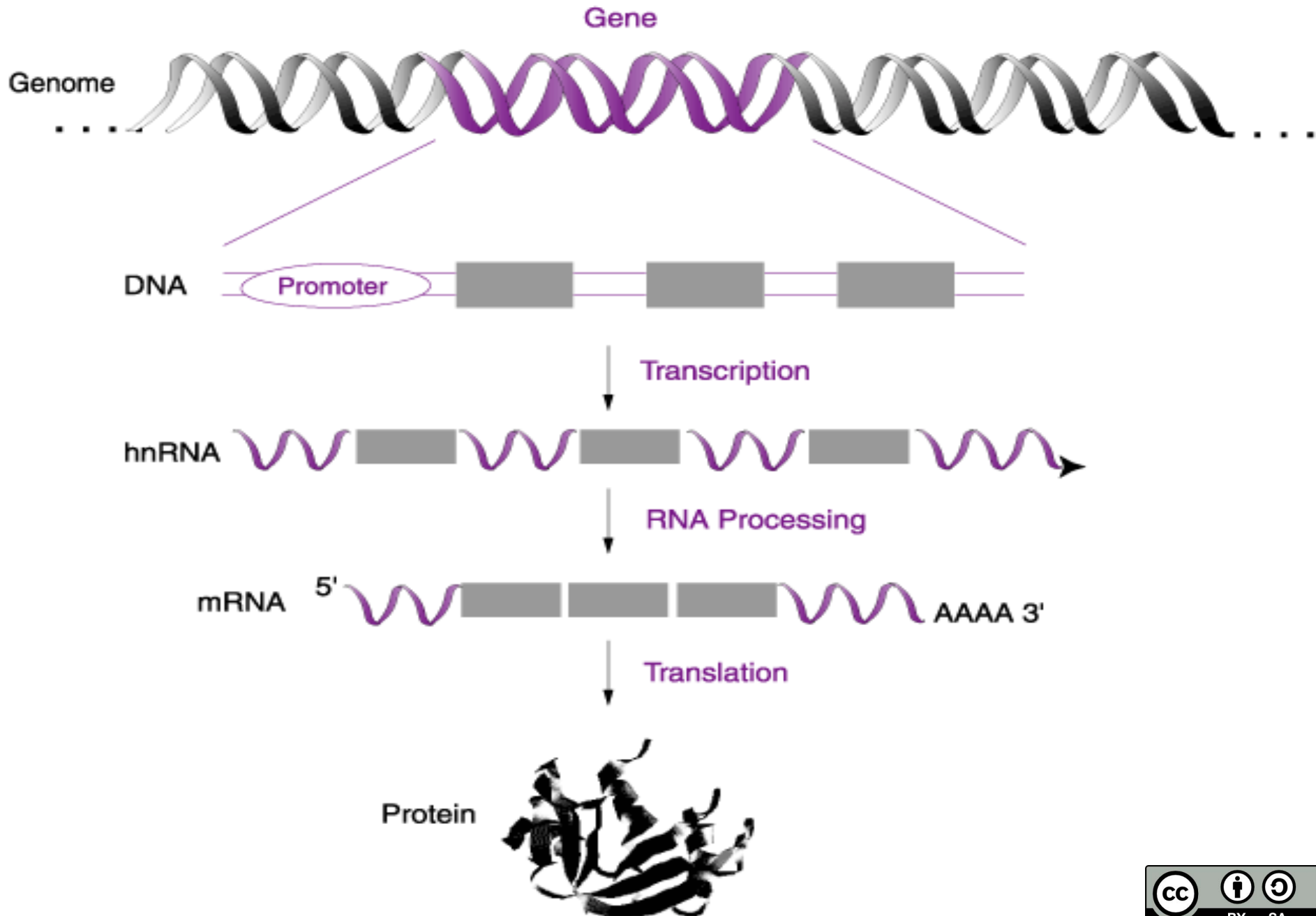


Plasmide	Chaperone	Promoter	Inducer	Marker
pG-KJE8	dnaK-dnaJ-grpE	araB	L-Arabinosis	Cm
	groES-groEL	Pzt1	Tetracycline	
pGro7	groES-groEL	araB	L-Arabinosis	Cm
pKJE7	dnaK-dnaJ-grpE	araB	L-Arabinosis	Cm
pG-Tf2	groES-groEL-tig	Pzt1	Tetracycline	Cm
pTf16	tig	araB	L-Arabinosis	Cm



## Signaling sequences of different secreted proteins

# EUKARYOTIC EXPRESSION SYSTEMS



# Yeasts

**Selection – auxotrophy (production strain purity control)**

**Shuttle vectors**

**Integrative plasmids – Yip**

- Without autonomous replication
- integration to chromosome

**Episomal plasmids – YEp - (50-200/cell), unstable**

- Autonomous replication

autonomous replicative sequence (ARS)

origin of replication of 2  $\mu$ m DNA plasmide (endogenous yeast plasmide)

**Centromere plasmids – YCp – insertion of centromeres**

# **Promoters**

**LDH – lactate dehydrogenase**

**ADH - alcohol dehydrogenase**

**AOX1 – alcohol oxidase - methanol**

**CUP1 - metallothionein -  $\text{Cu}^{2+}$**

**GAL1 - inducible with galactose (repression with glucose)**

**PGK - fosfoglycerate kinase – constitutive**

**PHO5 - secreted acidic phosphatase**

**- low levels or absence of phosphate**

# **Terminators**

## **Sequences encoding the signaling peptides**

- secretion of recombinant proteins**



# *Saccharomyces cerevisiae*

## **Benefits:**

**Grows to high densities, fast growth, ability to secrete proteins**

## **Drawbacks:**

**Low-level expression, hyperglycosylation**

***Kluyveromyces lactis***

**– cheap substrates**

***Pichia pastoris, Hansenula polymorpha***

**Methylotrophic yeasts**

**- methanol – carbon source**

**Promoter and terminator – AOX1**

**Glycosylation – similar to human cells**

**Secretion – sequence - acidic phosphatase**

**$\alpha$ -factor *S. cerevisiae***

# *Pichia pastoris*, *Hansenula polymorpha*

## **Benefits :**

Fast growth on methanol – cheap medium

Fast production (8 h) of protein

High density

Controlled expression via AOX1 promoter (alcohol oxidase I)

Higher expression than *E. coli* (selected proteins)

Better folding than *E. coli* – presence of chaperons

Production of big proteins (>50 kD)

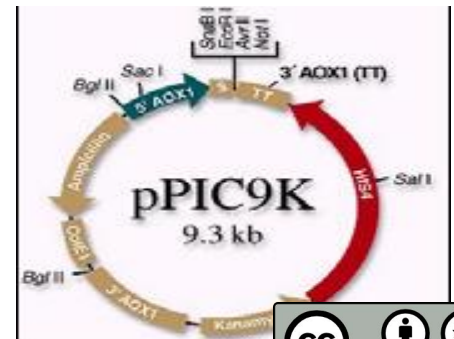
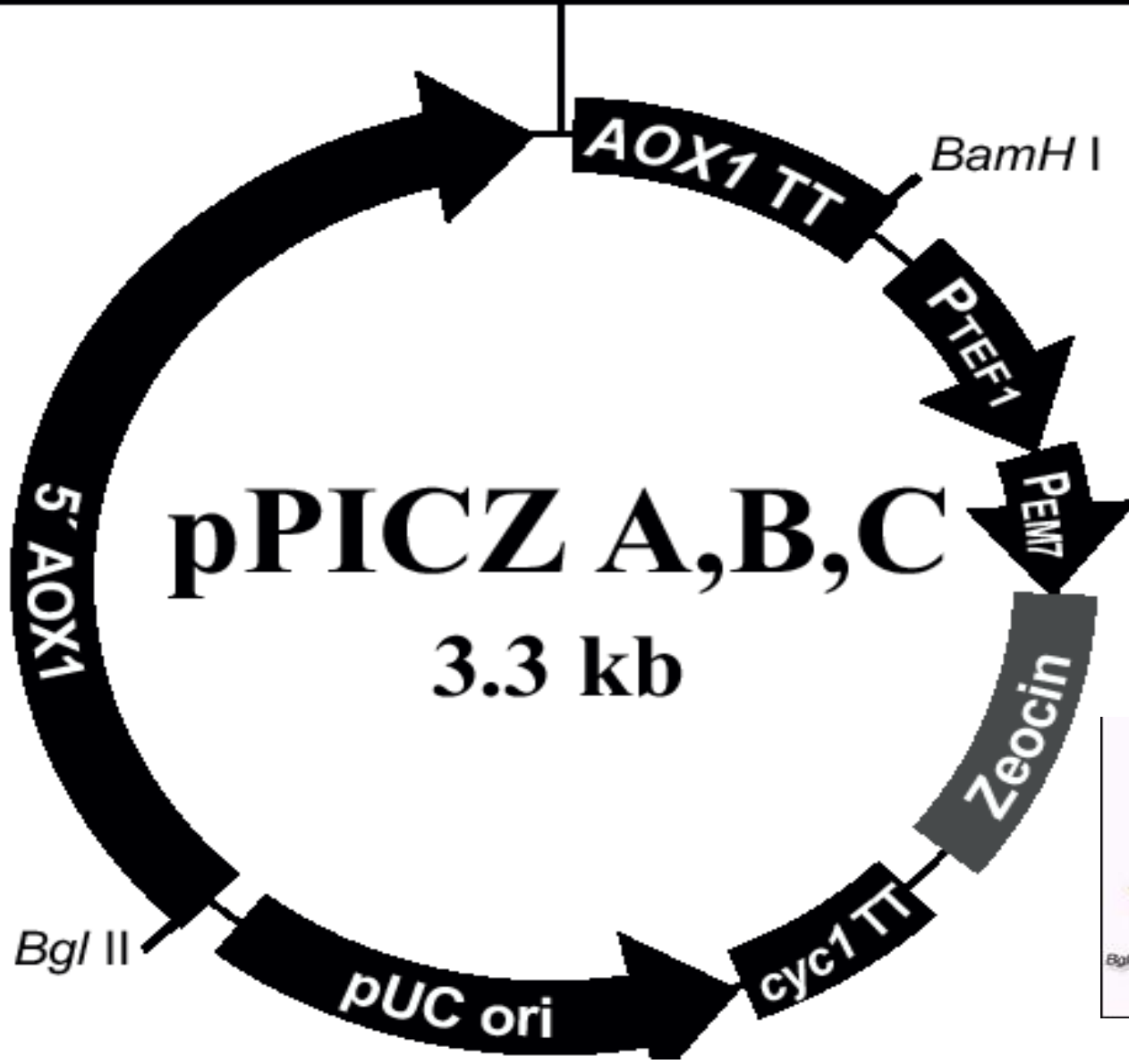
## *Pichia pastoris*

### **Drawbacks:**

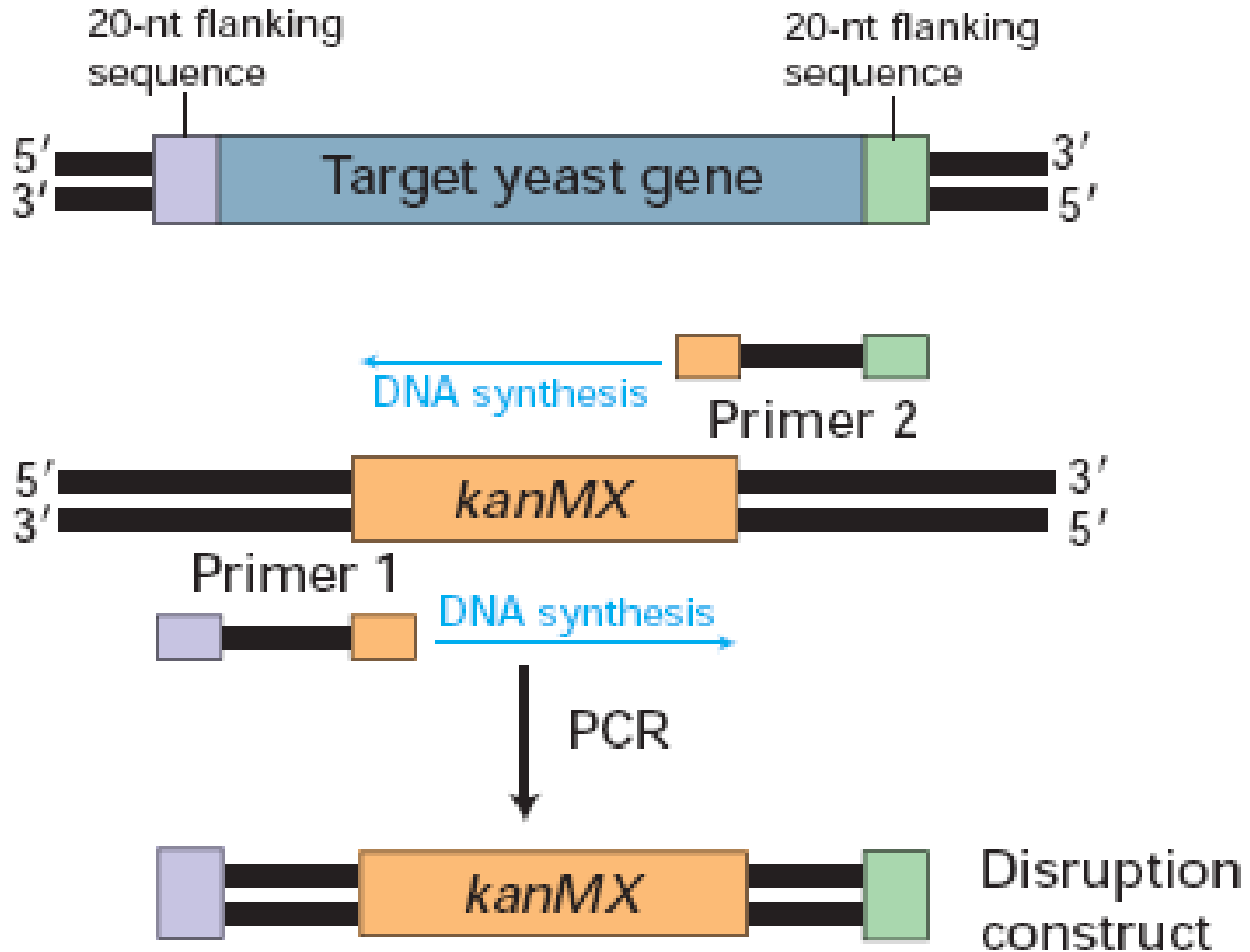
High demand of aeration

methanol is flammable

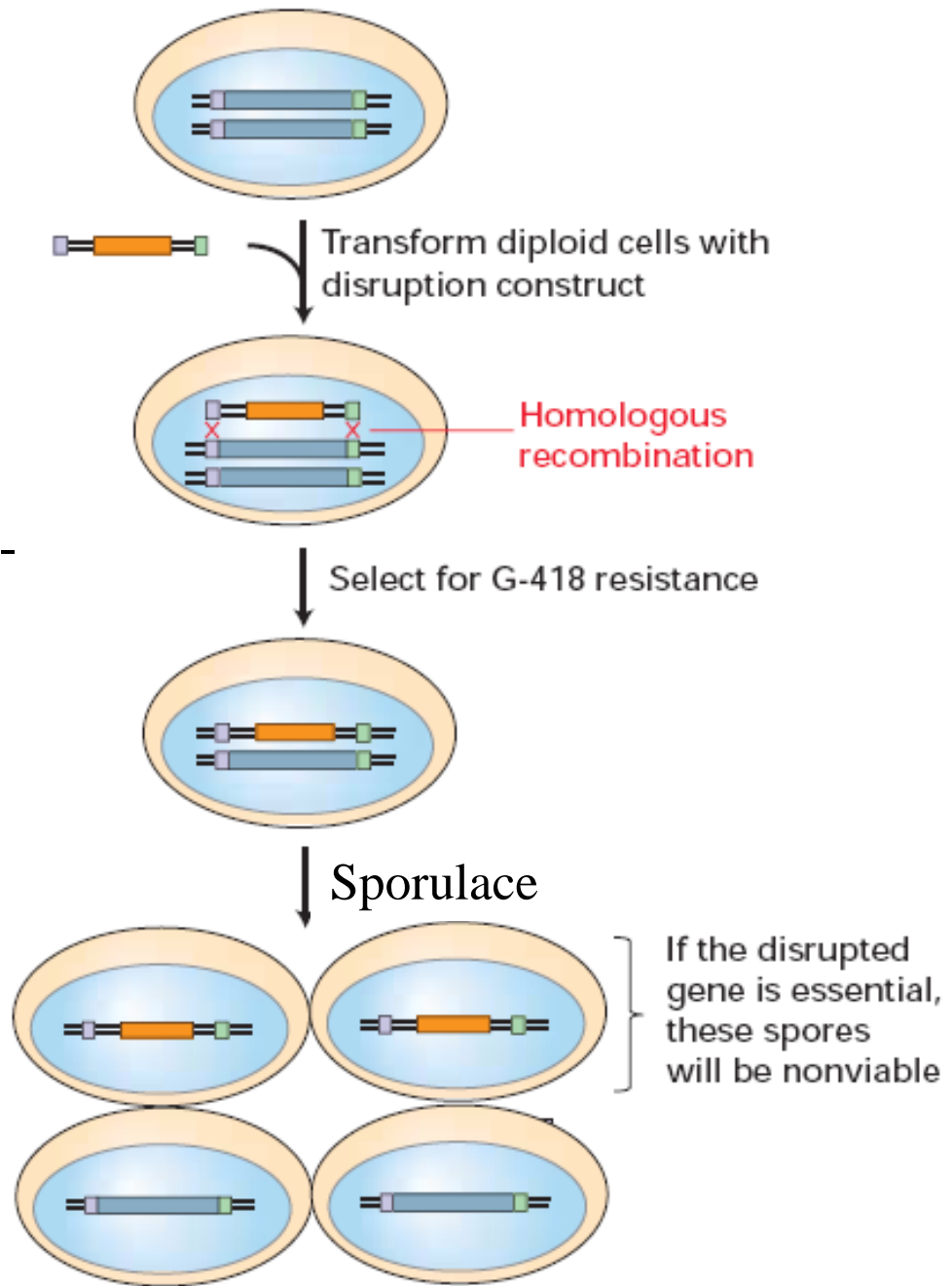
Too high levels of methanol are inhibitory



# Targeted gene disruption



Integration of *kanMX* –  
replacement of target sequence -  
recombinant diploid

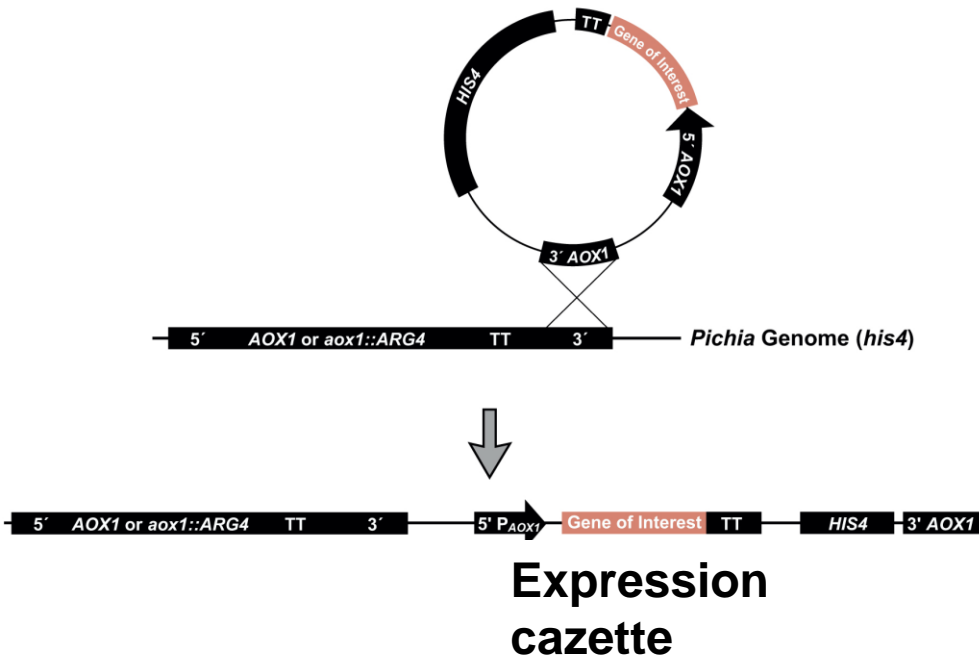


# Transformation and integration

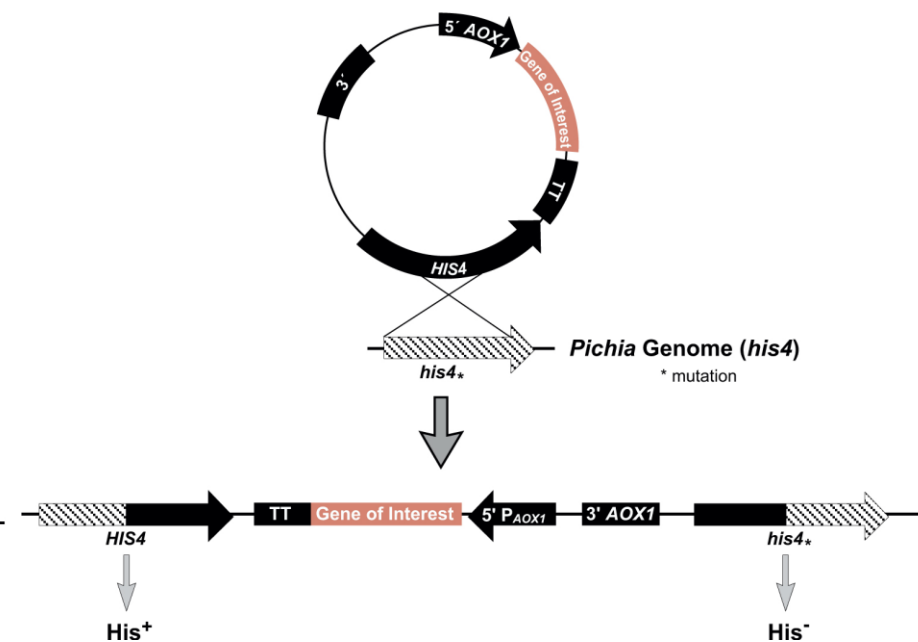
## Stable transformants

- homologous recombination in homologous regions

### Integration in *AOX1* locus



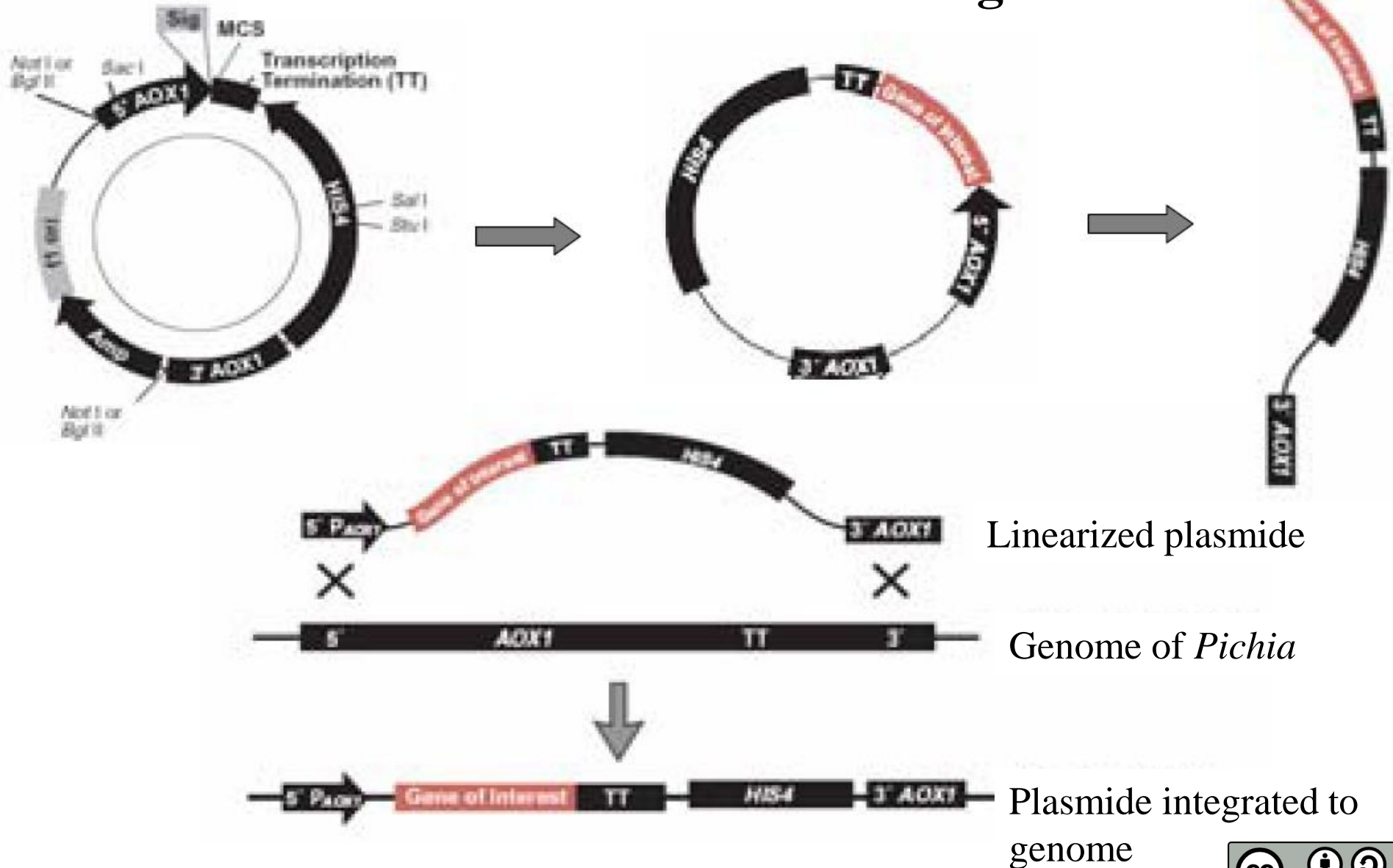
### Integration in *his4* locus

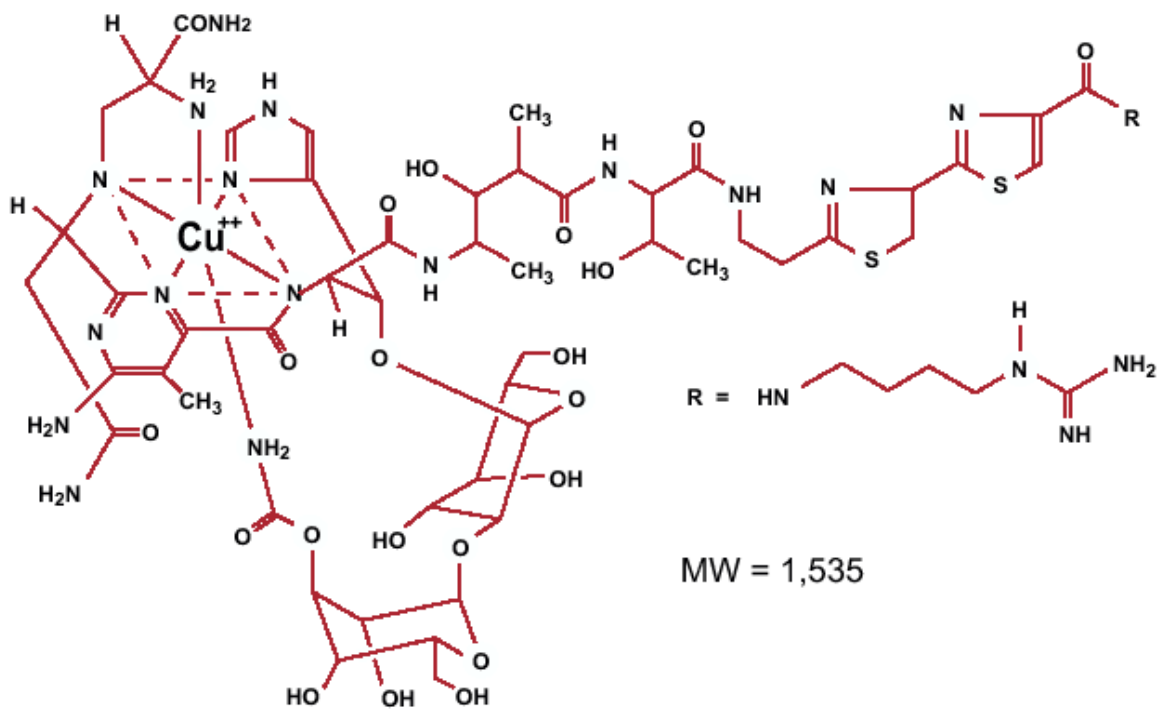




# Pichia

## Linearization of vector for integration

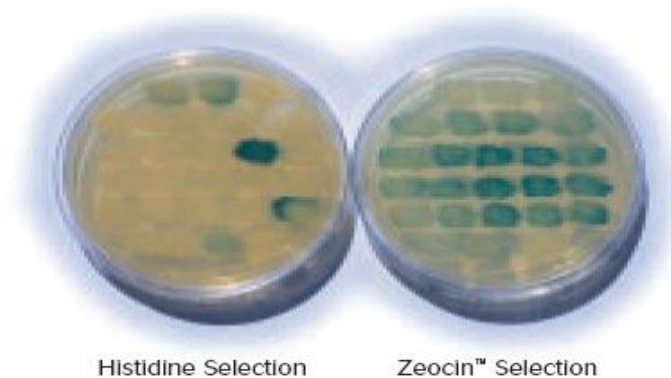




MW = 1,535

**Zeocin**

Figure 3 – Improved selection with Zeocin™



# **Some heterologous proteins - proteolysis**

## **Elimination:**

- Inhibitors**
- Protease mutants**
- pH below 3 or above 6**
- Presence of amino acids**

## **Micromycetes:**

*Aspergillus nidulans* and *Trichoderma reesei*

Higher ability to secrete proteins than *S. cerevisiae*.

Growth in fibers – troublesome culture scale-up

# **Tissue cell cultures**

- defined environment – growth factors, carcinogens**
- specific cell line (even genetically modified)**
- defined growth parameters**

**Posttranscription and posttranslation modifications**

**Model system – mammalian, human cells**

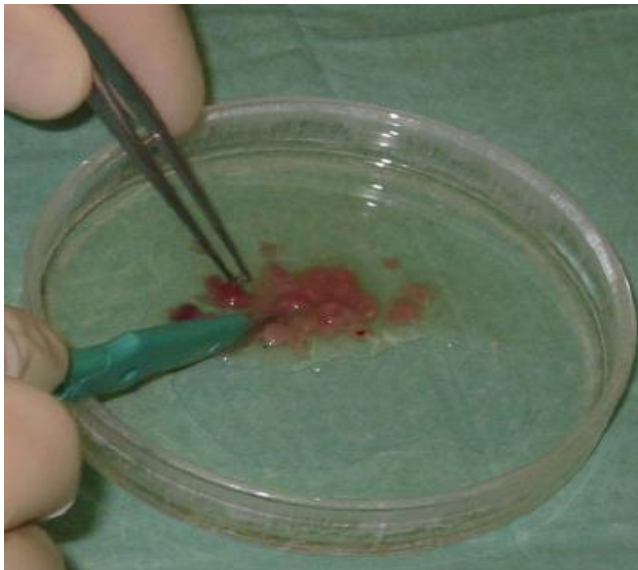
**– localization, regulation, interaction**

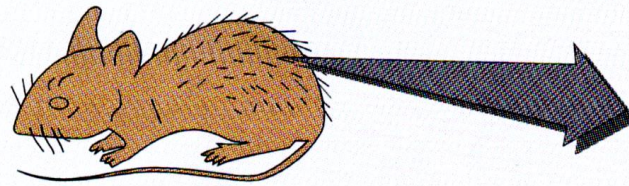
**Cells in growth medium (37°C, serum 5-20%, CO<sub>2</sub> 3-5%)  
inoculation of germ cells (desintegrated biopsy, enzymatic  
digestion) – culture; limited number of divisions**

**Mutants – permanent cell lines – „immortal “**

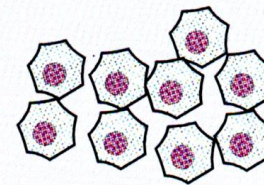
**human cells – permanent only cancer cells**

**rodents – embryonal cells – permanent lines (non-cancer)**

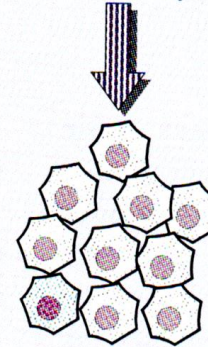




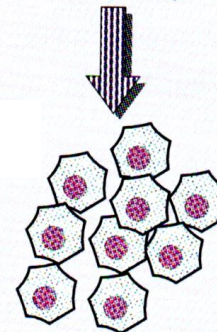
Growth of cells is restricted by properties of individual tissue



**Primary cells**  
Culture divides for several divisions

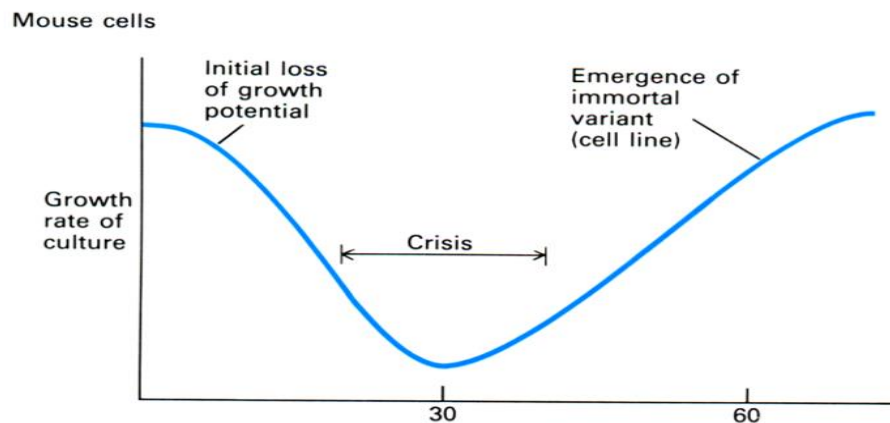
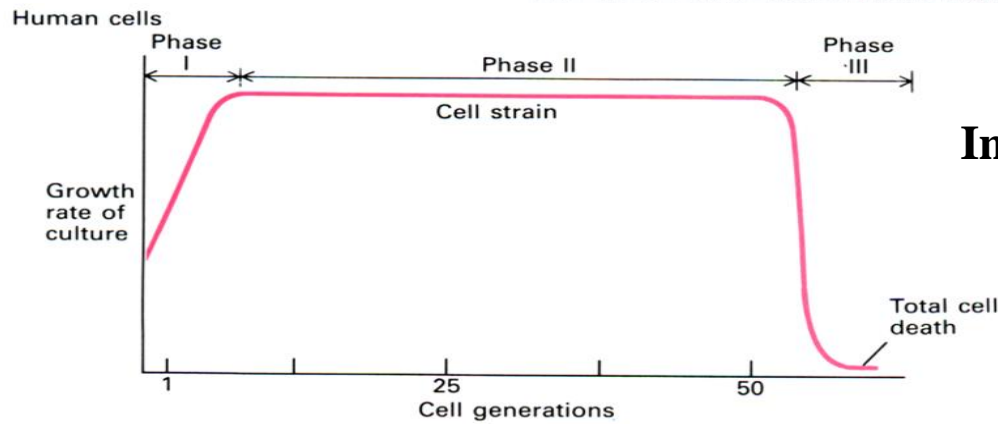


**Crisis**  
Majority of cells die; a few cells grow out

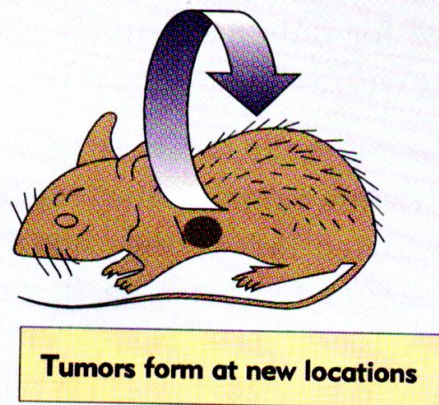
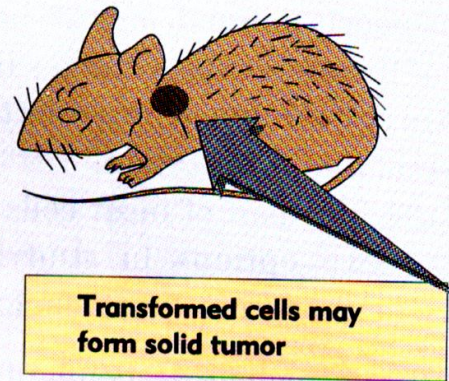


**Established cell line**  
Divides indefinitely, but cells continue to adhere to substratum, require serum, and are inhibited by contact

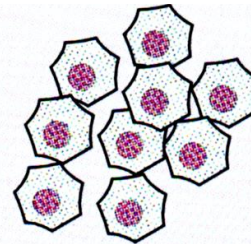
## Imortalizace



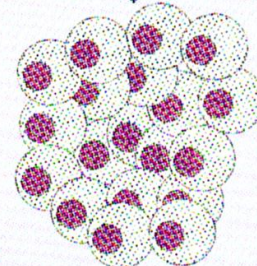




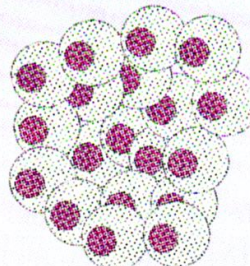
### Immortalization



### Transformation



### Metastasis



### Established cell line

Divides indefinitely, but cells continue to adhere to substratum, require serum, and are inhibited by contact

### Transformed cells

Independent of anchorage, serum, contact inhibition; change shape, round up, and grow into a focus

### Fully tumorigenic

Cells become mobile, and can migrate to start new colonies



# **Transformed cells**

**Make tumors in host**

**Immortalized**

**Changed transcription of some genes**

**Secretion of transforming growth factor**

**Lower dependence on growth factors – autocrine**

**Suicide growth – continuous growth even without nutrients**

# **Changes in the growth and surface**

**Loss of actin microfilaments**

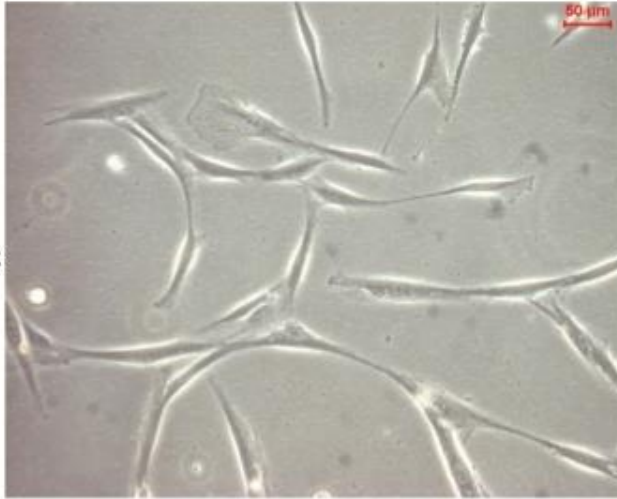
**Changed morphology → round-shaped**

**Loss of the surface fibronectin – lower adherence**

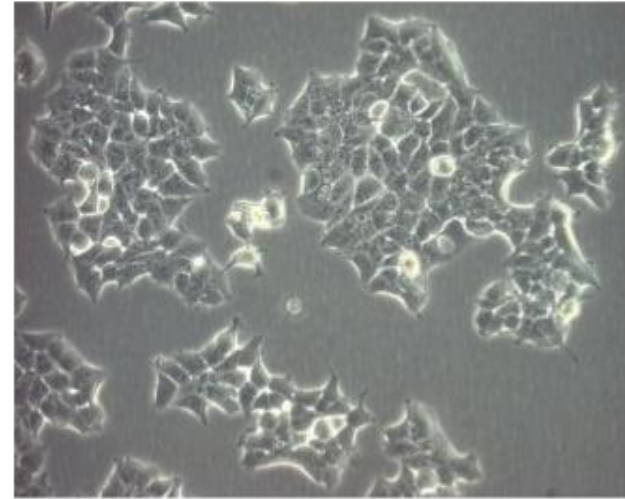
**Lower substrate dependence**

# CELL MORPHOLOGIES VARY DEPENDING ON CELL TYPE

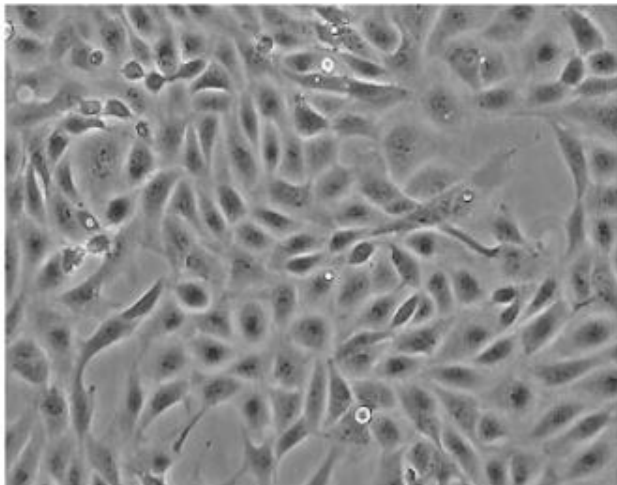
**Fibroblastic**



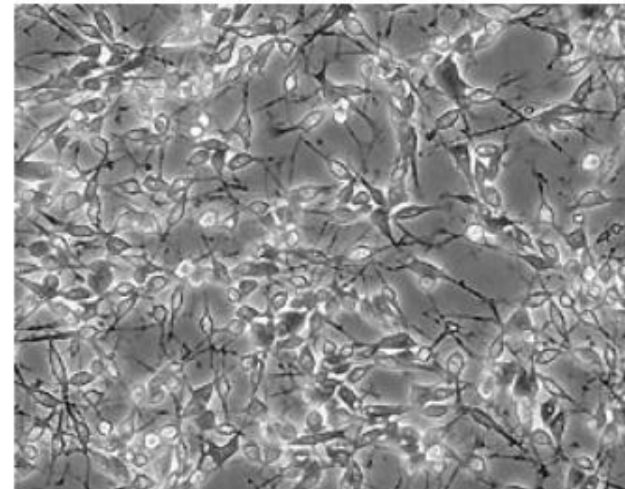
**Epithelial**



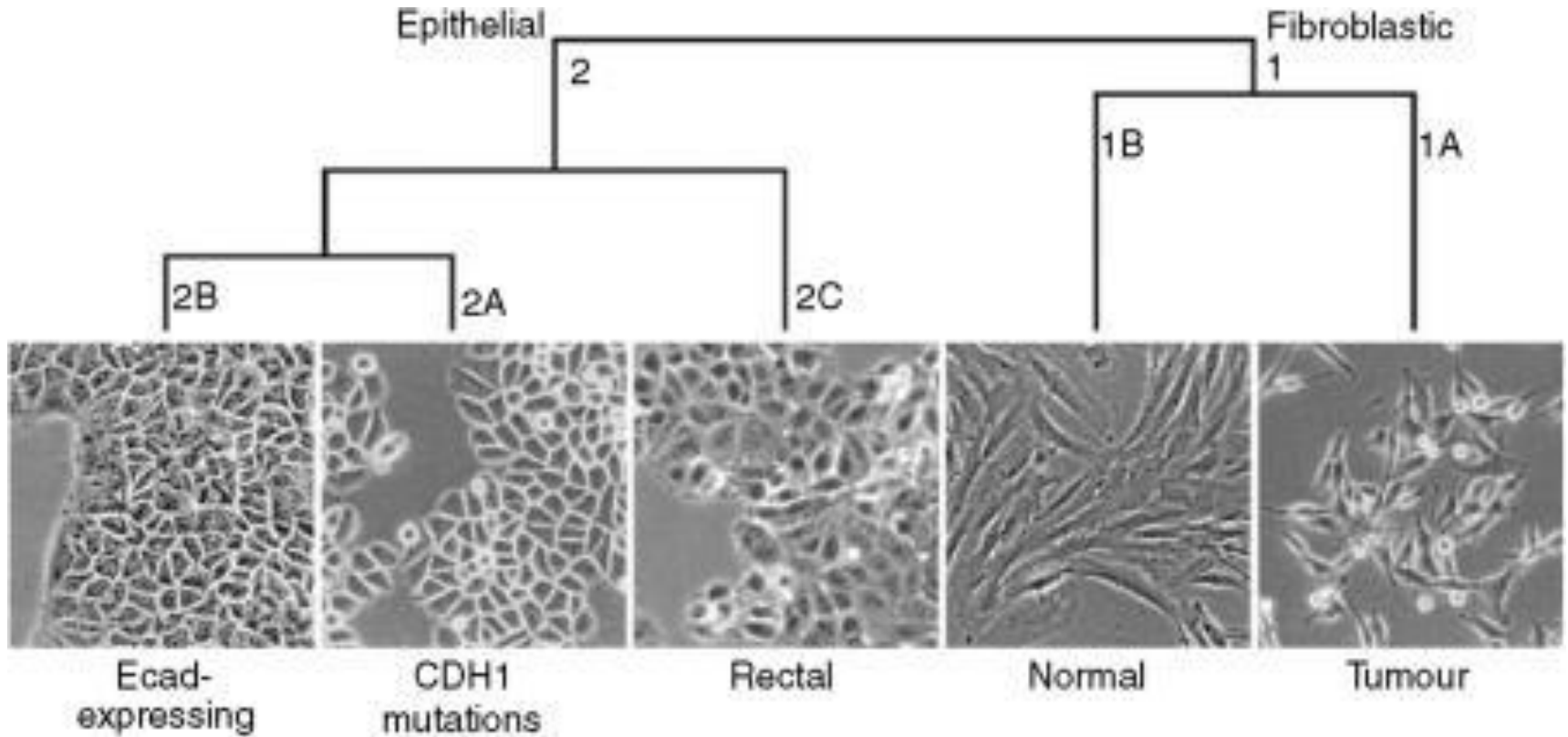
**Endothelial**



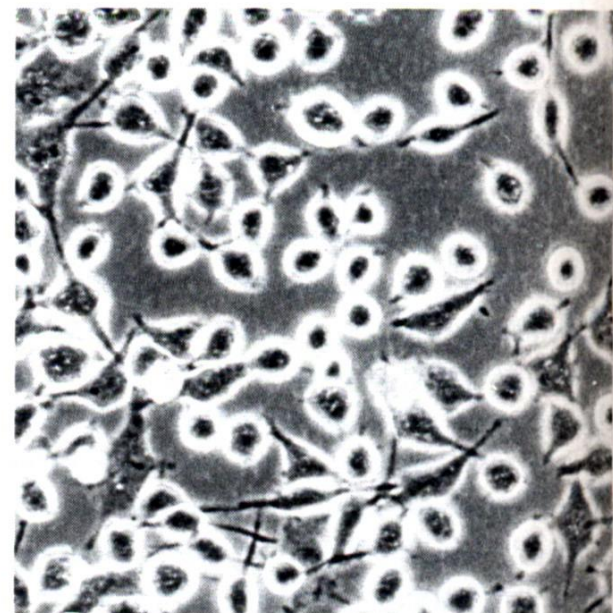
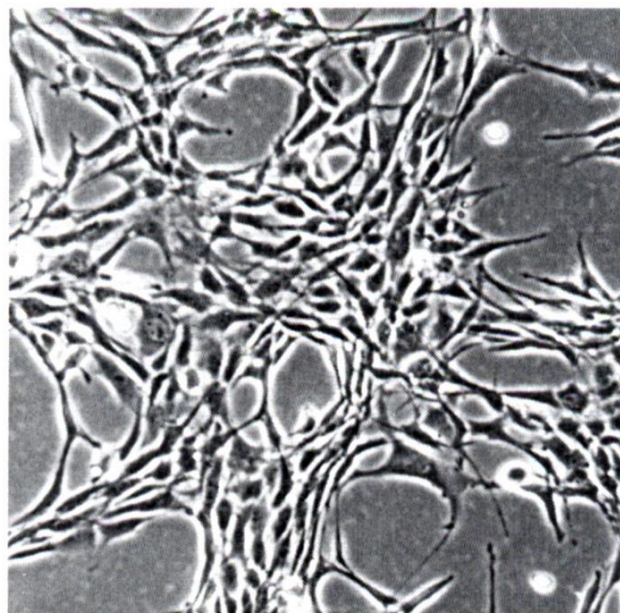
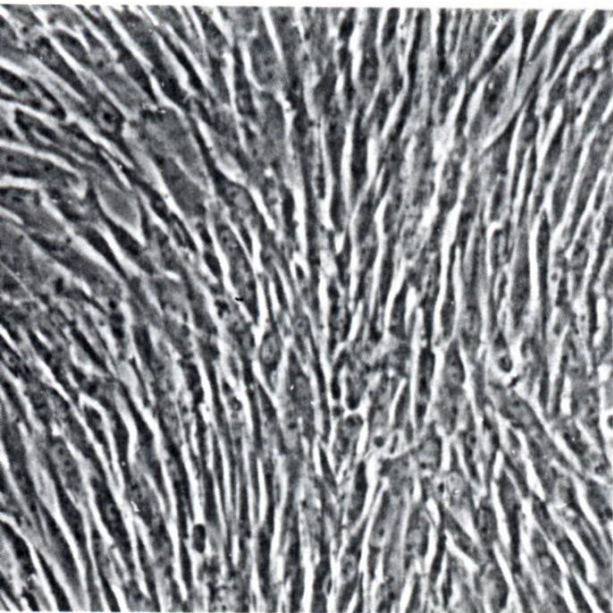
**Neuronal**



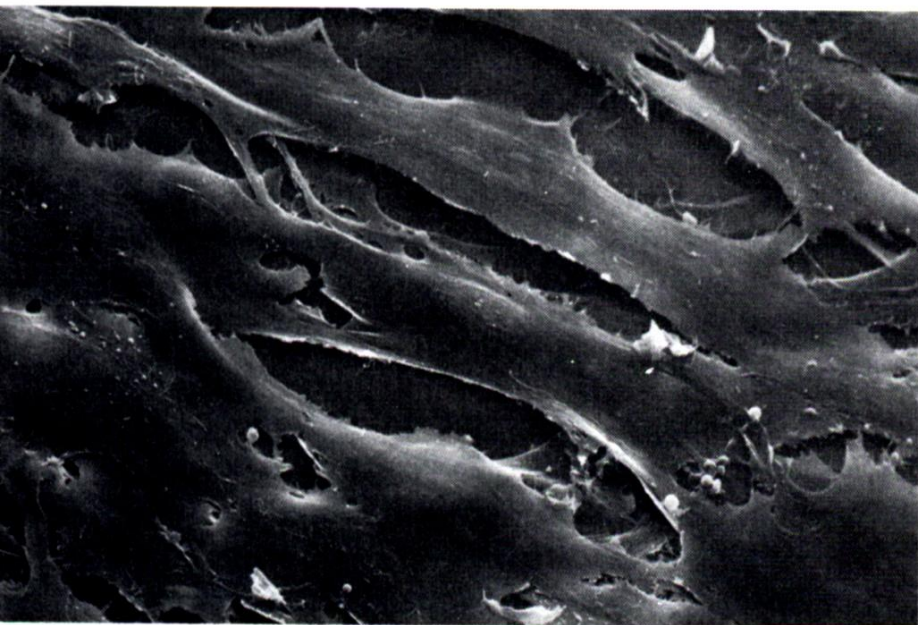
# Morphology of representative cell lines







(d)



(e)



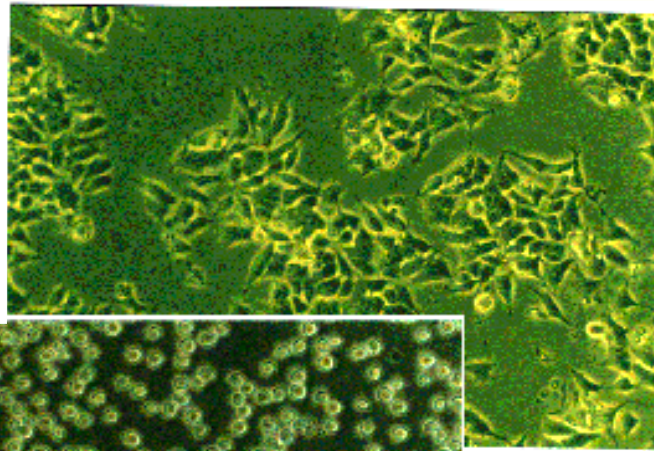


# Cultures a) attachment-dependent x b) -independent

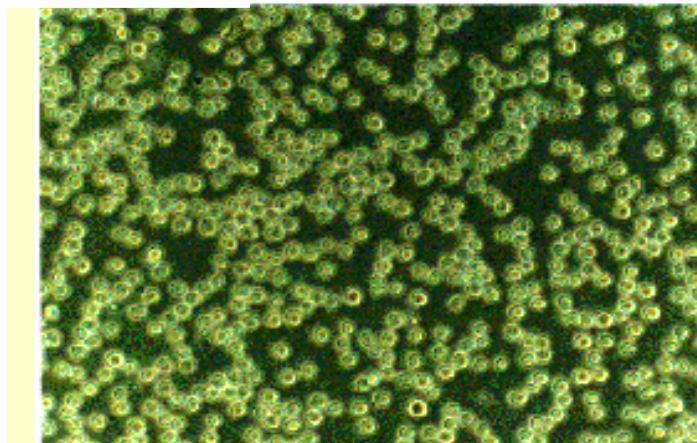
a) Support – proteins mediating attachment  
Cells from tissues

b) Cells from lymphocytes  
easy scale-up: important for industrial application

a)



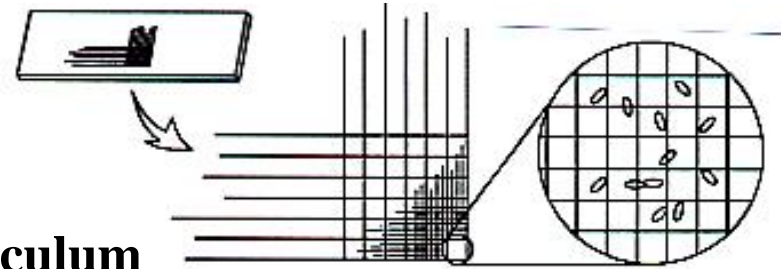
b)



# Cell culture

Important parameter – concentration of inoculum

- Cell interactions



Adherence - fibronectin, laminin, collagen

Passage – trypsinization

Growth demands

Generation time

Liquid nitrogen -  $-196^{\circ}\text{C}$



# Culture dishes

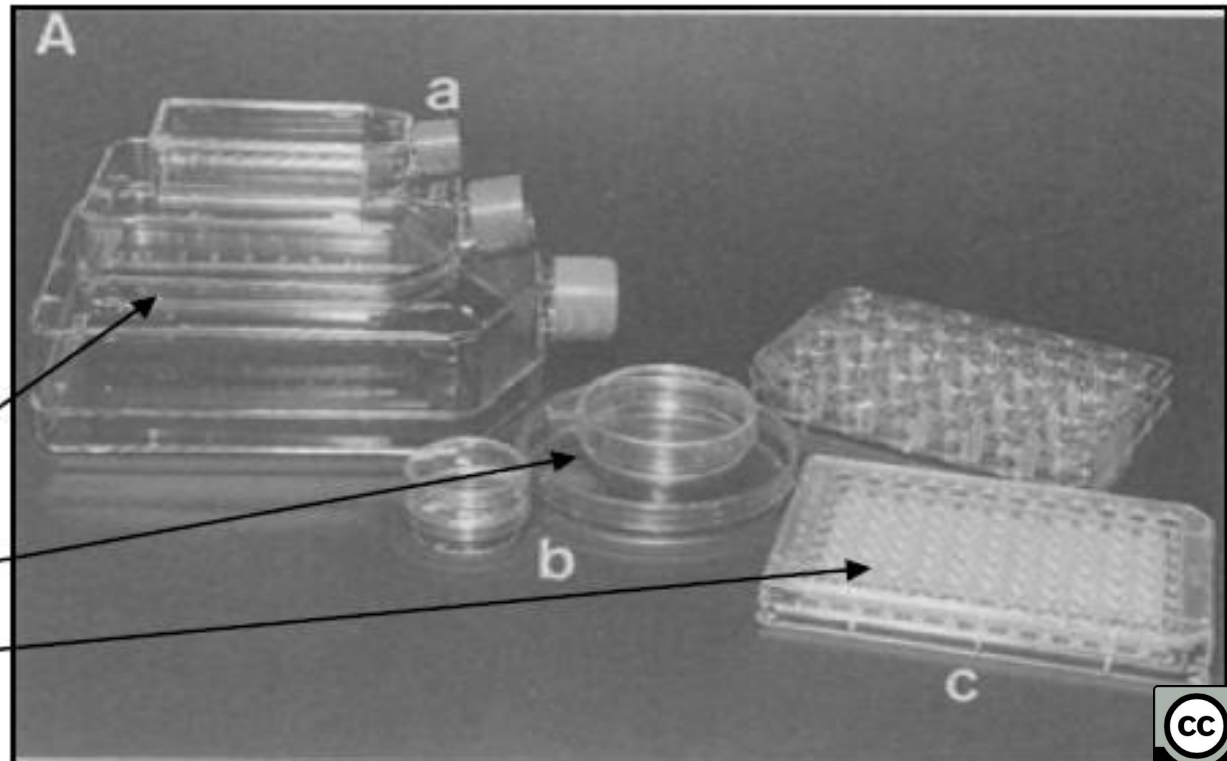


*T-flasks*



*Well-plates*

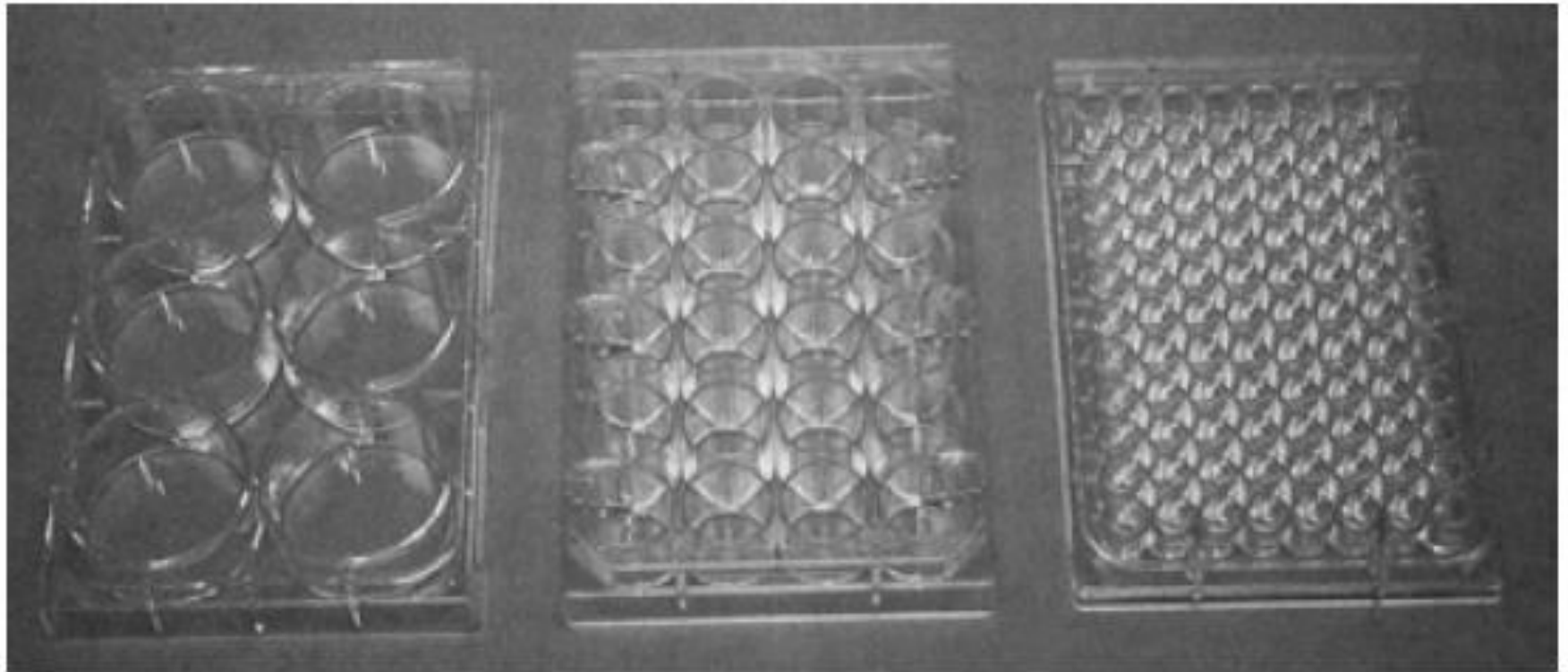
- Flasks
- Dishes
- Multiwell plates





# Different multiwell plates

---



6

24

96

# Complex media RPMI 1640, CMRL 1066, F12

**Basic medium**

**More than 40  
components**

**AA, vitamins,..**

**Incubation at  
constant pH under  
5% CO<sub>2</sub>**

RPMI1640	mg/liter			mg/liter
Amino Acids			Salts	
I-Arginine	200		Ca(NO3)2.4H2O	100
I-Asparagine	50		KCL	400
I-Aspartic acid	20		MgSO4	100
I-Cystine	50		NaCl	6,000
I-Glutamic acid	20		NaHCO3	2,000
I-Glutamine	300		Na2HPO4	1,512(7H2O)
Glutathione, reduced	1		Vitamins	
Glycine	10		p-Aminobenzoic acid	1
I-Histidine	15		Biotin(H)	0.2
I-Hydroxyproline	20		Choline	3(Cl)
I-Isoleucine	50		Cyanocobalamin(B12)	0.005
I-Leucine	50		Folic acid(M)	1
I-Lysine	40(HCl)		Inositol	35
I-Methione	15		Nicotinamide	1
I-Phenylalanine	15		Pantothenic acid	0.25(Ca)
I-Proline	20		Pyroxidine(B6)	1(HCl)
I-Serine	30		Riboflavin(B2,G)	0.2
I-Threonine	20		Thiamine(B1)	1(HCl)
I-Tryptophan	5		Others	
I-Tyrosine	20		Glucose(Dextrose)	2,000
I-Valine	20		Phenolsulfonphthalein	5

# **MEM (Eagle's medium)**

## **Dulbecco's Modified Eagle's medium (DMEM)**

**- essential amino acids, source of energy, vitamins and salts**

**surface 2 – 5 mm**

**Inoculum –  $10^4$  –  $10^5$  cells/ml (1:2 – 1:4)**

**Storage – fresh medium,**

**50% FBS,**

**DMSO – kryoprotective,**

**Serum FBS, FCS**

**- growth factors**



**Harry Eagle**

**Renato Dulbecco**

# **Growth hormones**

**PDGF (platelet-derived growth factor) – principal growth factor of the serum**

**FGF (fibroblast growth factor)**

**EGF (epidermal growth factor)**

**Insulin-like growth factor (IGF-1 a IGF-2)**

**Fibronectin and fetuin (in fetal serum)**

**- Enable the attachment of the cell to surface**

**$\alpha$ 2-macroglobulin - inhibits trypsin**

**Transferrin binds iron forming less toxic and biologically more available form.**

# Suspension cultures

Control of the process, higher yields

Requires adaptation of adherent cells

- culture flask surface
- Depletion of serum (shock or gradual)



*Spinner flasks*



*Teflon bags*

Uveřejněné materiály jsou určeny studentům Vysoké školy chemicko-technologické v Praze

jako studijní materiál. Některá textová i obrazová data v nich obsažená jsou převzata

z veřejných zdrojů. V případě nedostatečných citací nebylo cílem autora/ů záměrně poškodit event. autora/y původního díla.

S eventuálními výhradami se prosím obračejte na autora/y konkrétního výukového materiálu, aby

bylo možné zjednat nápravu.

The published materials are intended for students of the University of Chemistry and Technology, Prague as a study material. Some text and image data contained therein are taken from public sources. In the case of insufficient quotations, the author's intention was not to intentionally infringe the possible author(s) rights to the original work.

If you have any reservations, please contact the author(s) of the specific teaching material in order to remedy the situation.

# Gene expression II

## Biotechnology



EUROPEAN UNION  
European Structural and Investing Funds  
Operational Programme Research,  
Development and Education



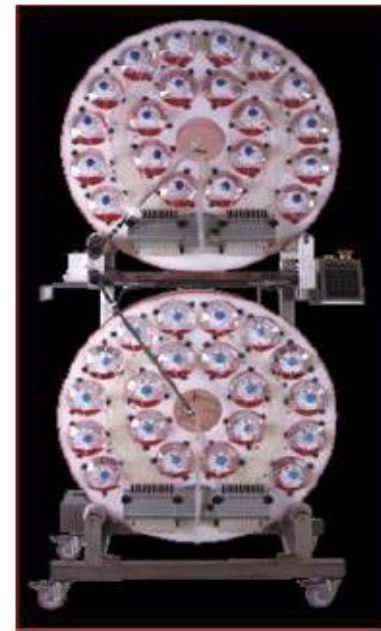
MINISTRY OF EDUCATION,  
YOUTH AND SPORTS

# Culturing in bigger volumes





**Fermentor**



*Roller bottles*



# Synthetic media

## Benefits -

stable defined composition

elimination of growth inhibitors

Elimination of potential risk of contamination with viruses

## Mycoplasmas - problem with contamination



# **Synthetic media drawbacks**

**Lack of unidentified components of serum,**

**Lower number of growth passages.**

# **pH control**

**Optimum - pH 7,4**

**below 7,0 - 6,5 retarded growth**

## **Phenol red**

**yellow                      pH 6,5**

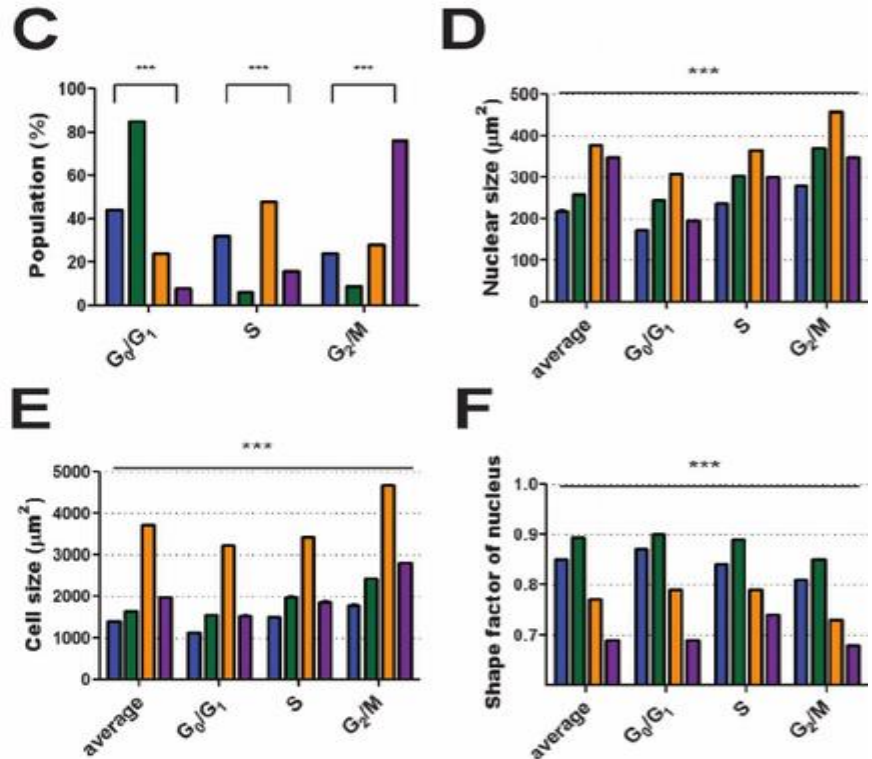
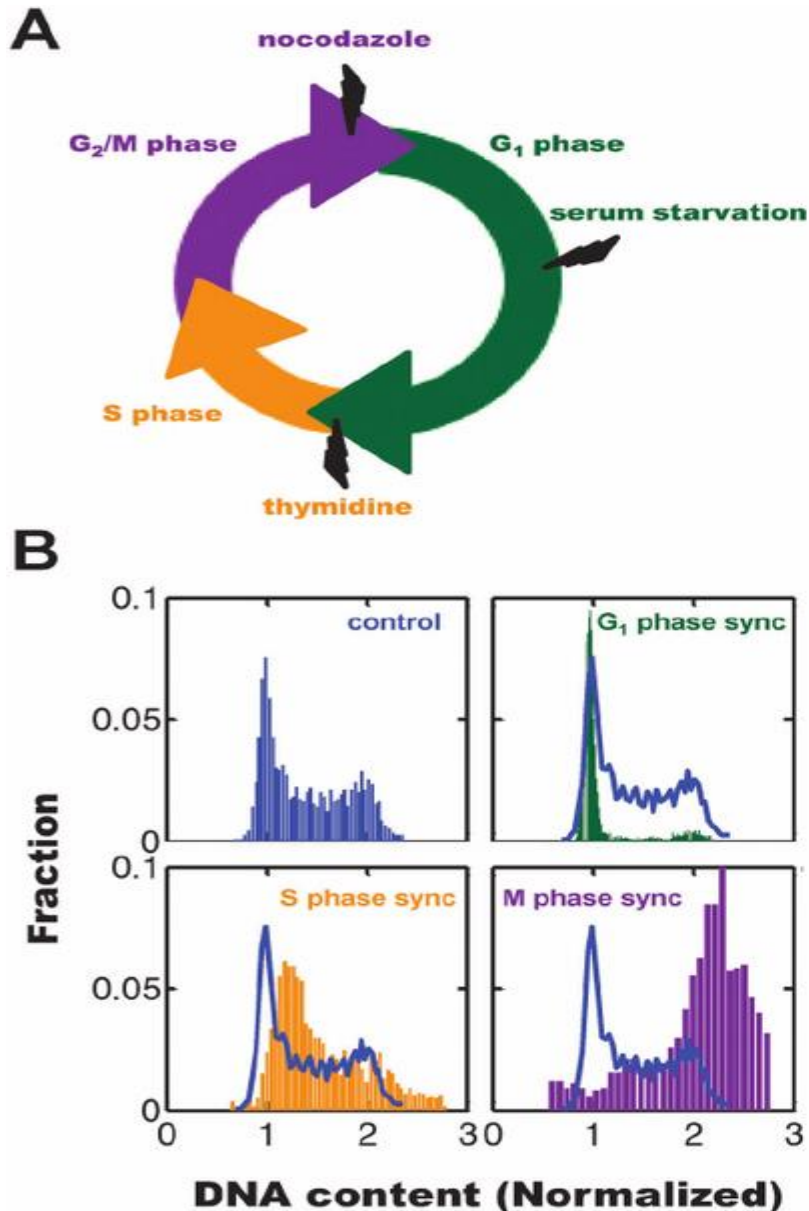
**orange                      pH 7,0**

**red                              pH 7,4**

**lila                             pH 7,6**

**purple                      pH 7,8**

# Synchronization of cell culture



thymidine or aphidicoline - early S phase block

nocodazole (depolymerization of microtubules) – block in G<sub>2</sub> / M

lovastatin or nutrient depletion - block in G<sub>1</sub>

Fluorescence-activated cell sorting - FACS

# TRANSFECTION

**Precipitation with  $\text{Ca}_3(\text{PO}_4)_2$  – easy to use**

**DNA in  $\text{CaCl}_2$  +  $\text{Na}_3\text{PO}_4$  buffered with HEPES**

*HEPES* (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)

**Colloid precipitate of DNA - cell adhesion and uptake**

**Higher efficiency with DMSO or glycerole**

**Very sensitive to pH - pH 7,05**

- Some types of cells too fragile
- Various efficiency of transformation
- The amount of DNA entering the cells is not tunable

## **DEAE Dextran** - cationic polysaccharide – DNA binding.

Complex - DNA/polysaccharide

– taken up by cells through unknown mechanism

Not universal for all the cell types

Peptides

poly(lysine)

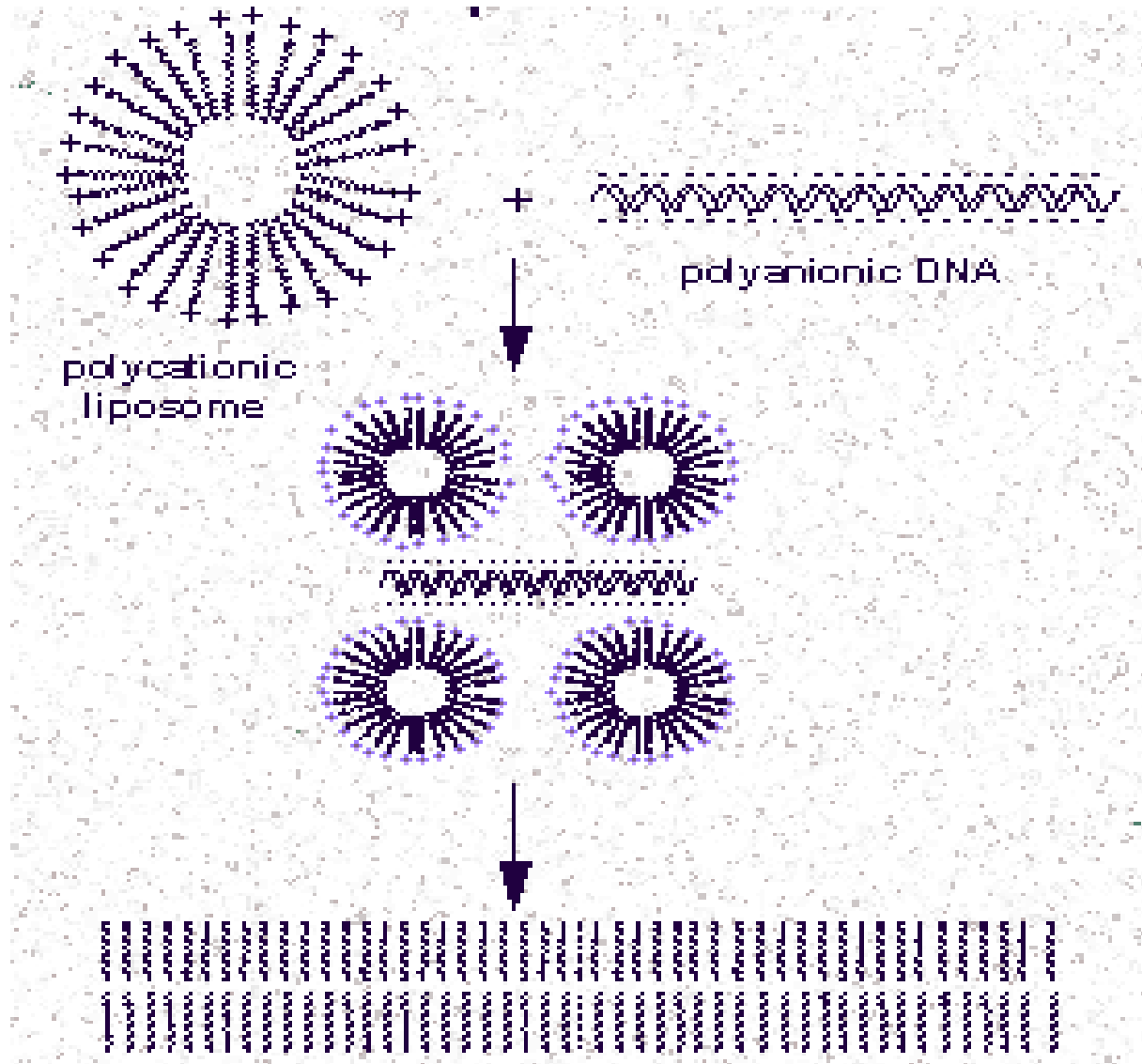
Some peptides – targeting through peptide sequence

Polyethylenimine – very cheap



# Lipofection – mixture of cationic lipid with DNA (neutral)

N-[1-(2,3-dioleyloxy)propyl]-N,N,N-trimethylammonium chlorid (DOTMA) a dioleylylphosphatidyl ethanolamin (DOPE)

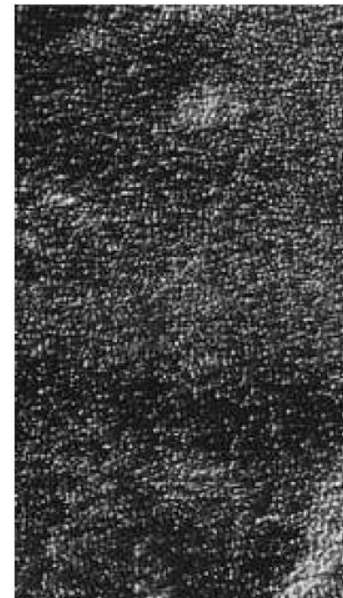
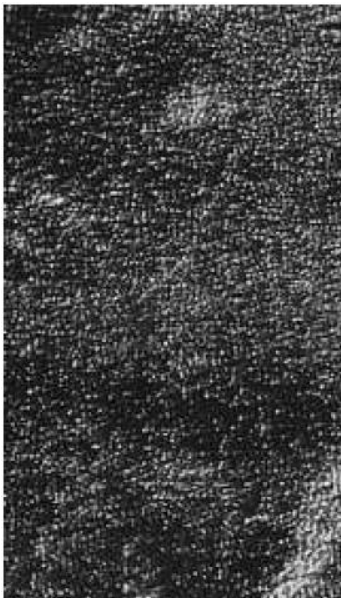




# ELECTROPORATION

**Exposition of DNA and cells to high-voltage puls**

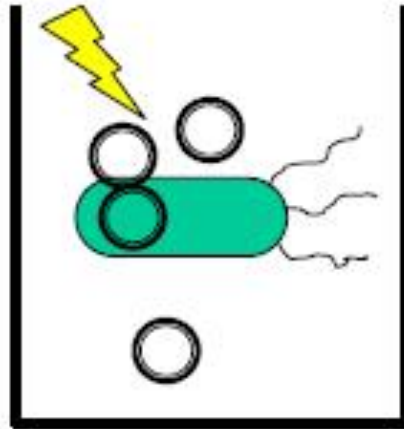
- relatively expensive equipment and process (special cuvettes)
- All cell types
- 1 500 – 2 500 V at 25 mF.
- highly efficient, possible control „copy number“ by change of conditions



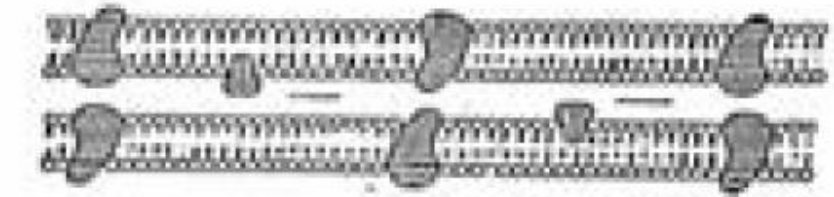
# Electroporator



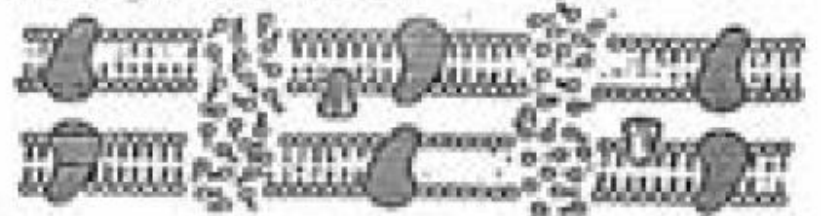
25 microfarads = 2500 V  
@ 200 ohms for 5 ms



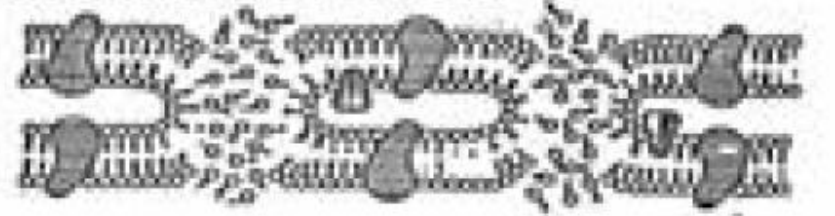
-Pores in membranes, DNA uptake and pore closing



Voltage Stimulated

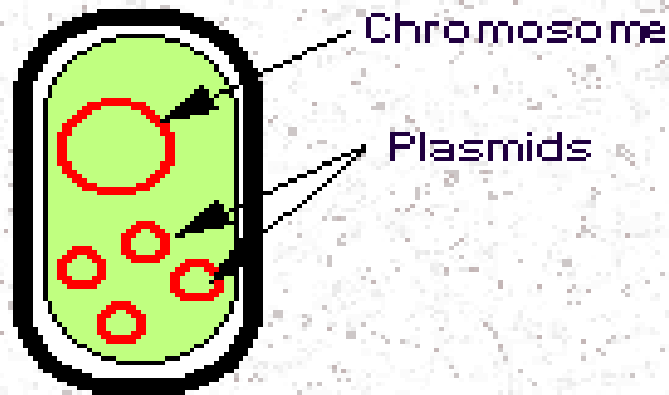


Reconstruction Process

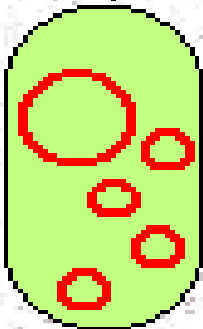


# Protoplast fusion

## Bacterial Cell

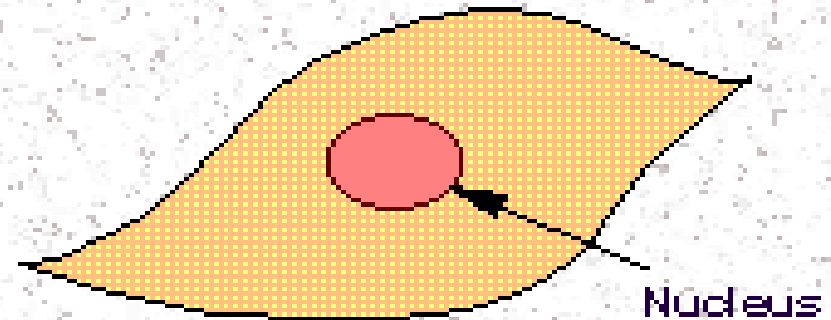


Cell wall enzymatically removed

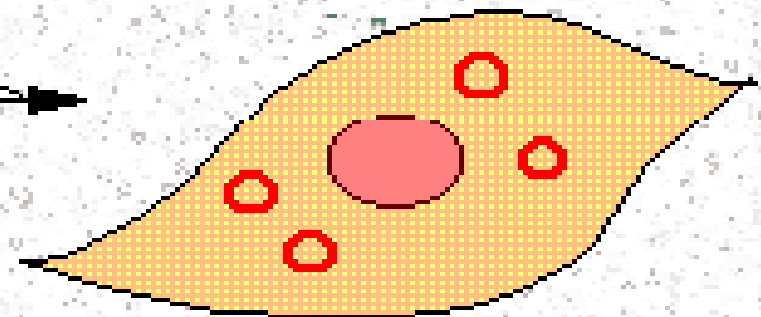


## Protoplast

## Mammalian Cell



PEG fusion,  
add antibiotics



Problems: bacterial contamination, irreproducible

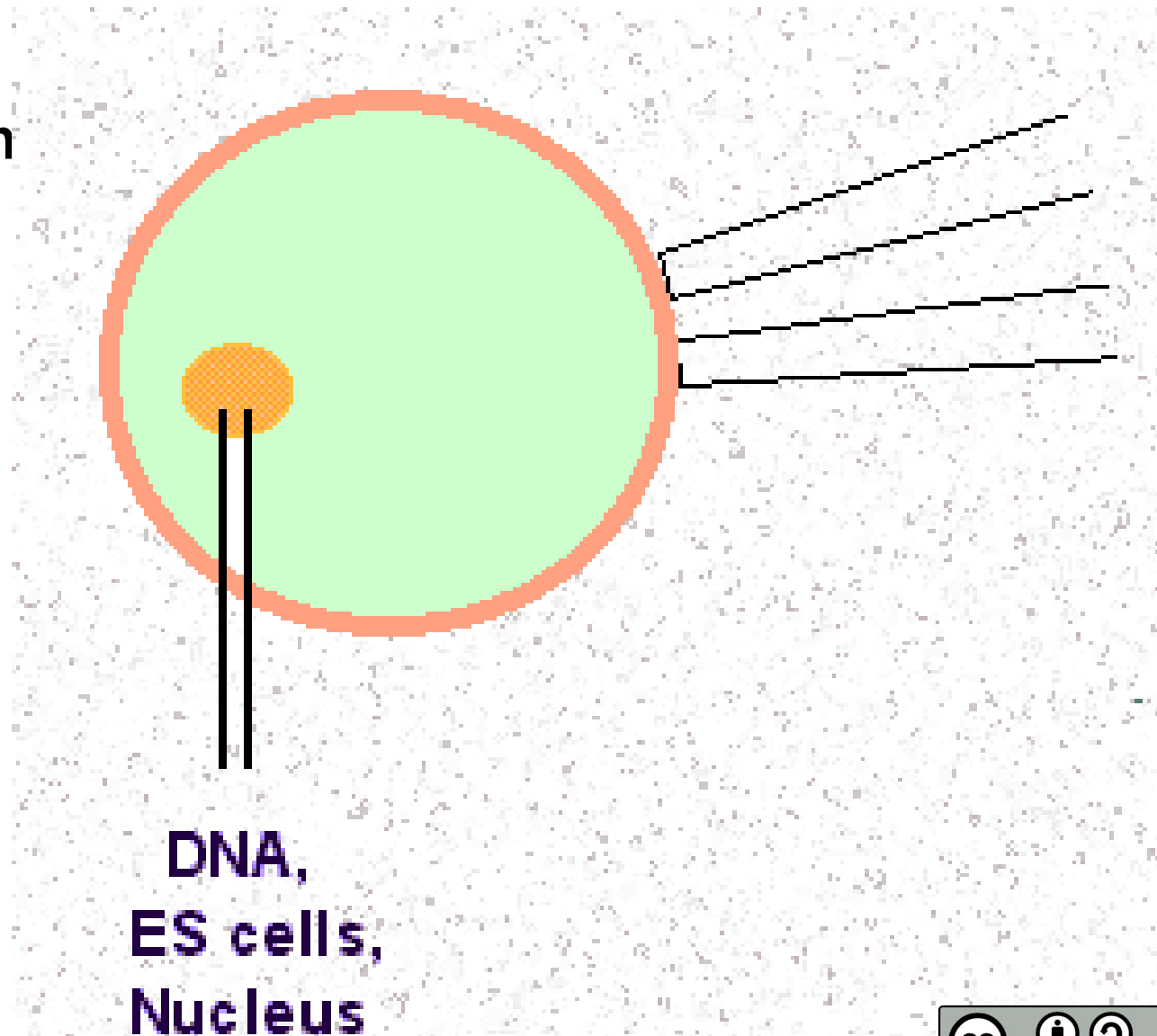
# MICROINJECTION

Controlled „copy number“

Whole nuclei

Special equipment

Trained staff





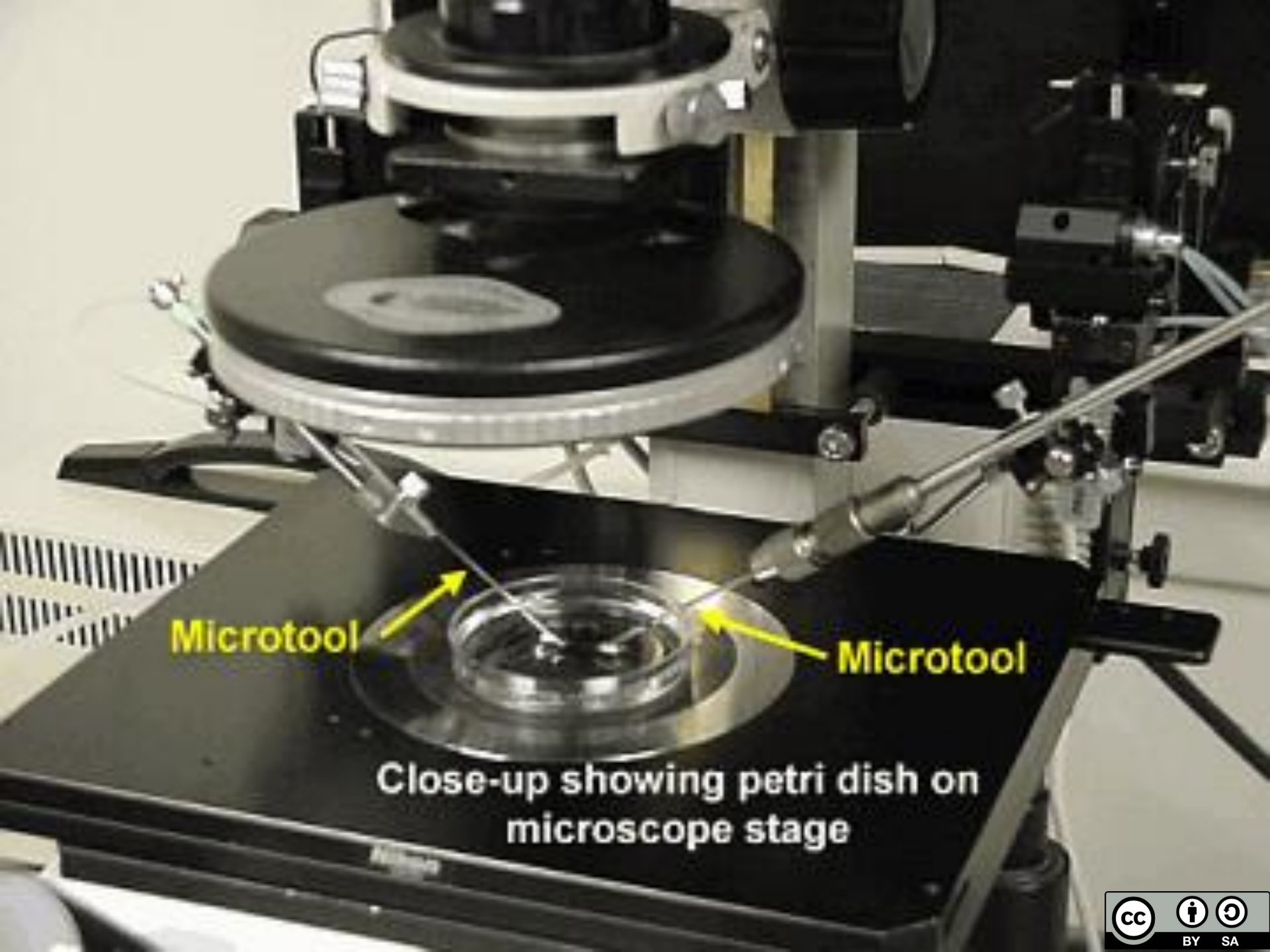
Micromanipulator Joystick

Microtools

Micromanipulator Joystick

Syringe system

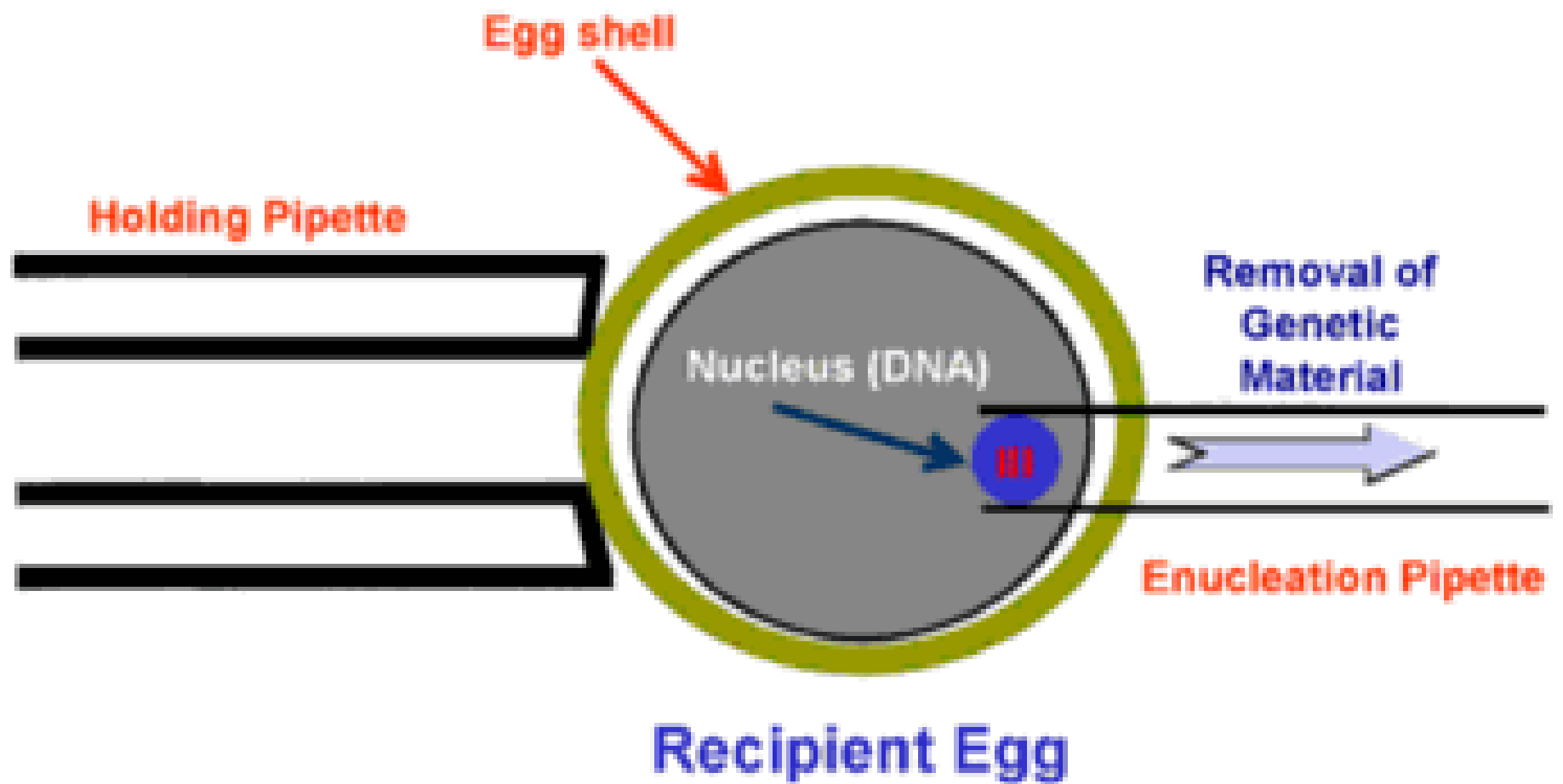
Micromanipulator Station

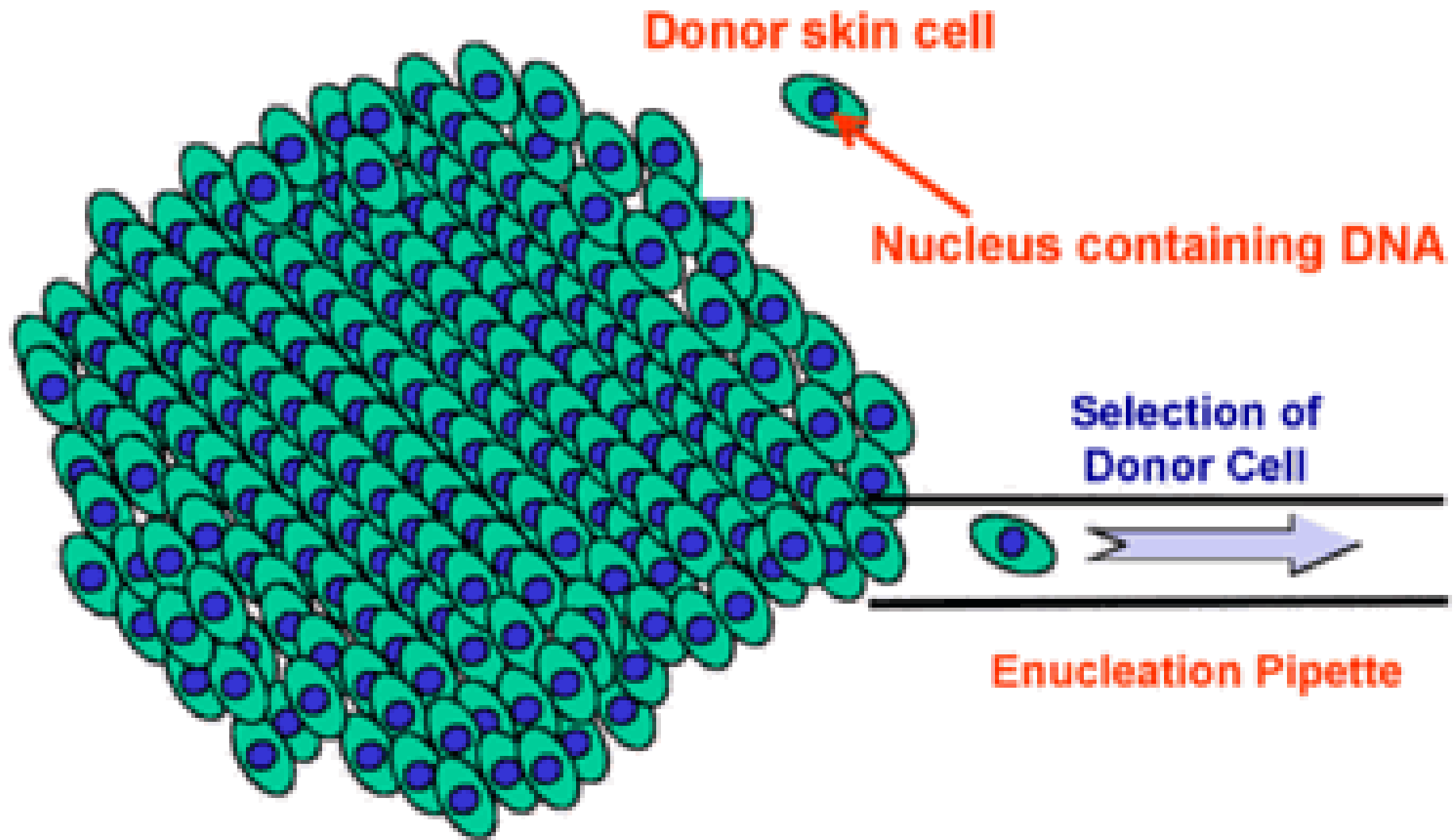


**Microtool**

**Microtool**

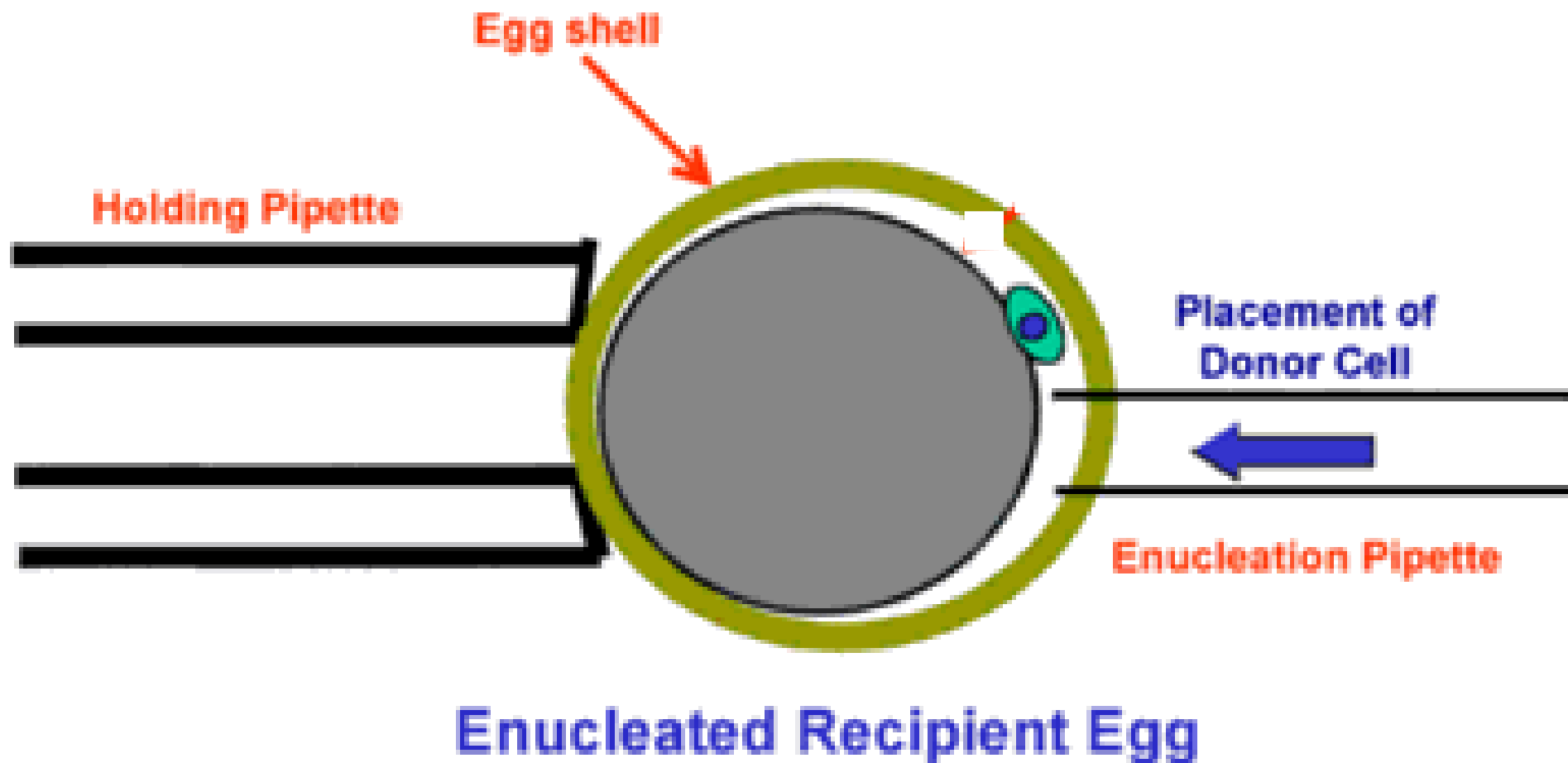
**Close-up showing petri dish on  
microscope stage**





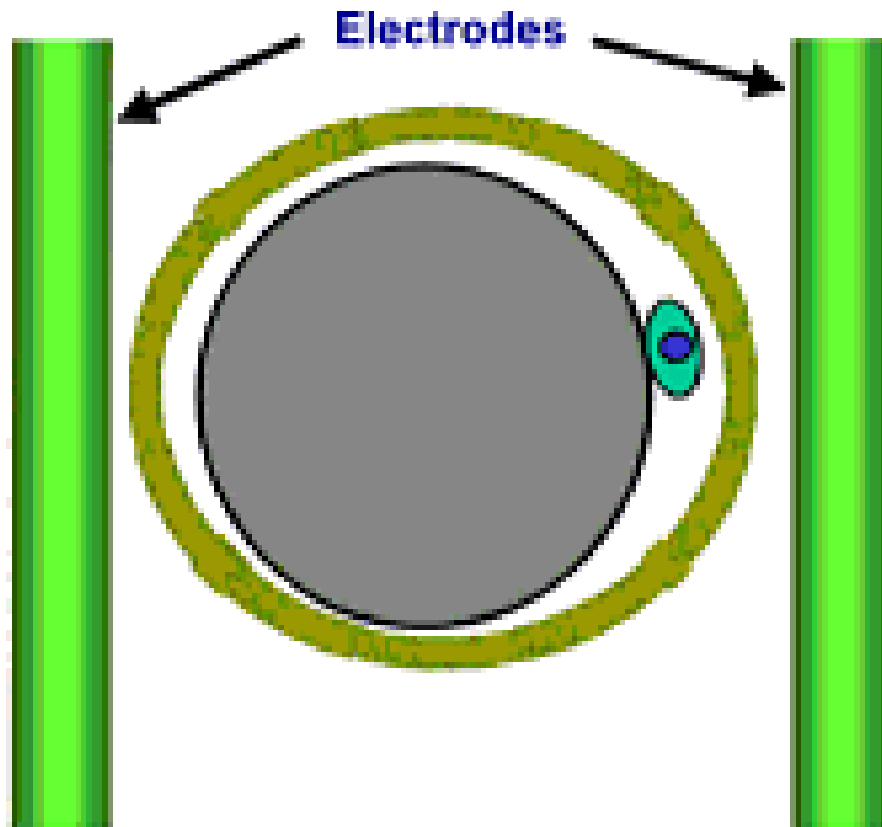
Donor skin cells growing  
in petri dish





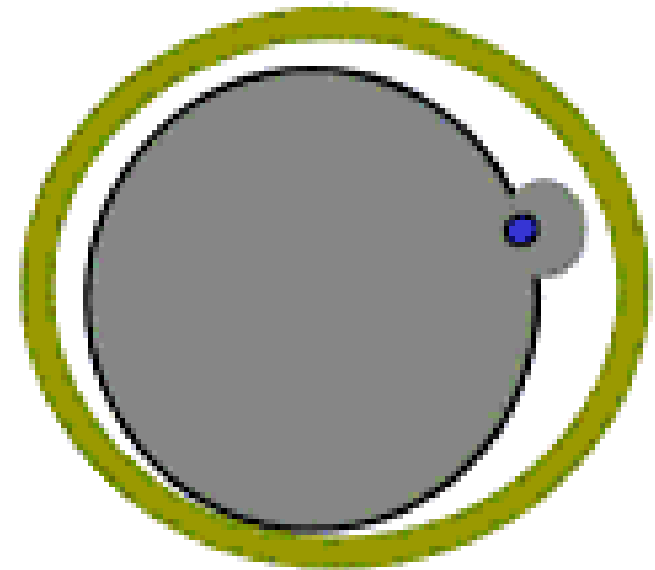


## Fusion Step 1



Enucleated recipient  
egg with donor skin  
cell ready for fusion

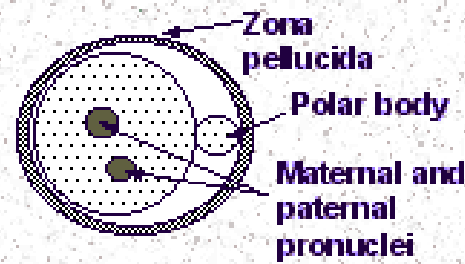
## Fusion Step 2



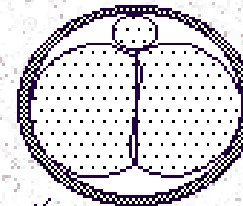
Membranes of enucleated  
egg and donor cell fused  
together

# Egg Development

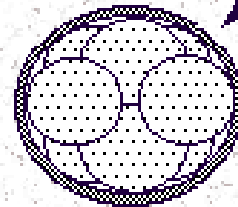
Fertilised egg.



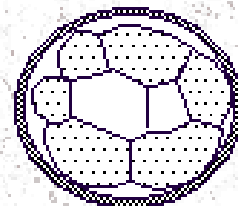
2 cell



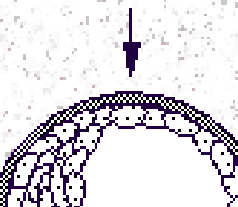
4 cell



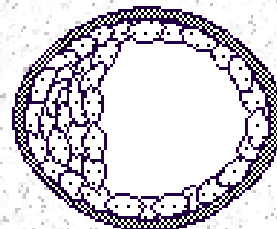
Morula (8 cell)



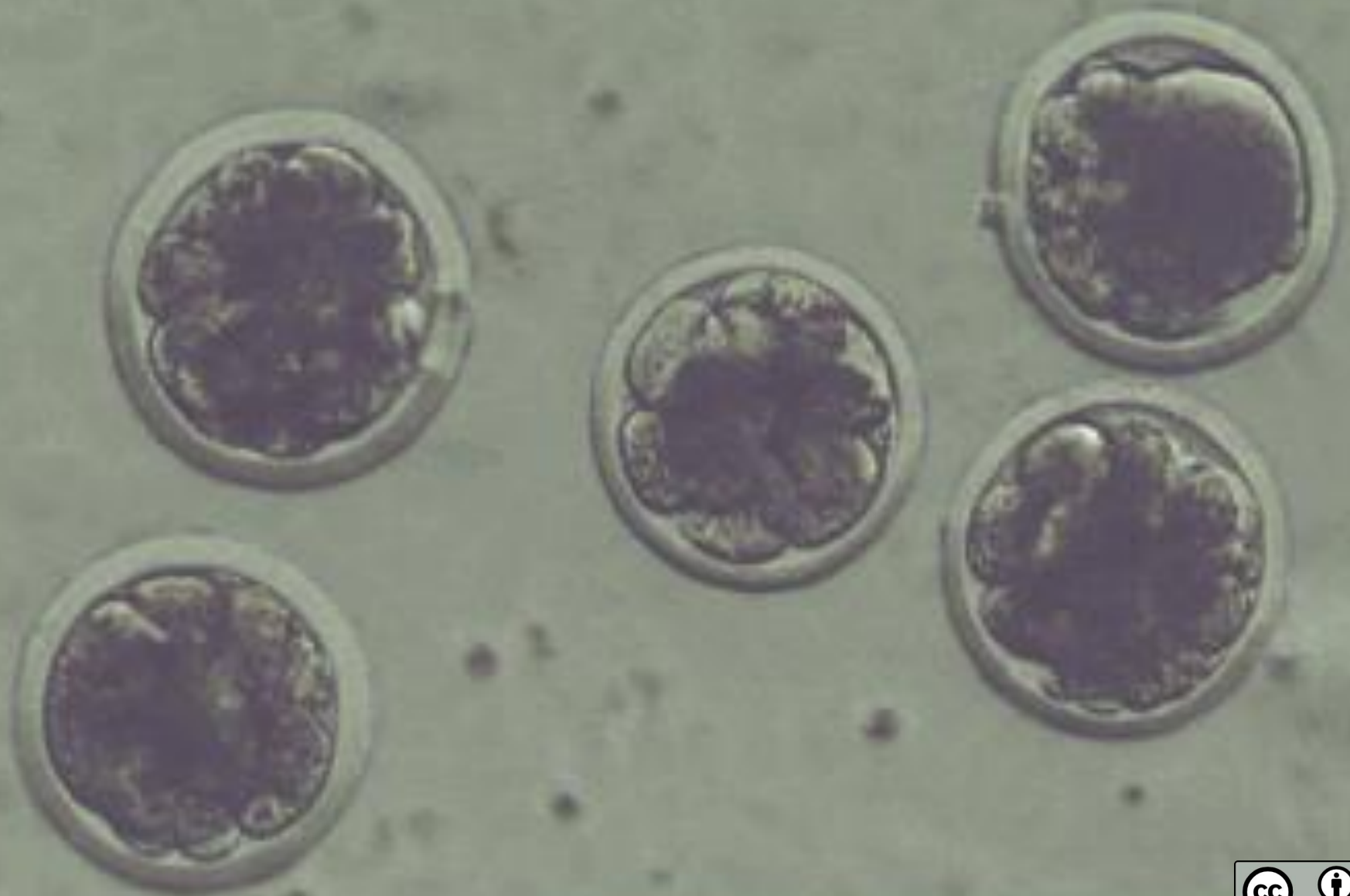
16 cell



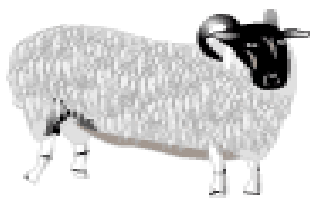
Blastocyst



# 5 cloned goat embryos ready for transfer



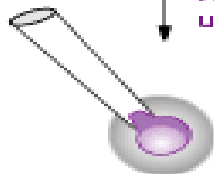
Scottish Blackface ewe



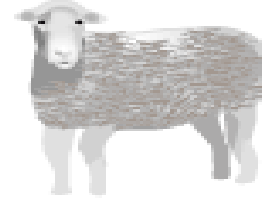
Isolate egg cells  
from ovaries



Remove haploid nucleus  
using micromanipulation



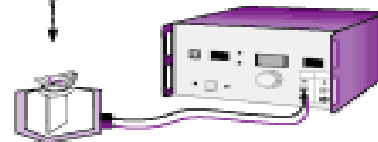
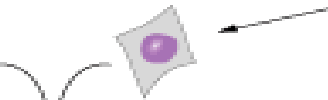
Finn Dorset ewe



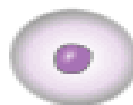
Isolate diploid somatic  
cells from mammary gland



Induce quiescence ( $G_0$ )  
by growth in low serum



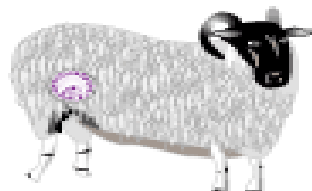
Fuse cell membranes  
by electroporation



6 days  
in culture



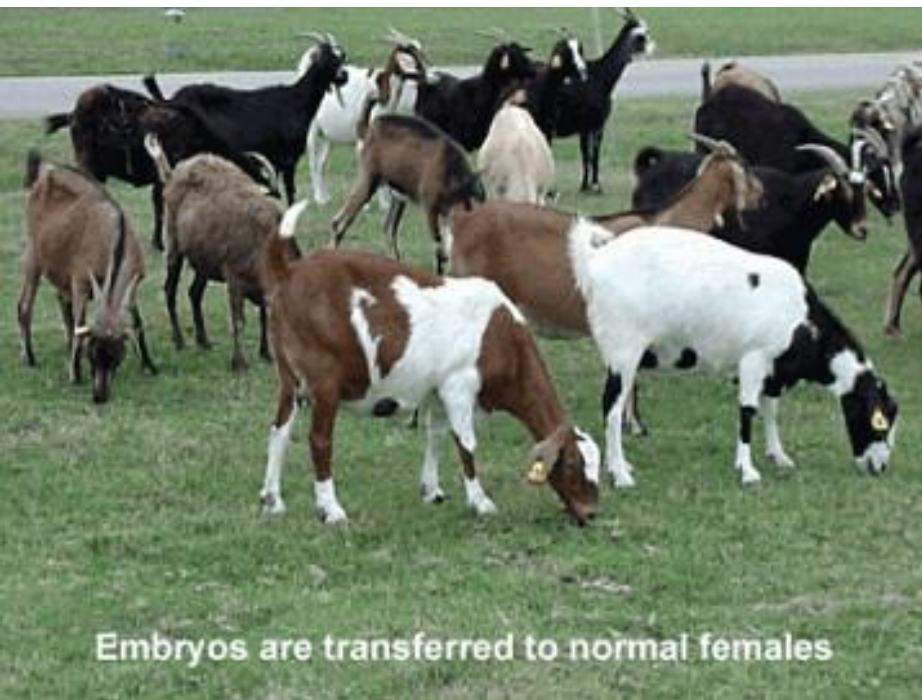
Uterine  
implantation



150 days  
gestation



Dolly







**Joseph Ratzinger,**

**- Cloning is a human arrogance**



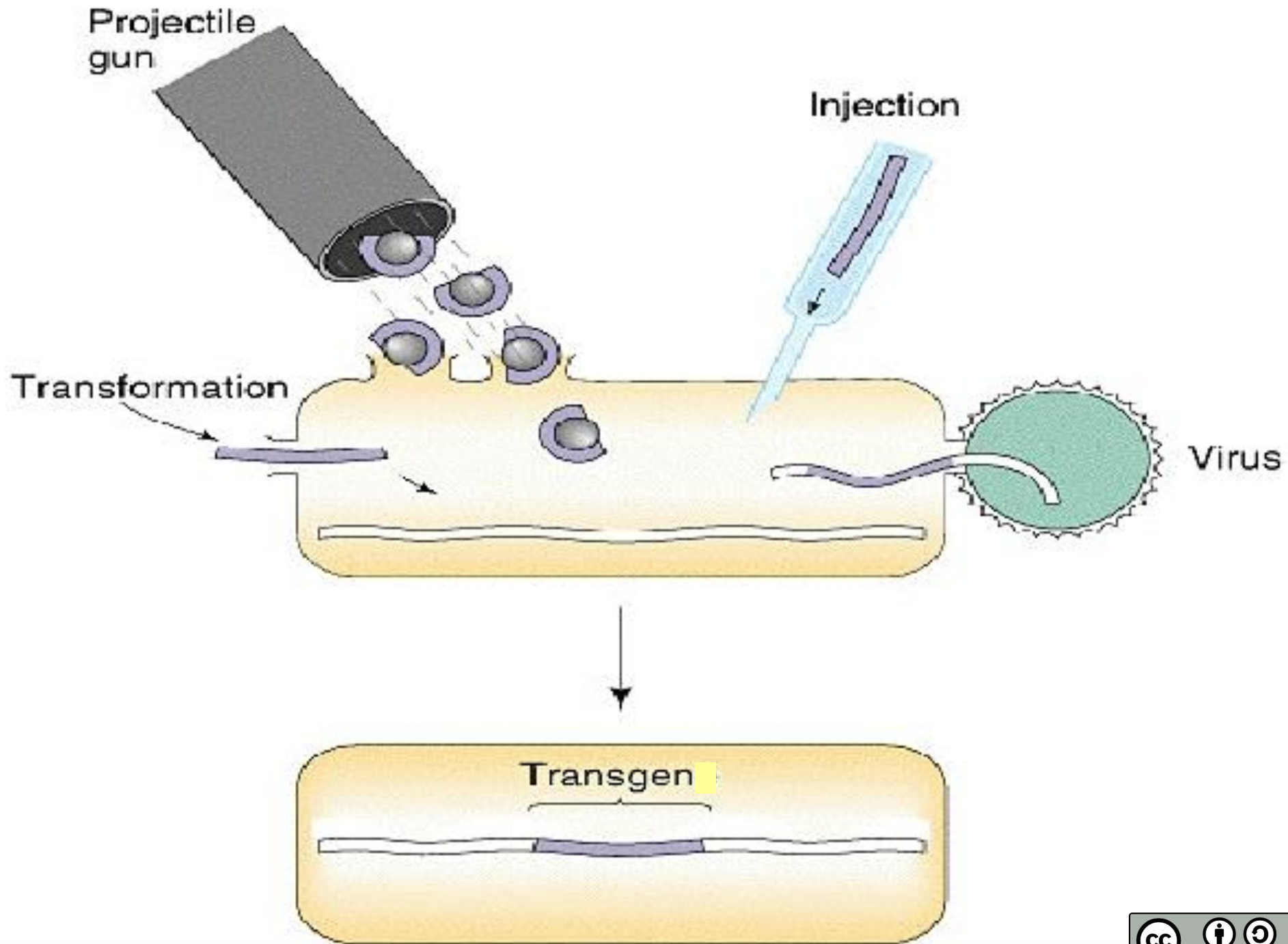
**2004**

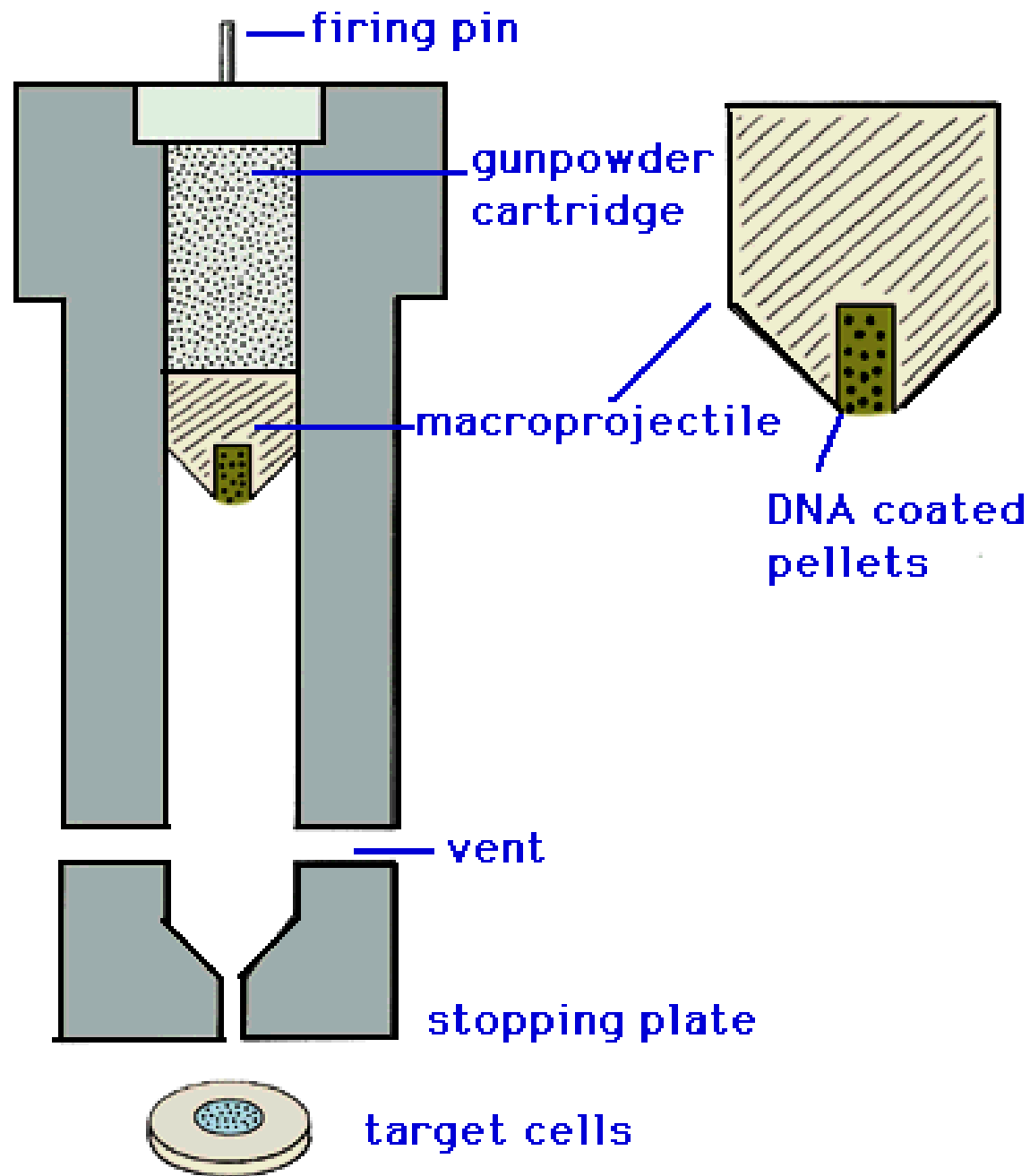
Woo Suk Hwany from Soul National University in Korea:

Successful cloning of healthy human embryos, grown in mice

Panos Zavos – cloning of people for the production of healthy children

Number of teams – cloning of germ cells for the investigation and treatment of diseases







# Infection

**Recombinant viral genomes – replacement of genome regions not essential for the growth in tissue cultures**

**(Vaccinia virus, Baculovirus, Herpesvirus, Retroviruses, HepDNA)**

**- Sometimes the encapsidation supported by helper virus**

## **Vectors based on recombinant viruses**

**Infection vs. Lipofection etc.**

**Infection efficiency often quite high**

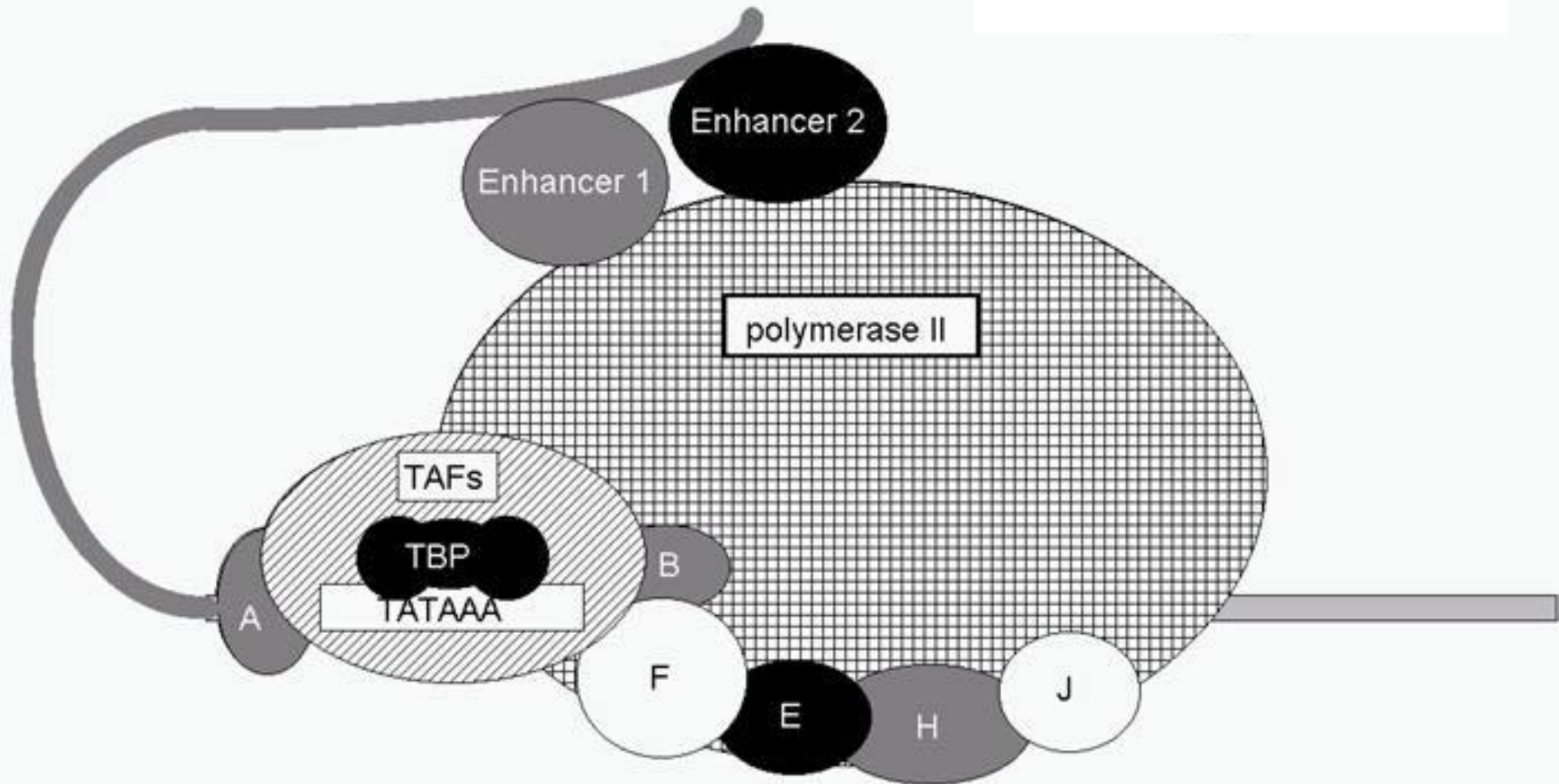
**Combined with high titer = ~ 100% of cells**

**Tissue specific targeting - 'Magic Bullet' transgenesis**

- **Targeting of genes to specific cells**
- **Specific receptors**



# Transcription initiating complex



**TATA box (TATAAAA) (Hogness box)**

**Promoter proximal elements, enhancers**

**Ribosom binding Kozak's sequence ACCAUGG**

# **Promoters of mammalian cells with enhancer sequences**

**HCMV - Human cytomegalovirus - early promoter**

**SV40 – early promoter**

**SV40 - late promoter**

**Adenovirus – main late promoter**

**Herpes Simplex Virus - TK promoter**

**Promoter of mouse metallothionein I (heavy metal induction)**



# **Retrovirus promoters**

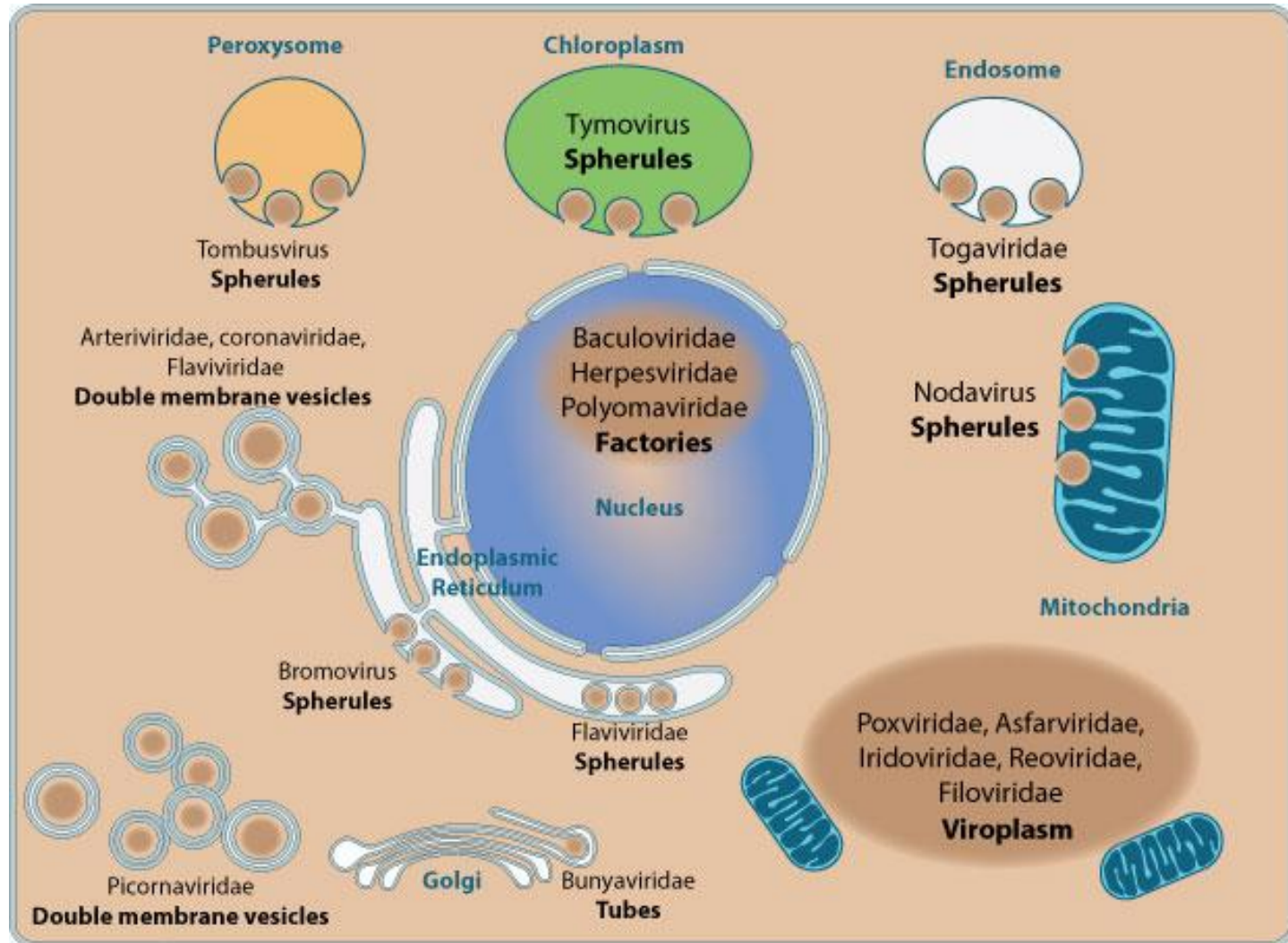
**Rous Sarcoma Virus - LTR promoter**

**Mouse Mammary Tumour Virus - LTR promoter (induced  
by glucocorticoids)**

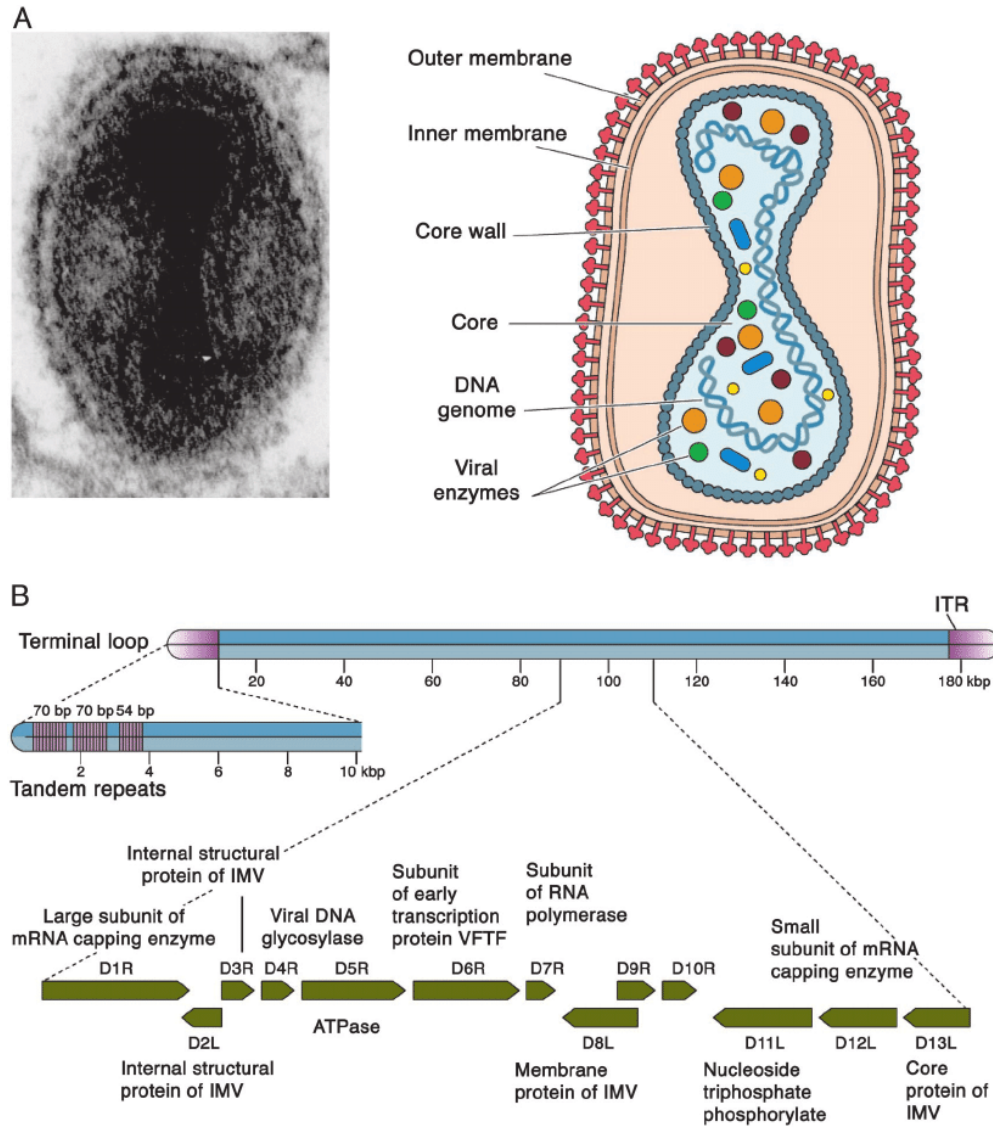
**Moloney Murine Sarcoma Virus - LTR promoter**

# Vaccinia virus

transcription and transient expression in cytoplasm (viral factors)



- Infects most mammalian and avian cell lines, no integration to cell chromosome
- No infection of CHO (Chinese hamster ovary), primary lymphocytes, macrophages
- Cloning of big fragments (20 kbp), stable infectivity, high expression, secretion, “suitable” posttranslation modification



# **Baculovirus** *Autographa californica*

big, enveloped, ds DNA virus,

Entry via endocytosis - transport to nucleus

Polyhedrin protein – protects against proteolytic inactivation

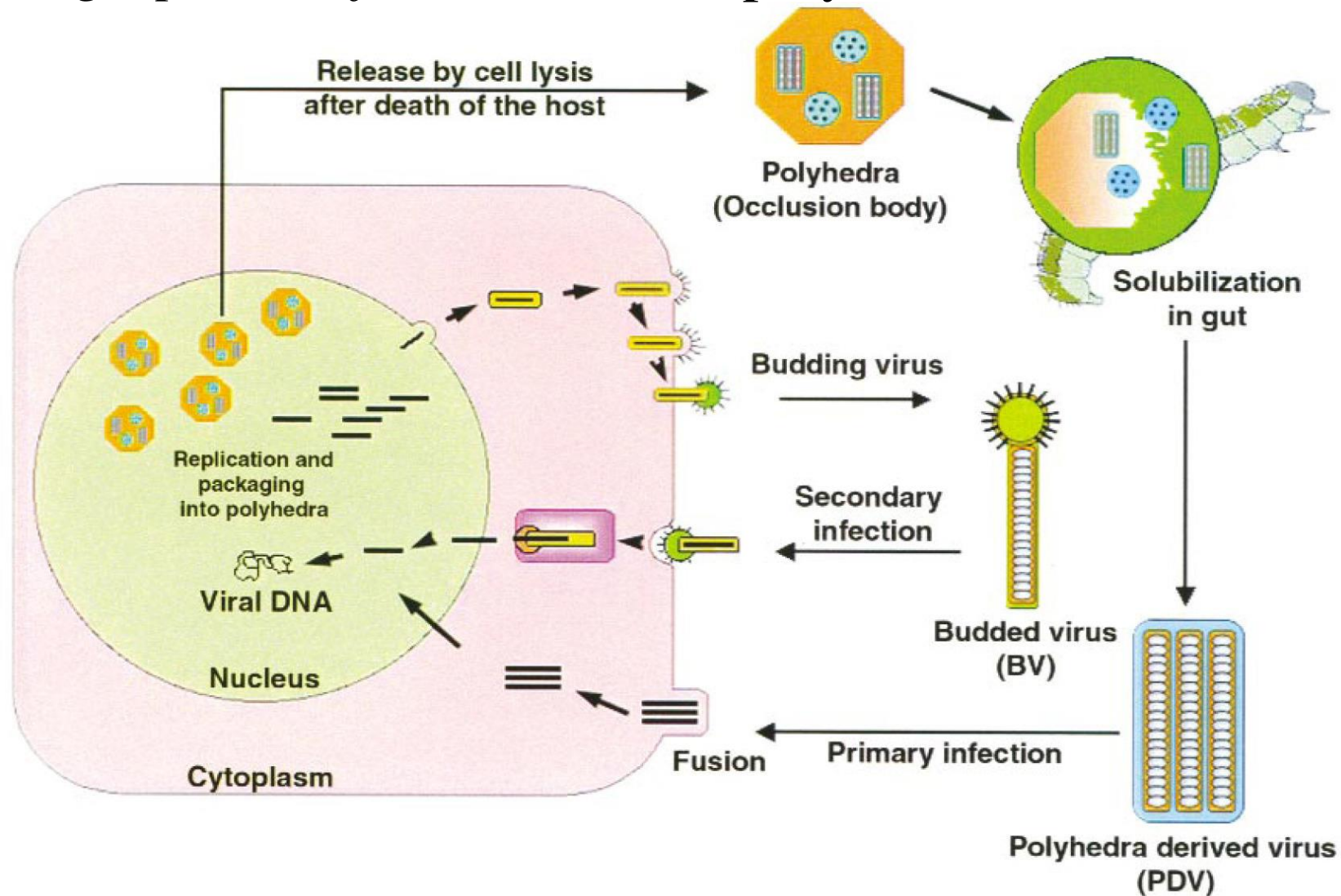
- Solubilized in intestine
- No replication in mammals and plants = safe
- recombinant baculoviruses with surface-exposed flu vaccine



**Host cells**

**Sf9** (*Spodoptera frugiperda*)

# *Autographa californica* nuclear polyhedrosis virus (AcNPV)



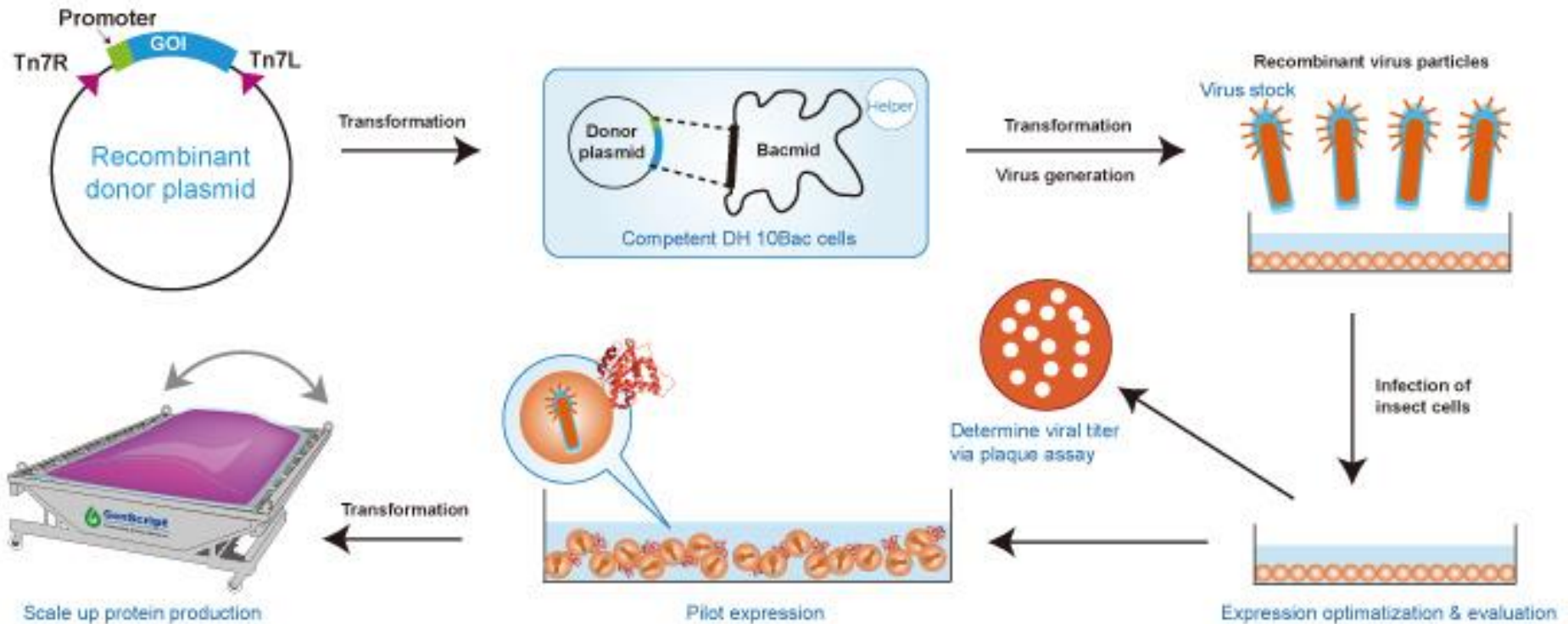
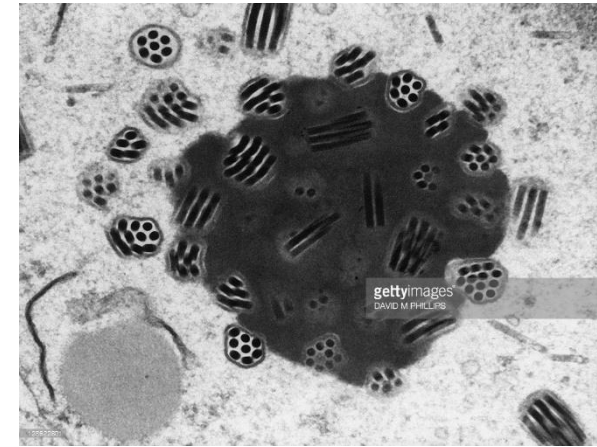


# Recombinant baculovirus

- replacement of polyhedrin sequence with cloned gene

- homologous recombination

- different morphology of plaques – identification



# Common cell lines

## Human

HeLa	cervix carcinoma (Henrietta Lacks)
HEK 293	embryonic kidney – transformed with adenovirus
Jurkat	leukemia T cells
MCF7	mammary carcinoma

## Subhuman primates

Vero	monkey kidney
COS	monkey kidney

## Rodents

3T3	mouse fibroblasts (transformed, non-cancer)
NS0 a Sp2/0	mouse myeloma
CHO	Chinese hamster ovary)



# Enzyme markers and antibodies for cell type identification

**TABLE 35.1 Some common enzyme markers for cell line identification**

<i>Enzyme</i>	<i>Cell type</i>
Tyrosine aminotransferase	Hepatocytes
Tyrosinase	Melanocytes
Glutamyl synthase	Brain (astroglia)
Creatine kinase (isoenzyme MM)	Muscle cells
Creatine kinase (isoenzyme BB)	Neurons, neuroendocrine cells
Non-specific esterase	Macrophages
DOPA-decarboxylase	Neurons
Alkaline phosphatase	Enterocytes, type II pneumocytes
Angiotensin-converting enzyme	Endothelium
Sucrase	Enterocytes
Neuron-specific esterase	Neurons

**TABLE 35.2 A selected list of antibodies used for the detection of cell types**

<i>Antibody</i>	<i>Cell type</i>
Cytokeratin	Epithelium
Epithelial membrane antigen	Epithelium
Albumin	Hepatocytes
$\alpha$ -Lactalbumin	Breast epithelium
Carcinoembryonic antigen (CEA)	Colorectal and lung adenocarcinoma
Prostate specific antigen (PSA)	Prostatic epithelium
Intracellular cell adhesion molecule (I-CAM)	T-cells and endothelium
$\alpha$ -Fetoprotein	Fetal hepatocytes
Human chorionic gonadotropin (hCG)	Placental epithelium
Human growth hormone (hGH)	Anterior pituitary
Vimentin	Mesodermal cells
Integrins	All cells
Actin	All cells

# CHO (Chinese hamster ovary) cells

Prototype of tissue cells

Mutants deficient in

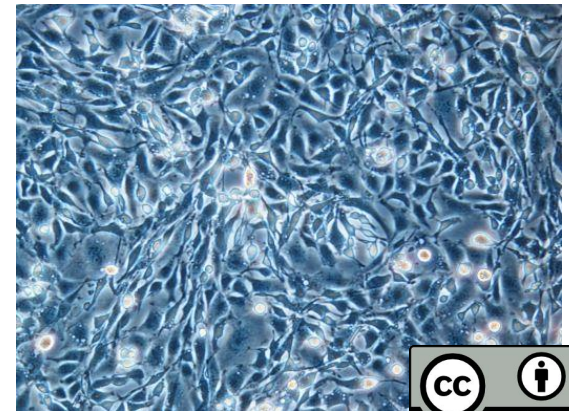
Adenine phosphoribosyl transferase (APRT)

Dihydrofolate reductase (DHFR) - selection

Require proline

Growth in suspensions – possible culturing in fermentors

Yields ~ 5 g protein/l culture



## **COS cells (monkey kidney cells)**

- from CV-1 cells (CV-1 in Origin with SV40 genes)

Common COS-1 a COS-7

**transformed with replication-defective virus SV40**

- **High concentration of SV40 T antigen, but not virus**
- **episomal replication of vector**
- **high production thanks to high copy-number - 10 000/cell**

**Plasmids with SV40 promoter (constitutive)**

# Stable transfection

- Integration to chromosome

## Selection

**Adenosin deaminase (ADA)**

decontamination **9- $\beta$ -D-xylofuranosyl adenin (Xyl-A)** to inosin derivative

**Xyl-A**  $\rightarrow$  Xyl-ATP incorporation to DNA  $\rightarrow$  **apoptosis**

ADA-deficient line

**Bleomycine binding protein**

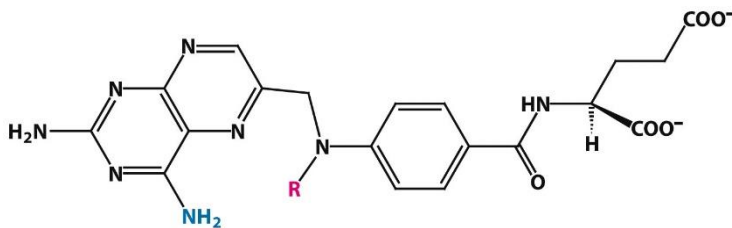
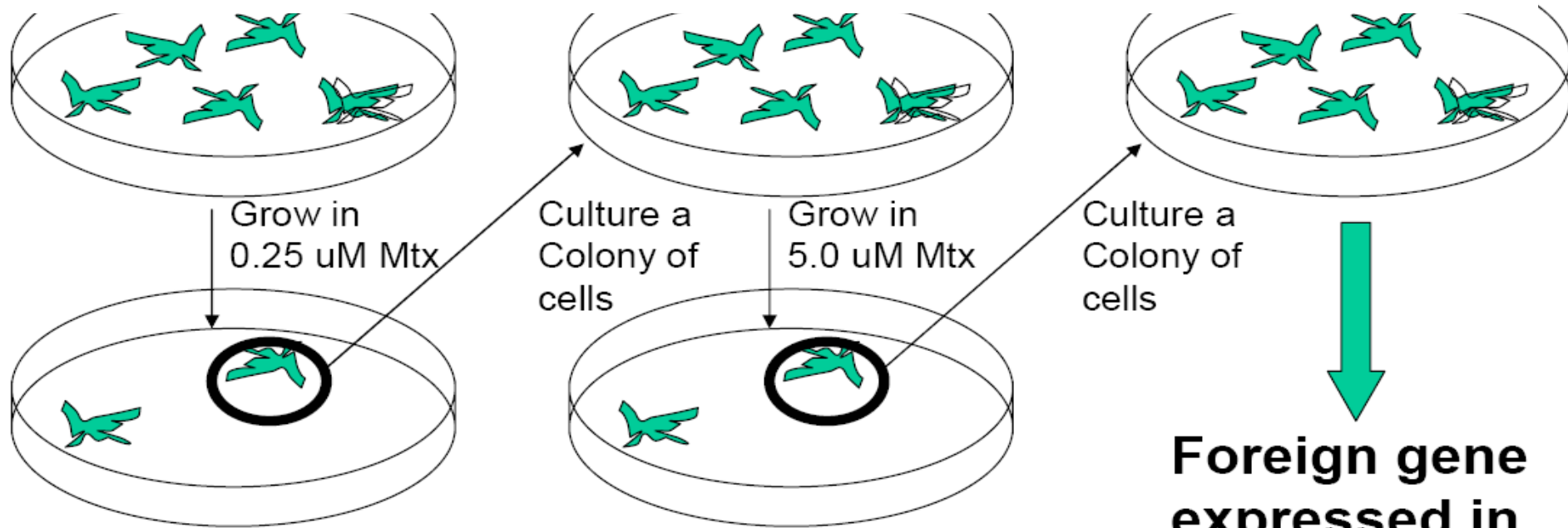
Bleomycin / glycopeptide antibiotics / stopped growth in v G<sub>2</sub> phase

**Dihydrofolate reductase (DHFR)** – biosynthesis of tetrahydrofolate => purins

competitive inhibitor methotrexate (MTX, 4-amino-10-methylfolate)

– absence of exogenous purines – addiction of cells on DHFR

→ growth of cell overexpressing DHFR

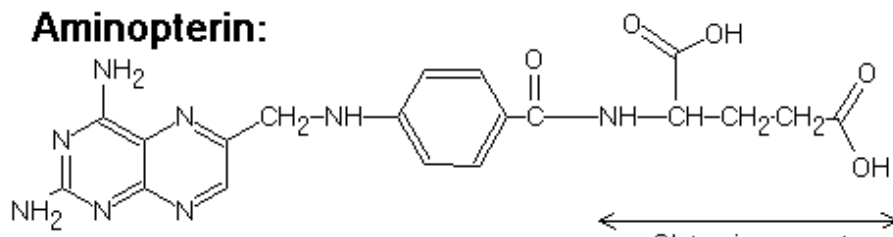


**Aminopterin (R = H) or methotrexate (R = CH<sub>3</sub>)**

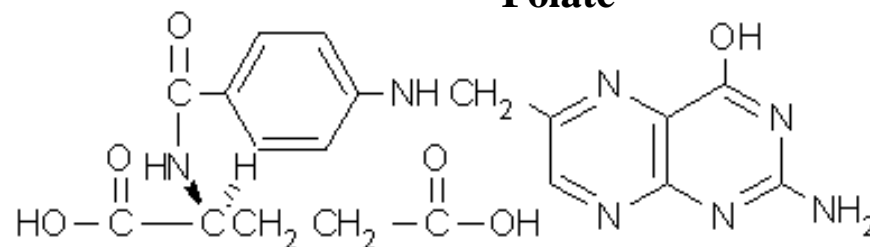
**Foreign gene  
expressed in  
high level in  
CHO cells**

**Aminopterin** (analogue of folate and *methotrexate*) - block of biosynthesis of tetrahydrofolate  
dihydrofolate reductase → block of biosynthesis of purines and TMP

**Aminopterin:**



**Folate**



**Possible recovery - alternative synthesis:**

dTTP from **thymidine** with **thymidine kinase (TK)**

dGTP from **hypoxanthine** (purine) and phosphoribosyl pyrophosphate  
**hypoxanthineguaninephosphoribosyl transferase (HGPRT)**

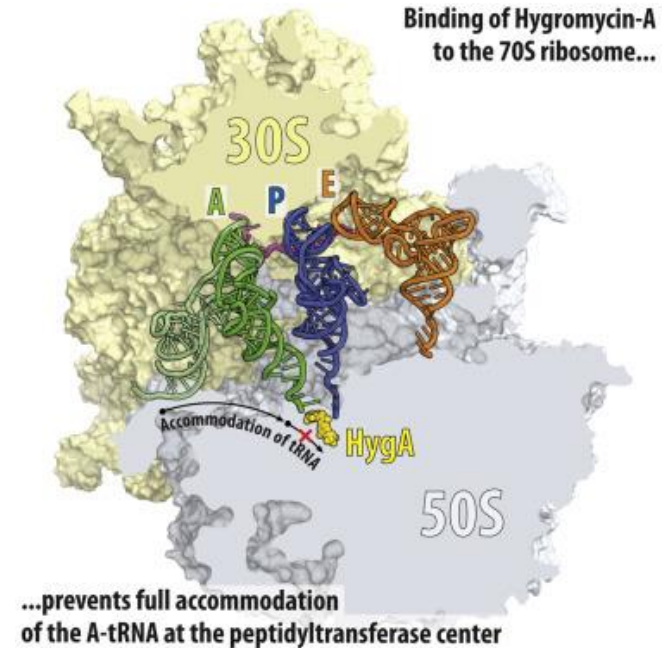
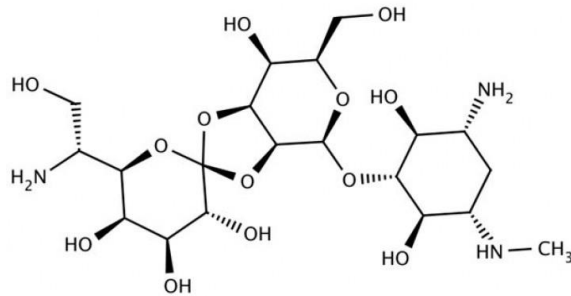
**Cells deficient in:**

**TK and HGPRT**

– **genes for these enzymes** in plasmids or in fused cells (hybridoma)

Selection of cells in HAT media containing aminopterin, thymidine and hypoxanthine





## Hygromycin

- aminoglycoside antibiotics

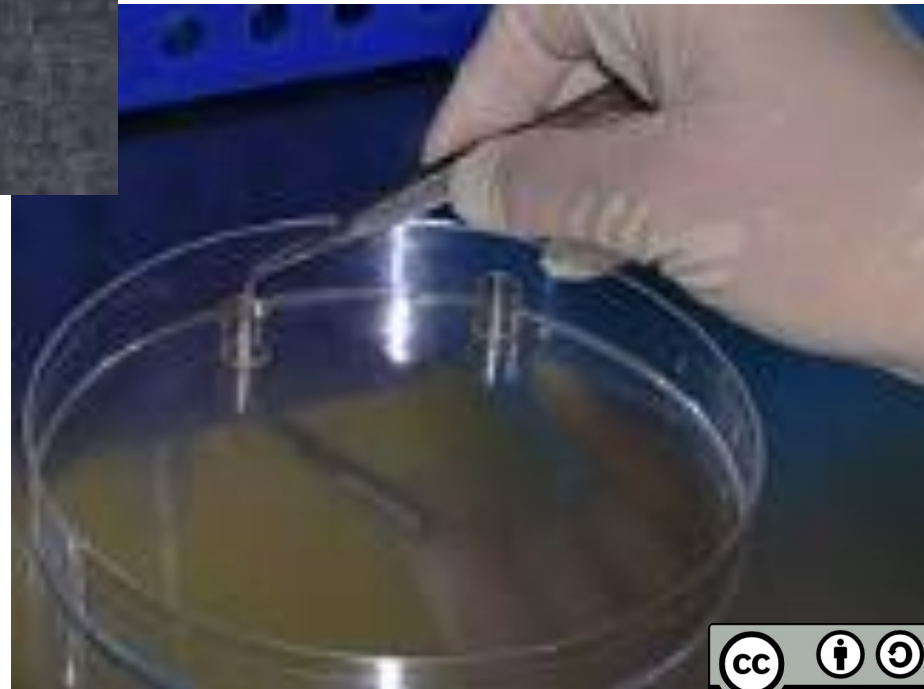
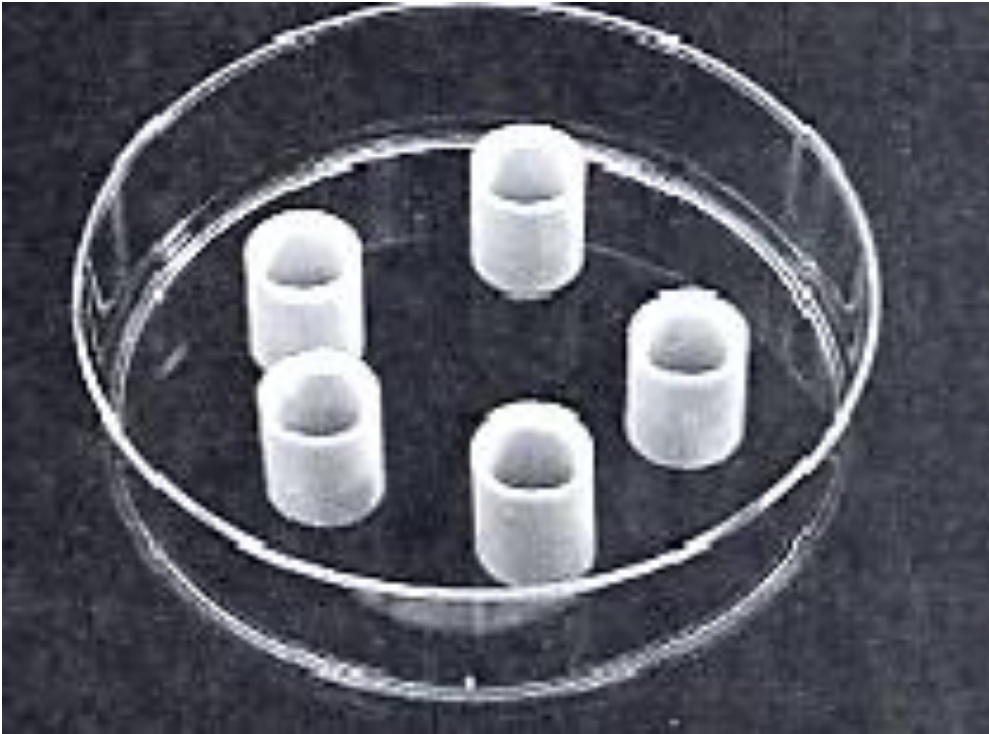
(from *Streptomyces hygroscopicus*) inhibits proteosynthesis

## Hygromycinphosphotransferase (HPH)

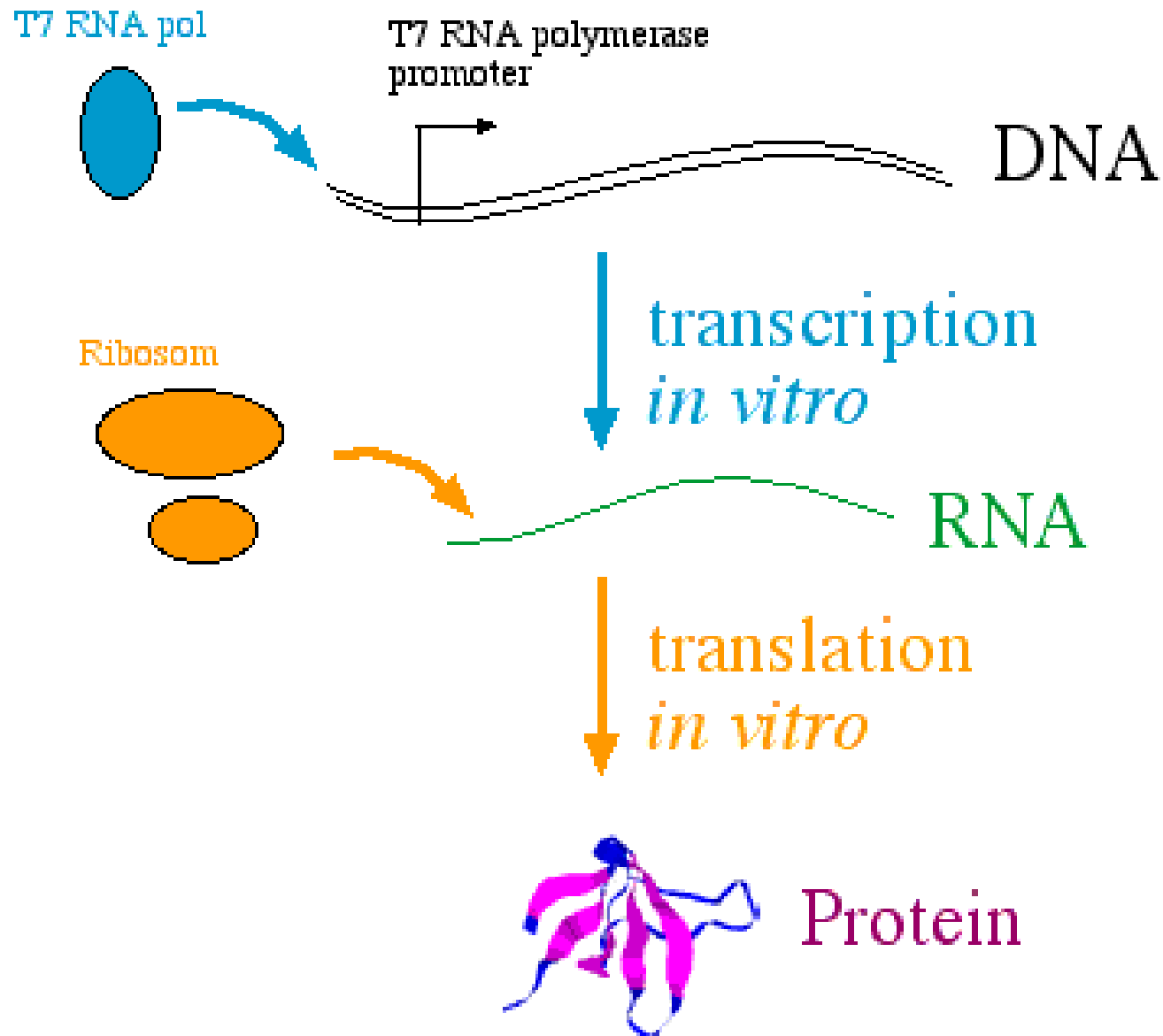
– phosphorylation inactivates hygromycin



# Clone isolation



# *In vitro* transcription/translation



# **Production of recombinant proteins**

**Targeting of genes to tissues – gene therapy**

**Research purpose**

**basic research**

**applied research**

**Medicine**

**treatment**

**prevention**

**diagnostics**

# RECOMBINANT PROTEINS

**Sales estimate - 56 billion \$ in 2006**

## **Blood derivatives**

Erythropoietins

Antihemophilic blood factors

Thrombolytic agents

Other recombinant agents related to blood

# **Interferon**

Interferon alpha

Interferon beta

other interferons

# **Recombinant hormones**

Insulin

Other recombinant hormones

Growth hormones

# **Recombinant vaccines**

**Vaccine against hepatitis**

**Other recombinant vaccines**

**Other recombinant proteins**

# **Monoclonal antibodies**

**Therapeutic**

**Anti-cancerogenic**

**Anti-inflammatory**

**Thrombolytic**

**Diagnostic**



## Comparison of different production systems for expression of recombinant proteins

System	Production cost	Time effort	Scale-up capacity	Product quality	Glycosylation	Contamination risk	Storage	Ethnic concerns
Bacteria	Low	Low	High	Low	None	Endotoxins	Medium/ -20 <sup>o</sup> C	Low
Yeast	Medium	Medium	High	Medium	Incorrect	Low	Medium/ -20 <sup>o</sup> C	Low
Mammalian cell cultures	High	High	Very low	Very high	Correct	Viruses, oncogenes	Difficult/ N <sub>2</sub>	Existing
Transgenic animals	High	High	Low	Very high	Correct	Viruses, oncogenes	Difficult	High
Plant cell cultures	Medium	Medium	Medium	High	Minor differences	Low	Medium/ -20 <sup>o</sup> C	Low
Transgenic plants	Low	High	Very high	High	Minor differences	Low	Easy/RT	Existing

S. Biemelt; U. Sonnewald (2004)

Presented in Credit Seminar (Division of Agricultural Physics, IARI, New Delhi) by Nirmal Kumar

Table 1: Recent FDA-Approved Products Using Animal/Yeast Alternative Expression Systems				
Drug Name	Indication	Stage of Development	Sponsor	Cell Line
<b>Xigris</b> (withdrawn in 2011)	Sepsis	Approved 2001	Eli Lilly & Company	HEK293 cells
<b>RotaTeq</b>	Rotavirus gastroenteritis	Approved 2006	Merck Sharp and Dohme Corp.	VERO cells
<b>ACAM2000</b>	Small Pox	Approved 2007	Sanofi Pasteur	VERO cells
<b>Rotarix</b>	Rotavirus gastroenteritis	Approved 2008	GlaxoSmithKline	VERO cells
<b>YF-Vax</b>	Yellow Fever	Approved 2008	Sanofi Pasteur	ALV-Free Chicken Embryos
<b>Kalbitor</b>	Hereditary angioedema (HAE)	Approved 2009	Dyax, Corp.	<i>P. pastoris</i> cells
<b>Ixiaro</b>	Japanese Encephalitis	Approved 2009	Intercell Biomedical	VERO cells
<b>Ceravix</b>	Human Papillomavirus (HPV)	Approved 2011	GlaxoSmithKline	Baculovirus Insect cell
<b>Benlysta</b>	Lupus	Approved 2011	Human Genome Sciences	NS0 cells

# **Production of proteins with baculovirus system**

- **Alpha and beta interferon**
- **Adenosin deaminase**
- **Erythropoietin**
- **Interleukin 2**
- **Poliovirus proteins**
- **Activator of plasminogen (TPA)**

+ GH



# **Analysis of efficient expression**

**RNA**

**RT-PCR, Northern blot, *in situ* hybridization**

**Protein**

**SDS-PAGE**

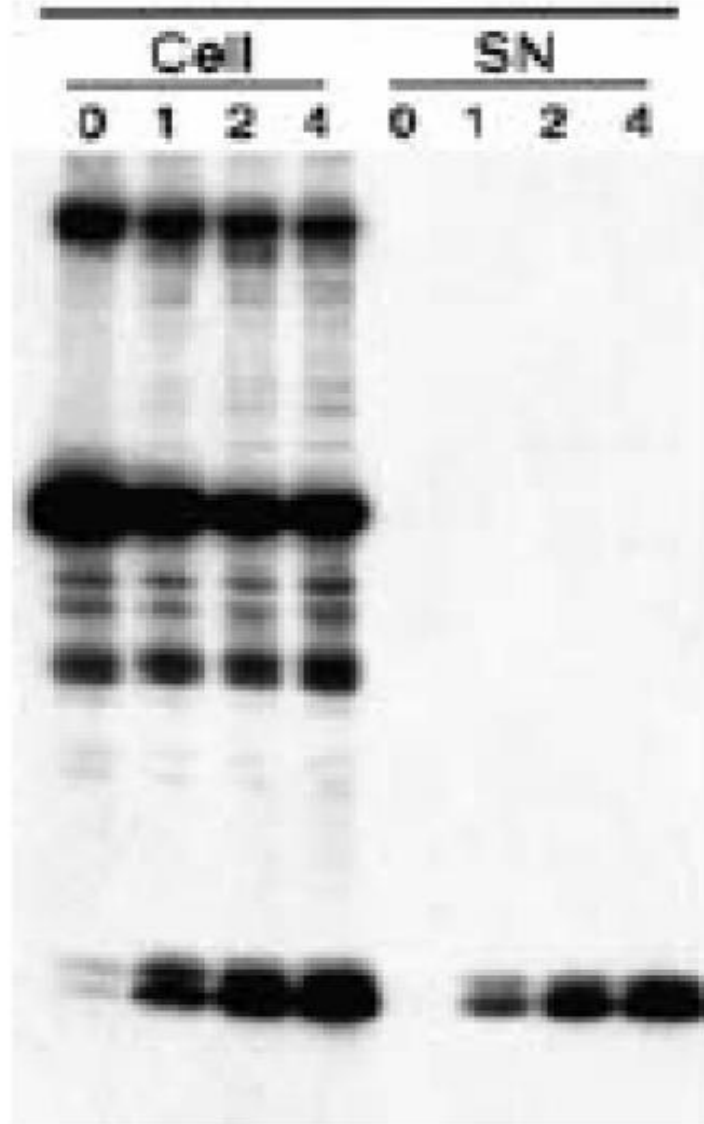
**Immunochemical methods**

**Pulse-chase**

**Metabolic labeling and immunoprecipitation**



# Pulse-chase



# Proteiny

**A**      *Staphylococcus aureus*

**G**      *Streptococci*

**L**      *Peptostreptococcus magnus*

A, G – Fc heavy chain; L – light chain

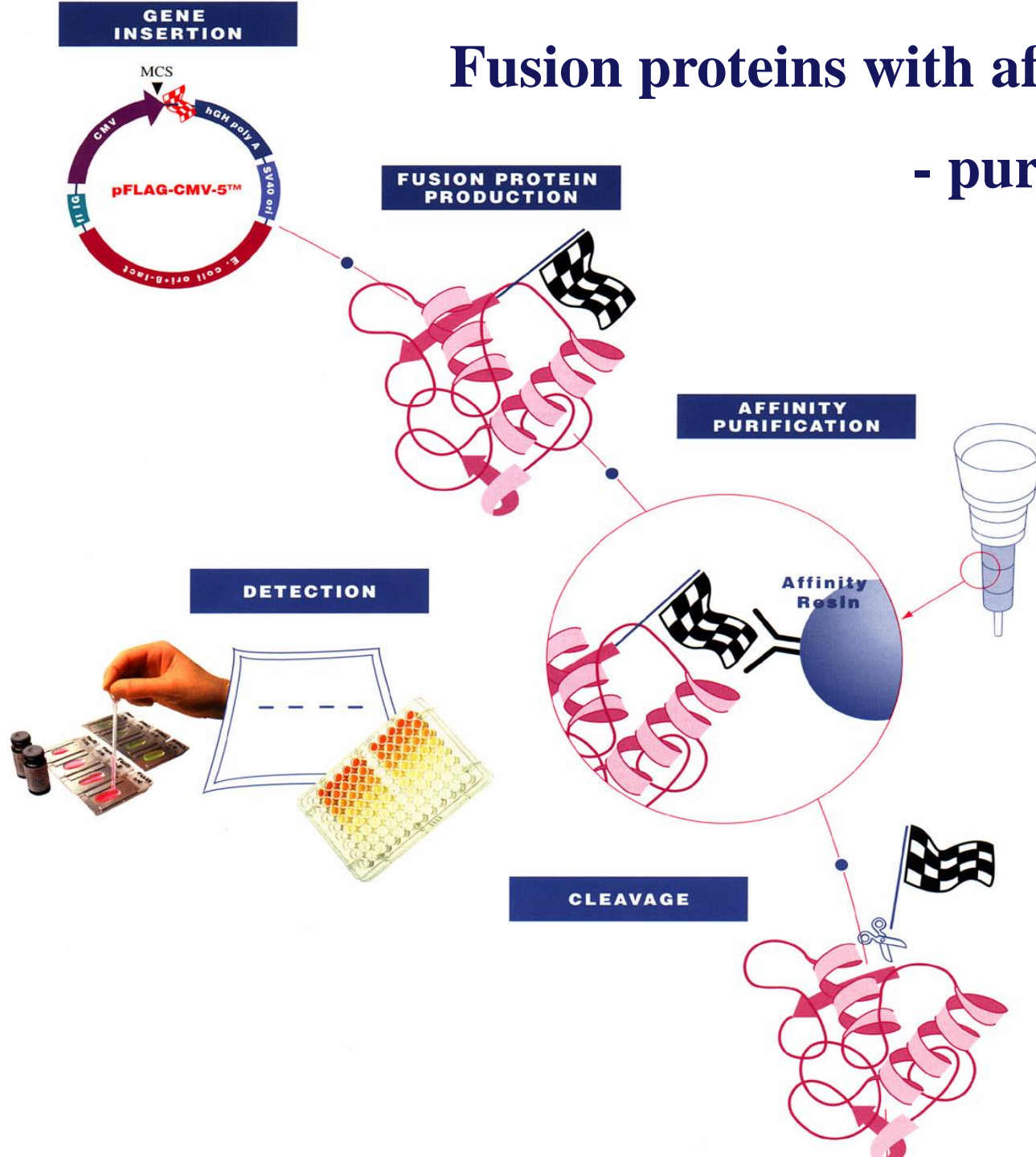
antibody	Protein A	Protein G
bovine	++	++++
human	++++	++++
equine	++	++++
goat	..	++
rabbit	++++	+++
rat	+/...	++
sheep	+/..	++
mouse	++	++
porcine	+++	+++



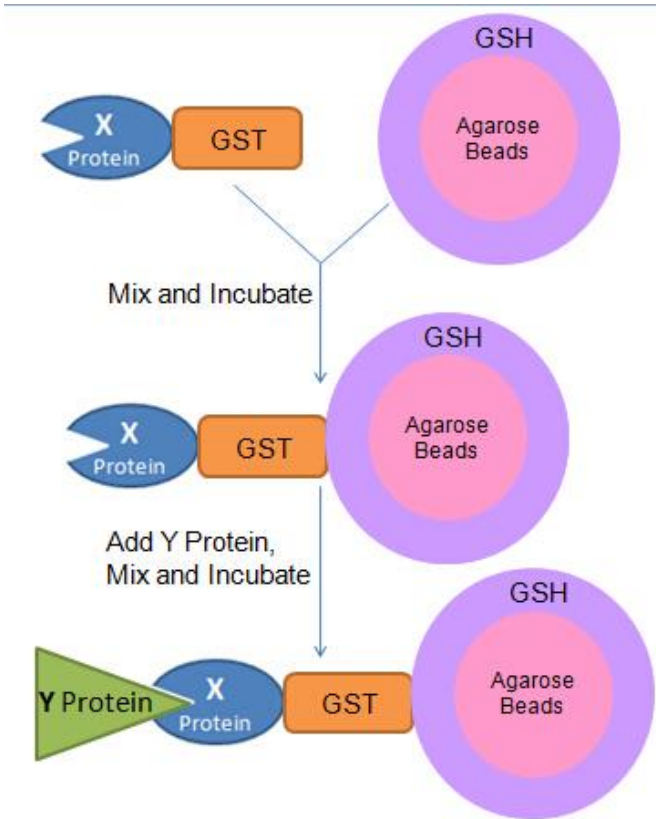
**Protein modification for higher yields, stabilization in soluble form, easier purification and detection**

# Fusion proteins with affinity tag

## - purification



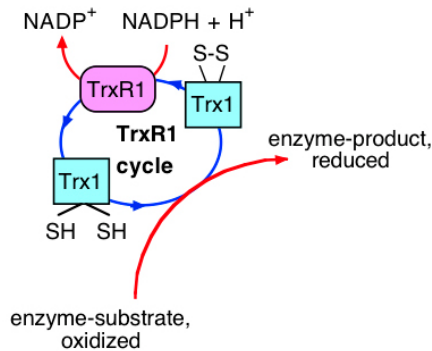
# Glutathion S-transferase (GST) 26 kDa, stabilizes protein, protects from proteolysis



**Pull down of protein interacting with GST-fusion protein  
immobilized on beads – identification of interacting partners**

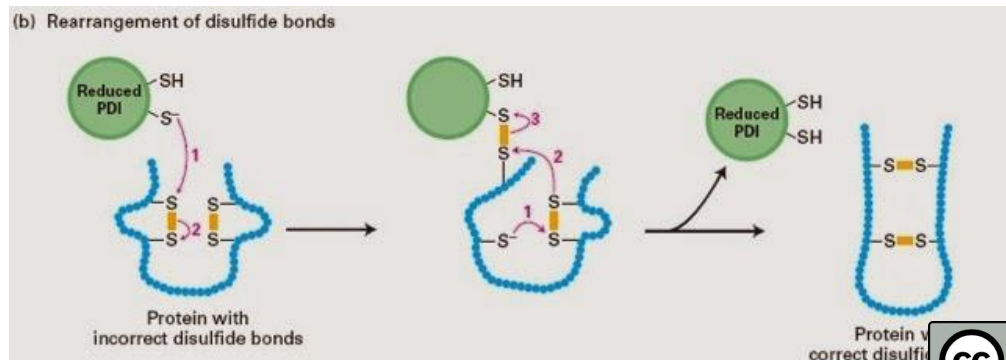
**Maltose-binding protein (*E. coli*, 42 kDa)**    maltose, increase in solubility

**Thioredoxin A (*E. coli* TrxA, 11,6 kDa)**    increased solubility, stability,  
promotes crystallization

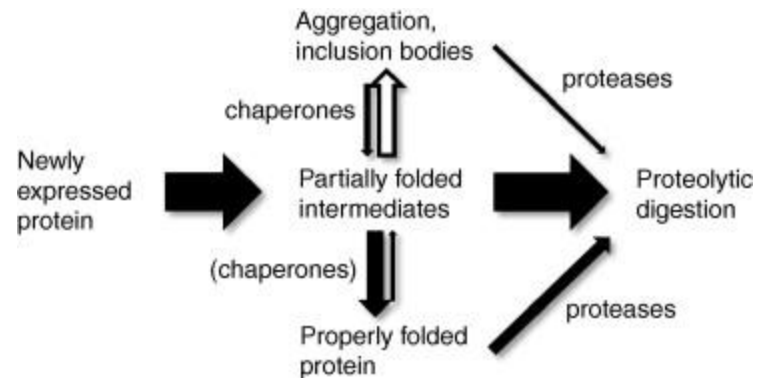


**SUMO (Small ubiquitin-related modifier)** supports protein folding – cleavage with  
Ulp protease from *S. cerevisiae* – more in other eukaryots..

**Protein disulfide isomerase** – up to 3-fold increase in solubility compared to TrxA

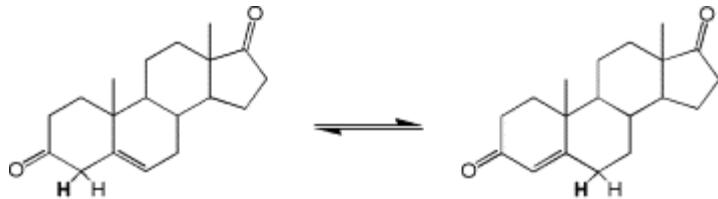


# Decreased solubility – supported formation of inclusion bodies



ketosteroid isomerase (*E. coli*, 13 kDa)

extremely insoluble protein



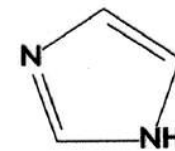
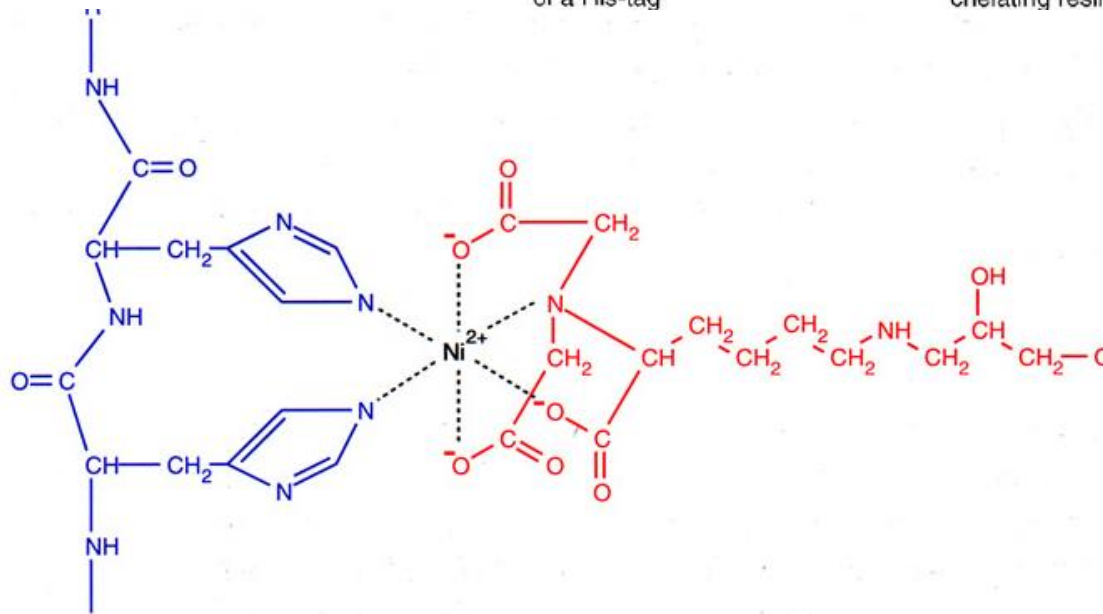
# Poly(hexa)histidine $\text{Ni}^{2+}$

## Immobilized-Metal Affinity Chromatography (IMAC)

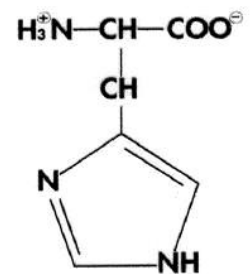


Two neighbouring residues  
of a His-tag

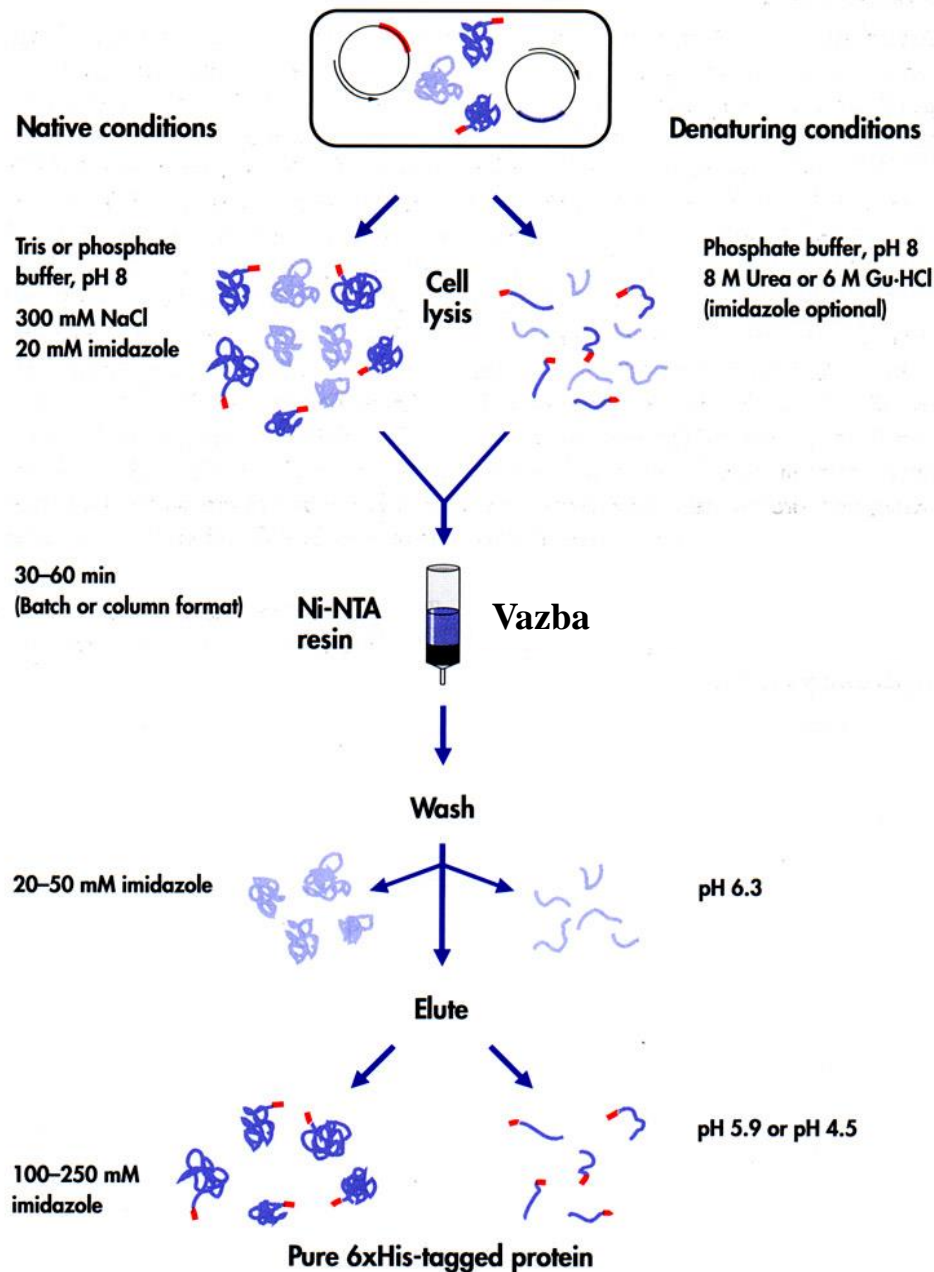
Nitrilotriacetic acid (NTA)  
chelating resin



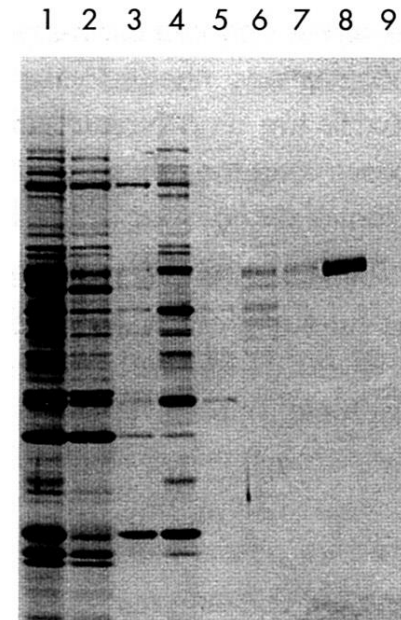
Imidazole



Histidine



# Purification of His-tag proteins



**Figure 18.** Purification under native conditions. Human serum response factor (SRF) was expressed from a vaccinia virus vector in HeLa cells and purified using Ni-NTA agarose with the indicated imidazole concentrations in the wash and elution steps. Proteins were visualized by Coomassie staining. **1:** cell lysate; **2:** flow-through; **3:** 0.8-mM wash; **4 & 5:** 8 mM wash; **6 & 7:** 40 mM wash; **8 & 9:** 80 mM elution.



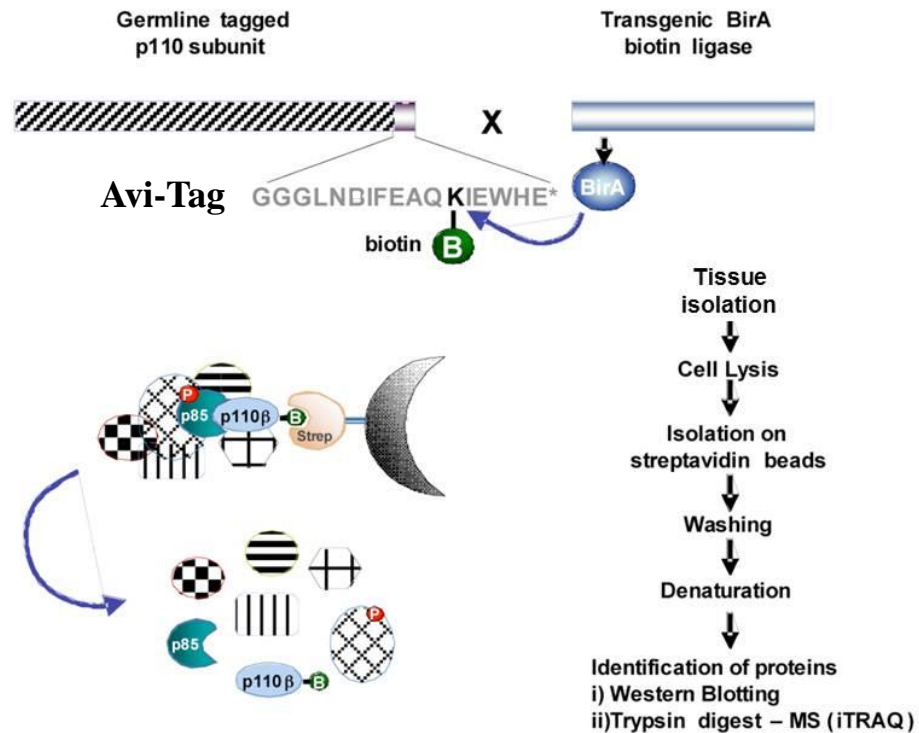
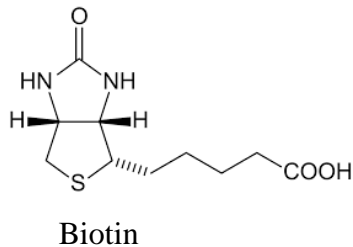
*In vivo* biotinylated peptide

avidin, streptavidin\*

Streptavidin-Binding Peptide (SBP)

streptavidin

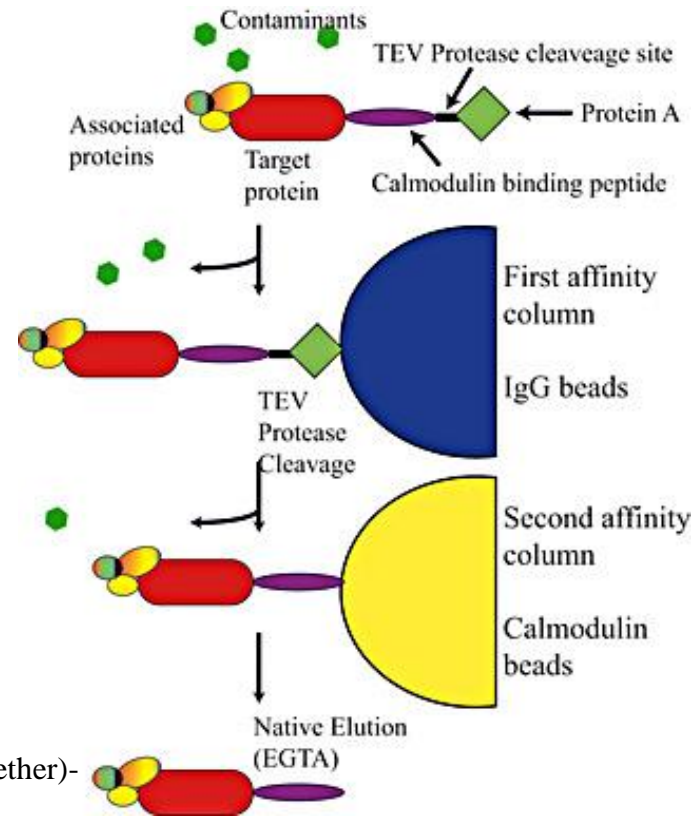
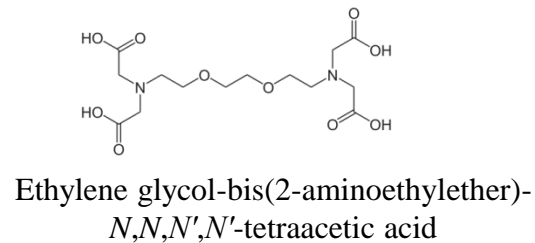
38-aa sequence



\*52.8 kDa protein *Streptomyces avidinii* - homo-tetramers high affinity to biotin

# Calmodulin binding protein (calcium-modulated protein)

Applicable only in bacteria



Tandem affinity chromatography

**Polyaspartate**

**Polyarginine**

**Polycysteine**

**anex**

**katex**

**thiol**

# Cleavage sites inserted into fusion proteins in *E. coli*

## Nonspecific cleavage

## Conditions/protease

↓  
Asp - Pro

acidic pH

↓  
Met - X

CNBr

Arg ↓ - X    nebo    Lys ↓ - X

Trypsin

Arg ↓ - X

Klostripain

# Cleavage sites inserted into fusion proteins in *E. coli*

## Recognized sequence

ENLYFQ ↓ S

Specific cleavage after SUMO (not Pro)

- in 2 M urea, pH 5,5-10,5)

Ile - Glu - Gly - Arg ↓ X

Leu - Val - Pro - Arg ↓ Gly - Ser

Leu - Glu - Val - Leu - Phe ↓ Gln ↓ Gly - Pro

Asp - Asp - Asp - Asp - Lys ↓ X

## Protease

TEV (Tobacco etch virus)

Ulp1 (SUMO proteasa)

Factor Xa

Thrombin

PreScission Protease™

(Amersham-Pharmacia, fusion of human rhinovirus 3C protease and GST)

Enterokinase

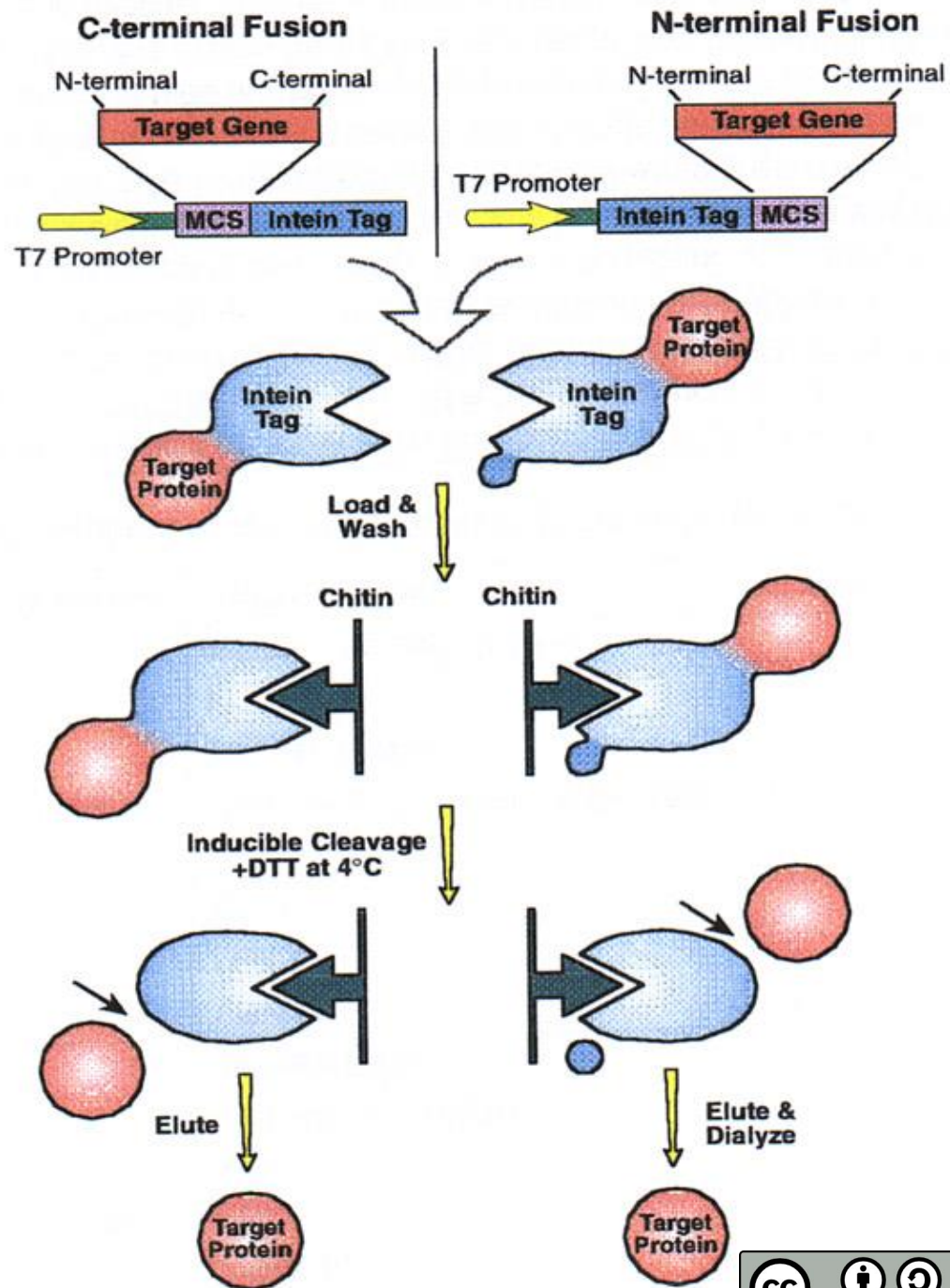
**IMPACT:**

**Purification**

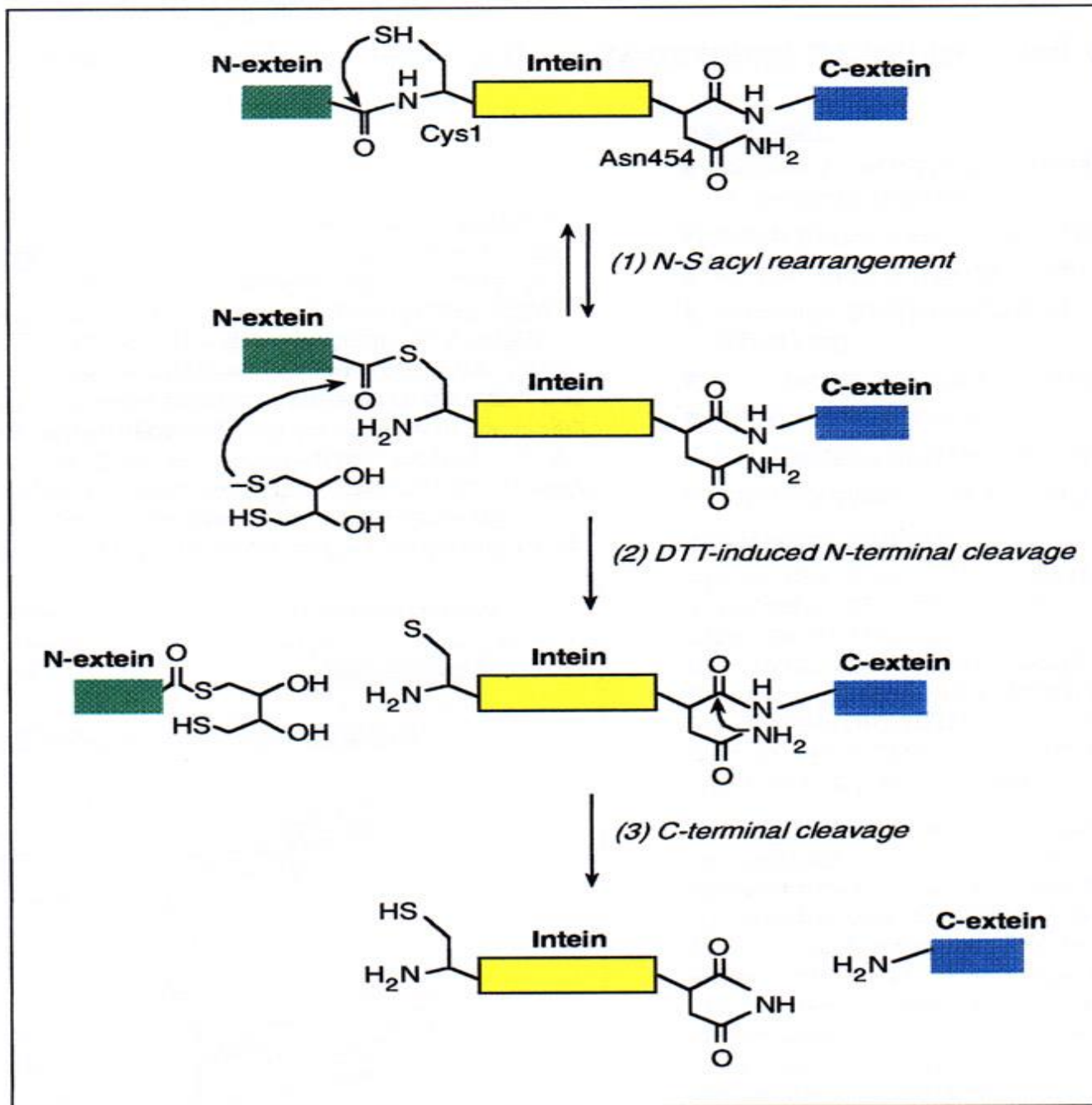
***Based on intein***

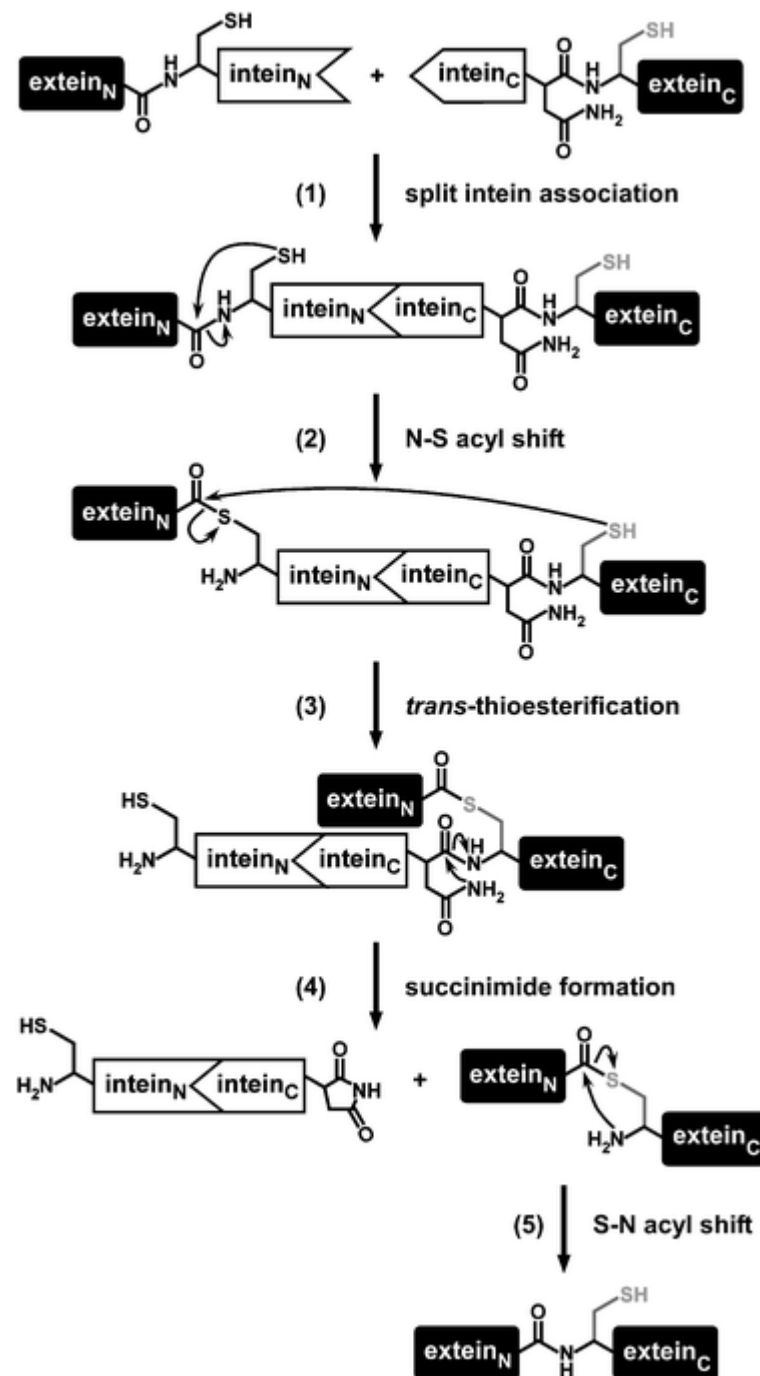
***Affinity chromatography***

***Chitin-binding protein***

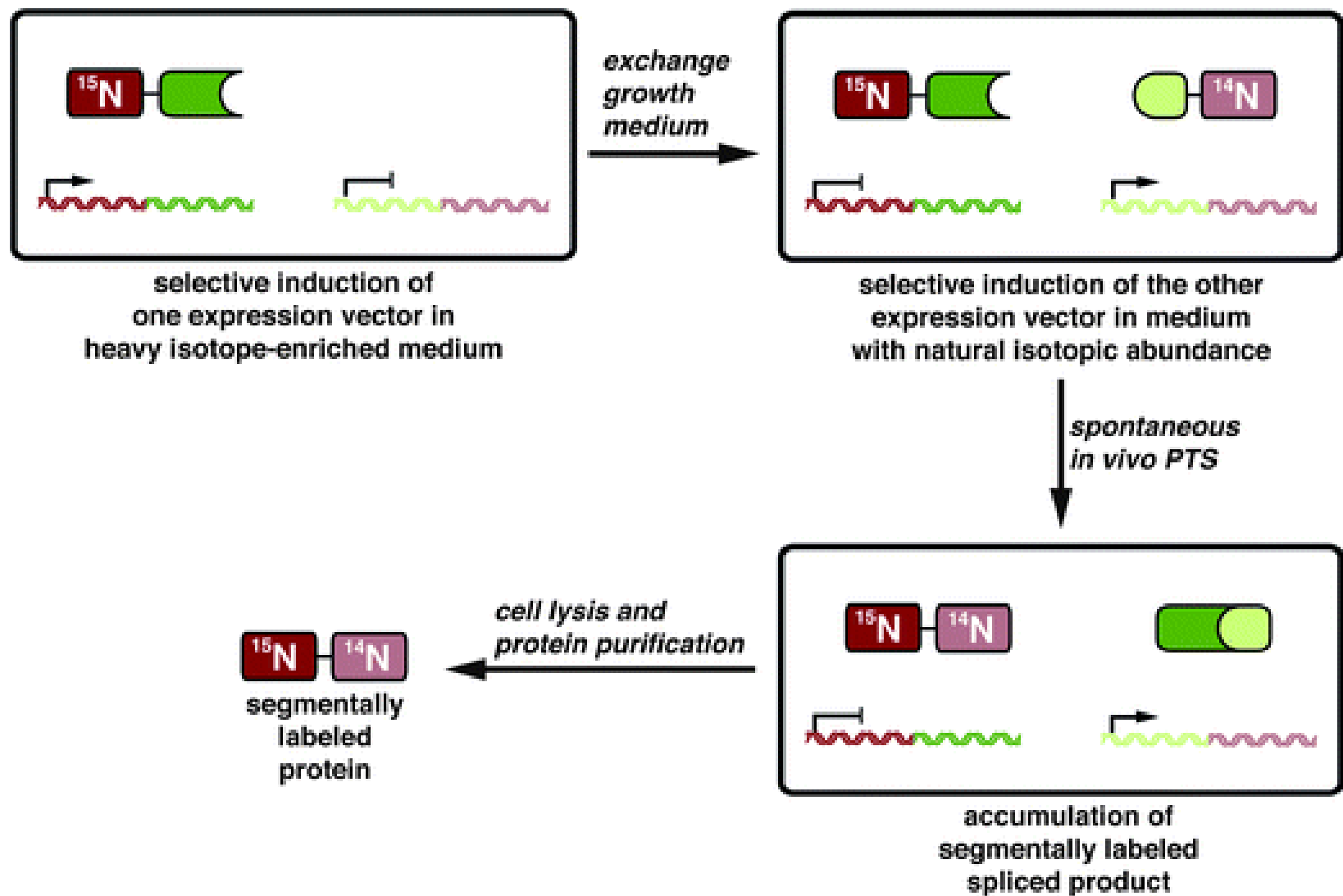


# Thiole-induced cleavage - modified intein









PTS - protein *trans*-splicing

# Epitope tagging

Fusion tagging of proteins with peptides – interaction with antibody

HA

YPYDVPDYA

c-myc

QVFFRNKLLF

FLAG (or 3xFLAG)

DYKDDDDK

HSV

QPELAPEDPED

T7

MASMTGGQQMG

# Fusion tagging with fluorescent protein

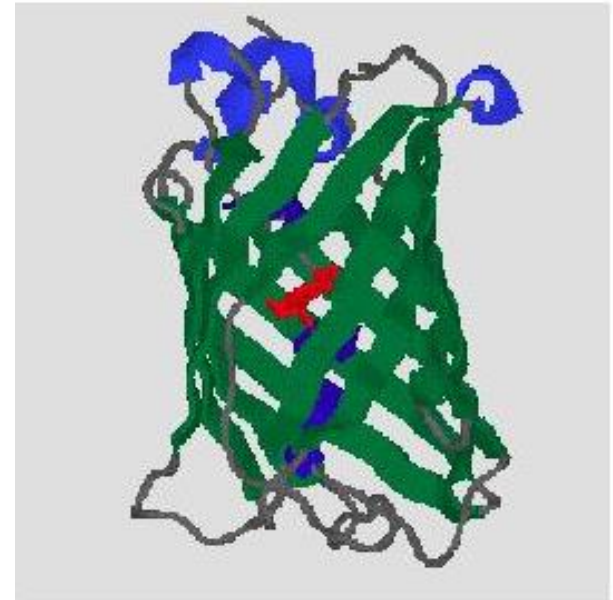
The inventor of GFP



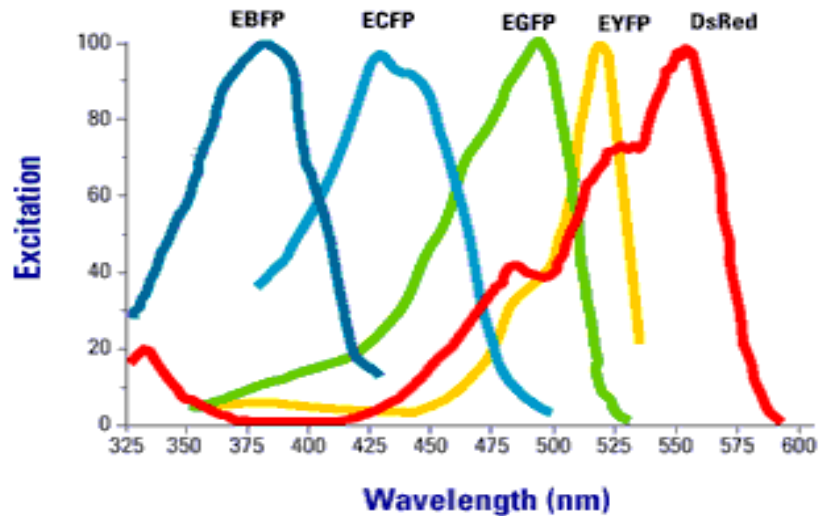
(GFP – Green fluorescent protein)

**Jellyfish – emit light through energy transfer to GFP**  
**GFP from *Aequorea victoria* fluoresce upon**  
**absorption of energy from photoprotein aequorine**  
**activated by  $\text{Ca}^{2+}$**

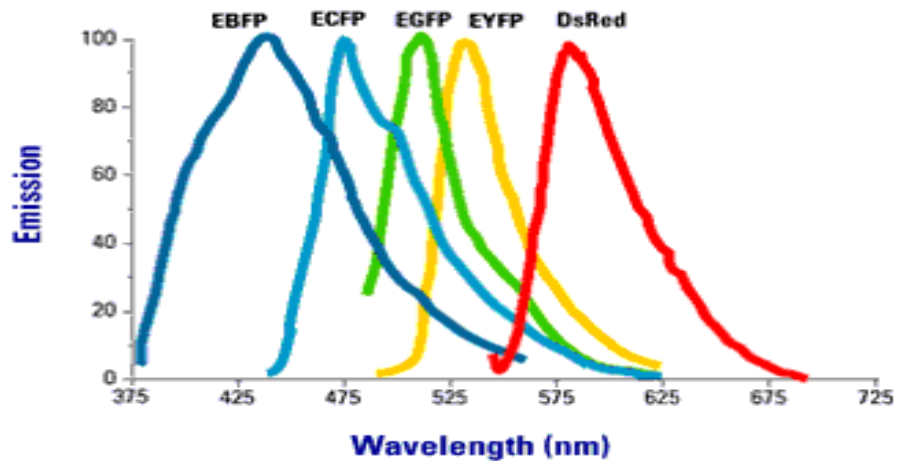
**Variants of GFP – various excitation/emission**  
**wavelengths**



## Excitation



## Emission

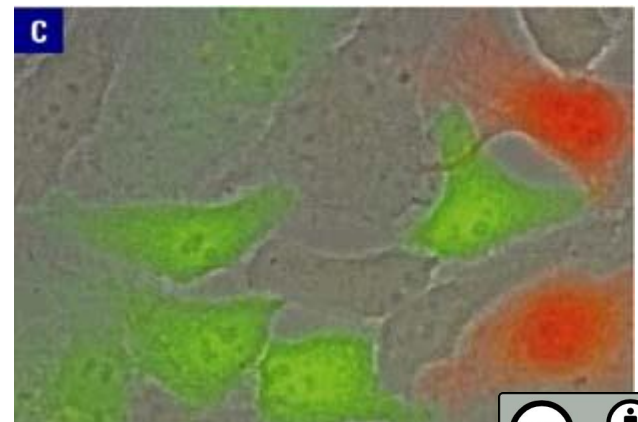
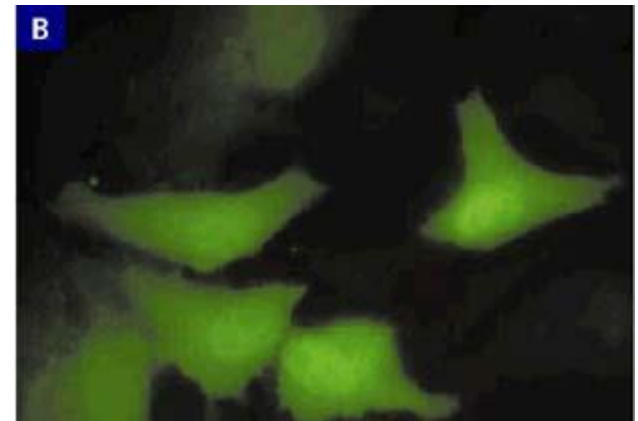
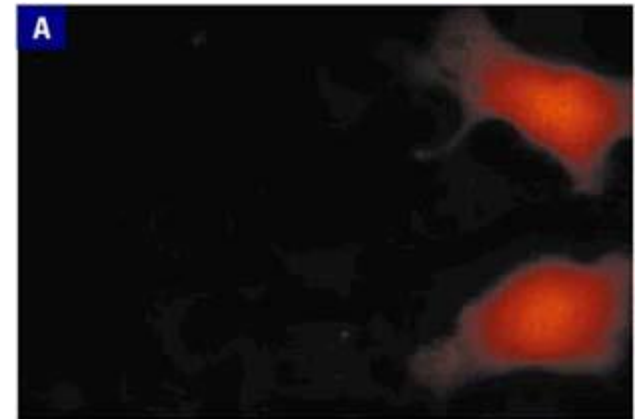


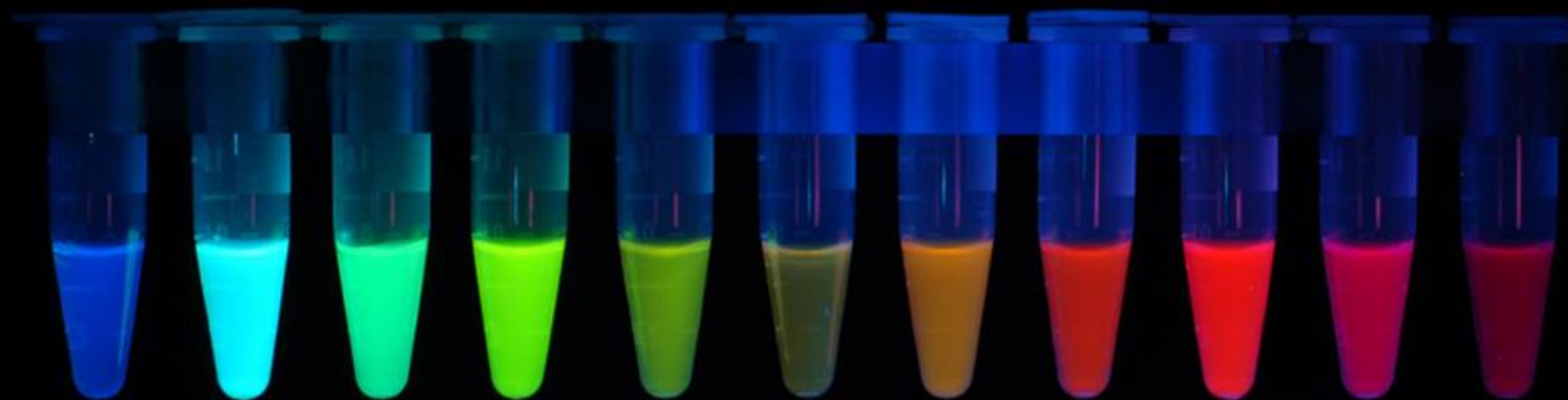
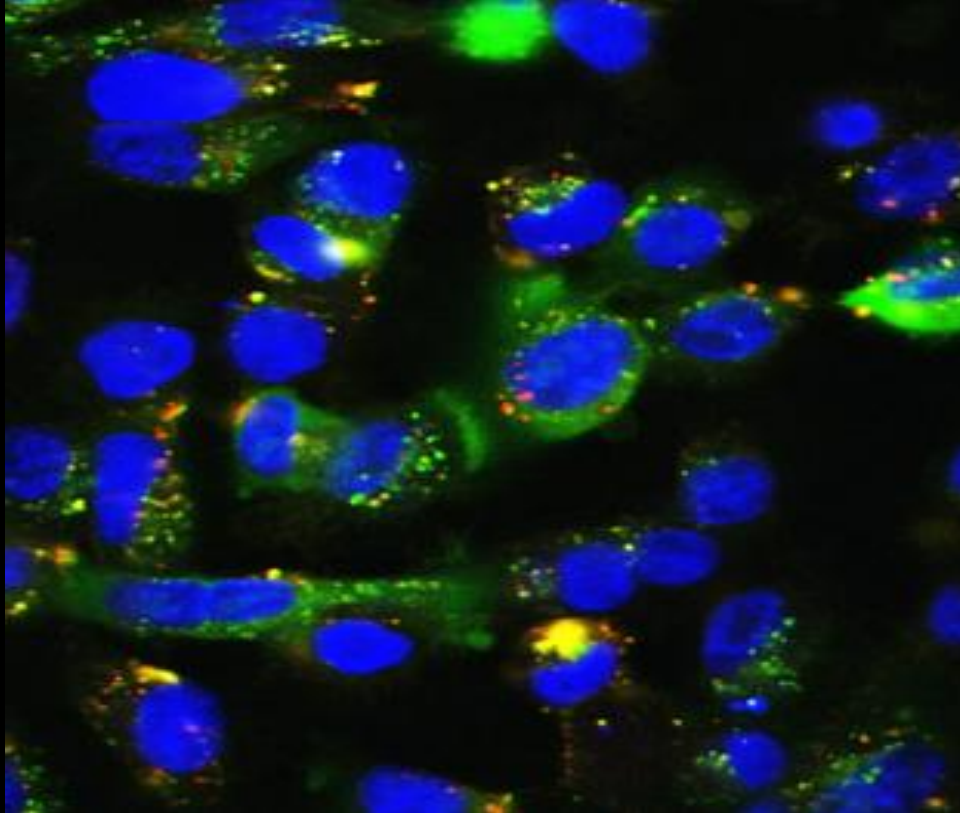
## GFP derivatives

EGFP: Enhanced GFP

YFP: Yellow fluorescent protein

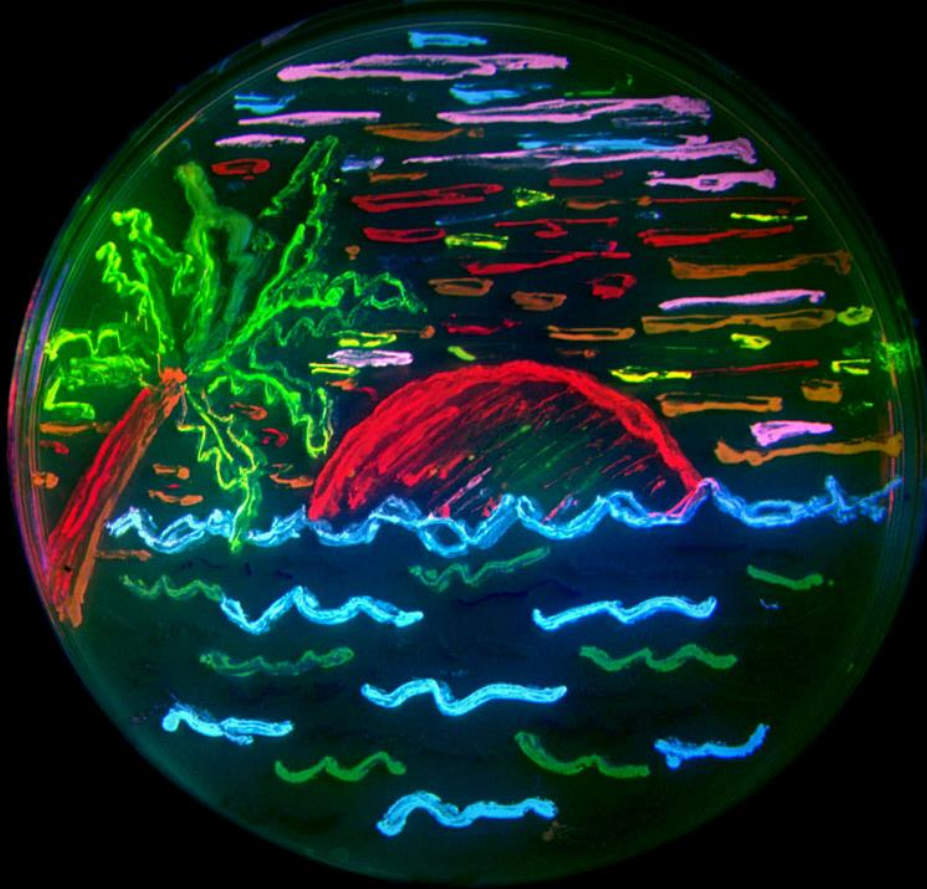
CFP: Cyan fluorescent protein







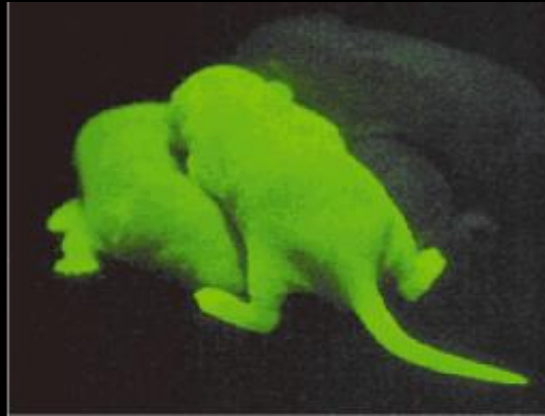
# Bacteria transformed with variants of GFP



visible



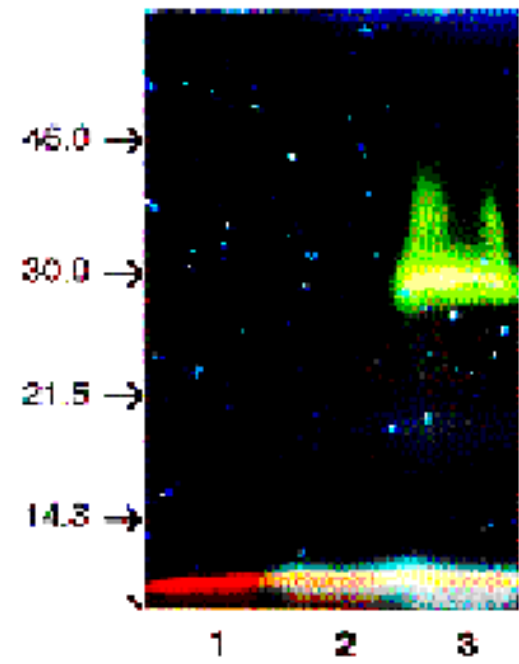
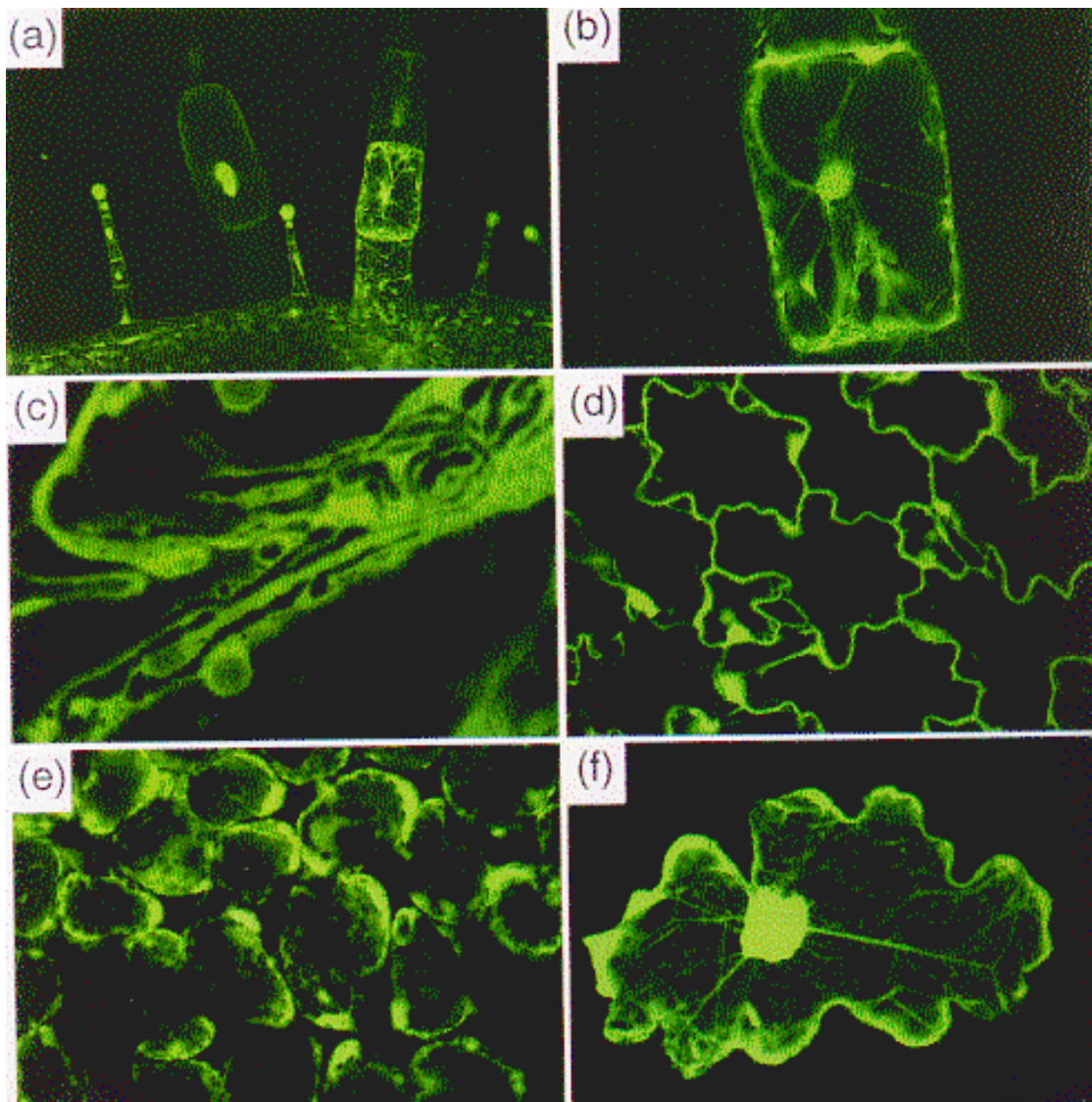
UV







- **GFP is visual marker**
- **Study of biological processes**
- **Regulation of gene expression**
- **Localization of product**
- **Cell transport**
- **Marker for identification of transgenic organisms**



**SDS PAGE**

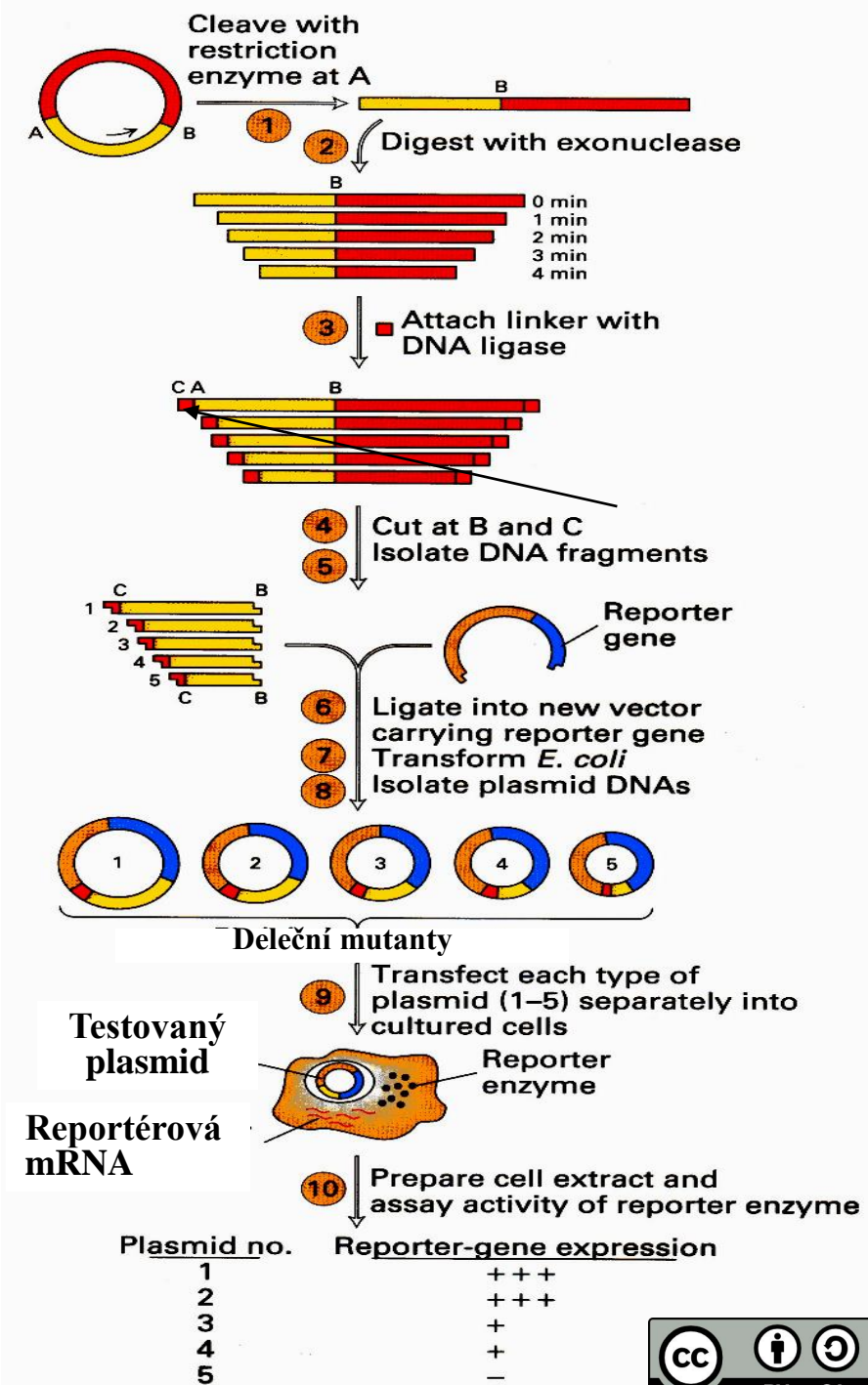
**EXTRACTS FROM TISSUES INFECTED  
WITH PVX (Potato virus) - GFP**

**Confocal laser scanning microscopy**

# REPORTER GENES

## – ACTIVITY of PROMOTERS

### Detection of fusion proteins



Terminální  
delece

Interní  
delece



**Reporter genes** – map the regulatory sequences *in vivo*

# Chloramphenicol acetyltransferase (CAT)

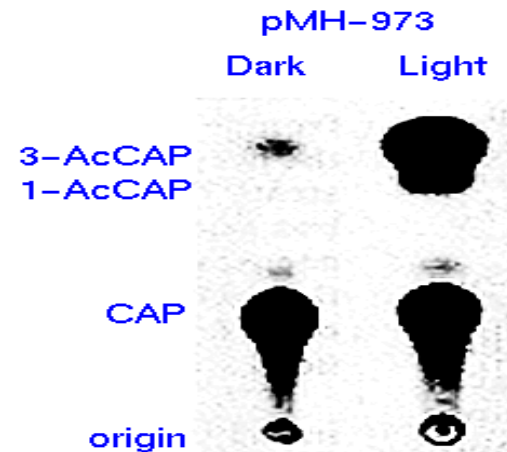
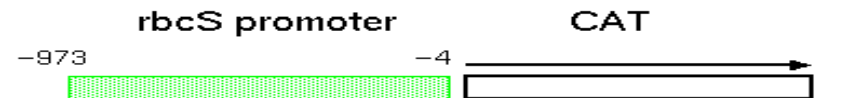
- Transfer of acetyl group from acetylcoenzyme A to chloramphenicol

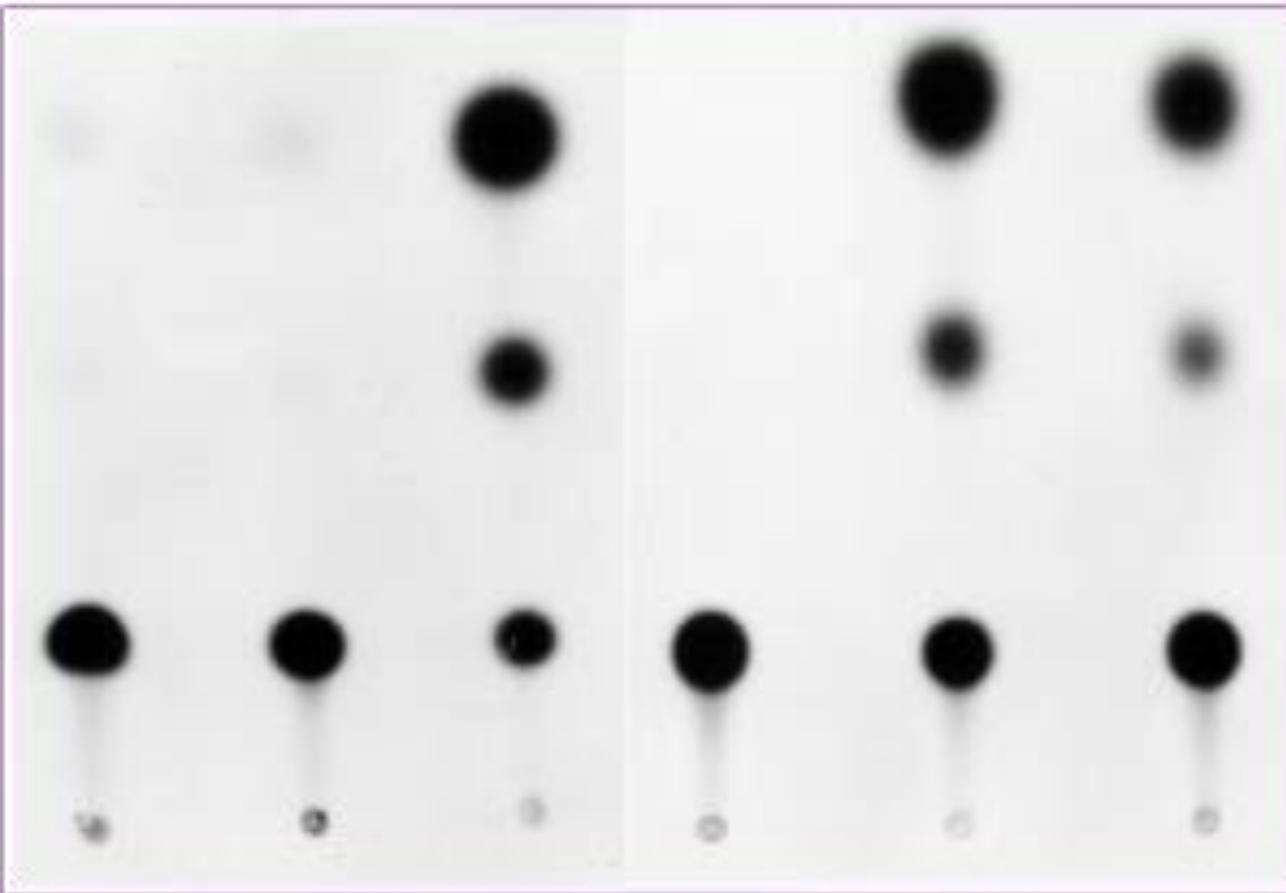
CAT experiment – incubation of lysates with  $^{14}\text{C}$  labeled chloramphenicol

Acetylated and non-acetylated form – thin-layer chromatography (silicone-covered glass plates)

Authoradiography or scintillation of scraped bends

CAP - chloramphenicol





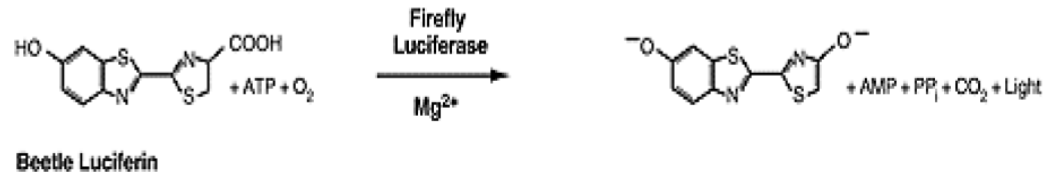
acetylated  
 $^{14}\text{C}$ -chloramphenicol

Unacetylated



# Luciferase

*luc* gene (*Photinus pyralis*)

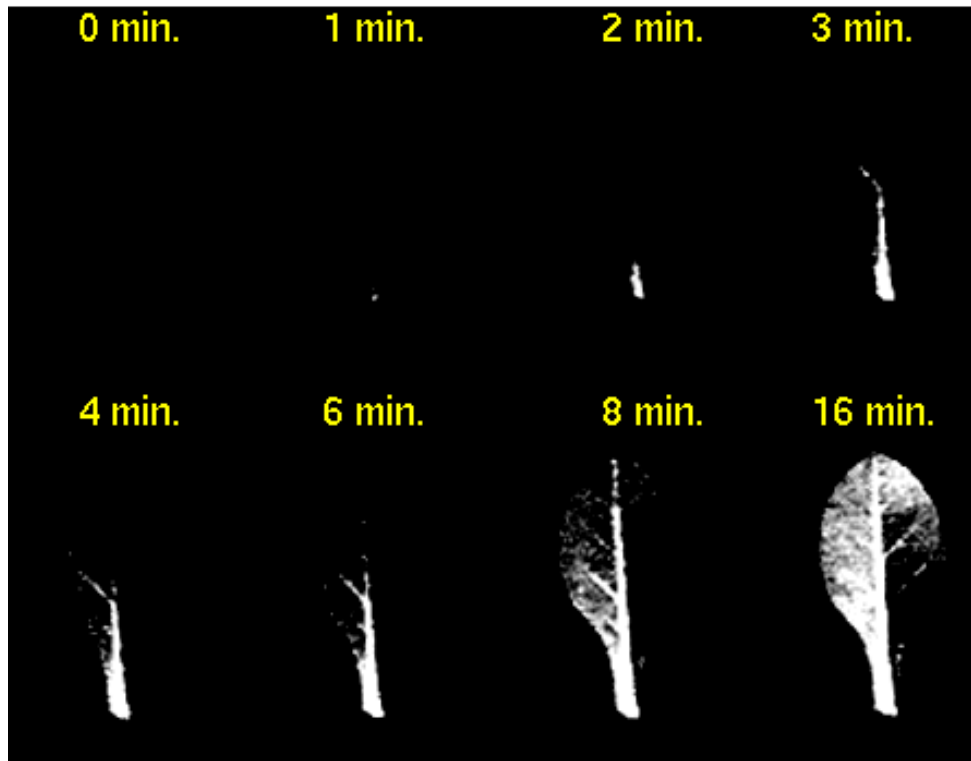


- First non-isotopic reporter system for mammalian cells.

Bioluminescent reaction catalyzed with luciferase

– luciferin, ATP, Mg<sup>2+</sup> and molecular oxygen.

Luciferase – sensitive to proteolysis  
– analysis of inducible promoters



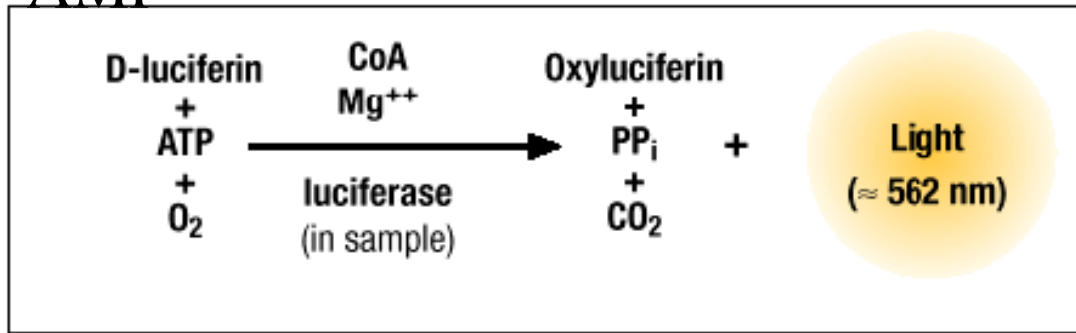


**2-phase reaction** – excitation and fading

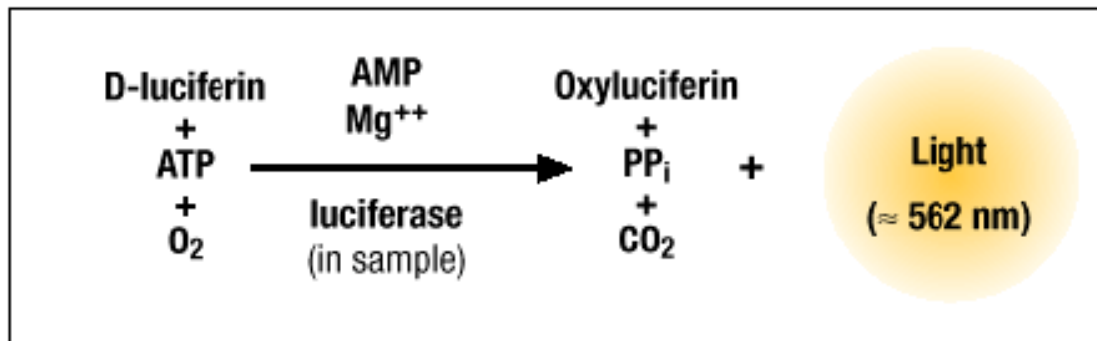
- Various types of experiments:

– addition of substrates and ATP – few seconds of intensive luminiscence

– followed by lower intensity of emission – fading with the use of AMP



**Intense luminiscence**



**Weaker, stable luminiscence - fading**

# **$\beta$ -glucuronidase (GUS)**

## **Fluorescent determination**

**4-MUG (4-methyl-umbeliferyl  $\beta$ -D-glukuronid)**

## **Luminiscent determination**

**adamantyl 1,2-dioxethanarylglukuronid (citlivější)**

## **Chromogenic substrate**

**X-gluc = 5-brom-4-chlorindolylglukuronid**

# Secreted reporter proteins

**Secreted alkaline phosphatase (SEAP)**

- chromogenic substrate p-nitrophenylphosphate

**Human growth hormone (hGH)**

- antibody labeled with radioactive iodine  $^{125}\text{I}$

# Specific fluorescent probes for organelles

<i>Probe</i>	<i>Site</i>	<i>Excitation</i>	<i>Emission</i>
BODIPY	Golgi	505	511
NBD	Golgi	488	525
DPH	Lipid	350	420
TMA-DPH	Lipid	350	420
Rhodamin 123	Mitochondria	488	525
DiO	Lipid	488	500
diI-Cn-(5)	Lipid	550	565
diO-Cn-(3)	Lipid	488	500

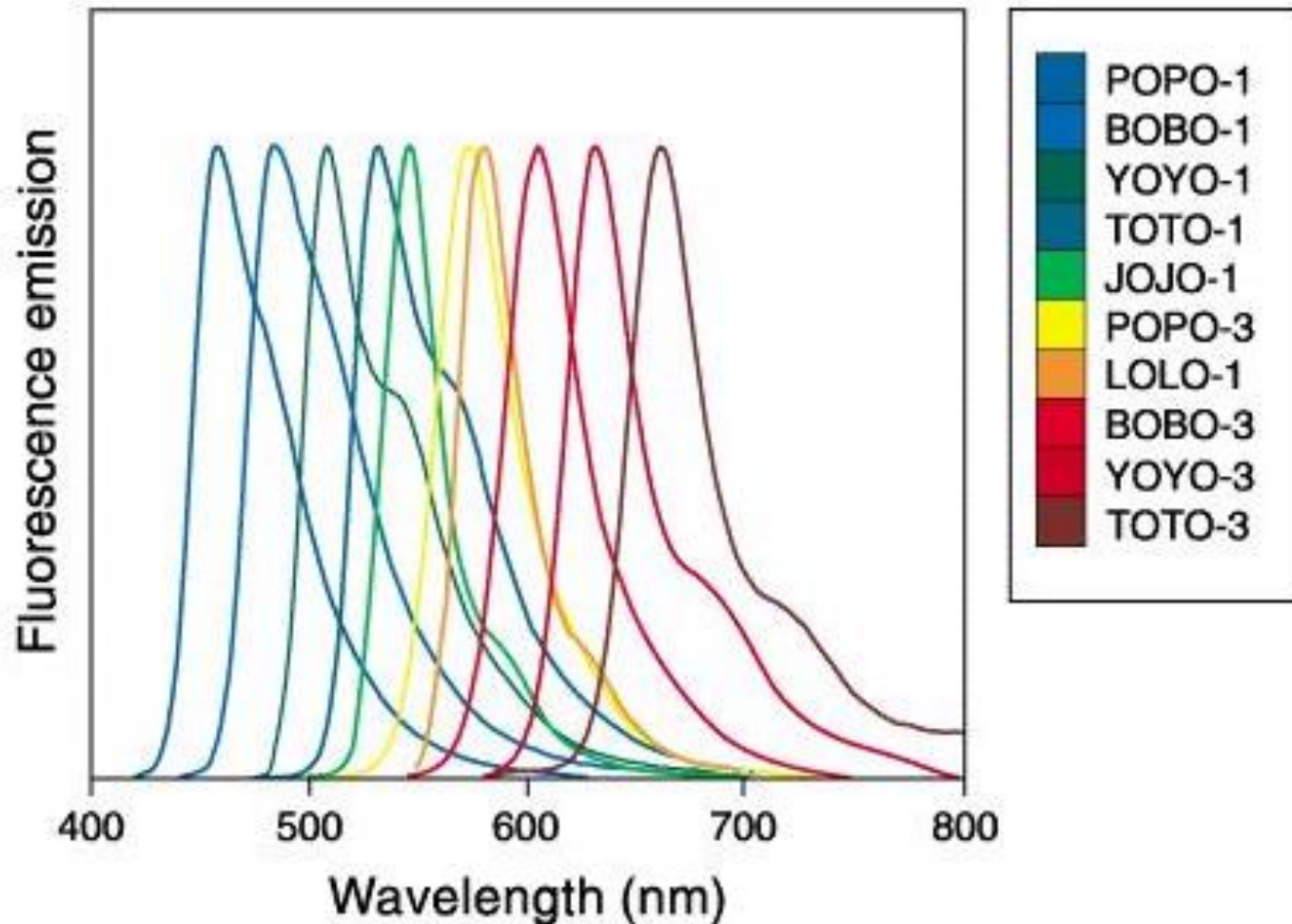
*BODIPY - borate-dipyrromethene complexes*

*NBD - nitrobenzoxadiazole*

*DPH - diphenylhexatriene*

*TMA - trimethylammonium*

# New generation DNA dyes (Molecular probes)



# Translation of purified mRNA

<b>Translation mixture</b>
<b><math>^{35}\text{S}</math>-Met (10,2mCi/mL;1175 Ci/mmol)</b>
<b>Purified RNA</b>
<b>10X buffer</b>
<b>DEPC-water</b>

## Inhibition of RNases

RNase free chemicals, gloves, autoclave,

DEPC – diethylpyrocarbonate (0,1%)

Inhibitors of RNases, work on ice

## ***E. coli* cell free system**

Crude extract (30S)

Endogenous mRNA depleted

Simple translation system

**Cheap!!!**

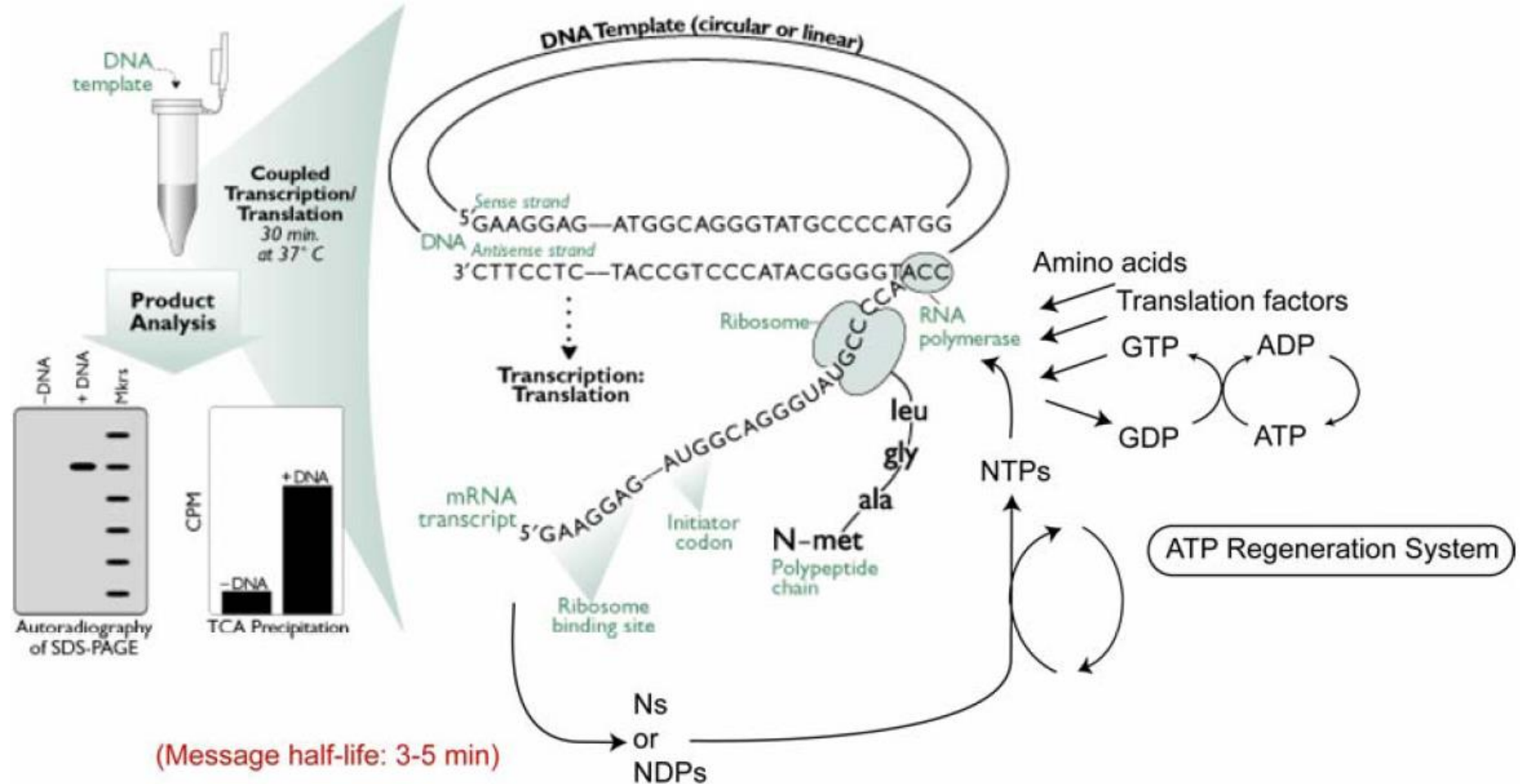
Problem:

Endogenous RNases degrade exogenous mRNA

Conjugated Transcription/Translation

Stabilising elements (hairpins) on mRNA





Problem: Phosphate released by ATP hydrolysis

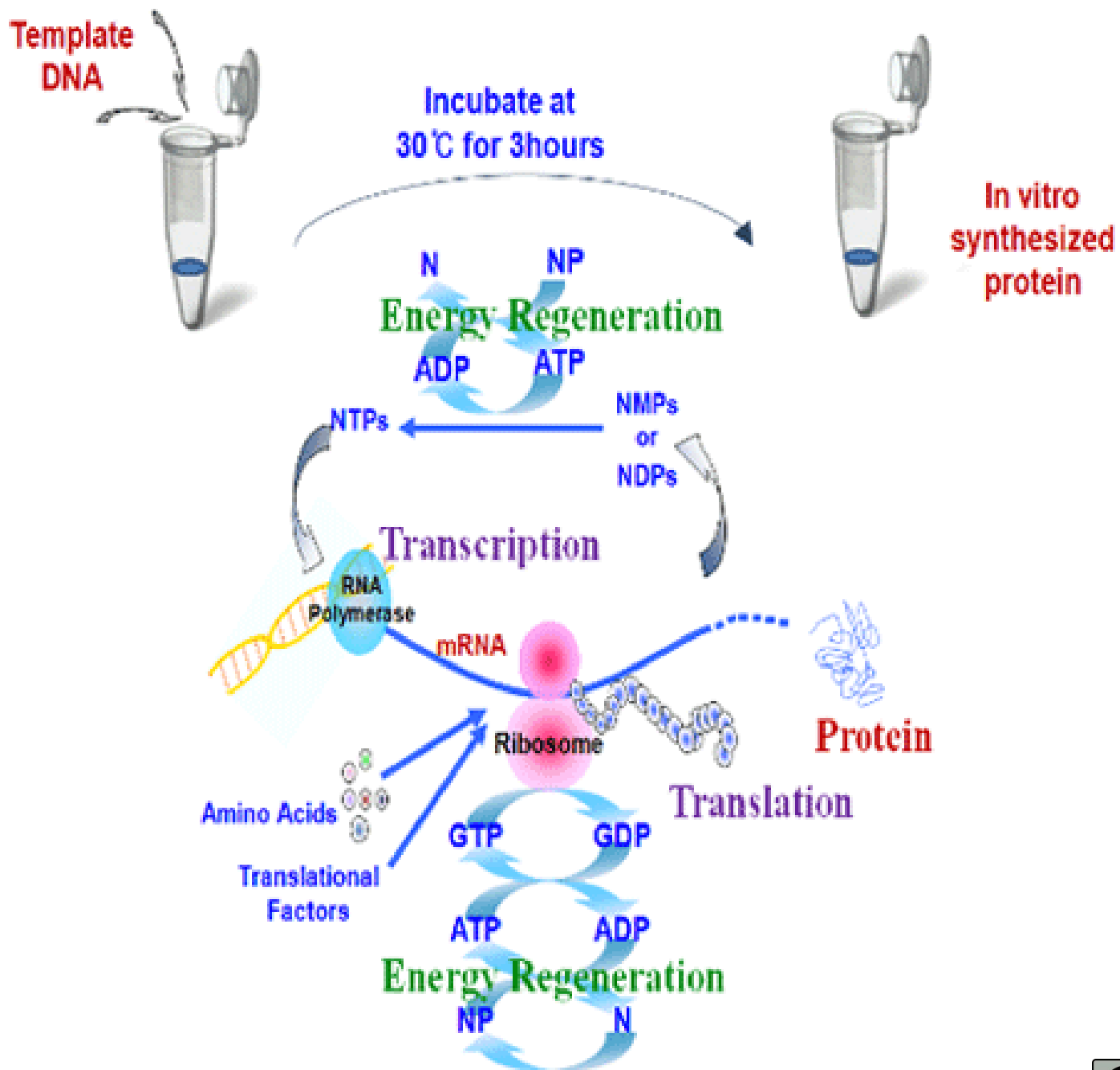
– inhibition of proteosynthesis

ATP regeneration required

Incomplete products

Suitable for proteins < 60 kDa

**Up to 0,5 mg/ml recombinant protein (batch)**



# Rabbit reticulocyte lysate

Efficient eukaryotic *in vitro* proteosynthetic system

Reticulocytes: immature red blood cells without nuclei, full translation machinery

Specialized on synthesis of hemoglobine

(Hb ~ 90% total protein)

Endogenous globin mRNA depleted with microccous nuclease

## **Heterologous proteins**

- synthesized with speed comparable to intact reticulocytes

Translated: capped (eukaryotic) and uncapped (viral) RNA

Necessary: Kozak konsensus and polyA signal

Synthesize uncut products

# Wheat germ extract

Alternative to rabbit reticulocytes

Low level of endogenous mRNA: low background

Micrococcus nuclease not necessary

Exogenous proteins (mammalian, viral, plant) efficiently synthesized

Uncut products

**Cheap!**



# **Suitable for synthesis of proteins:**

**Toxic to cells**

**Forming inclusion bodies in *E. coli***

**Sensitive to intracellular proteases**

**For structure biology:**

**Selective labeling of isotops for NMR**

**Incorporation of modified amino acids (Se) for crystallography**

Uveřejněné materiály jsou určeny studentům Vysoké školy chemicko-technologické v Praze

jako studijní materiál. Některá textová i obrazová data v nich obsažená jsou převzata

z veřejných zdrojů. V případě nedostatečných citací nebylo cílem autora/ů záměrně poškodit event. autora/y původního díla.

S eventuálními výhradami se prosím obračejte na autora/y konkrétního výukového materiálu, aby

bylo možné zjednat nápravu.

The published materials are intended for students of the University of Chemistry and Technology, Prague as a study material. Some text and image data contained therein are taken from public sources. In the case of insufficient quotations, the author's intention was not to intentionally infringe the possible author(s) rights to the original work.

If you have any reservations, please contact the author(s) of the specific teaching material in order to remedy the situation.



# Study of protein interactions

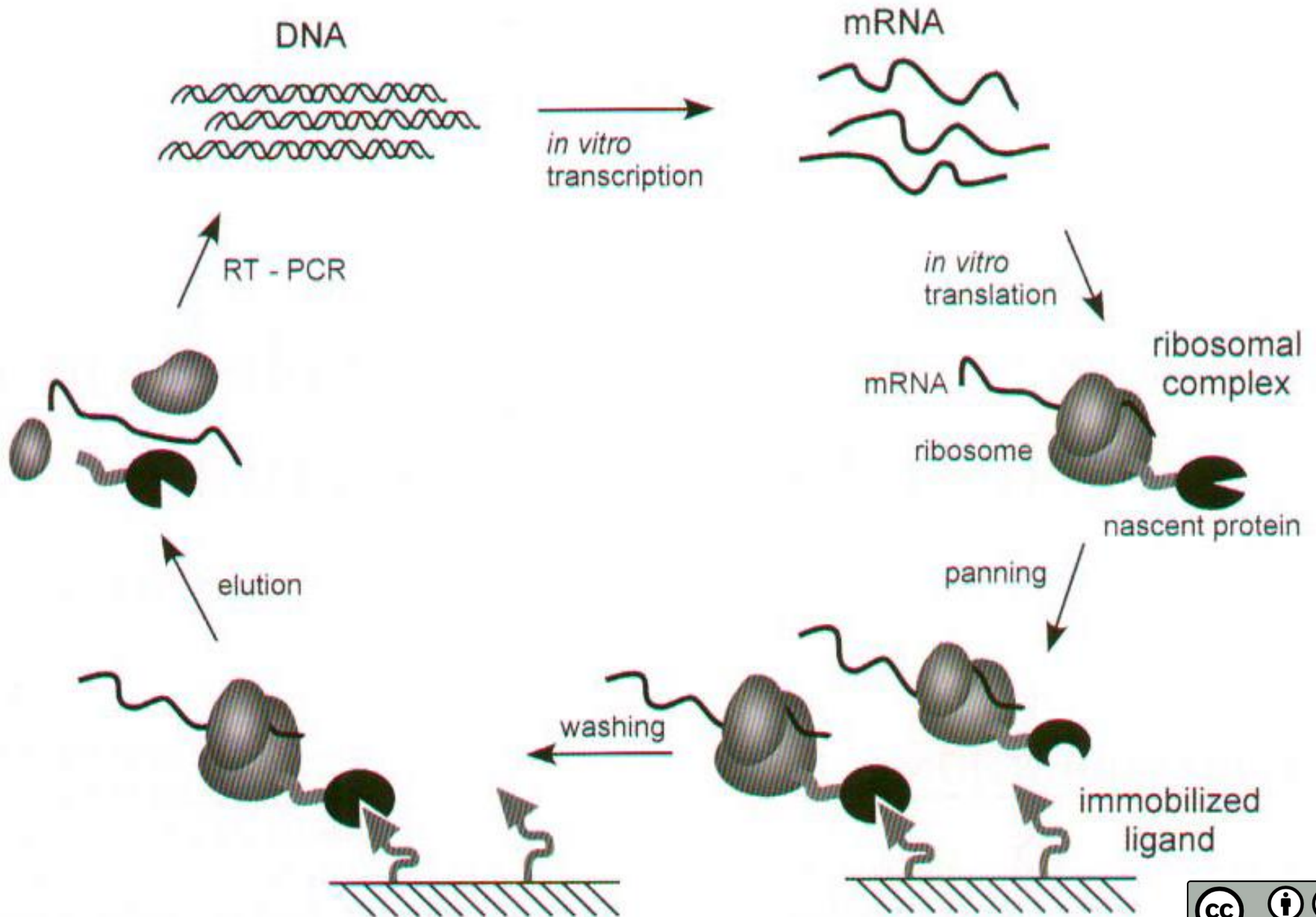


EUROPEAN UNION  
European Structural and Investing Funds  
Operational Programme Research,  
Development and Education



MINISTRY OF EDUCATION,  
YOUTH AND SPORTS

# Ribosomal display

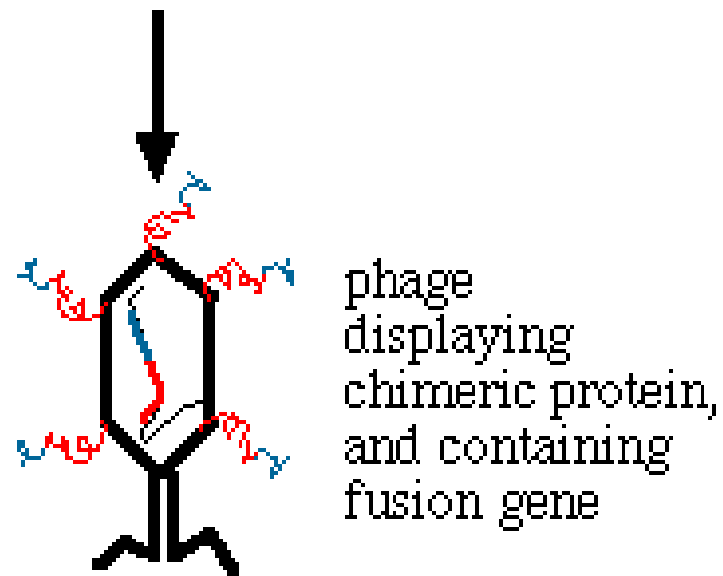
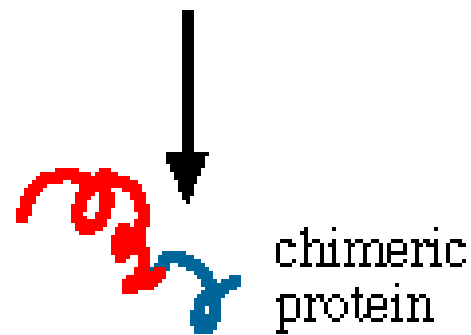
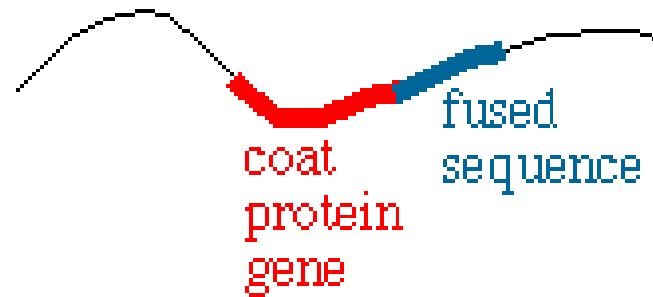


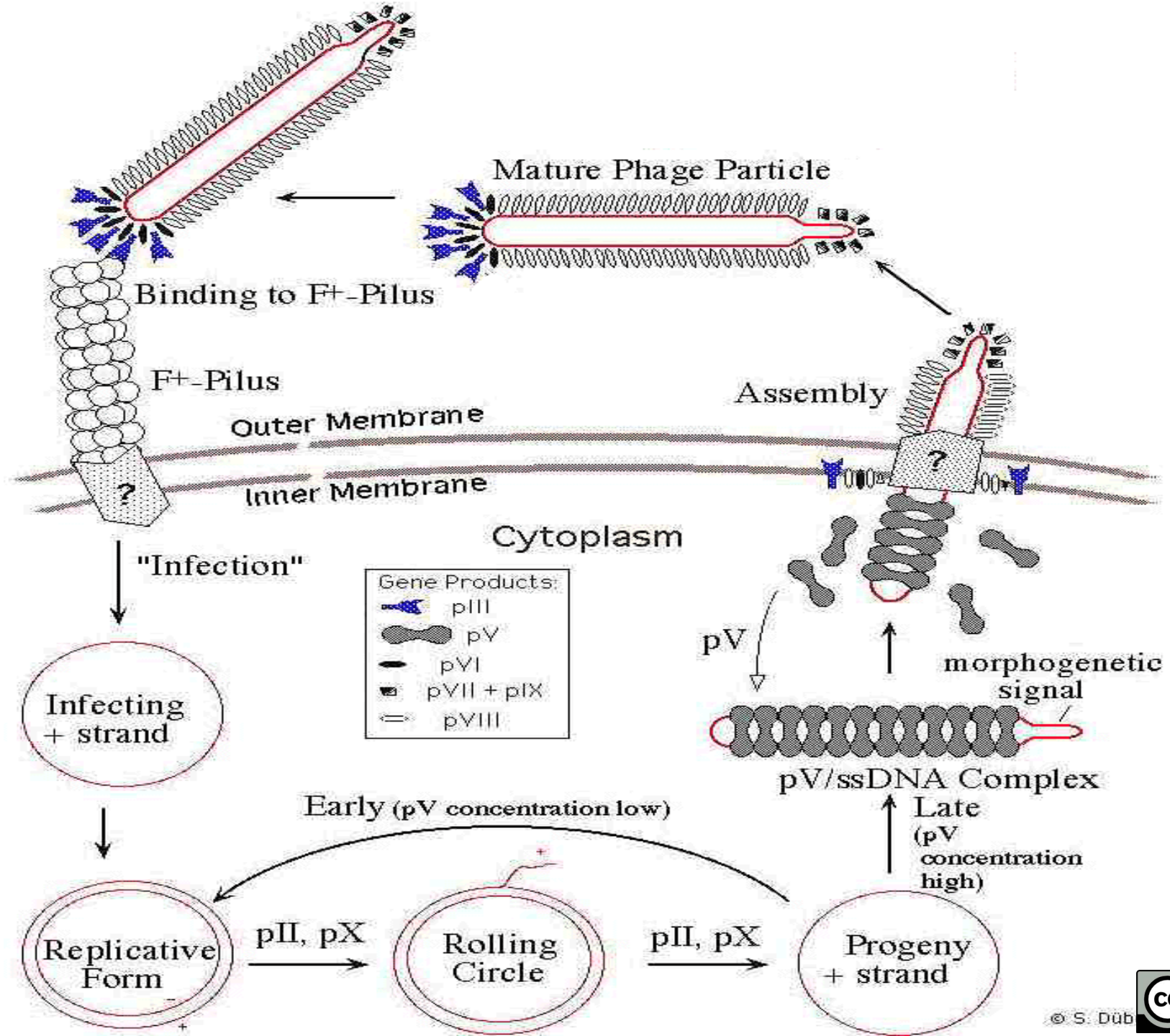
# Phage display

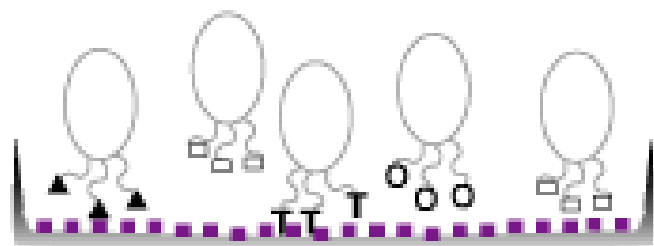
**Expression library of peptides / proteins**

**Expressed in fusion with envelope protein of filamentous phage (e.g. M13)**

T7 phage  
vector

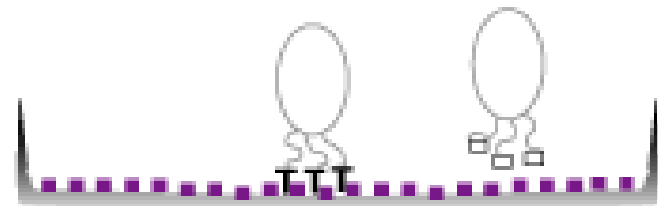






Mixed pool of recombinant phage

Wash-off  
unbound  
phage



Recover phage by  
adding *E. coli* host  
and repeat panning

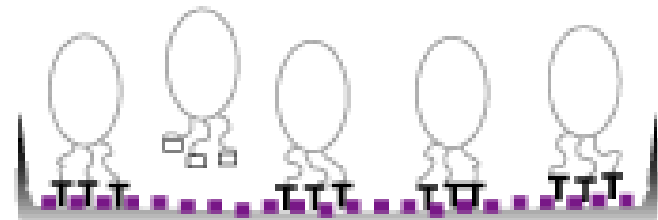
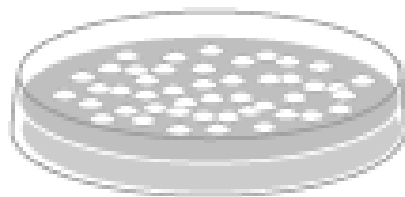
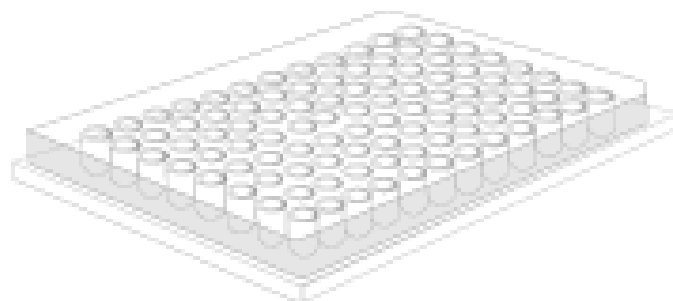


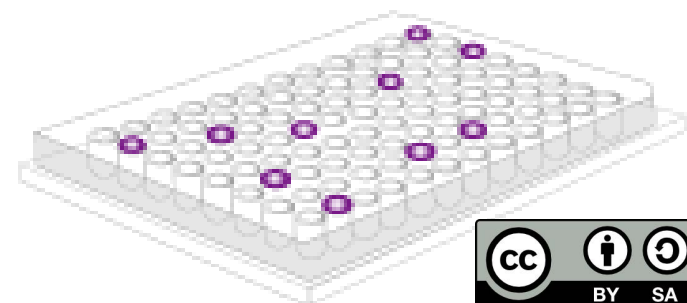
Plate phage-infected  
*E. coli* on LB+amp



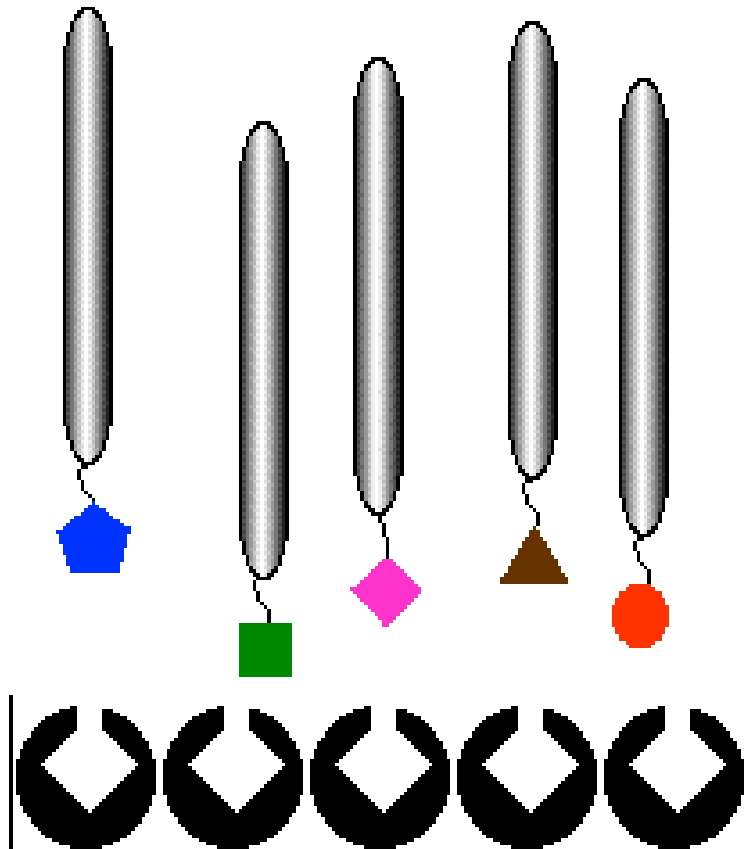
Pick single colonies  
into wells and add  
M13 helper phage



Assay for binding  
activity by ELISAs



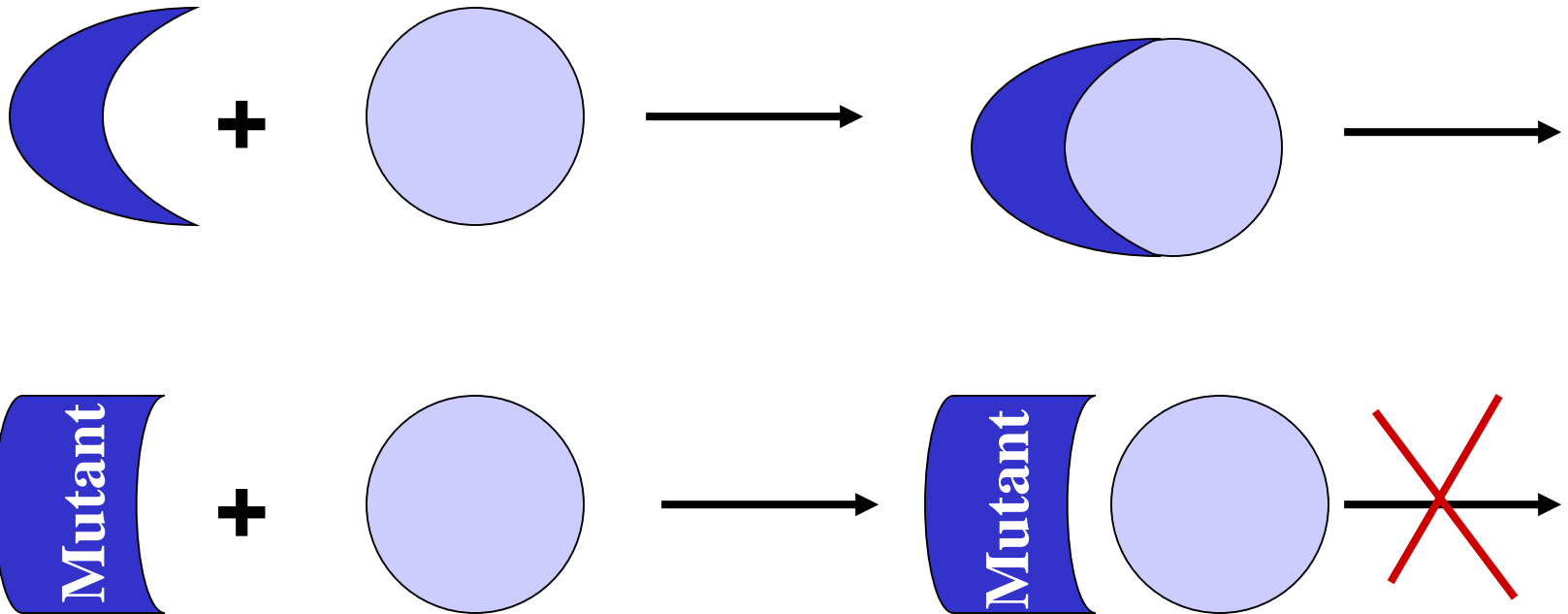
# Model "biopan"



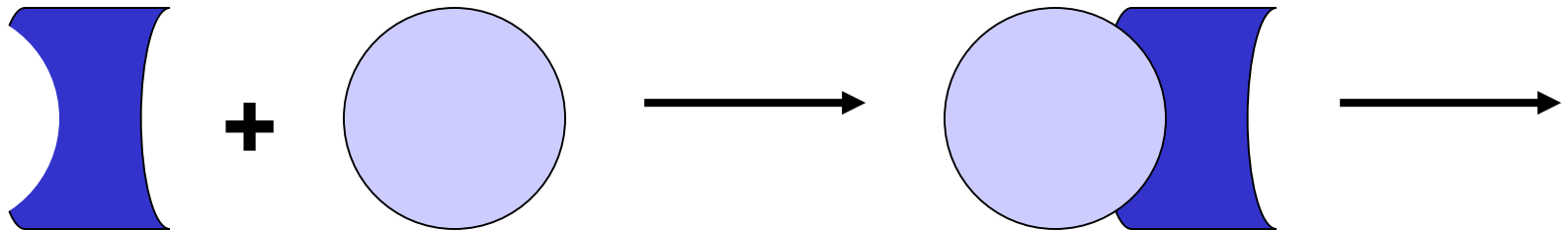
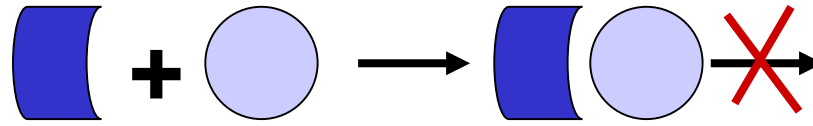
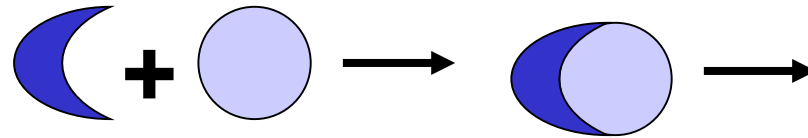


# Study of protein interaction

# Genetic approaches: suppression



# Elimination of the effect of mutation: intragenic suppression



## Co(immuno)precipitation

- direct method – „pull down“
- immuno – agarose beads or *S. aureus* - protein A

## Immunofluorescence

- interaction of proteins - localization in cell
- known proteins – antibodies

## FRET - fluorescence resonance energy transfer

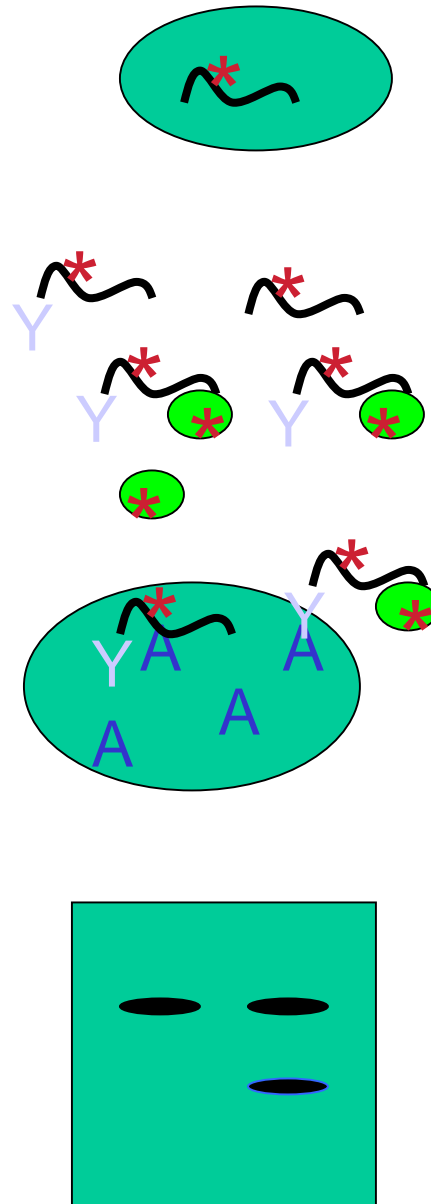
## Chemical crosslink

- stable product on SDS PAGE

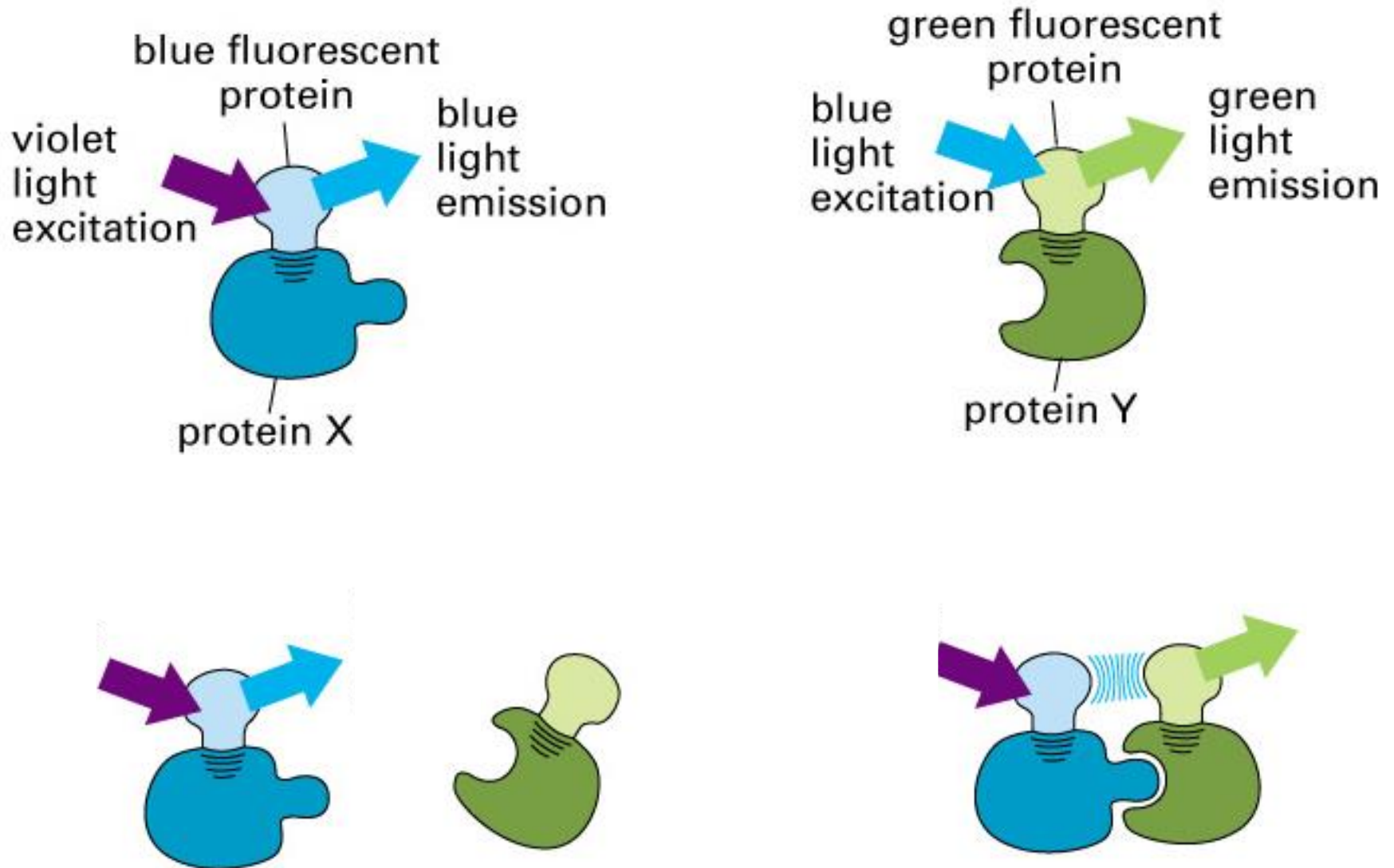
## Far Western

- interaction on membranes

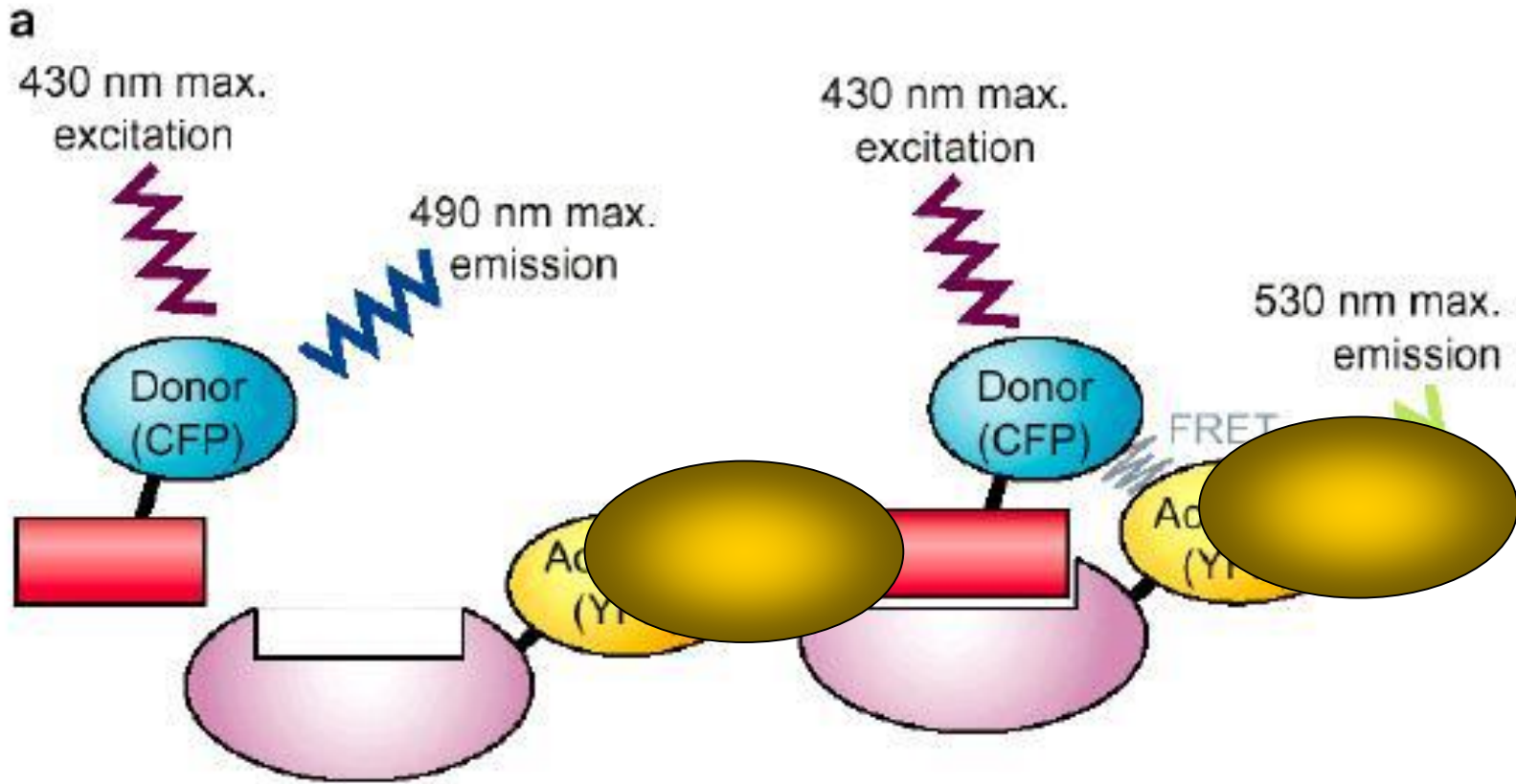
# Coimmunoprecipitation



# Fluorescence resonance energy transfer - FRET

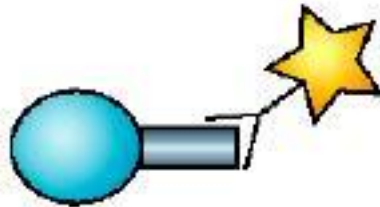


# FRET principle

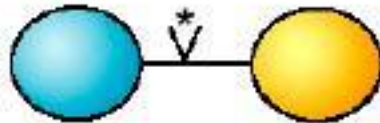




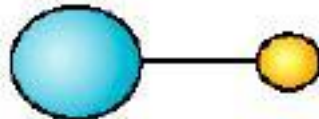
# FRET applications



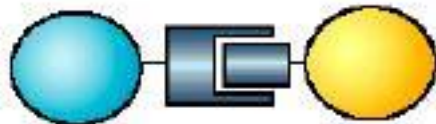
GFP fusion protein and fluorophore-coupled primary antibody undergo FRET<sup>46–48</sup>



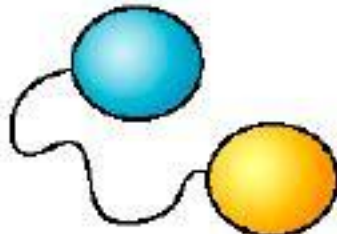
Proteolytic cleavage between two fluorescent proteins eliminates FRET<sup>70–72</sup>



FRET acceptor fluorescent protein is sensitive to chemical environment<sup>62</sup>



GFP fusion proteins interact and FRET<sup>40,41,73–75</sup>



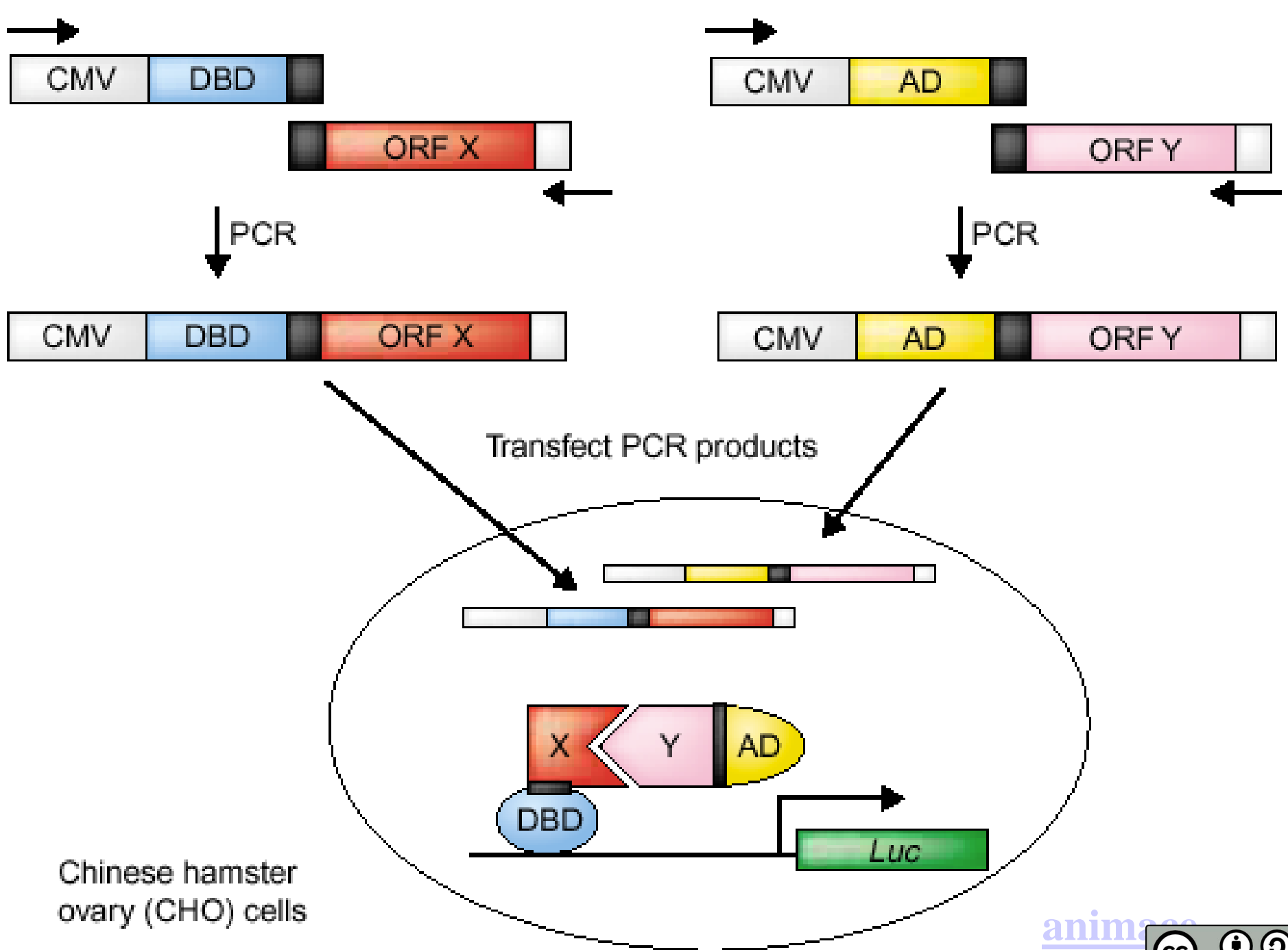
FRET efficiency varies with linker sequence conformation<sup>39,52,66–69</sup>

# **Yeast two hybrid**

**transcription activator – structuraly independent regions**

**(DNA binding domain and activation domain)**

- connection – necessary for transcription activation**
- mediated with proteins of interest**



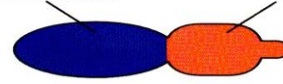
### (a) Hybrid proteins

DNA-binding domain

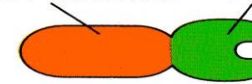
Bait domain

Activation domain

Fish domain

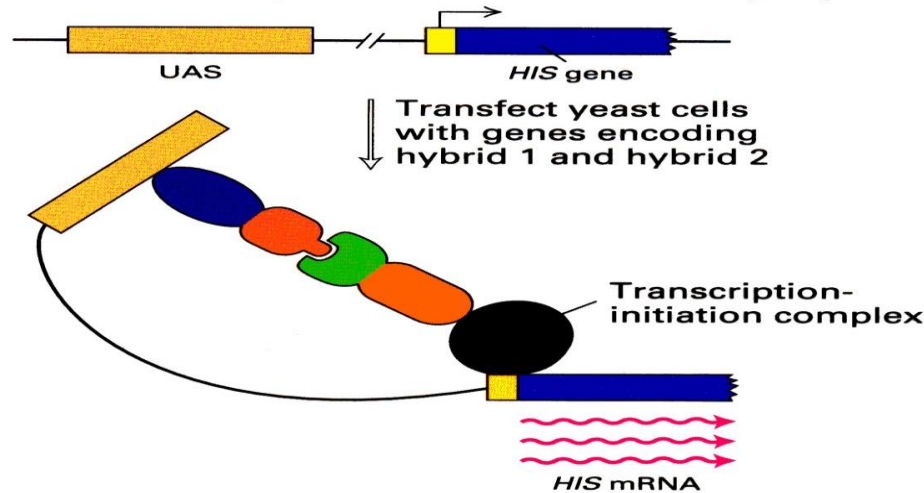


**Hybrid 1**



**Hybrid 2**

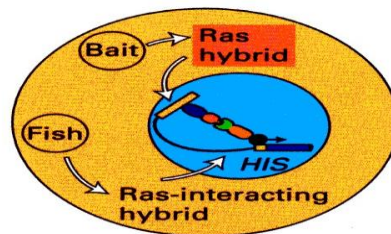
### (b) Transcriptional activation by hybrid proteins in yeast



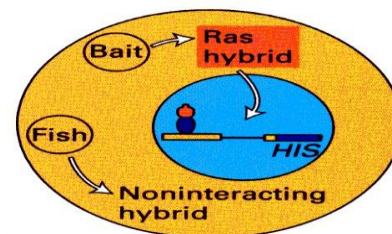
### Selekcce proteinů interagujících s Ras



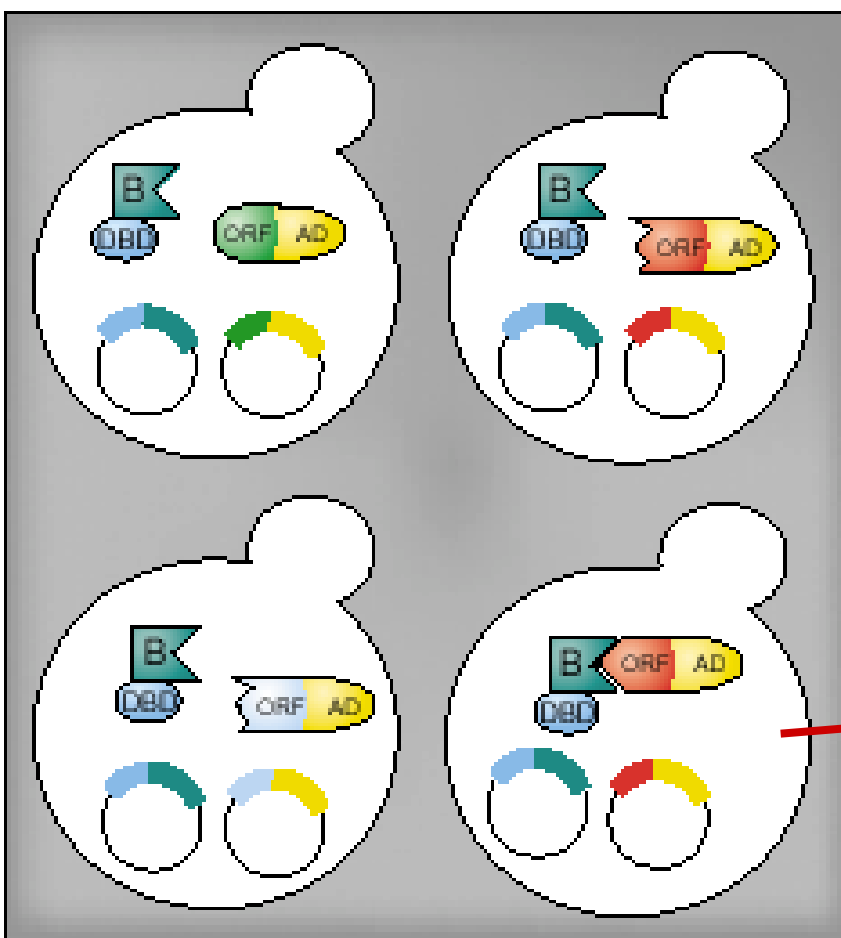
1. Transfect into *trp*, *leu*, *his* mutant yeast cells
2. Select for cells that grow in absence of tryptophan and leucine
3. Plate selected cells on medium lacking histidine



Colony formation



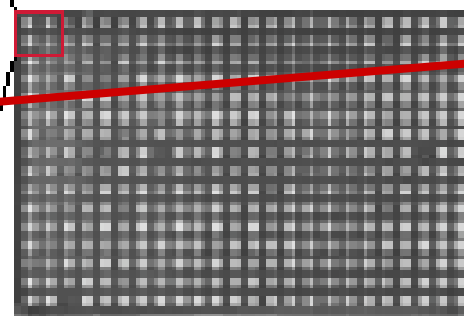
No colony formation



384-Pin replica tool

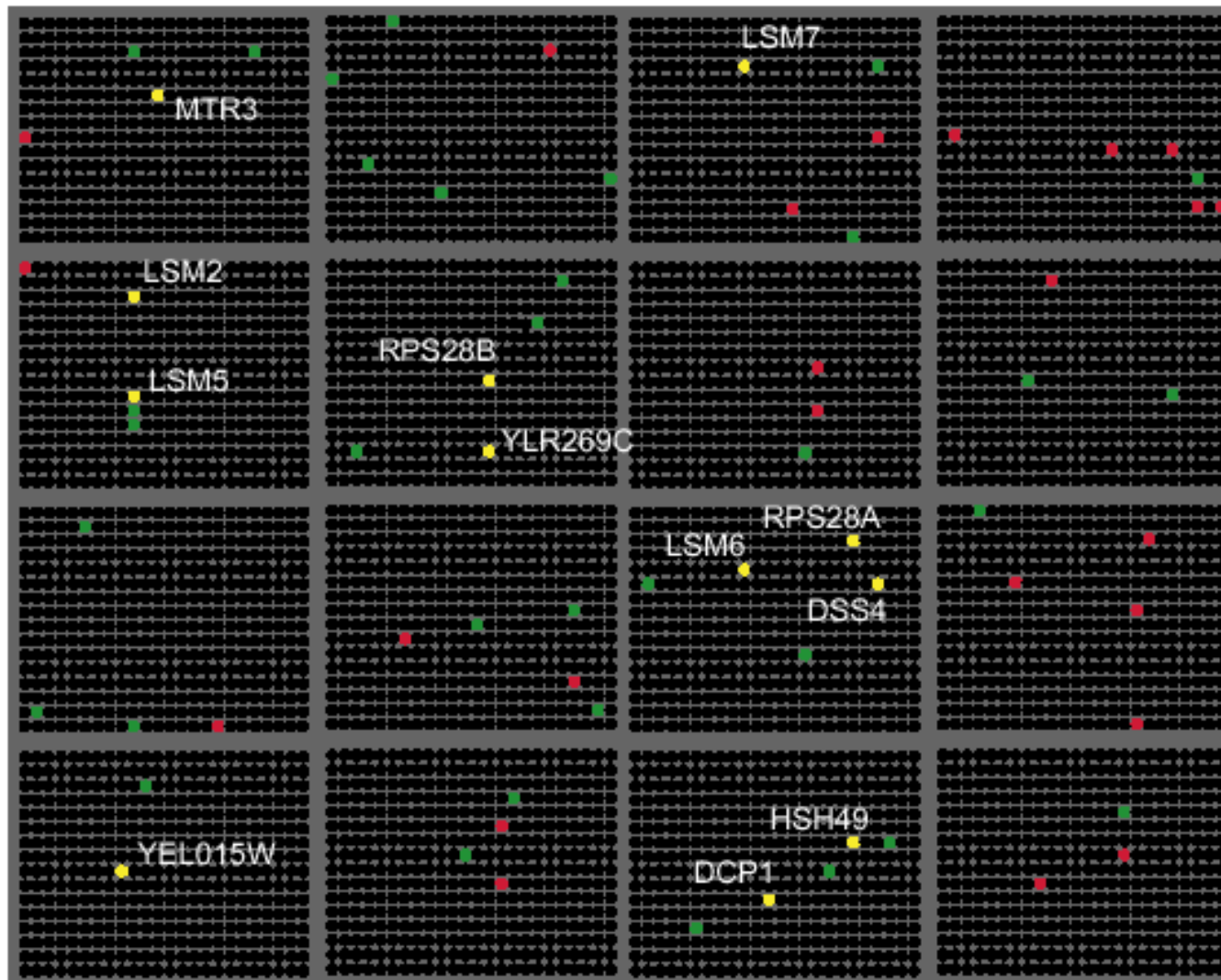


1 Bait + 384 AD-ORF  
diploids



Two-hybrid selective  
plate





**Elimination of false-positive results**

# Proximity ligation assay

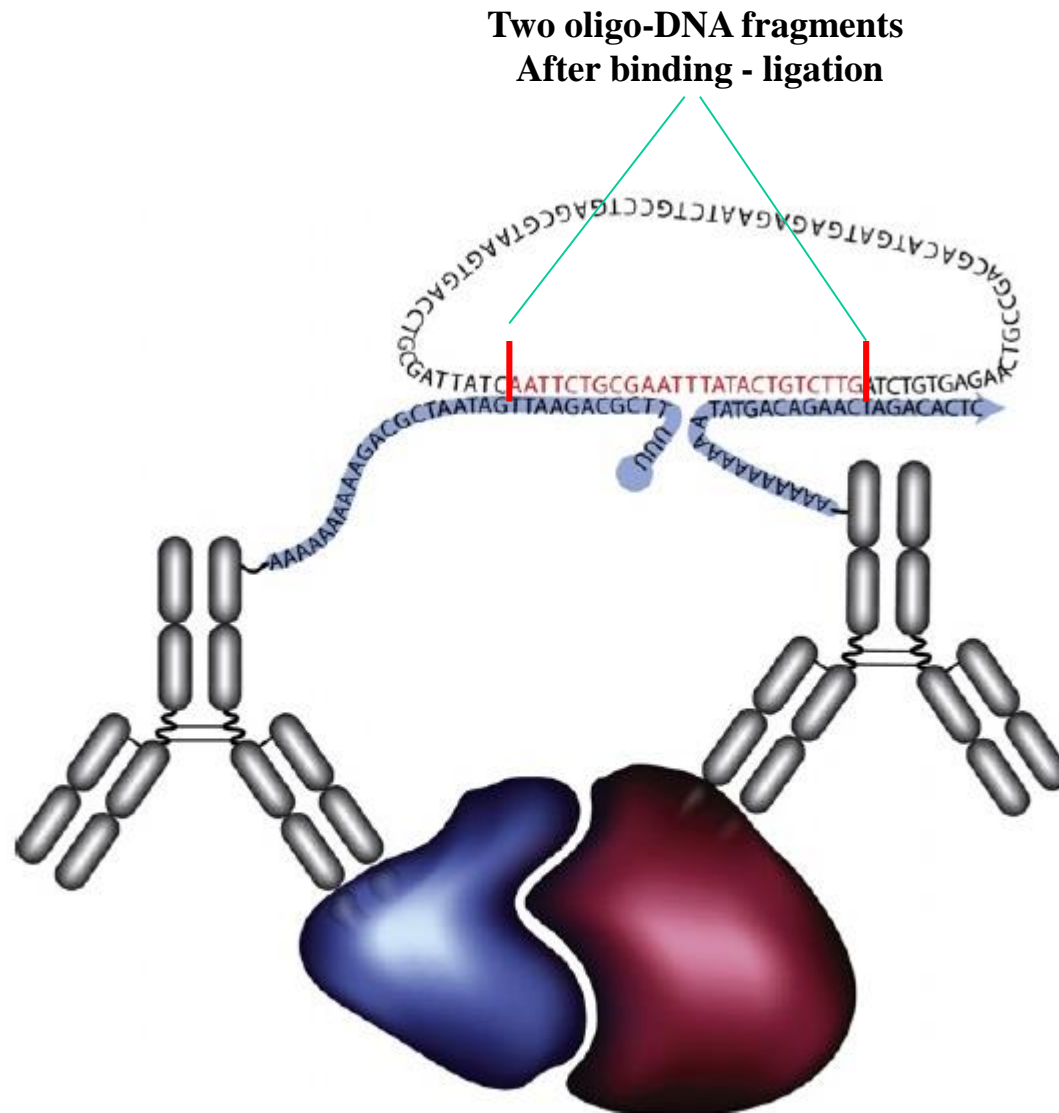
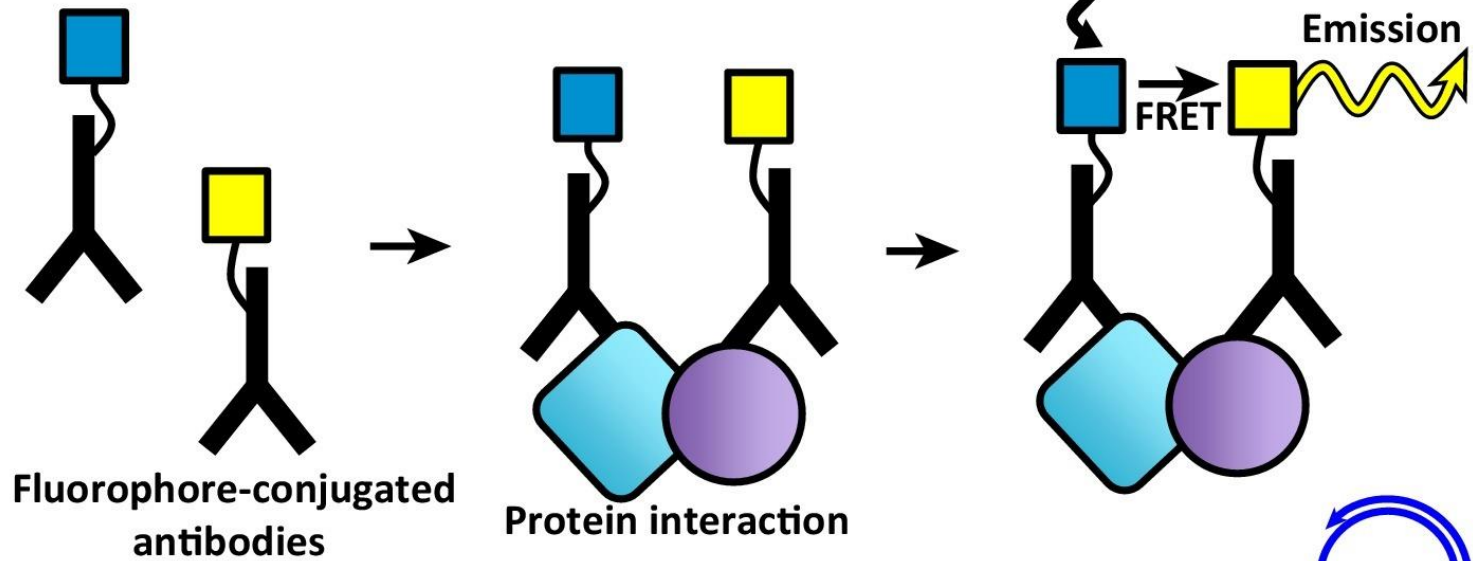


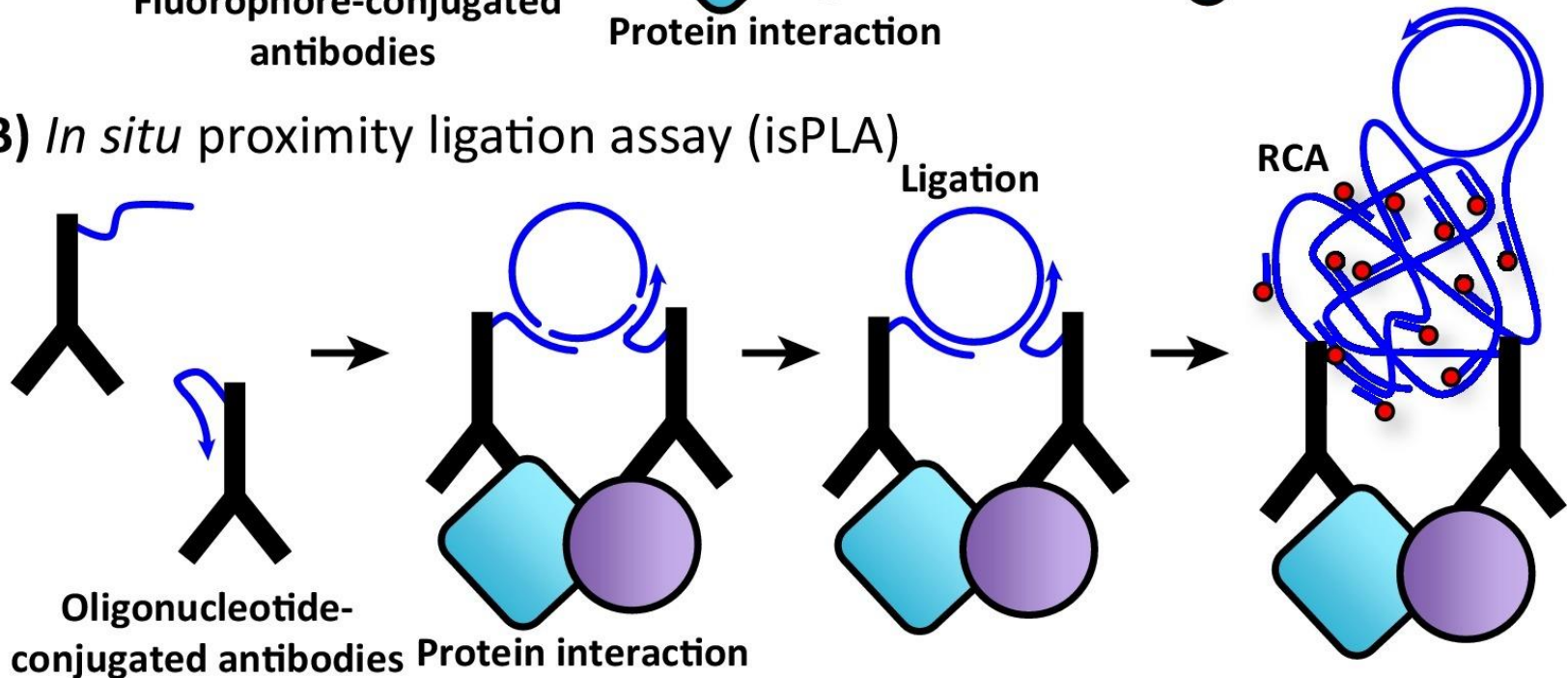
Fig. 1. A presentation of the oligonucleotide design used for *in situ* PLA



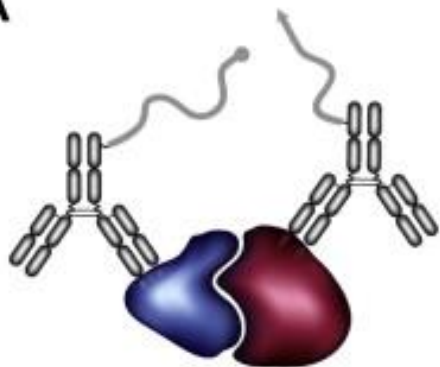
**(A) Förster resonance energy transfer (FRET)**



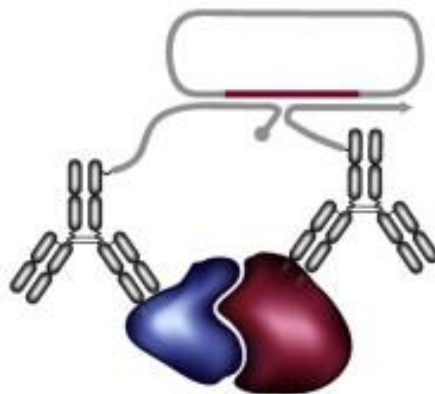
**(B) *In situ* proximity ligation assay (isPLA)**



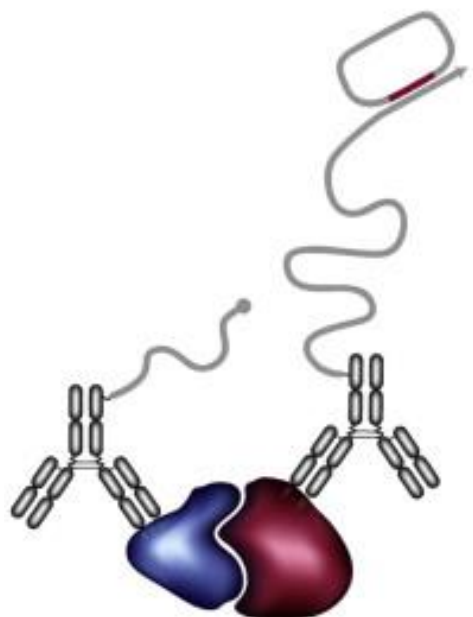


**A**

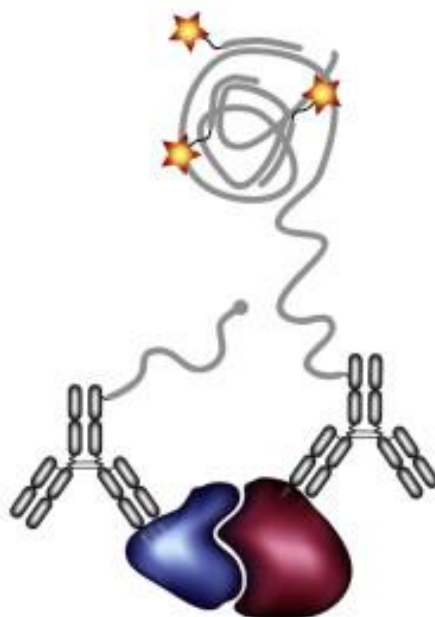
Proximity probe binding



Circularization and ligation of  
connector oligonucleotides



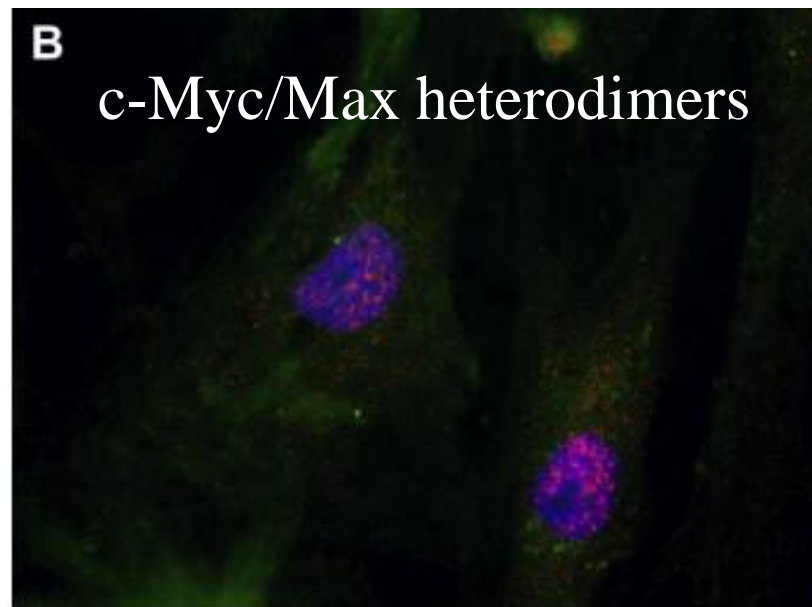
Rolling circle amplification



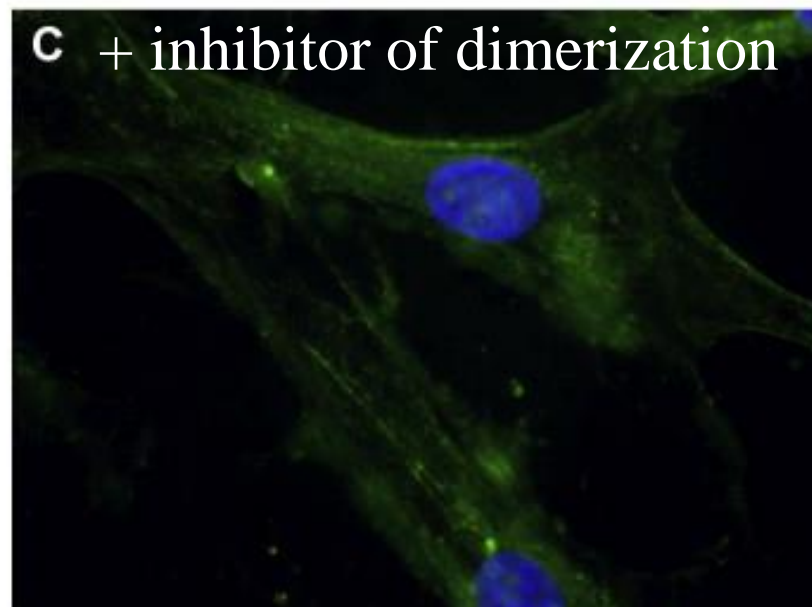
Detection of rolling circle product

**B**

c-Myc/Max heterodimers

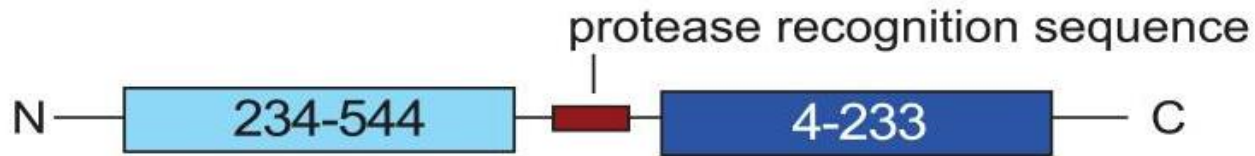
**C**

+ inhibitor of dimerization

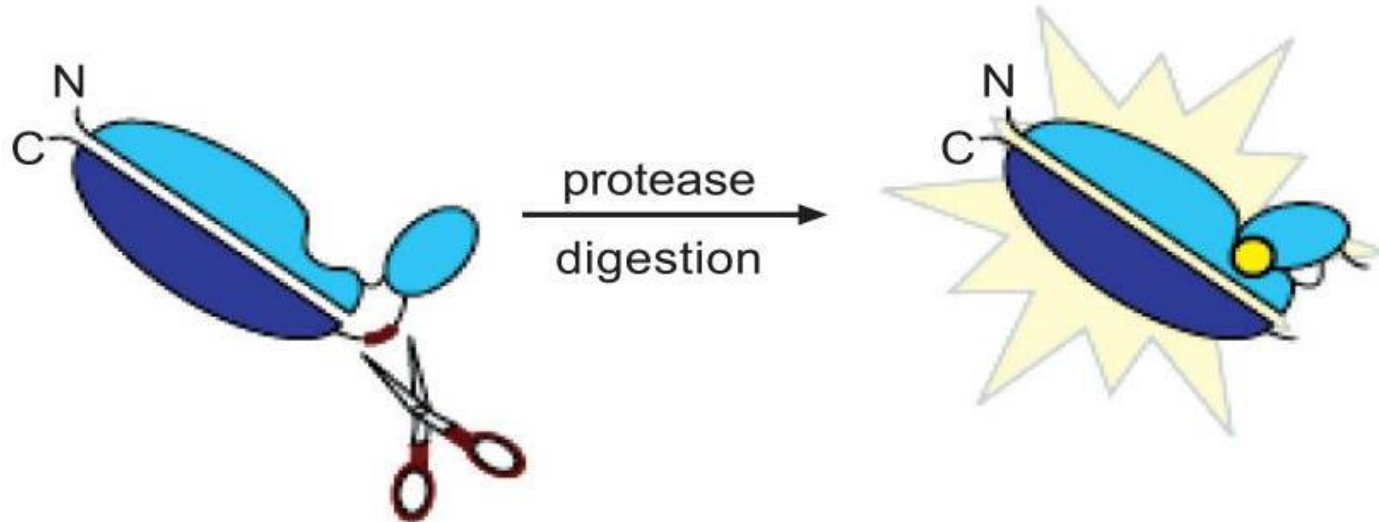


# LinkLight™ technology

A.



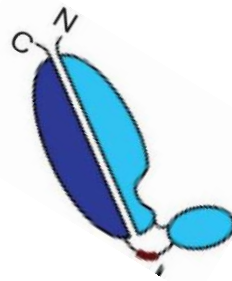
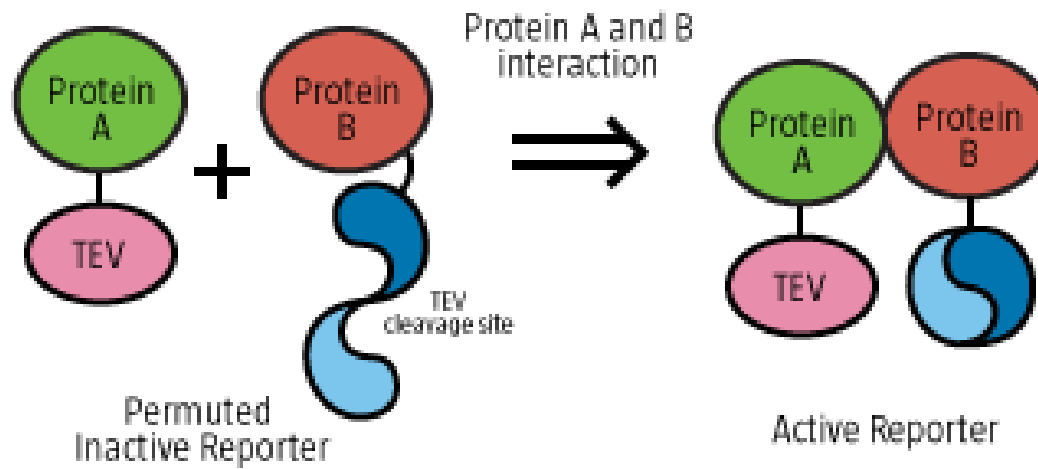
B.



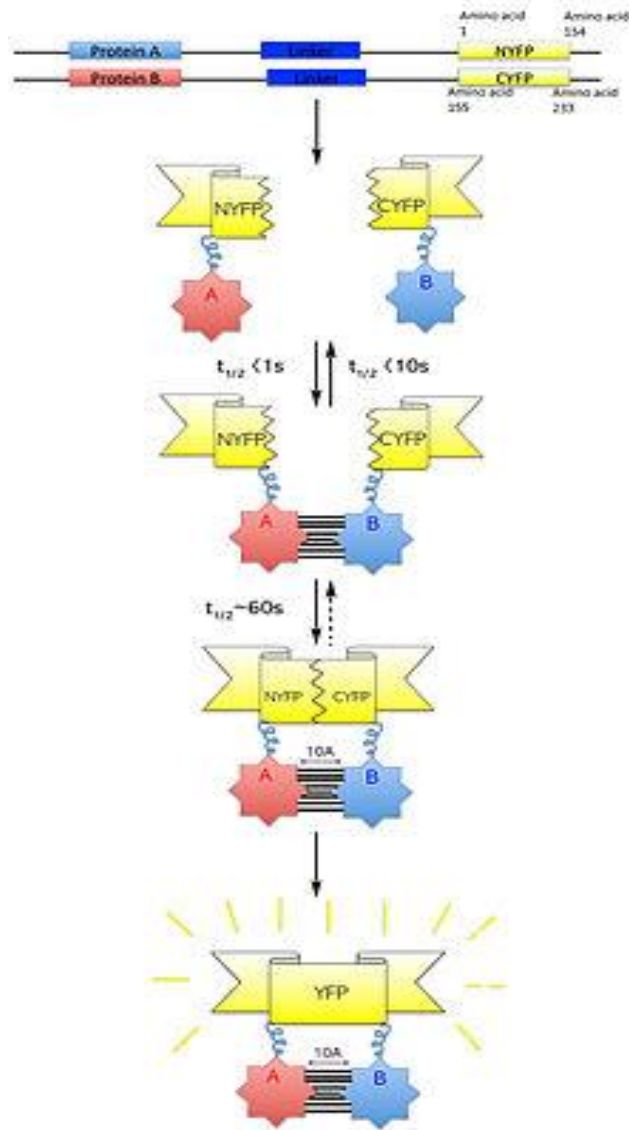
Inactive CP<sub>234</sub>-Luc protein  
(low luciferase activity)

Active CP<sub>234</sub>-Luc protein  
(high luciferase activity)

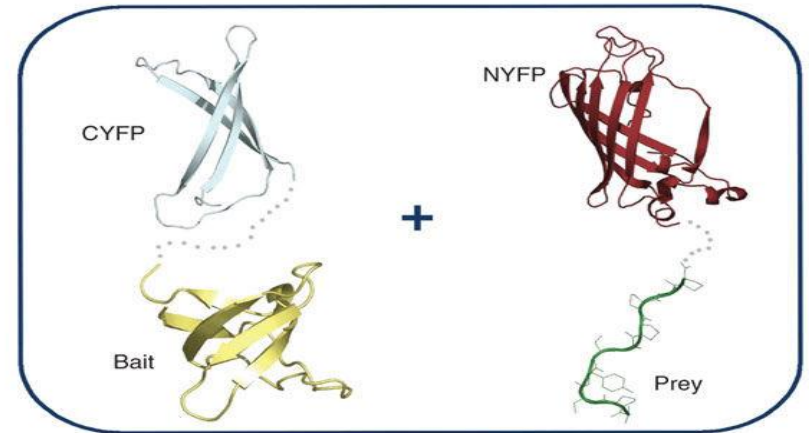
# LinkLight™ technology



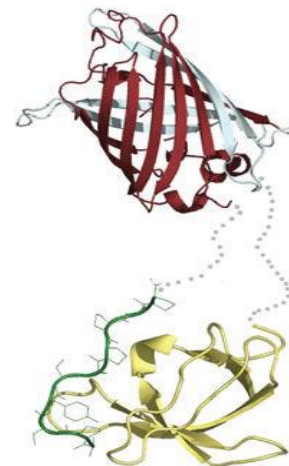
# Two domains of fluorescent proteins



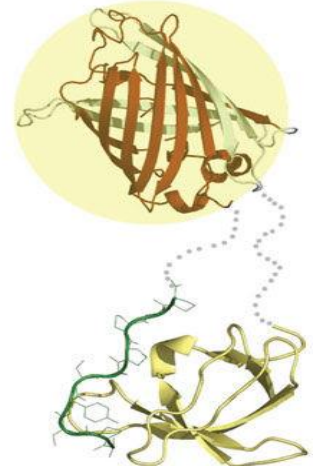
a



b



c

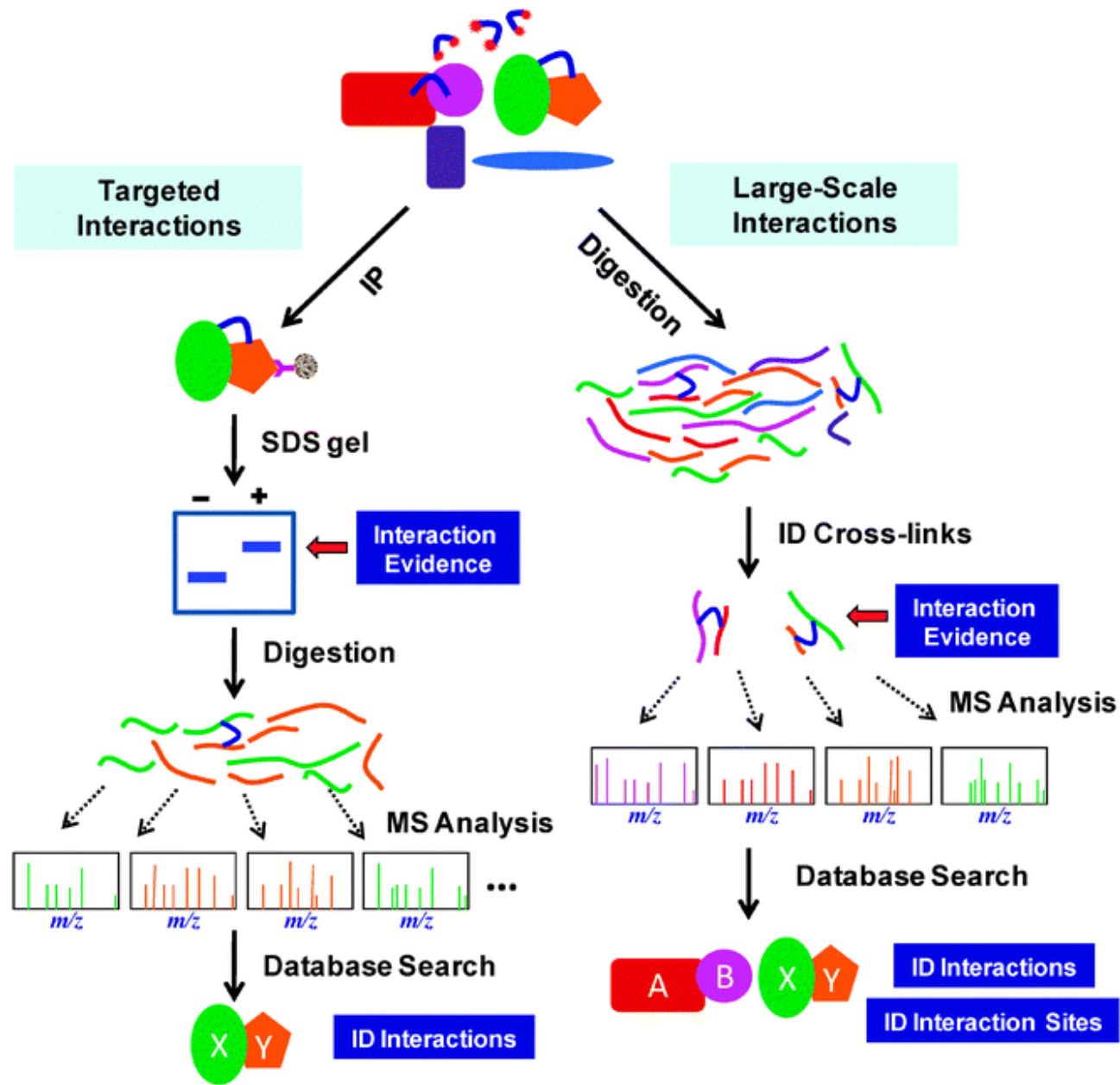


Functional (fluorescing) complex only upon interaction



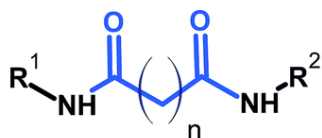
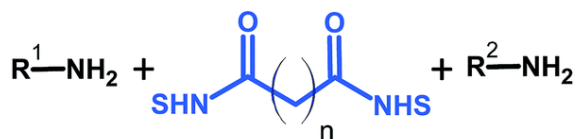


# Chemical crosslink



### AMINE-specific (Lys)

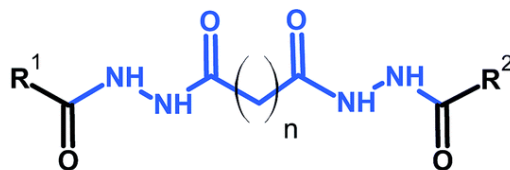
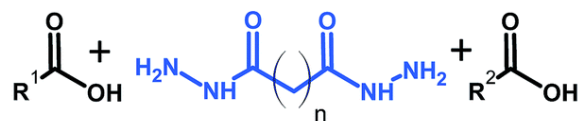
Succinimide esters\*



\*Alternatives: Sulfosuccinimidyl esters, imidates, carbamates

### CARBOXYL-specific (Asp/Glu)

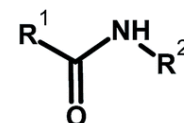
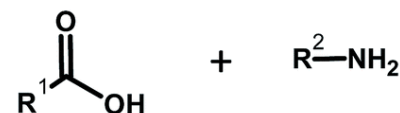
Coupling reagent\*\* + hydrazide



\*\* e.g. DMTMM, EDC

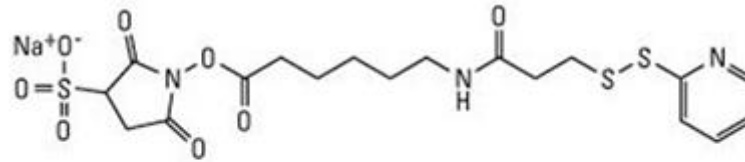
### ZERO-LENGTH (Lys + Asp/Glu)

Coupling reagent\*\* alone



\*\* e.g. DMTMM, EDC

# Homobifunctional probes



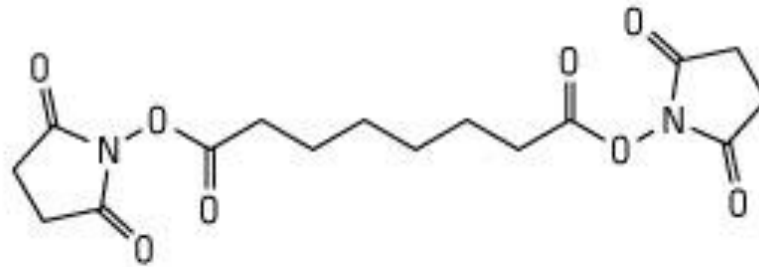
**Sulfo-LC-SPDP**

Sulfosuccinimidyl 6-[3'-(2-pyridyldithio)propionamido]hexanoate

MW 527.57

Spacer Arm 15.7 Å

**Prokřížuje SH skupiny**



**DSS**

Disuccinimidyl suberate

MW 368.34

Spacer Arm 11.4 Å

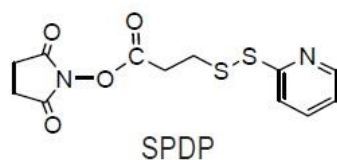
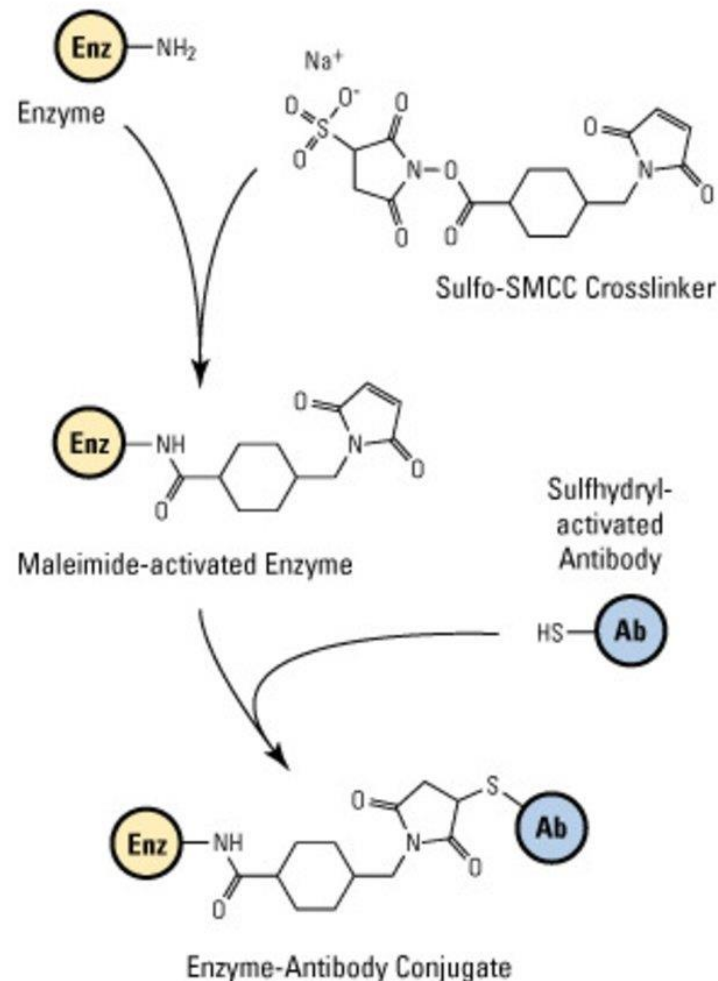
**Prokřížuje NH<sub>2</sub> skupiny**



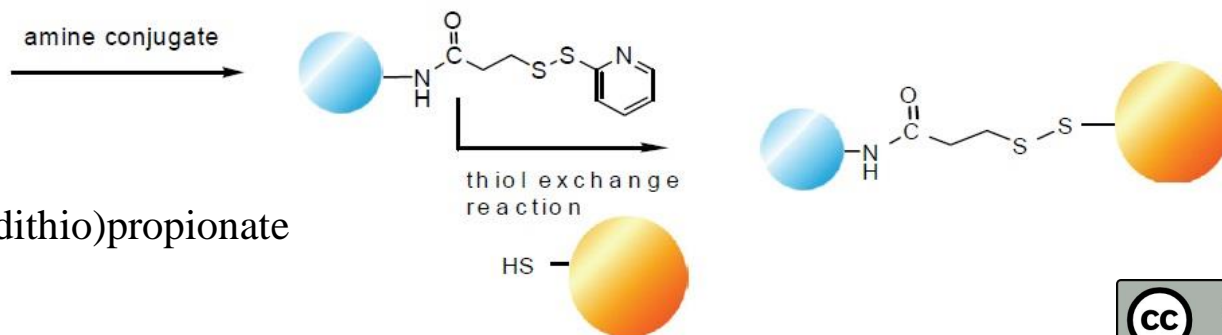
# Heterobifunctional probes

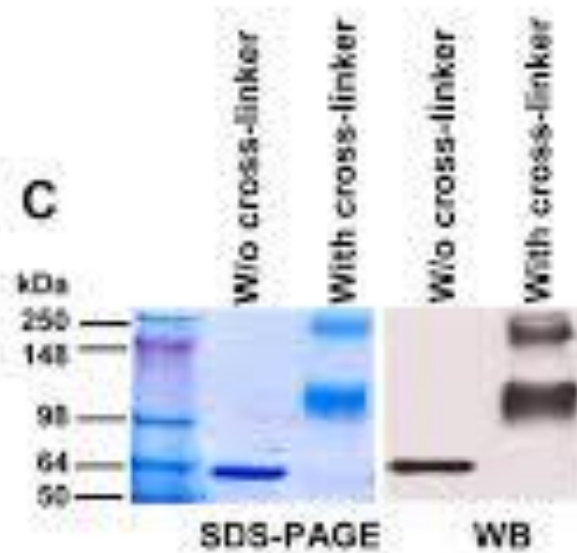
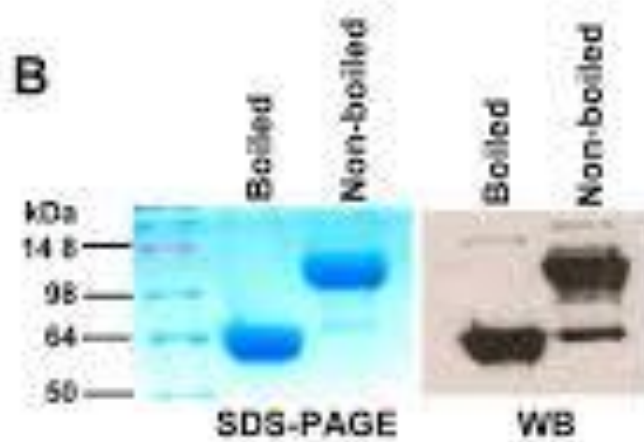
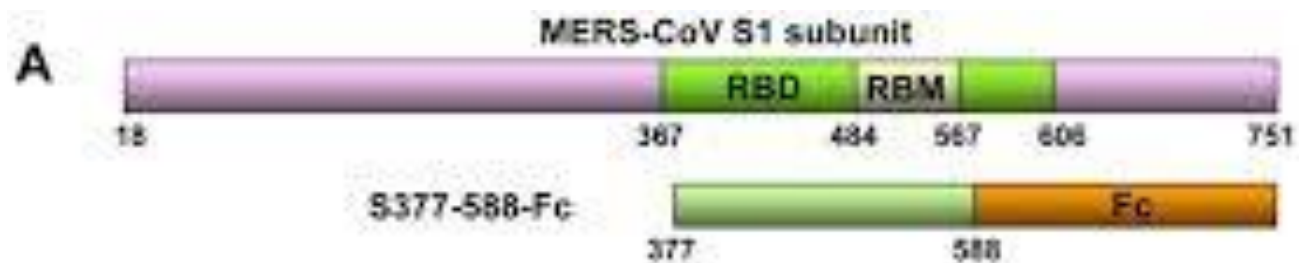
Sulfo-SMCC (sulfosuccinimidyl-4- (N-maleimidomethyl) cyclohexane-1-carboxylate)

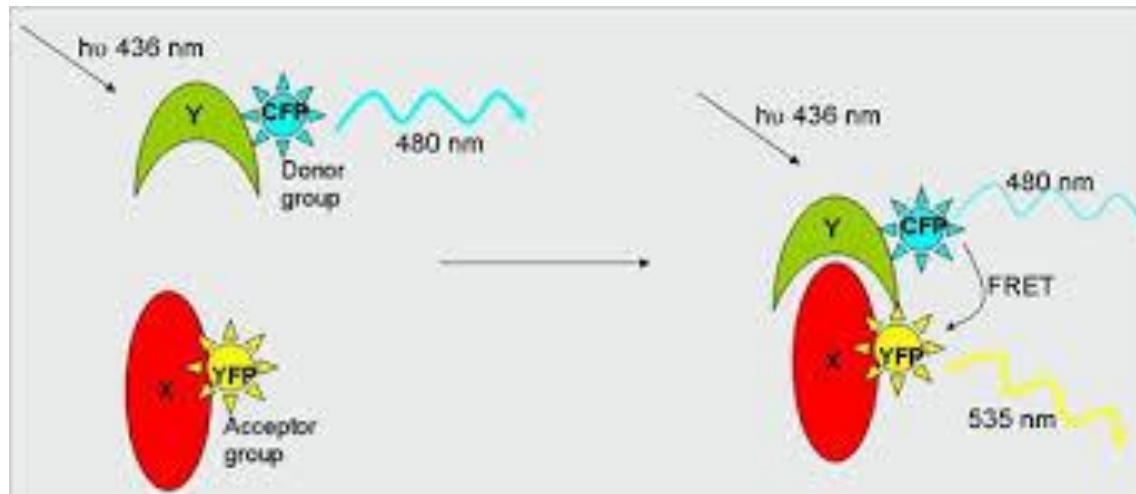
amino-reactive sulfo-NHS-ester group (left)  
sulfhydryl reactive maleimide group (right)

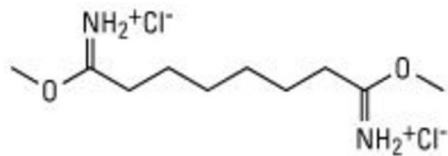


N-Succinimidyl 3-(2-pyridyldithio)propionate







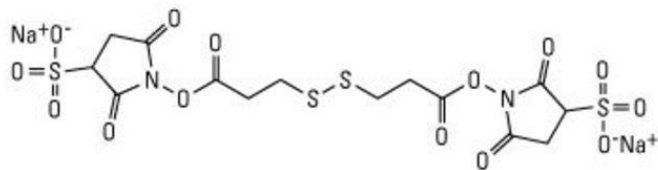


**DMS**

Dimethyl suberimidate • 2HCl

MW 273.20

Spacer Arm 11.0 Å

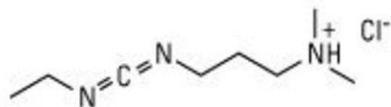


**DTSSP**

3,3'-Dithiobis(sulfosuccinimidylpropionate)

MW 608.51

Spacer Arm 12.0 Å

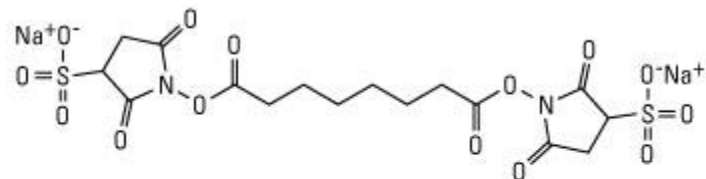


**EDC**

1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide • HCl

MW 191.70

Spacer Arm 0.0 Å

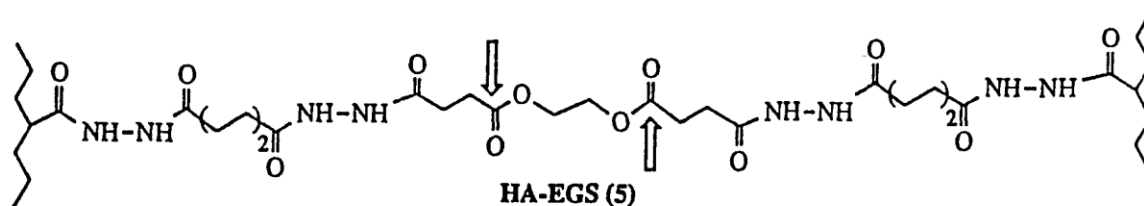
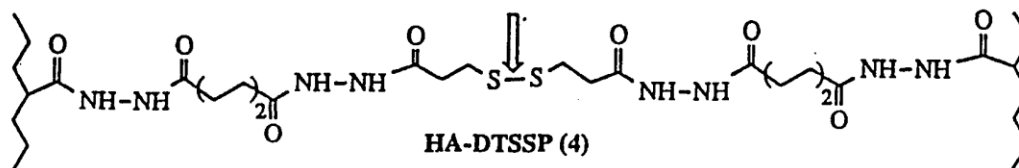
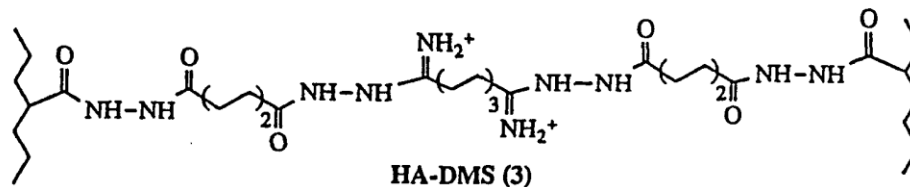
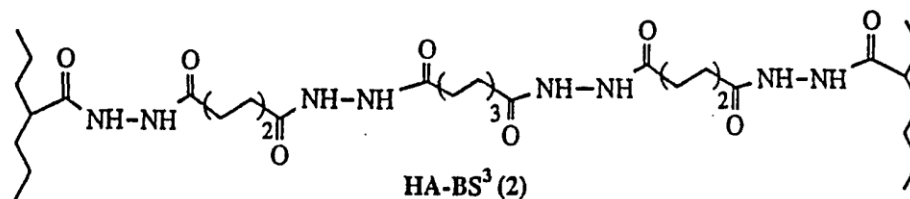


**BS3**

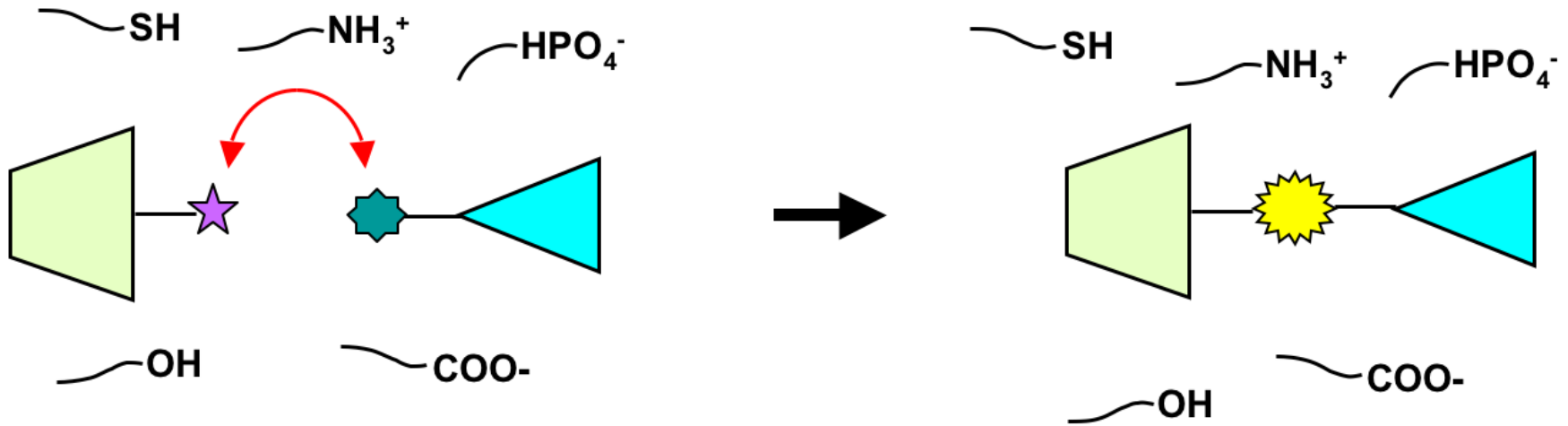
Bis(sulfosuccinimidyl) suberate

MW 572.43

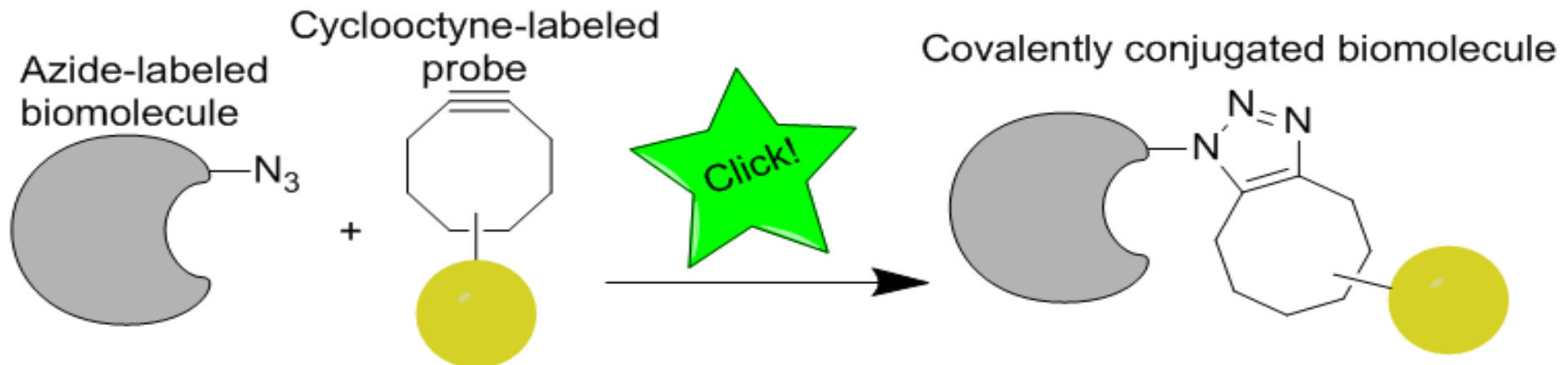
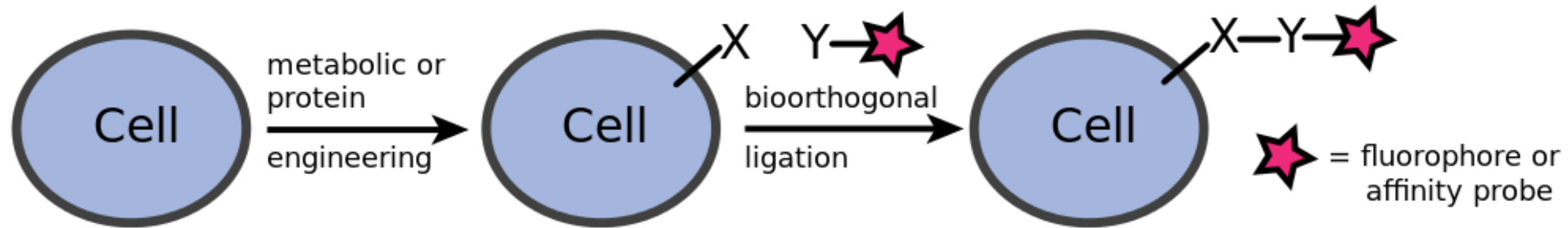
Spacer Arm 11.4 Å

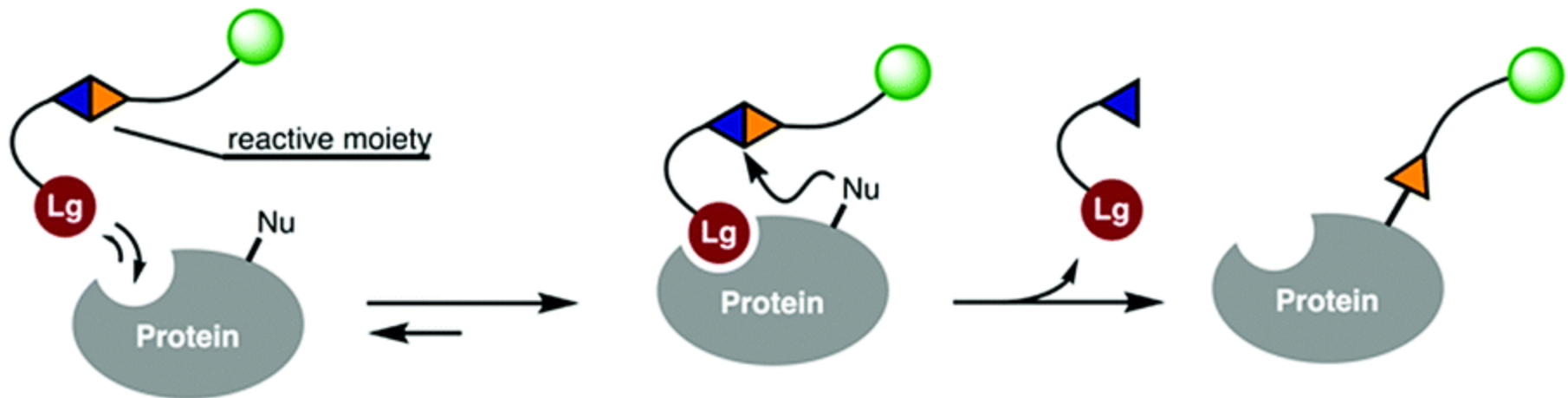


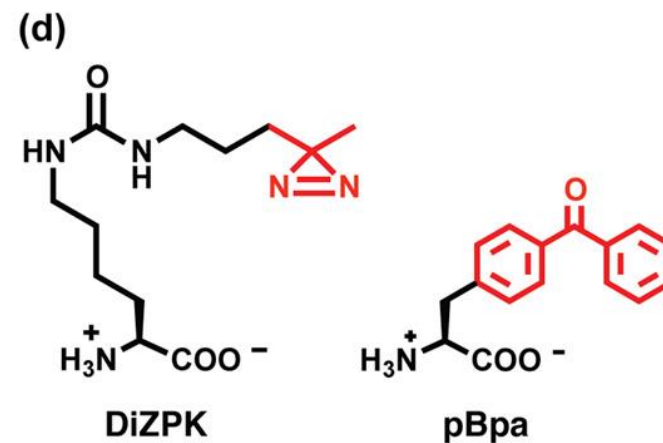
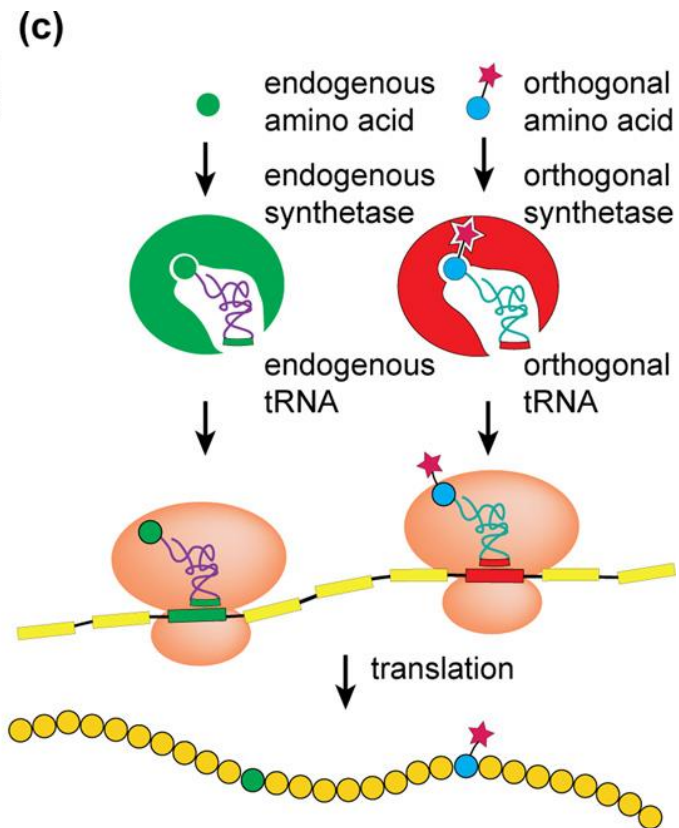
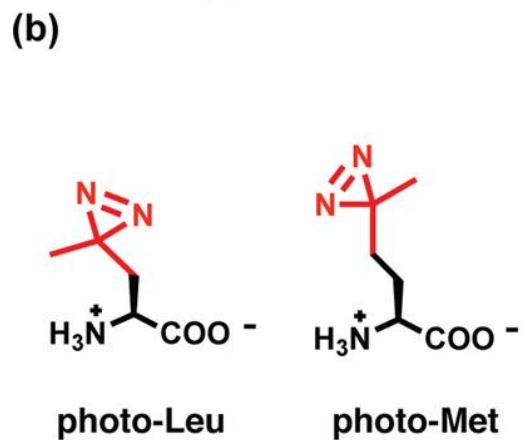
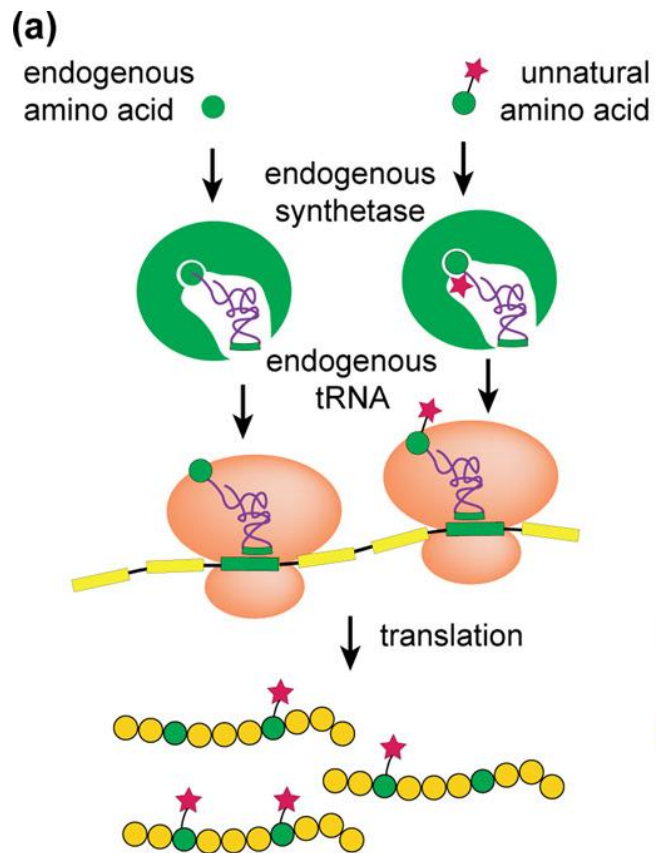
# Bioorthogonal chemistry - study of biomolecules in their natural environment



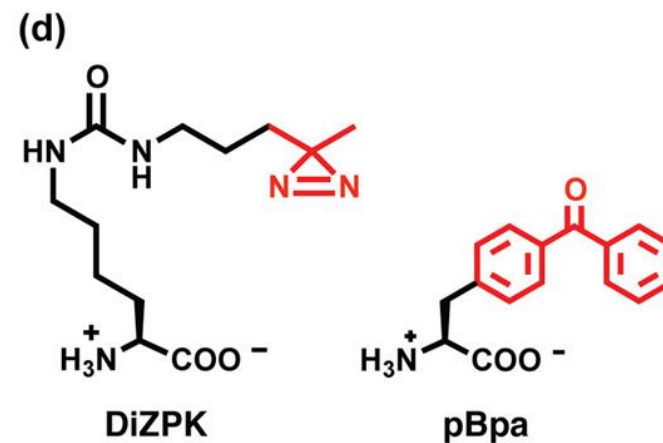
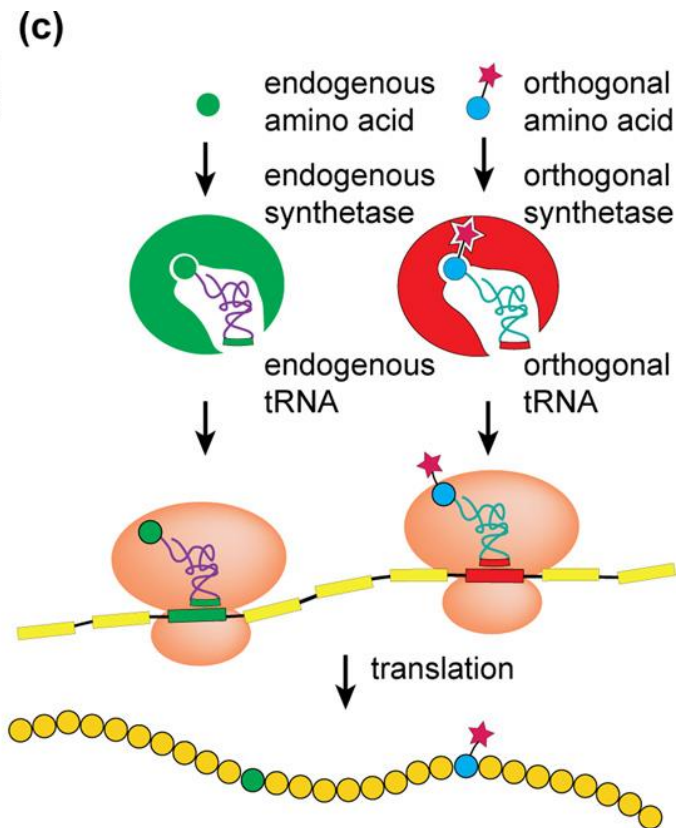
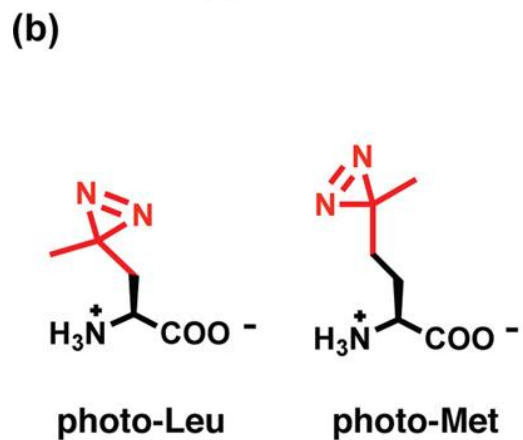
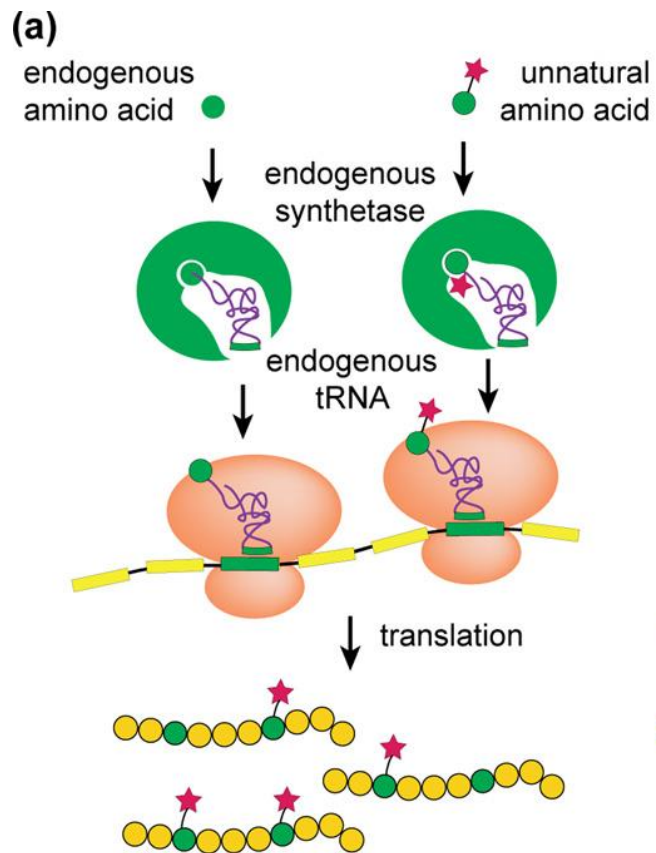
# Bioorthogonal chemistry - study of biomolecules in their natural environment

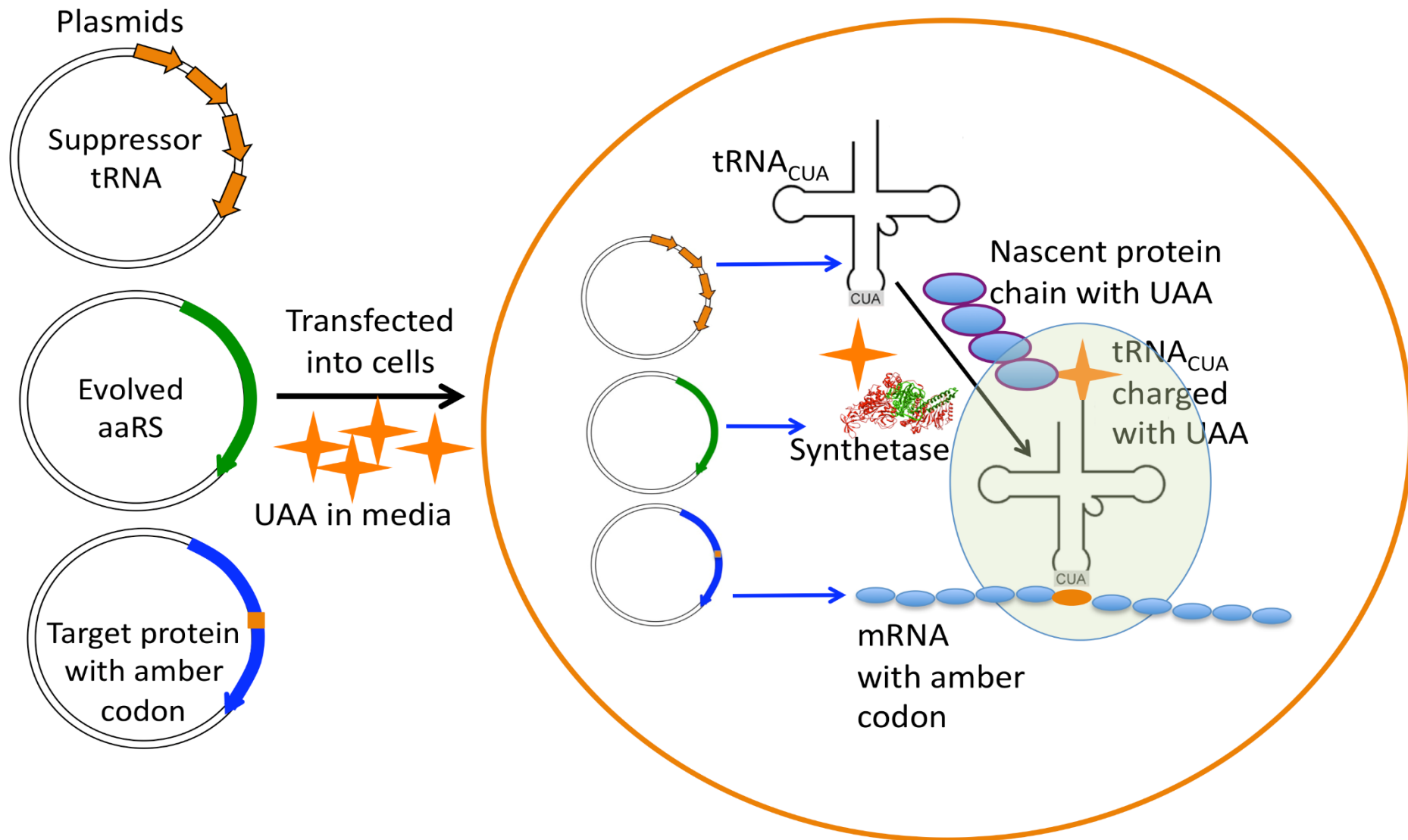


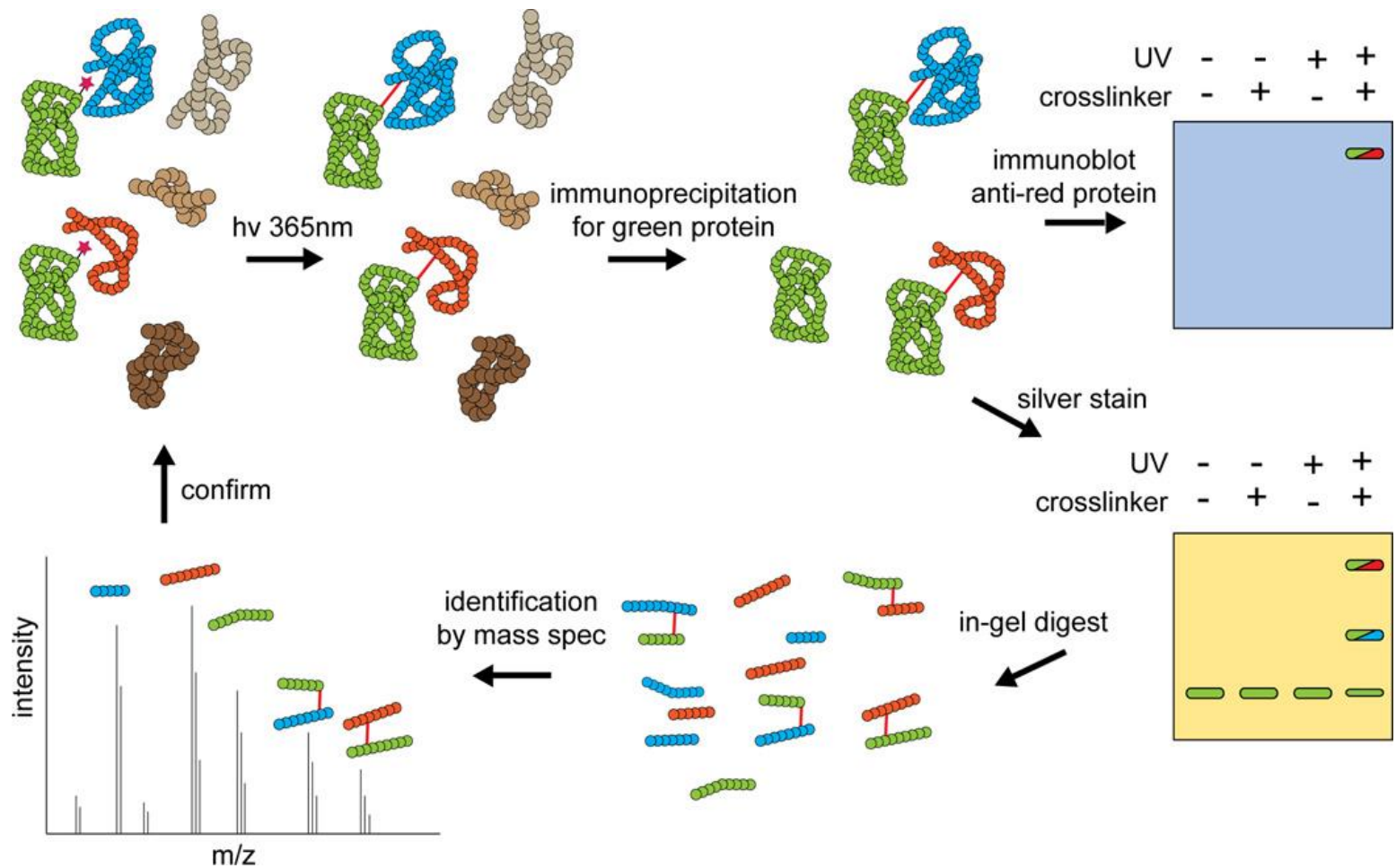




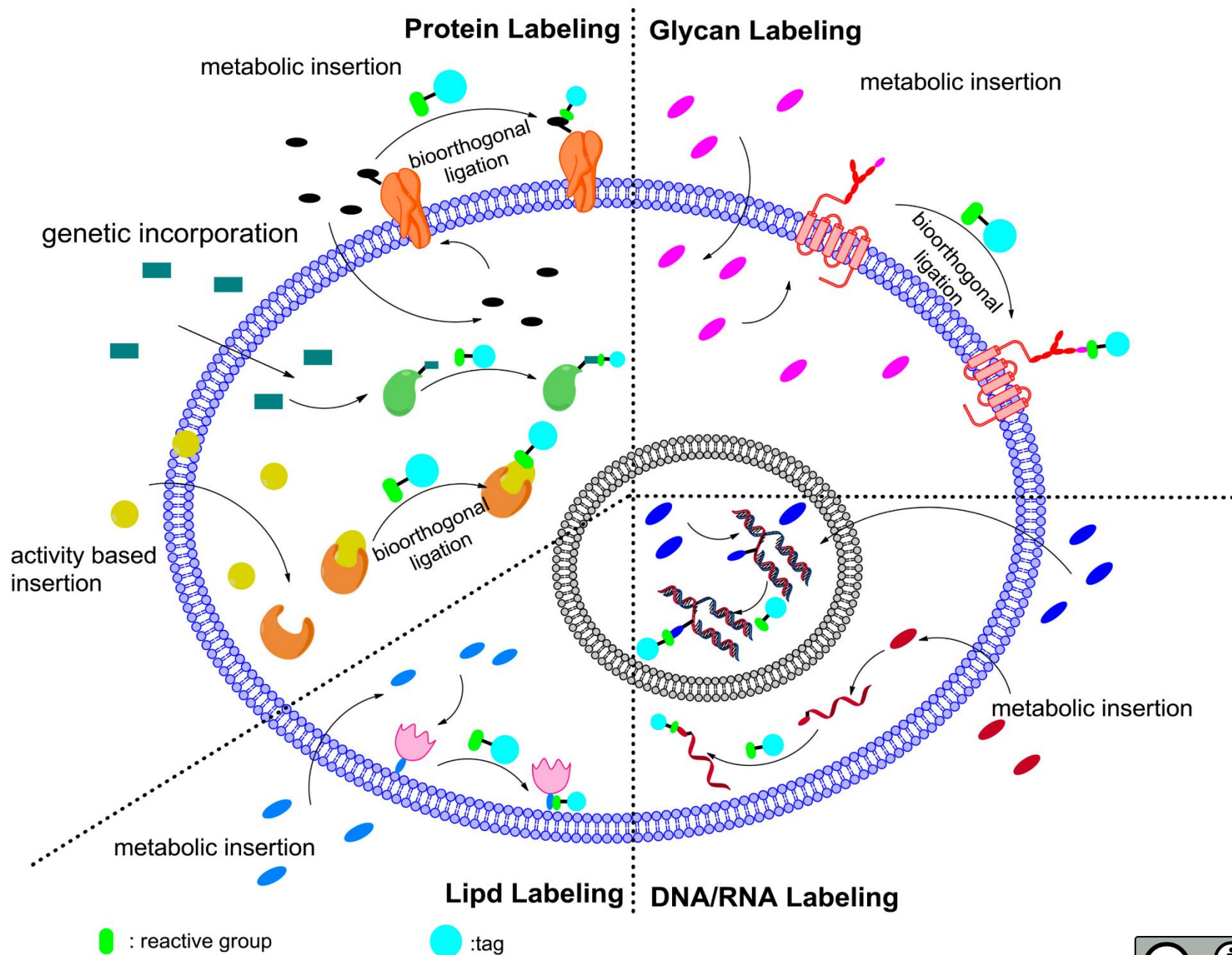


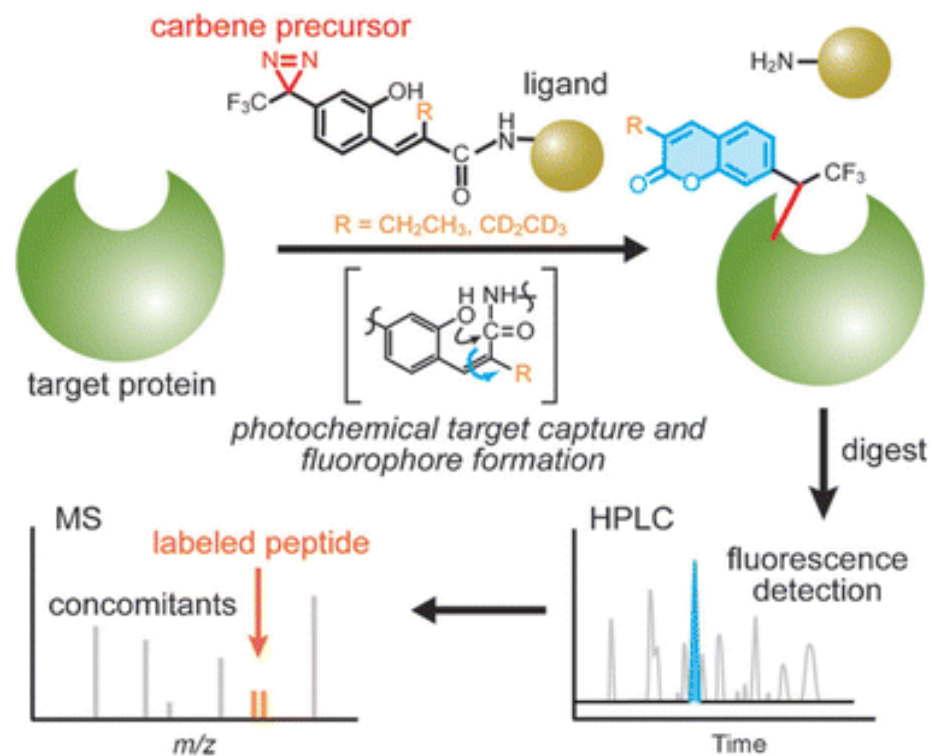
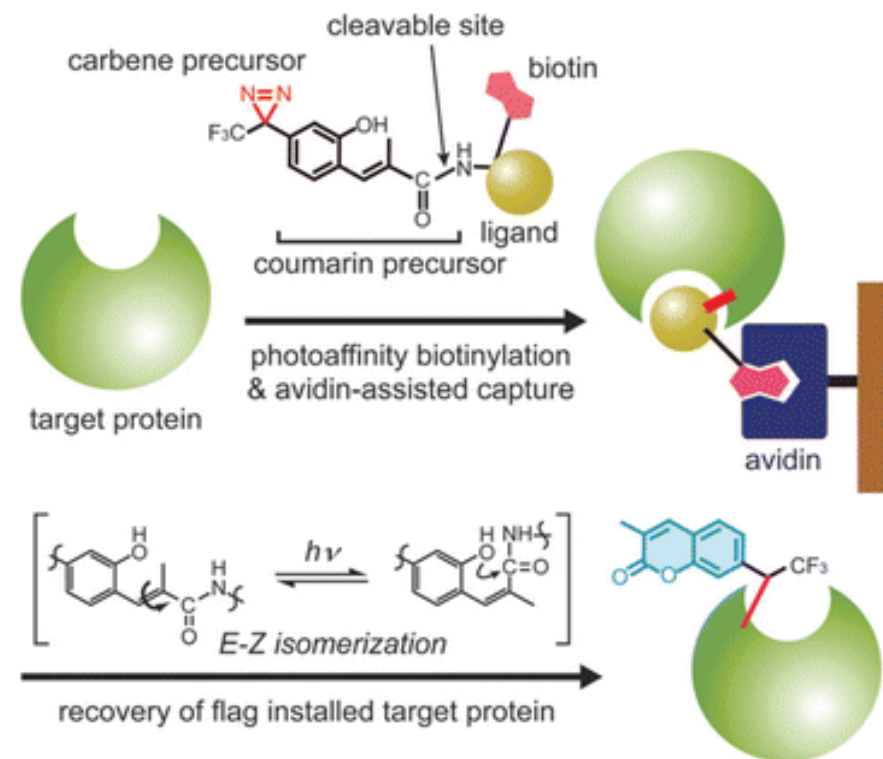




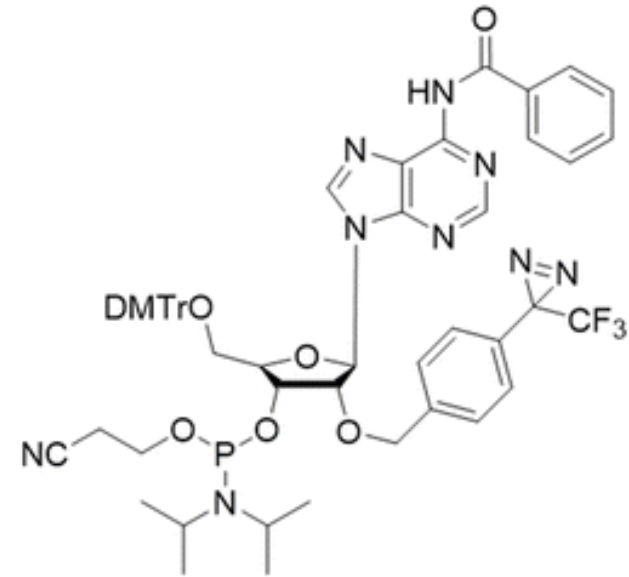
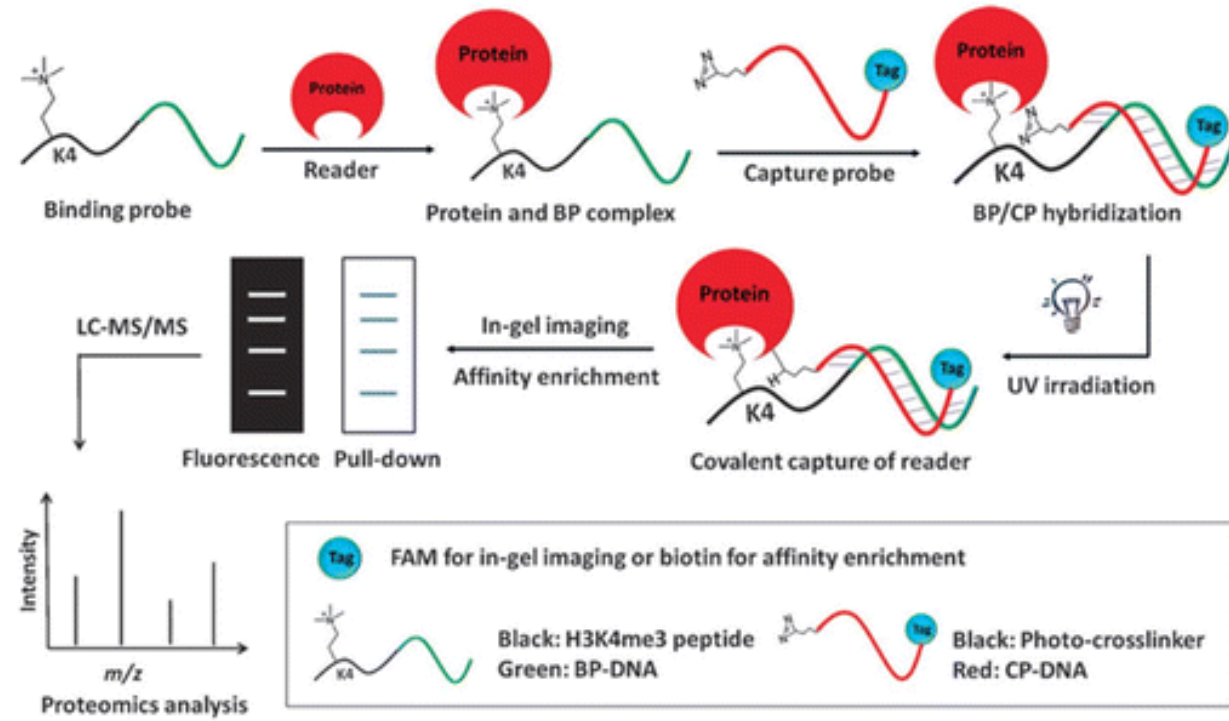


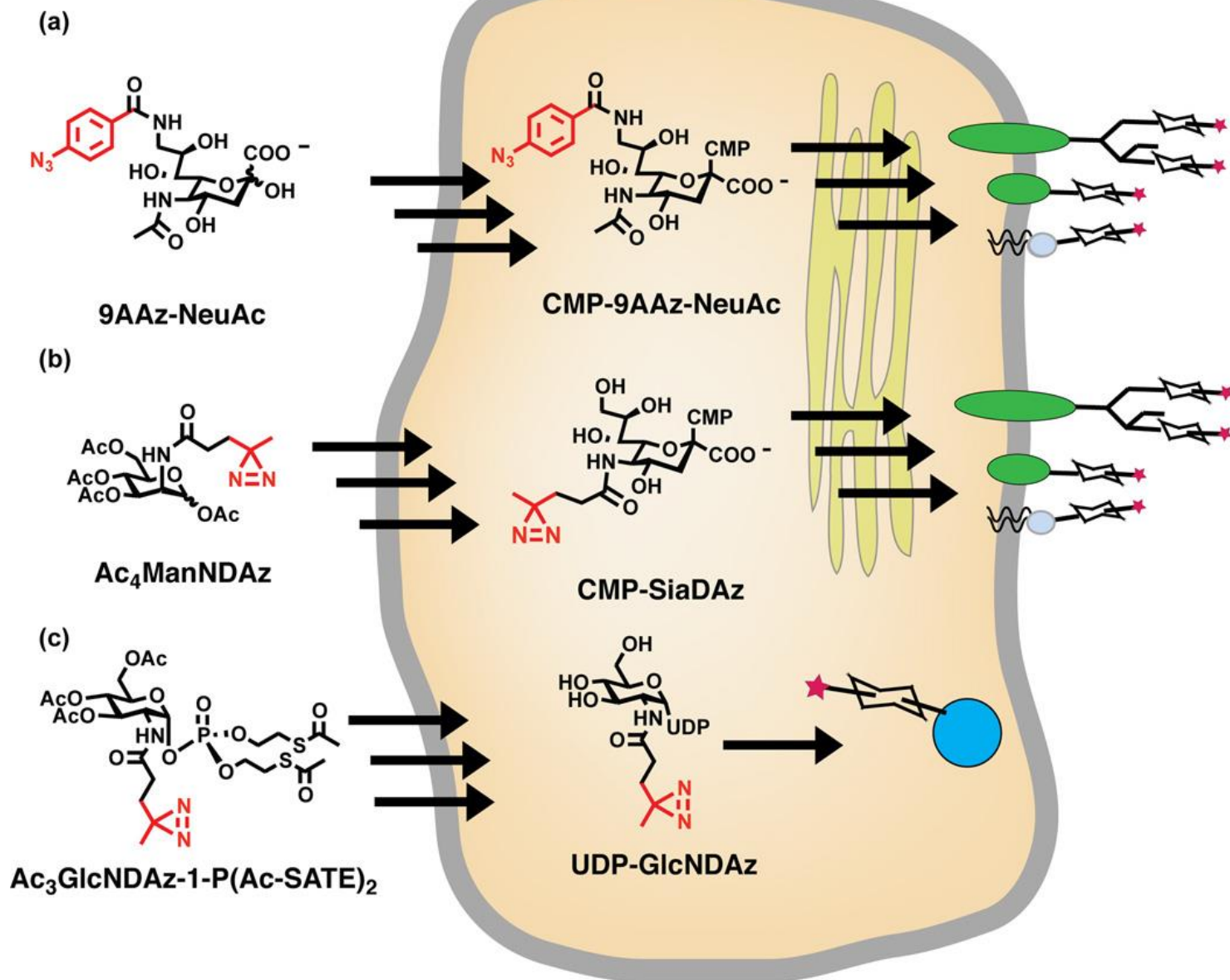
Analysis of interactome with mass spectrometry







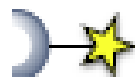




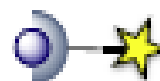
chemical reporter



detection reagent



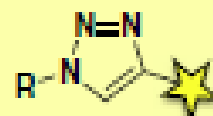
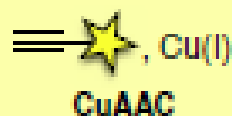
product



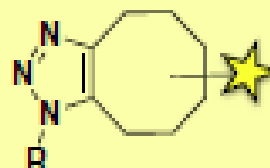
comments



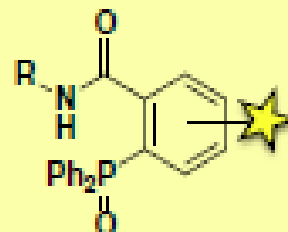
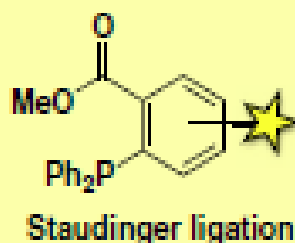
azide



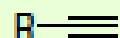
optimized reagents and conditions  
are compatible with live cells



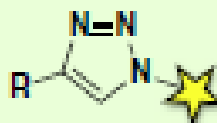
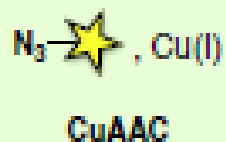
no metal catalyst required;  
some cyclooctynes are amenable  
to in vivo work



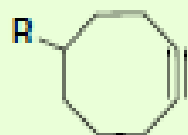
among the most selective in vivo  
reactions despite its slow rate



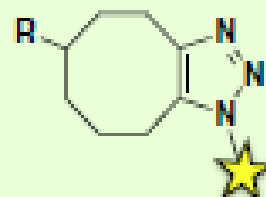
terminal alkyne



optimized reagents and conditions  
are compatible with live cells



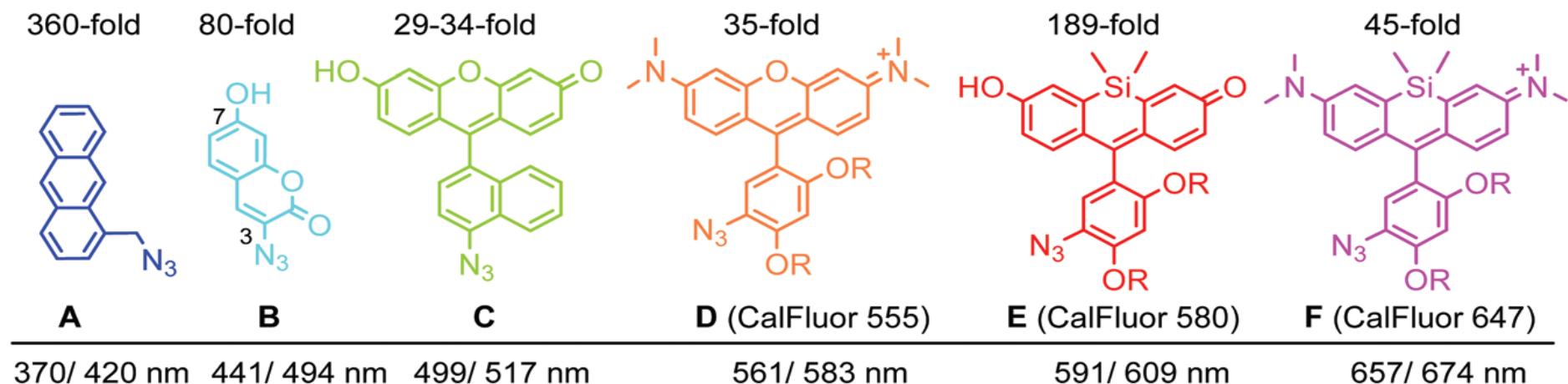
strained alkyne



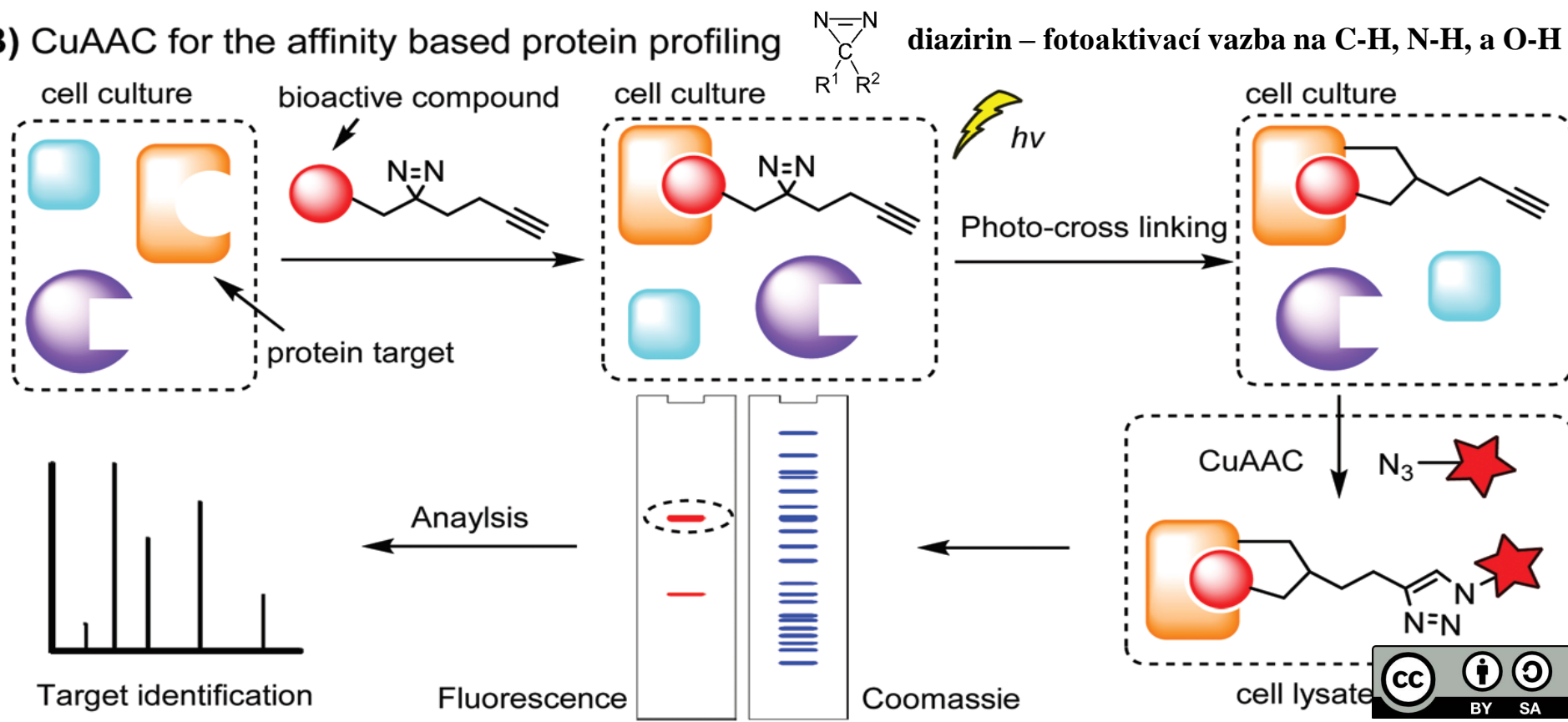
no metal catalyst required;  
some cyclooctynes are amenable  
to in vivo work

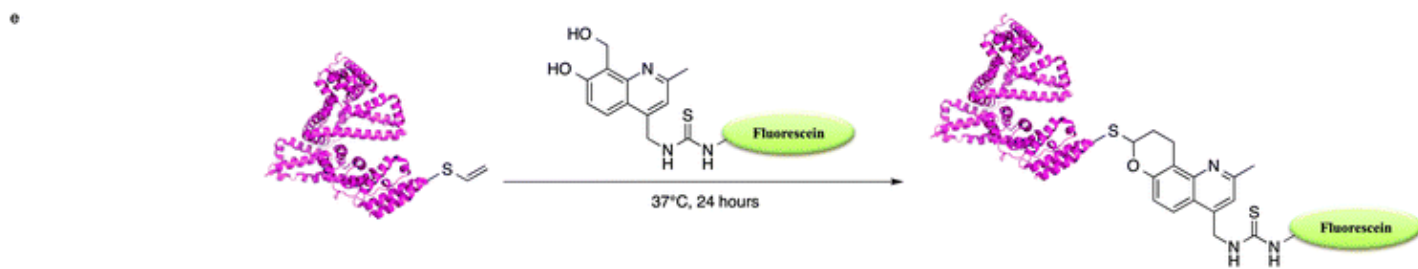
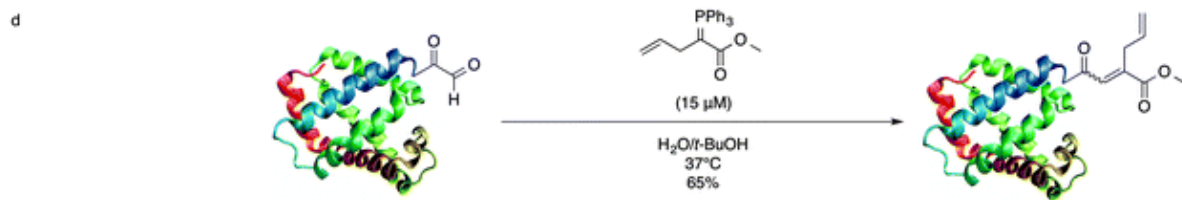
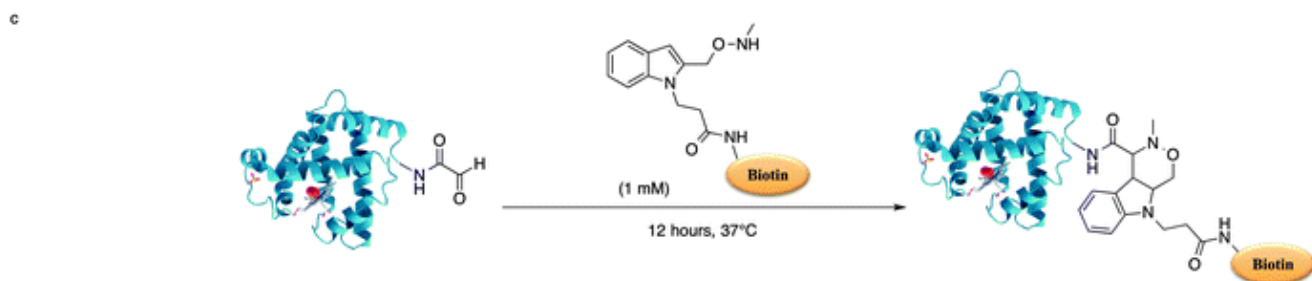
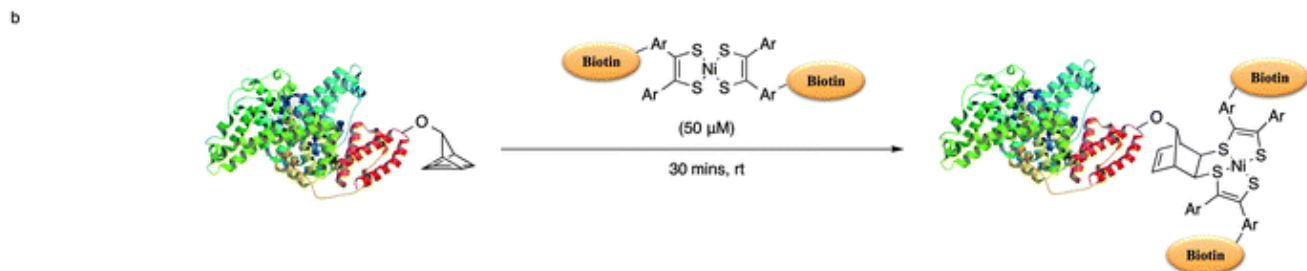
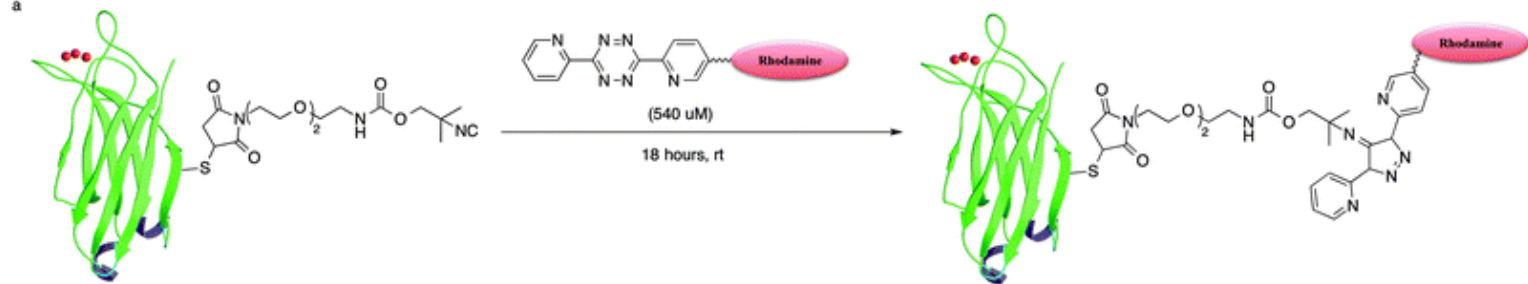


## A) A panel of highly fluorogenic azide probes

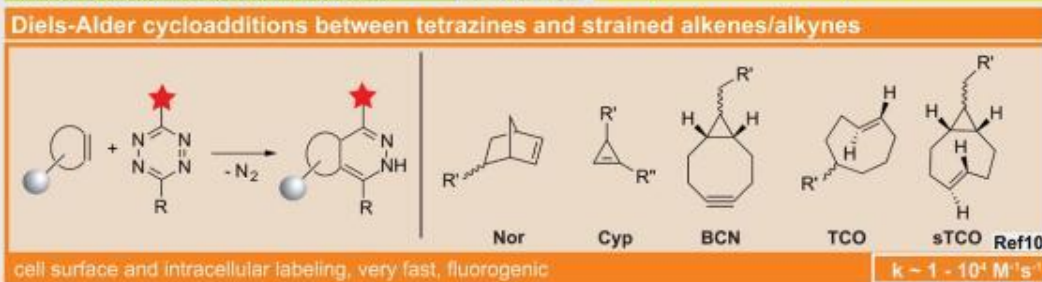
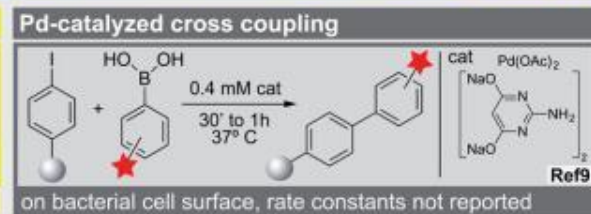
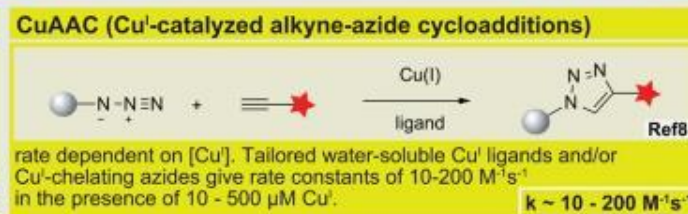
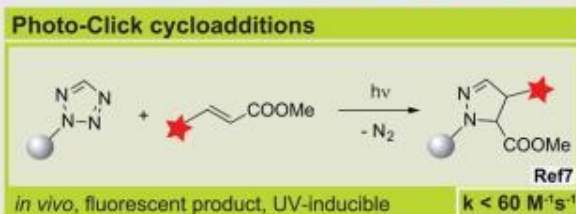
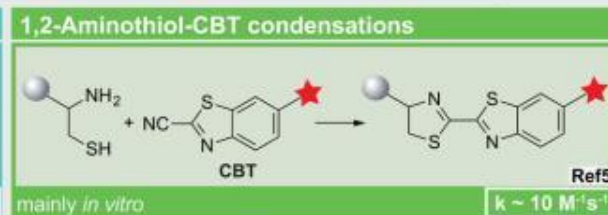
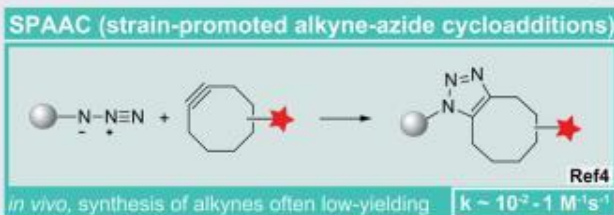
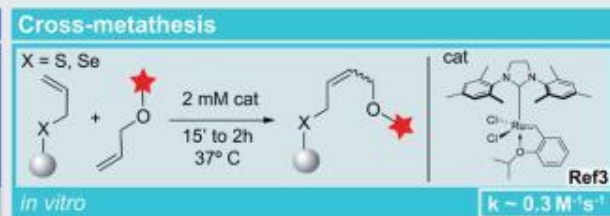
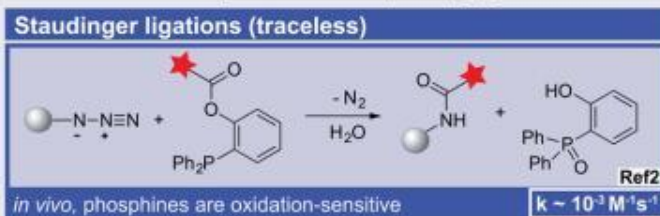
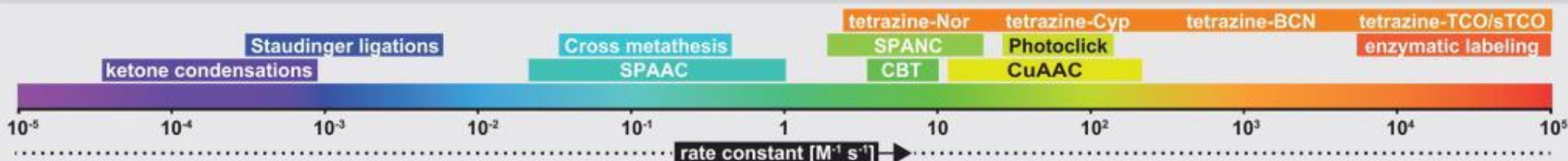


## B) CuAAC for the affinity based protein profiling





# BIOORTHOGONAL REACTIONS FOR LABELING PROTEINS



## References

- (1) a) Rideout, D. *Science* **1986**, 233, 561; b) Mahal, L.K. et al., *Science* **1997**, 276, 1125; c) Brustad, E.M. et al., *JACS* **2008**, 130, 17664; d) Chen, I. et al., *Nat. Methods* **2005**, 2, 99.
- (2) a) Saxon E. et al., *Science* **2000**, 287, 2007; b) Saxon E. et al., *Organic Letters* **2000**, 2, 2141; c) Prescher, J.A. et al., *Nature* **2004**, 430, 873; d) Tsao, M.L. et al., *ChemBiochem* **2005**, 6, 2147; e) Klick, K.L. et al., *PNAS* **2002**, 99, 19.
- (3) a) Lin, Y.A. et al., *JACS* **2008**, 130, 9642; b) Lin, Y.A. et al., *JACS* **2013**, 135, 12156.
- (4) a) Agard N.J. et al., *JACS* **2004**, 126, 15046; b) Baskin J.M., *PNAS* **2007**, 104, 16793; c) Ning, X. et al., *Angew. Chem. Int. Ed.* **2008**, 47, 2253; d) Plass, T. et al., *Angew. Chem. Int. Ed.* **2011**, 50, 3878.
- (5) a) Liang, G. et al., *Nature Chem.* **2010**, 2, 54; b) Nguyen, D.P. et al., *JACS* **2011**, 133, 11418.
- (6) a) Ning, X. et al., *Angew. Chem. Int. Ed.* **2010**, 49, 3065; b) McKay, C.S. et al., *ChemComm* **2011**, 47, 10040.
- (7) a) Song, W. et al., *JACS* **2008**, 130, 9654; b) Song, W. et al., *Angew. Chem. Int. Ed.* **2008**, 47, 2832; c) Yu, Z. et al., *Angew. Chem. Int. Ed.* **2012**, 51, 10600.
- (8) a) Wang, Q. et al., *JACS* **2003**, 125, 3192; b) Nguyen, D.P. et al., *JACS* **2009**, 131, 8720; c) Presolski, S.I. et al., *JACS* **2010**, 132, 14570; d) Uttamapinant, C. et al., *Angew. Chem. Int. Ed.* **2012**, 51, 5852.
- (9) a) Chalker, J.M. et al., *JACS* **2009**, 131, 16346; b) Spicer, C.D. et al., *JACS* **2012**, 134, 800.
- (10) a) Blackman, M.L. et al., *JACS* **2008**, 130, 13518; b) Devaraj, N.K. et al., *Bioconjugate Chem.* **2008**, 19, 2297; c) Lang, K. et al., *Nature Chem.* **2012**, 4, 298; d) Plass, T. et al., *Angew. Chem. Int. Ed.* **2012**, 51, 4166; e) Lang, K. et al., *JACS* **2012**, 134, 10317; f) Yang, J. et al., *Angew. Chem. Int. Ed.* **2012**, 51, 7476; g) Seitchik, J.L. et al., *JACS* **2012**, 134, 2898; h) Elliott, T. et al., unpublished data.

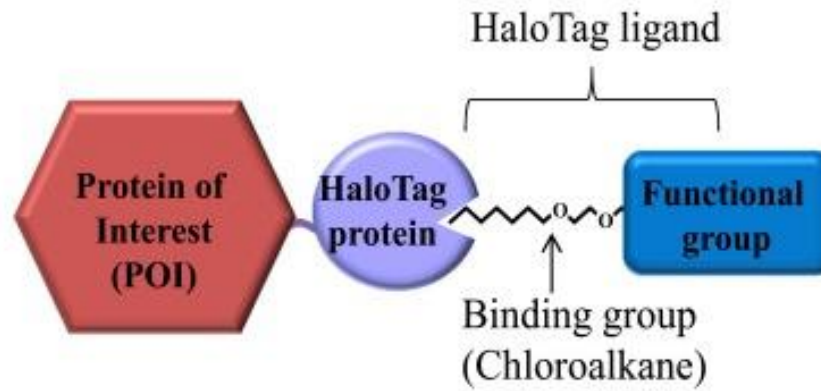
Kathrin Lang and Jason W. Chin

Medical Research Council, Laboratory of Molecular Biology, Center for Chemical and Synthetic Biology, Division for Protein and Nucleic Acid Chemistry, Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge, CB2 0QR, UK

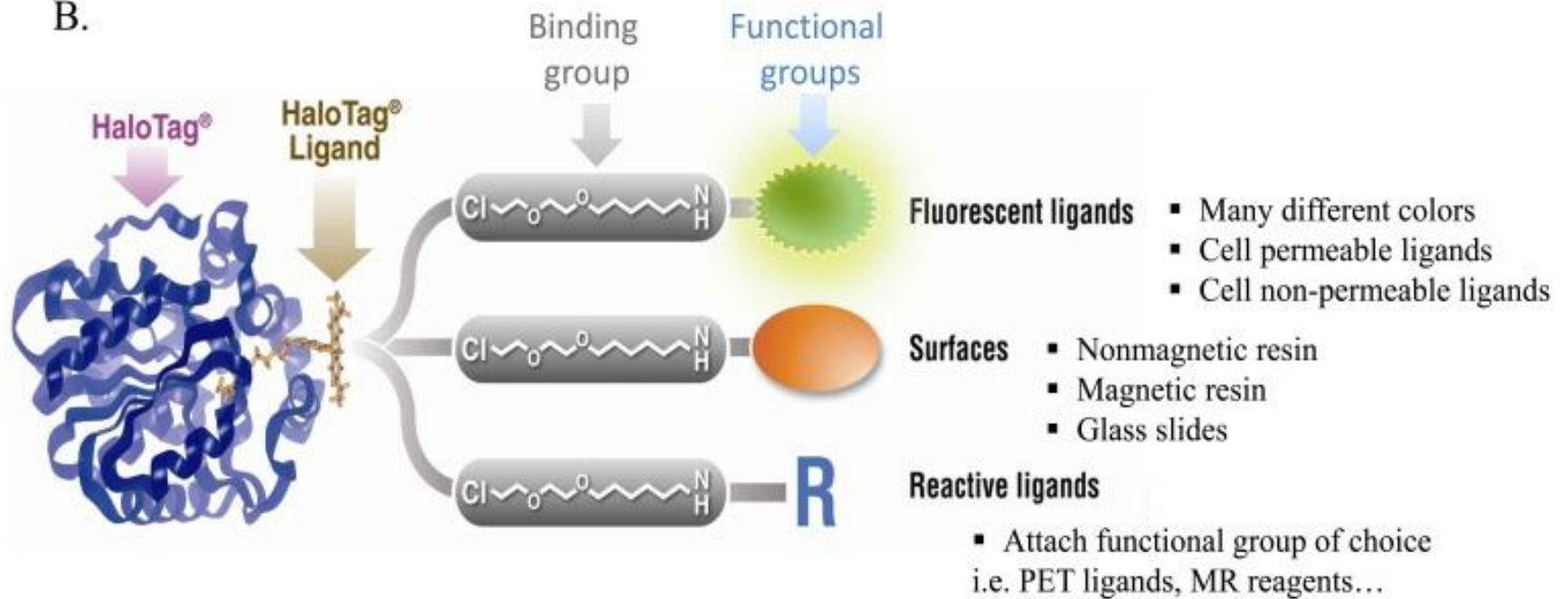




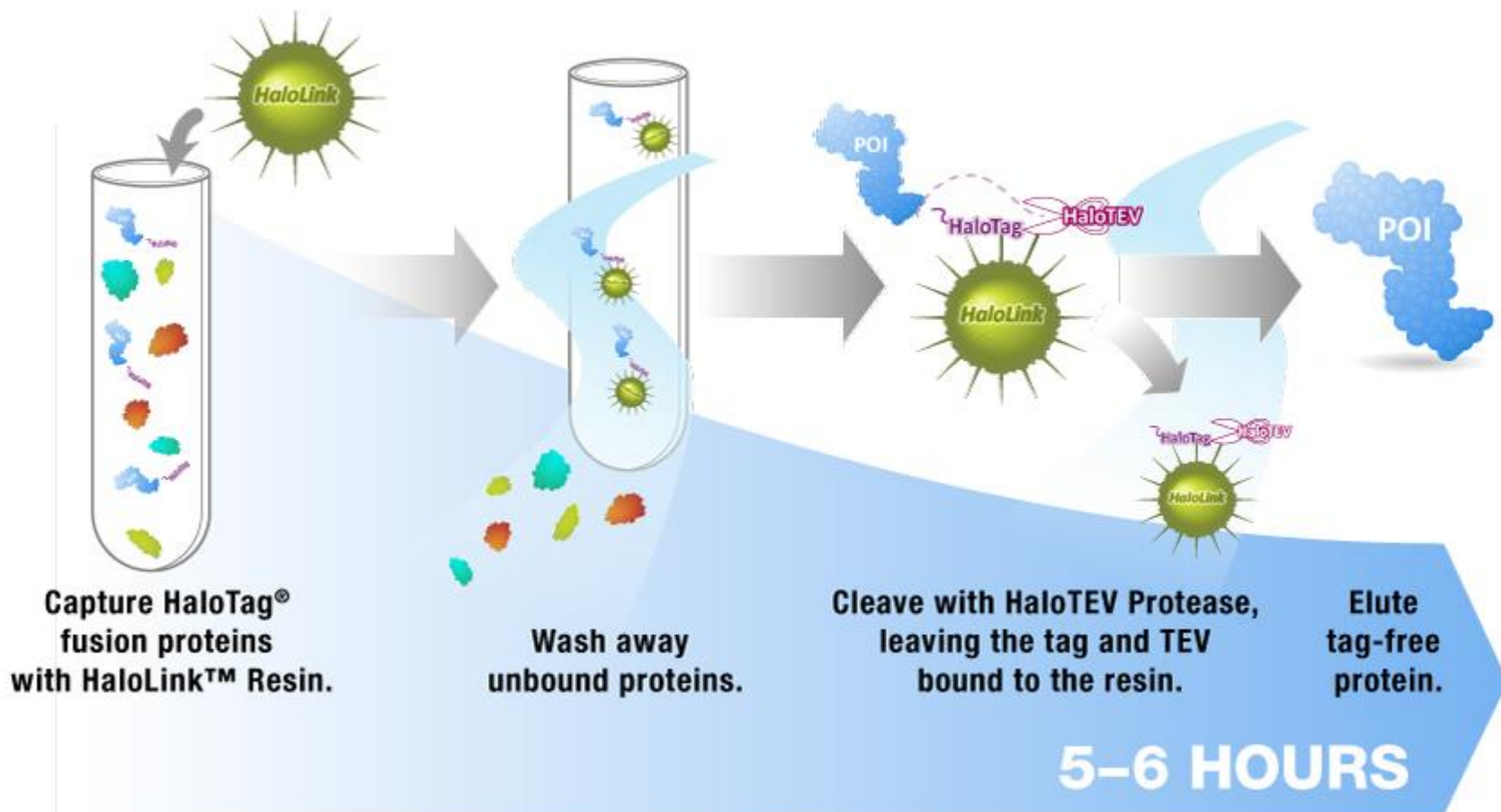
A.



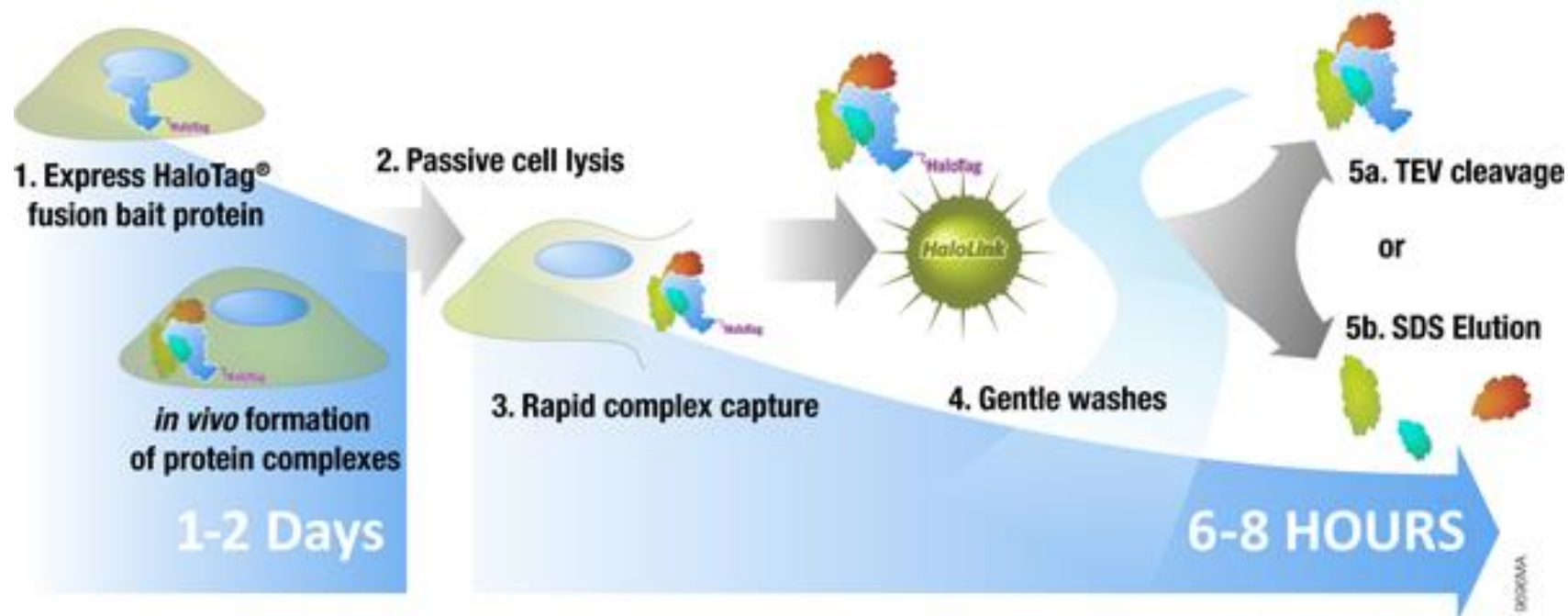
B.







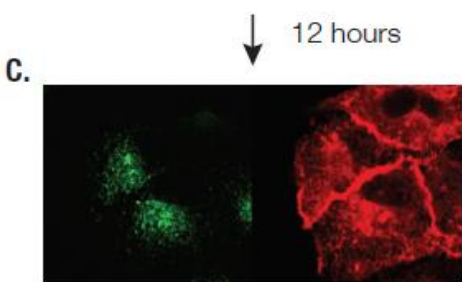
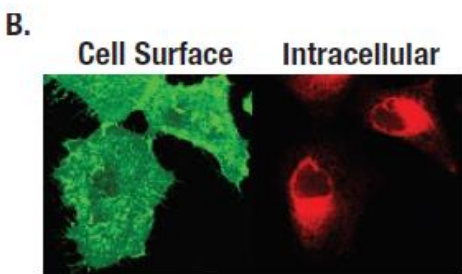
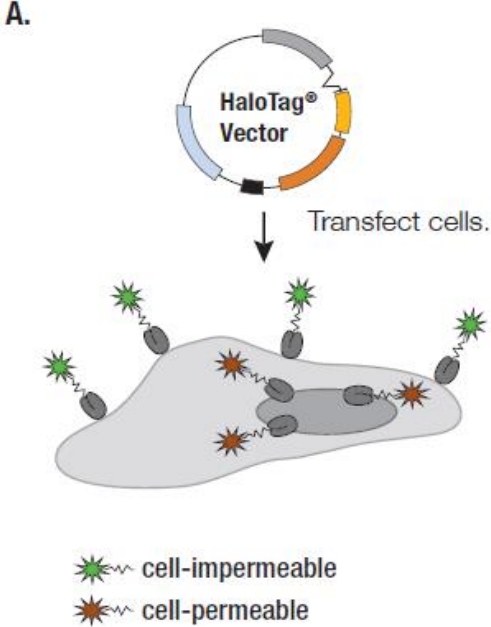
9640118





# Transport of $\beta$ -integrin fragment fused with HaloTag®

copyright BMC Cell Biology

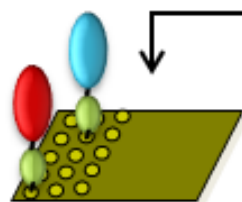




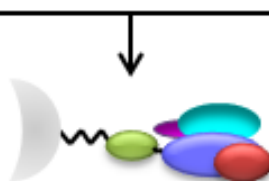
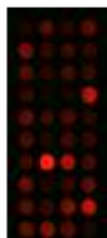
Expression of HaloTag fusion



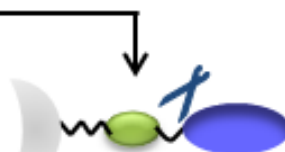
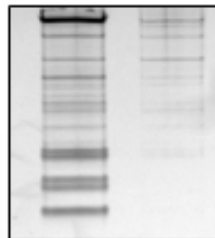
**Covalent capture**  
HaloTag surfaces



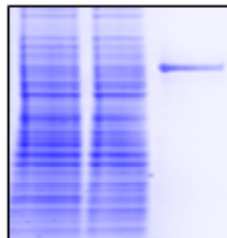
**Protein display**



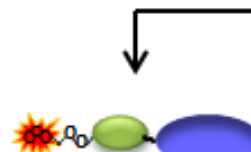
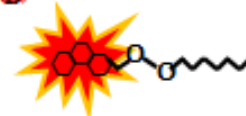
**Interactions**



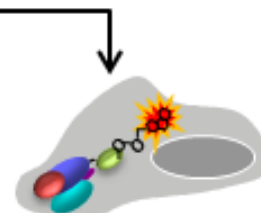
**Purification**



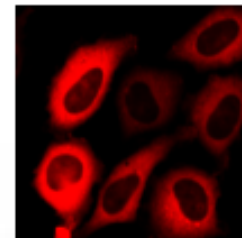
**Protein labeling**  
HaloTag fluorescent ligands



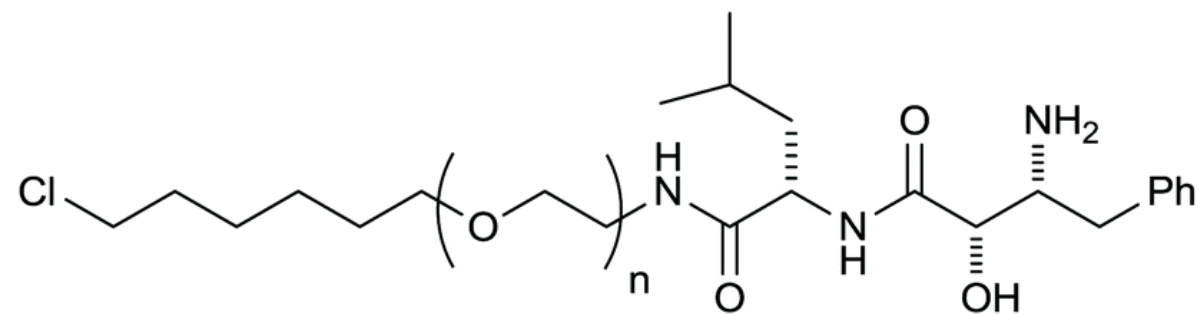
**Detection**



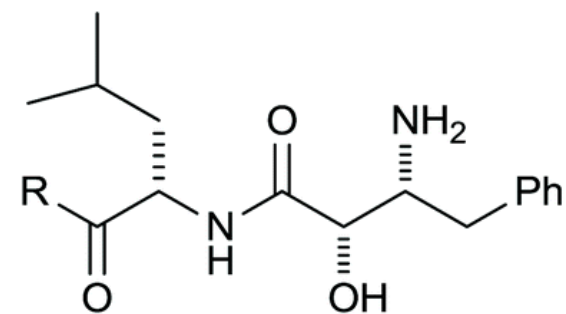
**Cellular imaging**



HaloTag-binding moiety    spacer    cIAP1-binding moiety

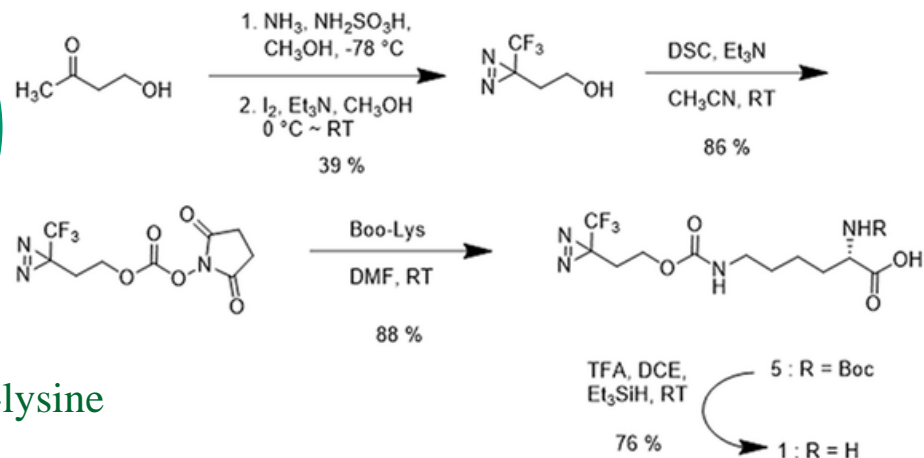
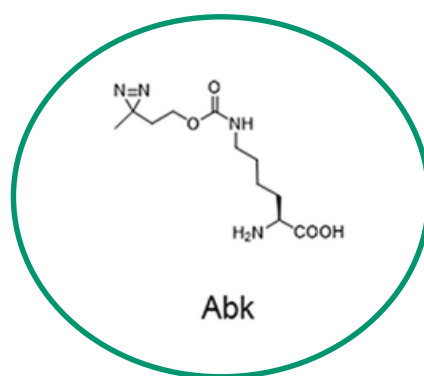


**1a-d** :  $n = 3-6$



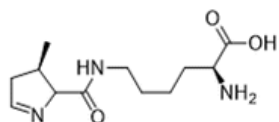
BE04 (**2**) : R = NHMe

MeBS (**3**) : R = OMe

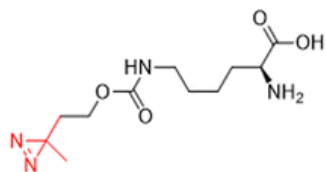


### 3'-azibutyl-N-carbamoyl-lysine (AbK)

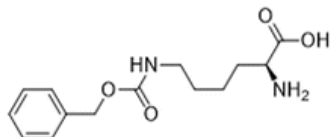
### Synthesis of the diazirine-modified lysine



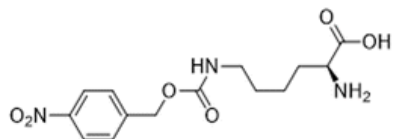
Pyrrolysine



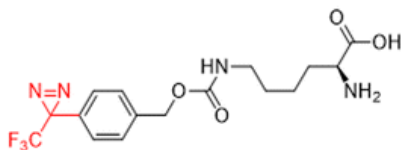
AbK



ZLys



pNO<sub>2</sub>ZLys



TmdZLys

Unnatural Amino Acid	Host	Unique property of the amino acid	Reference
<i>p</i> -iodo-L-tyrosine	<i>E. coli</i> , yeast mammalian	The heavy atom iodine facilitates structural studies of proteins. This UAA can also used for determination of tyrosine phosphorylation sites in a protein	[52], [53], [7]
<i>p</i> -acetyl-L-phenylalanine	<i>E. coli</i> , yeast,	The acetyl group can be modified bioorthogonally with any other probe with a hydroxylamine group.	[54], [46], [7]
<i>p</i> -benzoyl-L-phenylalanine, (Bpa)	<i>E. coli</i> , yeast, mammalian	Cross-links with nearby C-H bonds when exposed to light between 350-360 nm.	[9], [26], [27], [10], [7]
<i>p</i> -azido-L-phenylalanine (AzPhe)	<i>E. coli</i> , yeast, mammalian	AzPhe has been used as a photocrosslinking UAA to determine the interactions sites on proteins. This UAA has also been used as an IR-active probe for detection of conformational change in proteins. AzPhe also cross-links with triarylphosphines, alkyne or DIBO alkynes in a bioorthogonal manner.	[29], [26], [44], [47], [48], [7]
O-Methyl L-tyrosine	<i>E. coli</i> , yeast, mammalian	Increases the bulk of tyrosine	[11], [55]
L-(7-hydroxycoumarin-4-yl) Ethylglycine	<i>E. coli</i>	Fluorescent amino acid	[38]
dansylalanine	Yeast	Fluorescent amino acid	[39]
(S)-1-carboxy-3-(7-hydroxy-2- oxo-2H-chromen-4-yl)propan-1- aminium (CouAA)	<i>E. coli</i>	Fluorescent amino acid	[40], [41]
3-(6-acetylnaphthalen-2-ylamino) -2-aminopropanoic acid (Anap)	Yeast, mammals	Fluorescent amino acid	[42], [56]
Photocaged tyrosine derivatives	<i>E.coli</i> , mammalian	Photocaged amino acid	[34]
Photocaged cysteine derivatives	Yeast	Photocaged amino acid	[17]
Photocaged lysine derivatives	<i>E. coli</i> , mammalian	Photocaged amino acid	[18]
Photocaged serine derivatives	Yeast	Photocaged amino acid	[36]
(2,2'-bipyridin-5-yl)alanine (Bpy- Ala) (A Fe <sup>2+</sup> /3 <sup>+</sup> , Cu <sup>2+</sup> , Co <sup>2+</sup> /3 <sup>+</sup> , and Ru <sup>2+</sup> /3 <sup>+</sup> chelating UAA)	<i>E. coli</i>	This UAA was site specifically introduced in a DNA binding protein. The UAA modified protein site-specifically cleaved double stranded DNA.	[57], [58]
2-nitrophenyl alanine (2-NPA)	<i>E. coli</i>	When irradiated at 365 nm, the 2-NPA residue photocleaves the protein specifically at the site of incorporation.	[59]
<i>p</i> -carboxymethyl-phenylalanine	<i>E. coli</i>	A stable phosphotyrosine analogue, resistant to hydrolysis by protein tyrosine phosphatase resulting in constitutively active proteins.	[60]

## A Domain tags

Autofluorescent **Protein**—**FP**

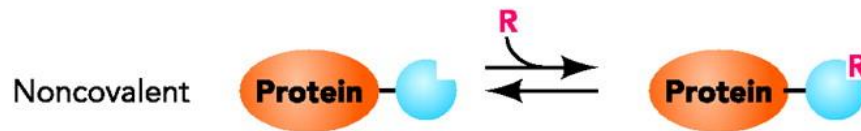
**R** Fluorescent, affinity, or reactive group  
**L** Leaving group

## Examples

CFP, GFP, YFP, RFP, etc.

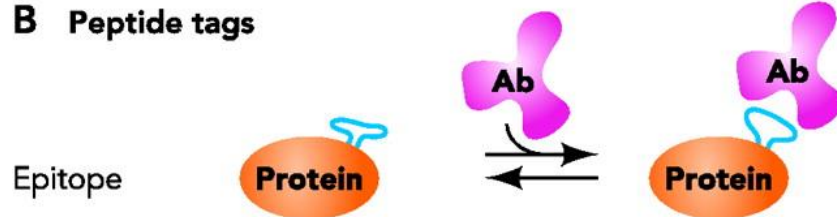


Dehalogenase, AGT, cutinase



FKBP, DHFR, single-chain antibodies

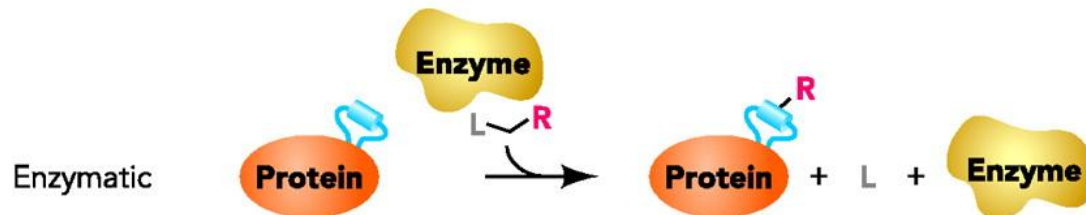
## B Peptide tags



HA, Flag, myc, etc.



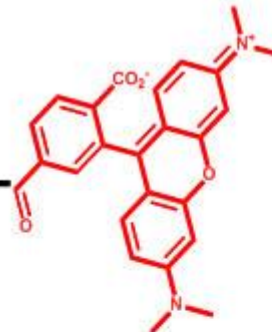
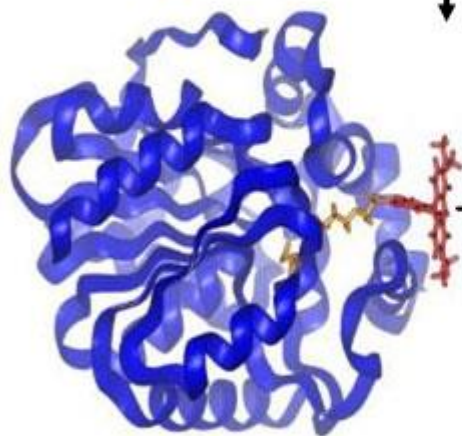
Tetracysteine, TR512, LBP



Biotin ligase, lipoic acid ligase,  
phosphopantetheinyl transferase,  
sortase, transglutaminase

## HaloTag

HaloTag ligand



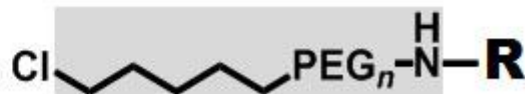
### ***HaloTag fluorescent ligands***

- Multiple fluorophores
- Cell-permeable ligands
- Cell-impermeable ligands
- Red-shifted ligands



### ***HaloTag surfaces***

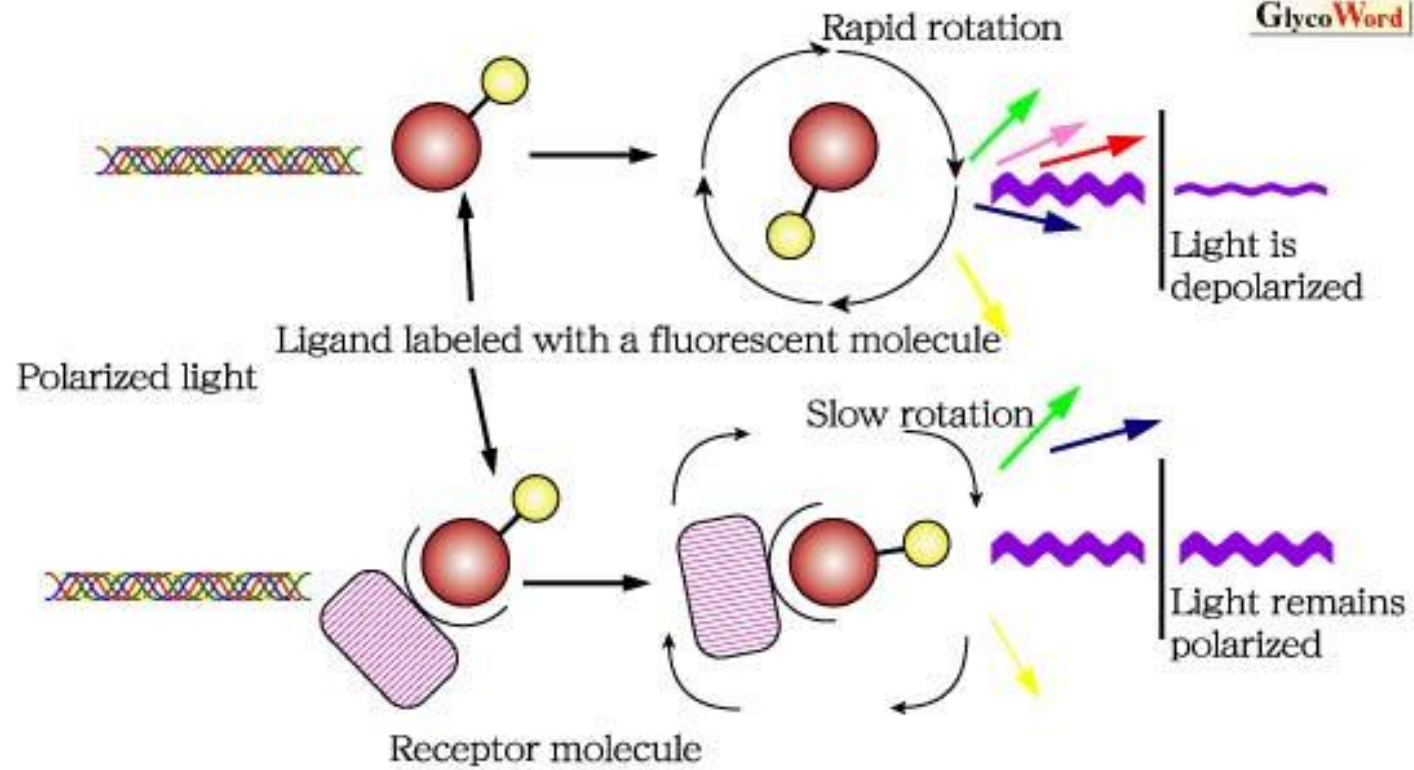
- Nonmagnetic resin
- Magnetic resin
- Glass slides
- 96-well plates



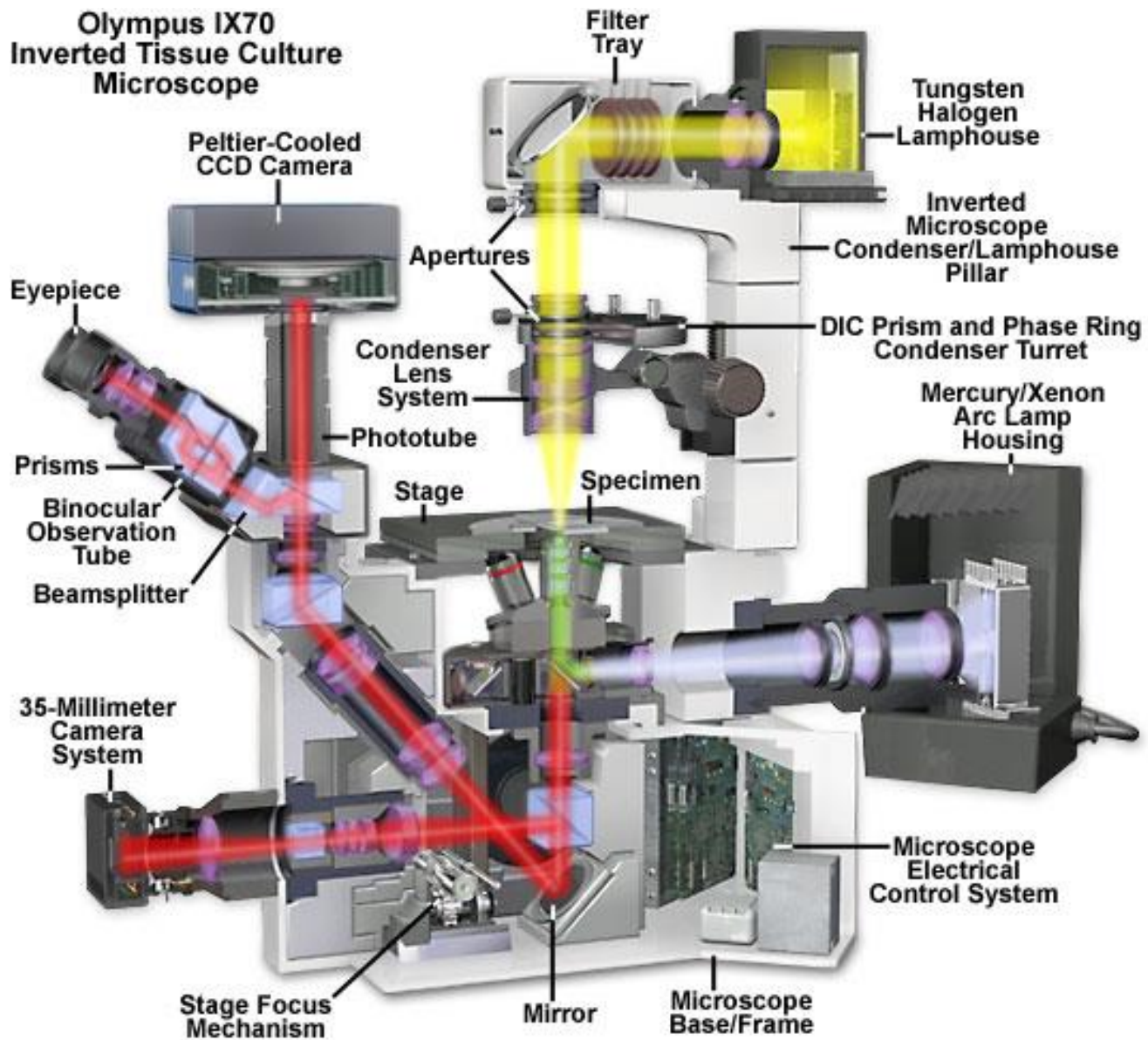
### ***HaloTag reactive ligands***

- Reactive chemistry (i.e. amine, thiol)
- Attach to functional groups of choice (i.e. Quantum dots, PET ligands)



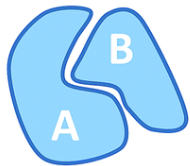


**Olympus IX70  
Inverted Tissue Culture  
Microscope**



## PROTEIN level

Native/  
non-denaturing MS



Detection of  
of intact complex

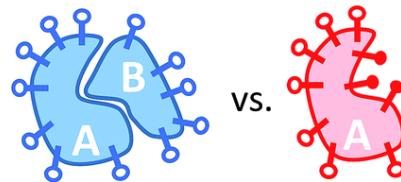
## PEPTIDE level

Hydrogen/deuterium  
exchange



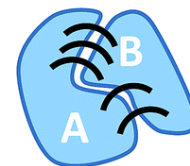
Protection from  
exchange in complex

Covalent labeling



Protection from labeling  
in complex

Chemical cross-linking



Formation of cross-links  
in complex

# Hydrogen–deuterium exchange

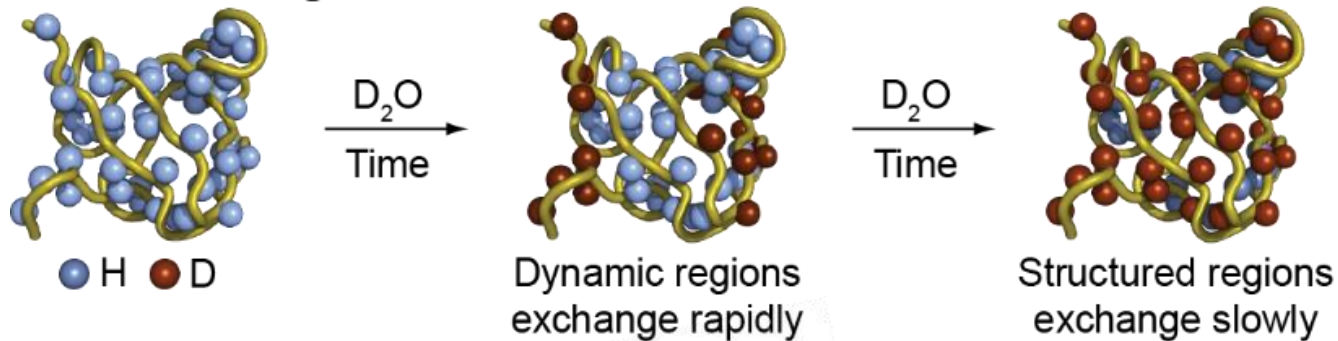
## (H–D or H/D Exchange)

deuterium is added to a protein in H<sub>2</sub>O by diluting the H<sub>2</sub>O solution with D<sub>2</sub>O

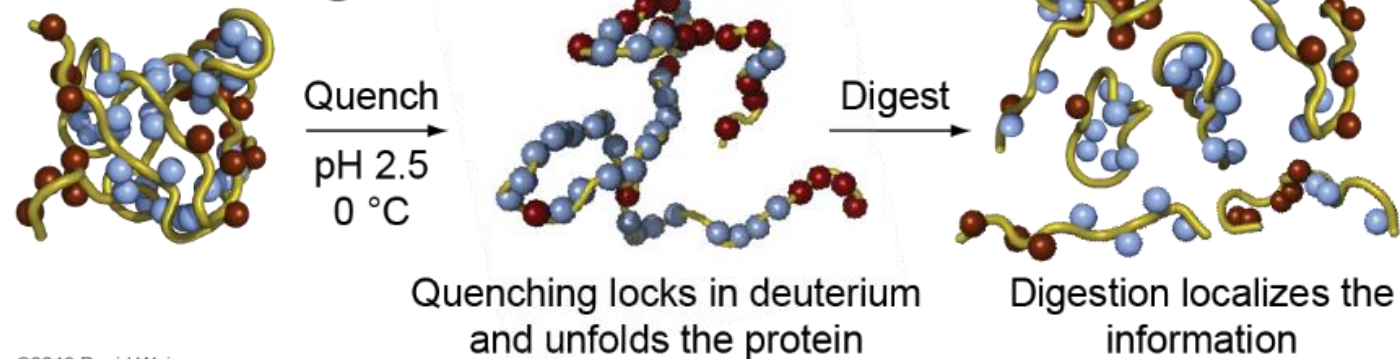
Mass spectrometry

NMR spectroscopy

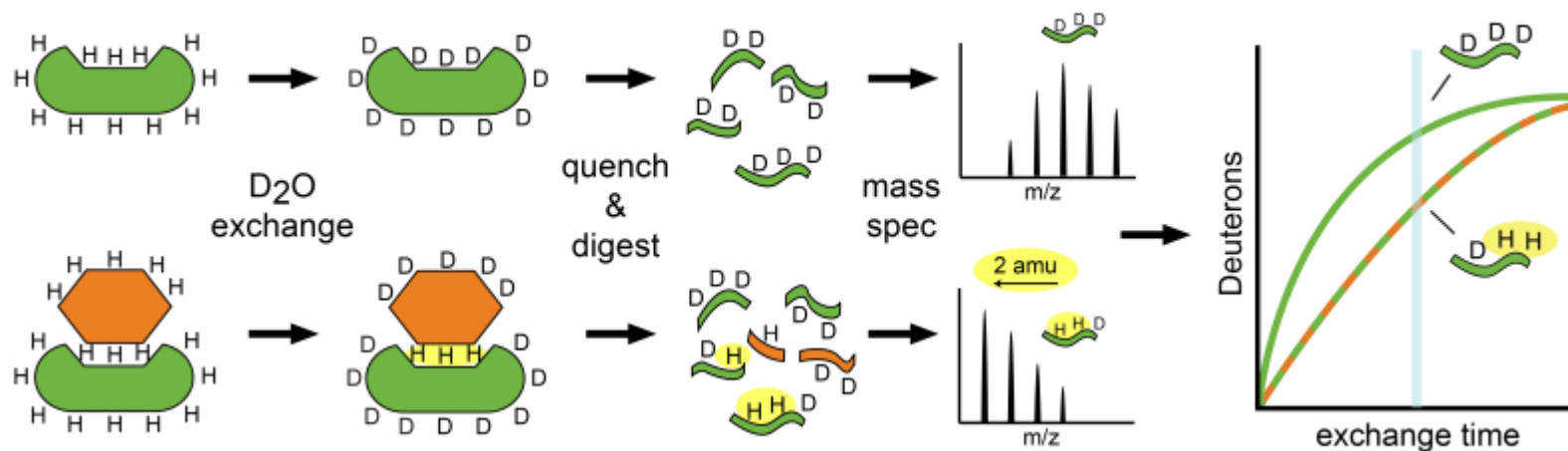
### H/D Exchange



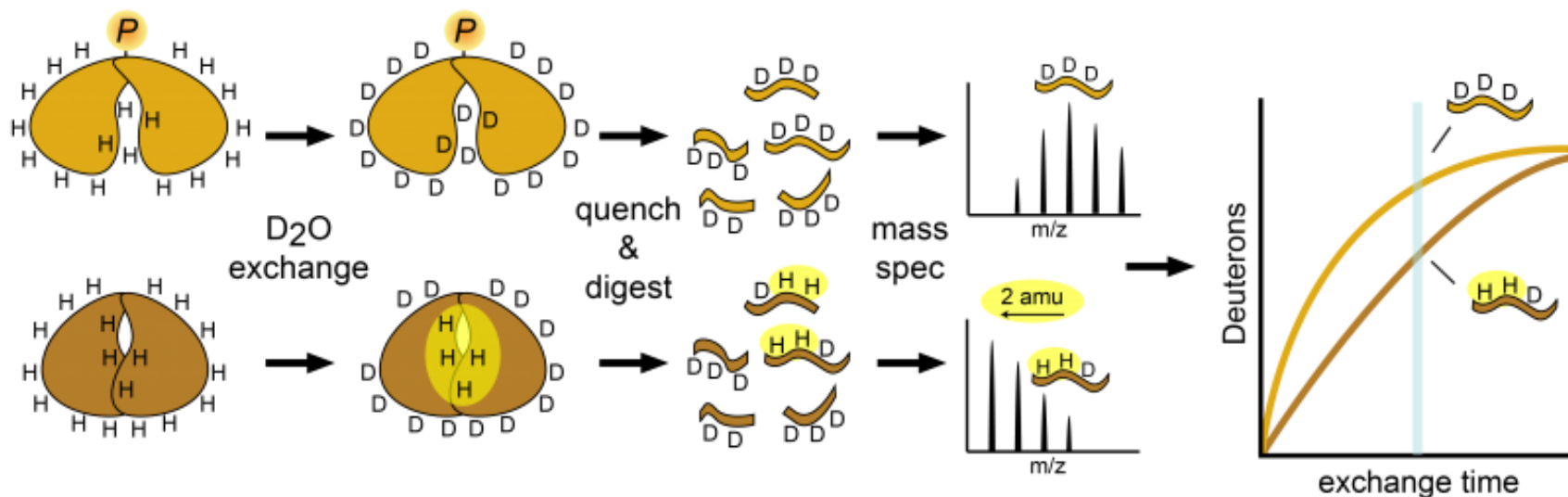
### Quench & Digest



## HDX-MS: Protein-Protein Interactions



## HDX-MS: Conformational Changes



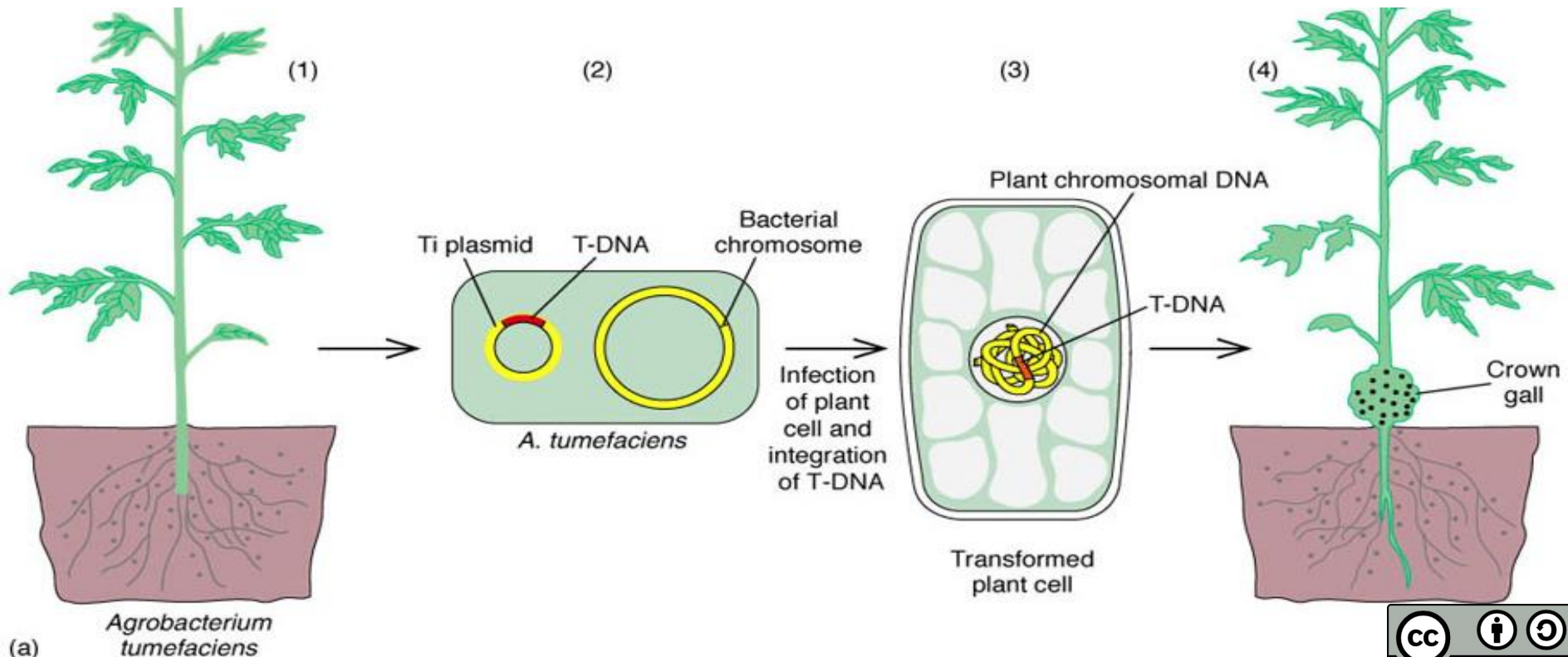


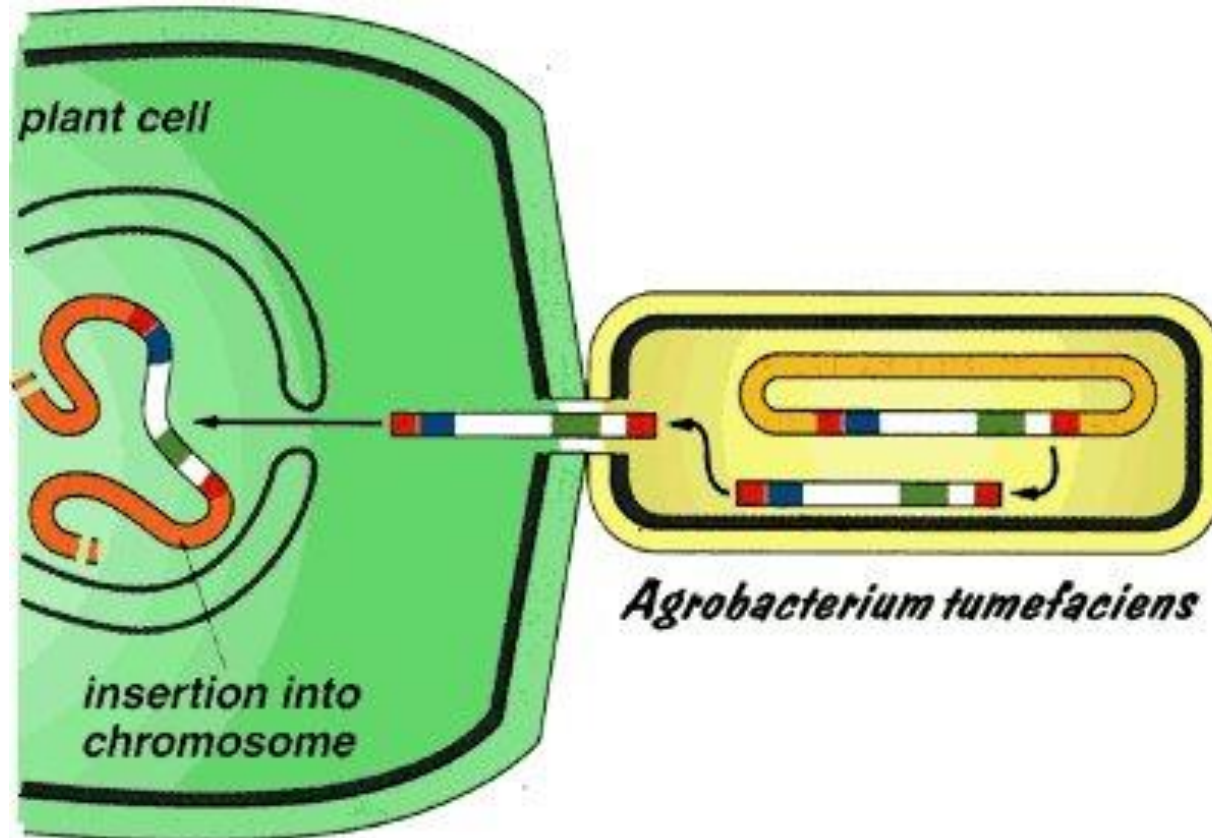
# Gene transfer to plants T-DNA (Ti - tumor inducing plasmid) and *Agrobacterium*

Natural plasmid of *A. tumefaciens* - tumors.

Infection with recombinant *A. tumefaciens*

T-DNA integration to host genome → proliferation of plant cells → crown gall.



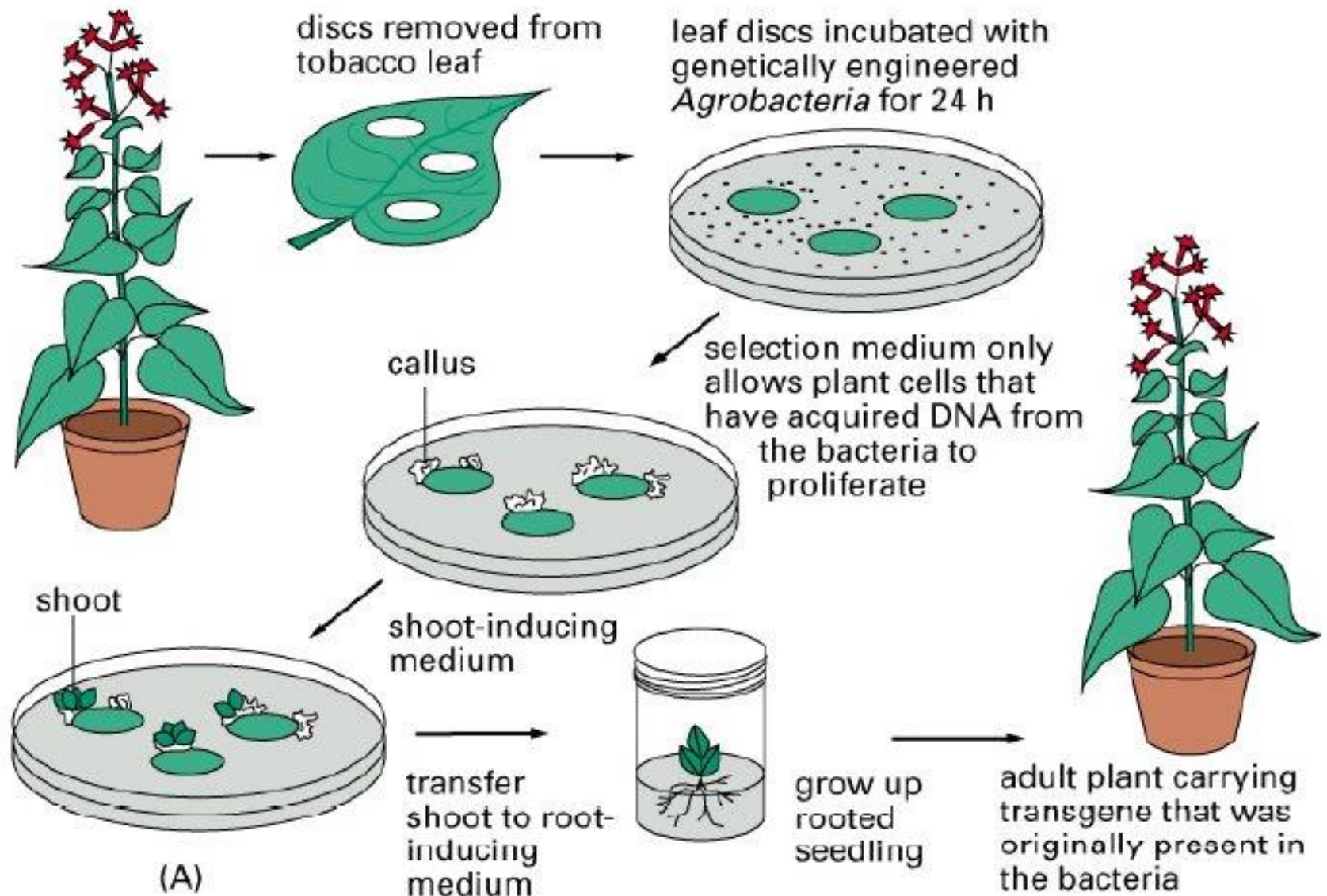


[www.plantsci.cam.ac.uk/Haseloff/SITEGRAPHICS/Agrotrans.GIF](http://www.plantsci.cam.ac.uk/Haseloff/SITEGRAPHICS/Agrotrans.GIF)

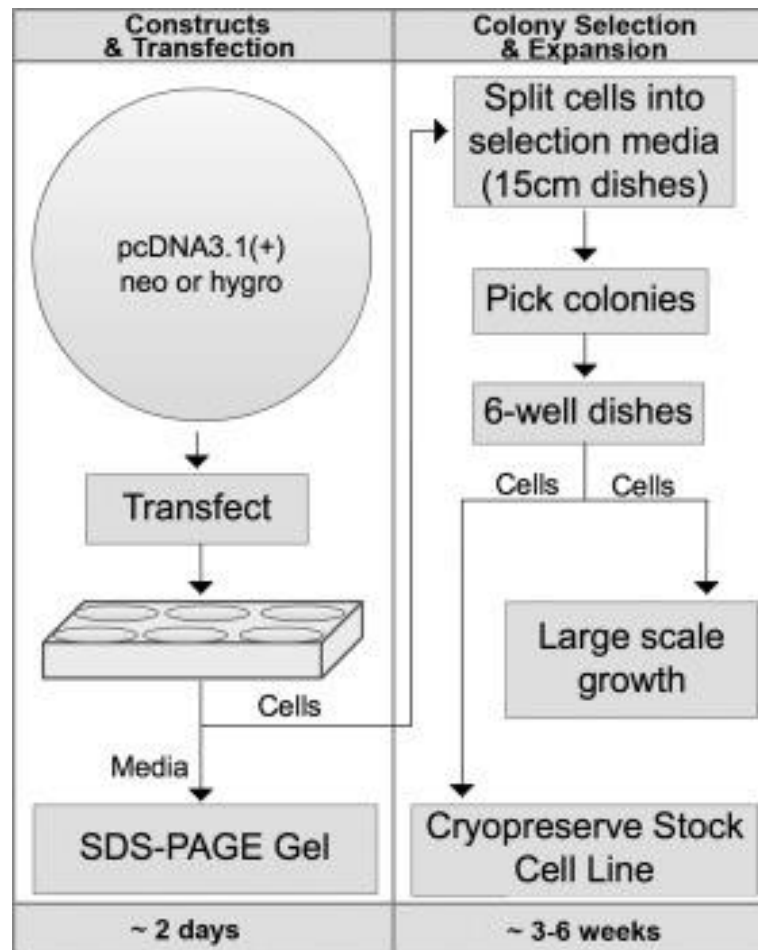
**Table 16.2. Reporter genes, assay and Identification method in transferred plants.**

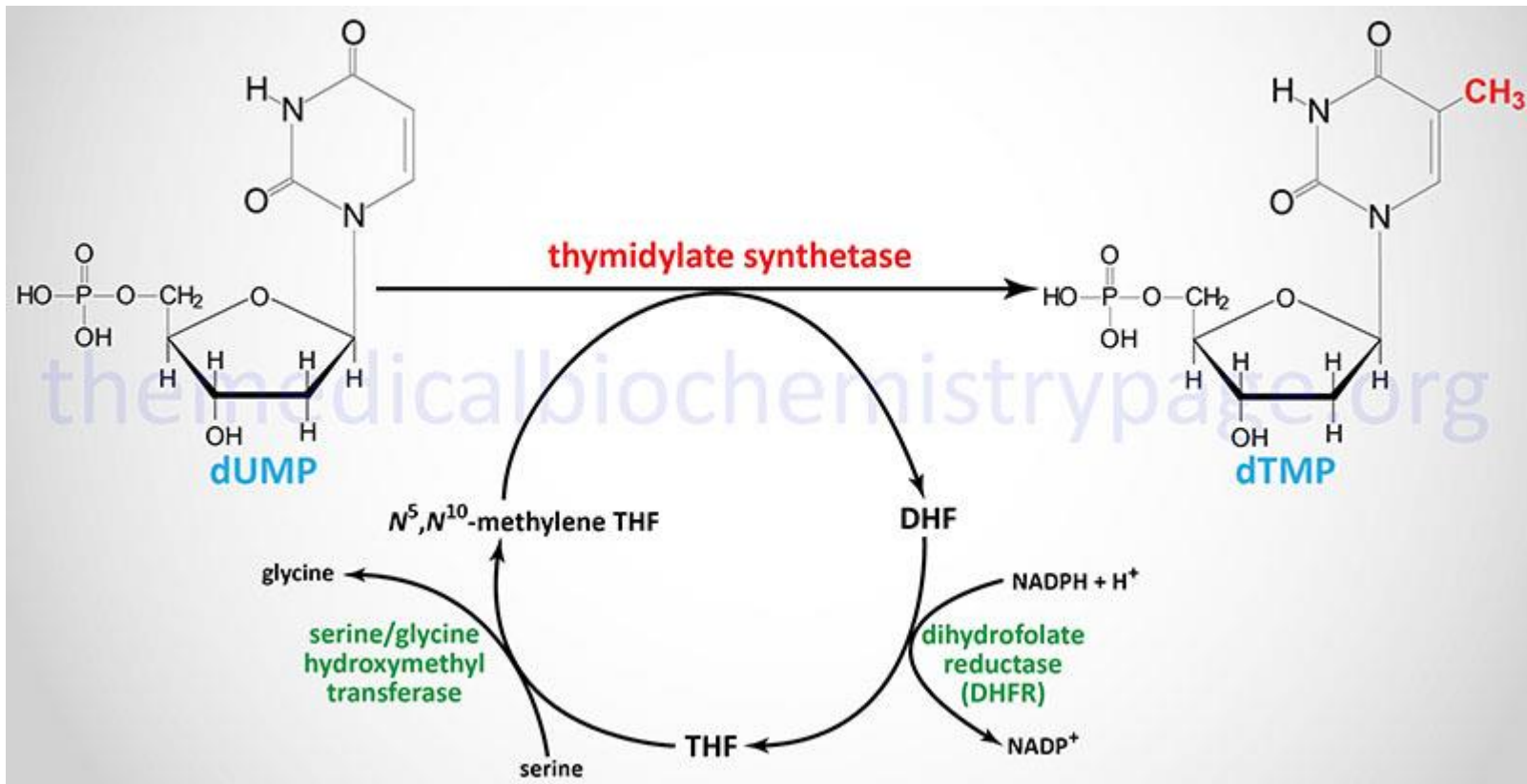
Reporter genes used for	Substrate and assay	Identification
Chloramphenicol acetyl transferase (CAT)	$^{14}\text{C}$ chloramphenicol + acetyl Co-A.	Detection of acetyl chloramphenicol by autoradiography
$\beta$ -glucuronidase (GUS)	TLC separation	Fluorescence detection colorimetric, fluorimetric and histochemical
$\beta$ -galactosidase (Lac Z)	Glucuronides (PNPG, X-GLUC, NAG, REG)	Colour of colony
Luciferase (LUC)	$\beta$ -galactoside (X-gal)	Bioluminescence (exposure of X-ray films)
Octopine synthase	Decanal and FMNH <sub>2</sub>	Electrophoresis
Nopaline synthase	ATP + O <sub>2</sub> + luciferin	
	Arginine pyruvate+NADH	
	Arginine+ketoglutaric acid+NADH	

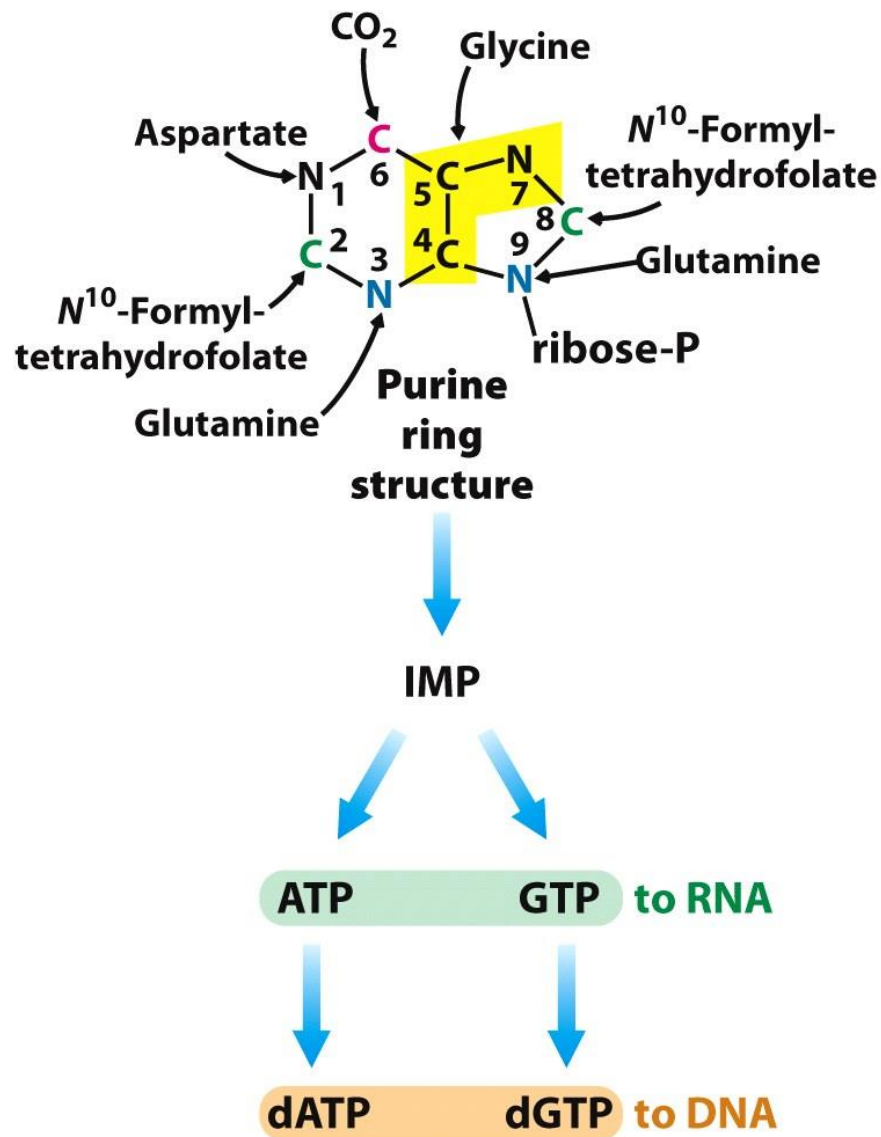




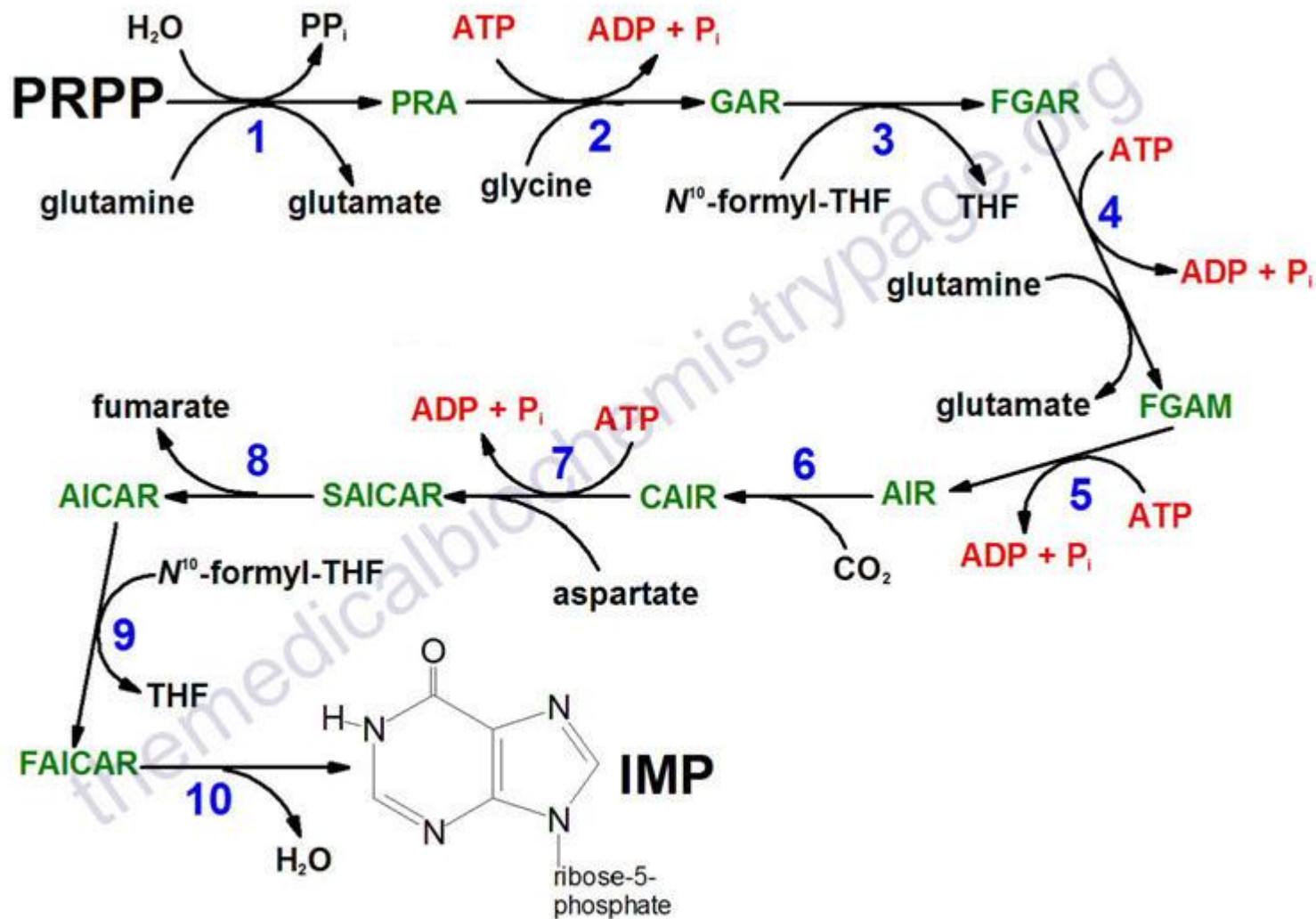
Alberts et al. (2002) Molecular Biology of the Cell 4/e



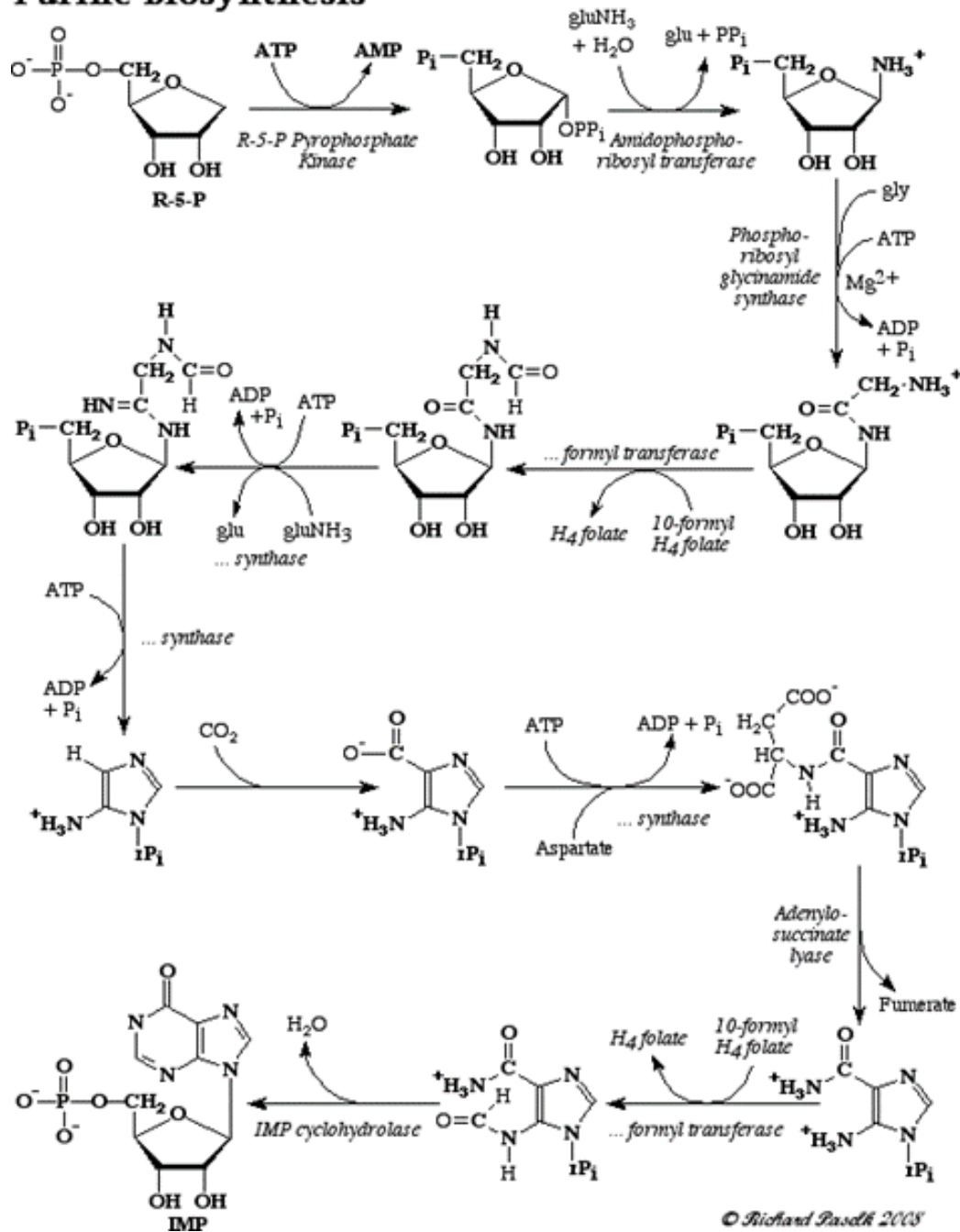




**Figure 25.5**  
*Biochemistry, Seventh Edition*  
© 2012 W. H. Freeman and Company



# Purine Biosynthesis

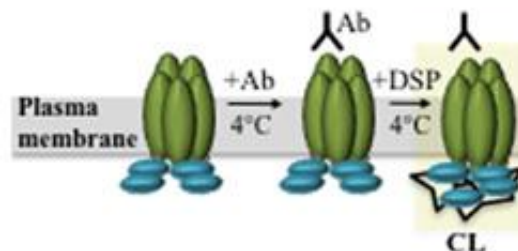


© Richard Pasdel 2008

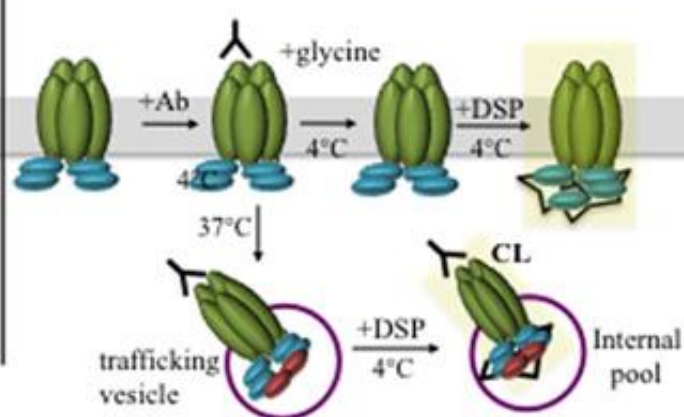


**A**

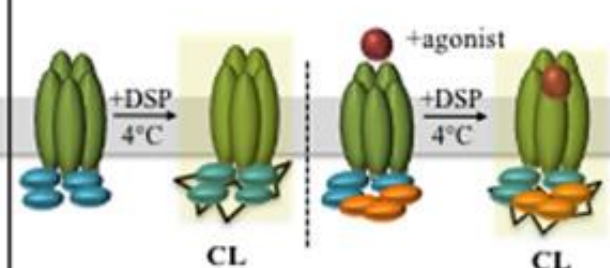
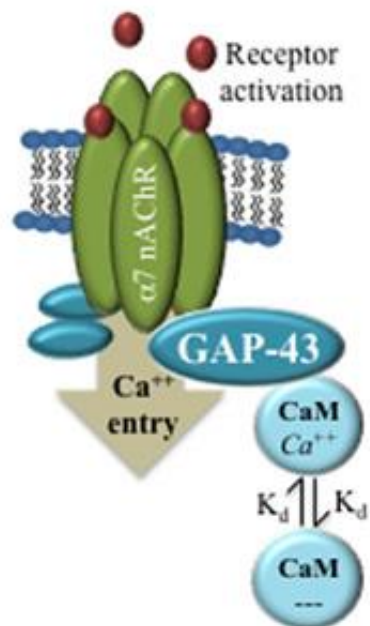
### Analysis of cell surface interactomes



### Comparison of cell surface vs. internal receptor pool interactomes

**B**

### Identification of proteins that associate with activated receptors

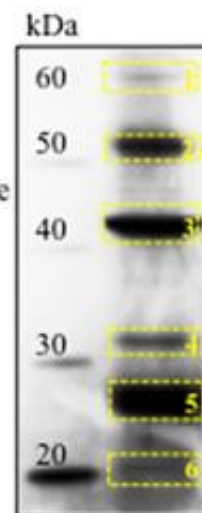
**C**

### 1) Receptor solubilization

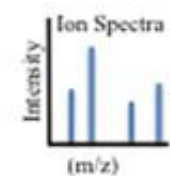
### 2) Crosslink proteins with BS<sub>3</sub>

then IP with an anti- $\alpha 7$  antibody

### 3) Separate the complex on SDS-PAGE

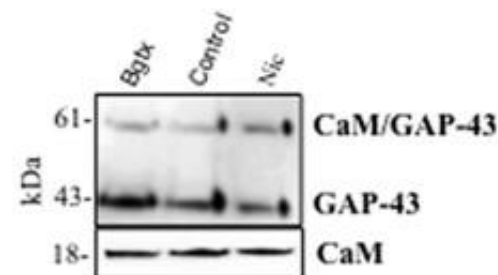


4) Coomassie stain and excise gel bands



Band	Protein
1	CaM/GAP-43
2	$\alpha 7$ nAChR
3	GAP-43
4	Calpain
5	IgG
6	CaM

Electrospray ionization (ESI) mass spectrometry



6) Western blot confirmation and quantification of protein



Uveřejněné materiály jsou určeny studentům Vysoké školy chemicko-technologické v Praze

jako studijní materiál. Některá textová i obrazová data v nich obsažená jsou převzata

z veřejných zdrojů. V případě nedostatečných citací nebylo cílem autora/ů záměrně poškodit event. autora/y původního díla.

S eventuálními výhradami se prosím obračejte na autora/y konkrétního výukového materiálu, aby

bylo možné zjednat nápravu.

The published materials are intended for students of the University of Chemistry and Technology, Prague as a study material. Some text and image data contained therein are taken from public sources. In the case of insufficient quotations, the author's intention was not to intentionally infringe the possible author(s) rights to the original work.

If you have any reservations, please contact the author(s) of the specific teaching material in order to remedy the situation.

# Sequence analysis



EUROPEAN UNION  
European Structural and Investing Funds  
Operational Programme Research,  
Development and Education



MINISTRY OF EDUCATION,  
YOUTH AND SPORTS

# Pyrosequencing

**Determination of sequence ~ 20 million bp  
in one apparatus in less than 6 h**

**Detection of released pyrophosphate during  
polymerization reaction – luminescence**

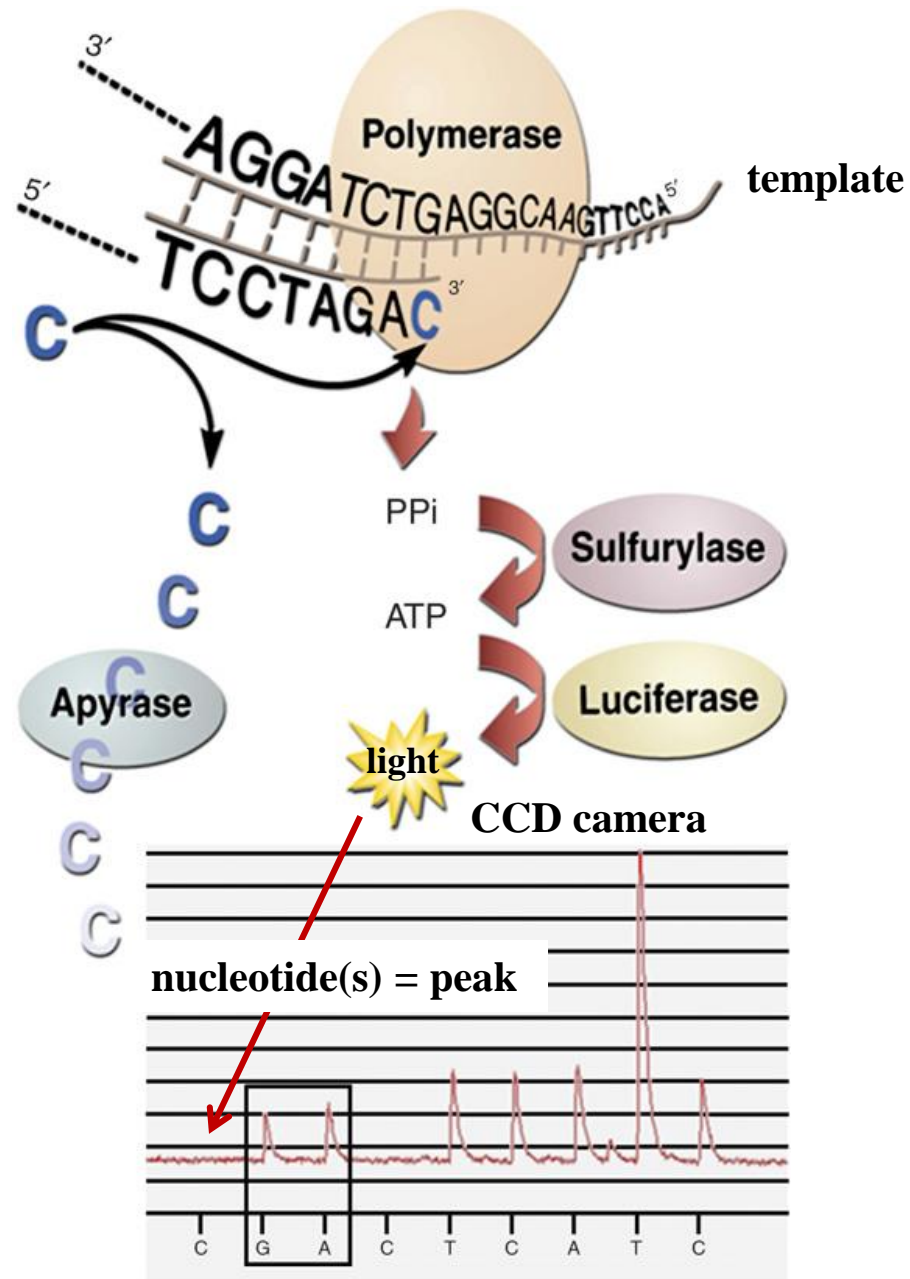
**DNA fragments ~100 bp denatured – individual ss fragments – attached to microscopic beads – individually separated.**

**PCR amplification ~ 10 millions identical copies.**

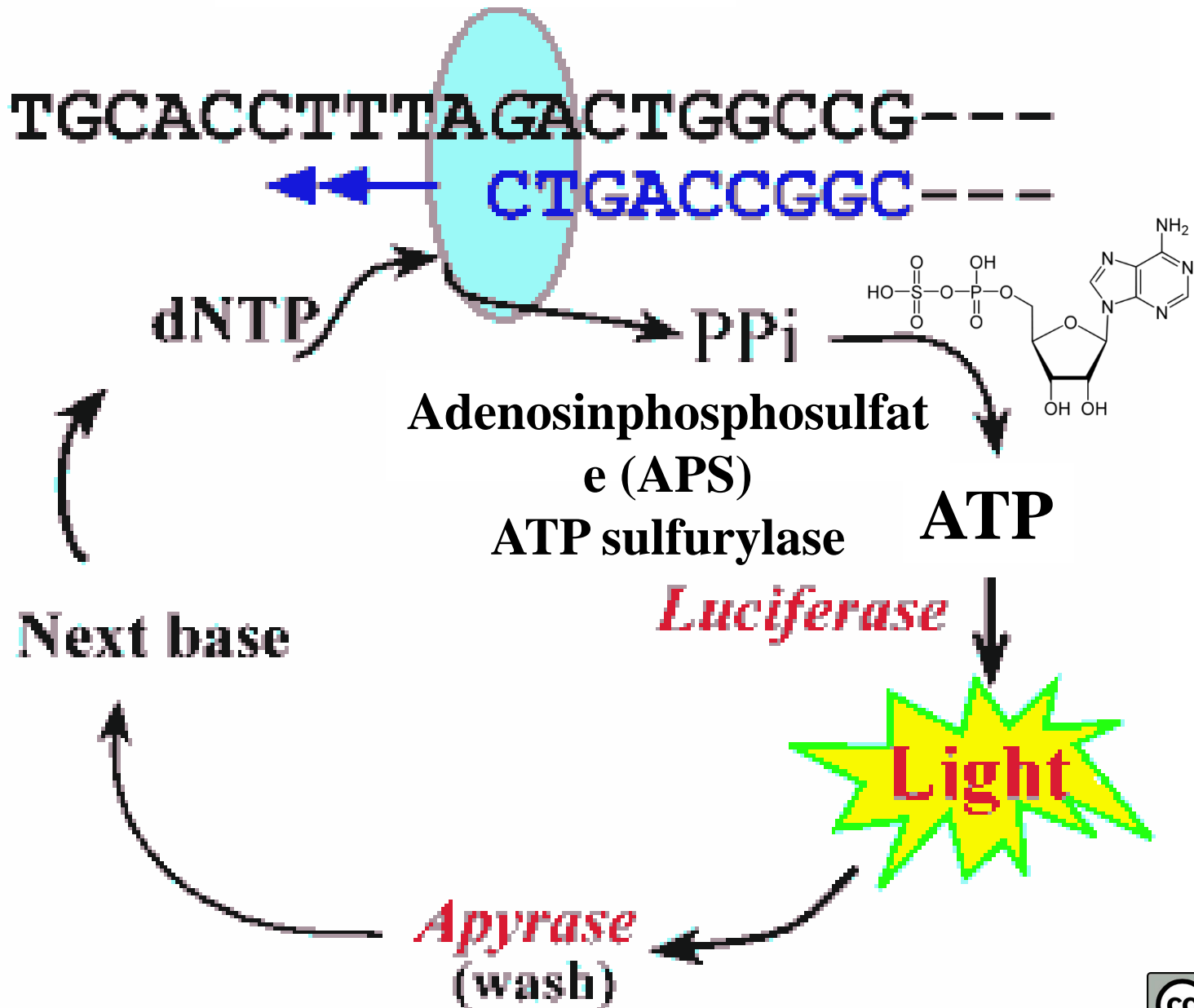
**Beads – distributed individually in microwells (~200,000 wells).**

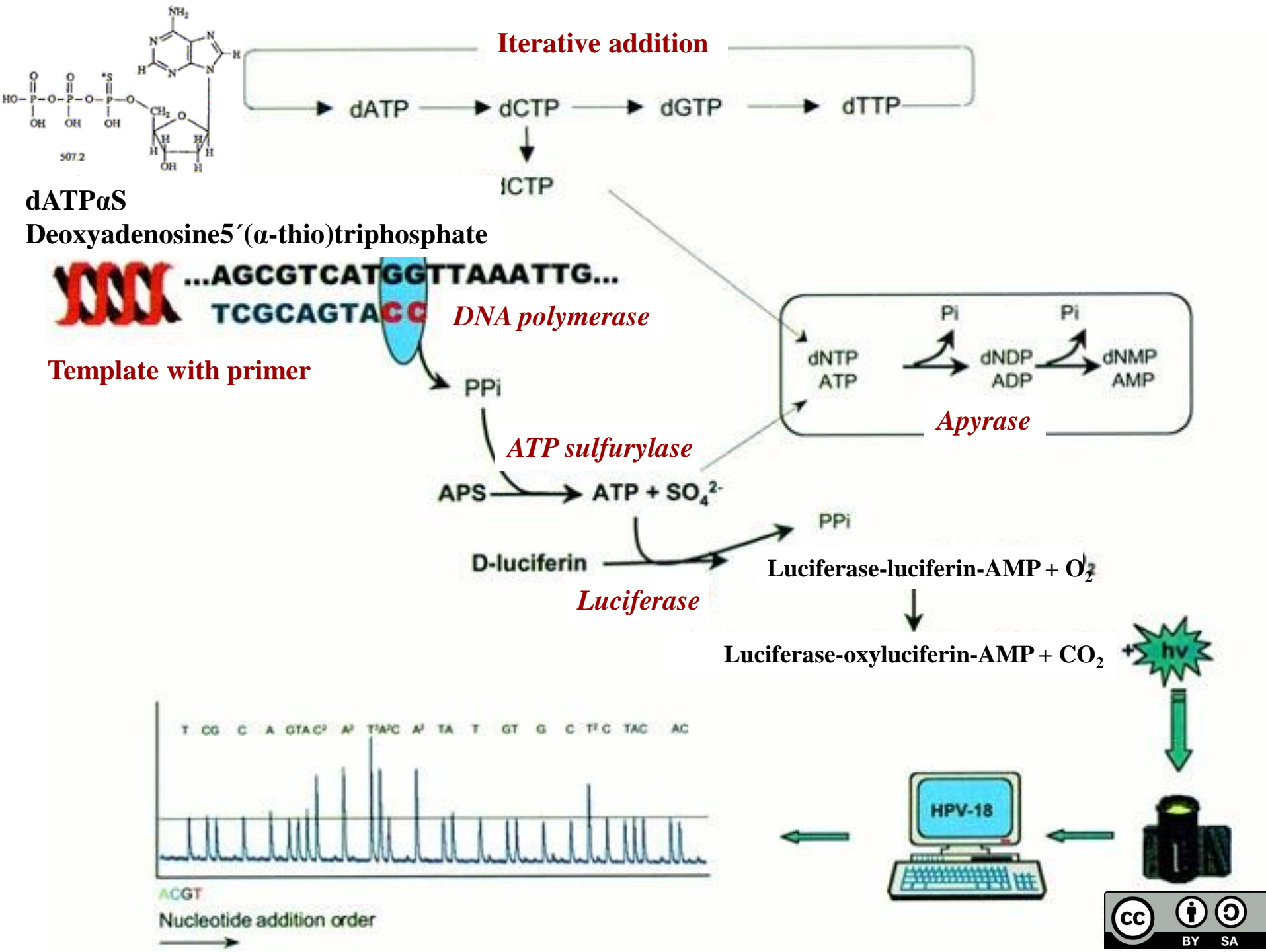
**Reaction mixture:**

- **DNA polymerase**
  - **adenosinephosphosulfate (APS)**
  - **ATP sulfurylase — generates ATP from APS and pyrophosphate (PPi).**
  - **luciferin + luciferase — conversion of luciferin to oxyluciferin**
- light emission**



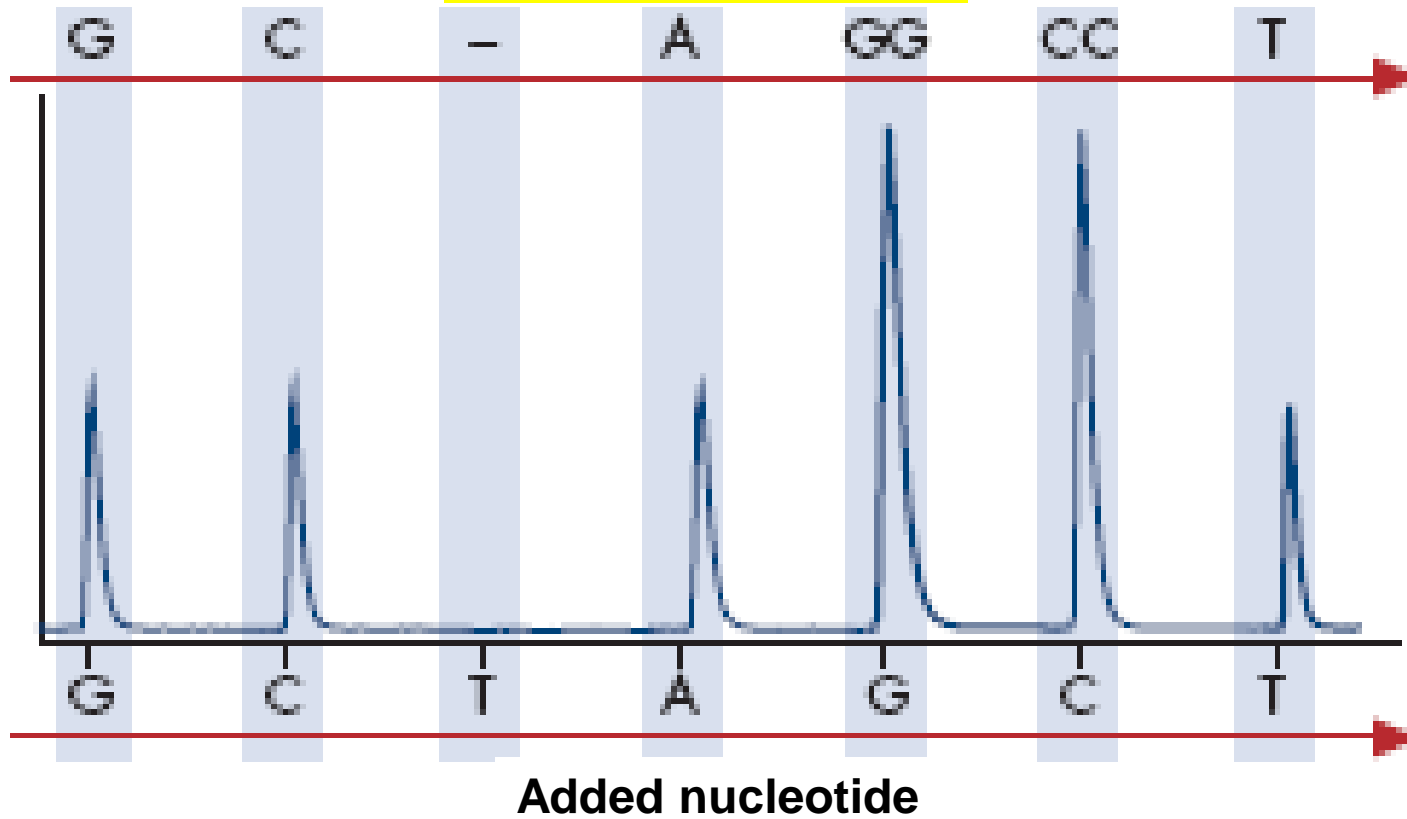
# DNA polymerase

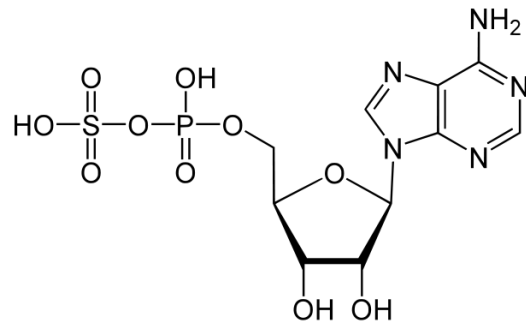
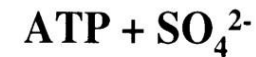




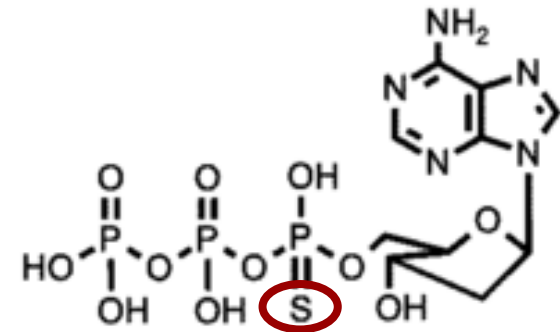


## Nukleotide sequence





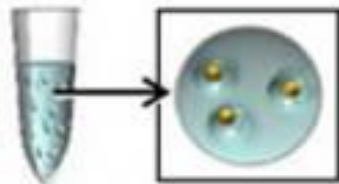
APS Adenosine phosphosulfate



Deoxyadenosine 5'( $\alpha$ -thio)triphosphate  
(no substrate for luciferase)

# Preparation of beads with clones of fragments

## Emulsion-based clonal amplification



Anneal sstDNA  
to an excess of  
DNA Capture  
Beads



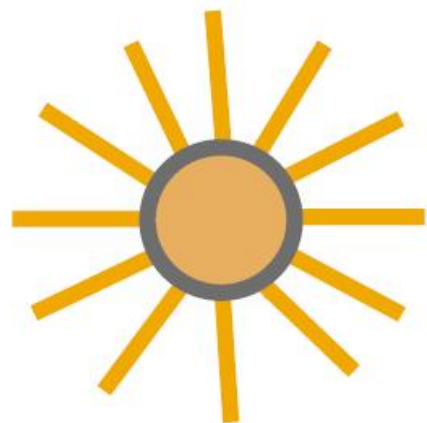
Emulsify beads  
and PCR reagents  
in water-in-oil  
micro reactors



Clonal amplification  
occurs inside micro  
reactors



Break micro  
reactors,  
enrich for  
DNA-positive



P1-coupled beads

+



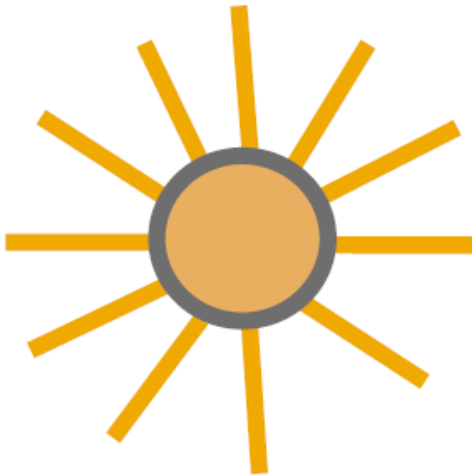
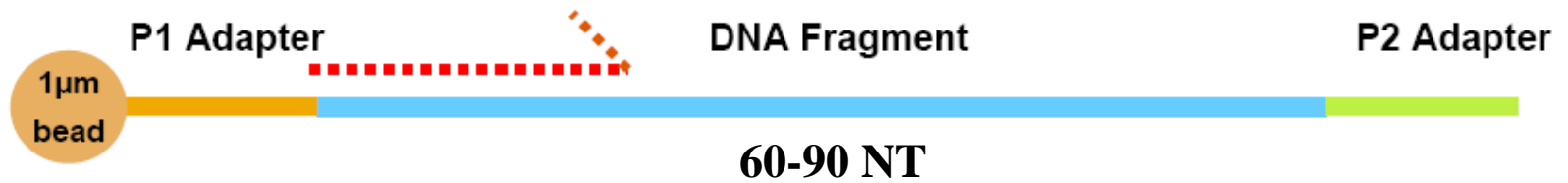
Templates



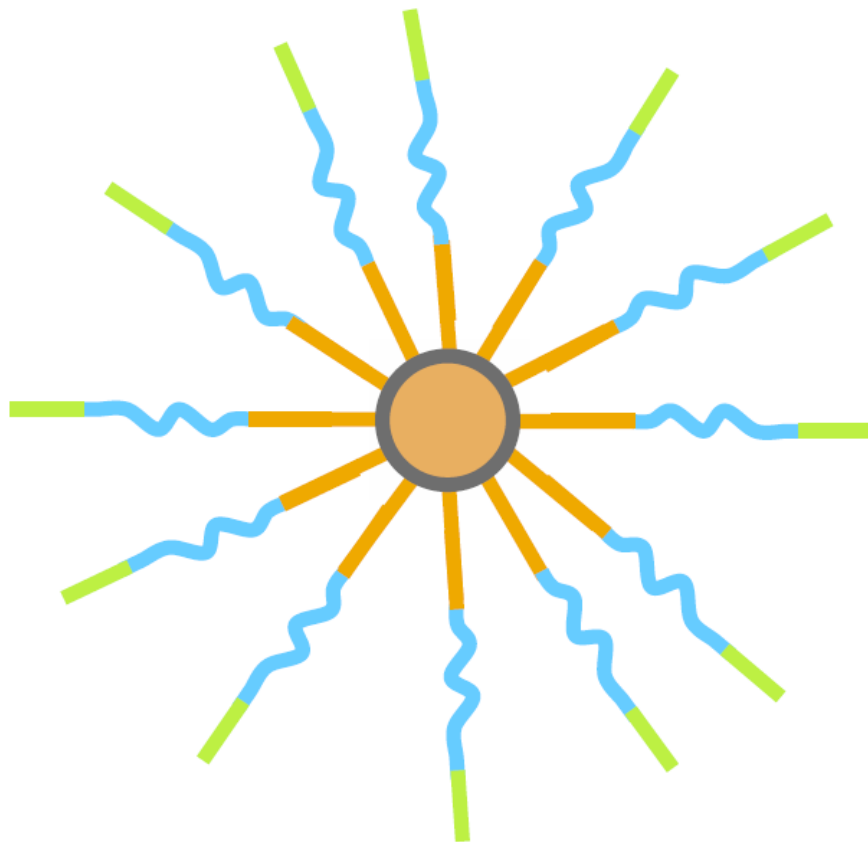
Primers P1 << P2



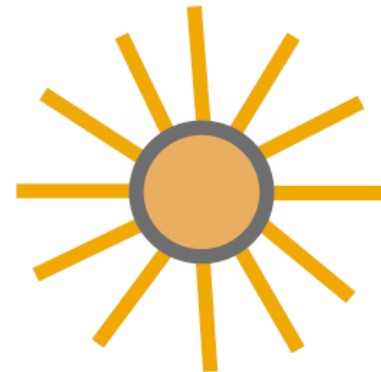
Enzyme



- 1) Annealing of template to P1
- 2) Extension of primer complementary to P1
- 3) Dissociation of template

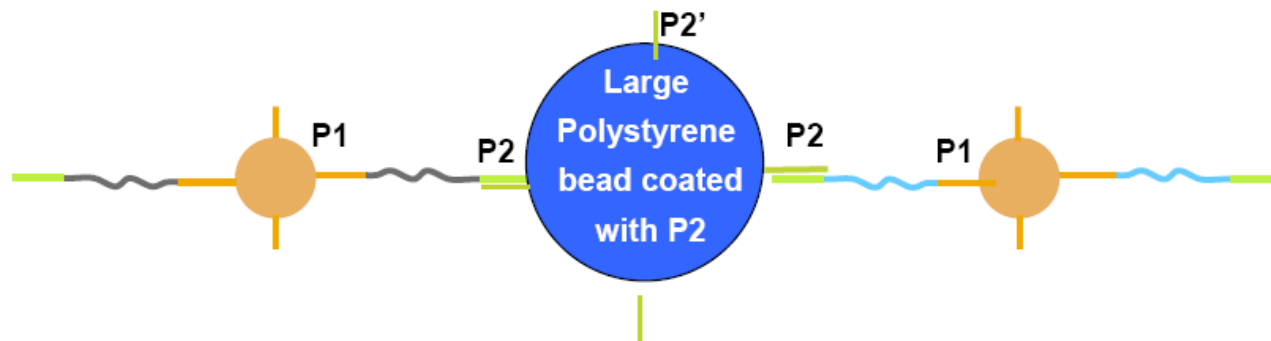


Bead contains ~20K amplified products from original single strand molecule

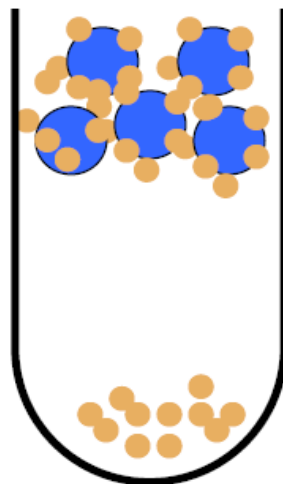


Beads with no product

3

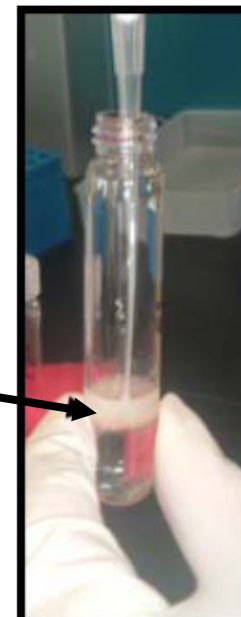


**Centrifuge in  
glycerol gradient**

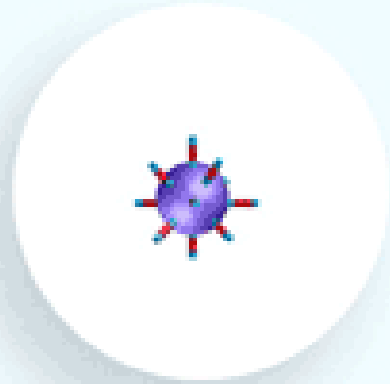


**Supernatant**  
*Captured beads with templates*

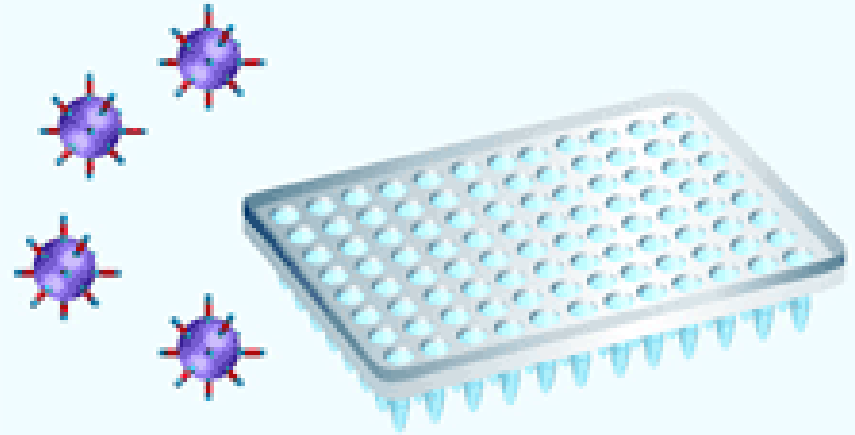
**Pellet**  
*Beads with no template*



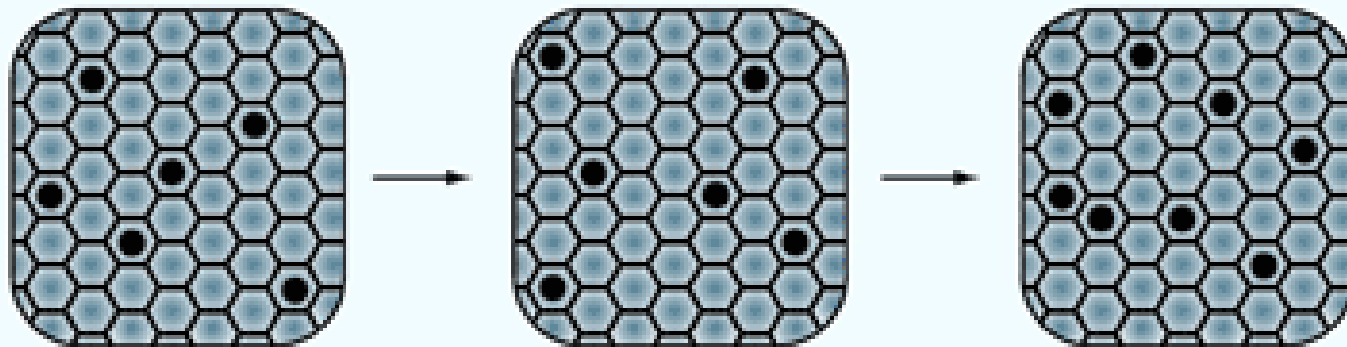




5. Clonal amplification

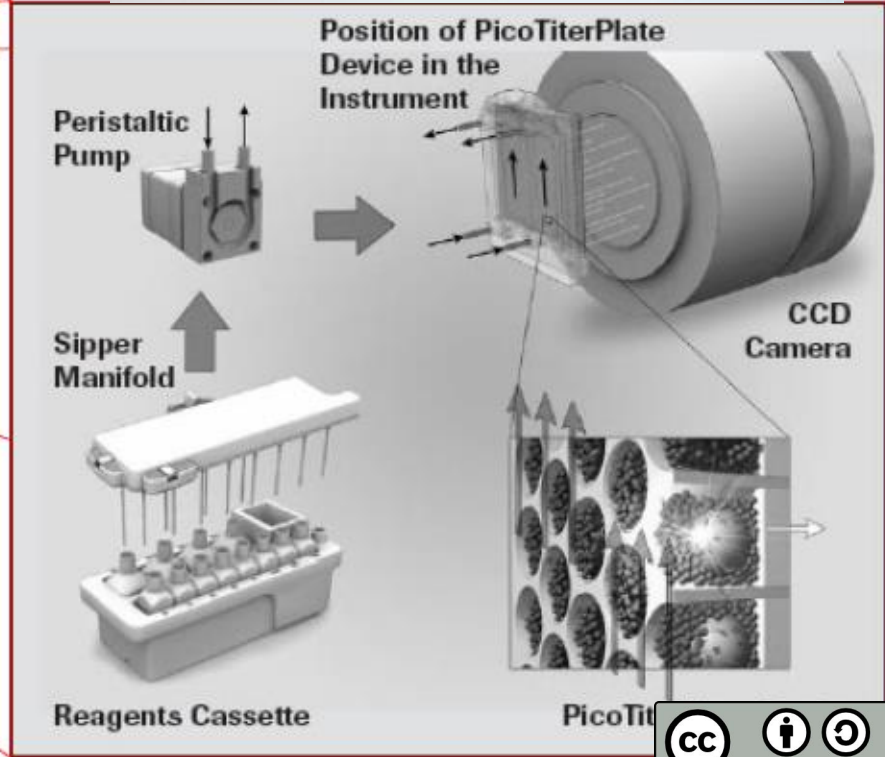
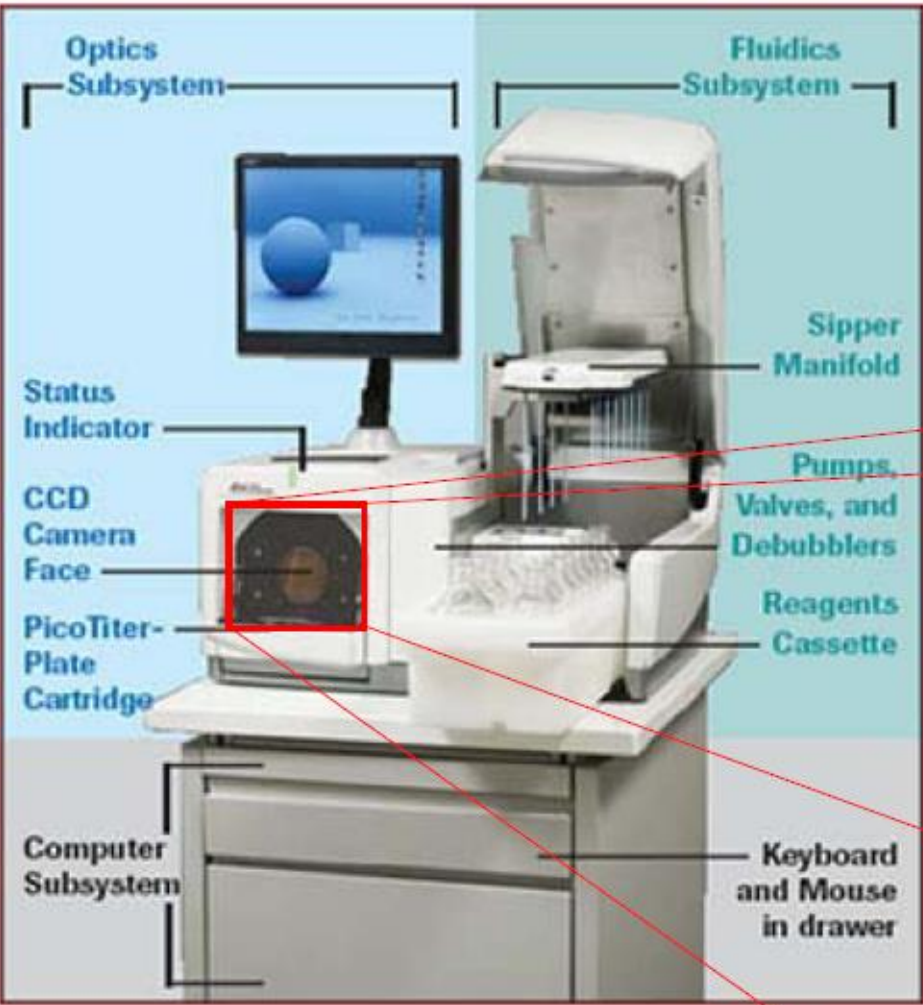
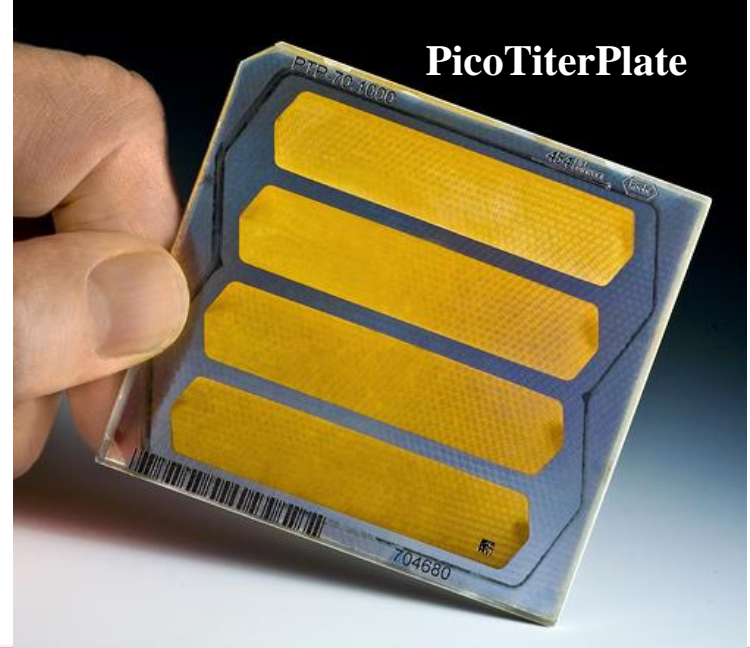


6. Beads deposited into wells on plate



7. Nucleotide complementary to template light signal recorded by CCD camera

# PicoTiterPlate



# Pyrosequencing 454 Biosciences (Roche)

**Fast** > 20 million bases ~ 5 h

**Cheap** – significantly cheaper than Sanger method

**Simple** – automated

**Efficient** - sequence of typical bacterial genome

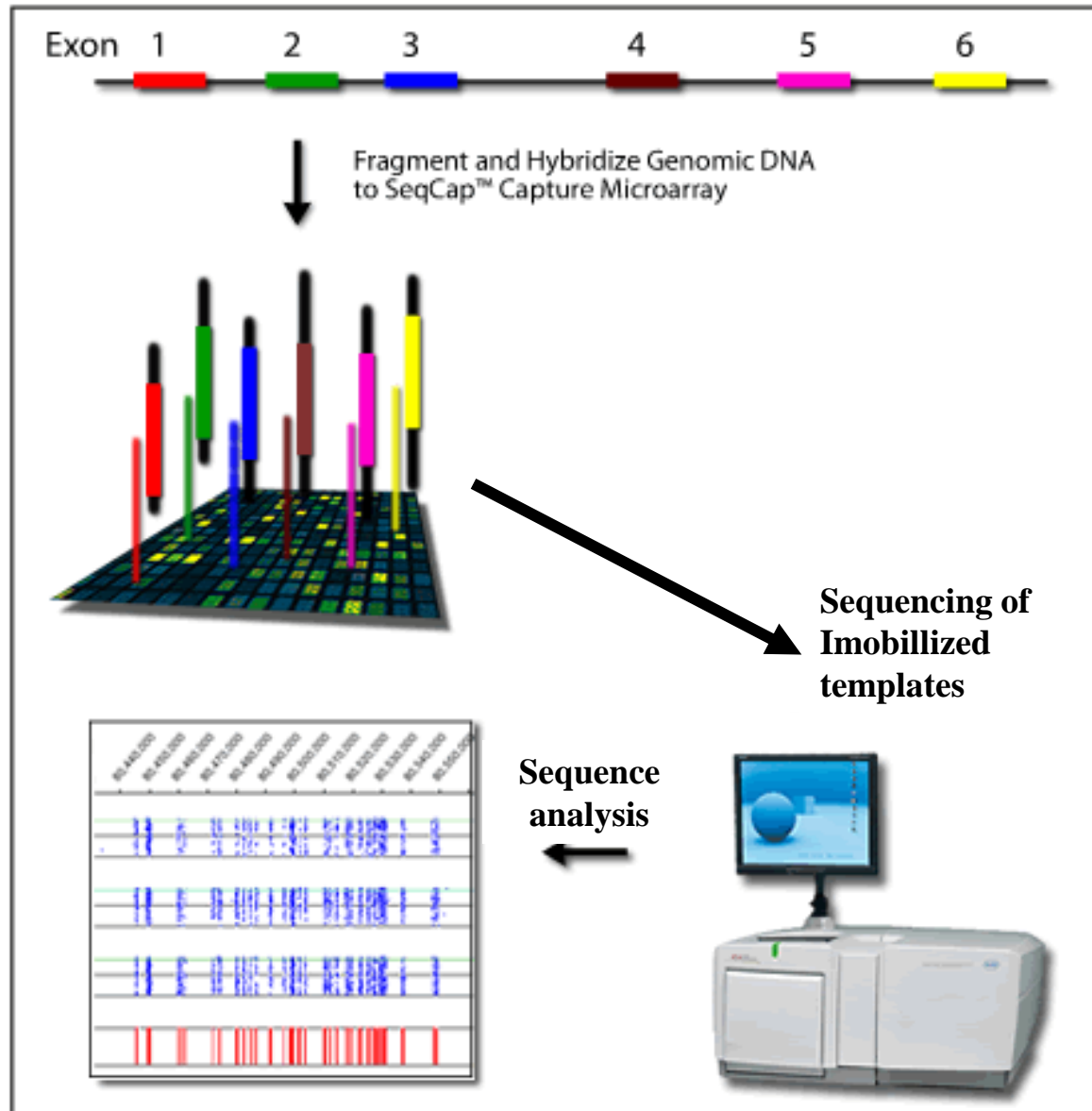
– one worker in several days, no cloning, no colony isolation

Yet not suitable **for** individual samples and large genomes

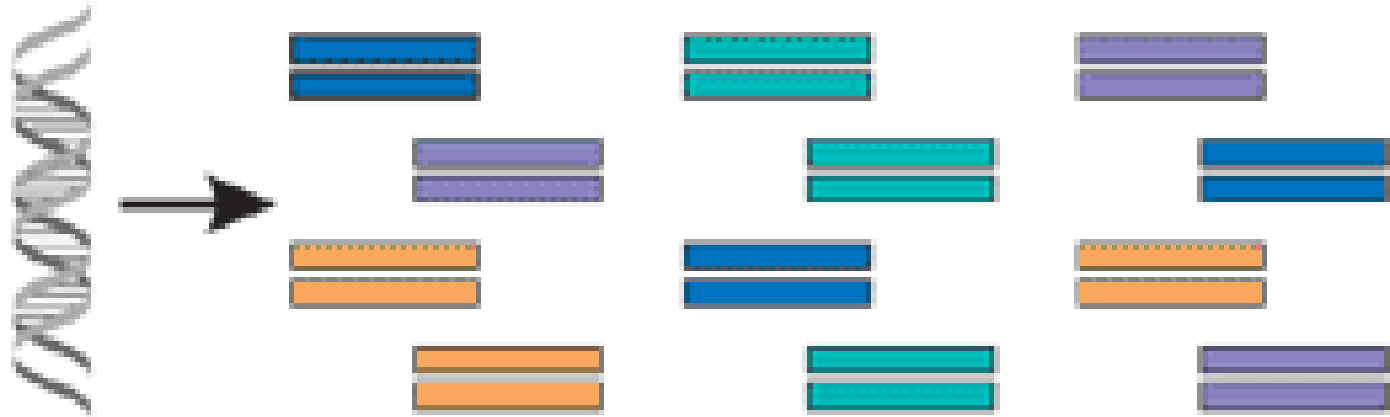
suitable for bacterial genomes and metagenomic studies

(identification of non-cultivable microorganisms)

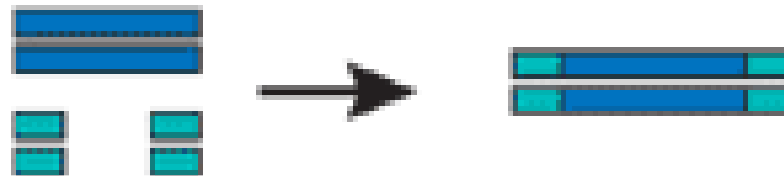
# Illumina Solexa sequencing

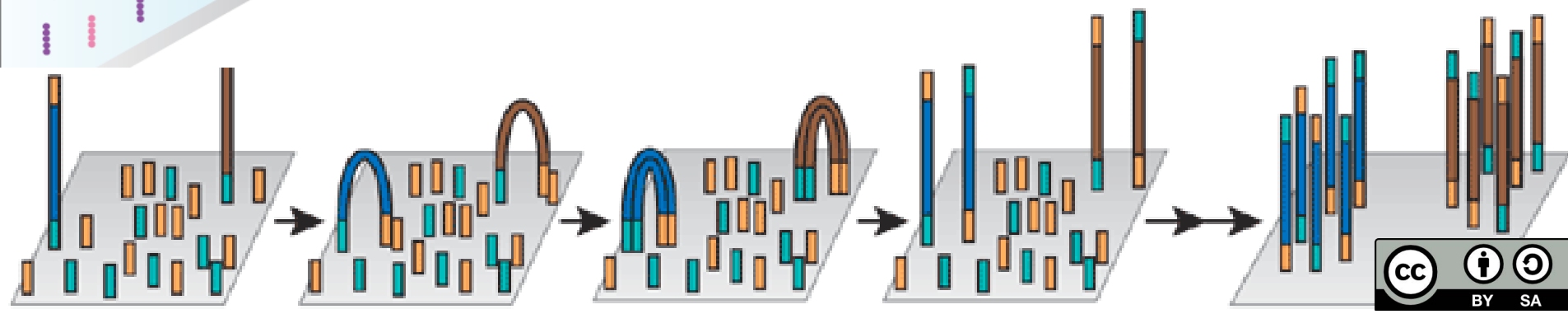
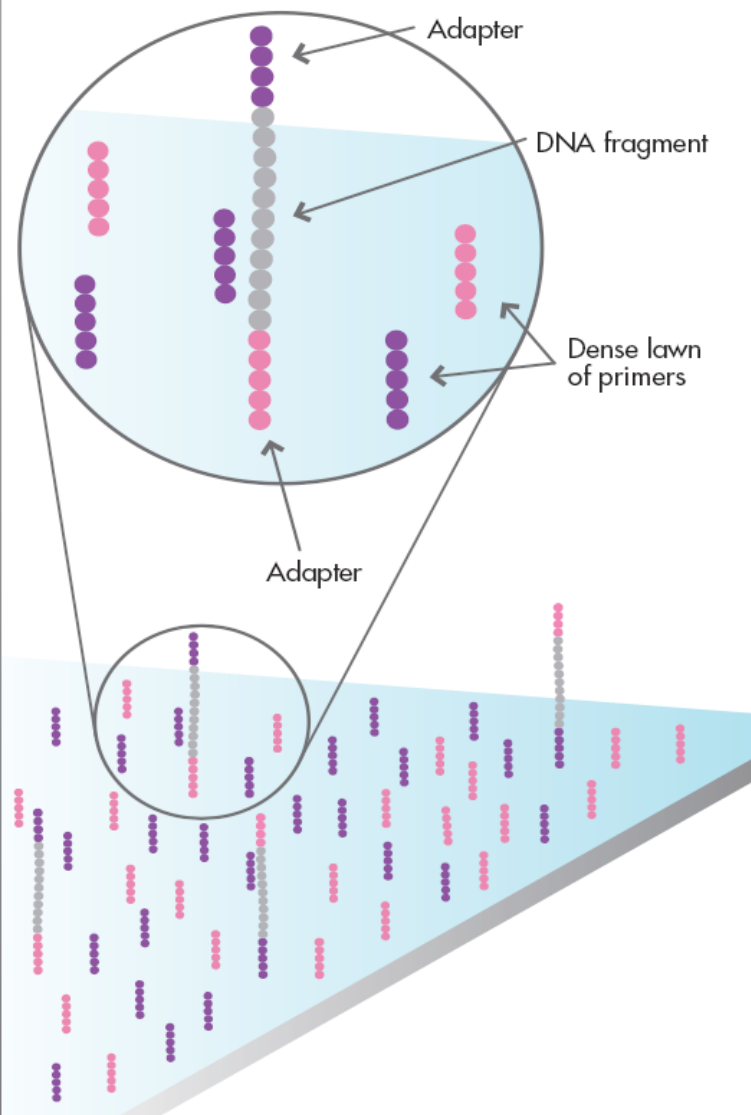


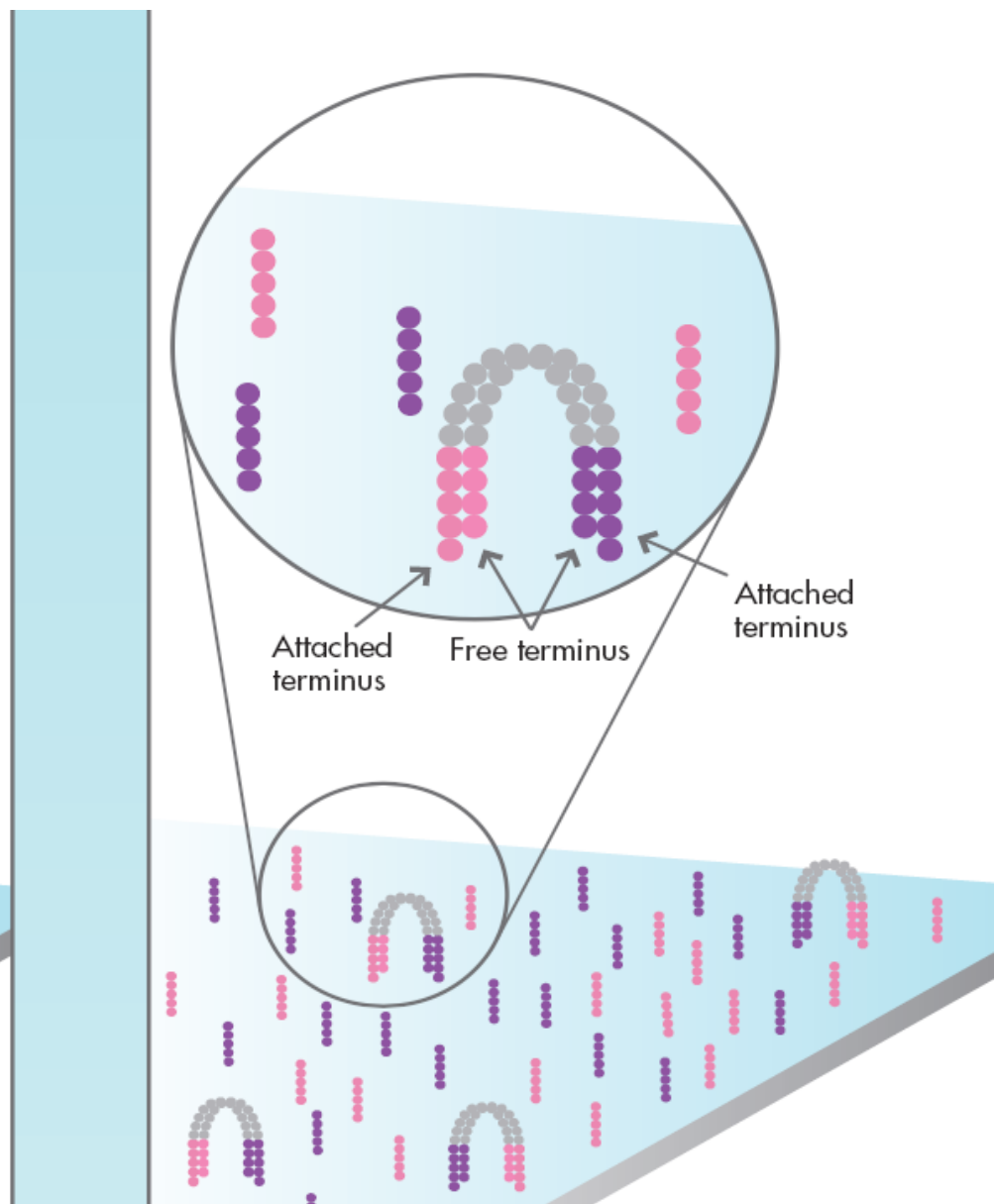
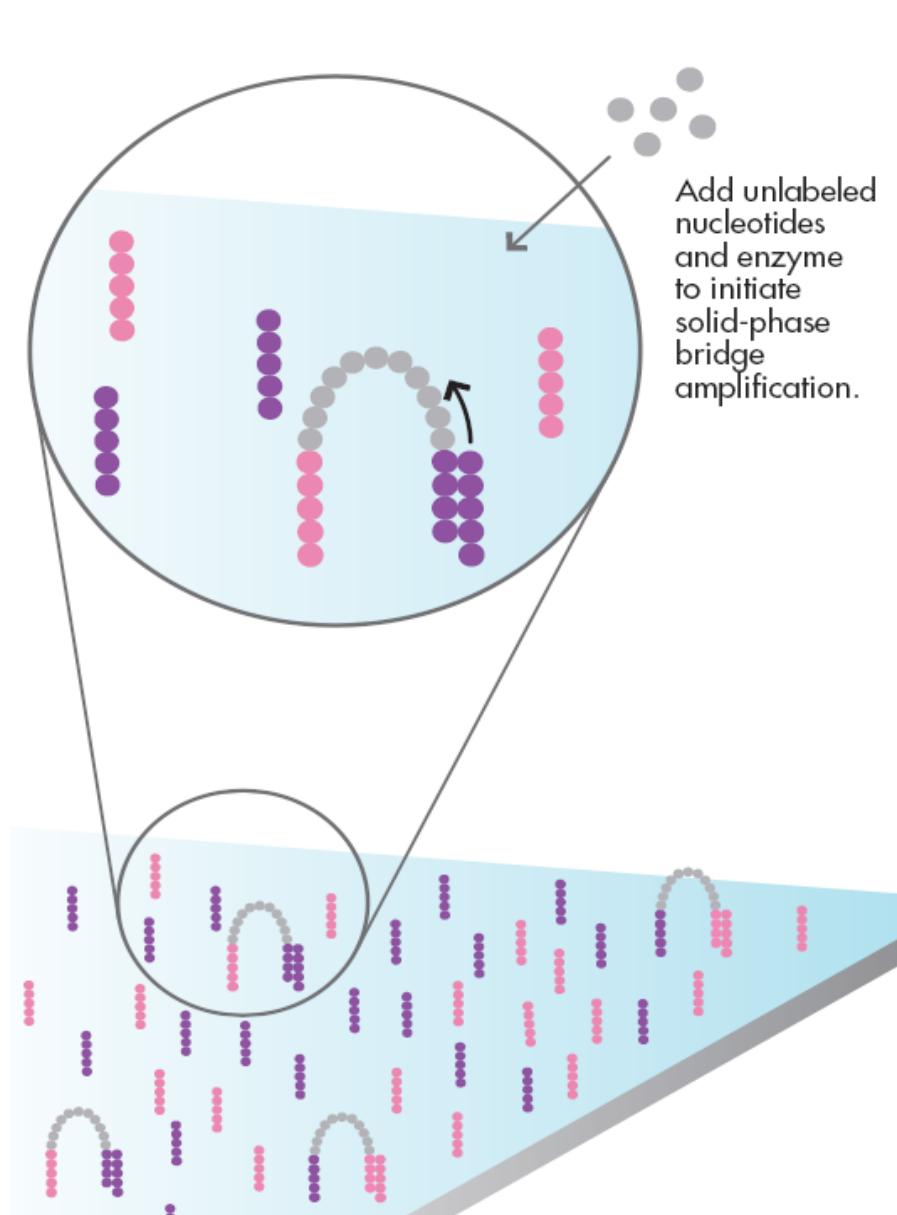
## DNA fragmentation



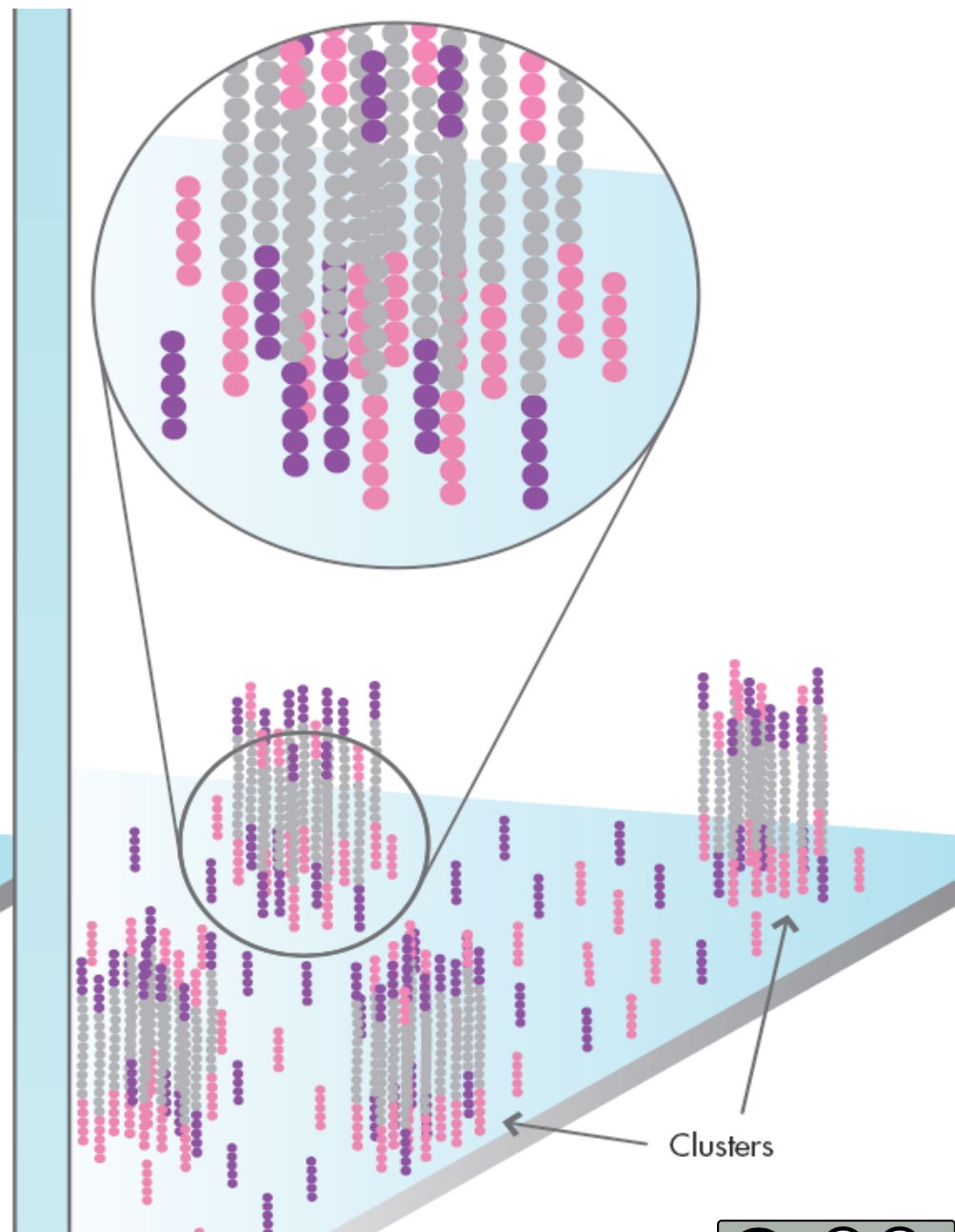
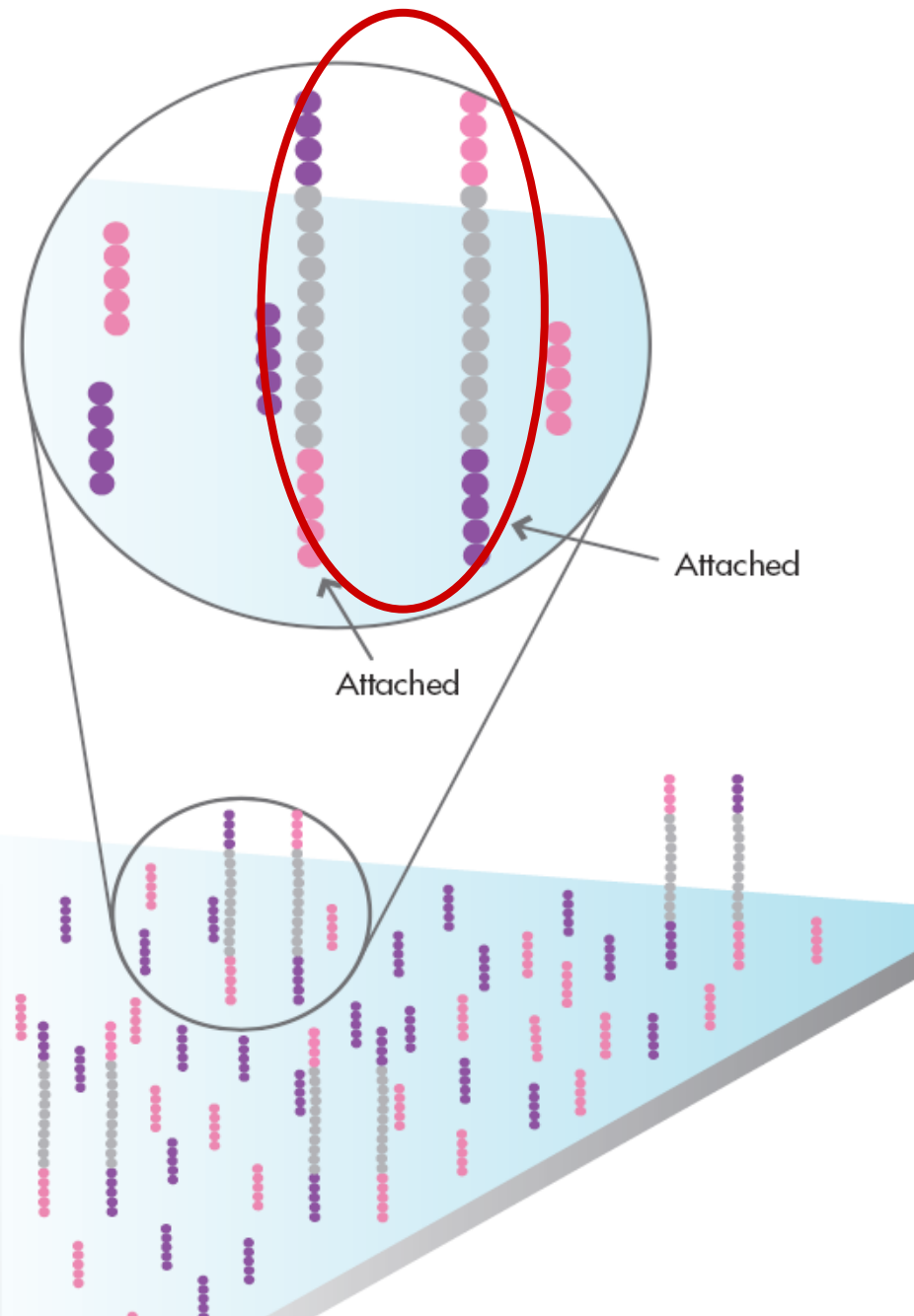
## *In vitro* adaptor ligation

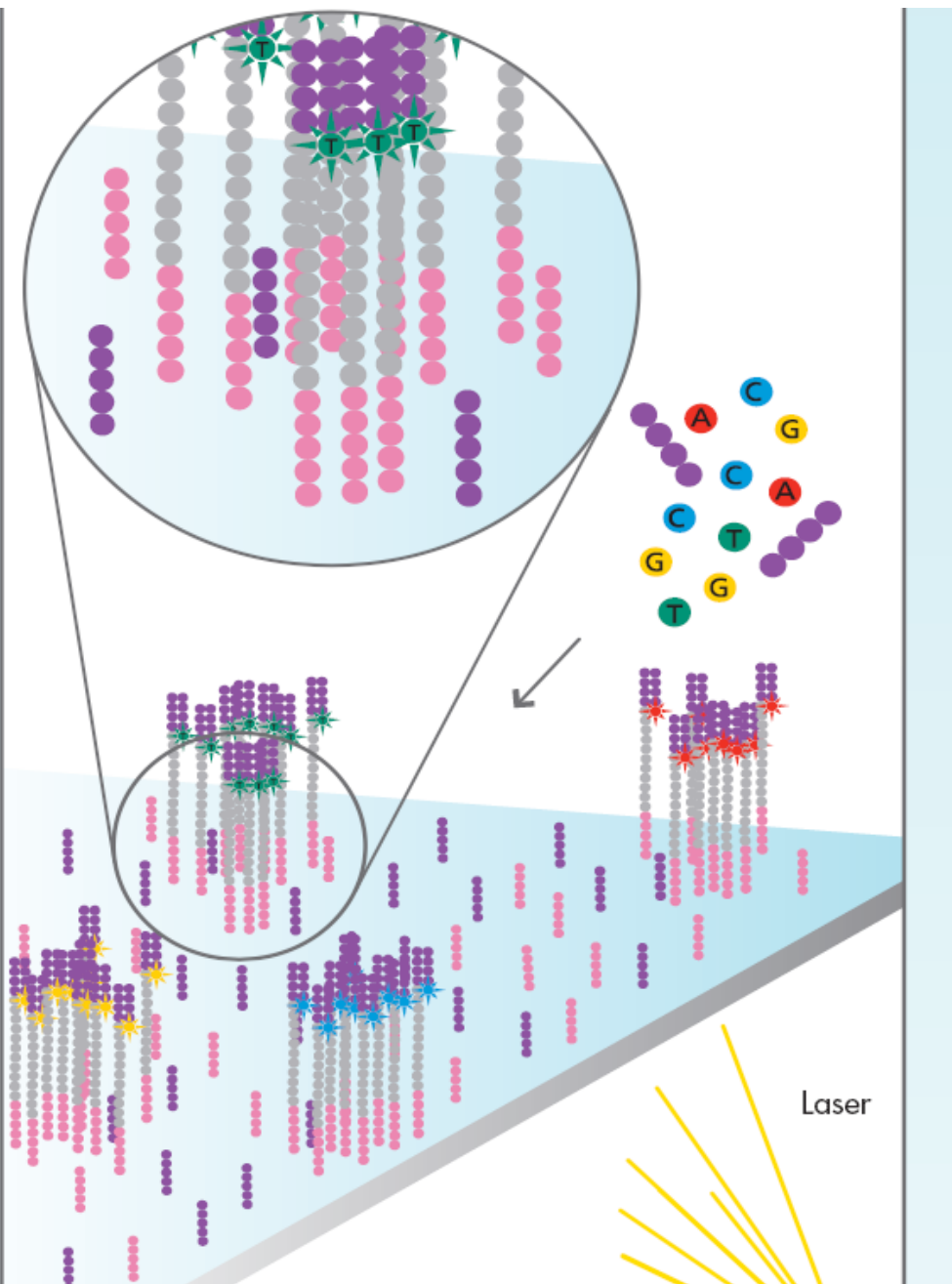




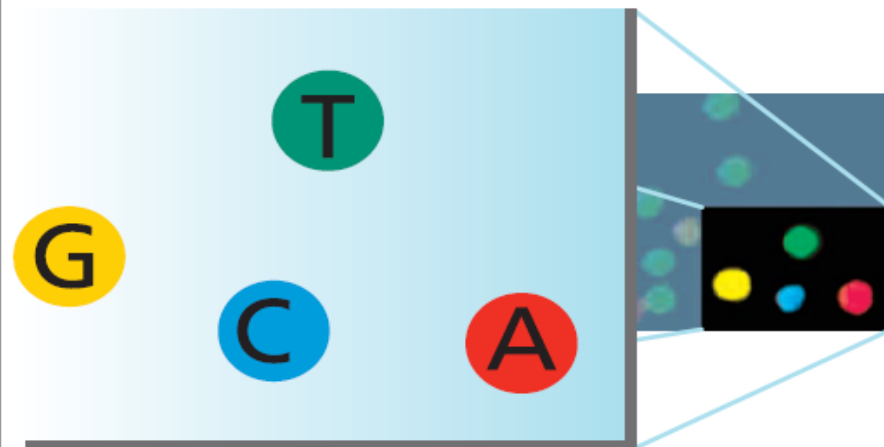




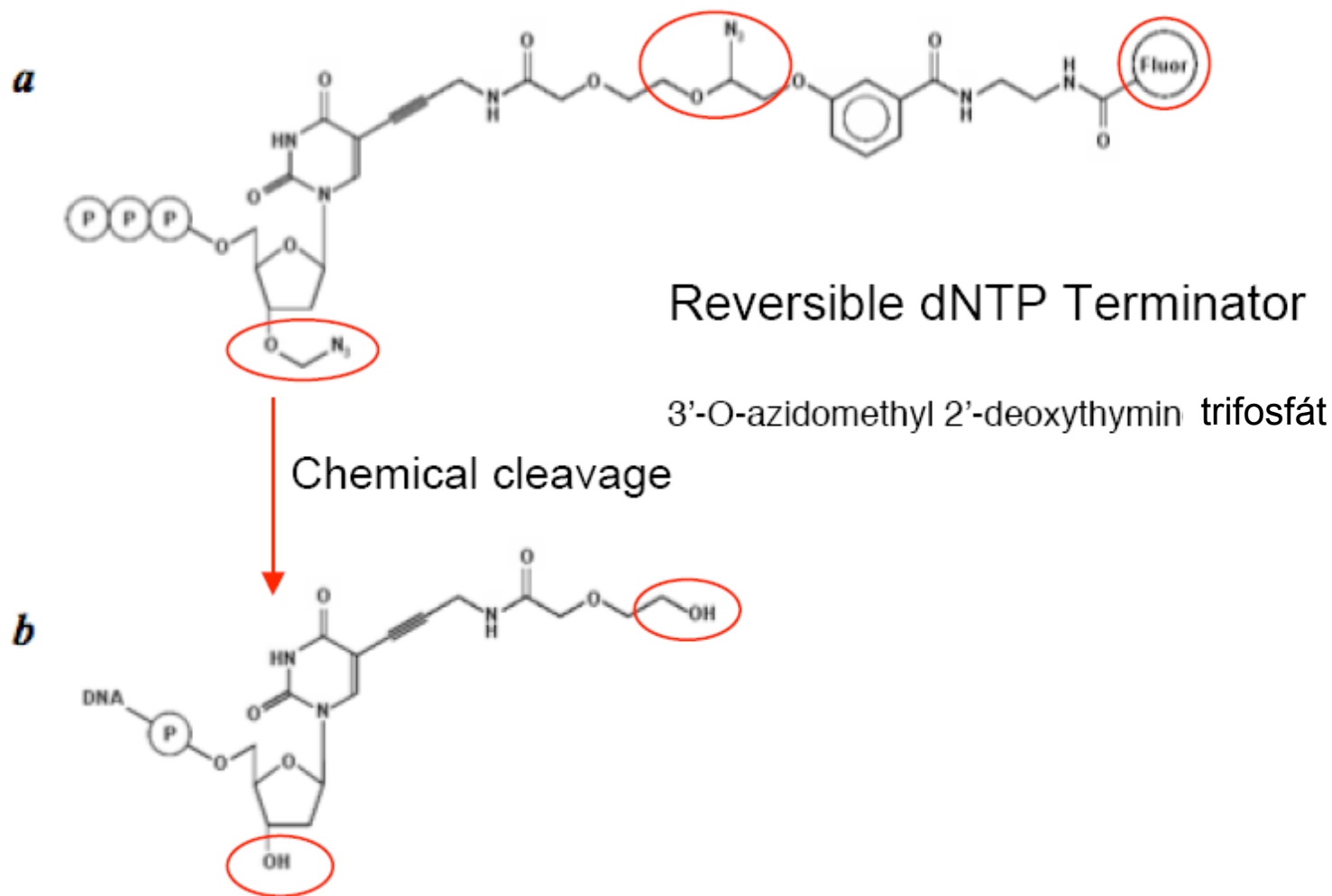




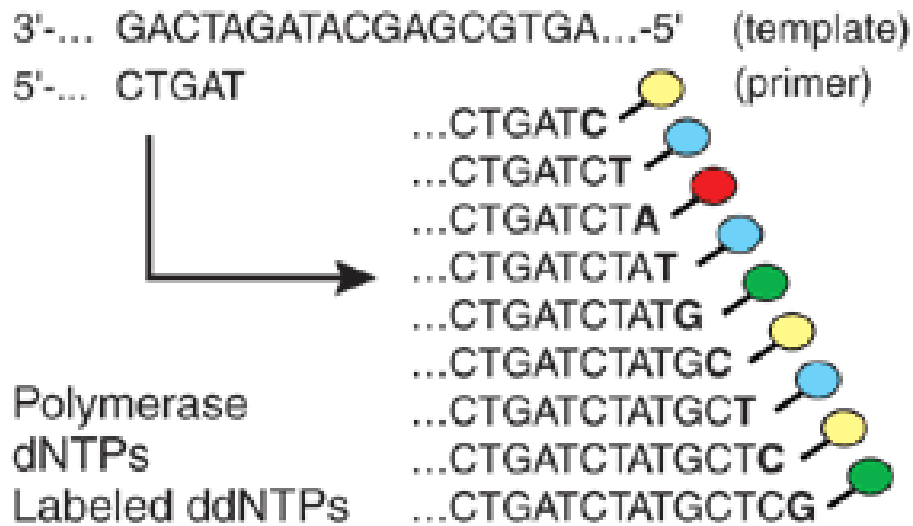
After laser excitation, capture the image of emitted fluorescence from each cluster on the flow cell. Record the identity of the first base for each cluster.



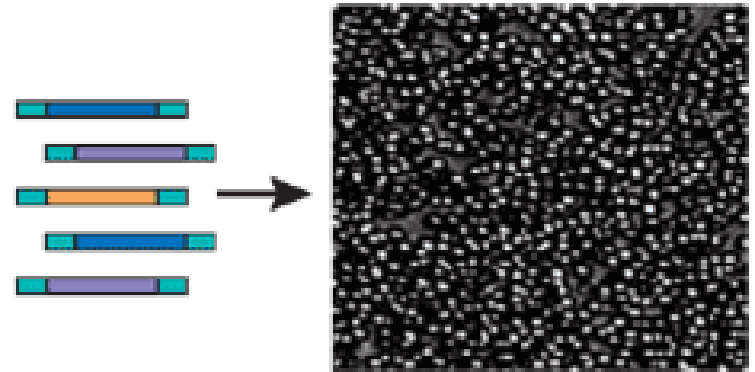
# Solexa High-Throughput Sequencing Protocol



## Cycle sequencing

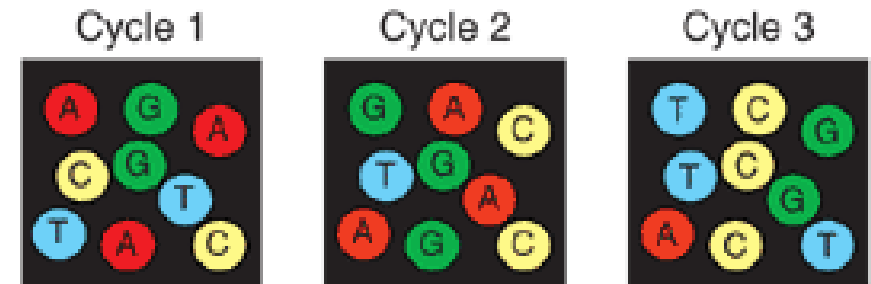


## Generation of polony array



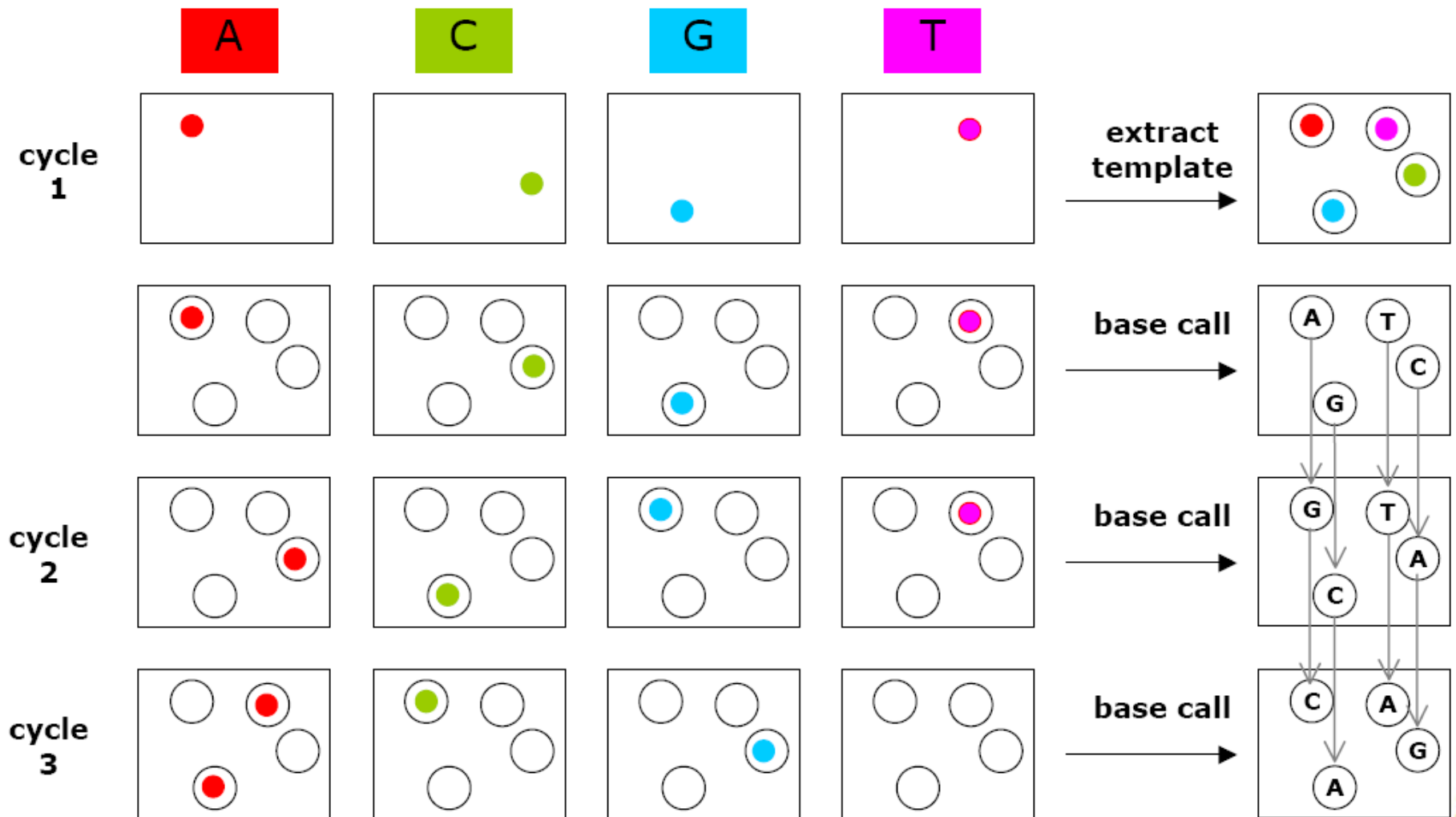
Electrophoresis  
 reading in each position

## Cyclic sequencing > 10<sup>16</sup> readings/polony



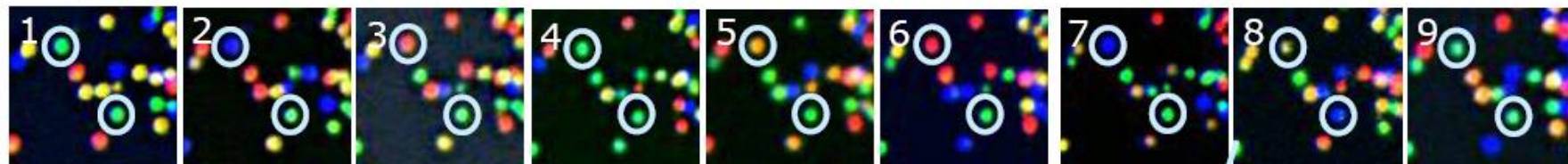
What is base 1? What is base 2? What is base 3?

# Sequence Determination From Four Colour Images



Sequences are AGC..., GCA..., TTA..., CAG...

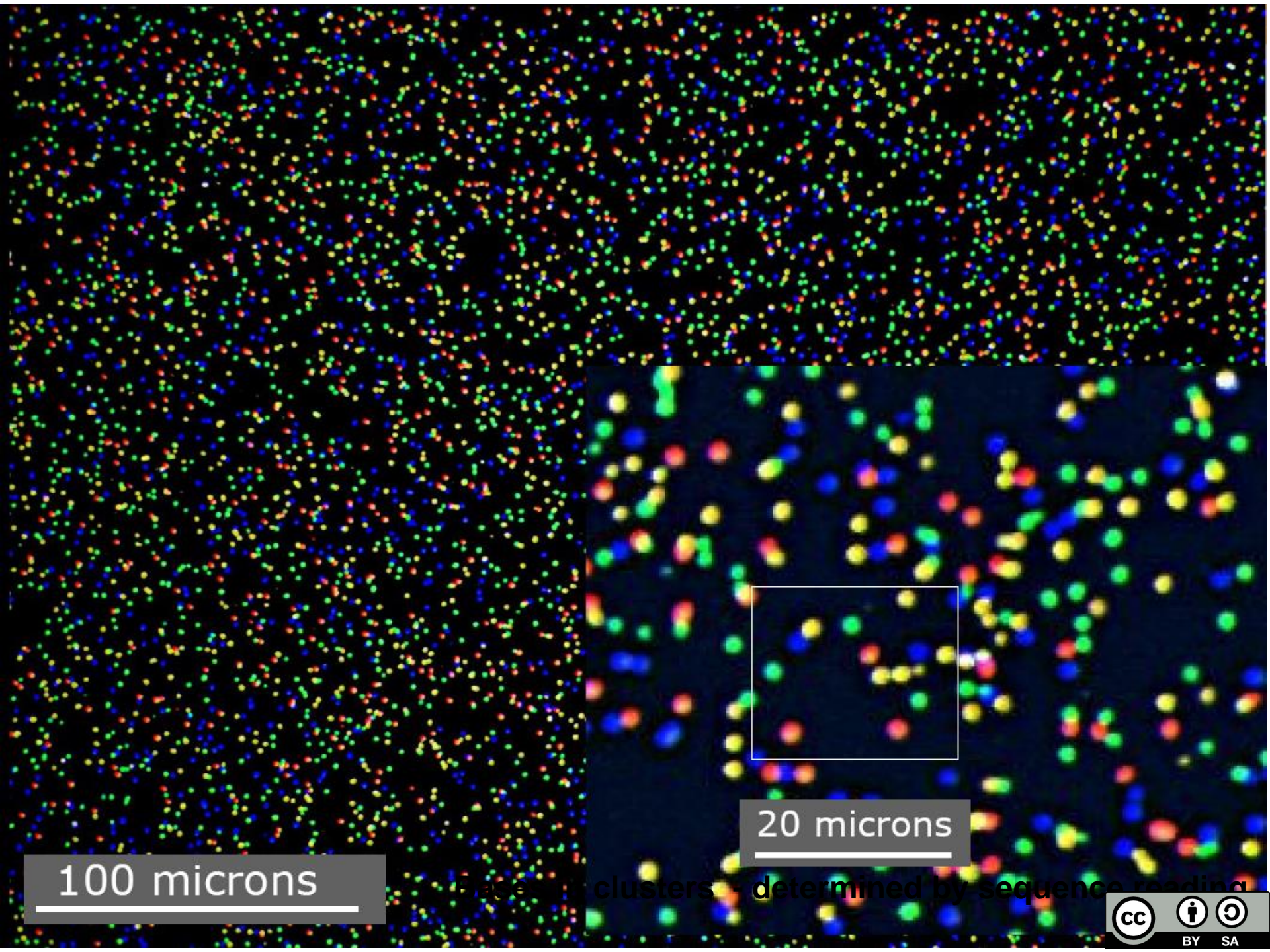
T G C T A C G A T ...



T T T T T T T G T ...

The identity of each base of a cluster is read off from sequential images





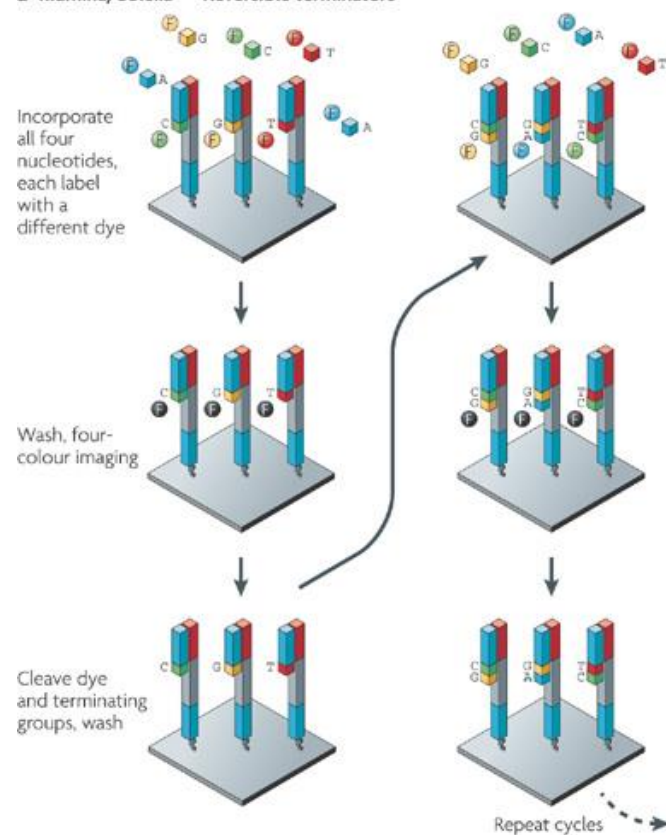
100 microns

20 microns

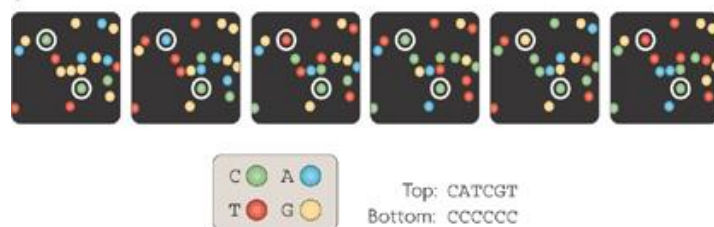
Face clusters - determined by sequence reading



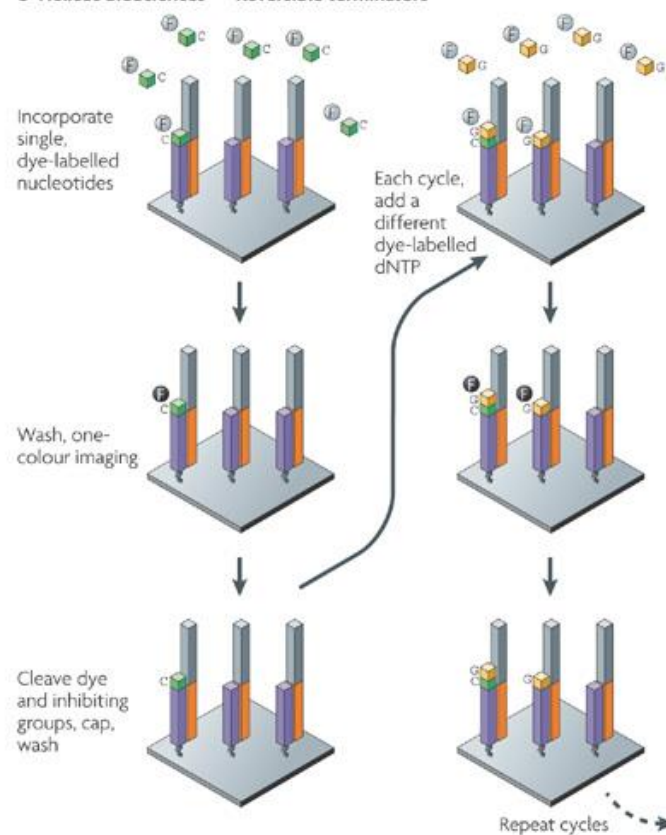
**a** Illumina/Solexa — Reversible terminators



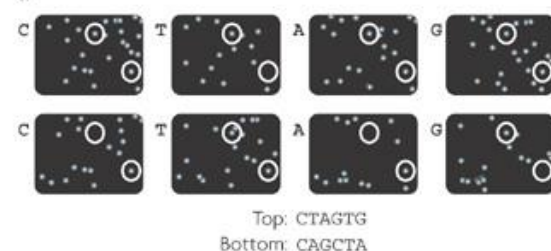
**b**



**c** Helicos BioSciences — Reversible terminators



**d**



Nature Reviews | Genetics

Four-colour cyclic reversible termination (CRT) **Illumina/Solexa's** 3'-O-azidomethyl reversible terminator chemistry; solid-phase-amplified template clusters. Following imaging, cleavage removes the fluorescent dyes and regenerates the 3'-OH group using the reducing agent tris(2-carboxyethyl)phosphine (TCEP)

**Helicos Virtual Terminators** - oligos labelled with the same dye are dispensed individually in a predetermined order. Following fluorescence imaging, a cleavage removes the dye and inhibitory groups using TCEP to permit the addition of the next Cy5-2'-deoxyribonucleoside triphosphate (dNTP). free sulphhydryl groups are then capped with iodoacetamide before the next nucleotide addition.



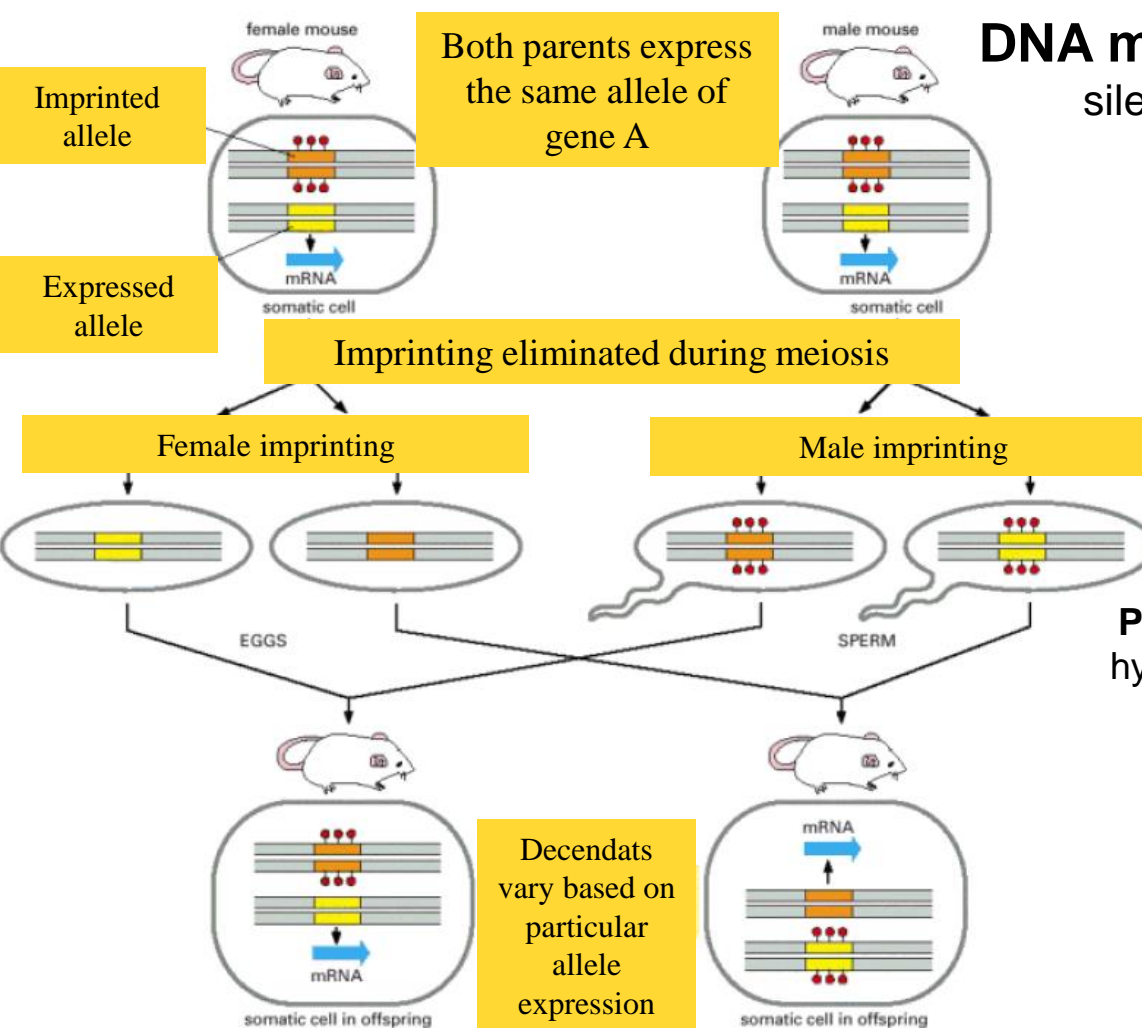
# Detection of methylated bases in genome

Methylation is epigenetic process

- determines monoallelic gene expression

# DNA methylation - genetic imprinting

silencing of one of homologous genes



**PWS:** obesity, low build, small hands and legs, hypotension, hypogonadism, mental retardation



## BWS

Large tongue, macrosomia – big children, defects of peritoneal wall, neonatal hypoglycemia

Mutation of gene imprinting may cause defect

Angelman syndrome

Prader-Willi syndrome

Beckwith-Wiedemann syndrome

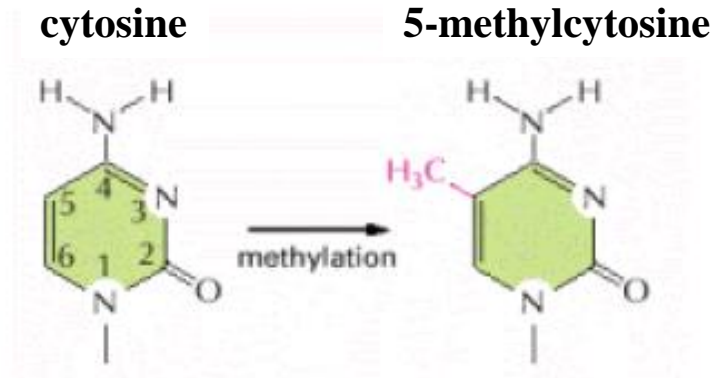


PWS

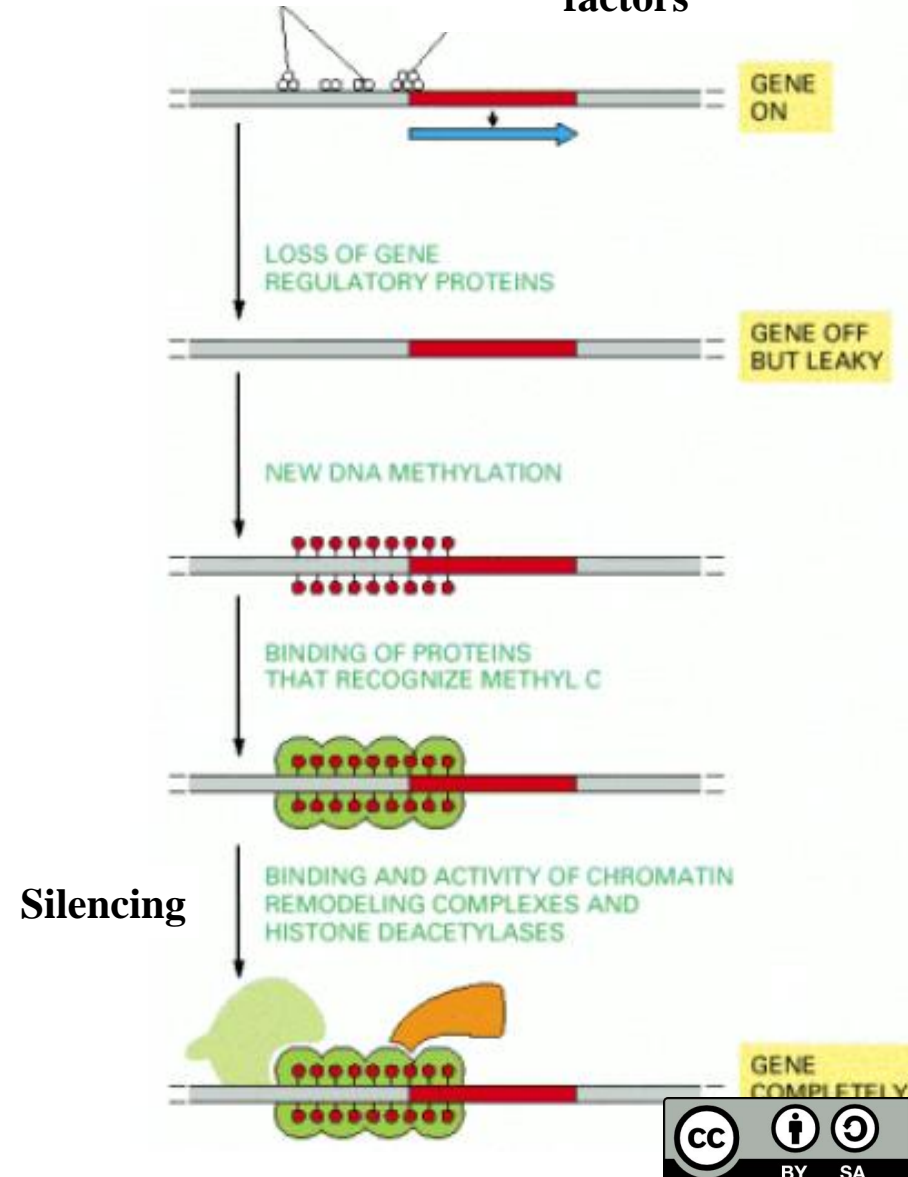
# How methylation turns off genes?

< 1% genes – imprinting (transcribed only one Copy)

Regulatory proteins      General transcription factors

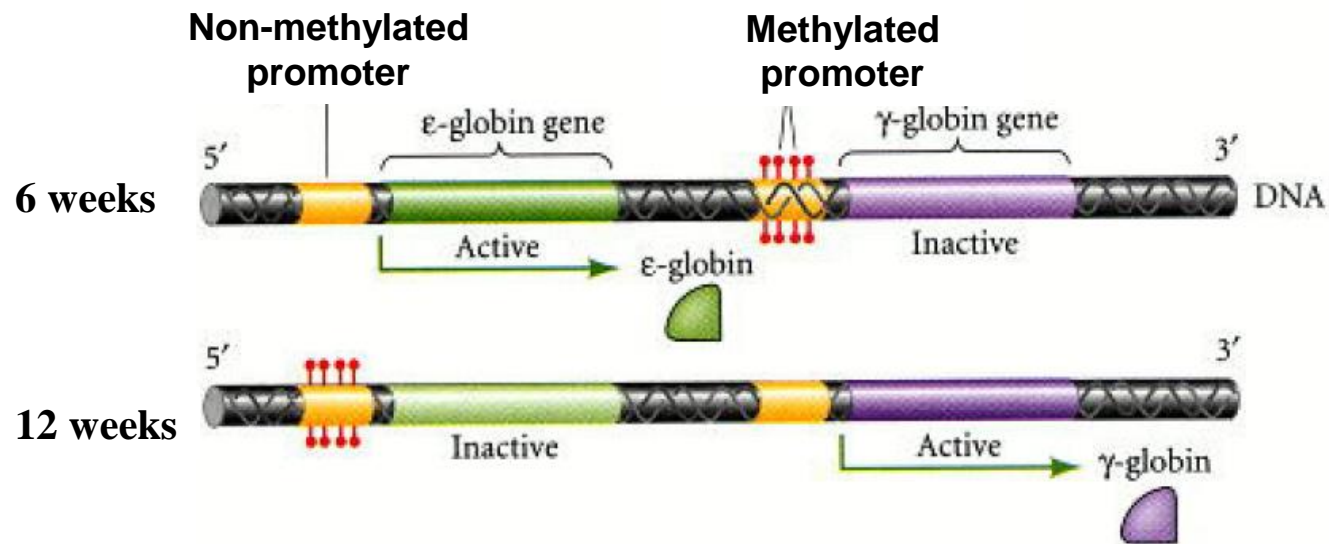
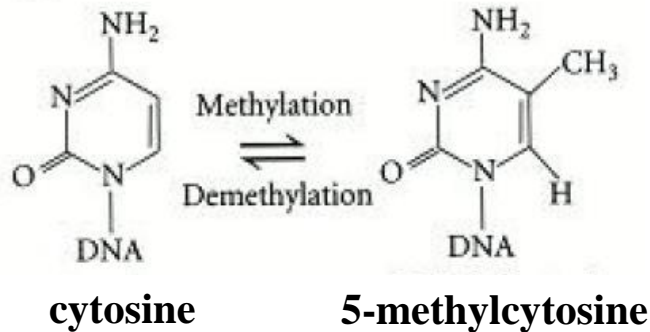


5-methylcytosine in CG nucleotides –  
no effect on DNA base-pairing



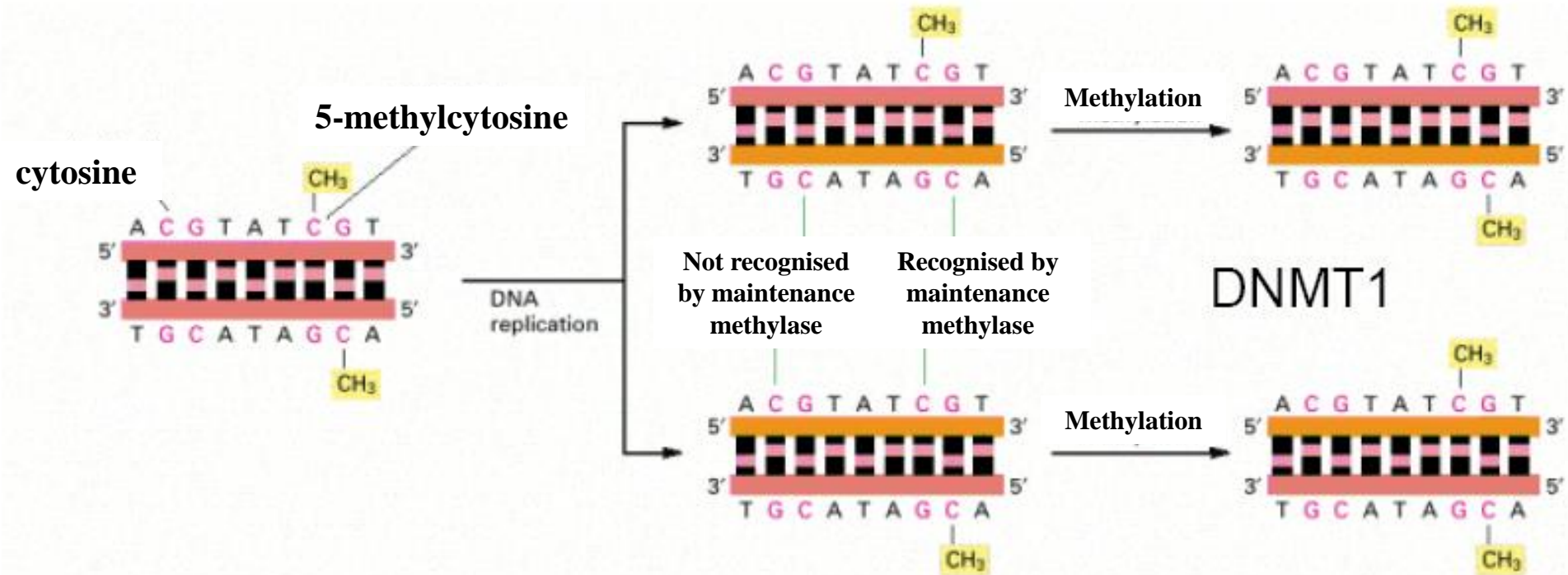
## Methylation of globin genes in human embryonic cells

Activity of globin genes negatively correlates with methylation of their promoters





# Methylation DNA pattern is **heritable** (methylation maintenance)



# **Aberant DNA methylation plays role cancerogenesis**

## **Gene specific aberrant DNA methylation**

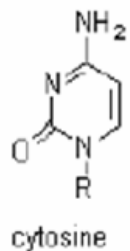
~30 000 CpG islands in human genome, CpG – low methylation in healthy tissue

- hypermethylation in v tumor suppressor genes or hypomethylation in oncogenes
- Hypermethylation of CpG islands

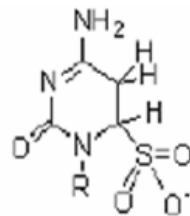
## **Hypomethylation of repetitive sequences**

(activation of retroviral and transposable elements => chromosome instability)

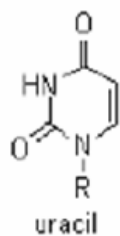




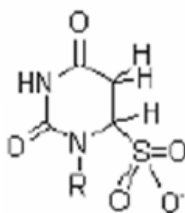
**Sulfonation**



**Hydrolytic  
deamination**



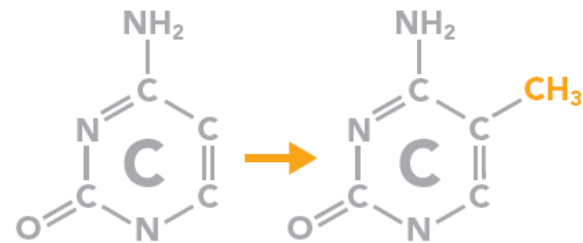
**Alcalic  
desulfonation**



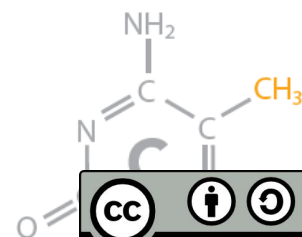
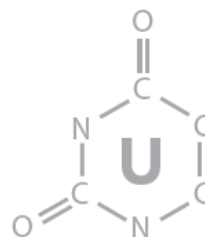
**Bisulfite sequencing**

**Non-methylated**

**Methylated**



**Bisulfite  
conversion**



**Non-methylated allele**

**Methylated allele**

— CG — CG —

— MeCG — MeCG —

**Bisulfite**

— UG — UG —

— MeCG — MeCG —

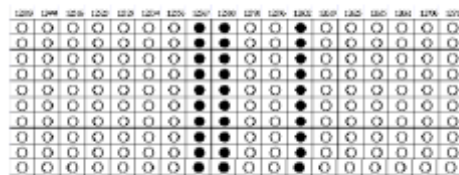
**PCR**

— **T**G — **T**G —

— **C**G — **C**G —

**Sequencing**

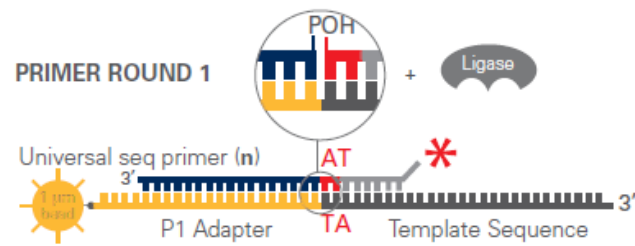
## Sequence data analysis



# **SOLiD (*Sequencing* by Oligonucleotide Ligation and Detection)**

Life Technologies – sold since 2008

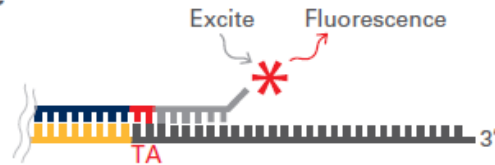
## 1. Primer annealing and ligation



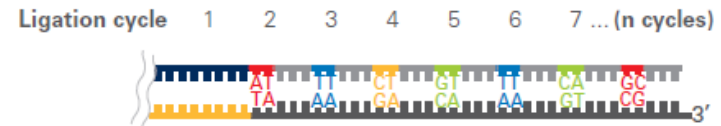
## 4. Cleavage of fluorophore



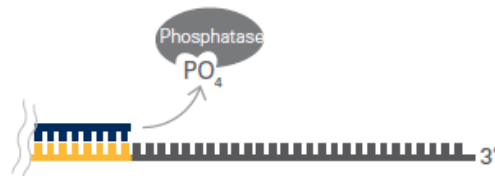
## 2. Image



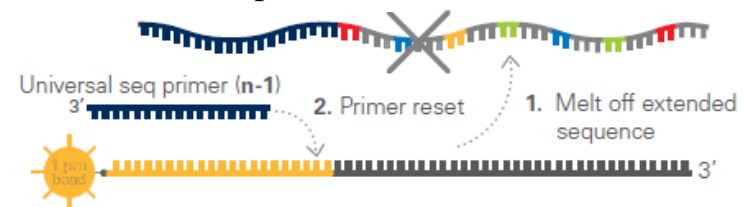
## 5. Repeat steps 1-4 to Extend Sequence



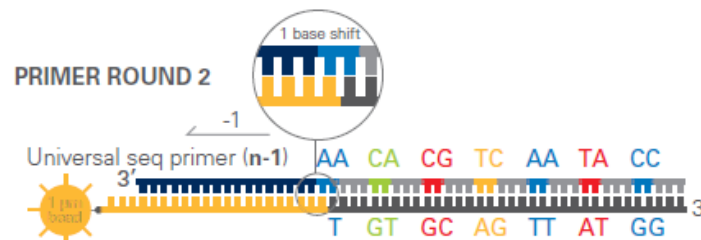
## 3. Blocking of unused primers



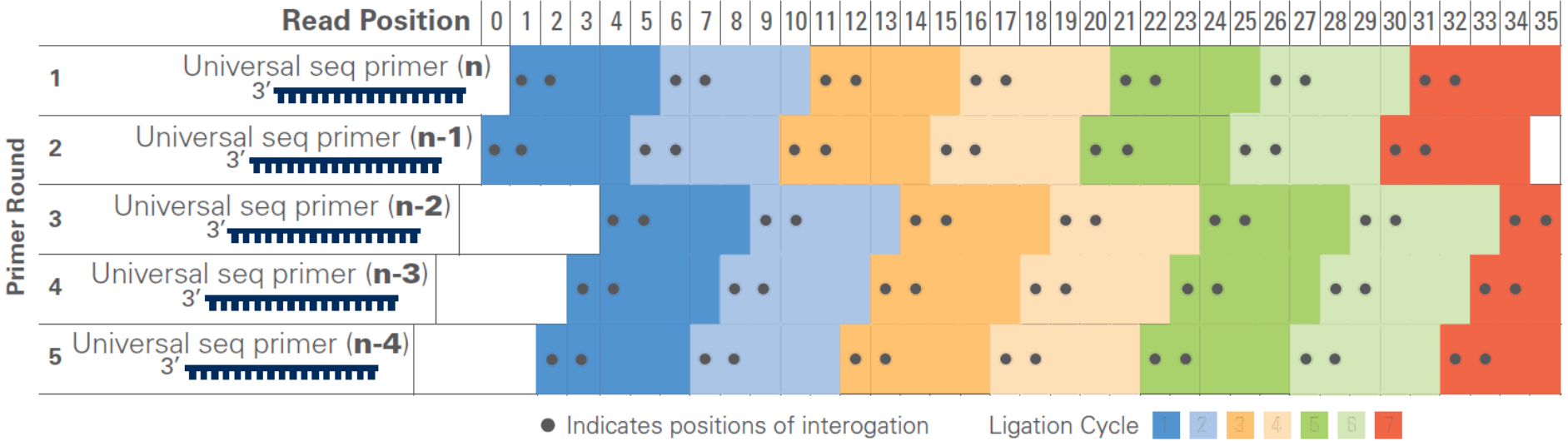
## 6. Next series: primer -1



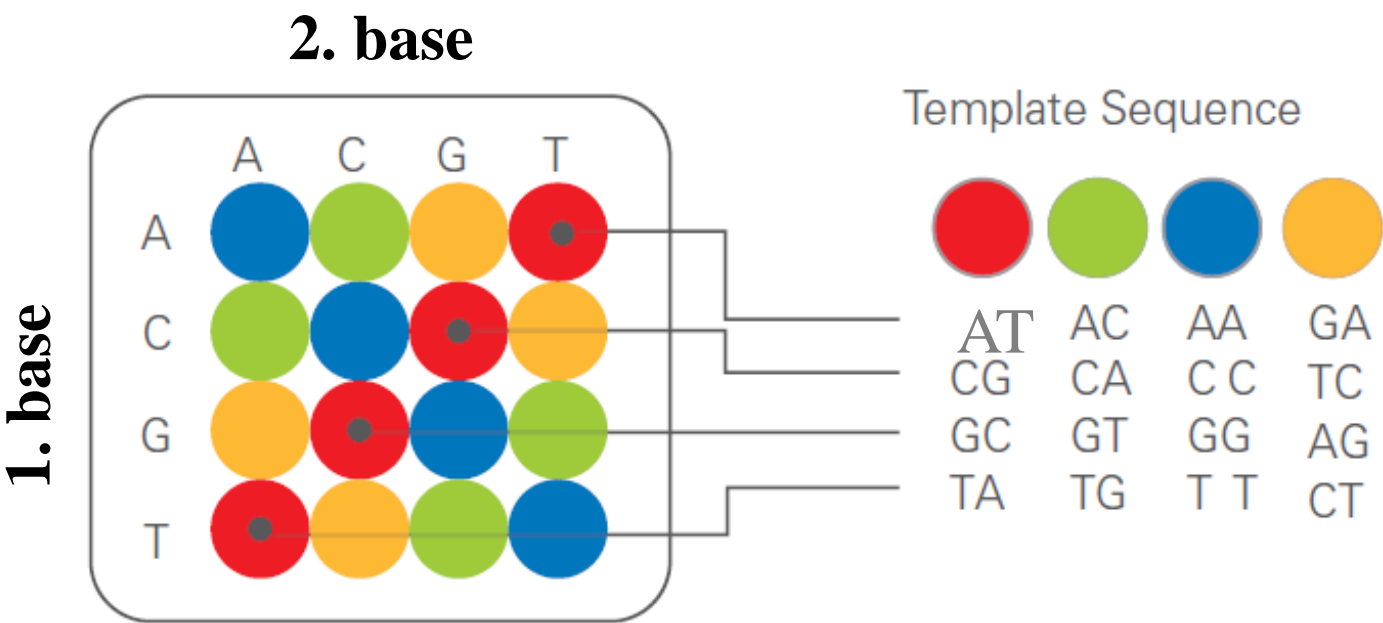
## 7. Repeat steps 1-5 with new primer



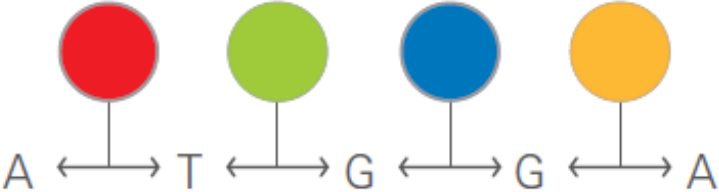
8. Repeat Reset with , n-2, n-3, n-4 primers



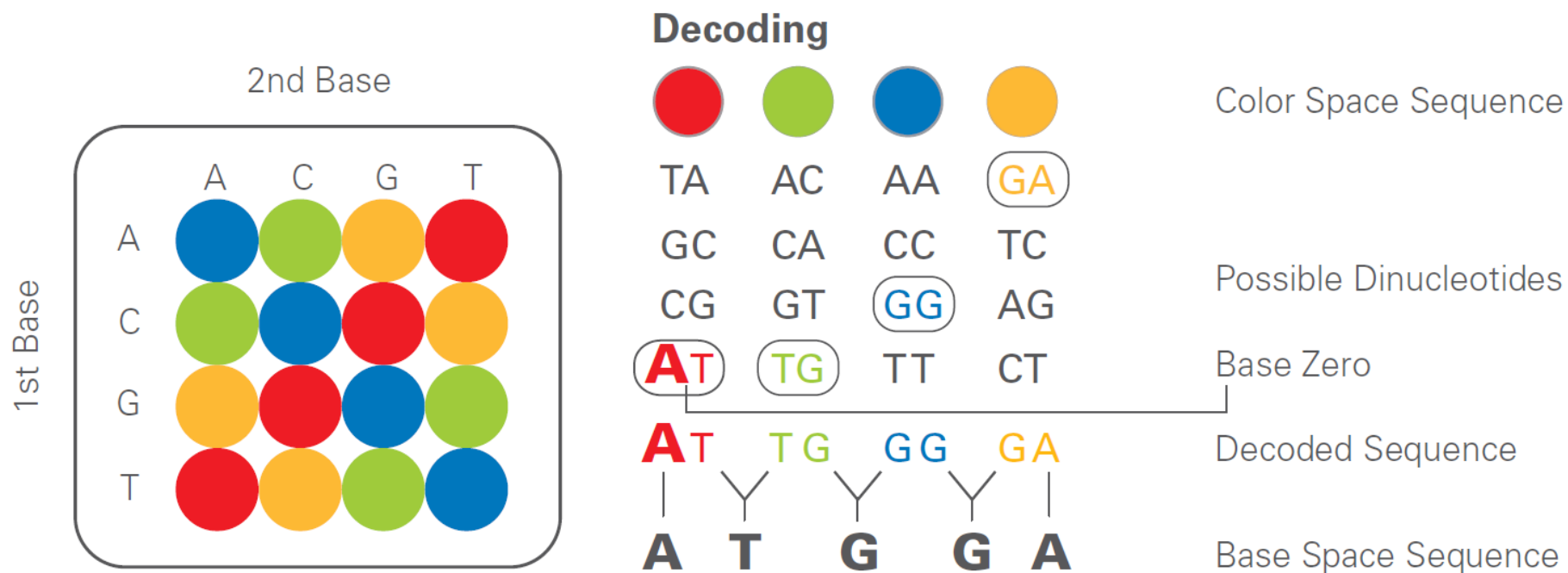
# Possible Dinucleotides Encoded By Each Color



With 2 base encoding each base is defined twice

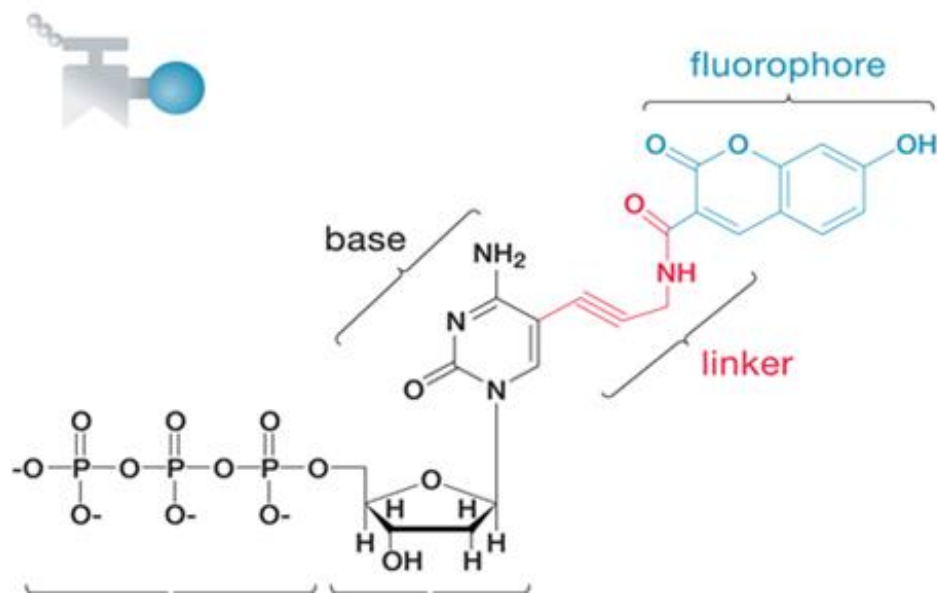




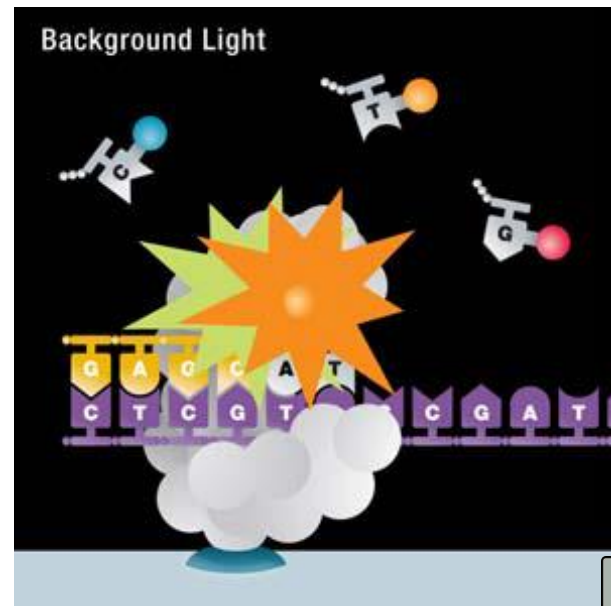
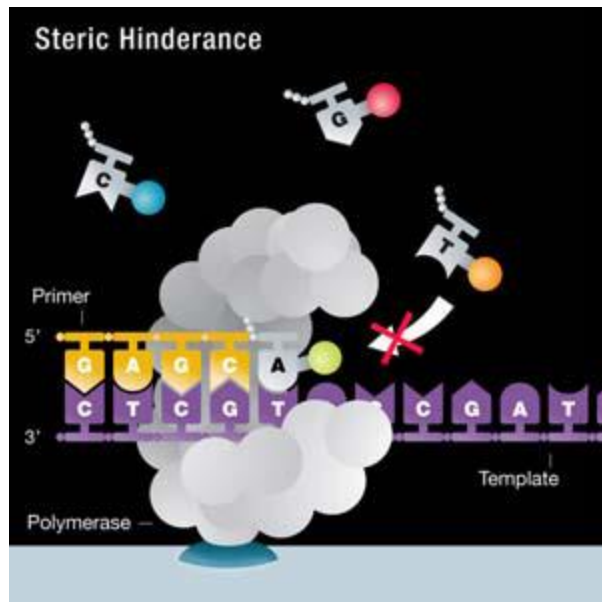


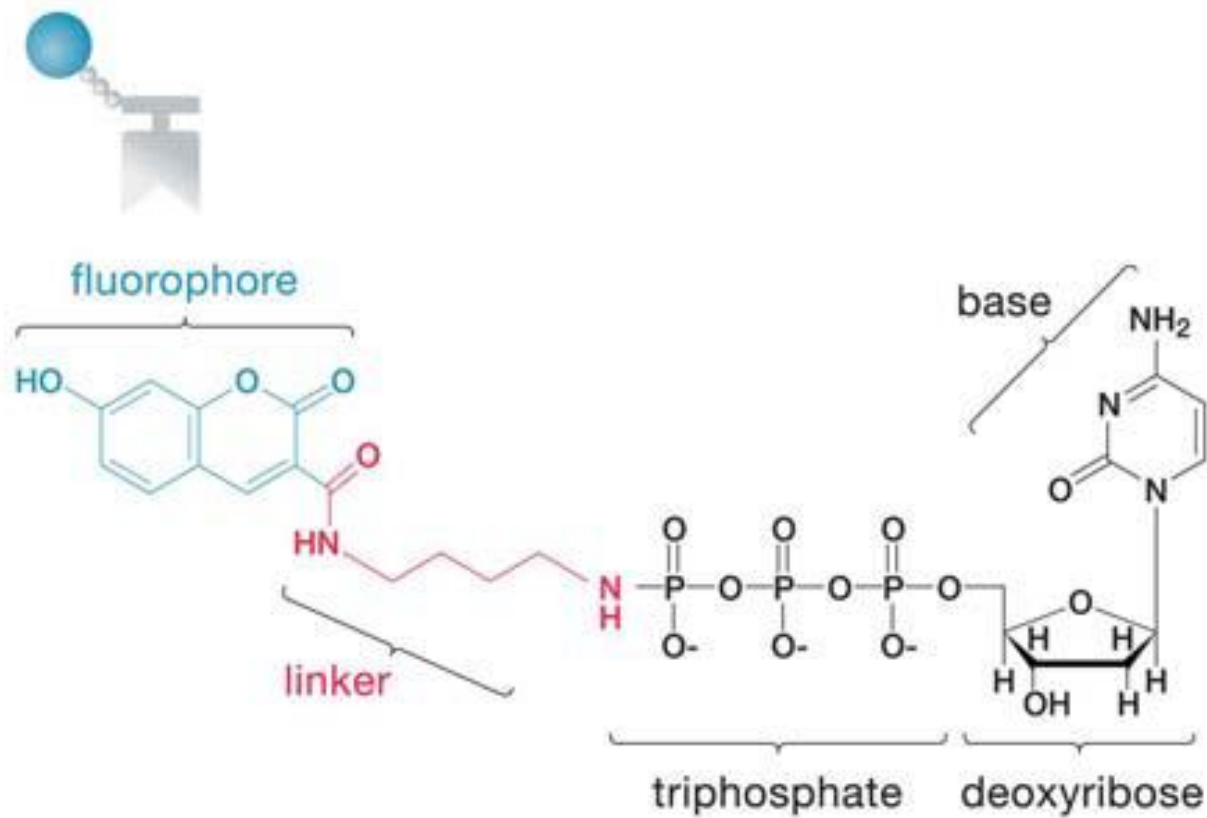
# **Single Molecule Real-Time Sequencing (SMRT™)**

**Pacific Biosciences**

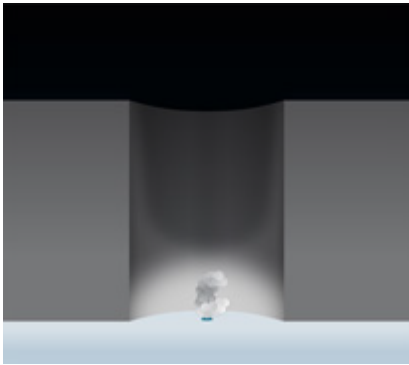


## Ordinary fluorescence nucleotide – base-labeled

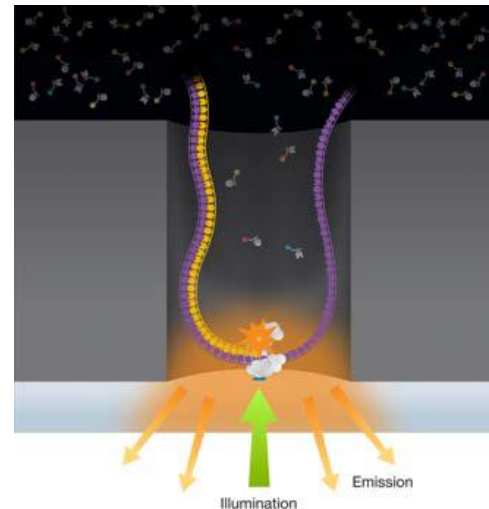
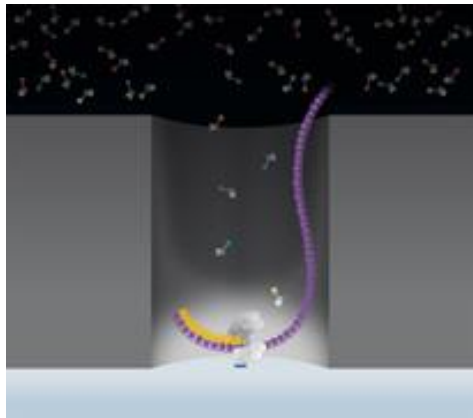




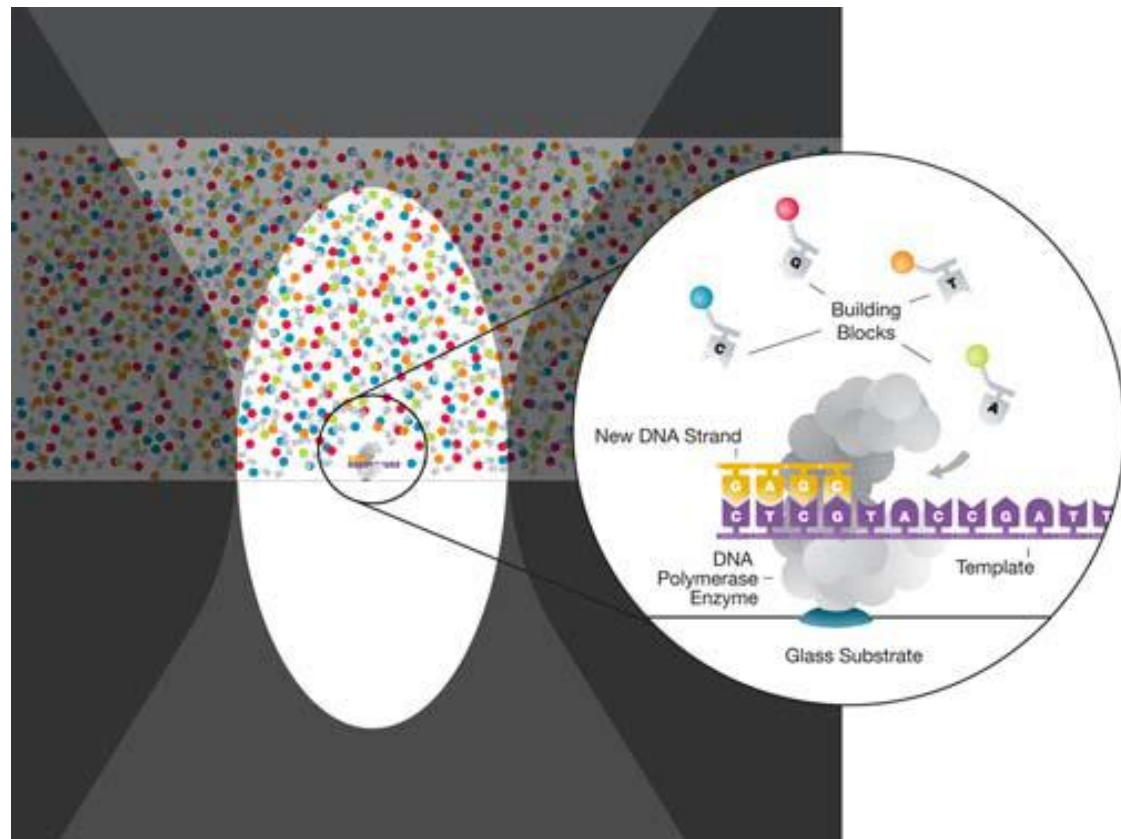
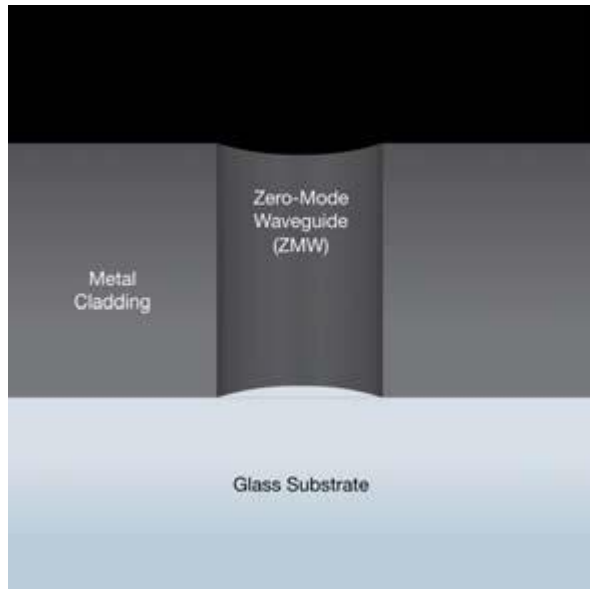
**SMRT utilizes fluorophore attached to phosphate**



One molecule of DNA polymerase attached to the bottom of each slot  
 During synthesis – fluorescence molecule retarded at the bottom –  
 signal detection



**Synthesis detected in thousands ZMWs simultaneously**  
**PacBio – various improvements to minimize background**

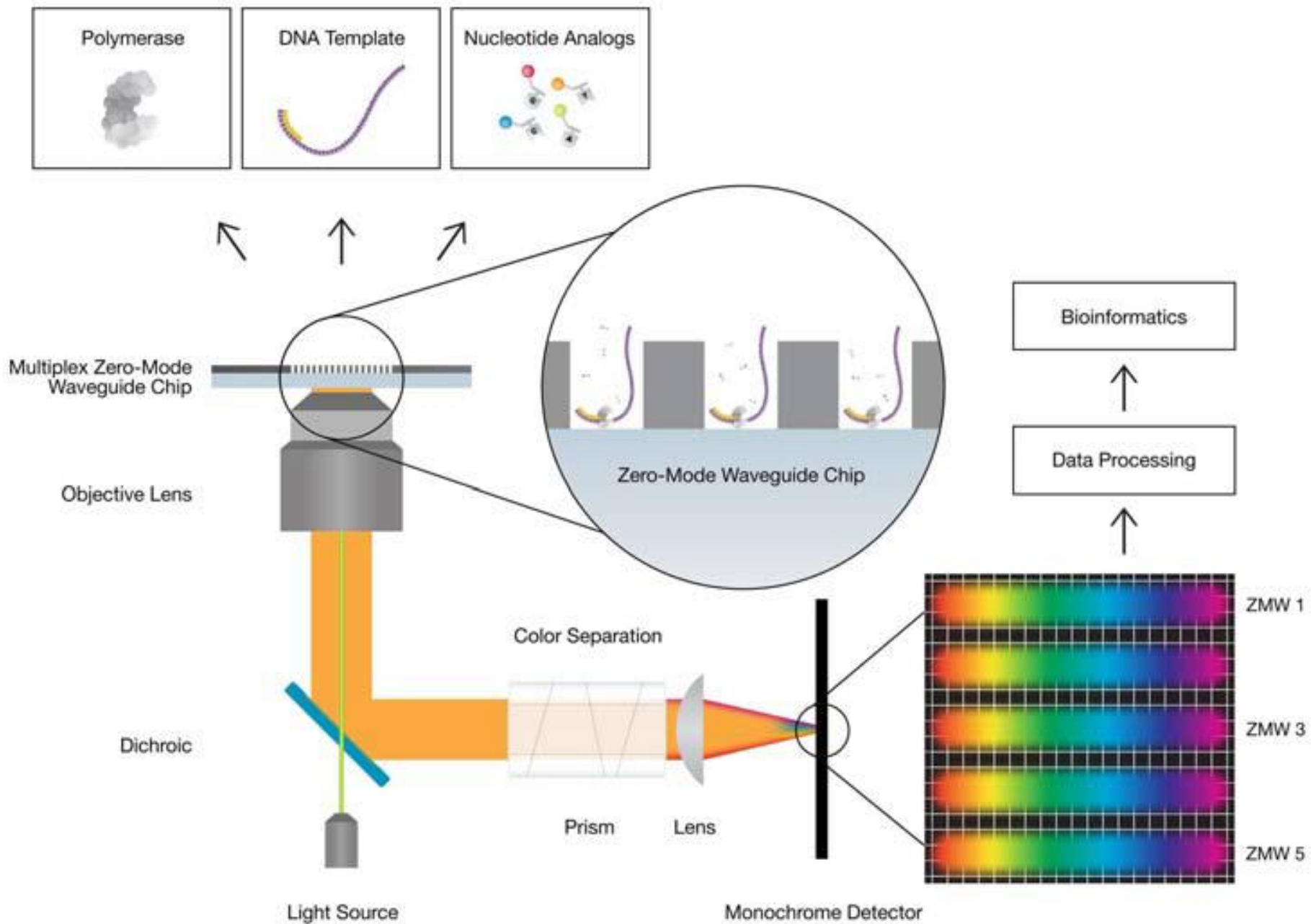


## Nanoscale - ZMW (zero-mode waveguide)

- cavity several nm in diameter
- perforated thin metallic film on transparent substrate
- Volume ~ 20 zeptoliters ( $10^{-21}$ l)

## Laser illumination of ZMW from the bottom

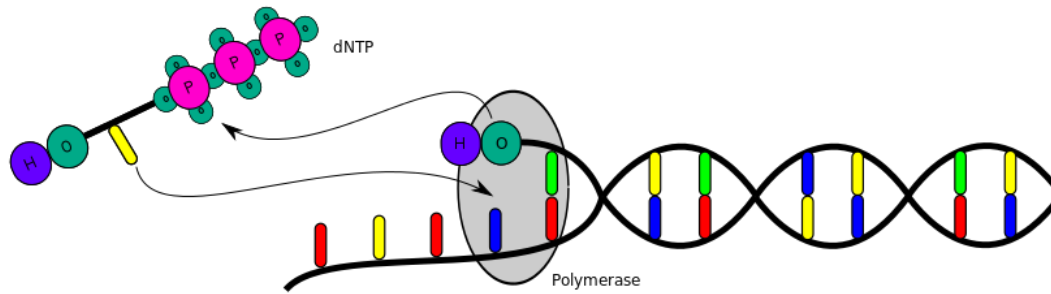
- Wavelength/diameter ZMW blocks passage through the slot
- only several nm over the bottom – detection of nucleotide added by polymerase



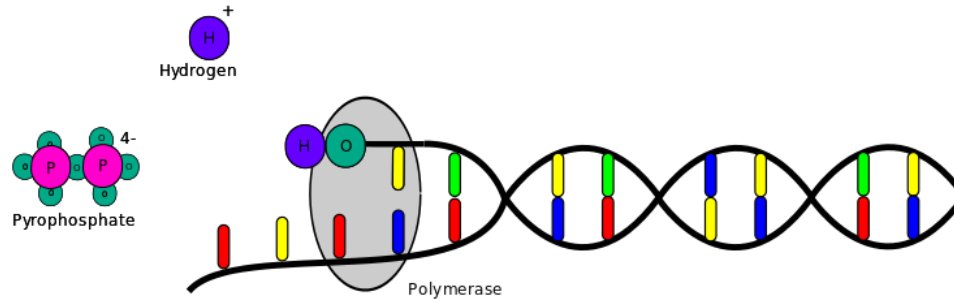


# Ion semiconductor sequencing

Detection of hydrogen ions that are released during the polymerization of DNA



Polymerase integrates a nucleotide.

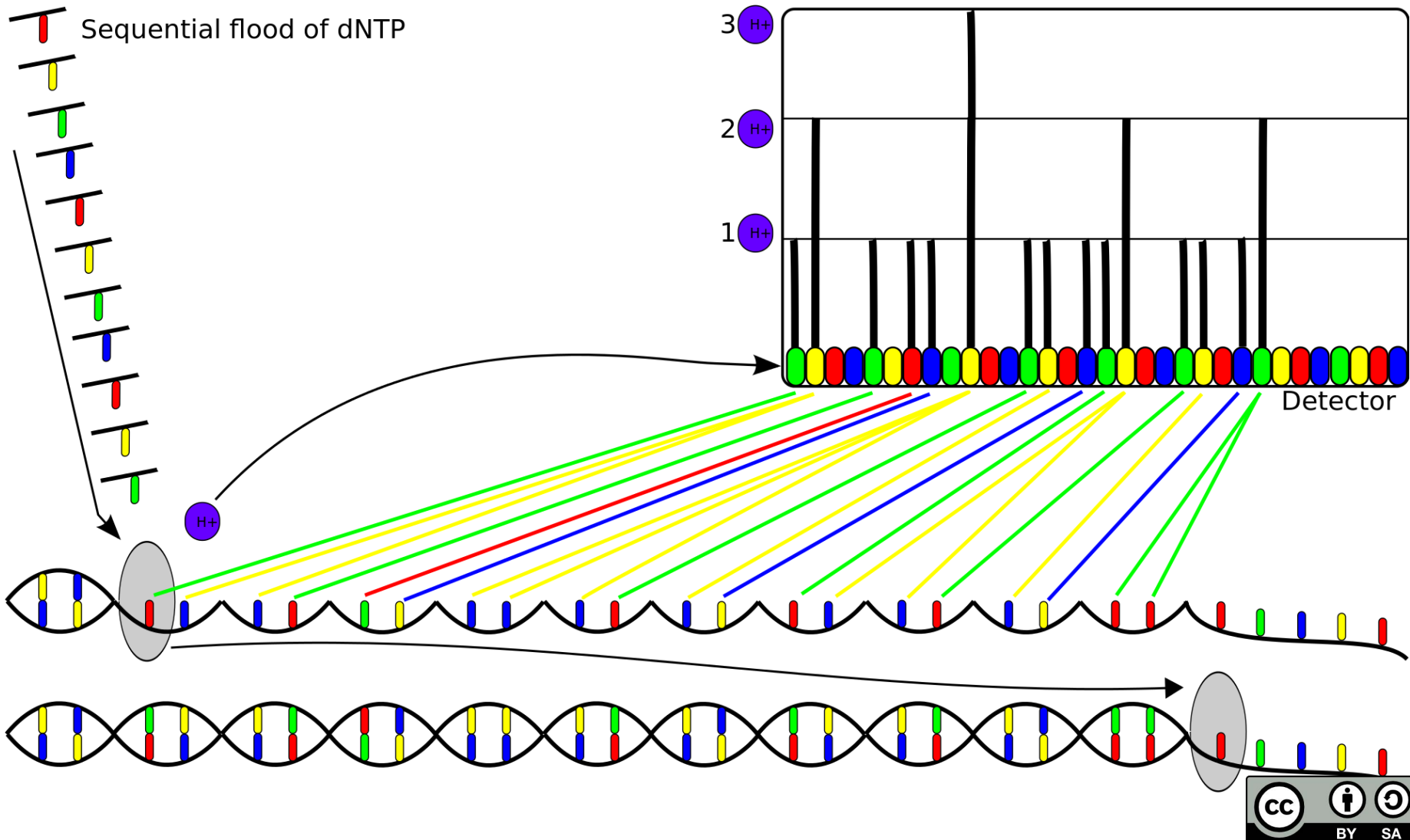


Hydrogen and pyrophosphate are released.

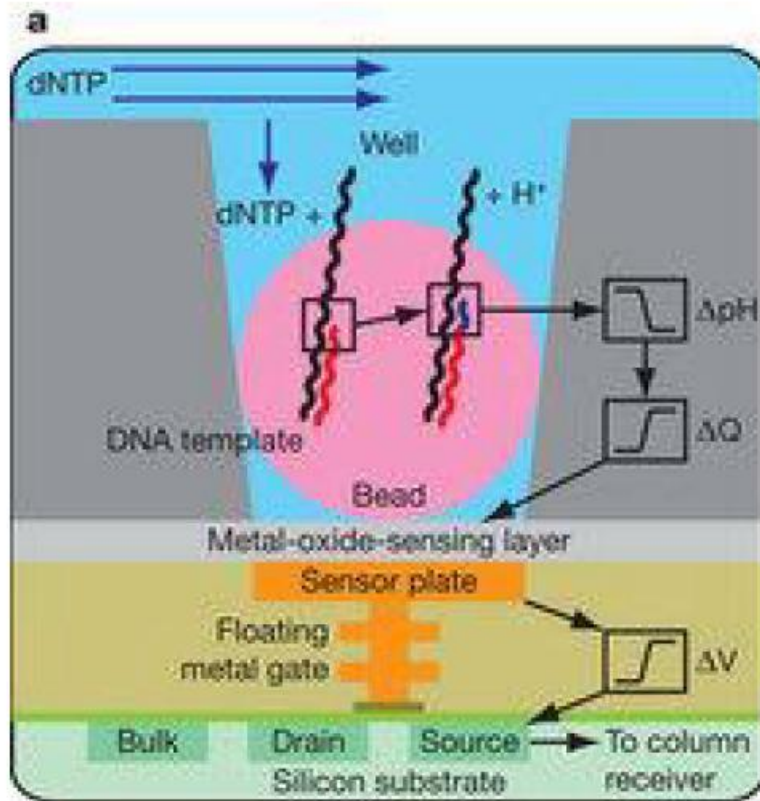


# Ion-sensitive field-effect transistor (ISFET)

- alternative electrode



# Ion Torrent sequencing technology



A simplified drawing of a well:  
a bead containing DNA template,  
and the underlying sensor and electronics.

1. Protons ( $H^+$ ) are released when nucleotides (dNTP) are incorporated on the growing DNA strands, changing the pH of the well ( $\Delta pH$ ).
2. This induces a change in surface potential of the metal-oxide-sensing layer, and a change in potential ( $\Delta V$ ) of the source terminal of the underlying field-effect transistor.



**Immobilization of biotin labeled complex on streptavidin coated microtiter plates**

**Wash**

**Release of probe or sample detection**



**Green color**

**Release of entrapped hybrids by antibody anti-digoxigenin with conjugated enzyme (Anti-DIG-peroxidase)  
After addition of chromogenic substrate.**



**Biotin labeled anchoring probe**

**Streptavidin**

ABTS - 2,2"-azin-bis(3-ethylbenzothiazolin-6-sulfonate)

**mRNA sequencing**

All reactions on 96-well plate

Oligo dT beads

- 1. Direct mRNA extraction**
- 2. Double-strand cDNA synthesis**
- 3. Enzymatic fragmentation of cDNA**
- 4. End repair and A addition**
- 5. Addition of adapters with 96 unique barcodes for multiplex samples**
- 6. Amplification by PCR**

**96 RNA-seq libraries**

"On beads" reaction

The diagram illustrates the process of mRNA sequencing. It begins with a green leaf representing the source of mRNA. The mRNA is extracted (Step 1) and converted into double-strand cDNA (Step 2). This cDNA is then enzymatically fragmented (Step 3). The fragments undergo end repair and A addition (Step 4), where 'A' stands for adenine. Adapters with 96 unique barcodes are added (Step 5) for multiplexing. Finally, the libraries are amplified by PCR (Step 6) to produce 96 RNA-seq libraries. All reactions are performed on a 96-well plate. Oligo dT beads are used to facilitate the initial steps.

**mRNA sequencing**

All reactions on 96-well plate

Oligo dT beads

**1. Direct mRNA extraction**

**2. Double-strand cDNA synthesis**

**3. Enzymatic fragmentation of cDNA**

**4. End repair and A addition**

**5. Addition of adapters with 96 unique barcodes for multiplex samples**

**6. Amplification by PCR**

**96 RNA-seq libraries**

"On beads" reaction

The diagram illustrates the mRNA sequencing workflow. It begins with a green leaf representing the source of mRNA. The process involves several steps: 1. Direct mRNA extraction, showing three wavy lines representing mRNA molecules with poly-A tails (AAAAA). 2. Double-strand cDNA synthesis, where the mRNA is converted into double-stranded cDNA (ds-cDNA) represented by four horizontal lines. 3. Enzymatic fragmentation of cDNA, resulting in smaller ds-cDNA fragments shown as short horizontal segments. 4. End repair and A addition, where the fragmented cDNA is prepared for ligation, shown as segments with single 'A' bases at their ends. 5. Addition of adapters with 96 unique barcodes for multiplex samples, where colored Y-shaped adapter sequences are added to the cDNA fragments. 6. Amplification by PCR, leading to the final stage: 96 RNA-seq libraries. A callout box labeled "On beads" reaction shows the initial step where oligo dT beads (represented by brown circles with TTTT sequences) bind to the poly-A tails of the mRNA. This entire process is noted to occur on a 96-well plate.

[illegible][illegible][illegible]

**mRNA sequencing**

All reactions on 96-well plate

Oligo dT beads

- 1. Direct mRNA extraction**
- 2. Double-strand cDNA synthesis**
- 3. Enzymatic fragmentation of cDNA**
- 4. End repair and A addition**
- 5. Addition of adapters with 96 unique barcodes for multiplex samples**
- 6. Amplification by PCR**

**96 RNA-seq libraries**

"On beads" reaction

The diagram illustrates the process of mRNA sequencing. It begins with a green leaf representing the source of mRNA. The mRNA is extracted (Step 1) and converted into double-strand cDNA (Step 2). This cDNA is then enzymatically fragmented (Step 3). The fragments undergo end repair and A addition (Step 4), where 'A' stands for adenine. Adapters with 96 unique barcodes are added (Step 5) for multiplexing. Finally, the libraries are amplified by PCR (Step 6) to produce 96 RNA-seq libraries. All reactions are performed on a 96-well plate. Oligo dT beads are used to facilitate the initial steps.

**mRNA sequencing**

All reactions on 96-well plate

Oligo dT beads

- 1. Direct mRNA extraction**
- 2. Double-strand cDNA synthesis**
- 3. Enzymatic fragmentation of cDNA**
- 4. End repair and A addition**
- 5. Addition of adapters with 96 unique barcodes for multiplex samples**
- 6. Amplification by PCR**

**96 RNA-seq libraries**

"On beads" reaction

The diagram illustrates the process of mRNA sequencing. It begins with a green leaf representing the source of mRNA. The mRNA is extracted (Step 1) and converted into double-strand cDNA (Step 2). This cDNA is then enzymatically fragmented (Step 3). The fragments undergo end repair and A addition (Step 4), where 'A' stands for adenine. Adapters with 96 unique barcodes are added (Step 5) for multiplexing. Finally, the libraries are amplified by PCR (Step 6) to produce 96 RNA-seq libraries. All reactions are performed on a 96-well plate. Oligo dT beads are used to facilitate the initial steps.

[illegible]

**mRNA sequencing**

All reactions on 96-well plate

Oligo dT beads

- 1. Direct mRNA extraction**
- 2. Double-strand cDNA synthesis**
- 3. Enzymatic fragmentation of cDNA**
- 4. End repair and A addition**
- 5. Addition of adapters with 96 unique barcodes for multiplex samples**
- 6. Amplification by PCR**

**96 RNA-seq libraries**

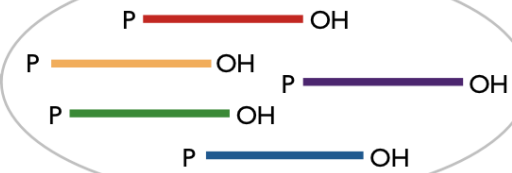
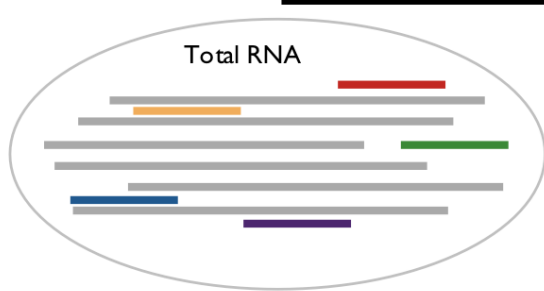
"On beads" reaction

The diagram illustrates the process of mRNA sequencing. It begins with a green leaf representing the source of mRNA. The mRNA is extracted (Step 1) and converted into double-strand cDNA (Step 2). This cDNA is then enzymatically fragmented (Step 3). The fragments undergo end repair and A addition (Step 4), where 'A' stands for adenine. Adapters with 96 unique barcodes are added (Step 5) for multiplexing. Finally, the libraries are amplified by PCR (Step 6) to produce 96 RNA-seq libraries. All reactions are performed on a 96-well plate. Oligo dT beads are used to facilitate the initial steps.

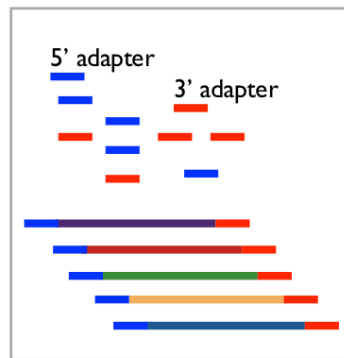
[illegible]

# miRNA sequencing

Total RNA isolation

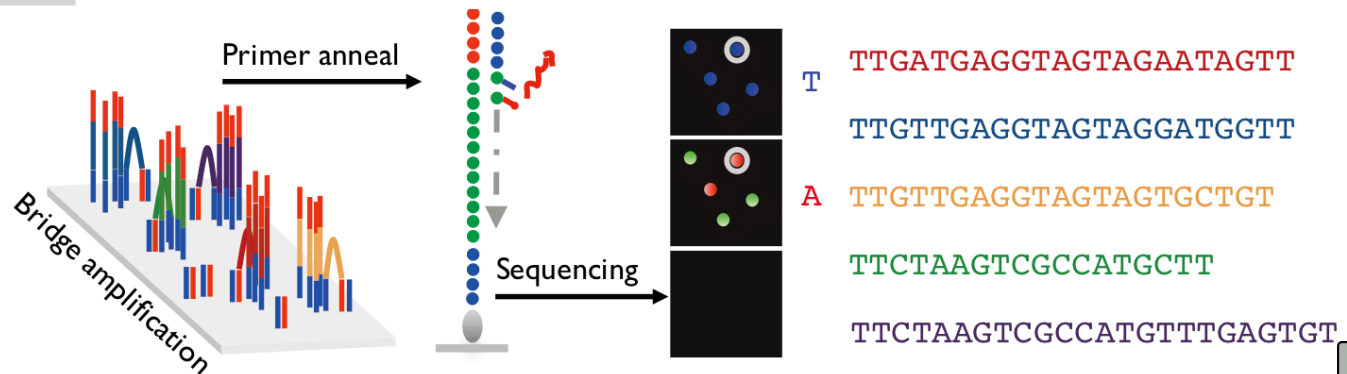


Adaptor ligation

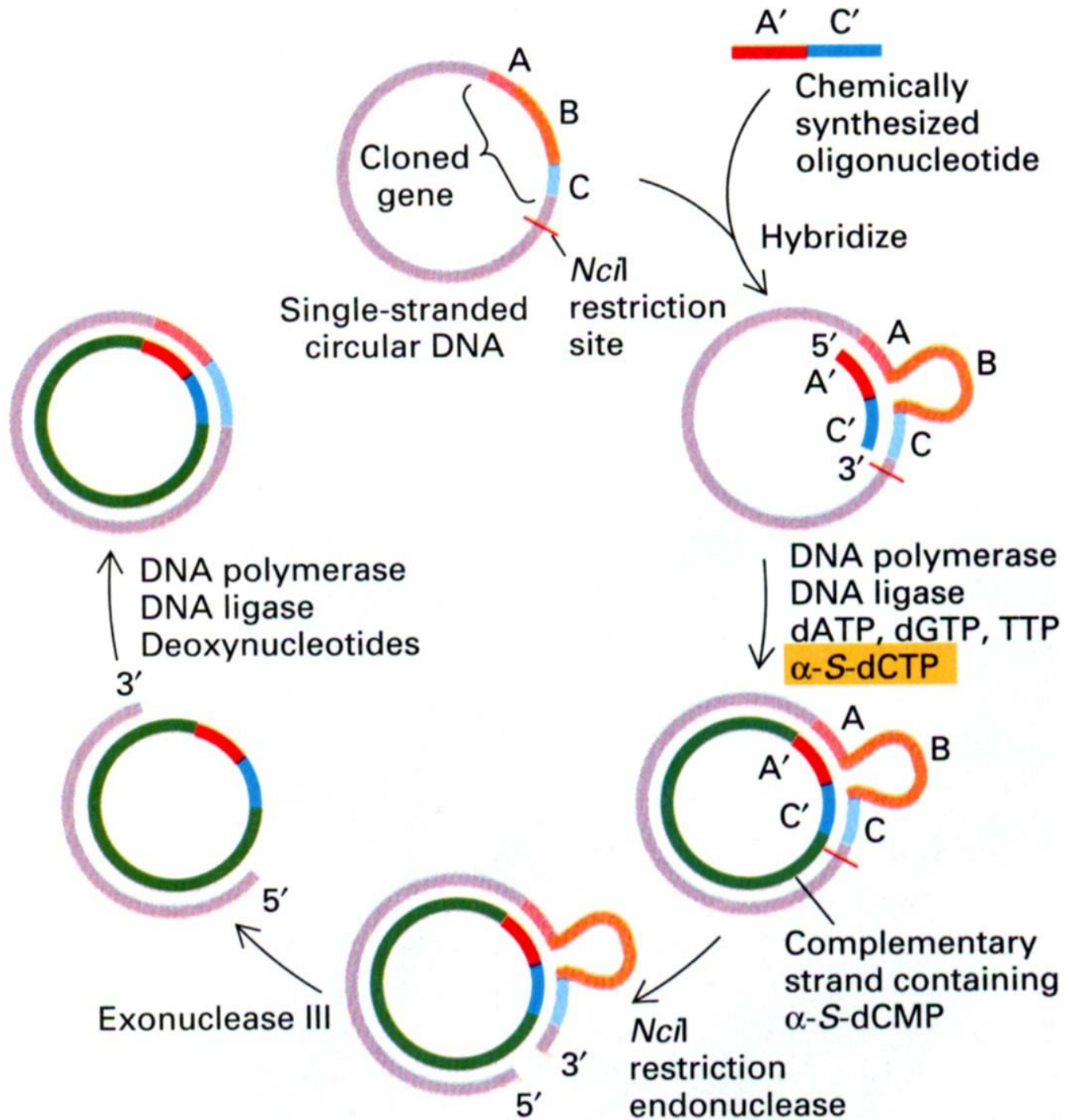


RT & PCR

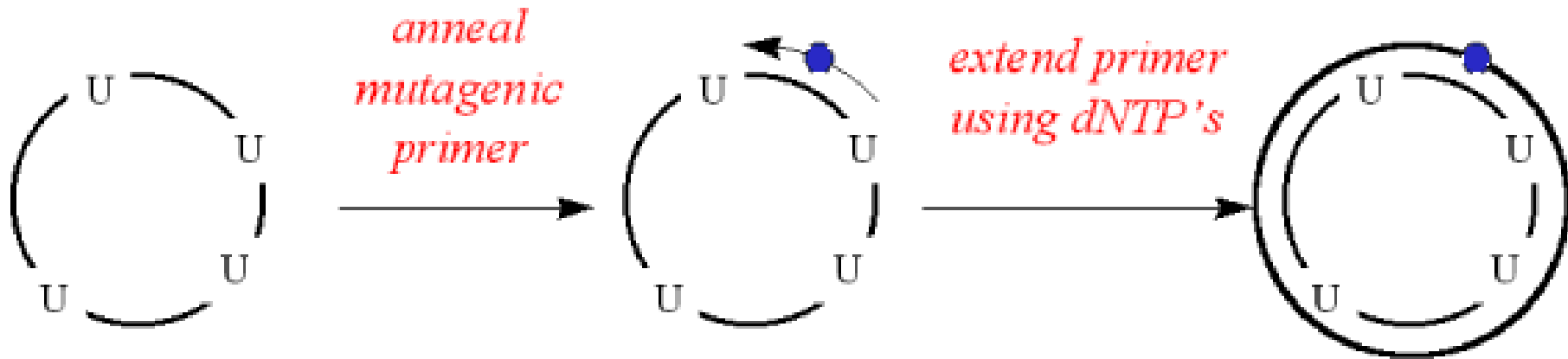
Sequencing











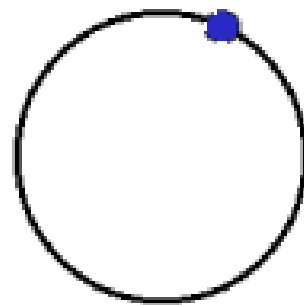
uracil containing  
M13 (Kunkel) Template  
(grown in *dut*<sup>-</sup>, *ung*<sup>-</sup> host)

*insert into*  
(*ung*<sup>+</sup>) *E. coli* host

**ung** – uracil N-glycosylase

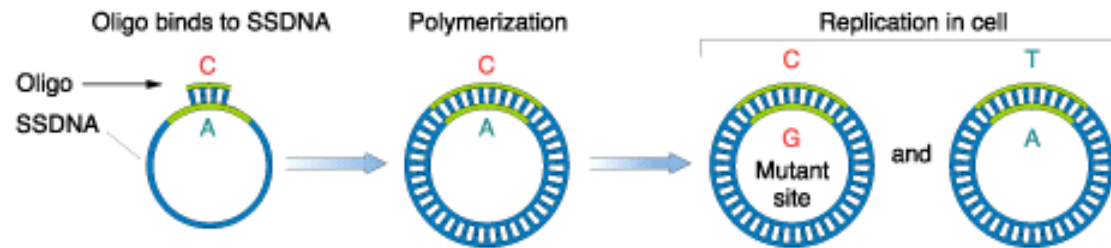
*parent strand*  
*preferentially*  
*degraded*

phage progeny  
with mutation

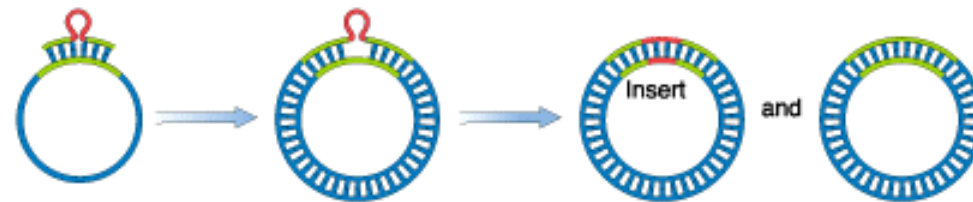


(a) Oligonucleotide-directed mutagenesis

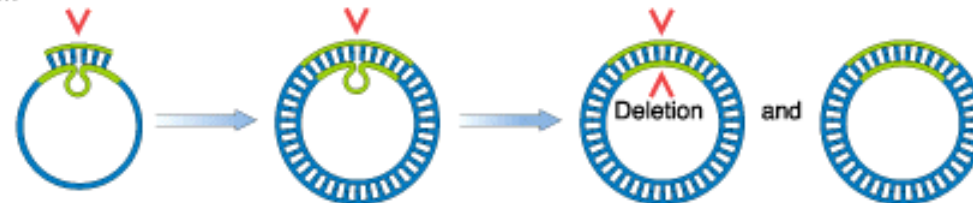
(i) Base-pair substitution



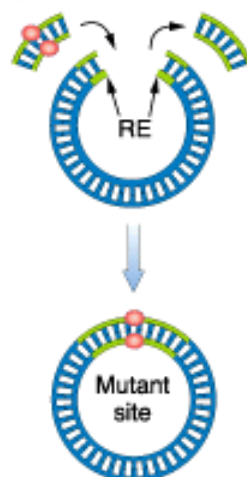
(ii) Insertion



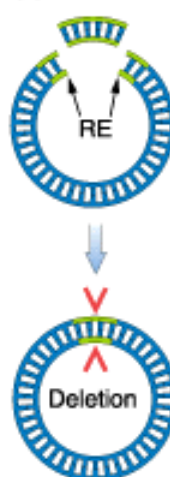
(iii) Deletion



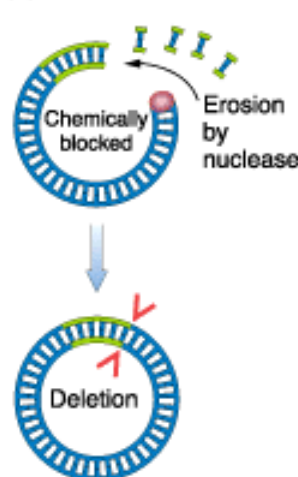
(b) Cassette replacement



(c) Deletion

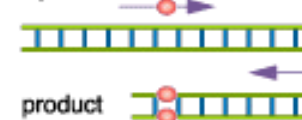


(d) Sets of deletions

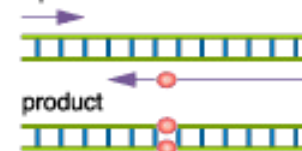


(e) PCR mutagenesis

- 1st PCR to obtain long primer



- 2nd PCR using long primer



Uveřejněné materiály jsou určeny studentům Vysoké školy chemicko-technologické v Praze

jako studijní materiál. Některá textová i obrazová data v nich obsažená jsou převzata

z veřejných zdrojů. V případě nedostatečných citací nebylo cílem autora/ů záměrně poškodit event. autora/y původního díla.

S eventuálními výhradami se prosím obračejte na autora/y konkrétního výukového materiálu, aby

bylo možné zjednat nápravu.

The published materials are intended for students of the University of Chemistry and Technology, Prague as a study material. Some text and image data contained therein are taken from public sources. In the case of insufficient quotations, the author's intention was not to intentionally infringe the possible author(s) rights to the original work.

If you have any reservations, please contact the author(s) of the specific teaching material in order to remedy the situation.

# In situ proximity ligation

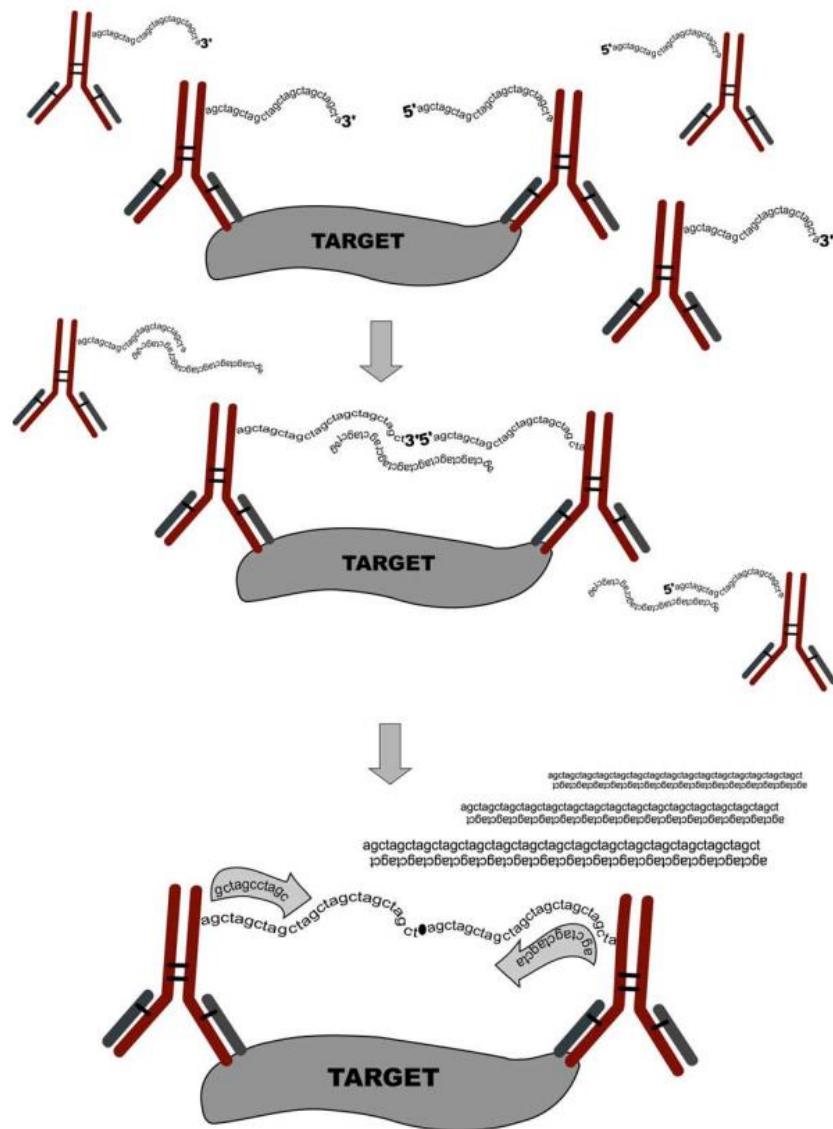


EUROPEAN UNION  
European Structural and Investing Funds  
Operational Programme Research,  
Development and Education



MINISTRY OF EDUCATION,  
YOUTH AND SPORTS

## Proximity ligation assay



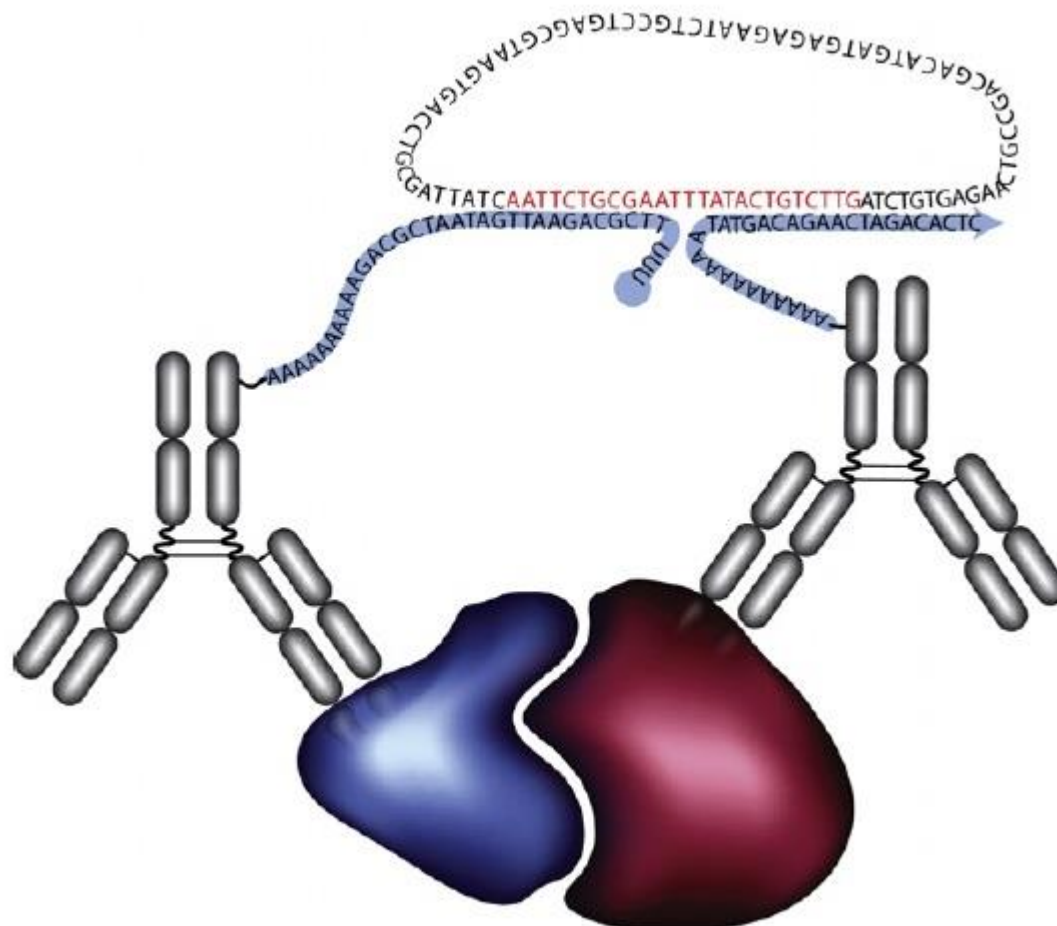
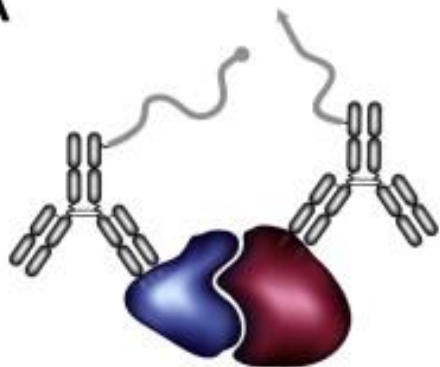
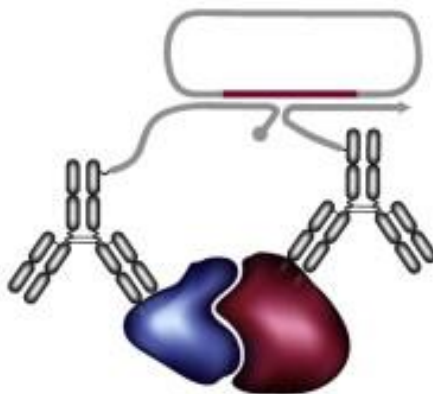


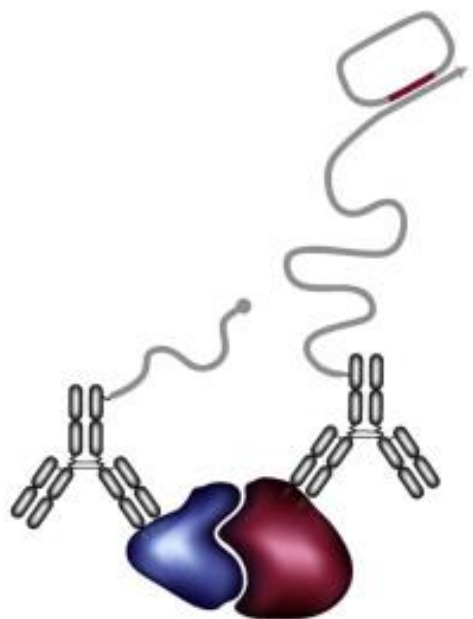
Fig. 1. A presentation of the oligonucleotide design used for *in situ* PLA

**A**

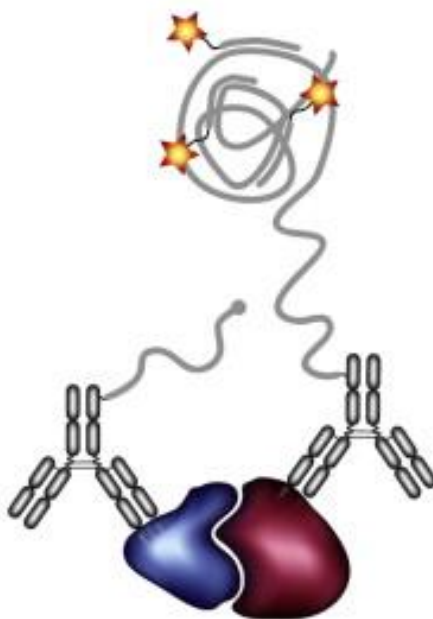
Proximity probe binding



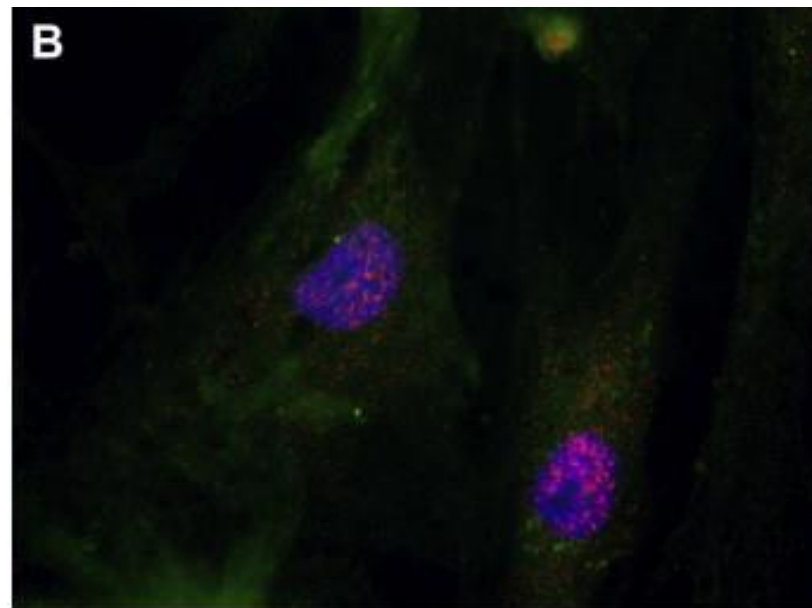
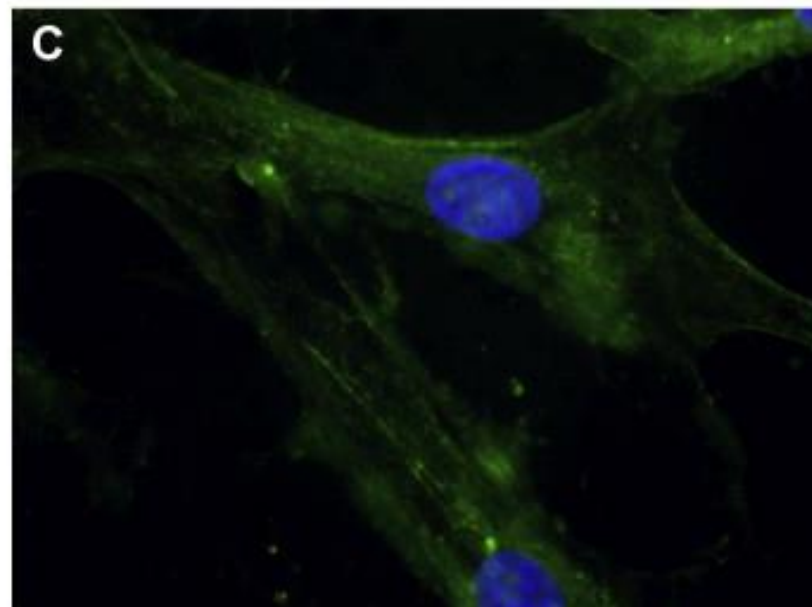
Circularization and ligation of connector oligonucleotides



Rolling circle amplification

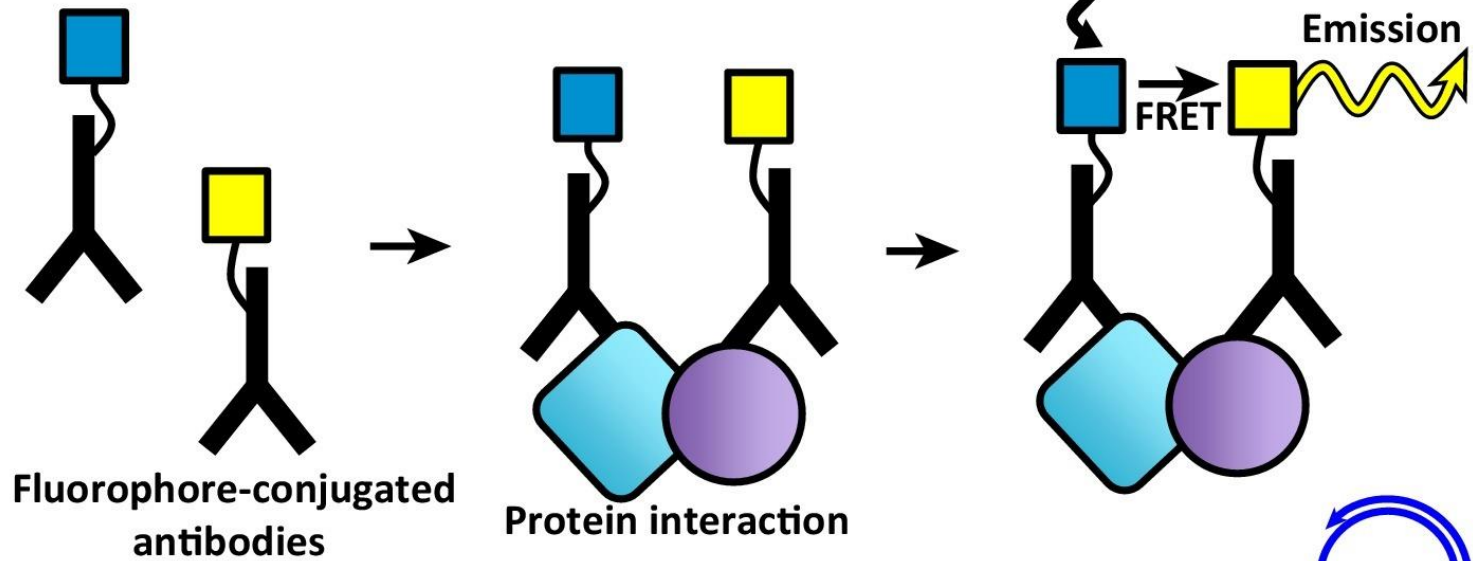


Detection of rolling circle product

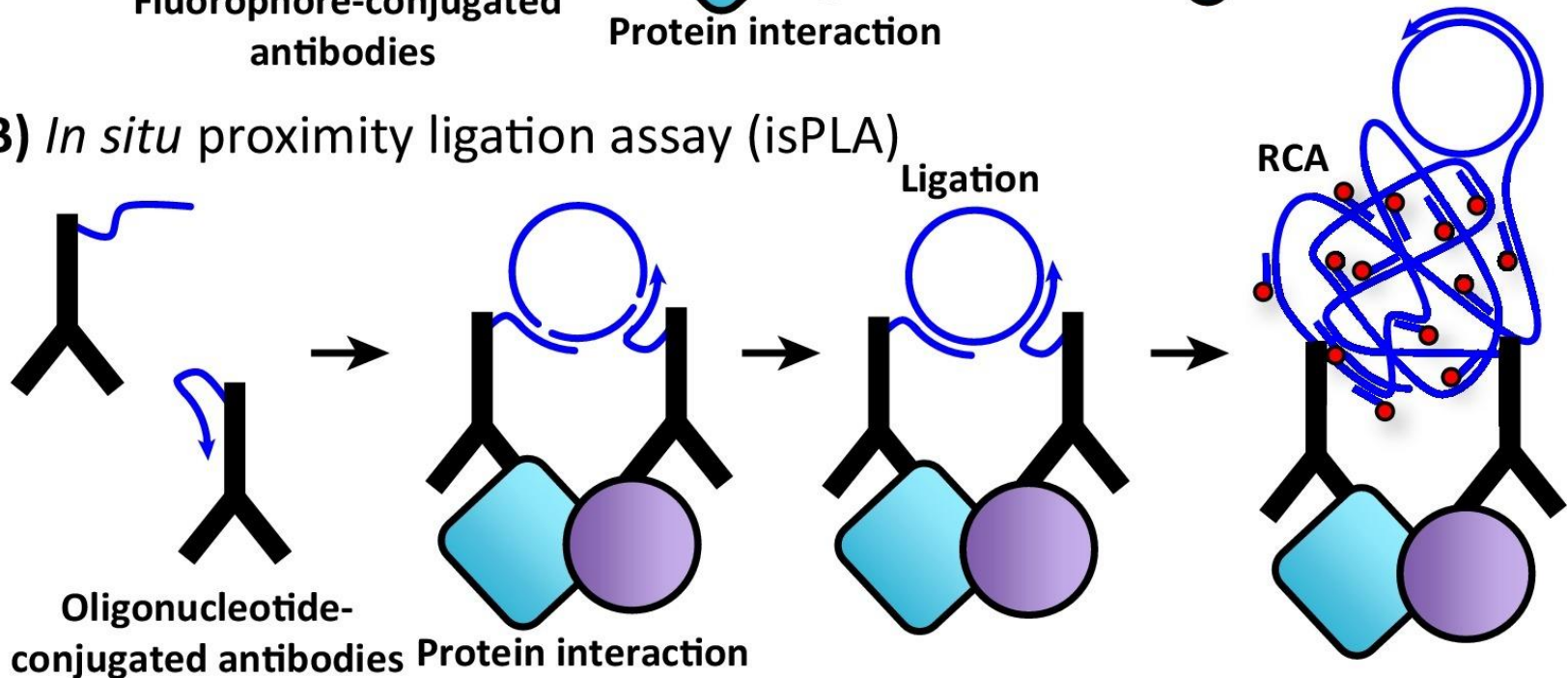
**B****C**

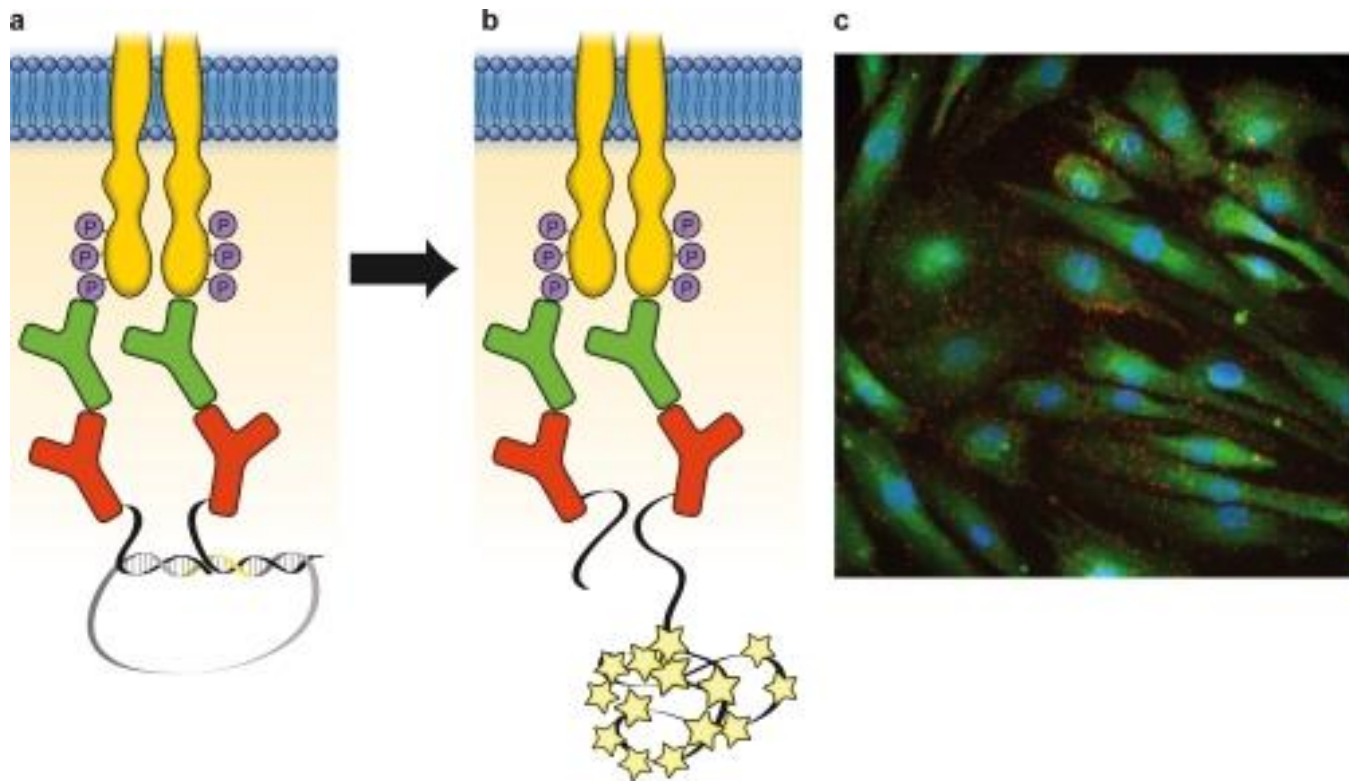


**(A) Förster resonance energy transfer (FRET)**



**(B) *In situ* proximity ligation assay (isPLA)**





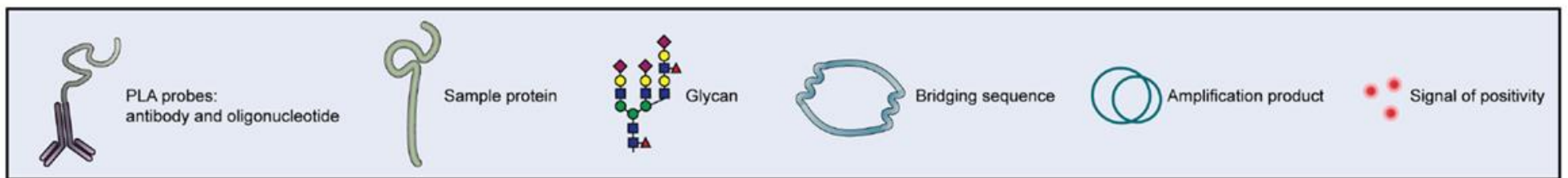
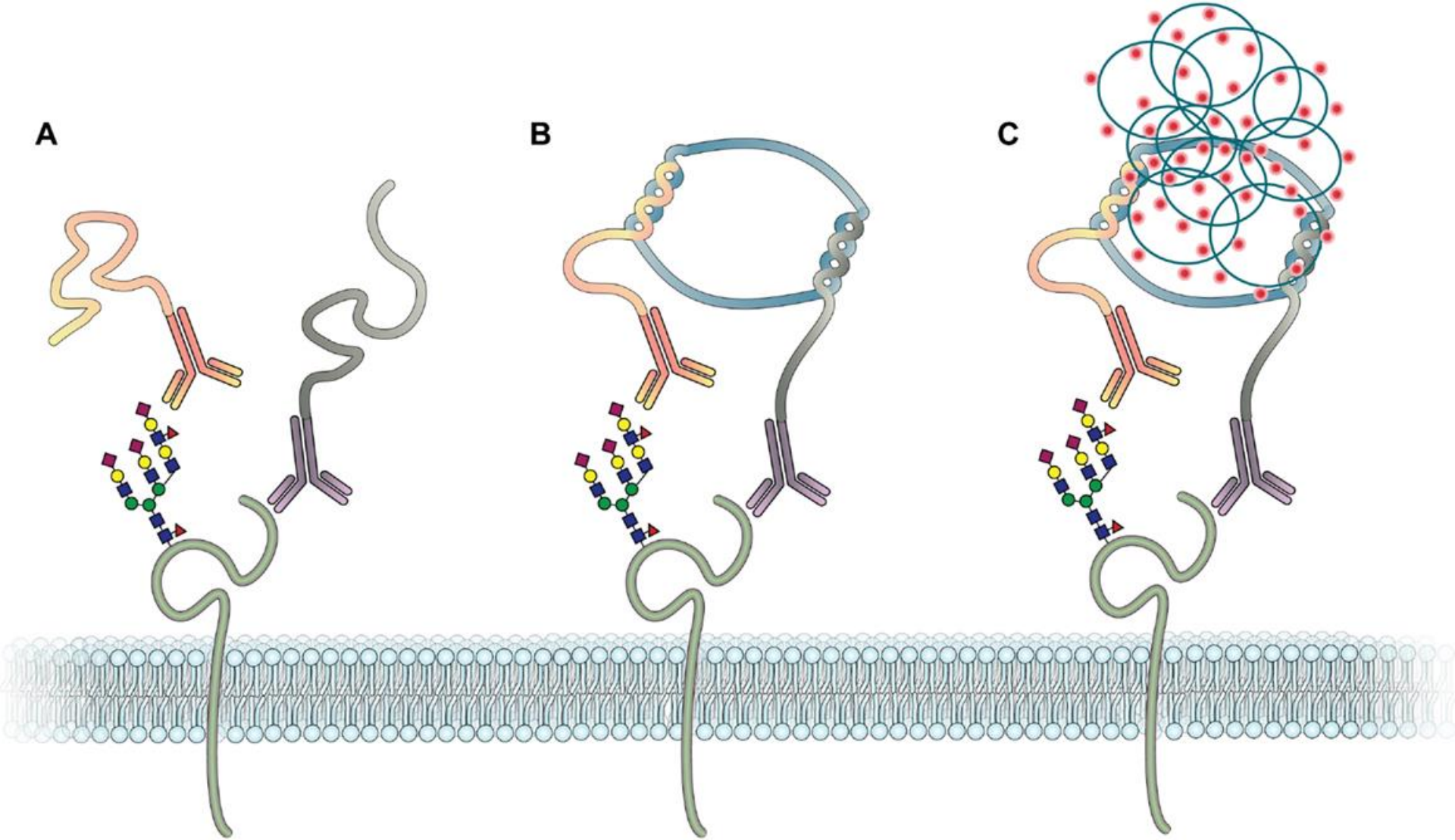
### ***In situ* proximity ligation assay**

$\alpha$ ,  $^{32}\text{P}$ -PDGFR $\beta$  – ( $\beta$ -platelet-derived growth factor receptor)

monoclonal antibodies: rabbit anti-receptor and anti-phosphotyrosine primary antibodies  
 - bound by species-specific antibodies conjugated to oligonucleotides.

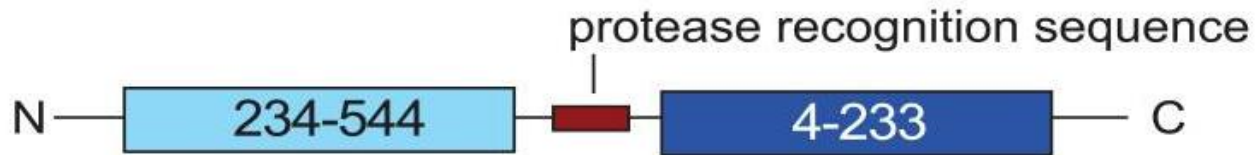
When in proximity - oligonucleotides serve as templates for the joining of two additional linear oligonucleotides into a DNA circle.

*Red* - *in situ*PLA signals, *green* - cytoplasmic staining, and *blue* - cell nuclei.

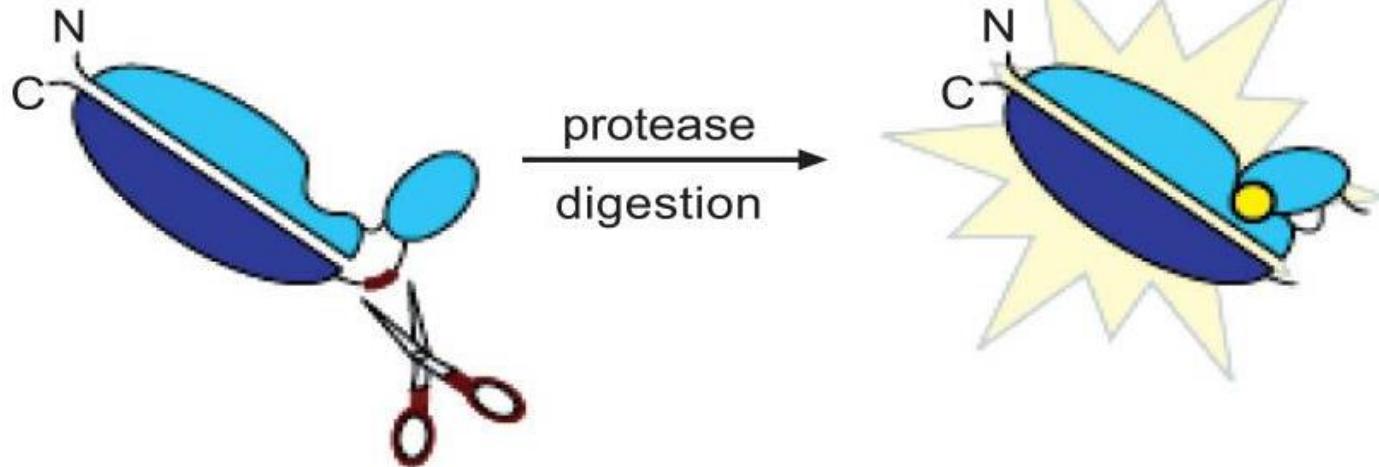


# LinkLight™ technology

A.



B.

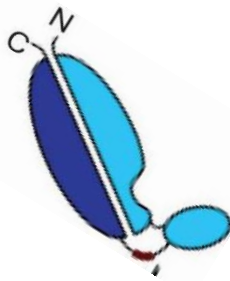
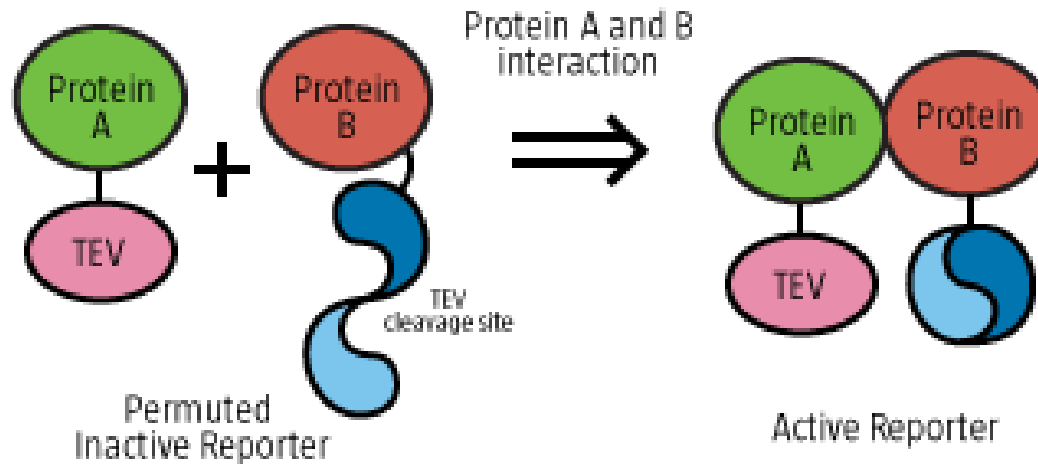


Inactive CP<sub>234</sub>-Luc protein  
(low luciferase activity)

Active CP<sub>234</sub>-Luc protein  
(high luciferase activity)

pLuc - luciferase in two fragments  
connected by a TEV protease cleavage sequence

# LinkLight™ technology



**LinkLight™** - two components.

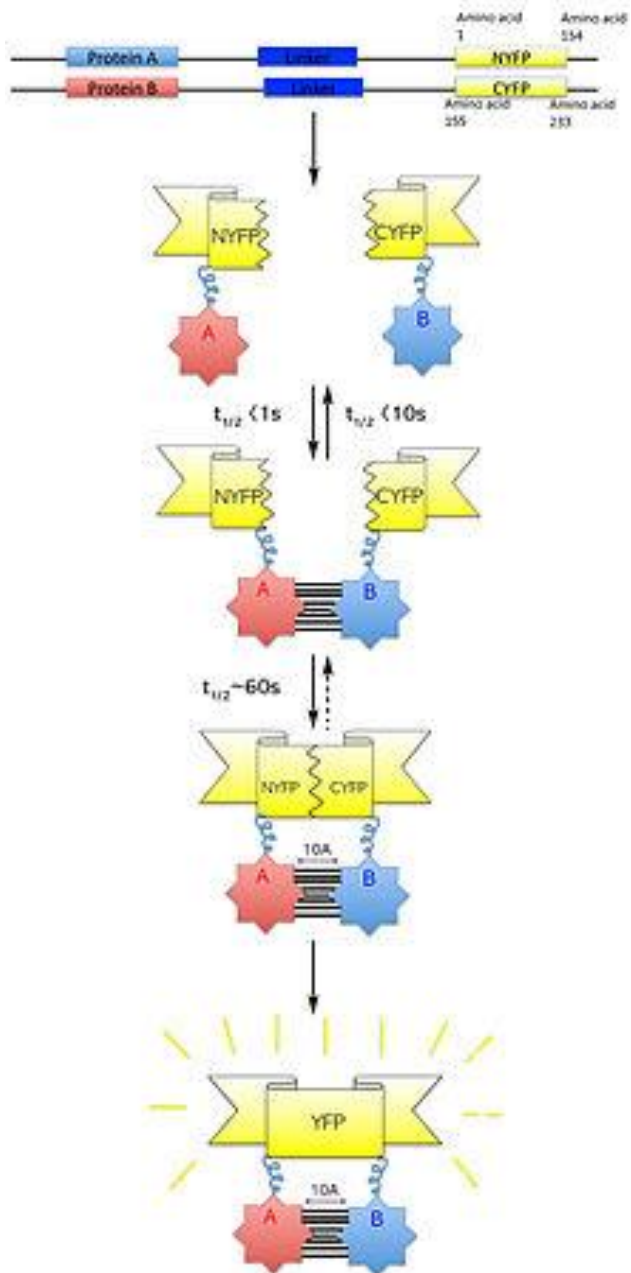
protein A linked to a Tobacco Etch Virus (TEV) protease

protein B is linked to a permuted luciferase (pLuc)

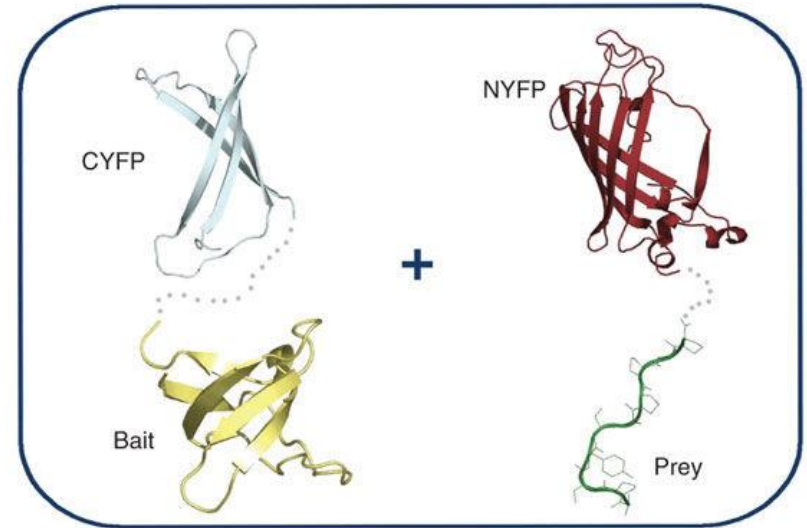
Interaction A + B → inactive pLuc cleaved: luciferase fragments spontaneously refold (fragment self-complementation affinity) → active luciferase



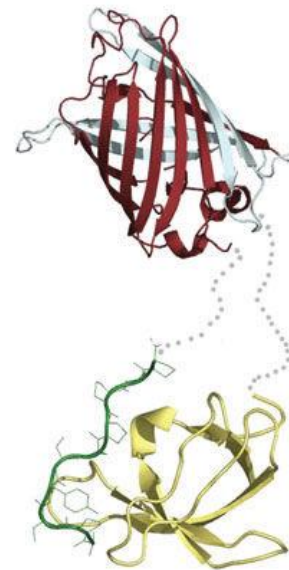
## Two domains of fluorescent domain



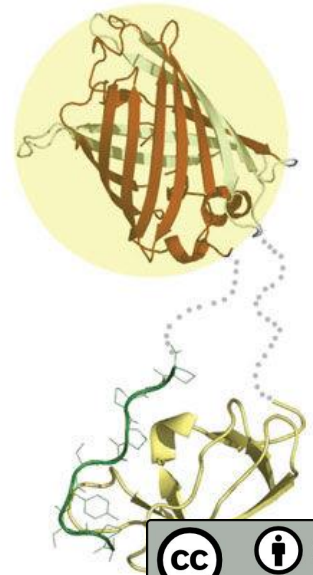
a



b

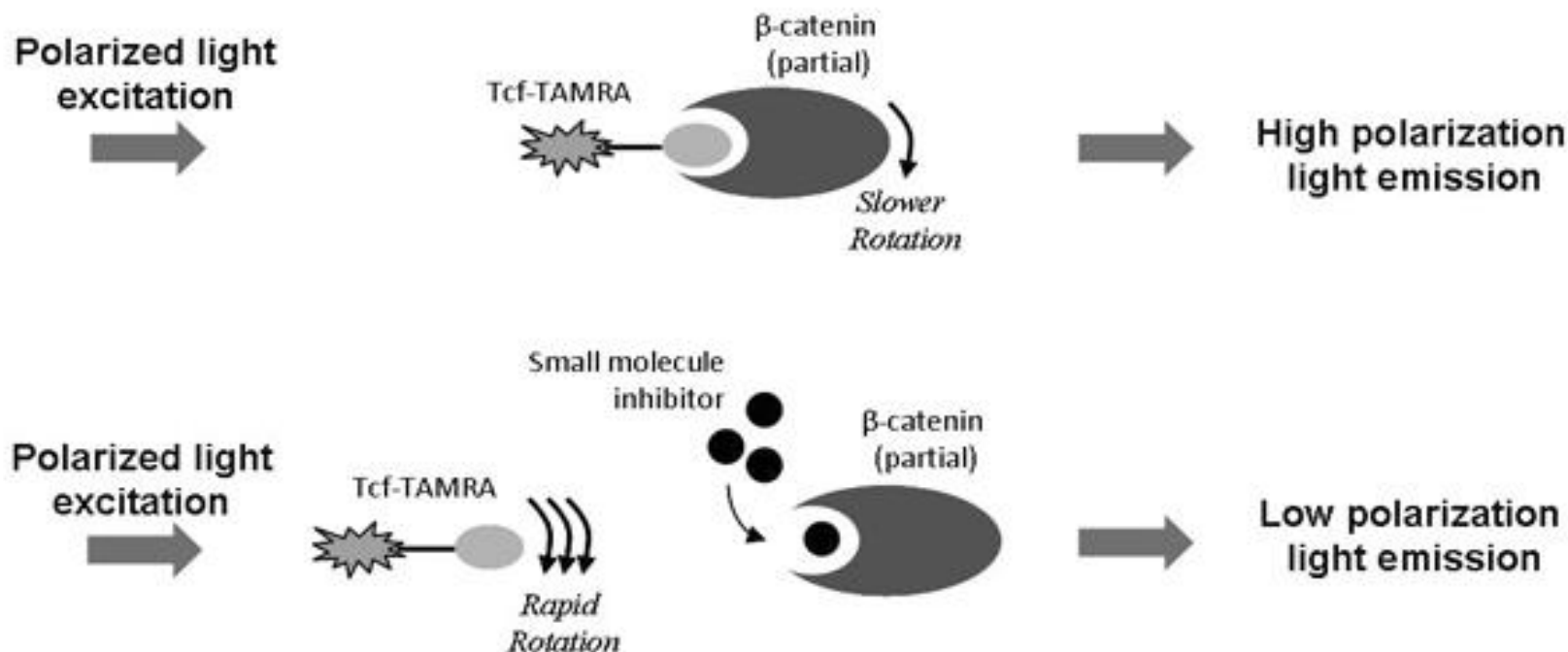


c



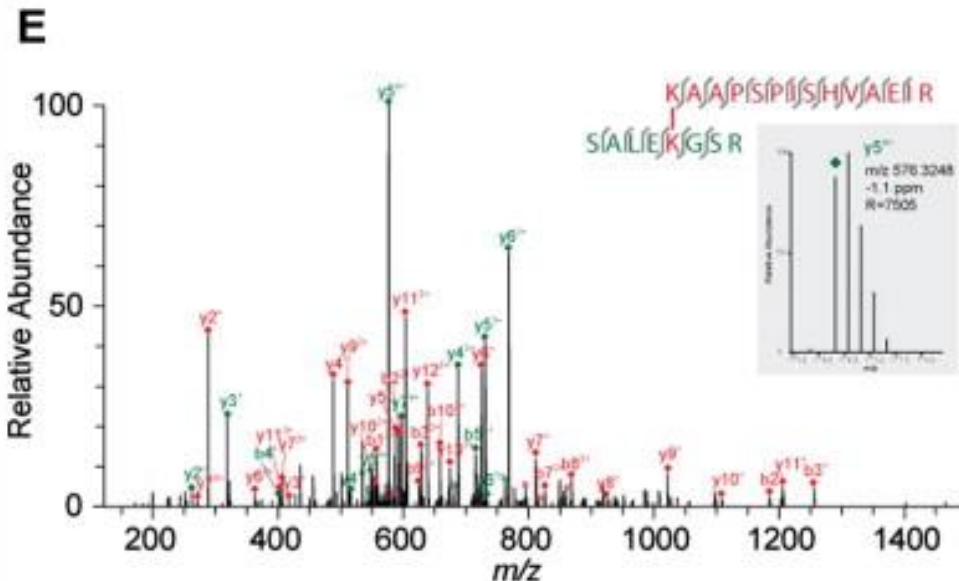
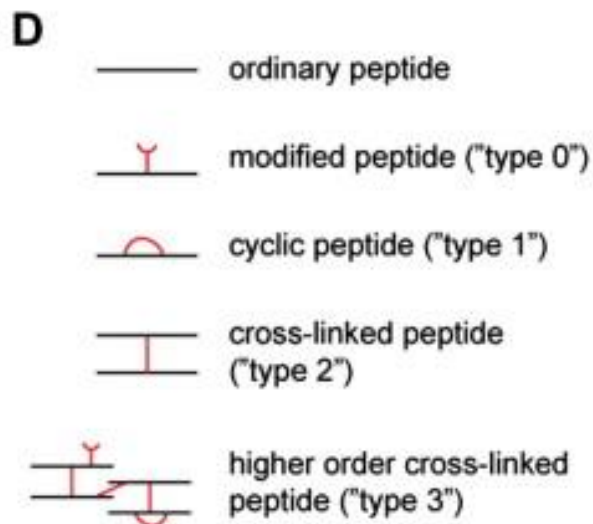
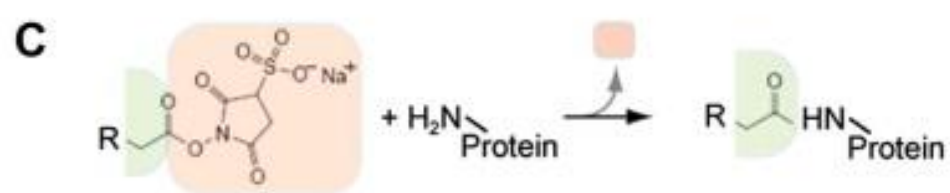
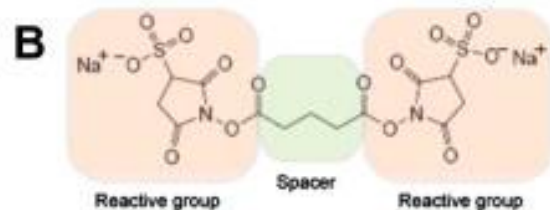
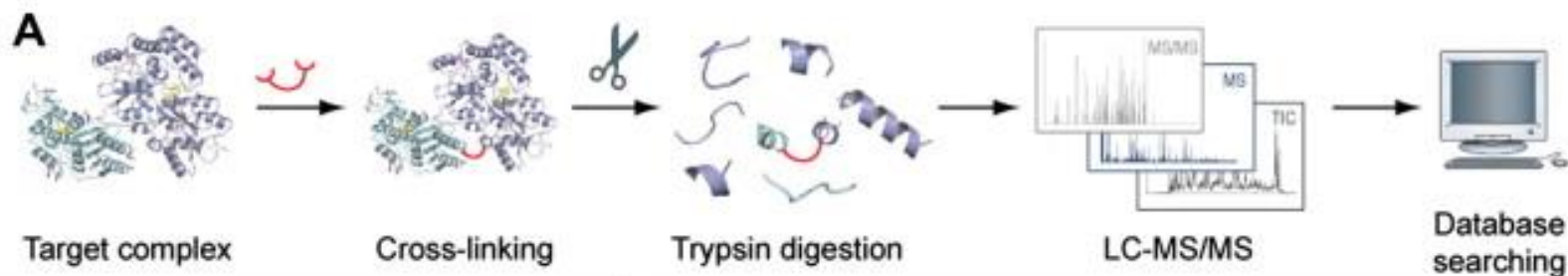
## $\beta$ -Catenin/Tcf Protein-Protein Interaction

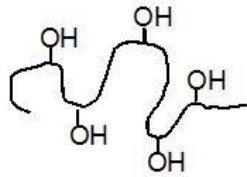
### Fluorescence Polarization Assay



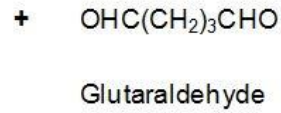
Binding of target protein to TAMRA fluorophore's rotational movement becomes slower due to the increase in molecular mass and thus the emitted light remains polarized



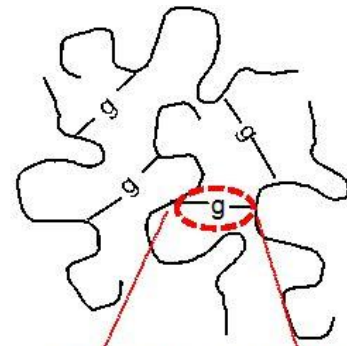




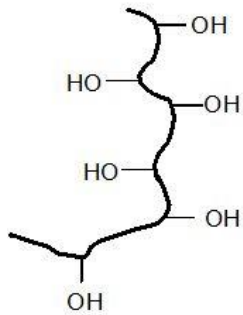
Polymers having functional hydroxyl group



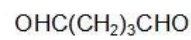
HCl



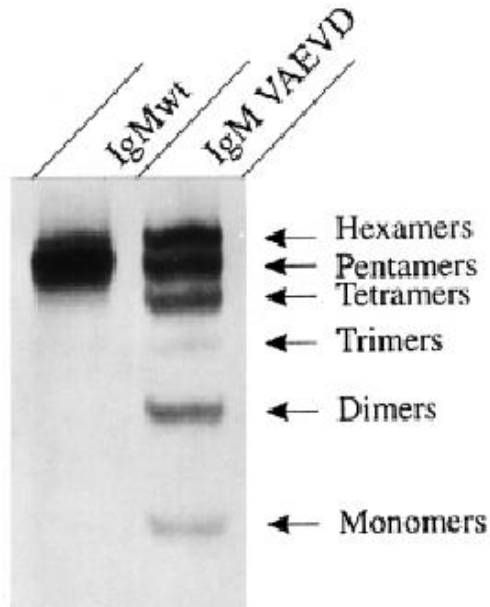
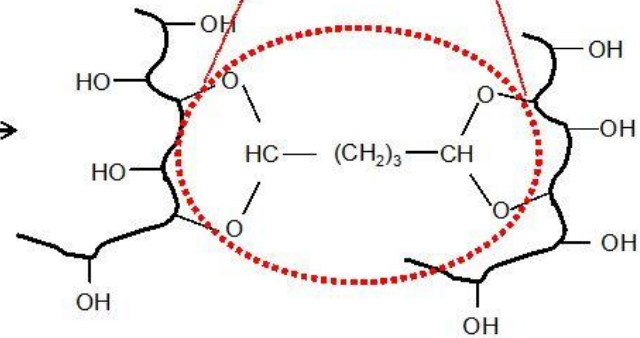
Chemically cross-linked network

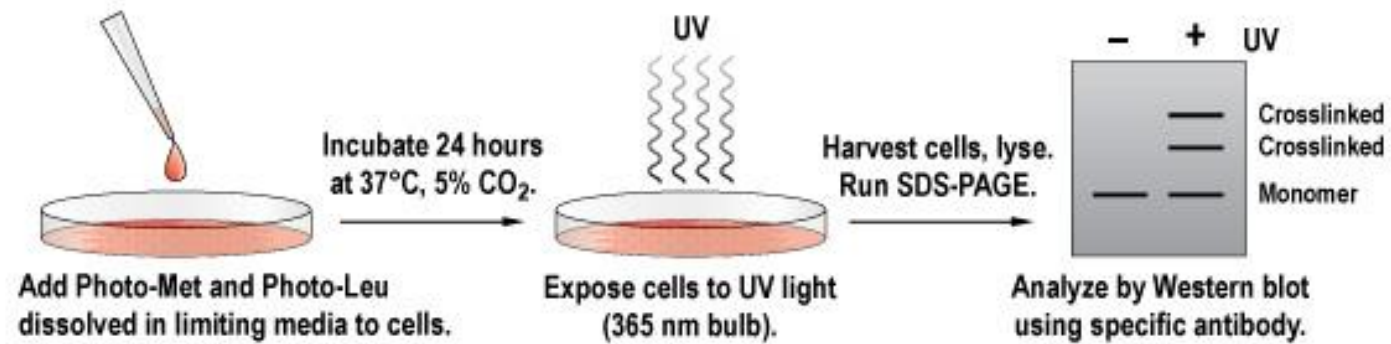


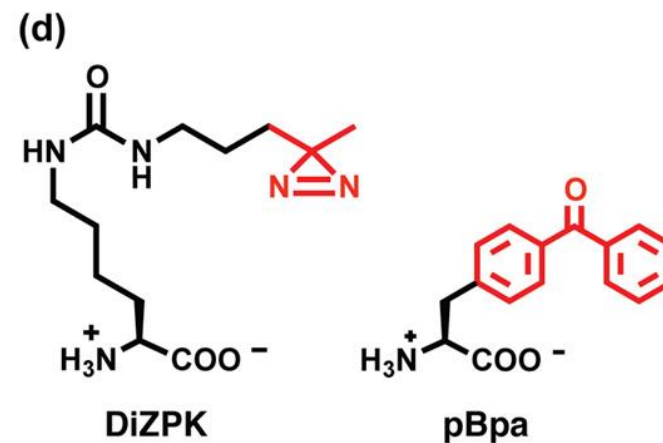
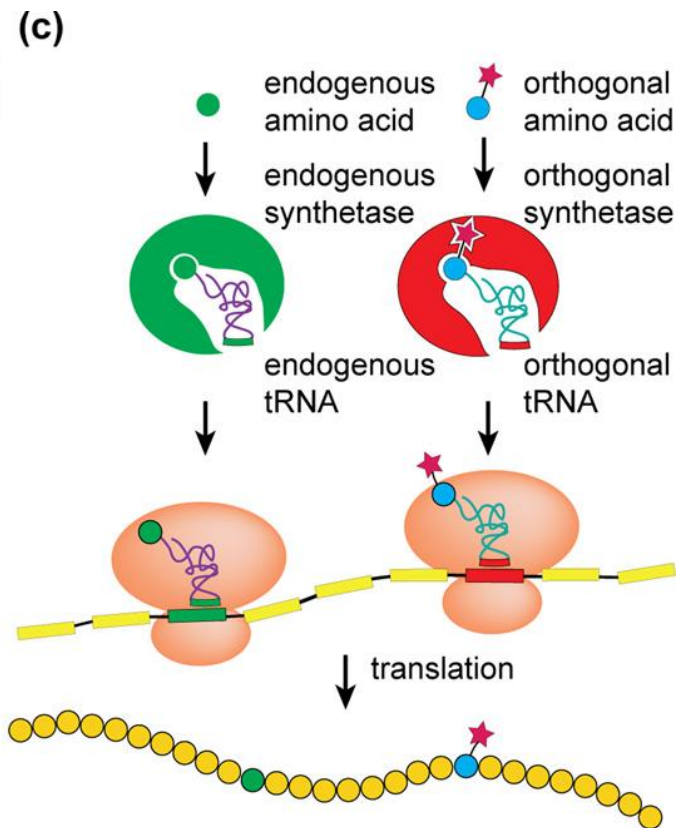
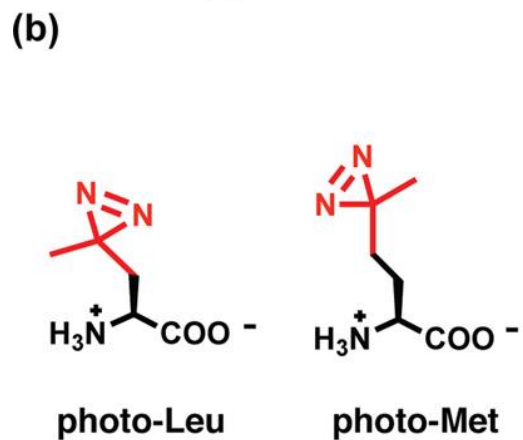
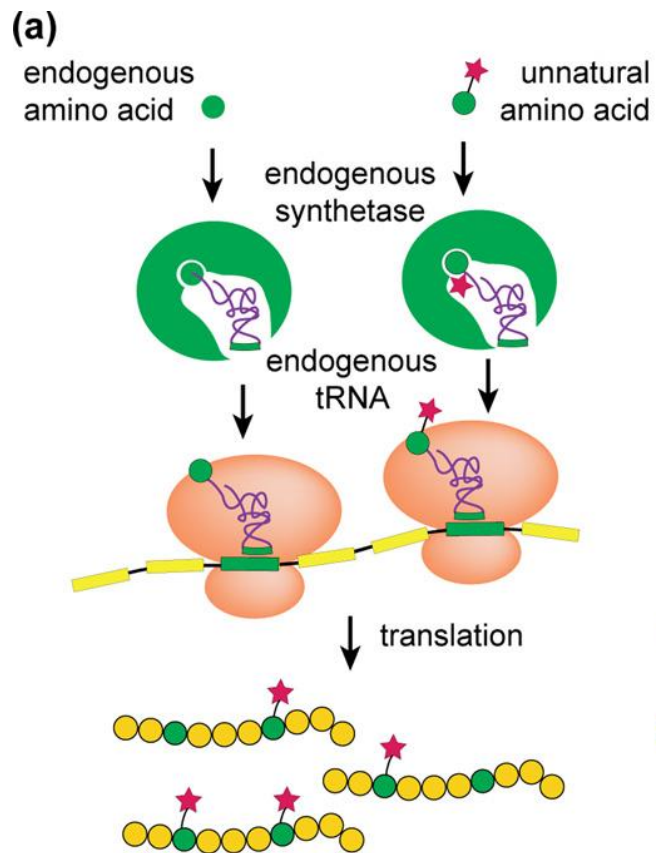
+



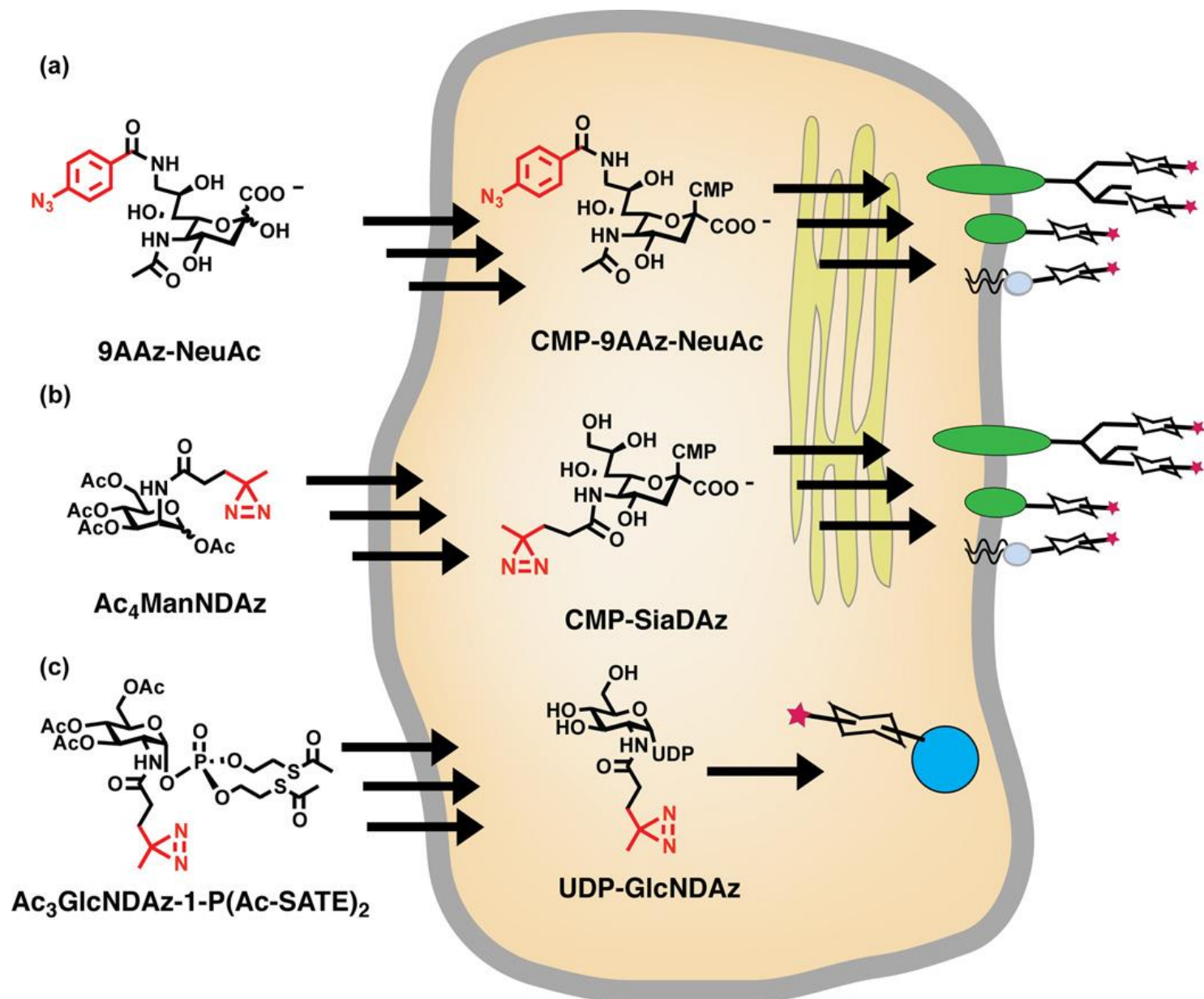
HCl





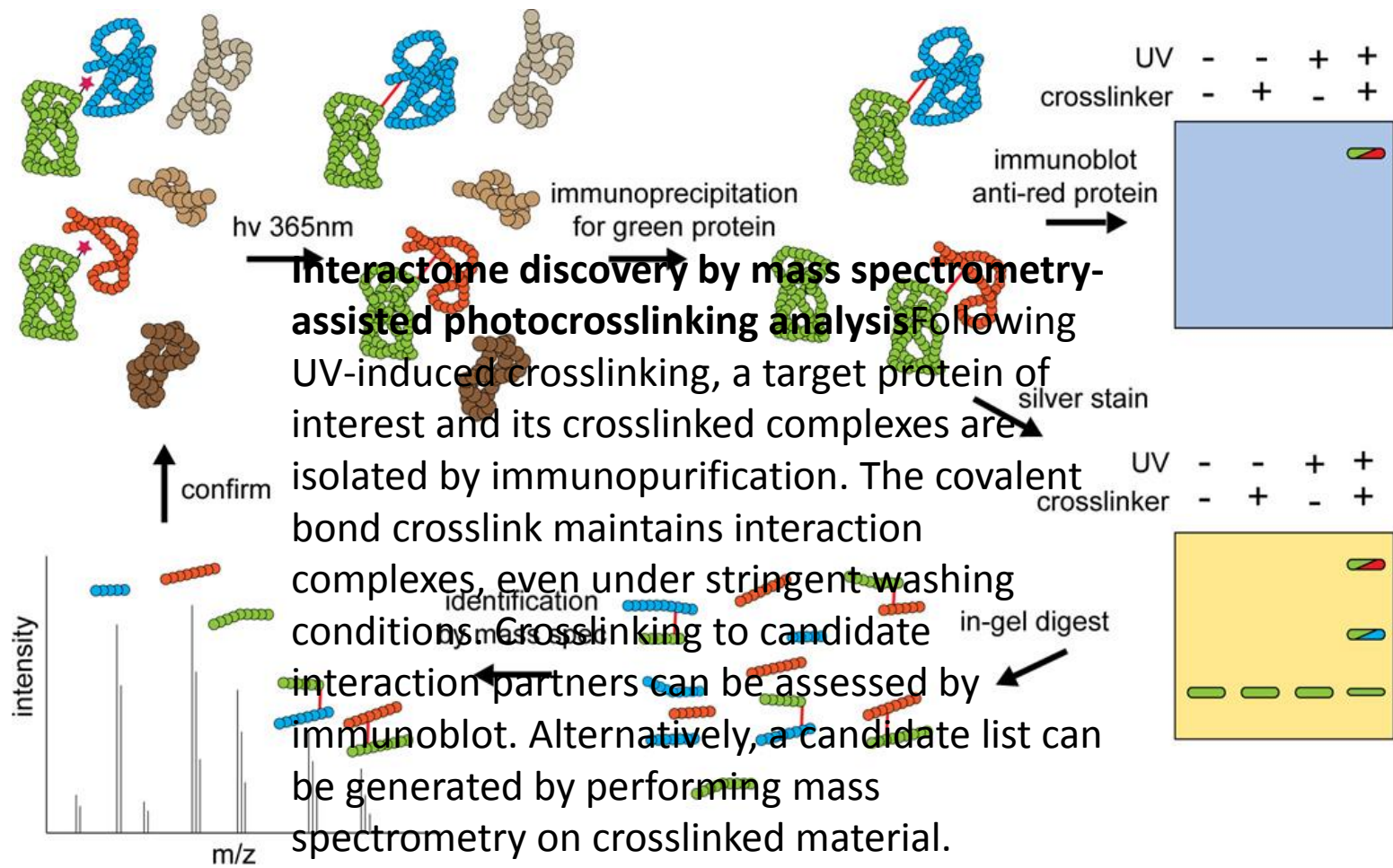


## Methods for incorporating photocrosslinking amino acids

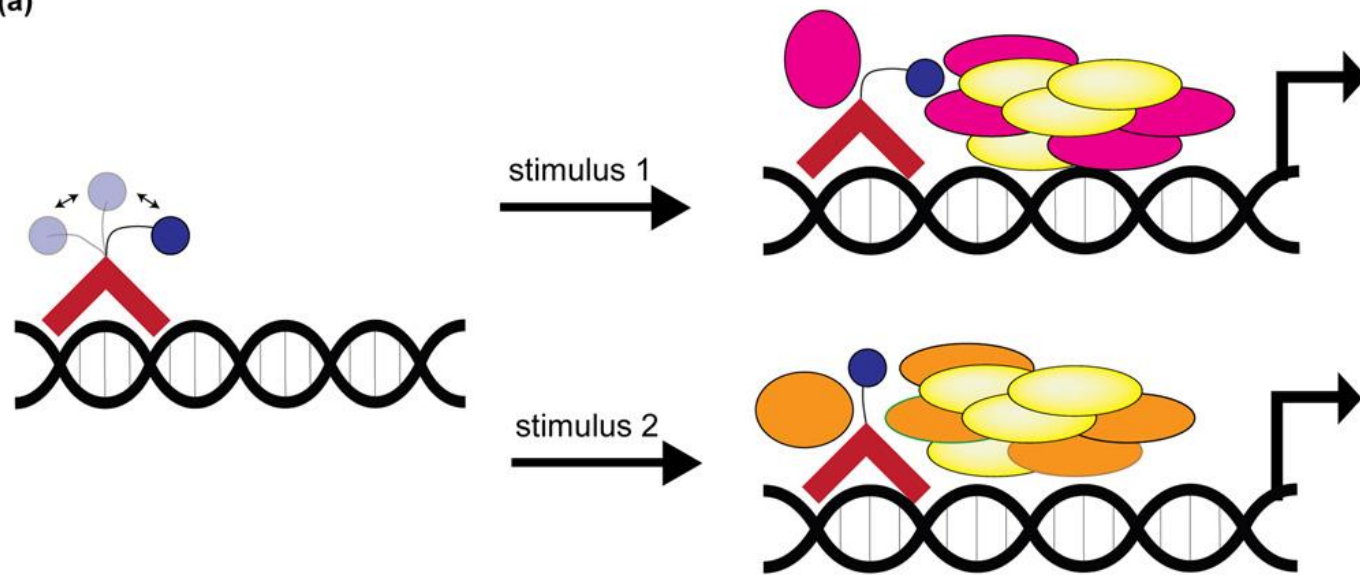


**Metabolic incorporation of photocrosslinking sugars**

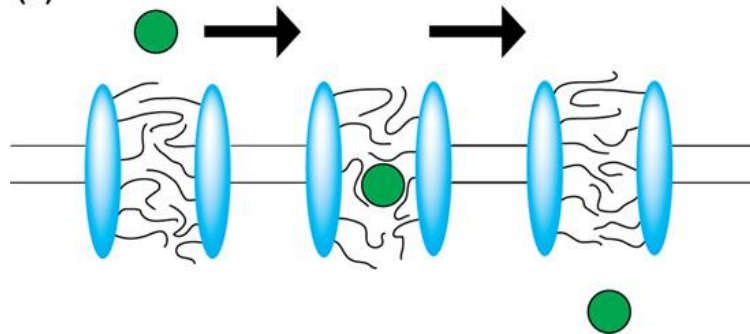




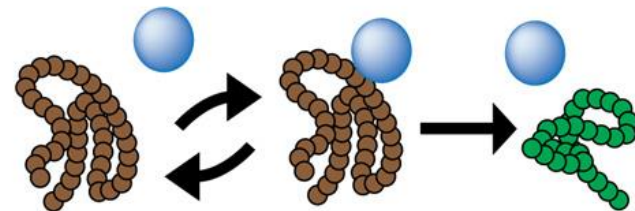
(a)



(b)

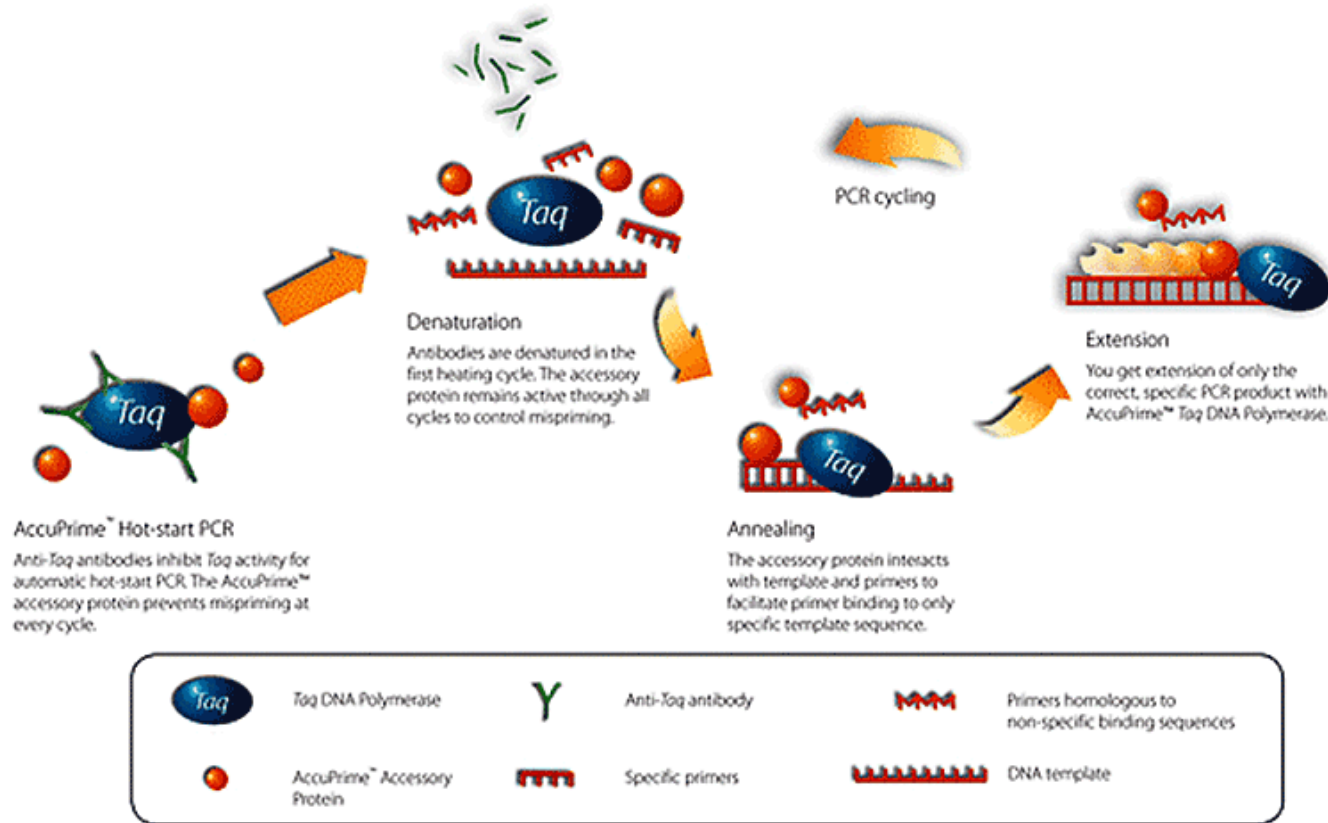


(c)



**Transient, dynamic interactions are fundamental to cellular biology**(a) Activation domains of transcription factors tend to be intrinsically disordered. This flexibility allows the activation domain to interact with different co-activators under different stimuli. (b) FG repeats of the nuclear pore complex form a dynamic sieve that acts as a gateway between the nucleoplasm and the cytoplasm. (c) Molecular chaperones undergo multiple interaction cycles with unfolded proteins. This process is critical for proper folding of nascent proteins and also of misfolded proteins that appear during times of stress.





# Hot-start PCR

DNA polymerase inactivated by a chemical modification or antibody that dissociates at high temperature.

- reduces nonspecific priming and primer-dimer formation and increases product yield.
- chemical hot-start: up to a 10-minute inactivation,
- antibody hot-start - less than two minutes inactivation.
- useful at low amount of DNA template, complex DNA templates or several pairs of primers used , as in multiplex PCR. New hot-start enzymes - good processivity, but no proofreading.

Uveřejněné materiály jsou určeny studentům Vysoké školy chemicko-technologické v Praze

jako studijní materiál. Některá textová i obrazová data v nich obsažená jsou převzata

z veřejných zdrojů. V případě nedostatečných citací nebylo cílem autora/ů záměrně poškodit event. autora/y původního díla.

S eventuálními výhradami se prosím obračejte na autora/y konkrétního výukového materiálu, aby

bylo možné zjednat nápravu.

The published materials are intended for students of the University of Chemistry and Technology, Prague as a study material. Some text and image data contained therein are taken from public sources. In the case of insufficient quotations, the author's intention was not to intentionally infringe the possible author(s) rights to the original work.

If you have any reservations, please contact the author(s) of the specific teaching material in order to remedy the situation.

# Forensic analysis



EUROPEAN UNION  
European Structural and Investing Funds  
Operational Programme Research,  
Development and Education



MINISTRY OF EDUCATION,  
YOUTH AND SPORTS

# Brief History of Forensic DNA Typing

- 1980 - Ray White describes first polymorphic RFLP marker
- 1985 - Alec Jeffreys discovers multilocus VNTR probes
- 1985 - first paper on PCR
- 1988 - FBI starts DNA casework
- 1991 - first STR paper
- 1995 - FSS starts UK DNA database
- 1998 - FBI launches CODIS database

# DNA Use in Forensic Cases

- Most are rape cases (>2 out of 3)
- Looking for match between evidence and suspect
- Must compare victim's DNA profile

## Challenges

- Mixtures must be resolved
- DNA is often degraded
- Inhibitors to PCR are often present

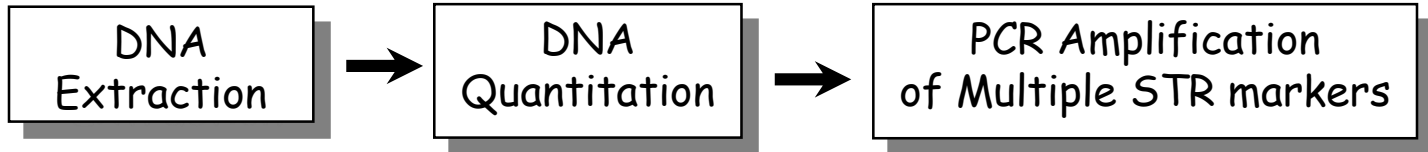
# Human Identity Testing

- Forensic cases -- **matching suspect with evidence**
- Paternity testing -- **identifying father**
- Historical investigations
- Missing persons investigations
- Mass disasters -- **putting pieces back together**
- Military DNA “dog tag”
- Convicted felon DNA databases

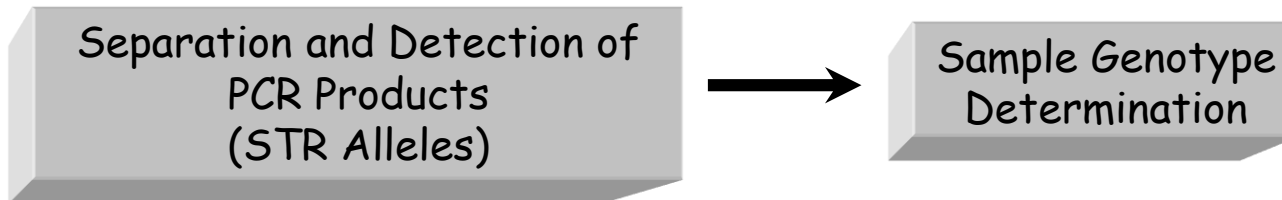
# Steps in DNA Sample Processing

Sample Obtained from  
Crime Scene or Paternity  
Investigation

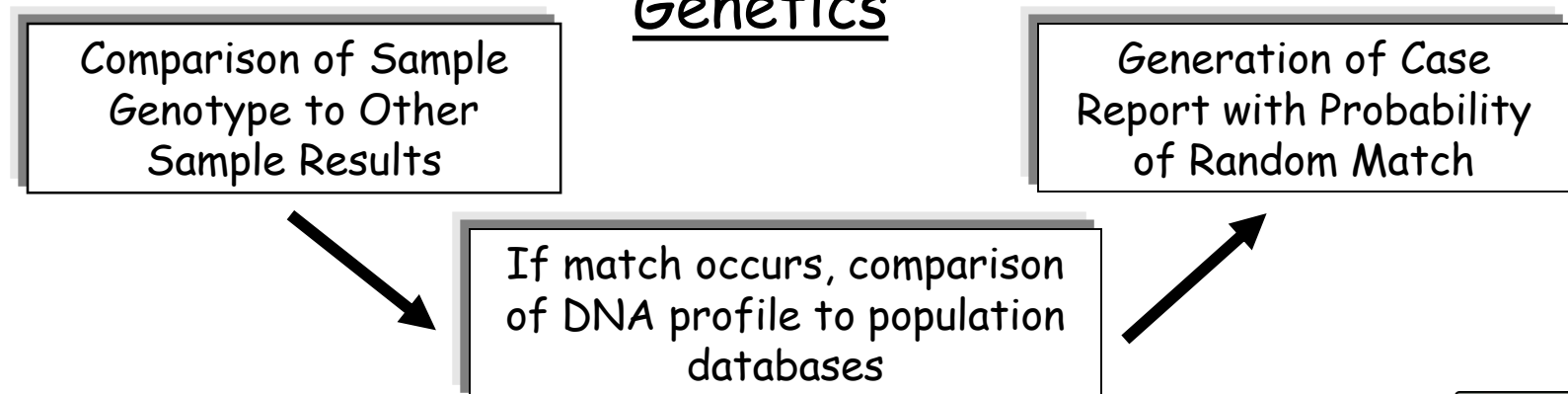
## Biology



## Technology



## Genetics



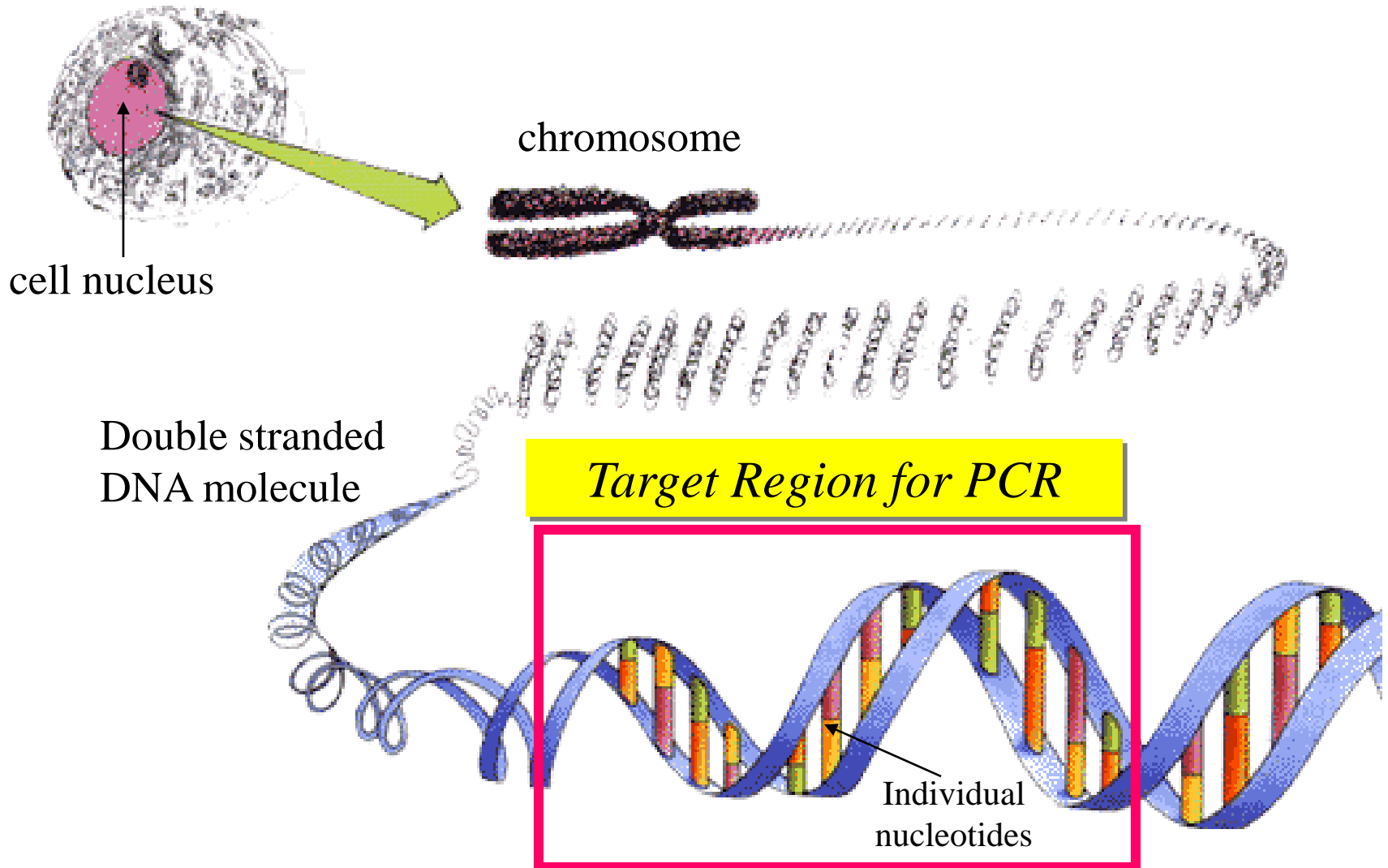


# Sources of Biological Evidence

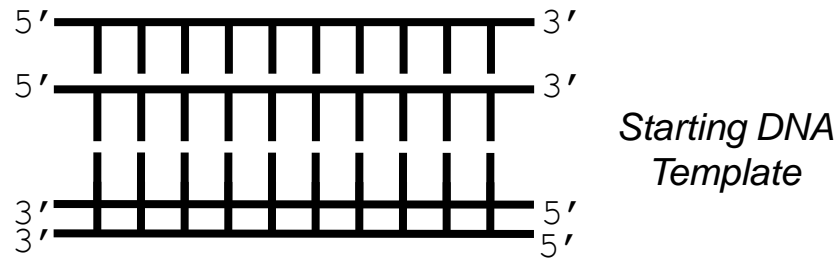
- **Blood**
- **Semen**
- **Saliva**
- **Urine**
- **Hair**
- **Teeth**
- **Bone**
- **Tissue**



# DNA in the Cell

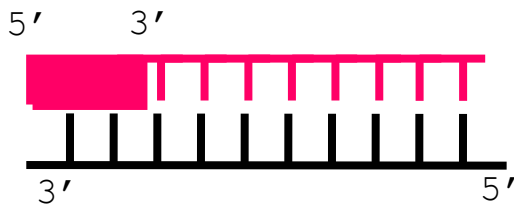


# DNA Amplification with the Polymerase Chain Reaction (PCR)

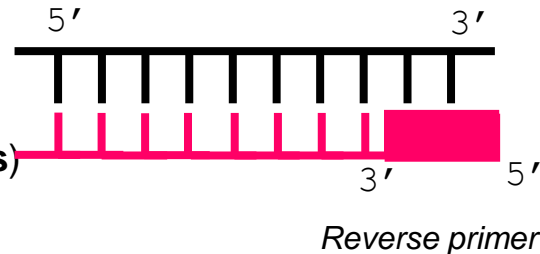


Separate  
strands  
(denature)

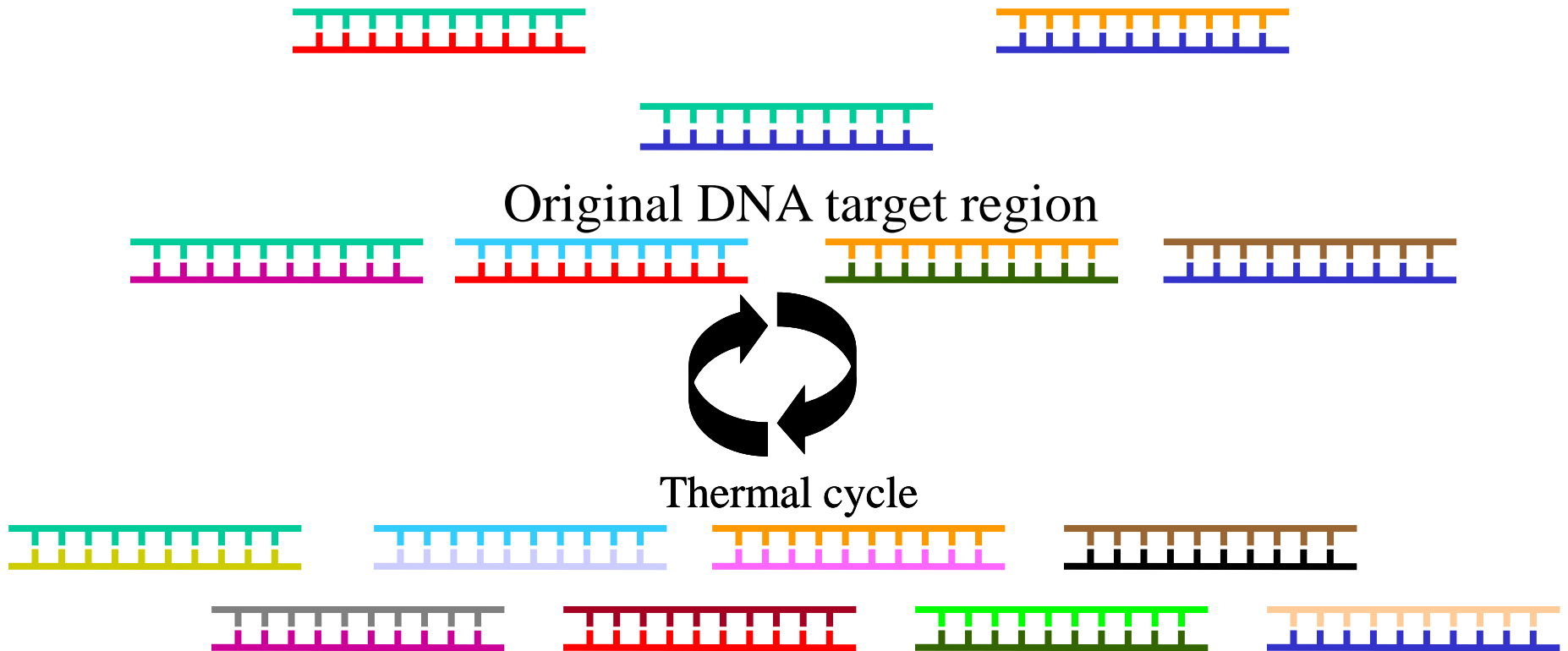
Forward primer



Make copies  
Add primers  
(extend primers)  
(anneal)

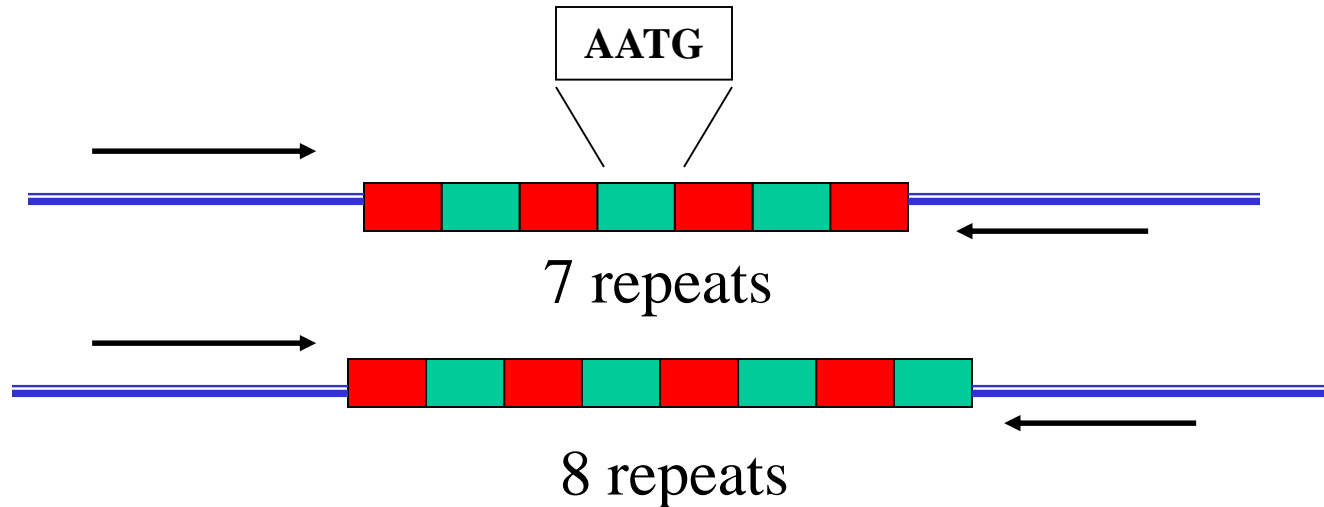


# PCR Copies DNA Exponentially through Multiple Thermal Cycles



*In 32 cycles at 100% efficiency, 1.07 billion copies of targeted DNA region are created*

# Short Tandem Repeats (STRs)

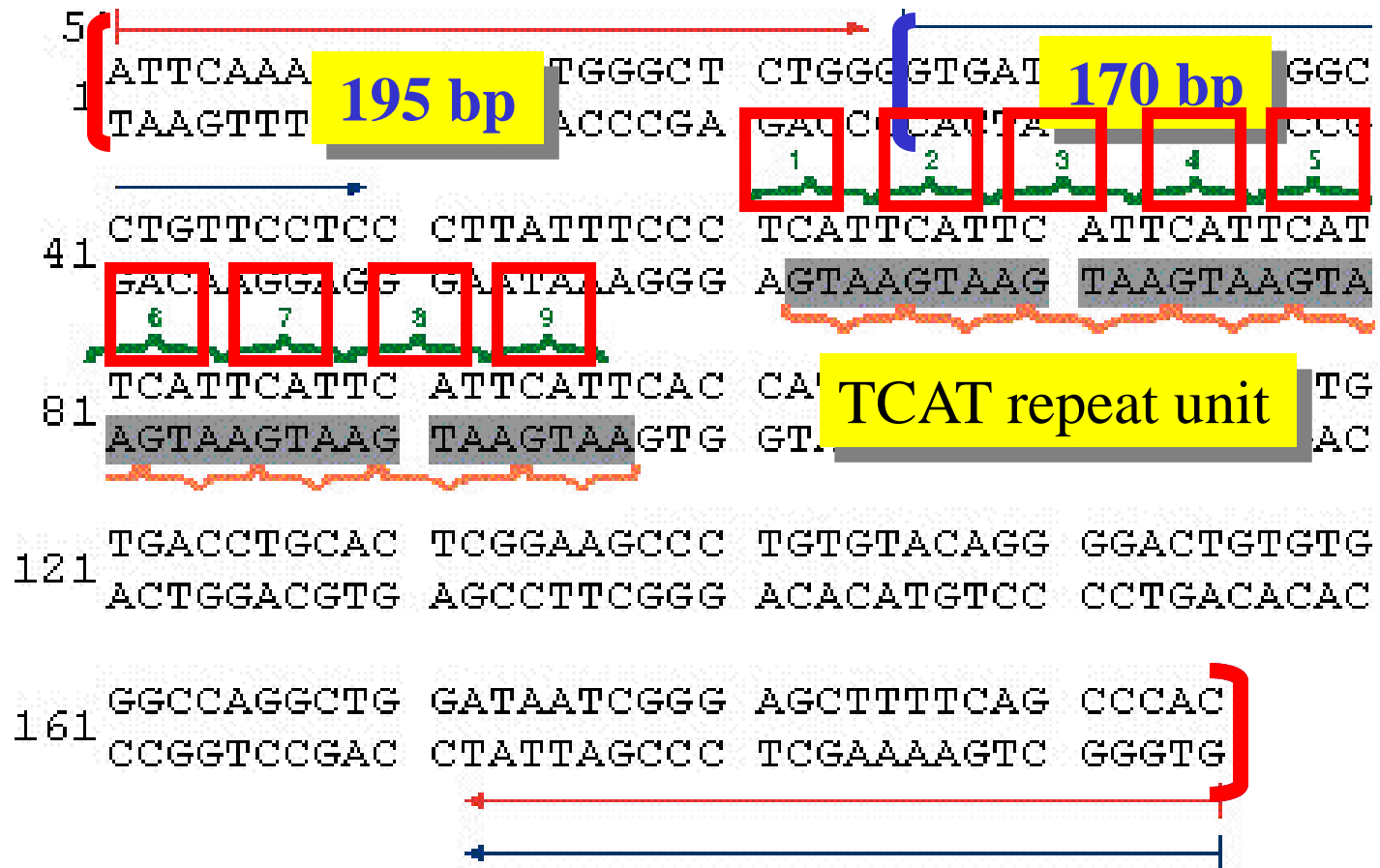


*the repeat region is variable between samples while the flanking regions where PCR primers bind are constant*

Homozygote = both alleles are the same length

Heterozygote = alleles differ and can be resolved from one another

# HUMTH01 Sequence from GenBank (Accession D00269)



*Different primer sets produce different PCR product sizes for the same STR allele*



# Multiplex PCR

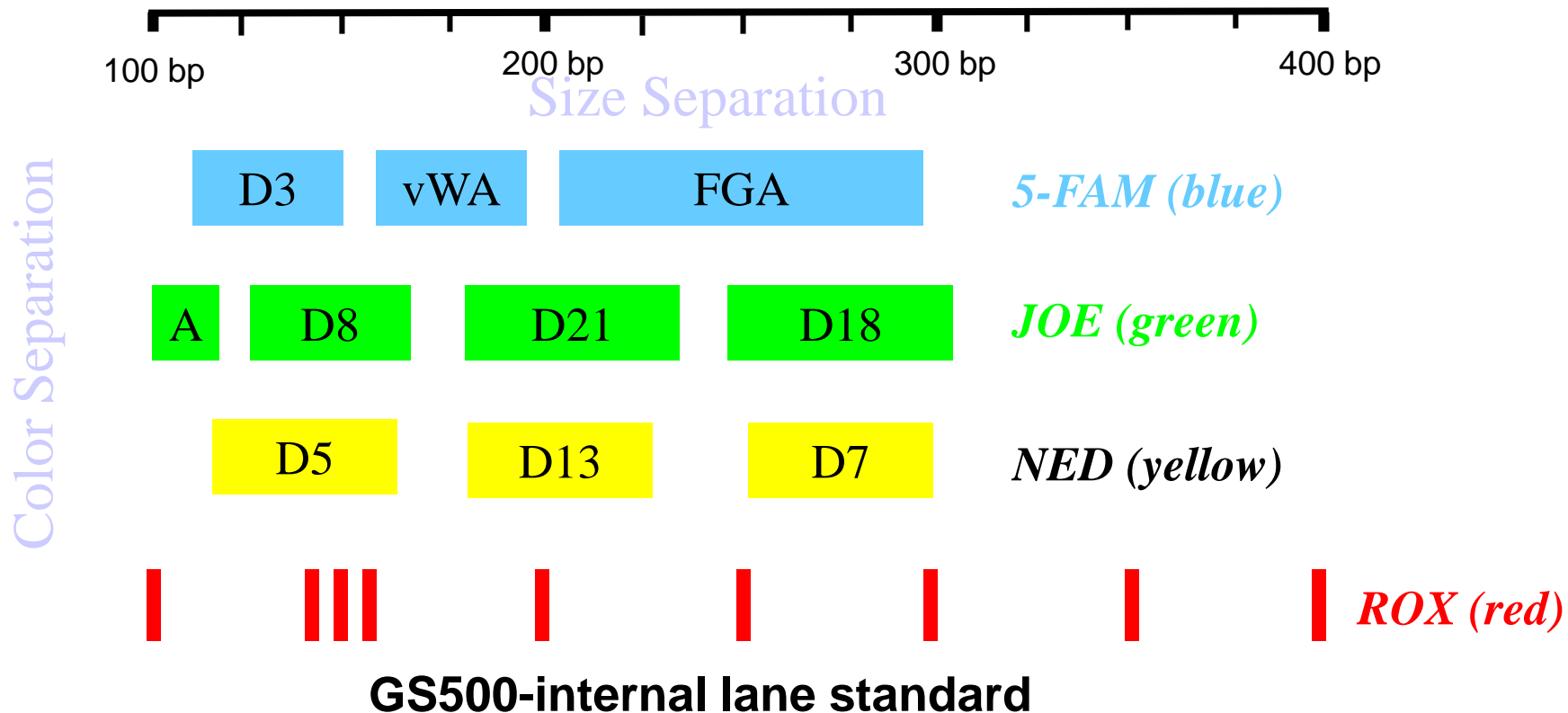
- Over 10 Markers Can Be Copied at Once
- Sensitivities to levels less than 1 ng of DNA
- Ability to Handle Mixtures and Degraded Samples
- Different Fluorescent Dyes Used to Distinguish STR Alleles with Overlapping Size Ranges



# An Example Forensic STR Multiplex Kit

## AmpFISTR® Profiler Plus™

*Kit available from PE Biosystems (Foster City, CA)*

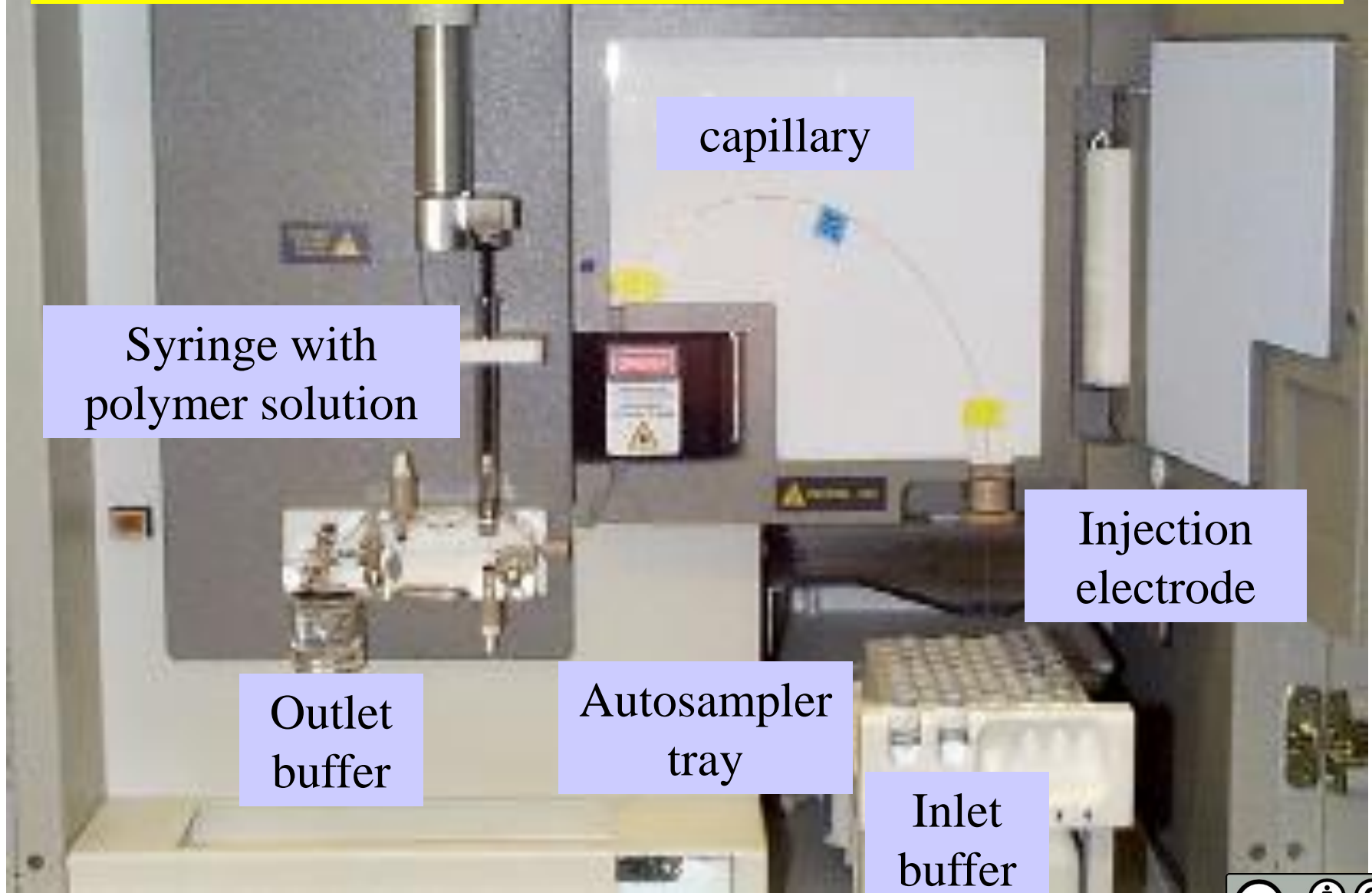


*9 STRs amplified along with sex-typing marker amelogenin in a single PCR*

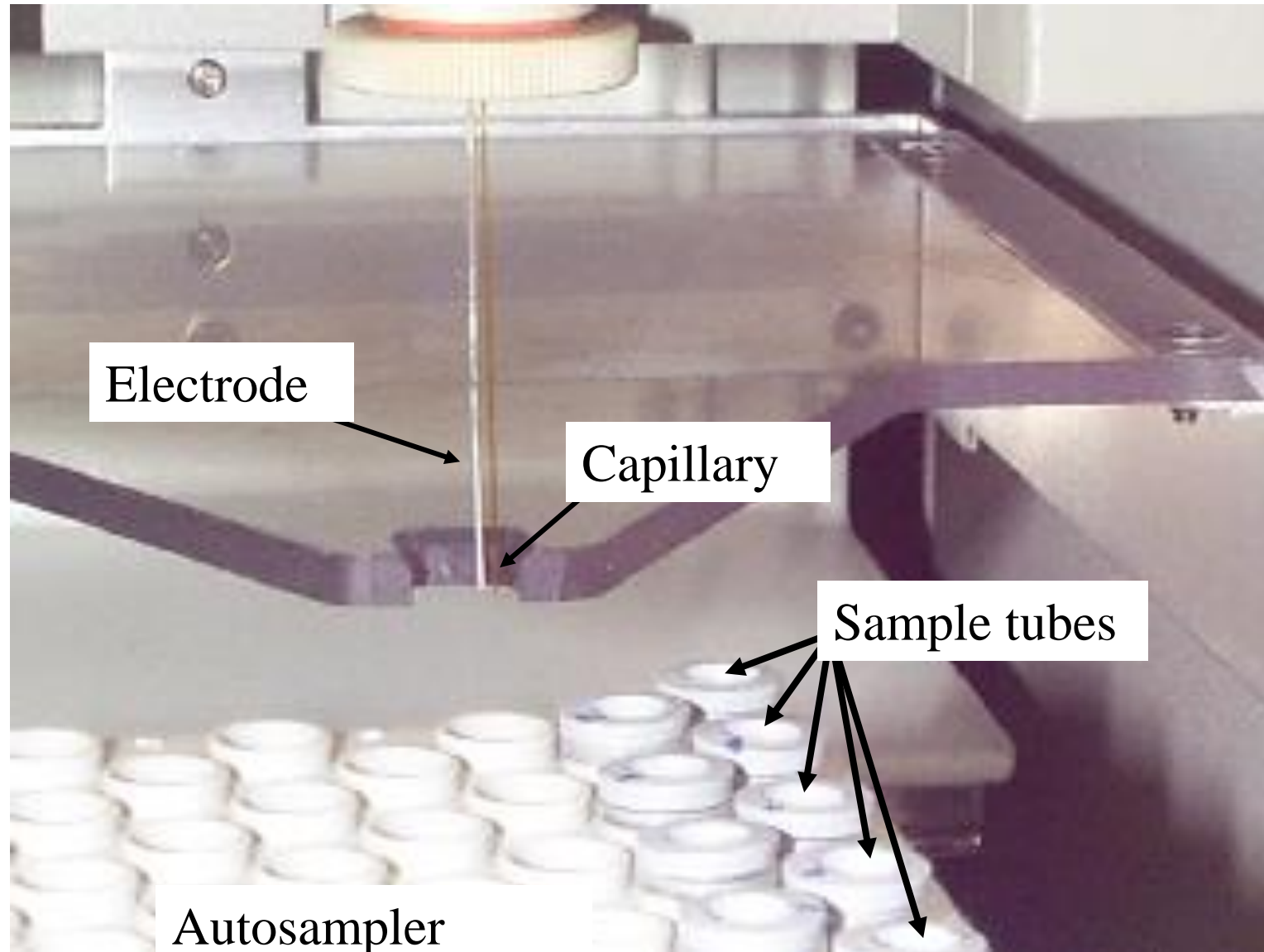
# Available Kits for STR Analysis

- Kits make it easy for labs to just add DNA samples to a pre-made mix
- 13 CODIS core loci
  - Profiler Plus and COfiler (PE Applied Biosystems)
  - PowerPlex 1.1 and 2.1 (Promega Corporation)
- Increased power of discrimination
  - CTT (1994): 1 in 410
  - SGM Plus<sup>TM</sup> (1999): 1 in 3 trillion
  - PowerPlex<sup>TM</sup> 16 (2000): 1 in  $2 \times 10^{17}$

# ABI Prism 310 Genetic Analyzer

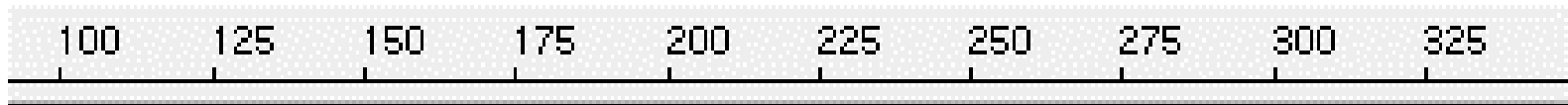


# ABI Prism 310 Sample Loading Area



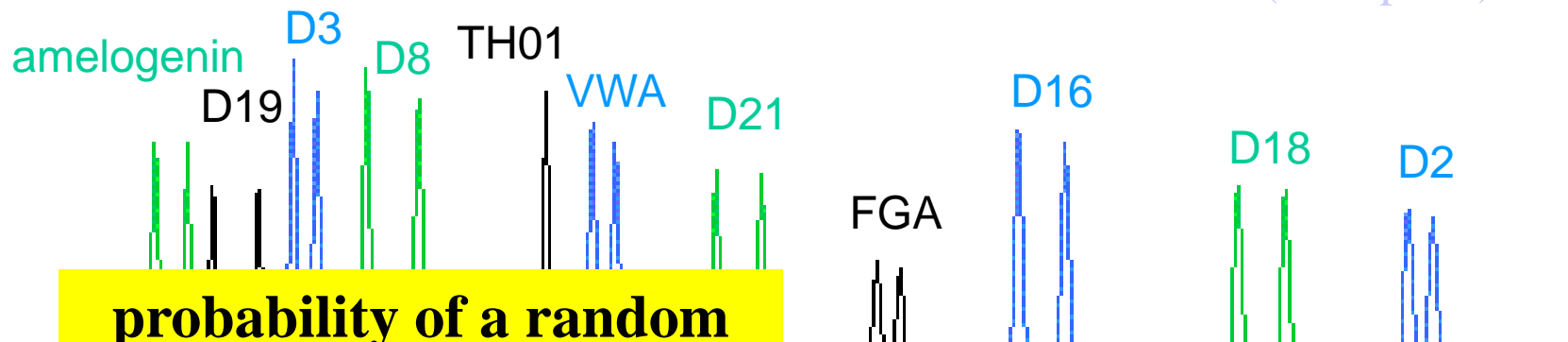
# Human Identity Testing with Multiplex STRs

AmpFlSTR® SGM Plus™ kit



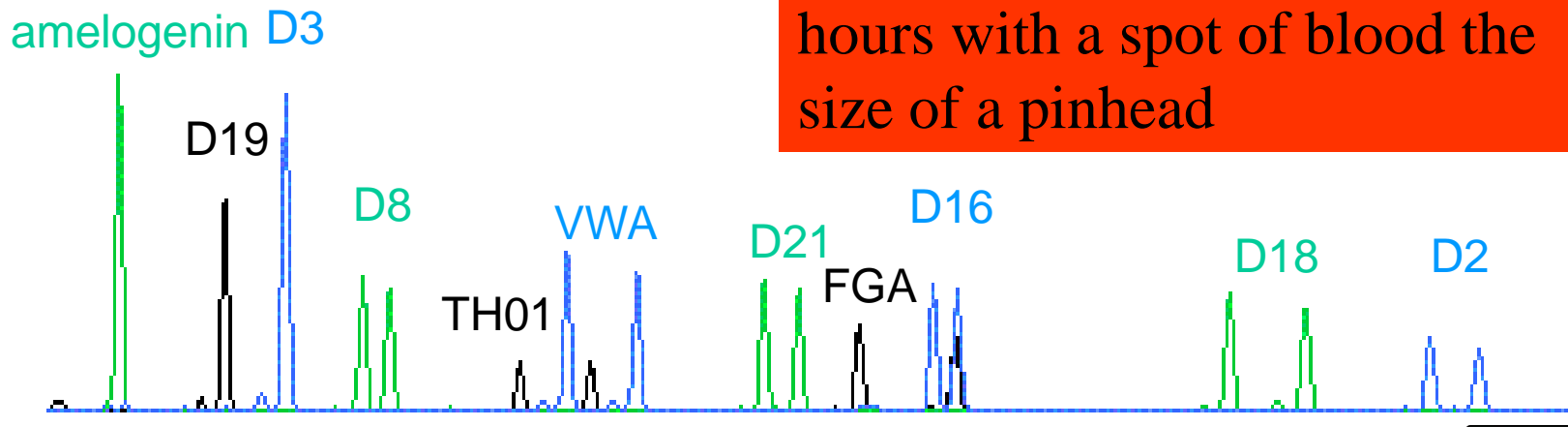
DNA Size (base pairs)

Two different individuals



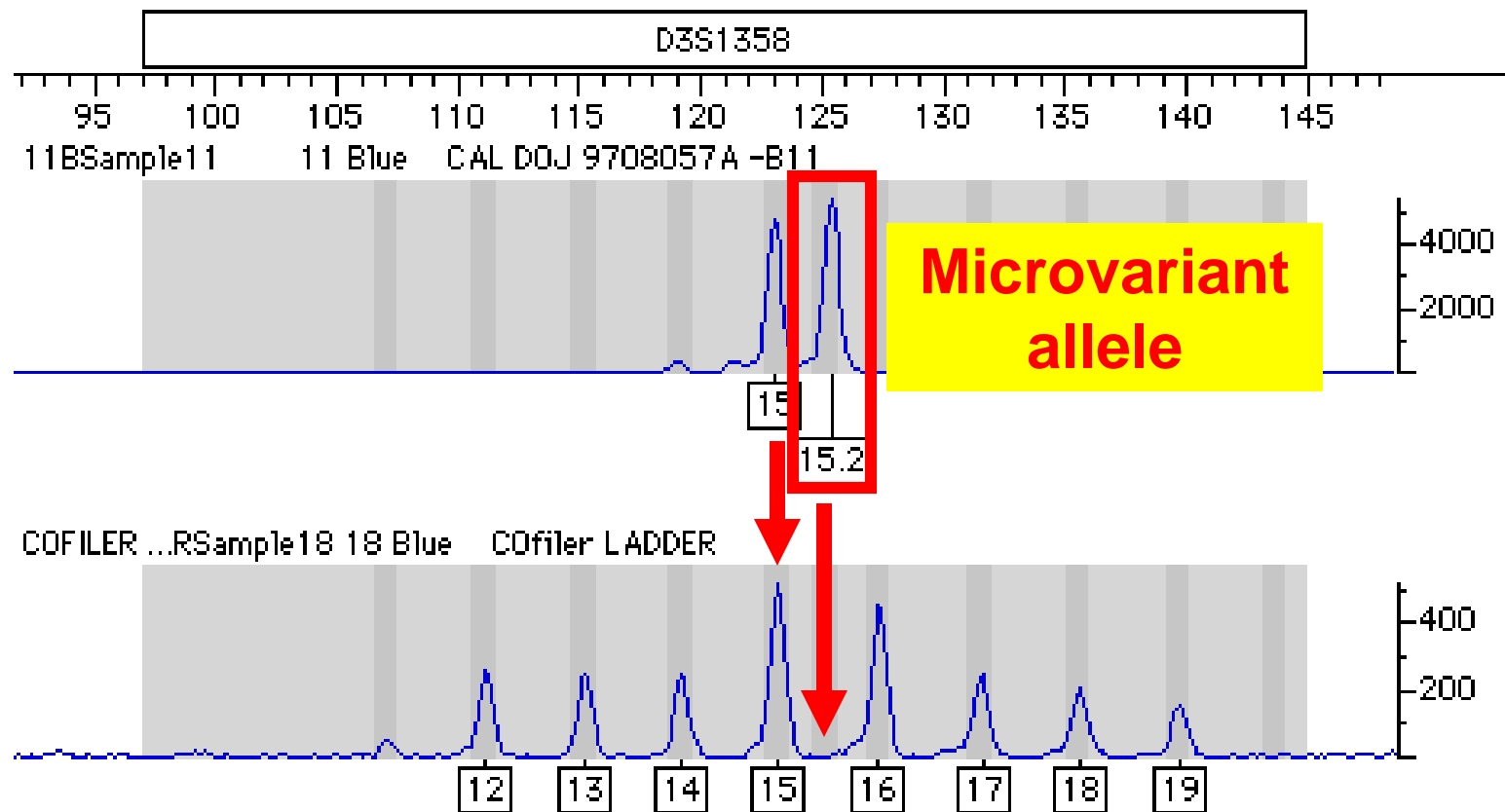
**probability of a random match: ~1 in 3 trillion**

**Results obtained in less than 5 hours with a spot of blood the size of a pinhead**

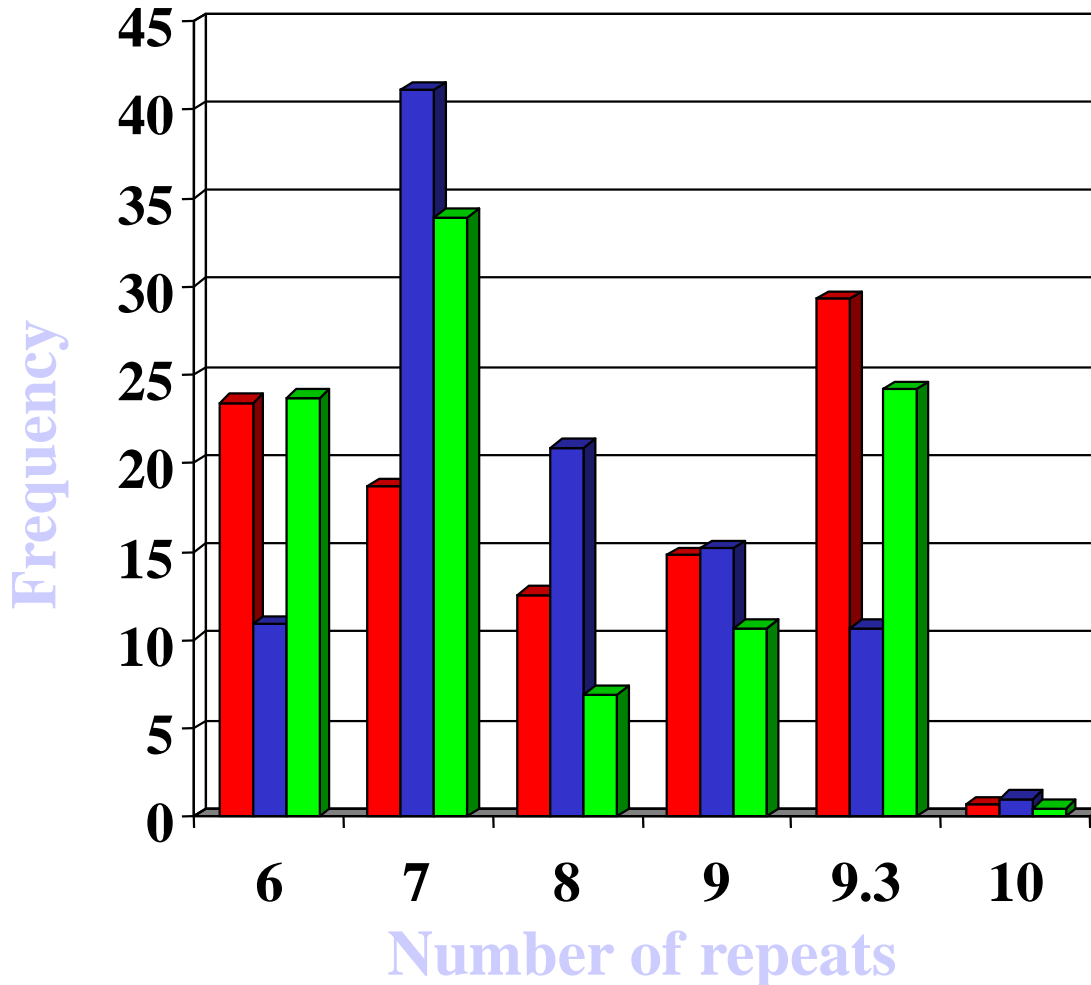


**Simultaneous Analysis of 10 STRs and Gender ID**

# STR genotyping is performed by comparison of sample data to allelic ladders



# STR Allele Frequencies



TH01 Marker

■ Caucasians (N=427)  
■ Blacks (N=414)  
■ Hispanics (N=414)

\**Proc. Int. Sym. Hum. ID*  
(Promega) 1997, p. 34



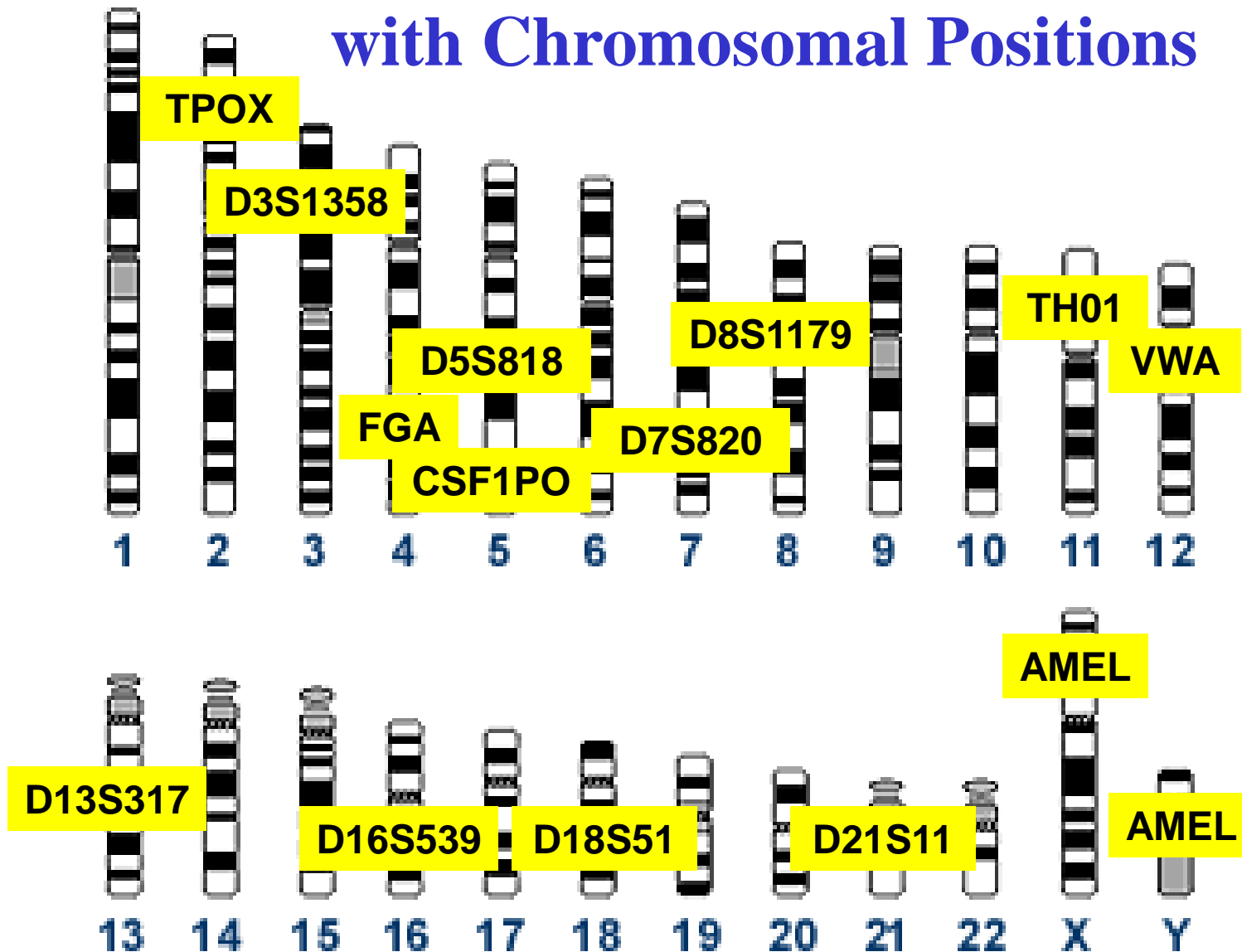
# FBI's CODIS DNA Database

## Combined DNA Index System

- Used for linking serial crimes and unsolved cases with repeat offenders
- Launched October 1998
- Links all 50 states
- Requires >4 RFLP markers and/or 13 core STR markers
- Current backlog of >600,000 samples



# 13 CODIS Core STR Loci with Chromosomal Positions



Uveřejněné materiály jsou určeny studentům Vysoké školy chemicko-technologické v Praze

jako studijní materiál. Některá textová i obrazová data v nich obsažená jsou převzata

z veřejných zdrojů. V případě nedostatečných citací nebylo cílem autora/ů záměrně poškodit event. autora/y původního díla.

S eventuálními výhradami se prosím obračejte na autora/y konkrétního výukového materiálu, aby

bylo možné zjednat nápravu.

The published materials are intended for students of the University of Chemistry and Technology, Prague as a study material. Some text and image data contained therein are taken from public sources. In the case of insufficient quotations, the author's intention was not to intentionally infringe the possible author(s) rights to the original work.

If you have any reservations, please contact the author(s) of the specific teaching material in order to remedy the situation.

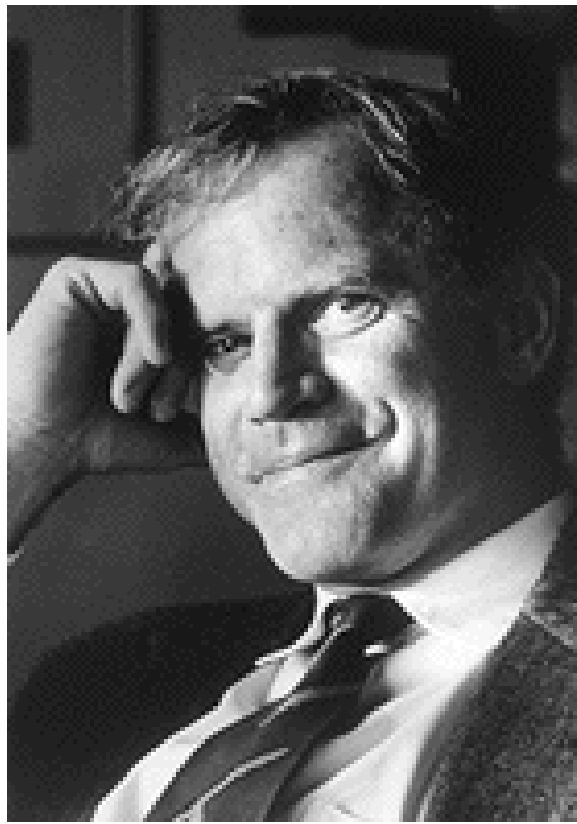
# PCR



EUROPEAN UNION  
European Structural and Investing Funds  
Operational Programme Research,  
Development and Education



MINISTRY OF EDUCATION,  
YOUTH AND SPORTS



## PCR 2 principles

### **DNA Hybridization**

**DNA oligonucleotide - duplex with complementary ssDNA sequence**

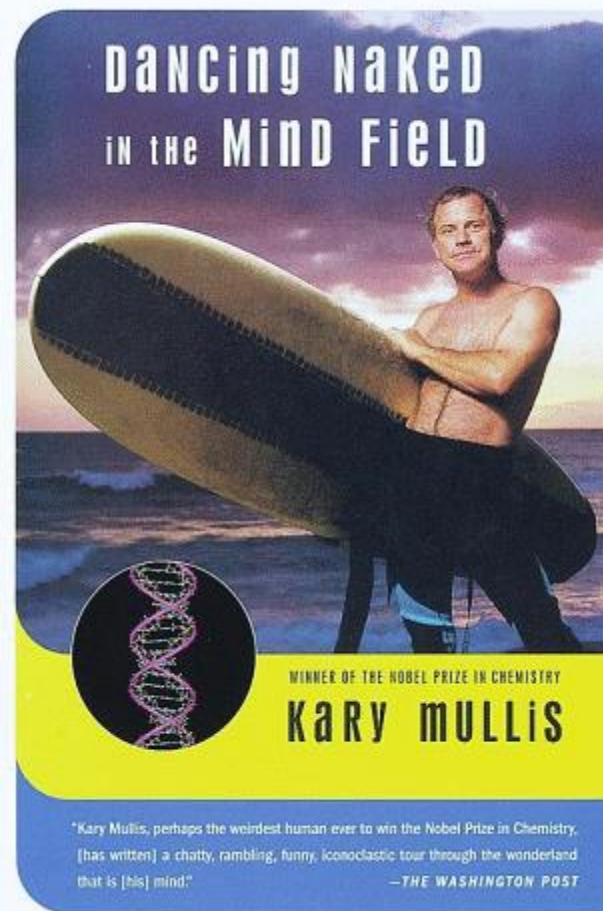
**DNA polymerase cannot initiate DNA replication**

**- but can extend a primer**



## **1993 Nobel prize for chemistry**

1991 Cetus sold to Hoffman-La Roche company PCR patent  
for 300 millions \$



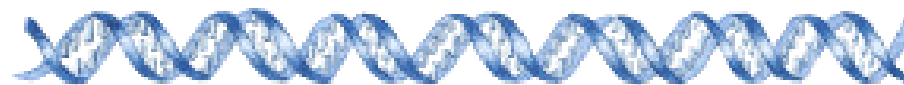
Mullis is no shy, socially inept bench chemist,

- full life as possible, opening himself to experiences like hallucinogenic drugs, surfing, casually handling dangerous chemicals, and taking shots at the sacred cows of science. *Dancing Naked in the Mind Field* is Mullis's own chronicle of his adventures, from wooing countless women to possibly being abducted by aliens, and it's a funny, shocking tale indeed.

This man certainly doesn't suffer from lack of self-esteem, and yet you might want him along on a trip to the astral plane, say, or a tour of the human genome. Mullis is a fascinating character and his autobiography will put to rest forever the stereotype of scientist as skeptical nerd.







Double-helical  
DNA

Denaturation

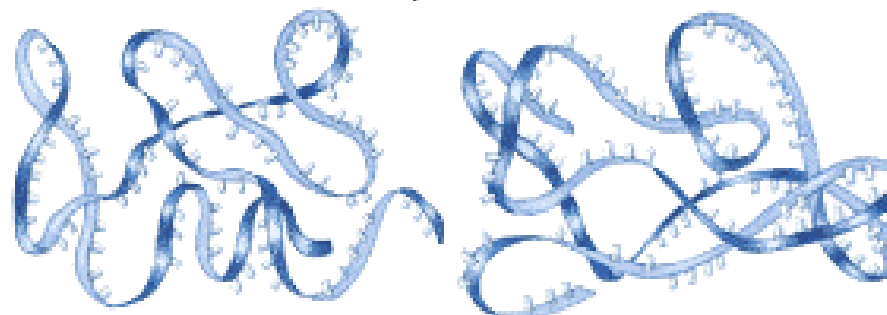
Annealing



Partially denatured  
DNA

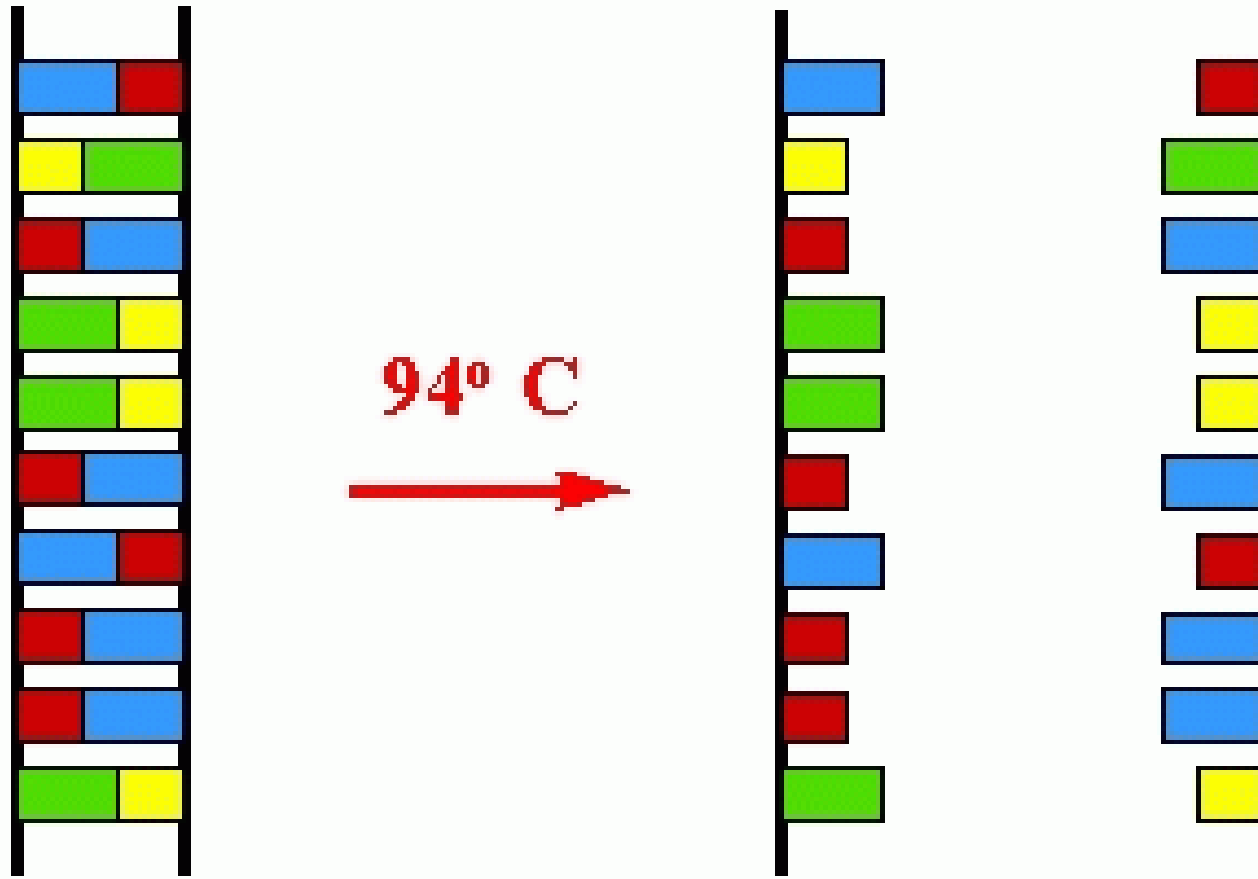
Separation  
of strands

Association of  
strands by base  
pairing

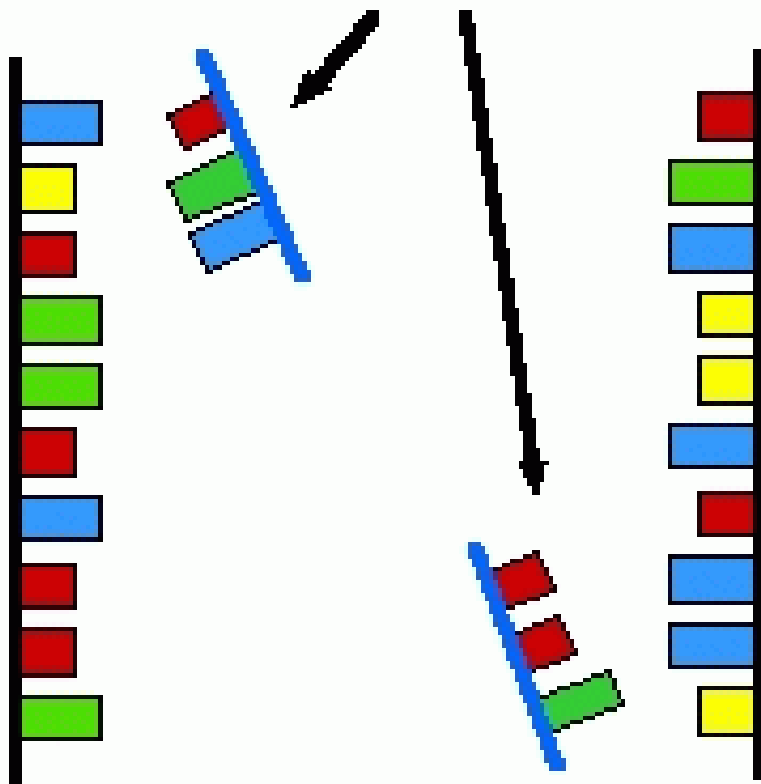


Separated strands  
of DNA in random coils

# DNA STRANDS ARE SEPARATED BY HEATING

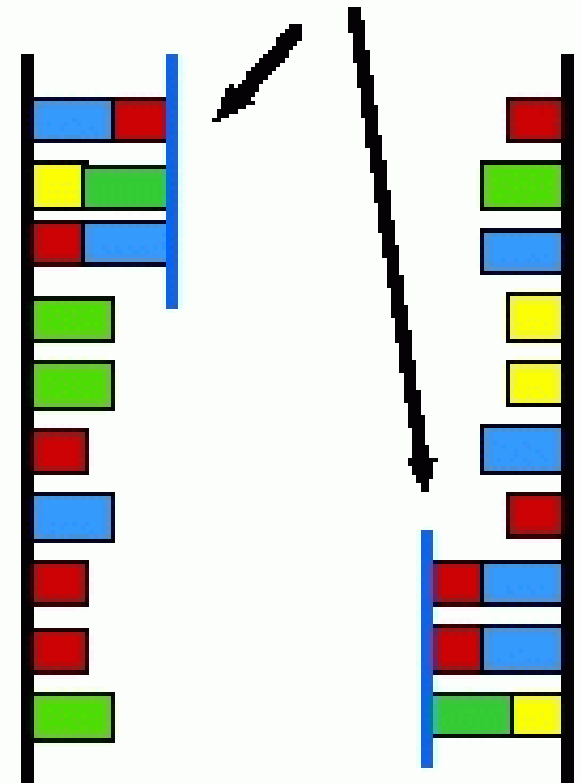


## UNATTACHED PRIMERS



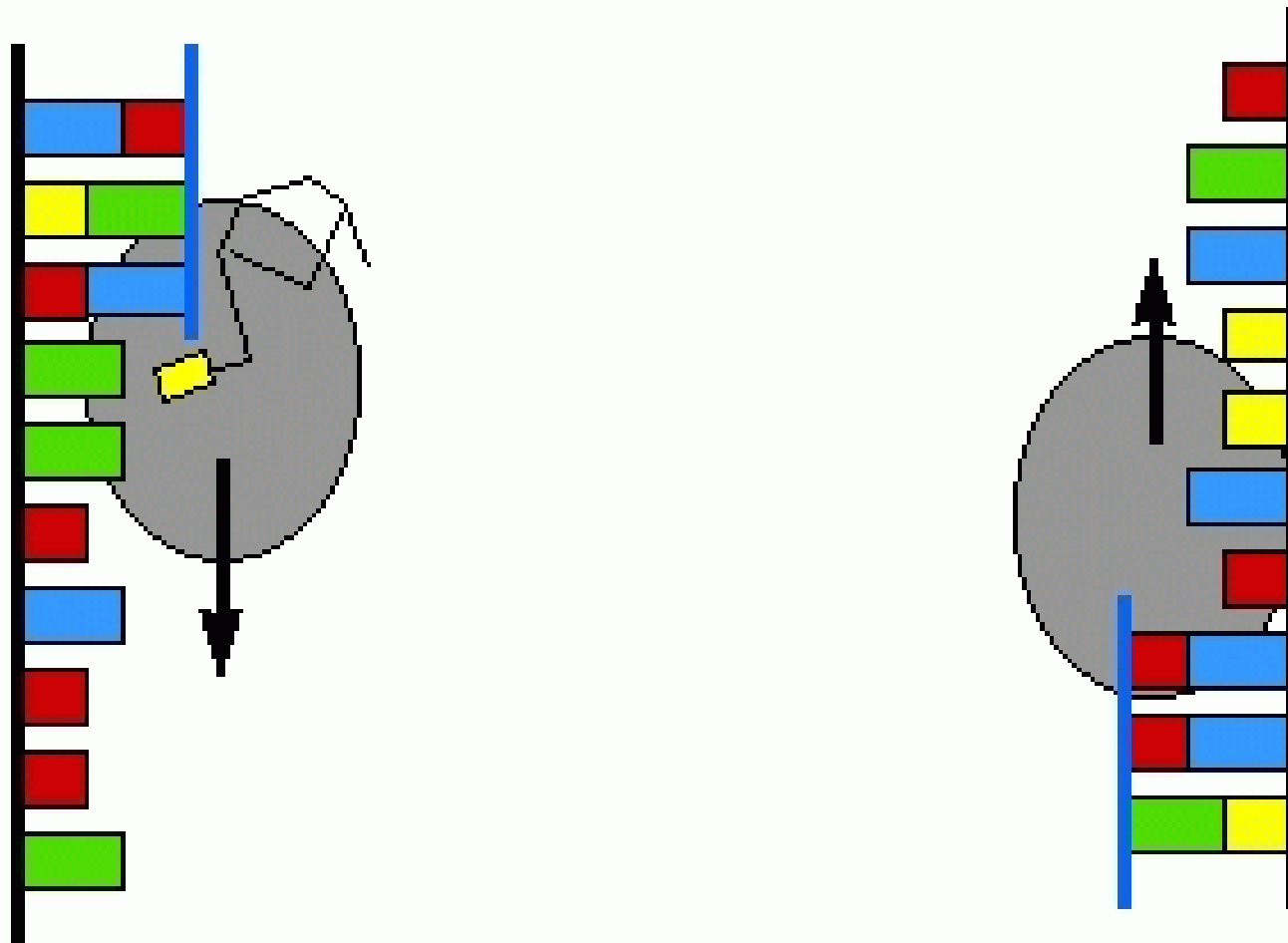
54° C

## ANNEALED PRIMERS

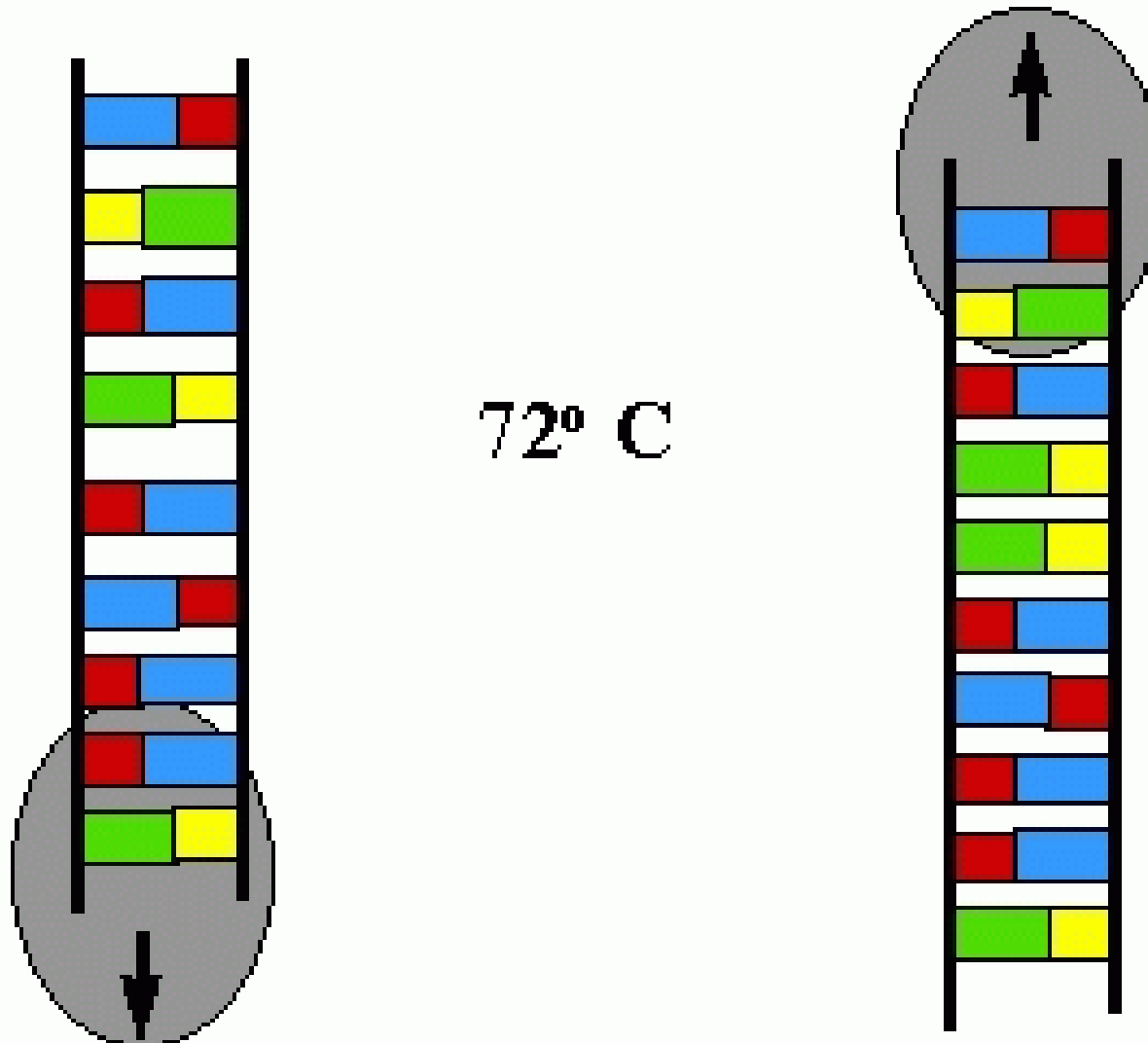


# MAKING NEW DNA MOLECULES

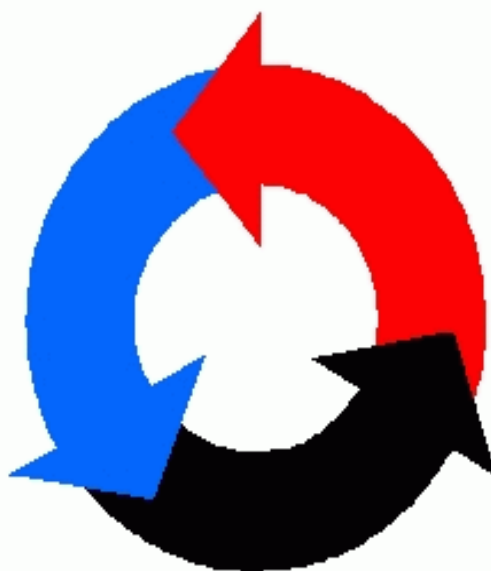
@ 72° C



# COPYING IS COMPLETED FOR EACH STRAND

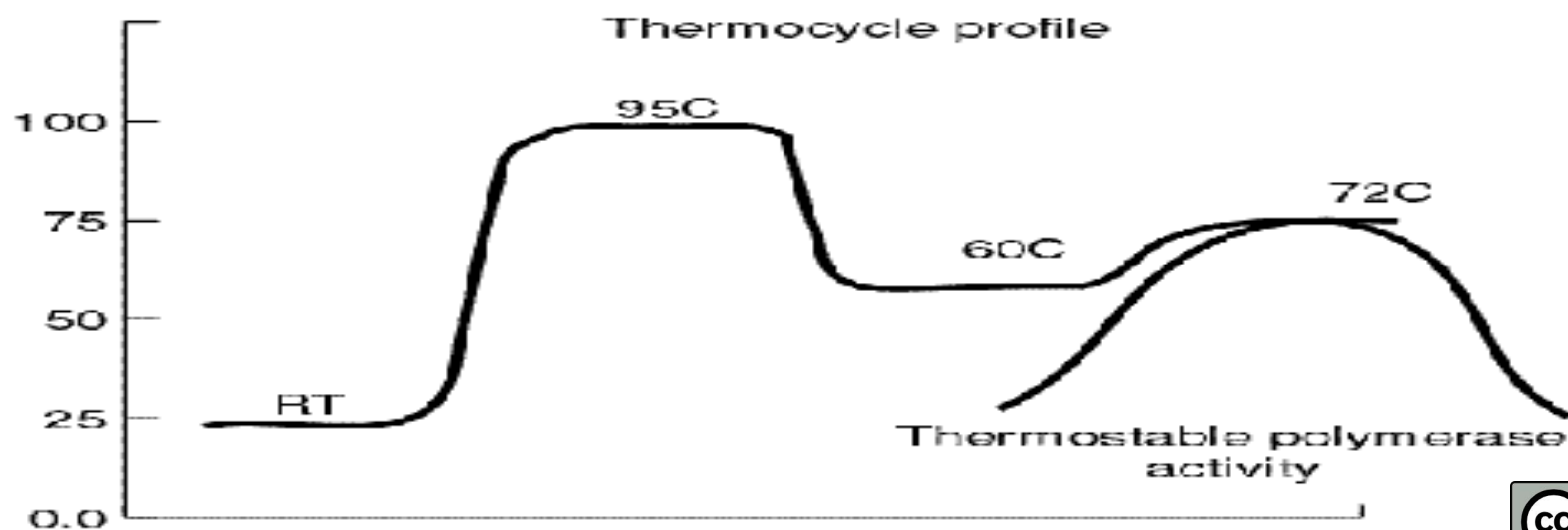


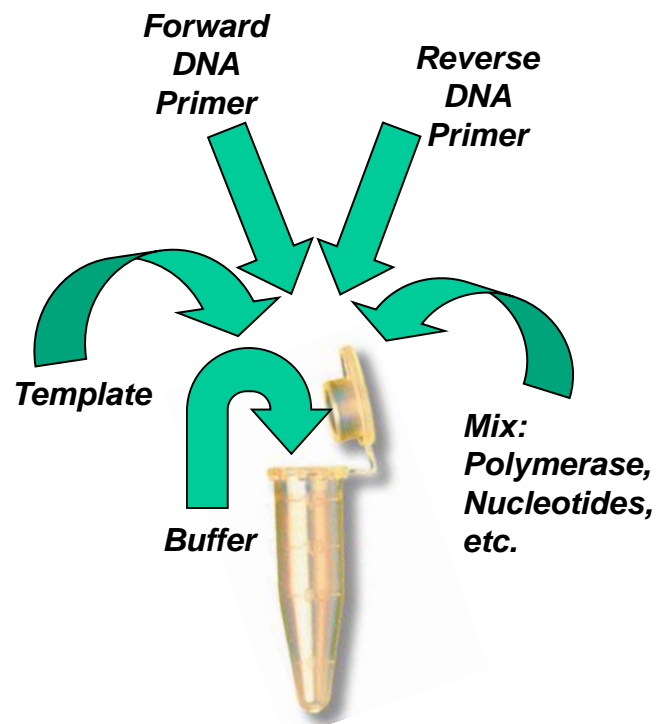
**BIND**  
55 C



**HEAT**  
94 C

**COPY**  
72 C







# CYCLES

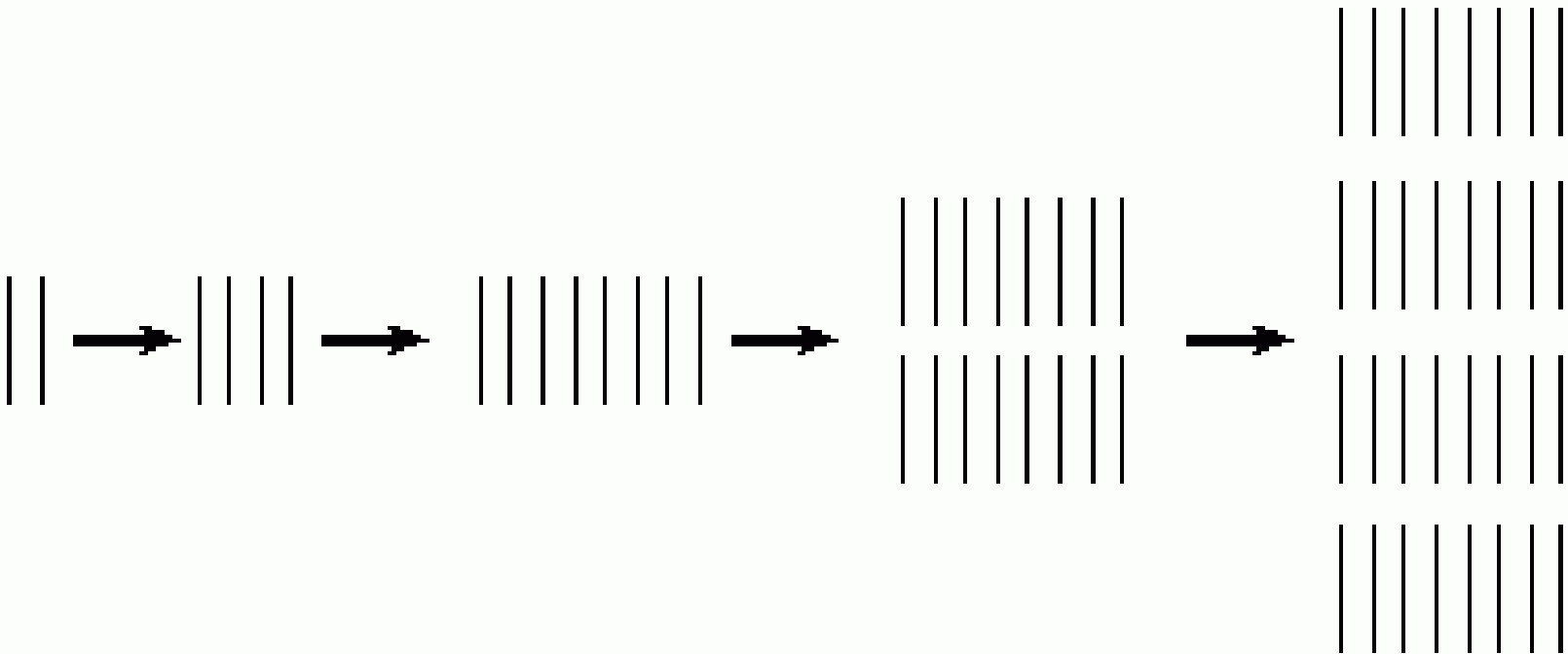
1

2

3

4

5



[animace](#)

[Length of fragments](#)



# PCR

**2 primers (~ 20 nucleotides) complementary to 3'-ends**

**amplified sequences**

**Target DNA - template**

**Thermostable DNA polymerase**

**Mixture of all 4 deoxyribonucleotides**

**Buffer with  $\text{Mg}^{2+}$  ions**

# Primers

- ~ 18 – 24 nucleotides
- no secondary structure
- balanced G/C and A/T ratio
- Not mutual complementarity
- $T_m$  (~ 55°C – 65°C) – higher  $T$  - higher specificity
- both primers ~ similar  $T_m$
- optimal concentration in reaction ~ 0,1 – 0,6  $\mu\text{M}$ ,
- higher concentration - nonspecific products
- 5'-end – allows addition of noncomplementary bases (e.g. restriction site)

$$T_m = 4^{\circ}\text{C} \times (\#G + \#C) + 2^{\circ}\text{C} \times (\#A + \#T)$$

$$T_{\text{anneal.}} = T_m - 4^{\circ}\text{C}$$

**For ~ 20 nucleotides:**

**annealing ~ 54°C**





# ***Taq* DNA (*Thermus aquaticus*)**

Optimum 75°C

Half life at 95°C about 40 min

5' → 3' polymerase activity

5' → 3' exonuclease activity

No 3' → 5' exonuclease activity

Incorporation of errors ~  $2 \times 10^{-4}$  errors/base

Processivity – up to 10 kb

terminal transferase - 3'-ends adenosin



# ***Pwo* and *Pfu* DNA polymerases**

**(*Pyrococcus woesei* and *Pyrococcus furiosus*)**

i 3' → 5' exonuclease activity

10 x higher fidelity compared to *Taq* polymerase.

lower processivity

# ***PfuTurbo* DNA polymerase**

(*Pfu* DNA polymerase and „new thermostable factor“)

Stratagene amplification of complex DNAs up to 10 kb  
vectors ~ 15 kb.

# ***Tth* DNA polymerase**

*(Thermus thermophilus)*

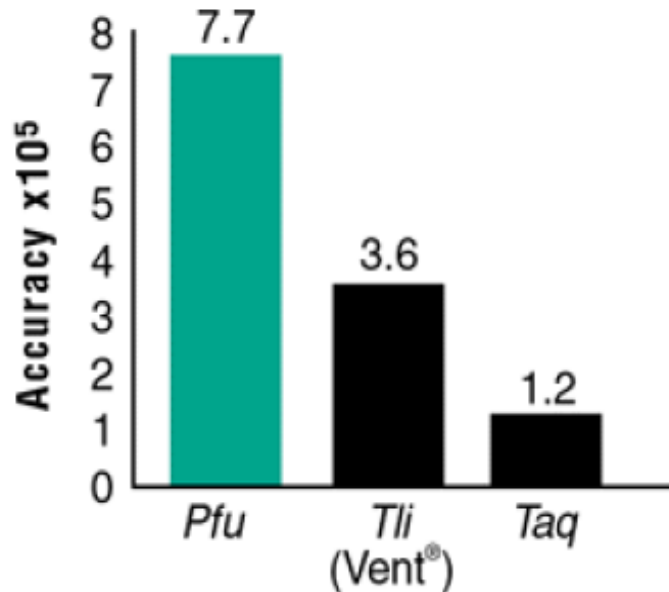
Temperature optimum as *Taq* DNA polymerase.

and reverse transcriptase - cDNA.

RT-PCR.

**Commercially available mixtures of thermostable DNA polymerases – combination of advantageous properties - processivity and accuracy.**

**Incorporation of modified bases – probes**



DNA Polymerase	Error Rate x 10 <sup>-6</sup>
Pfu	1.3 ± 0.2 SD
Deep Vent®	2.7 ± 0.2 SD
Tli (Vent®)	2.8 ± 0.9 SD
Taq	8.0 ± 3.9 SD
Ultma®	55 ± 2

# Forensic uses of PCR

- DNA Fingerprints can be amplified from
  - ✓ hair
  - ✓ blood
  - ✓ semen
  - ✓ teeth
  - ✓ skin samples
  - ✓ saliva (from cigarette butts)
- DNA fingerprints can unambiguously identify criminals and free innocent persons

# Amplification and recovery of prehistoric DNA.

- Using a mixture of short primers, it is possible to amplify whole DNA.
- In this way libraries of extinct animals can and have been constructed.
- Dodo bird DNA from feathers.
- Saber toothed DNA from bones preserved in Tar.
- Mammoth DNA from frozen tissue.

## **Healthcare**

- **Genome mapping; characterization of genes**
- **Prenatal diagnostics of genetic diseases**
- **Analysis of allelic changes**
- **Preimplantation diagnostics and typization**
- **Detection infectious microorganisms and viruses**



# Diagnostics

Infection with a pathogenic microorganism

Standard procedure:

1. Collect sample from a patient
2. Cultivate microorganisms
3. Identify the pathogen

**Time consuming; usually takes several days: delays targeted treatment**

PCR-based procedure:

1. Collect sample from a patient.
2. Isolate DNA and PCR amplify DNA using pathogen-specific primers.
3. Analyze amplified DNA

**Fast; can be performed in one day.  
Extremely sensitive**

# **CONTROL OF PRODUCTS**

**GMO**

**Microbiologic control**

**Control of identity**

# NESTED PCR

Primer 1.1

1<sup>st</sup> Round PCR



Primer 1.2

Primer 2.2

2<sup>nd</sup> Round PCR



Primer 2.2



Amplification product

# NESTED PCR

The advantages of n-PCR are:

- Its increased **specificity**  
(specific binding of 2<sup>nd</sup> primer pair).
- Increased **sensitivity**  
(2<sup>nd</sup> round of PCR amplification)

n-PCR – detection of organisms in low copy numbers

- Viruses in CSF (herpes simplex, JCV)
- Eye samples (adenovirus, herpes simplex)

# Multiplex PCR

m-PCR - detecting multiple targets in a single reaction

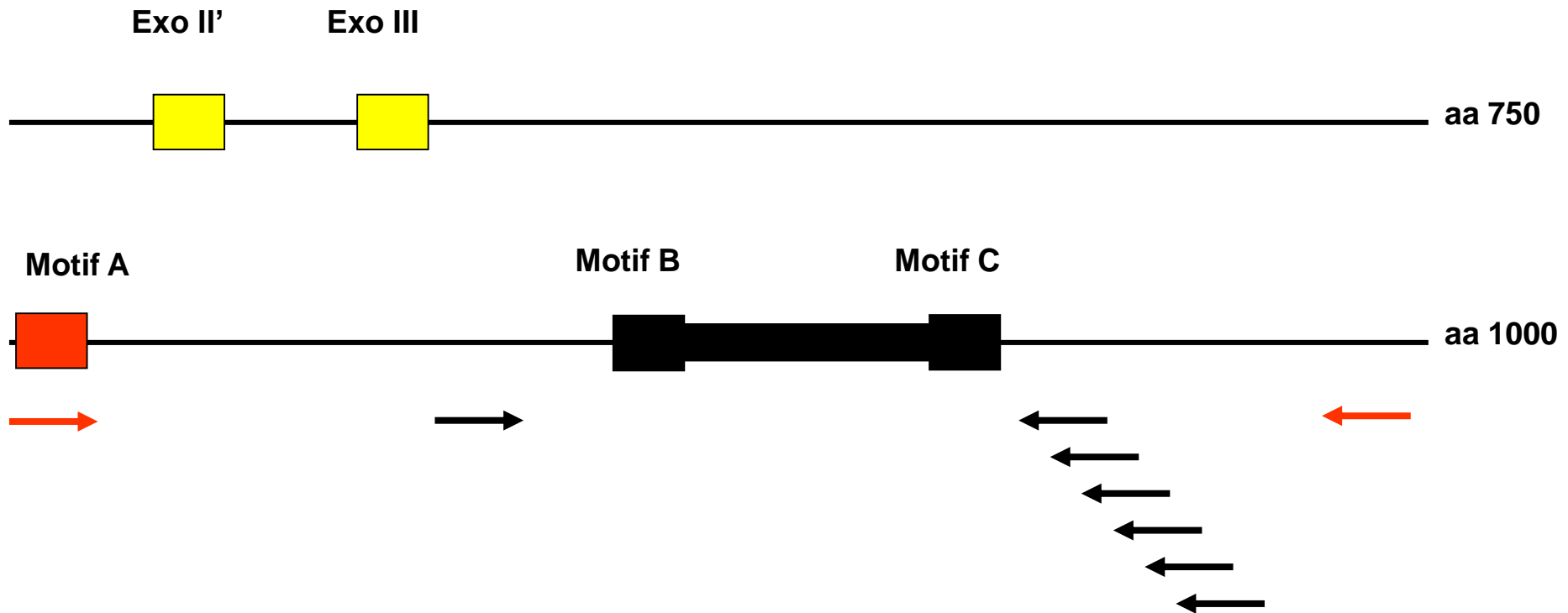
PCR reagent mix contains multiple sets of primers  
Any one of these may be amplified during the PCR

- Primer sets to multiple organisms
- Primer sets to multiple target genes in the same organism

Major advantage - reduction in test processing time

# Herpes virus Multiplex Primers

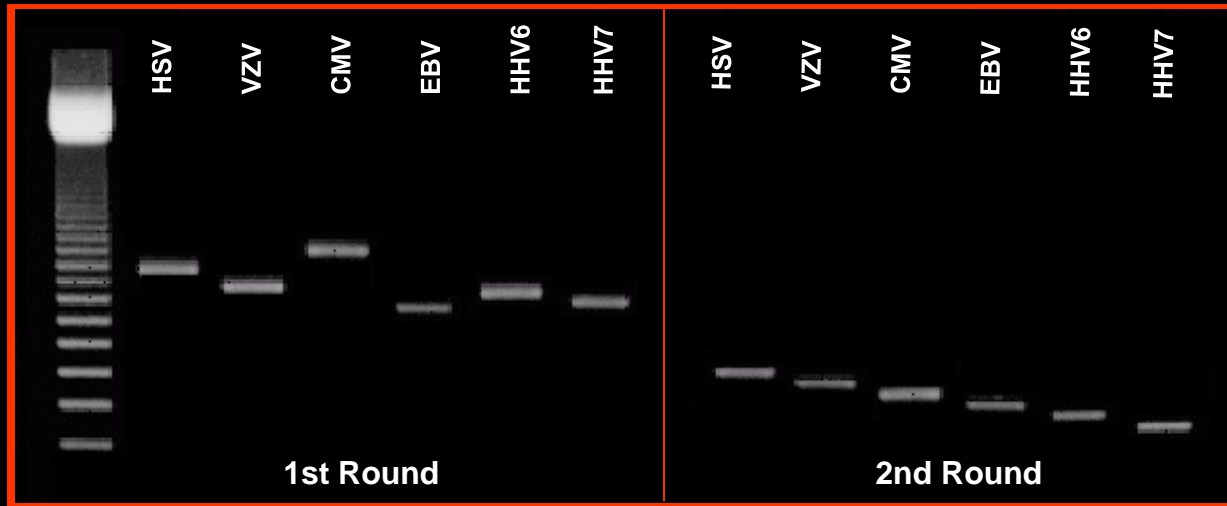
## Herpesvirus DNA polymerase gene



Reference: Heringa and Argos 1994

# HERPES MULTIPLEX PCR

Results of PCR amplification with external and internal primers

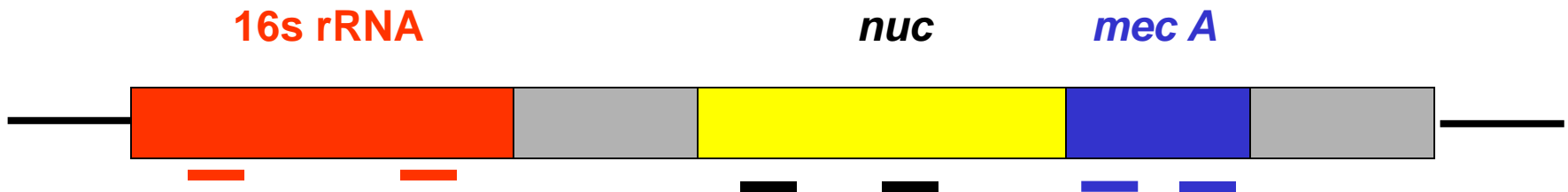


150-250 bp



# APPLICATION OF m-PCR

## *Staphylococcus aureus* Genome

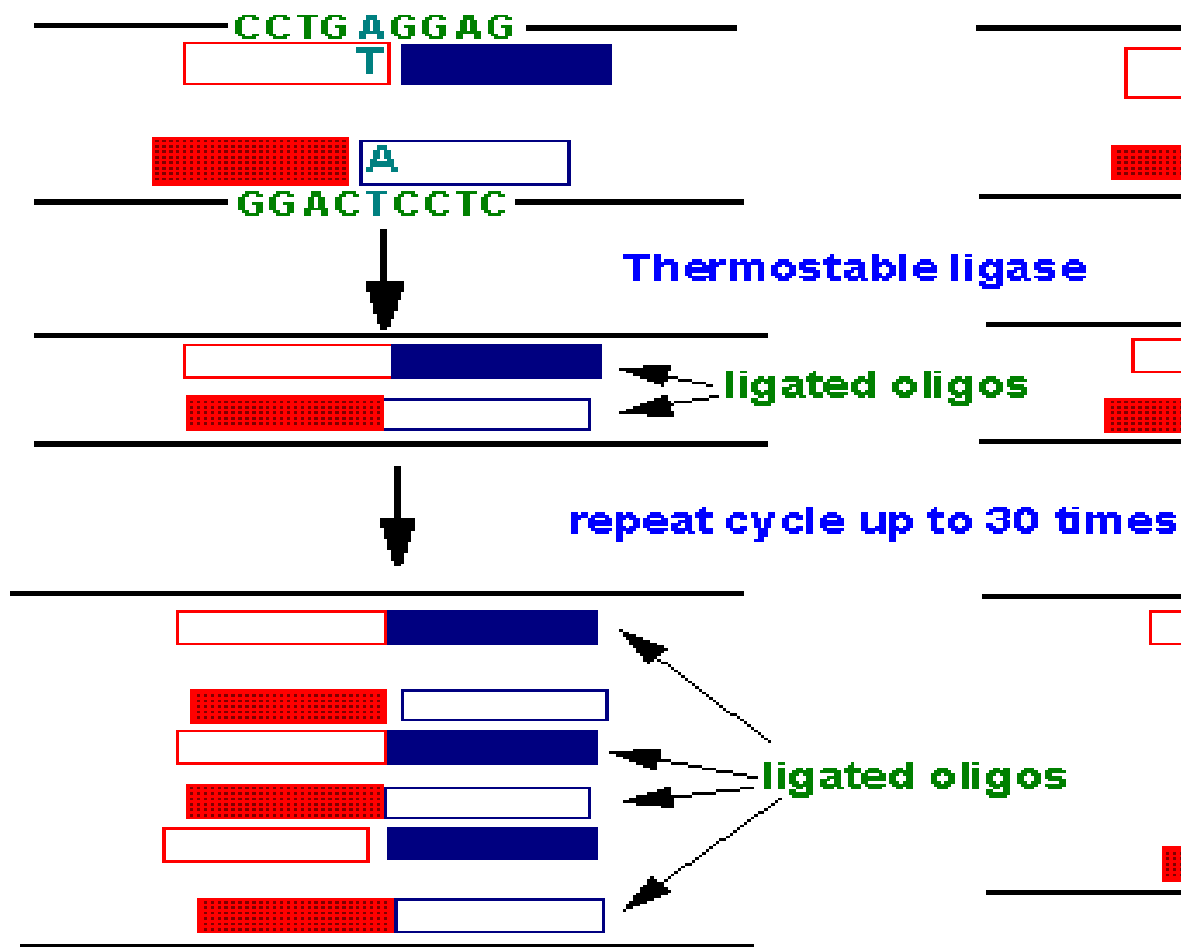


- Detection of 16s rRNA gene - common to ALL bacteria

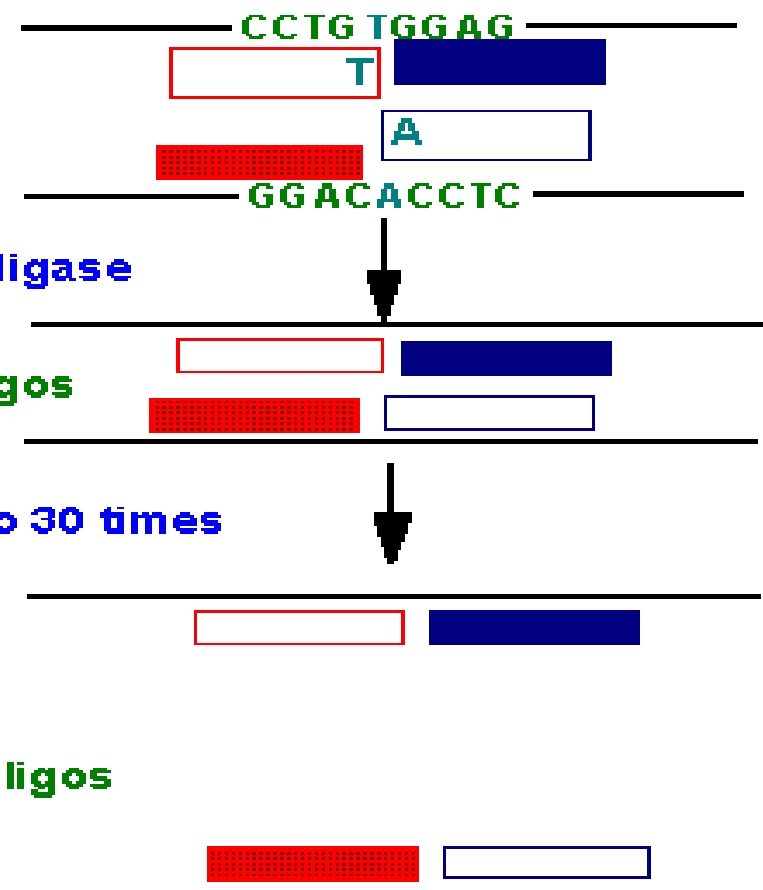
- Detection of *genus-specific* gene sequences
  - *nuc* gene is specific for ALL *Staph aureus*

- Detection of drug resistance
  - *mec A* gene confers methicillin resistance in MRSA

## Normal $\beta$ -Globin



## Sickle $\beta$ -Globin



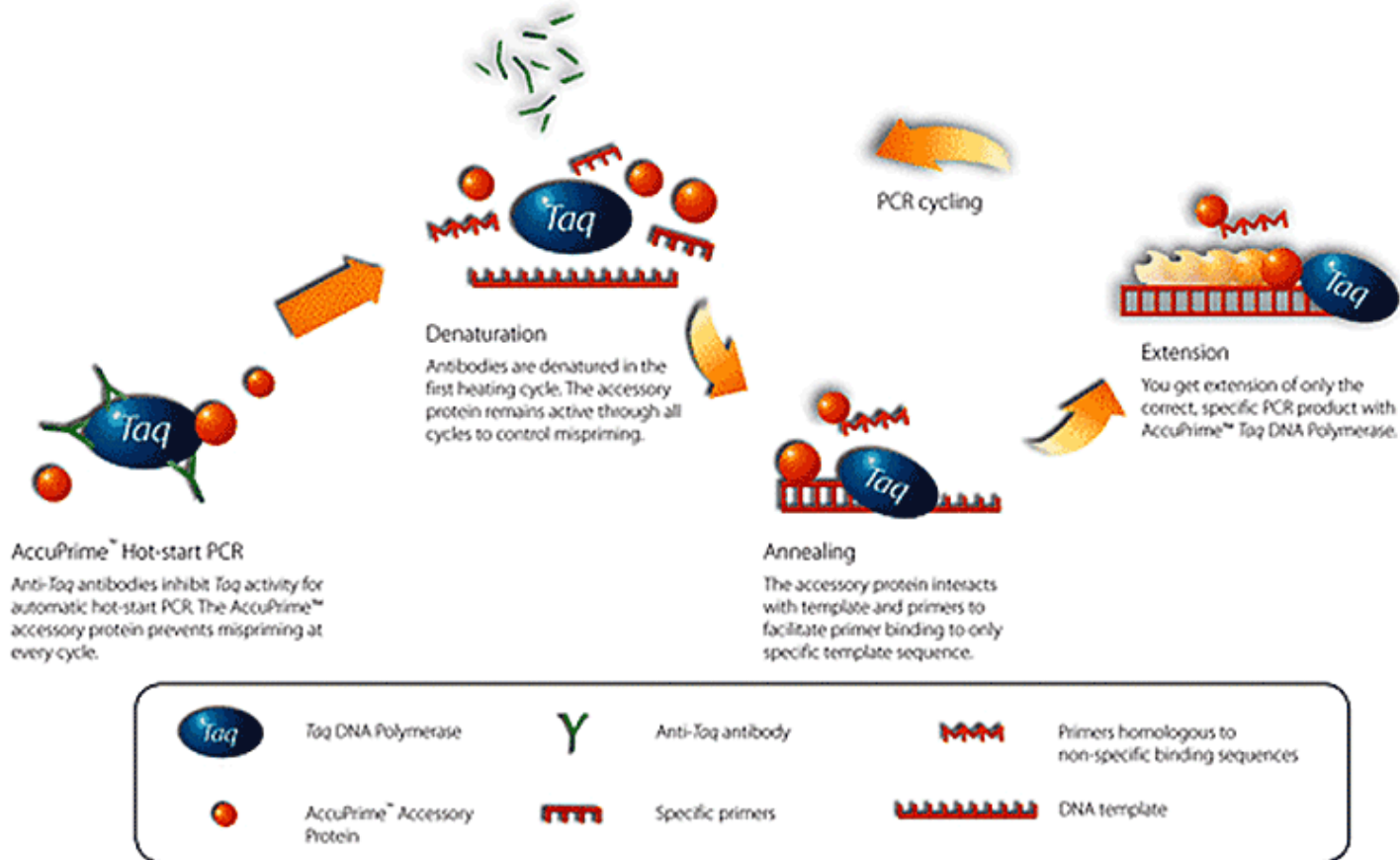
copyright M.W.King 1996

## The Ligase Chain Reaction (LCR) – detection of point mutations

specific primers to the wt

sickle-cell mutation - 3' nucleotide – mismatch

Primer does not hybridize DNA ligase will not ligate the two oligos of each pair together.



## Hot-start PCR

DNA polymerase inactivated by a chemical modification or antibody that dissociates at high temperature.

- reduces nonspecific priming and primer-dimer formation and increases product yield.
- chemical hot-start: up to a 10-minute inactivation,
- antibody hot-start - less than two minutes inactivation.

## Viral infection

### Old procedure:

1. Collect sample from a patient
2. Detect viral proteins by virus-specific antibodies

**Low sensitivity: can detect virus only at the late stages of the infection**

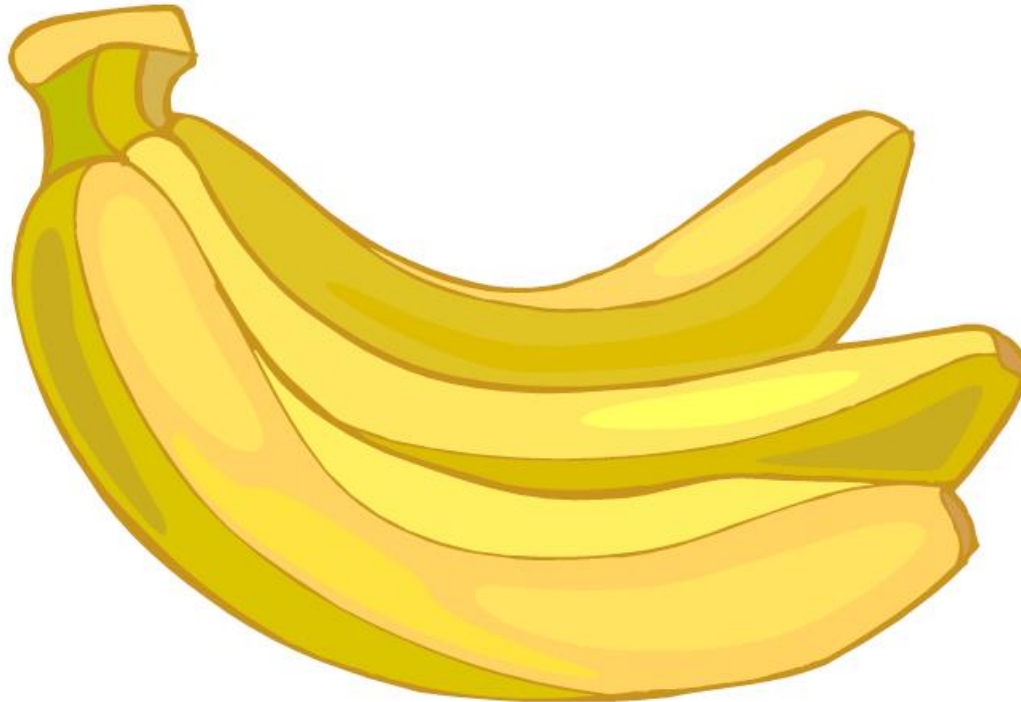
### New PCR-based procedure:

1. Collect sample from a patient.
2. Isolate DNA (or RNA) and PCR amplify using specific primers.
3. Analyze amplified DNA

**Extremely sensitive: can detect virus early after the infection**

# Forensic PCR

**Your Genetic Code Is 99.9% the Same As Your Neighbour's And....**

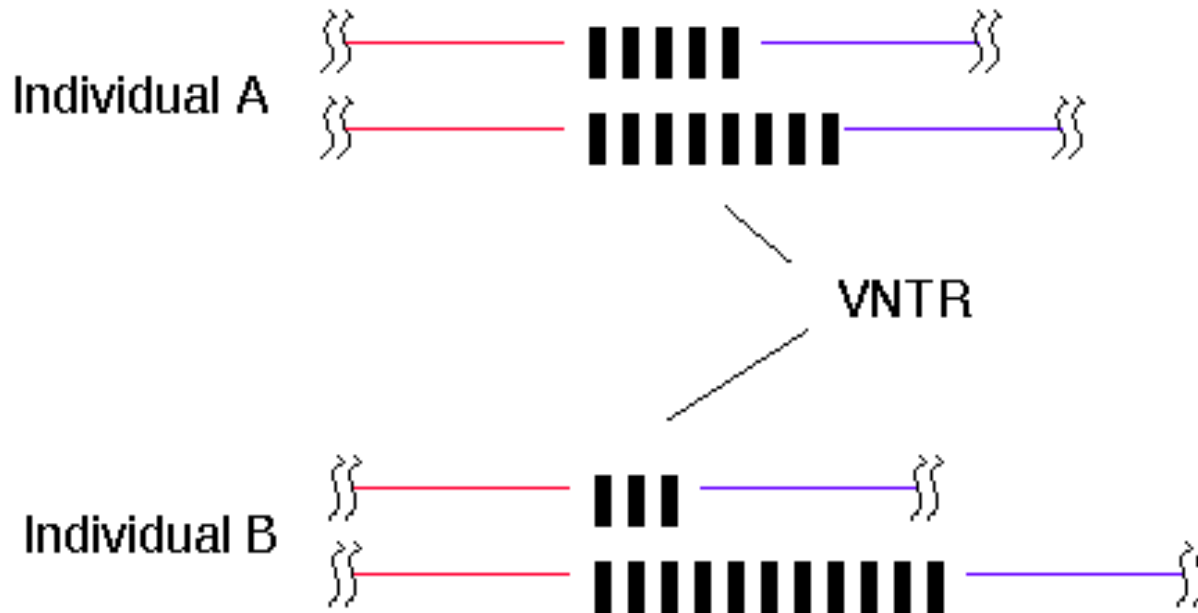


**You share half your genes with a banana!**

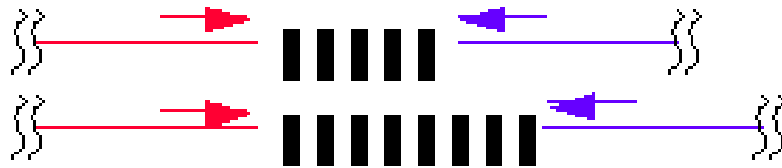
# Forensic PCR

## Human DNA contains

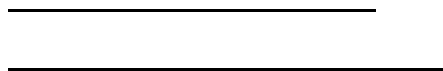
- Numerous repetitions (30-300 bp, practical use ~ 30-60 bp)
- Individual variations – VNTR (variable number of tandem repeats)
- use of VNTR – DNA fingerprinting



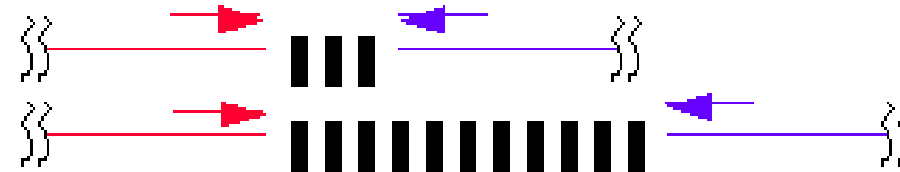
Individual A



PCR



Individual B



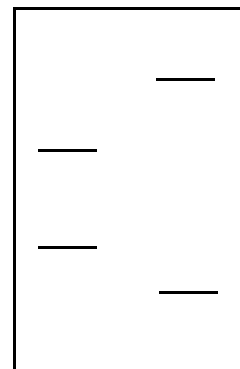
PCR



Electrophoretic  
separation of  
fragments by  
using PCR



A B

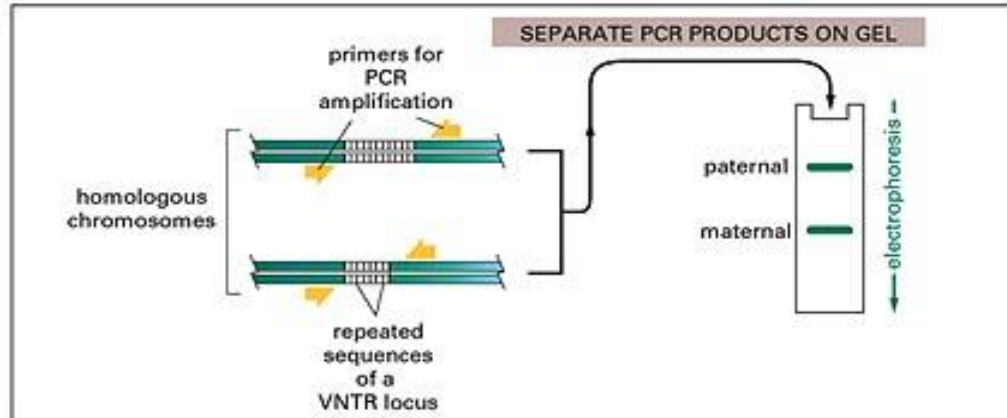


High sensitivity  
(theoretically single cell)

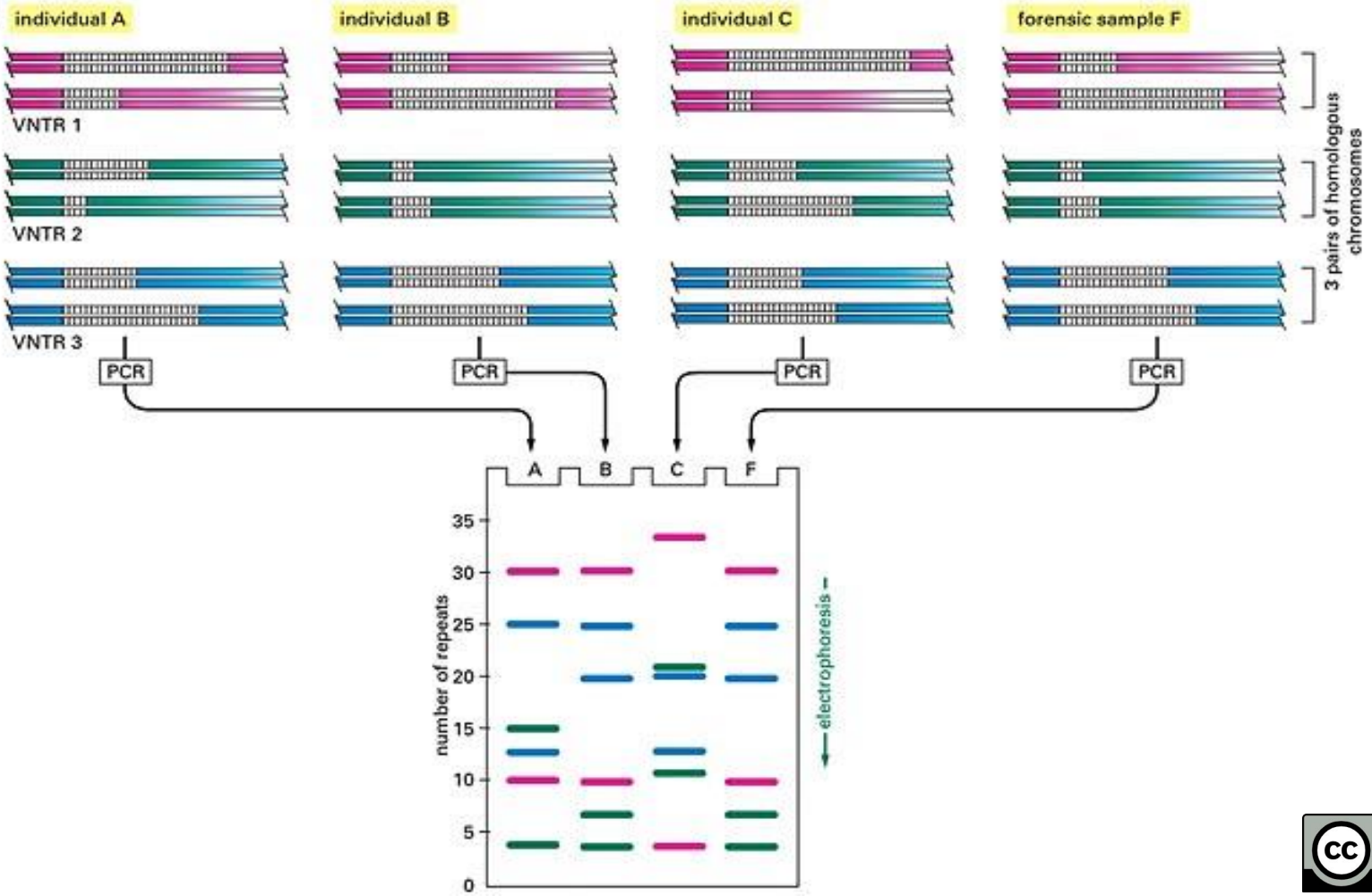
Sample – specific for each  
individual



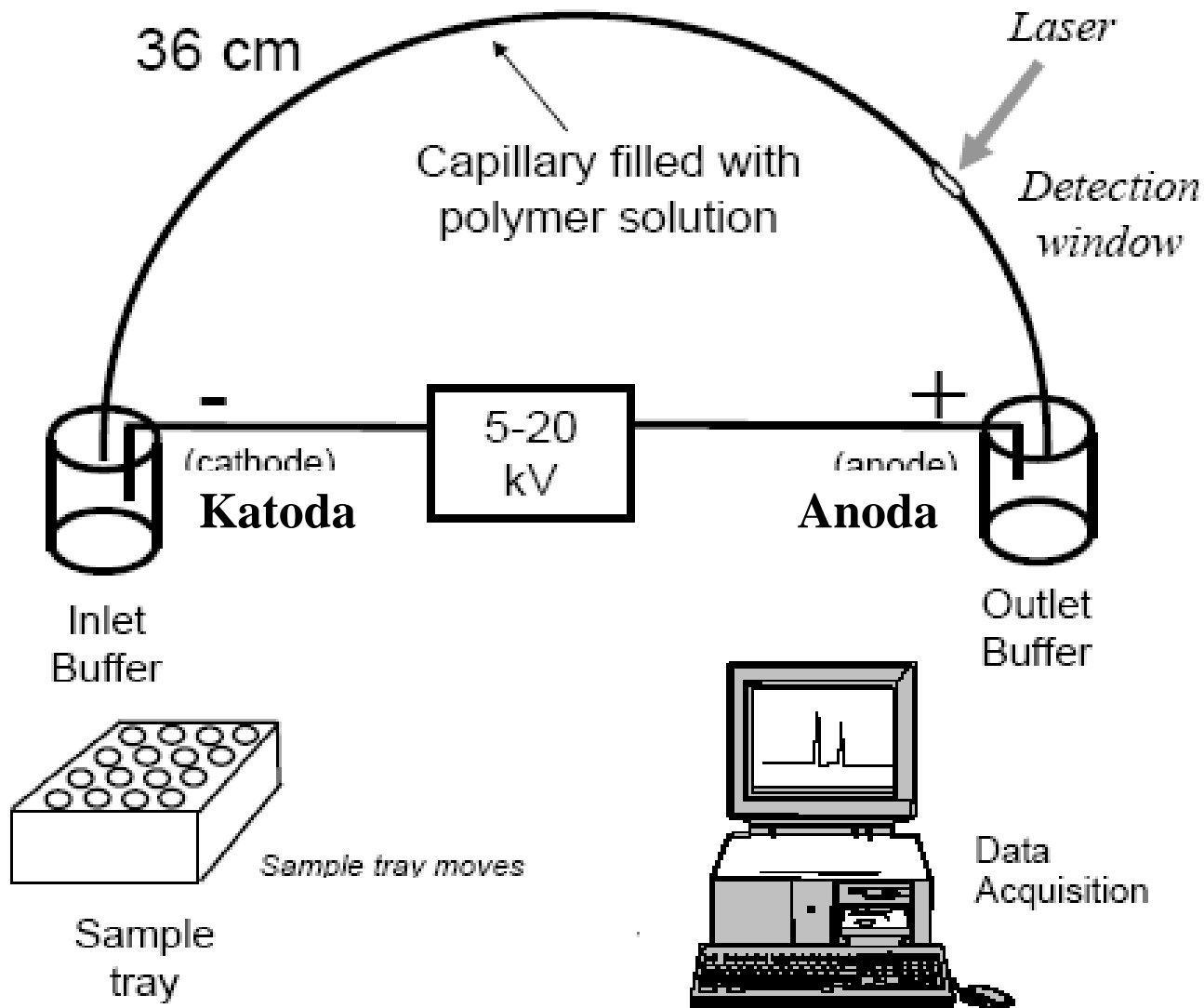
(A)



(B)

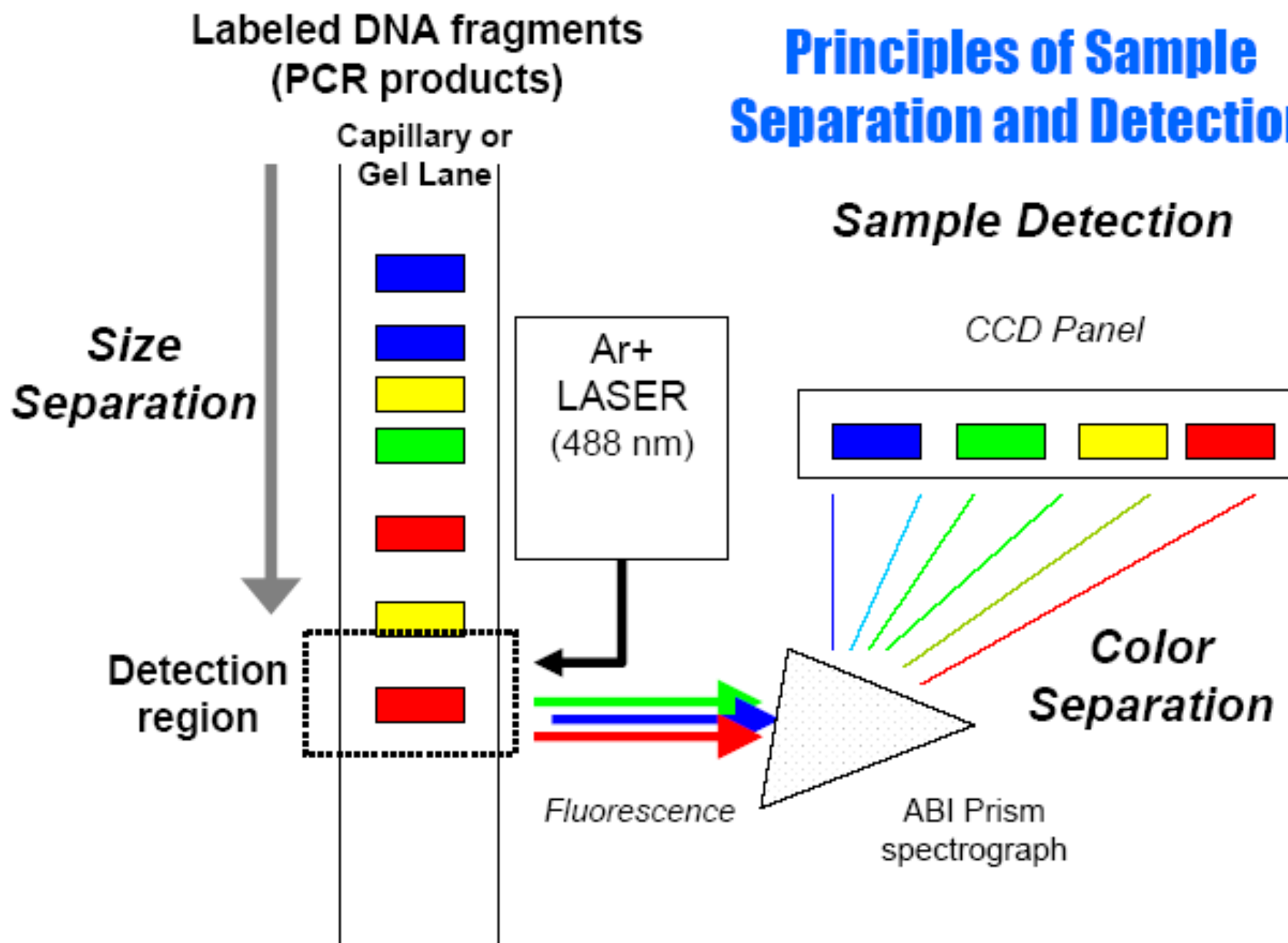


# Capillary Electrophoresis System



Butler, J.M. (2001), *Capillary Electrophoresis*, Figure 9.3, ©Academic Press

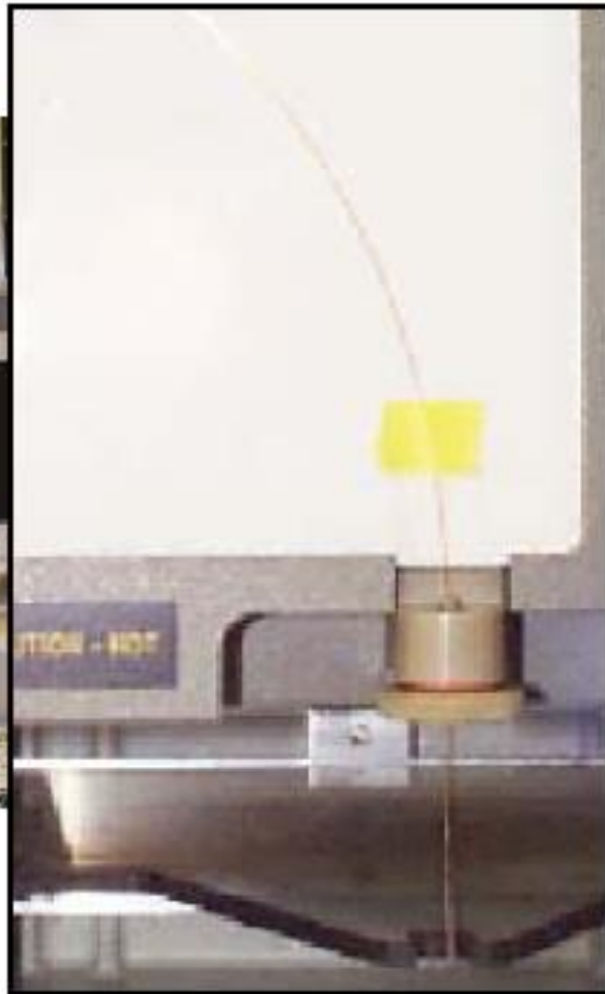
# Principles of Sample Separation and Detection



Butler, J.M. (2001) *Forensic DNA Typing*, Figure 10.8, ©Academic Press

# Capillary Electrophoresis Instrumentation

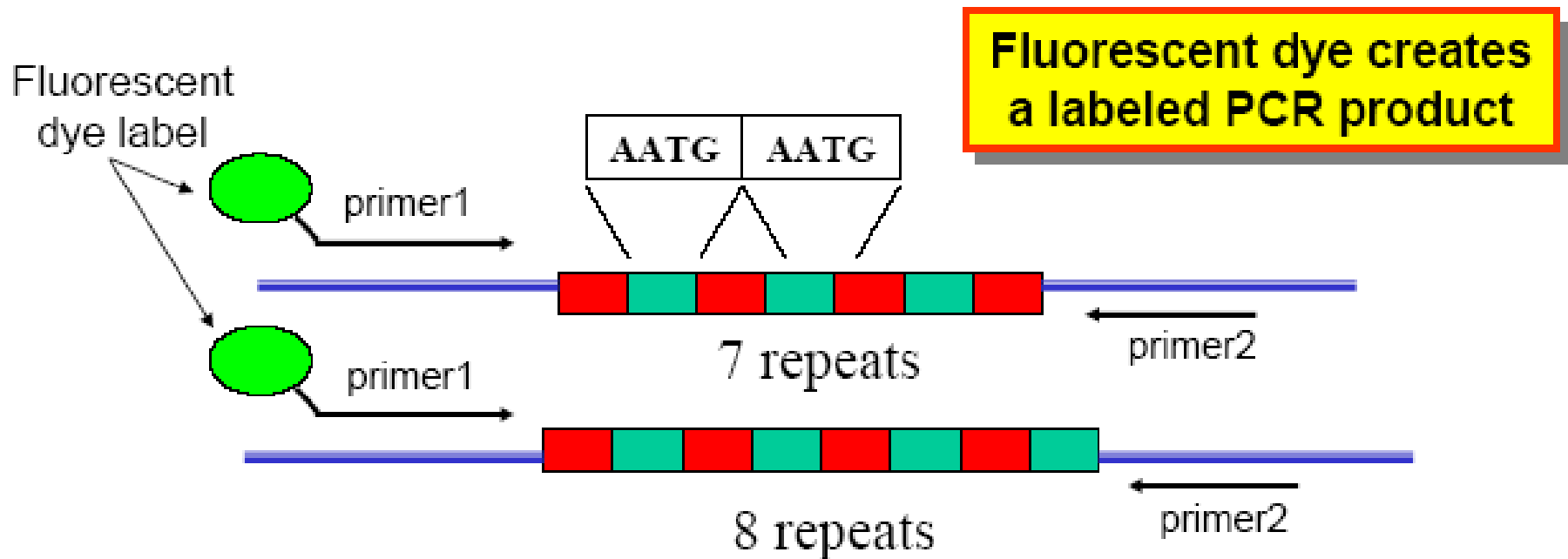
**ABI 310**  
single capillary



**ABI 3100**  
16-capillary array



# Short Tandem Repeats (STRs)



*the repeat region is variable between samples while the flanking regions where PCR primers bind are constant*

Homozygote = both alleles are the same length

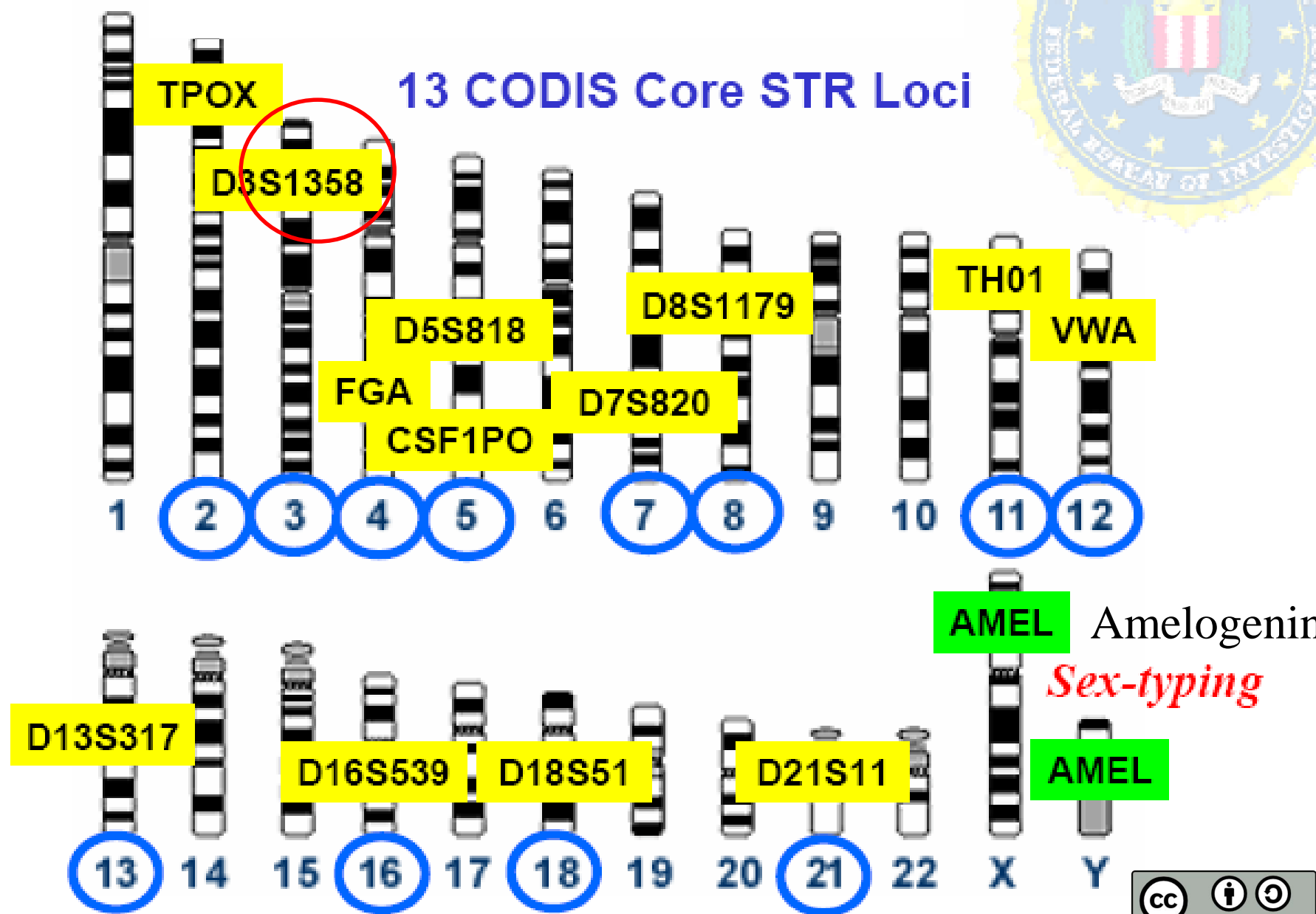
Heterozygote = alleles differ and can be resolved from one another

*Primer positions define PCR product size*

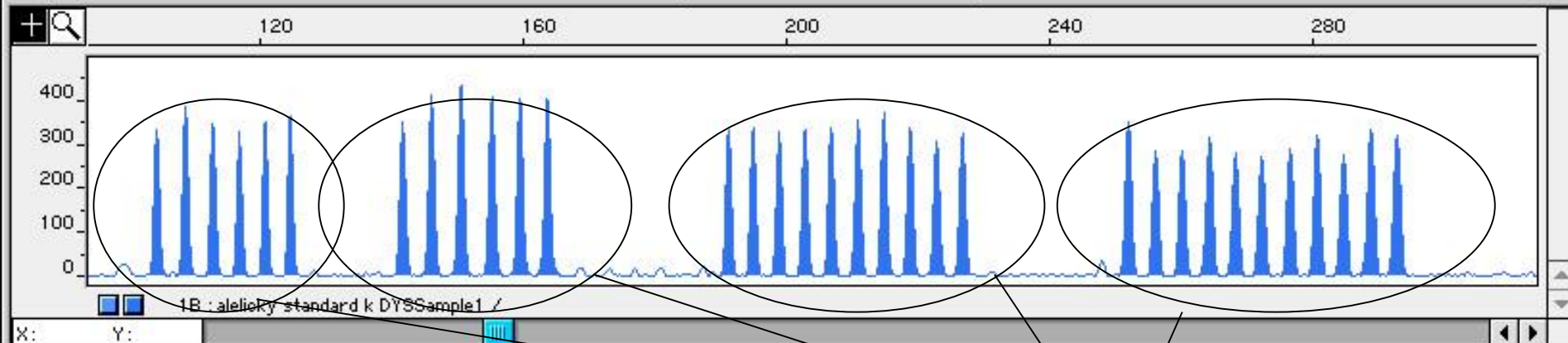
# Position of Forensic STR Markers



## 13 CODIS Core STR Loci







Dye/Sample Peak	Minutes	Size	Peak Height	Peak Area	Data Point
1B, 34	14.90	141.91	353	2236	4063
1B, 35	15.03	146.25	419	2606	4097
1B, 36	15.15	150.77	435	2851	4132
1B, 37	15.29	155.40	412	2660	4168
1B, 38	15.41	159.62	409	2672	4201
1B, 39	15.54	163.87	409	2802	4236
1B, 40	16.38	191.28	337	2185	4467
1B, 41	16.50	195.12	341	2178	4500
1B, 42	16.63	199.07	334	2211	4534
1B, 43	16.75	202.98	337	2212	4567
1B, 44	16.87	206.94	343	2276	4600
1B, 45	17.00	211.03	362	2420	4634
1B, 46	17.12	215.02	375	2522	4667
1B, 47	17.23	218.91	343	2336	4699
1B, 48	17.35	222.94	314	2128	4732
1B, 49	17.47	226.87	329	2227	4764
1B, 50	18.21	252.11	356	2511	4966
1B, 51	18.33	256.18	292	2024	4998
1B, 52	18.45	260.27	288	2079	5030
1B, 53	18.56	264.38	320	2310	5062
1B, 54	18.68	268.38	285	2023	5093
1B, 55	18.79	272.40	274	2001	5124
1B, 56	18.91	276.69	296	2131	5157
1B, 57	19.03	280.75	327	2339	5188
1B, 58	19.14	284.83	279	2070	5219
1B, 59	19.25	288.93	335	2543	5250
1B, 60	19.36	292.91	326	2440	5280

STR loci

„Unlimited“ number of STRs in genome

For analysis - suitable

4 nt STRs

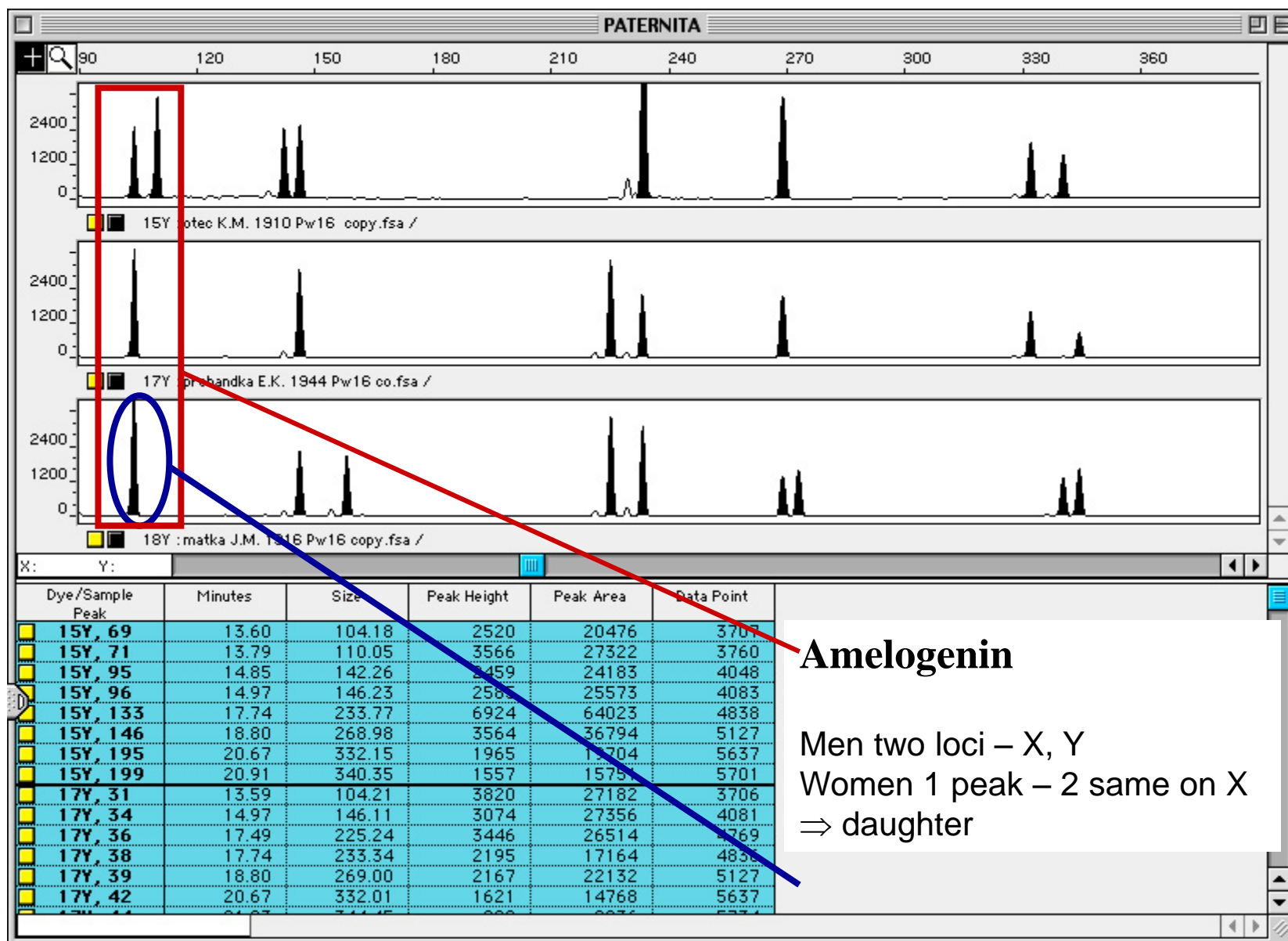
For one locus usually nine alleles

Maximally 27 alleles per locus

# Paternity – match both with father and mother







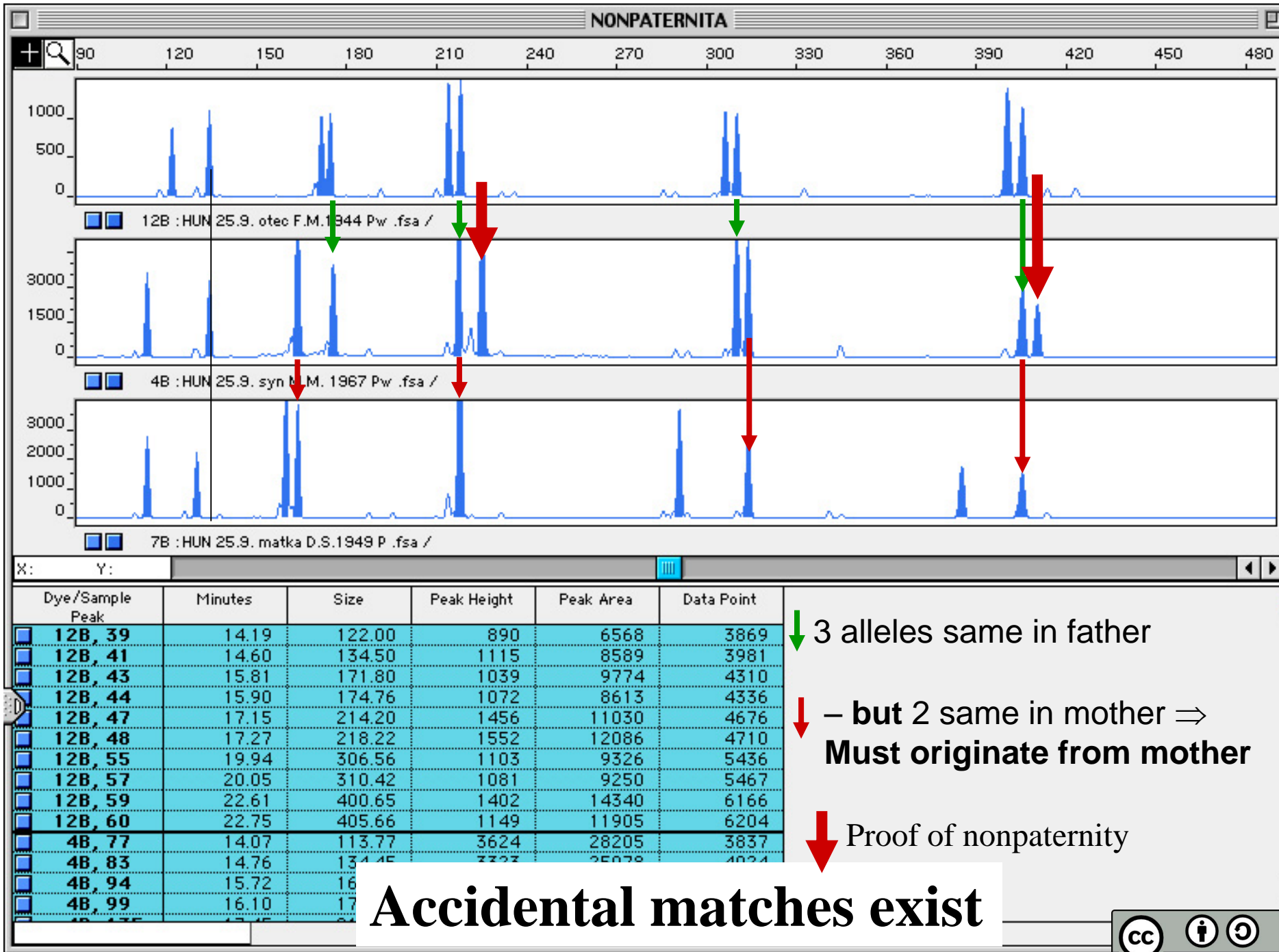
**Amelogenin**

Men two loci – X, Y

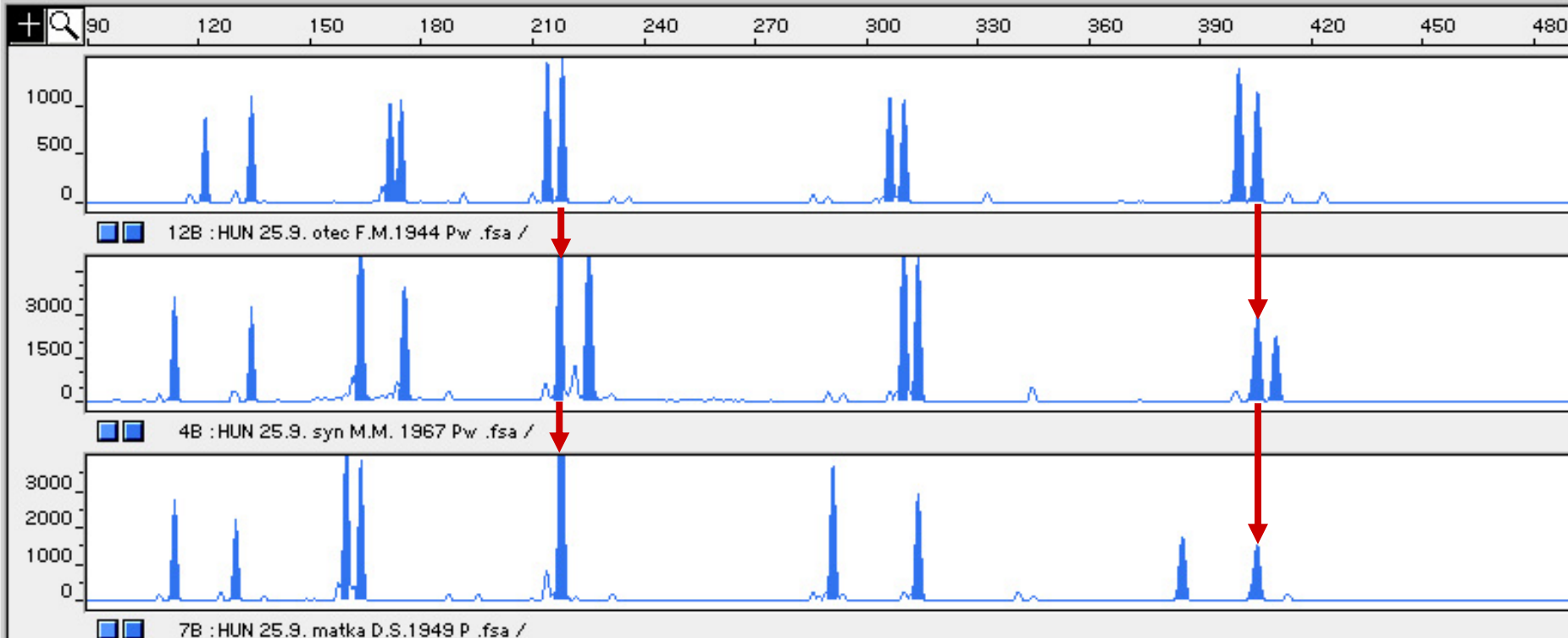
Women 1 peak – 2 same on X  
 ⇒ daughter



**↓ Father in**



# NONPATERNITA



Dye/Sample Peak	Minutes	Size	Peak Height	Peak Area	Data Point
12B, 39	14.19	122.00	890	6568	3869
12B, 41	14.60	134.50	1115	8589	3981
12B, 43	15.81	171.80	1039	9774	4310
12B, 44	15.90	174.76	1072	8613	4336
12B, 47	17.15	214.20	1456	11030	4676
12B, 48	17.27	218.22	1552	12086	4710
12B, 55	19.94	306.56	1103	9326	5436
12B, 57	20.05	310.42	1081	9250	5467
12B, 59	22.61	400.65	1402	14340	6166
12B, 60	22.75	405.66	1149	11905	6204
4B, 77	14.07	113.77	3624	28205	3837
4B, 83	14.76	134.45	3323	25078	4024
4B, 94	15.72	163.95	6866	60584	4286
4B, 99	16.10	175.81	3938	39295	4391

Both alleles same with father

– however same also with mother  
 ⇒ **inherited from mother**

# Prenatal diagnostics genetic testing

## $\alpha$ -fetoprotein

Formed in liver of embryo

Present in amniotic fluid and mother blood

High level indicates opened neuronal canal of the fetus or twins or older fetus than expected

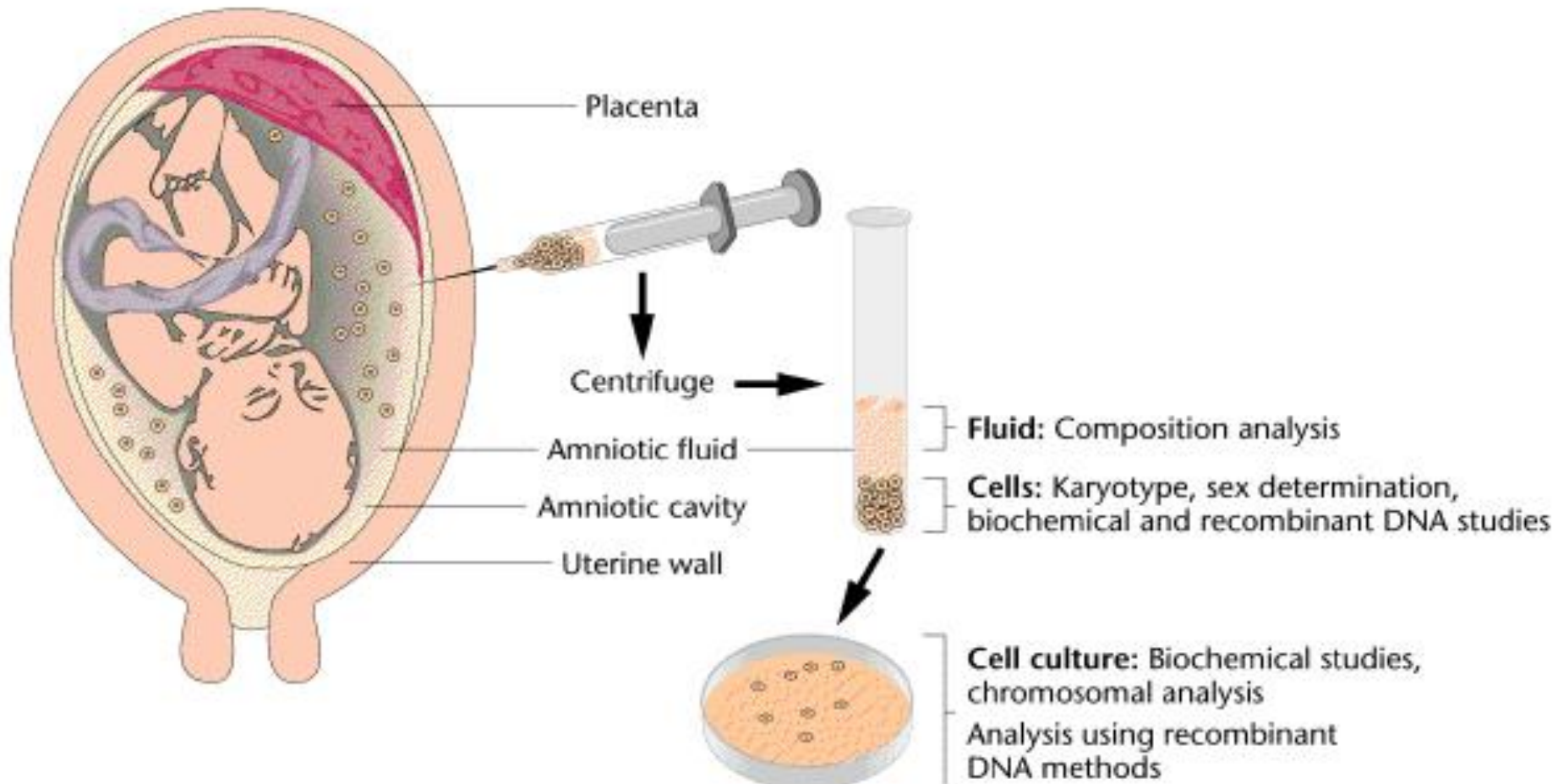
Very low level may indicate chromosomal abnormality e.g. Down syndrome

# Triple testing

- Analyses performed in elderly women:
- Besides alfa-fetoprotein (AFP)
  - Nonconjugated estriol (uE3)
  - Human beta-chorionic gonadotropin - (b-HCG)



# Testing of genetic disorders amniocentesis



# Down syndrome

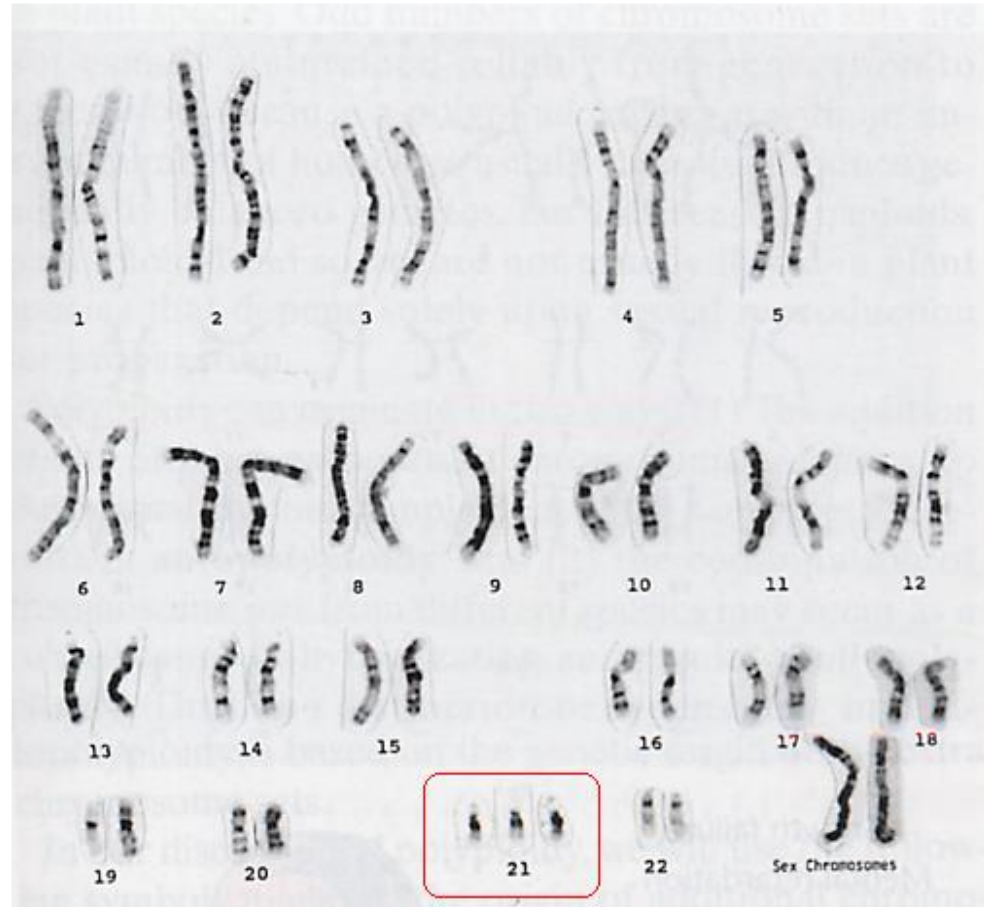
## 3 copies of chromosome 21

~ 3 from 2000 newborns

The risk of Down syndrome increases with mother's age

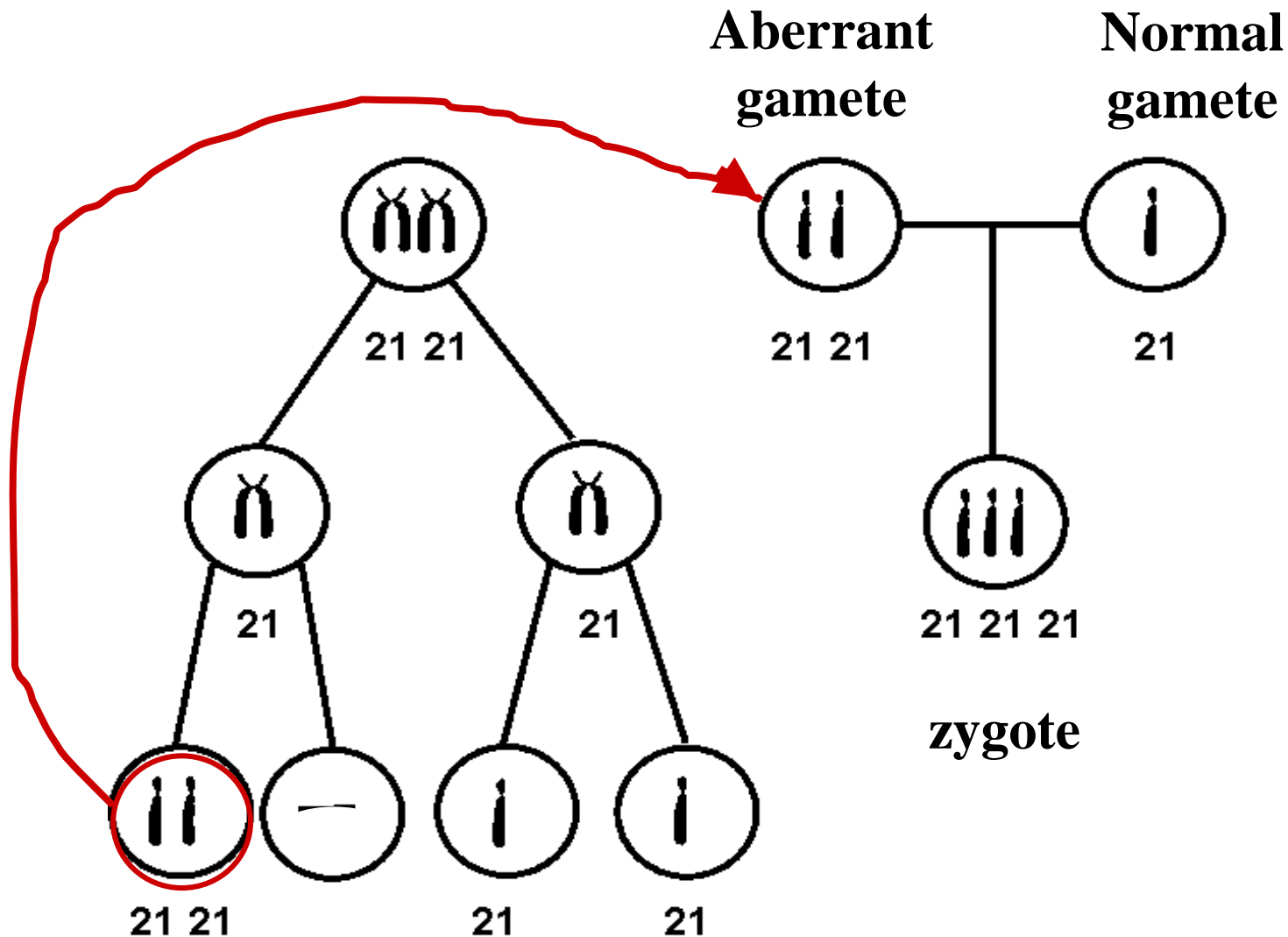
~ 1/200 mothers over 35

~ 1/50 mothers over 50

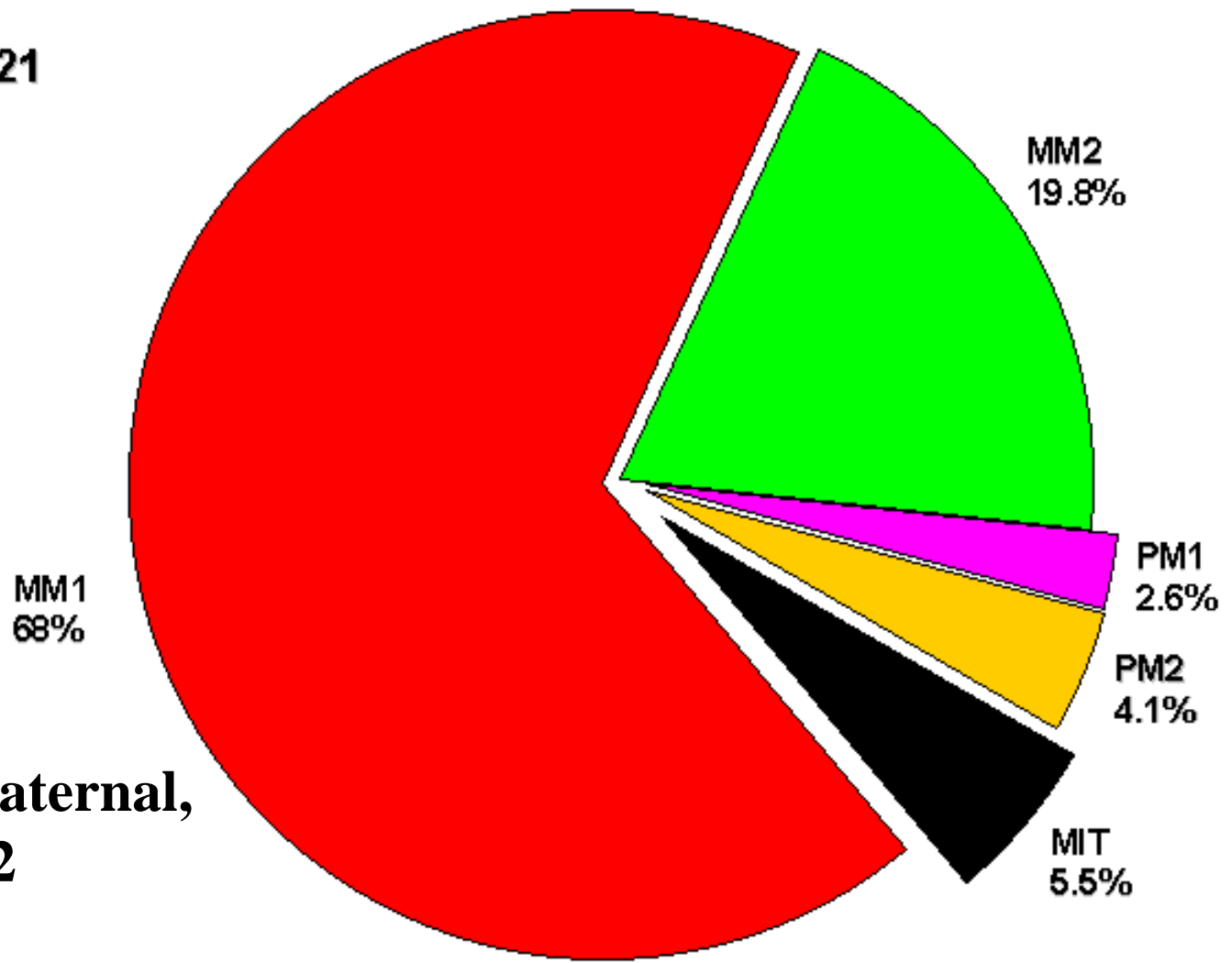




## Down syndrome (simple trisomy) - nondisjunction in meiosis II



**Origin of the extra  
chromosome 21  
in 510 families  
with free trisomy 21**

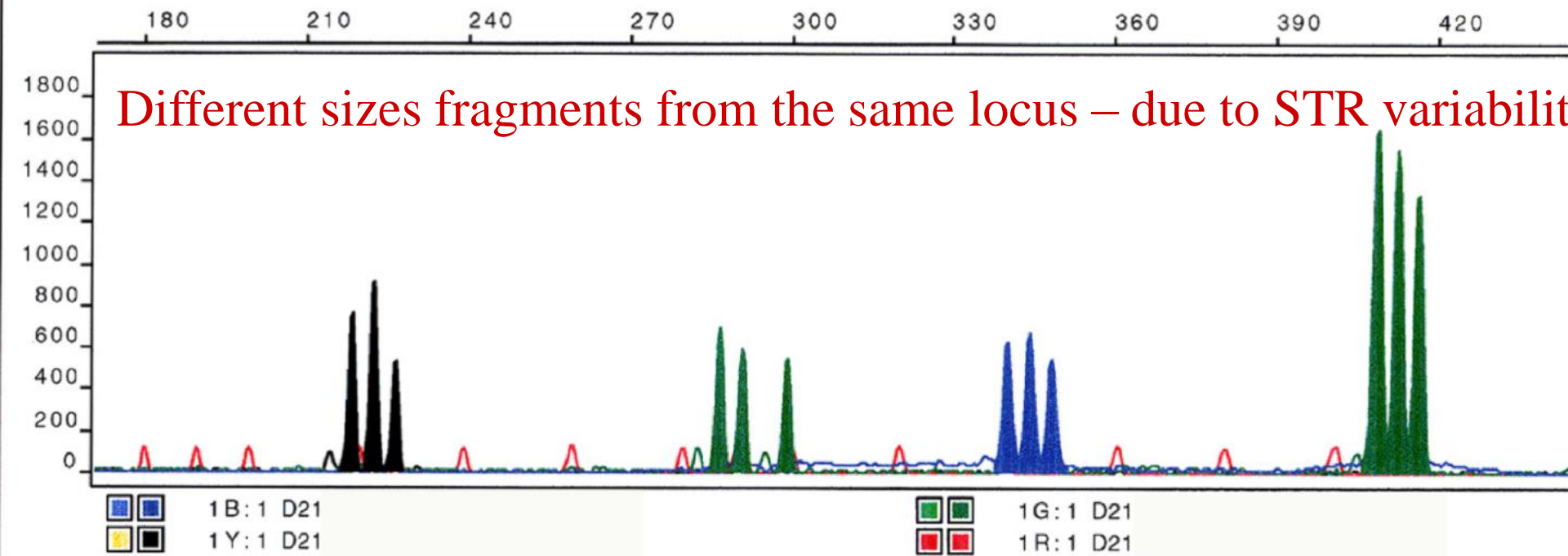


**M – maternal P-paternal,  
M1, 2: meiosis 1, 2  
Mit: mitosis**

Data from the Antonarakis and Hassold laboratories

se33109

# Down - 3 chromosomes detected in each locus of chromosome 2



Dye/Sample Peak	Minutes	Size	Peak Height	Peak Area	Data Point
1B, 10	20.33	340.13	631	8388	5544
1B, 11	20.44	344.14	672	8610	5574
1B, 12	20.55	348.15	555	7380	5604
1G, 19	18.88	286.87	701	7141	5147
1G, 20	18.99	290.92	602	6091	5178
1G, 22	19.22	299.21	549	5717	5241
1G, 26	22.13		1657	19677	6034
1G, 27	22.23		1556	18278	6061
1G, 28	22.33		1345	15609	6088
1Y, 17	16.93	218.77	767	7639	4616
1Y, 18	17.05	222.85	929	9398	4648
1Y, 19	17.16	227.02	537	5236	4680



# Unusual case: accidental match - the same STRs

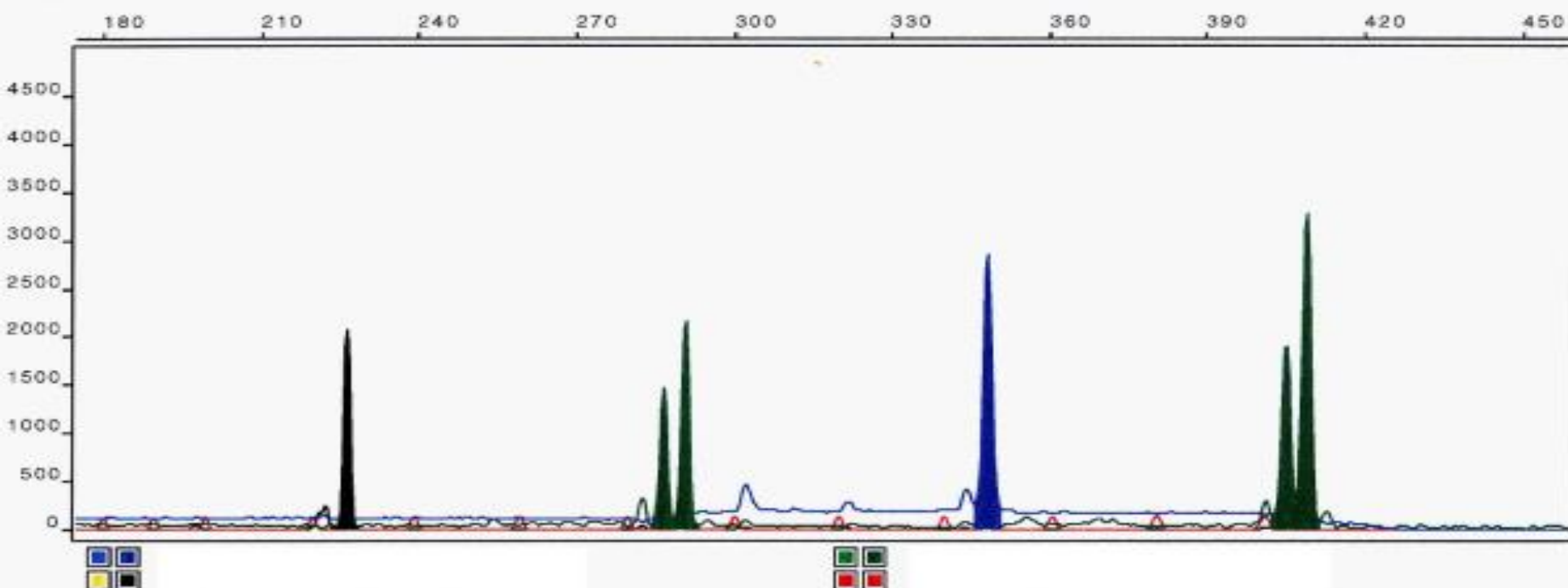
**Down - 3 chromosomes**

– ratio 2:1 in two alleles tested

- Two alleles are the same  $\Rightarrow$  more alleles need to be tested



Page 1 of 1



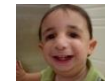
Dye/Sample Peak	Minutes	Size	Peak Height	Peak Area	Data Point
11B, 221	20.31	347.98	2858	37903	5537
11G, 41	18.66	286.93	1483	15985	5089
11G, 42	18.77	290.92	2193	23347	5119
11G, 49	21.77		1911	23650	5935
11G, 50	21.86		3311	39323	5961
11Y, 17	16.98	226.75	2079	18813	4629

# Trisomy

21 (Down syndrome)

18 (Edwards syndrome)

13 (Patau syndrome)



9

8 (Warkany syndrome 2)

22

Trisomy 21 and 18 most frequent. Rarely trisomy 13 fetus survives. Other trisomy - only in the case of mosaicism or trisomy of chromosome fragment

Trisomy of sex chromosomes can also occur and include:

XXX (Triple X syndrome)

XXY (Klinefelter syndrome)



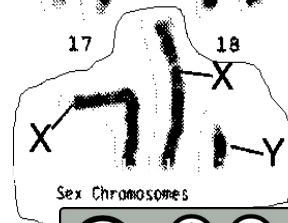
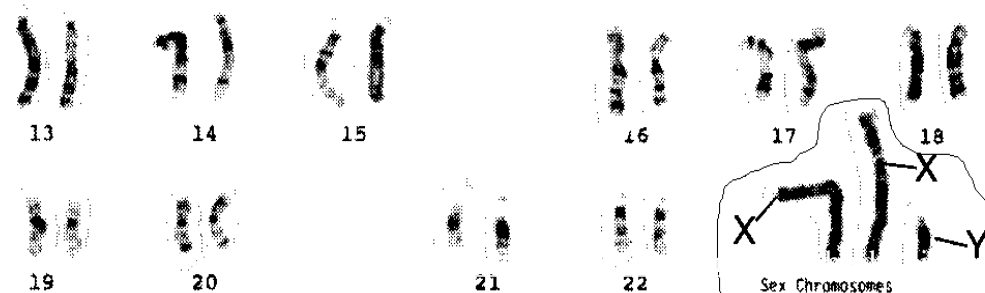
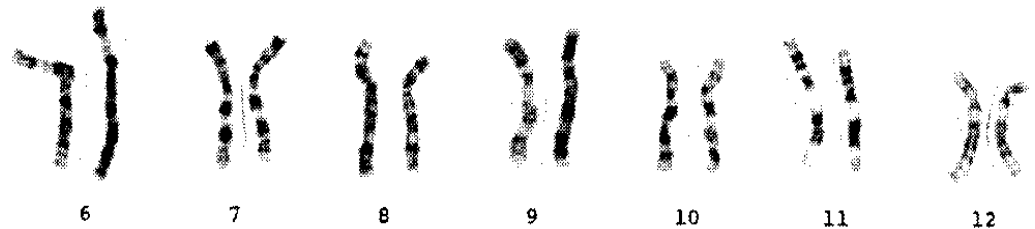
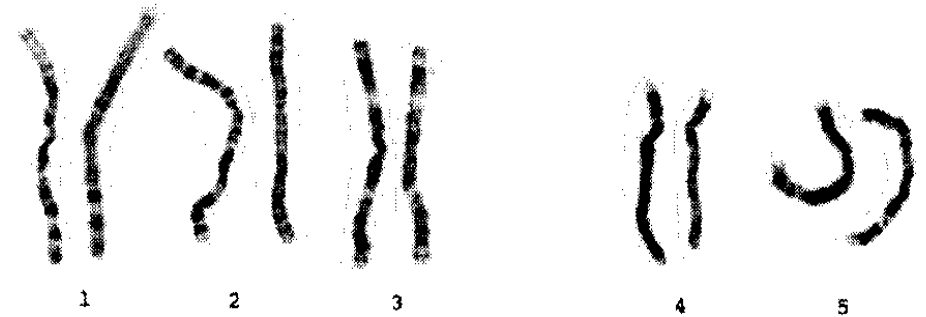
XYY



# Klinefelter syndrome

One extra X chromosome in males

(2 X chromosomes and 1 Y chromosome)

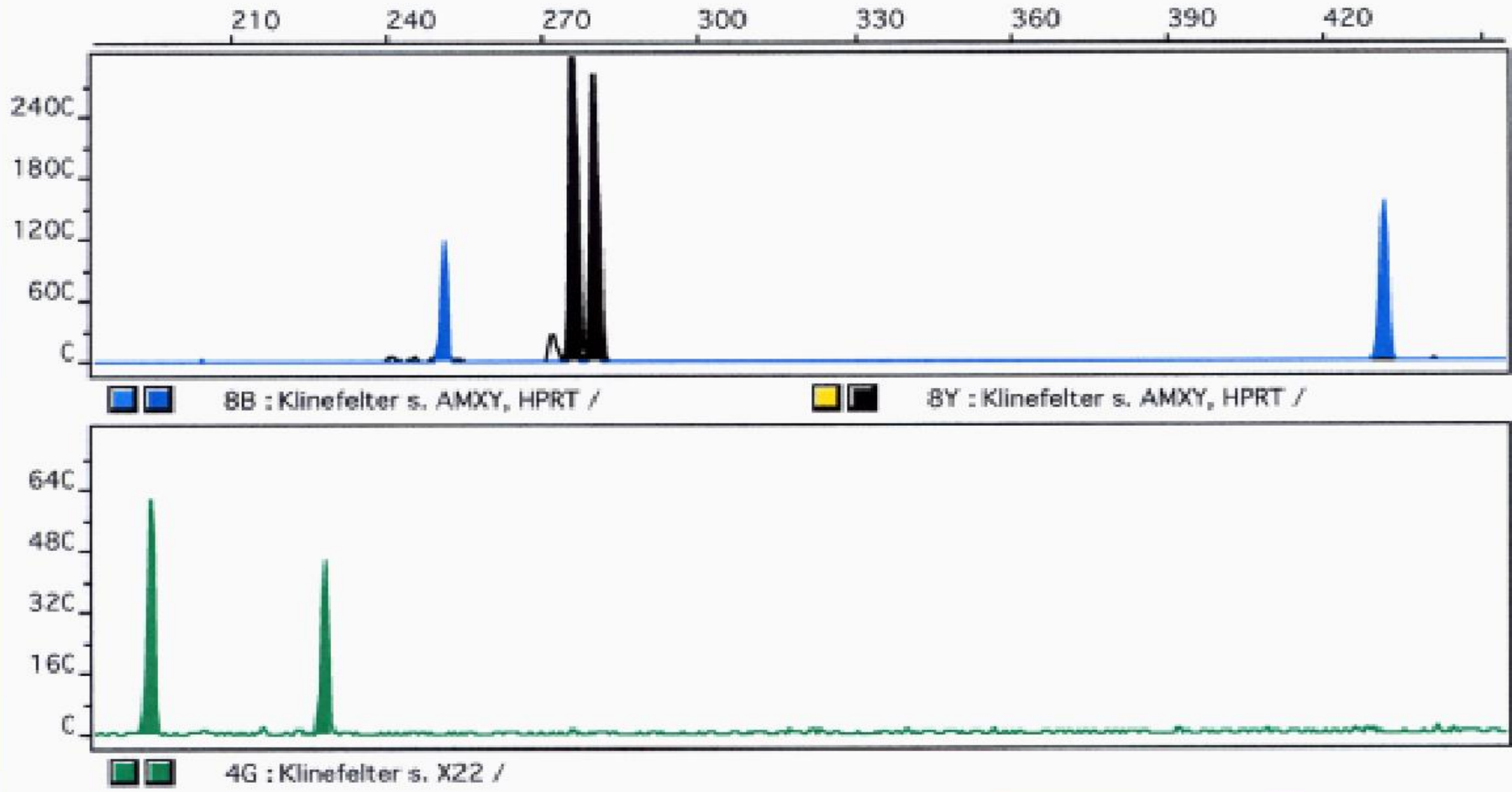


Genotype XXY – insufficient development of male secondary sex characteristics, infertility

Error during meiosis in mother

⇒ 2 X chromosomes instead of one





## Klinefelter XXY

Blue - gender: amelogenin **not STR** (is both on X and Y)  $\Rightarrow$  X+Y  
 the other on the X chromosome - 2 peaks  $\Rightarrow$  2X chromosomes  
 (locus differs in STR numbers)





# GMO detection

**regulatory sequences for gene expression in plants (transgenes)**

- **cowliflower virus promotor P35S**
- **terminator Tnos (for termination of transcription of nopaline synthase from *Agrobacterium tumefaciens*)**

# Transgenes

**resistance to herbicides (gene for fosfinotricine acetyltransferase from *Streptomyces hygroscopicus*)**

**resistance to insecticides (e.g. gene for endotoxine cryIA(b) from *Bacillus thuringiensis*)**

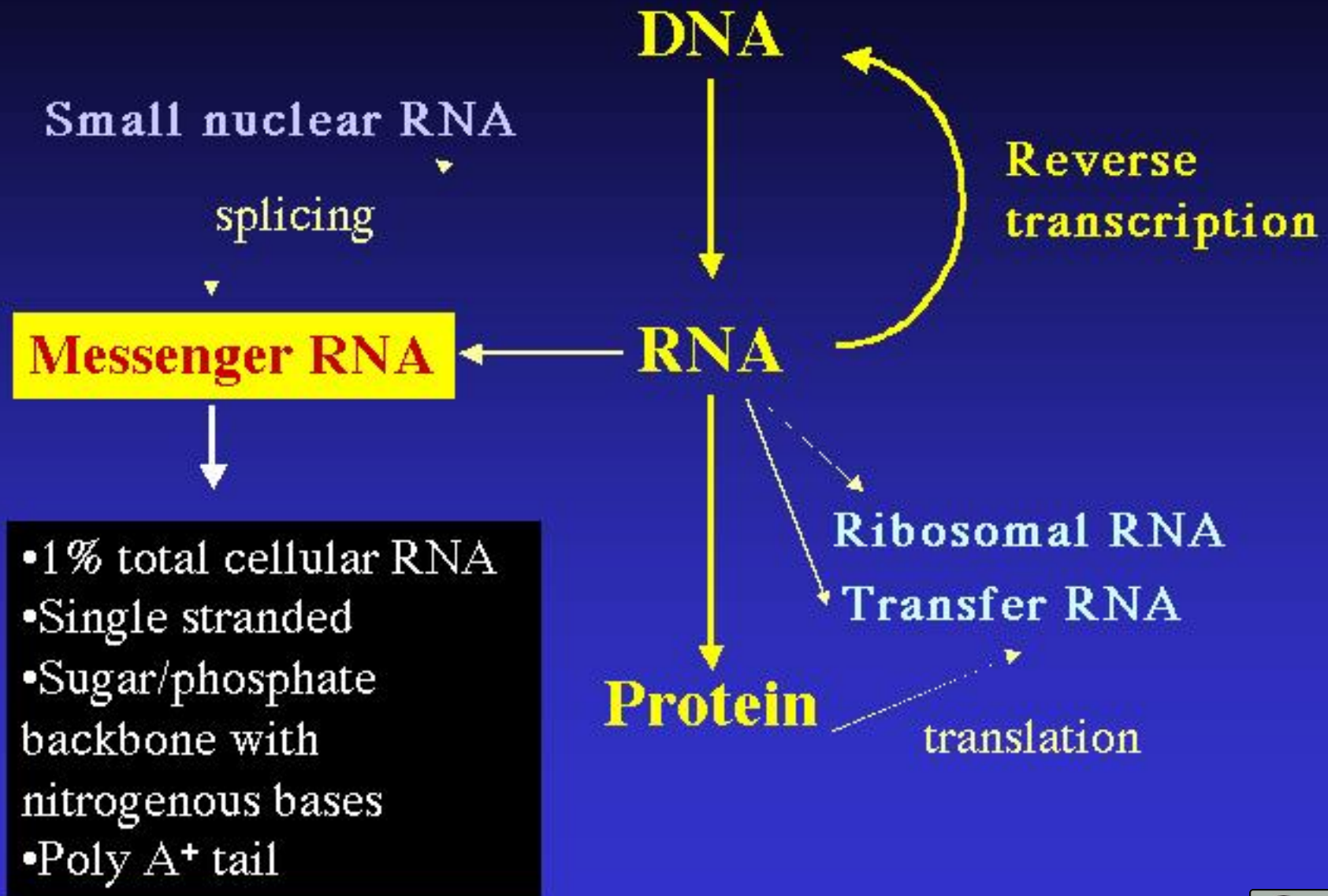


## Positive control

- Yield of reaction

## Negative control

- for contamination of reaction mixtures with similar sequences



# RNA processing

DNA



Transcription start



Transcription stop

Pre -RNA



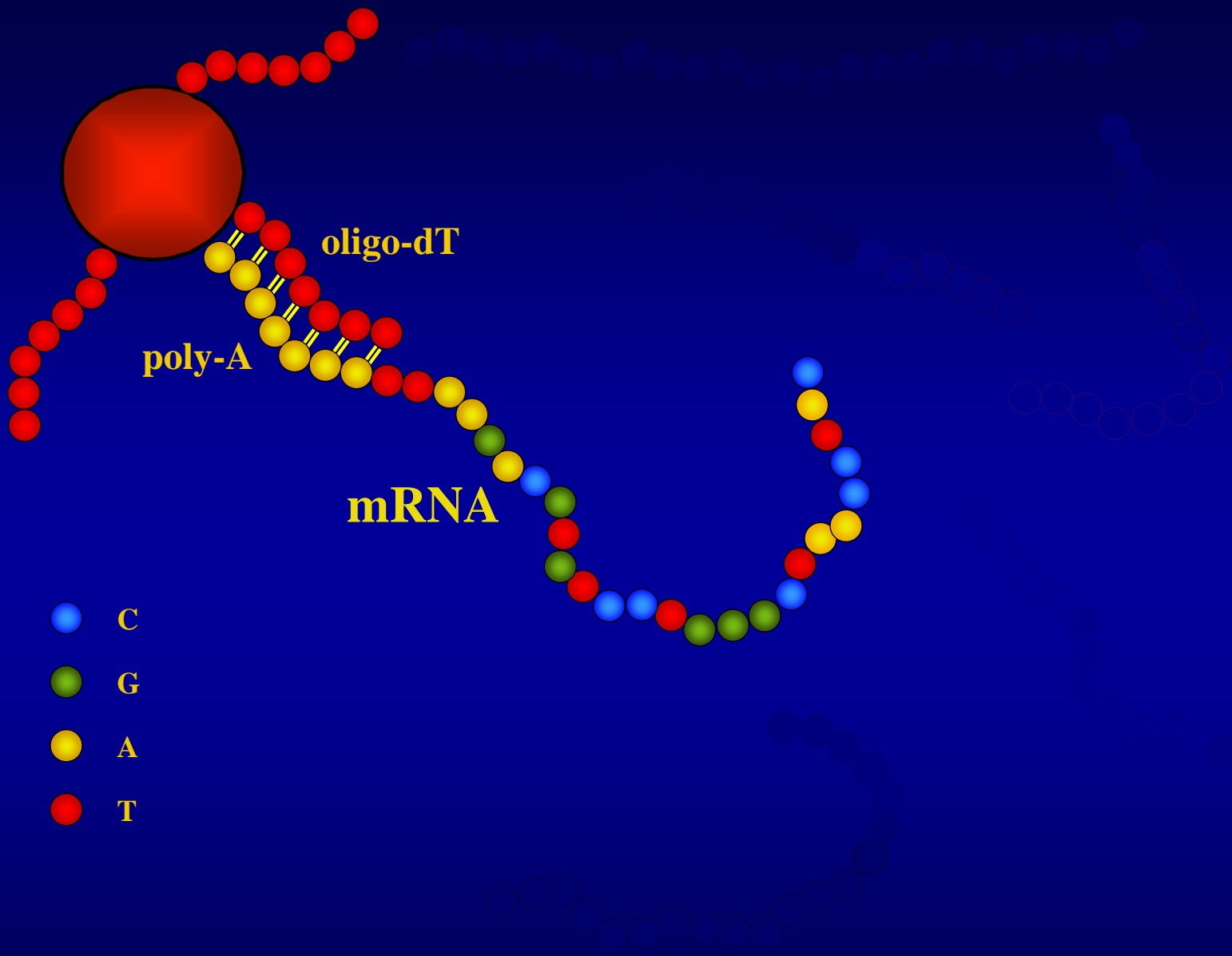
mRNA



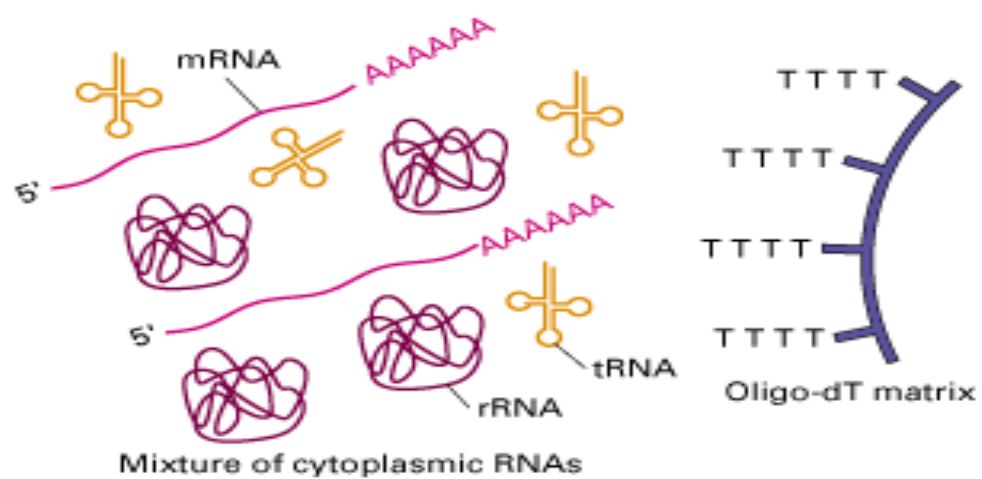
Translation start



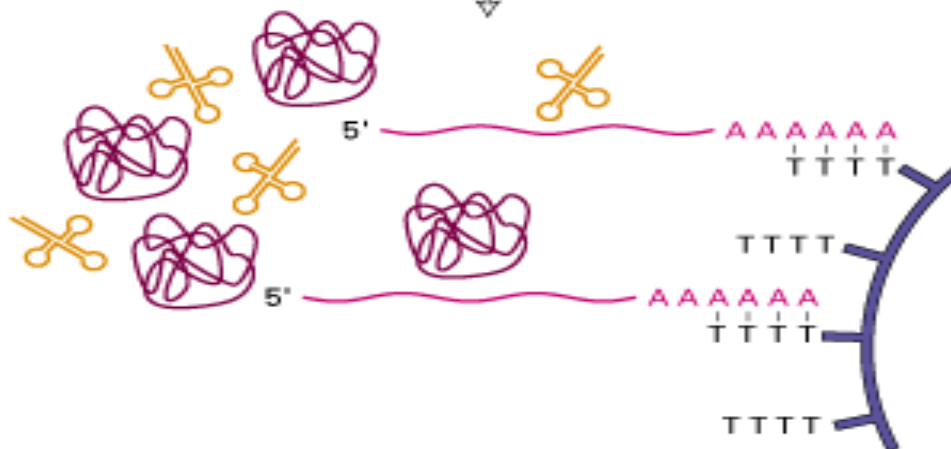
Translation stop



- C
- G
- A
- T



Mix under hybridization conditions



Wash away rRNA and tRNA

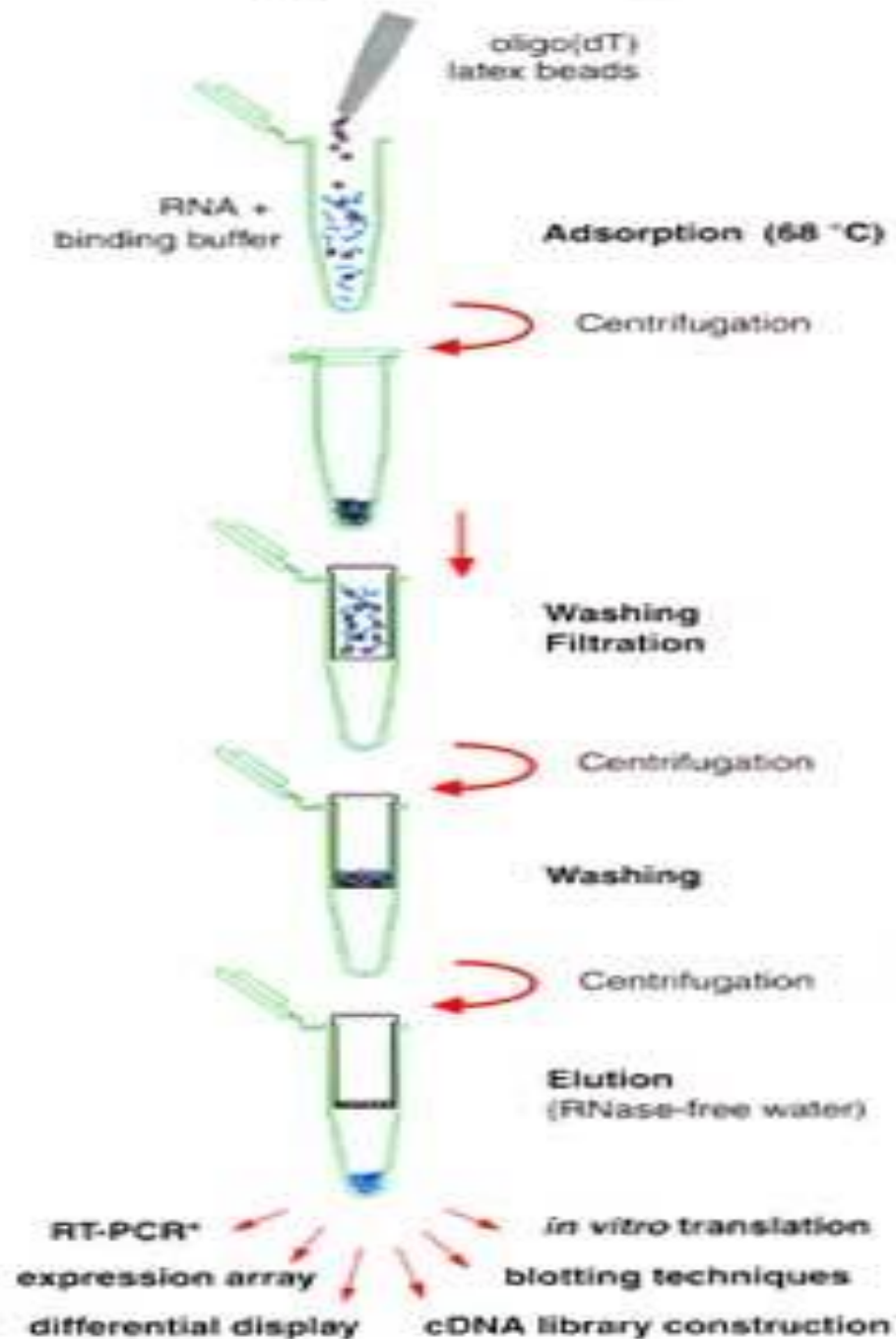
Elute column in low-salt buffer



Purified mRNA preparation



## NucleoTrap<sup>®</sup> mRNA working procedure



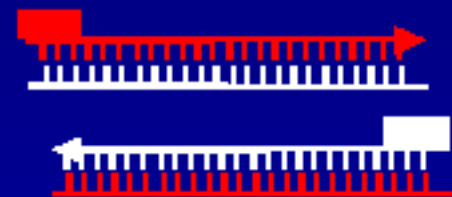
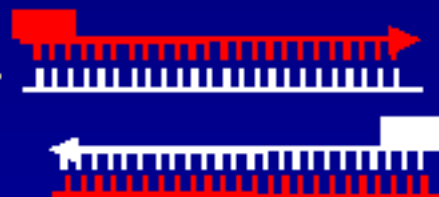
# Reverse transcription Polymerase Chain Reaction (RT-PCR)

Make a cDNA copy of mRNA with RT and poly dT primer



Copy of first strand with Taq polymerase

Denature, anneal primers, synthesis of DNA copies



Denature, anneal, synthesis, for many cycles to amplify

Visualise DNA

**CIP treatment to remove 5'-PO<sub>4</sub> from degraded mRNA, rRNA, tRNA and DNA**



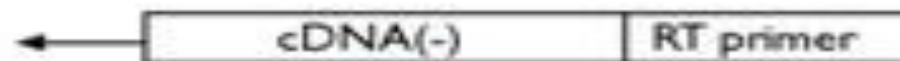
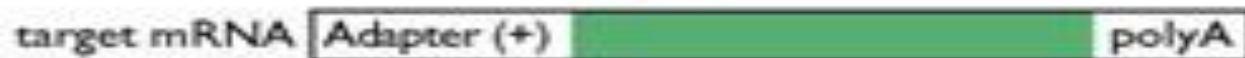
**TAP treatment to remove cap from full-length mRNA**



**RNA Adapter Ligation to Decapped mRNA**



**Reverse Transcription**



**Outer PCR**

# **The use of RT PCR**

**Cloning**

**Analysis of transcribed fragments**

# Subtraction library

RNA from  
experimental tissue

RNA from  
control tissue

Reverse transcription

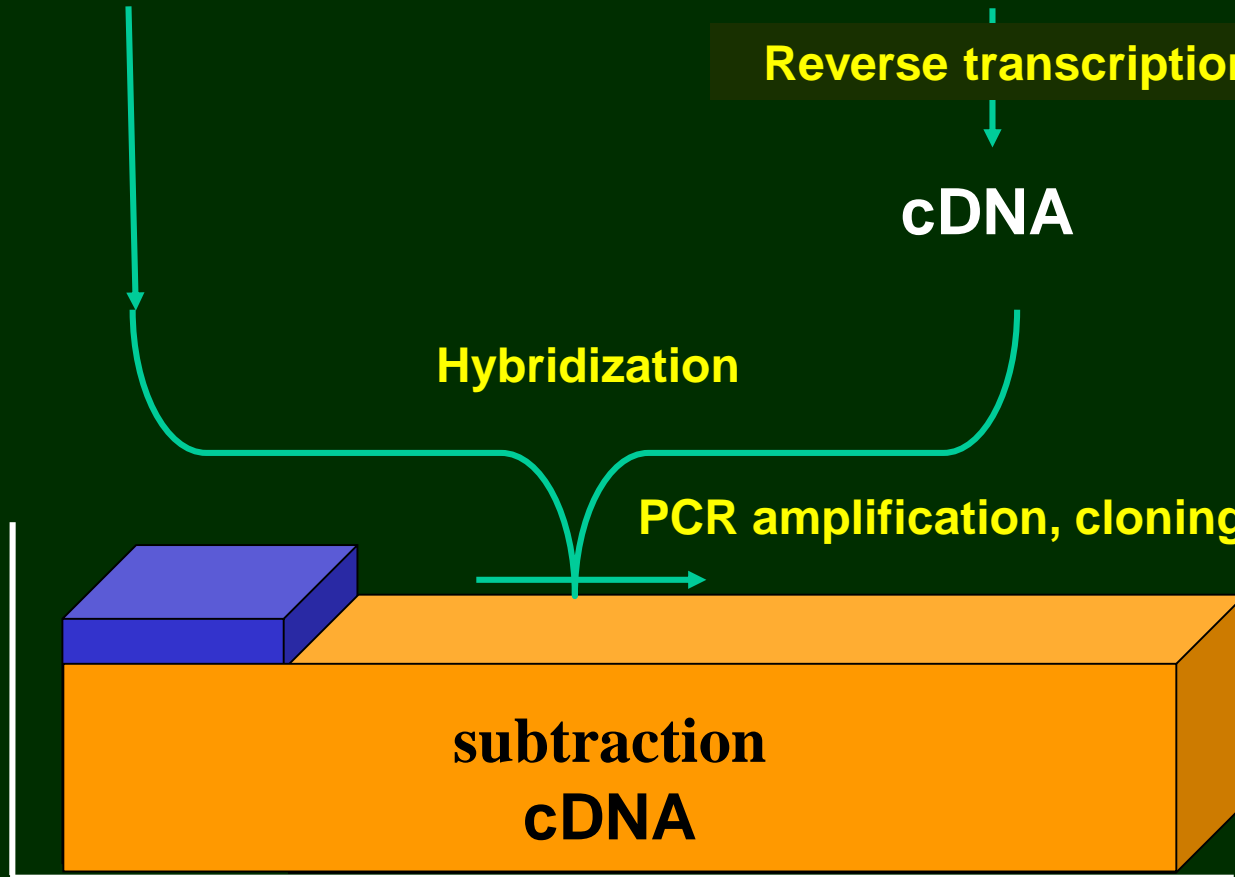
cDNA

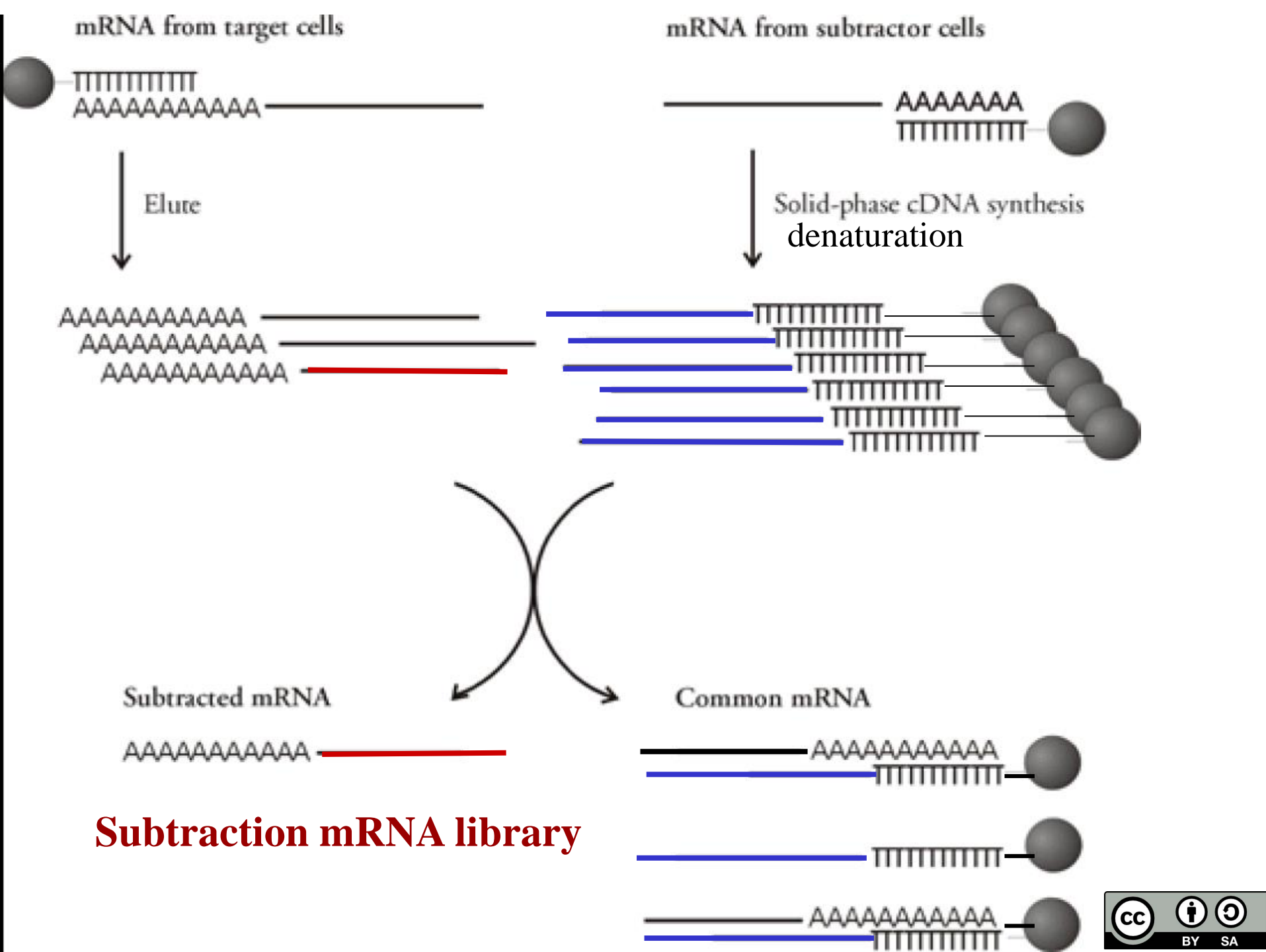
Hybridization

PCR amplification, cloning, sequencing

Amt

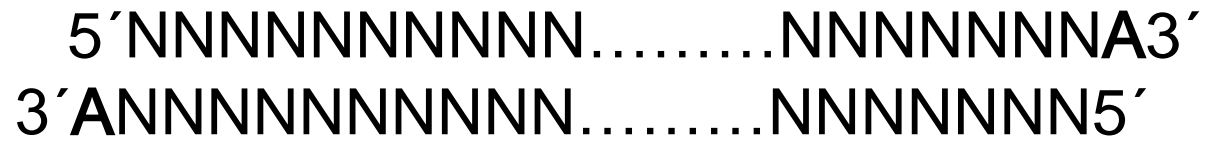
subtraction  
cDNA





# TA cloning

*Taq* polymerase joins 3' adenosine overhang:



**vector pCRII-topo – covalently joined topoisomerase I "**  
**→ activated" vector**

**Binds ds DNA in specific sequences**

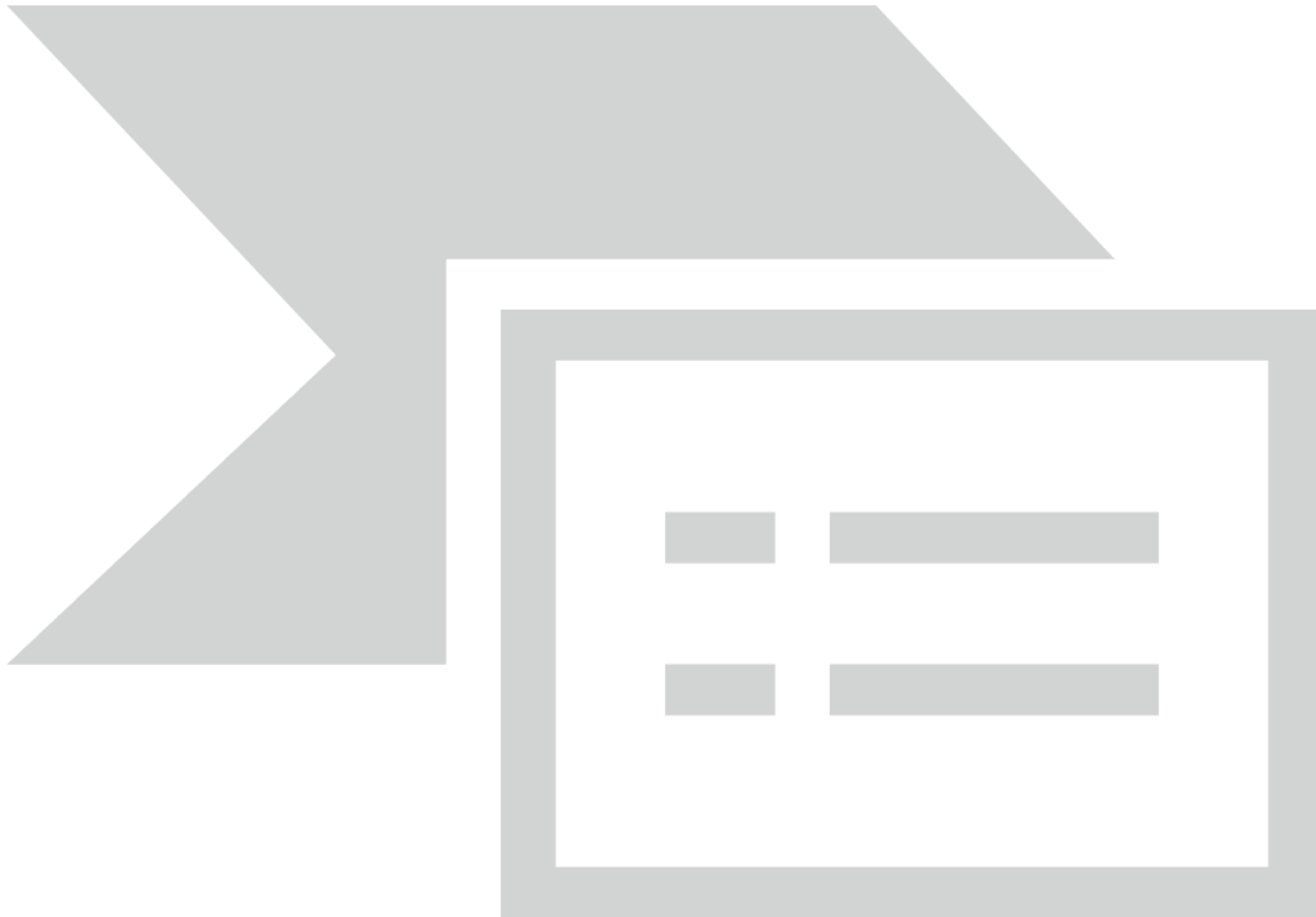
**Cleaves after 5' -CCCTT in one strand**

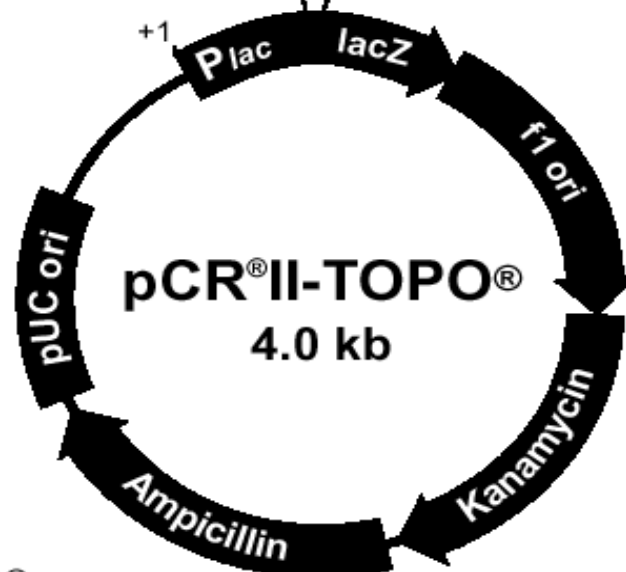
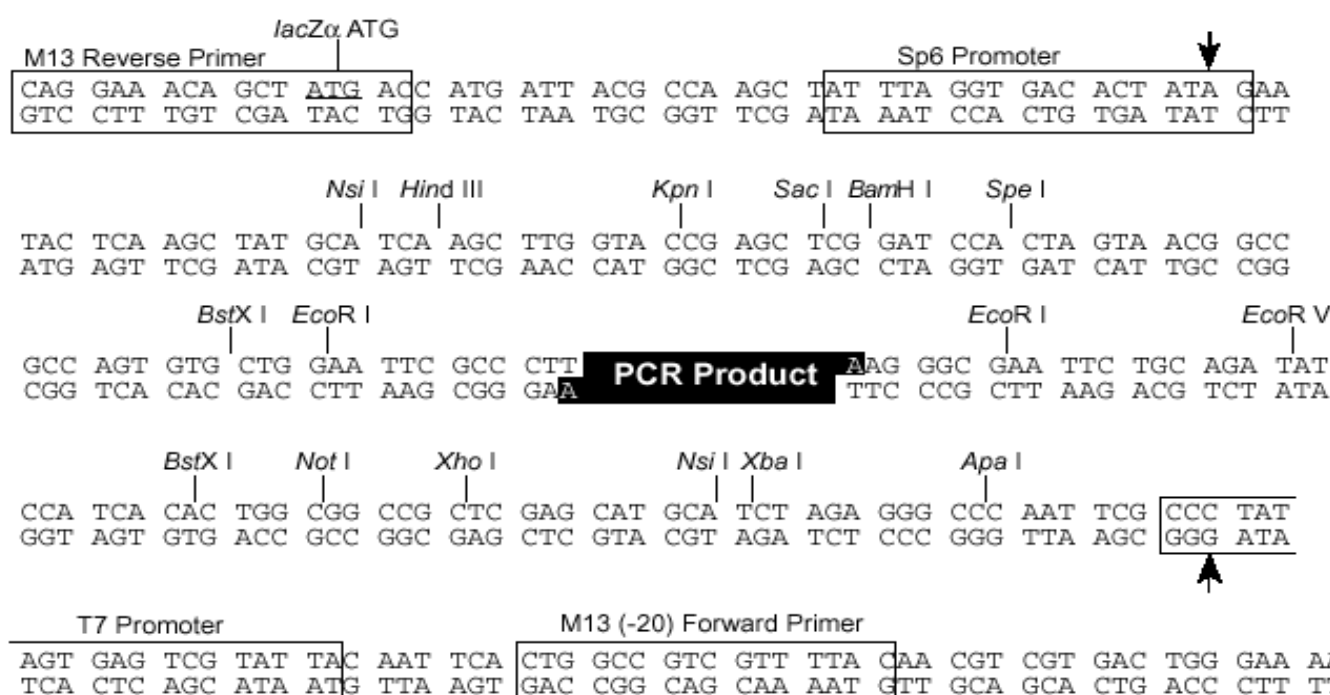
**E – covalent bond between 3' phosphate of cleaved chain and tyrosyl residue (Tyr-274) of topoisomerase I**

**Phosphotyrosyl bond between DNA and enzyme – attacked by 5' hydroxyl of ligated chain and enzyme is released**



NNNNNNNNNN.....NNNNNNNA  
ANNNNNNNNN.....NNNNNNN





# "Ligation Independent Cloning"

**Fragment insertion without phosphodiester bonds  
Joined with overhanging DNA ends**

**PCR – generated "T-less" complementary chain - long 3'-A-less region**

**3'-5' exonuclease**

**(*Pfu* polymerase, T4 DNA polymerase, Klenow fragment, etc.)**

**- In presence of dATP – degradation of "A-less" region – stop at the first A nucleotide**

**pLIC – complementary sequence + 3'-5' exonuclease and dTTP**

**- (13-14 nt) 5' overhanging end**

**stably joined – hydrogen bonds of long overhang**

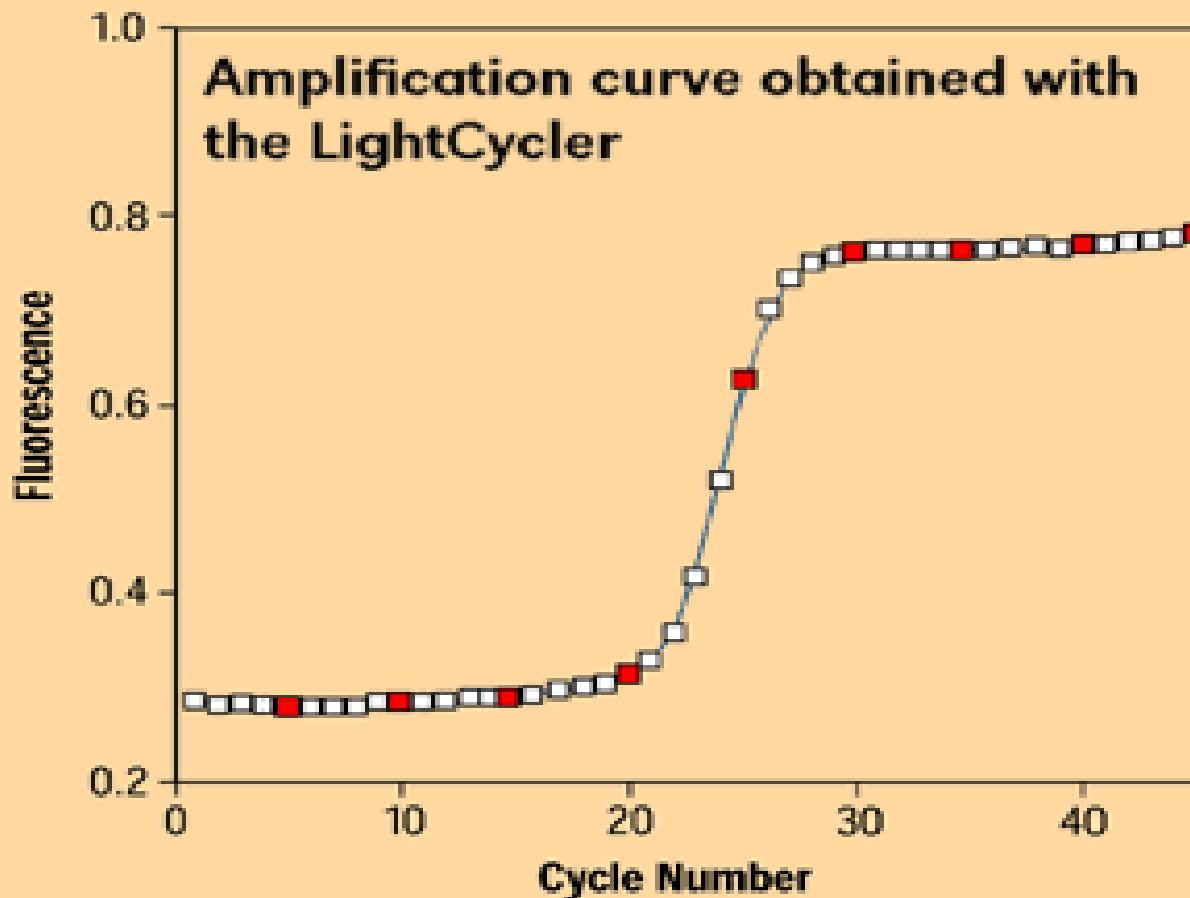
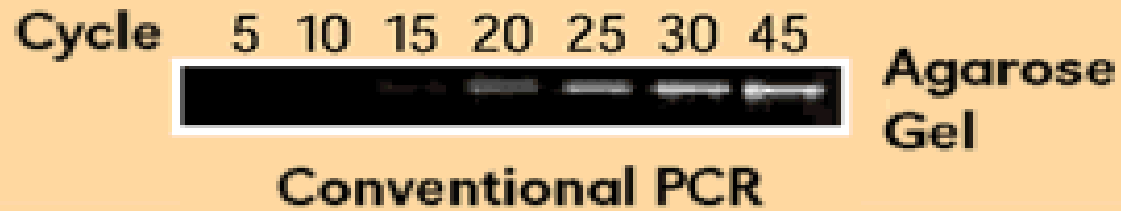
**- in *E. coli* – gaps repaired**

**LIC – cloning**  
**3'-5' exonuclease**  
**(*Pfu* polymerase, T4**  
**DNA polymerase,**  
**Klenow..)**

**Presence of dATP**  
**"A-less" sequence**  
**removal**  
**- stop at first A**



# Quantitative PCR



## **SYBR Green I**

- fluorescence dye**
- binds minor groove in dsDNA**

## **Unbound dye**

- low fluorescence**

## **SYBR Green I**

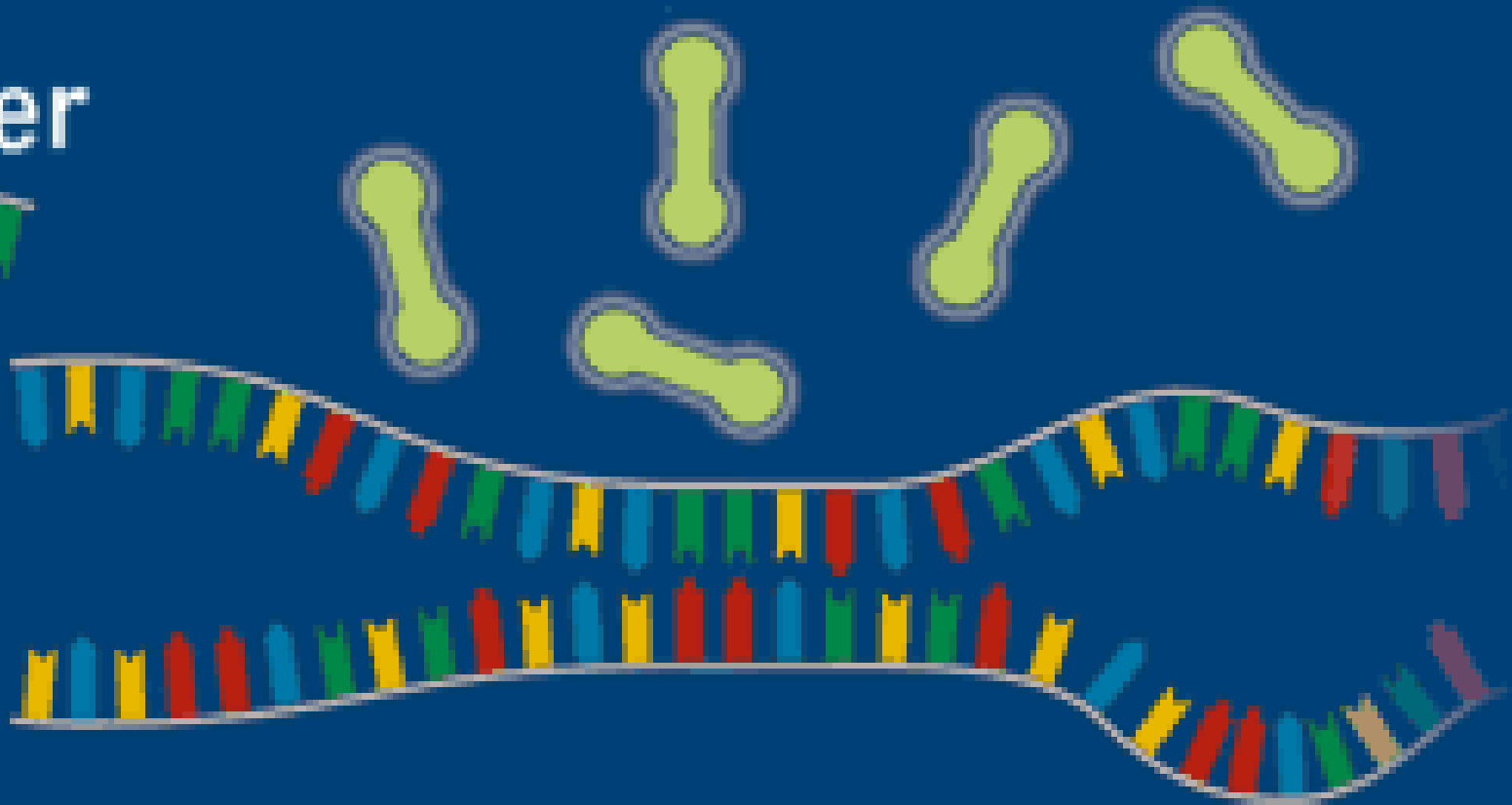
- very stable**

**(only 6% activity lost during 30 amplification cycles)**

**LightCycler – corresponding excitation and emission range**

# SYBR Green

## Primer

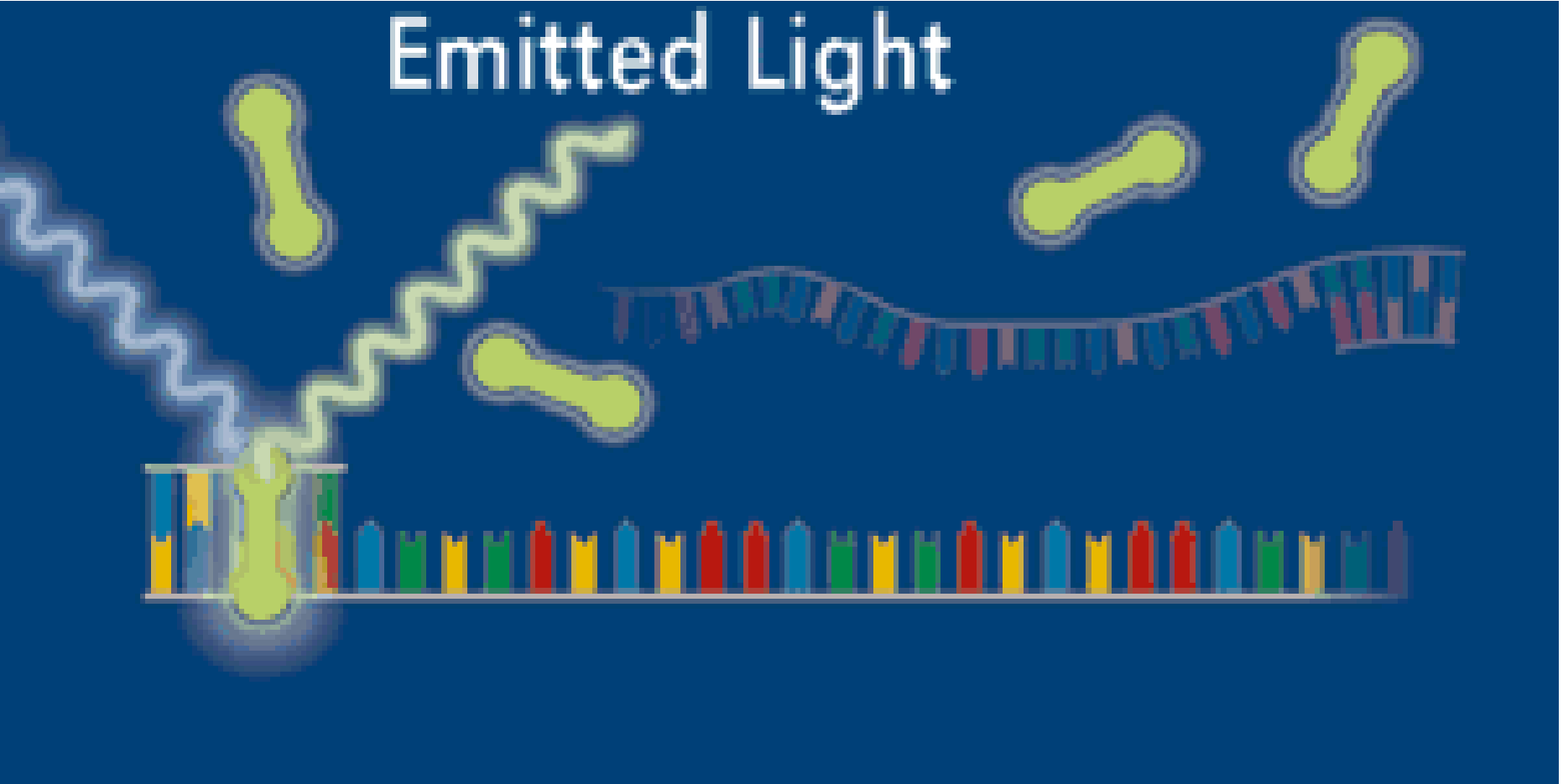


Reaction mixture DNA, primers, dye.

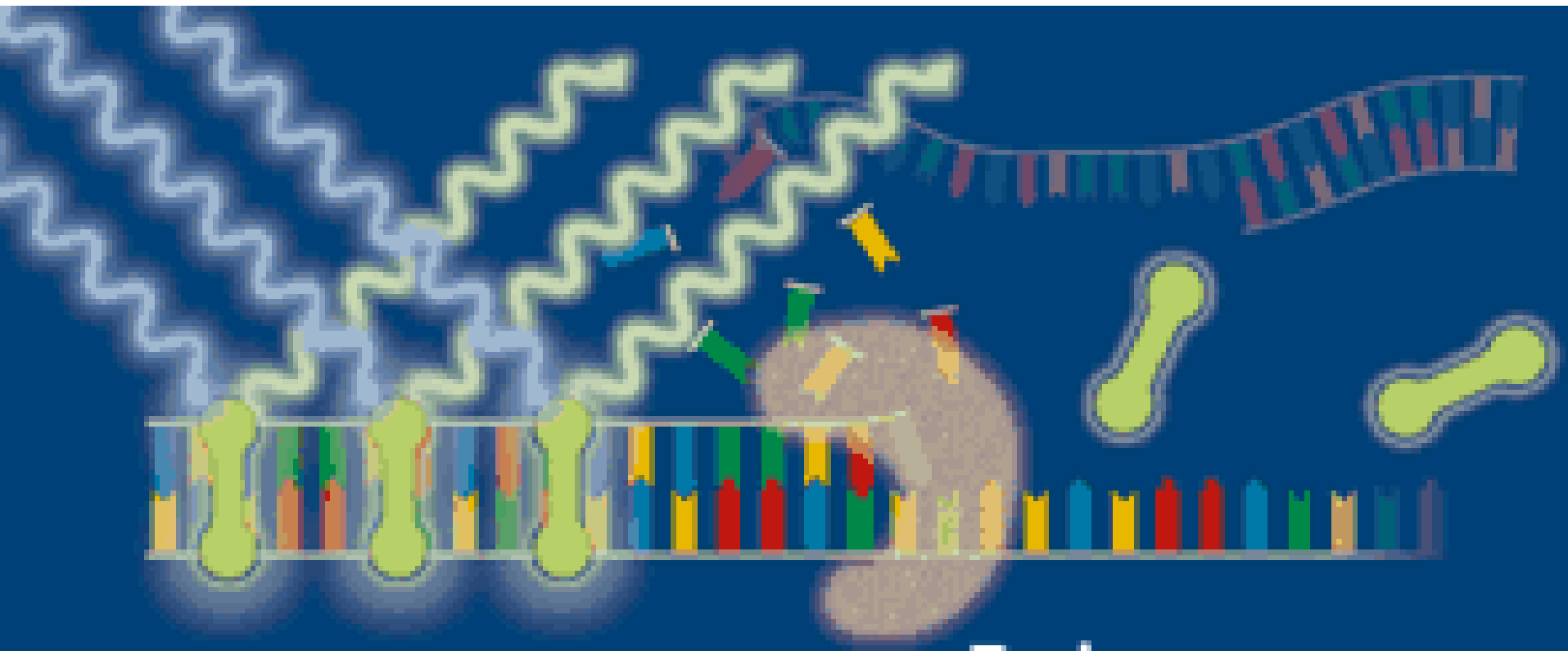
Unbound molecules – weak fluorescence



# Emitted Light

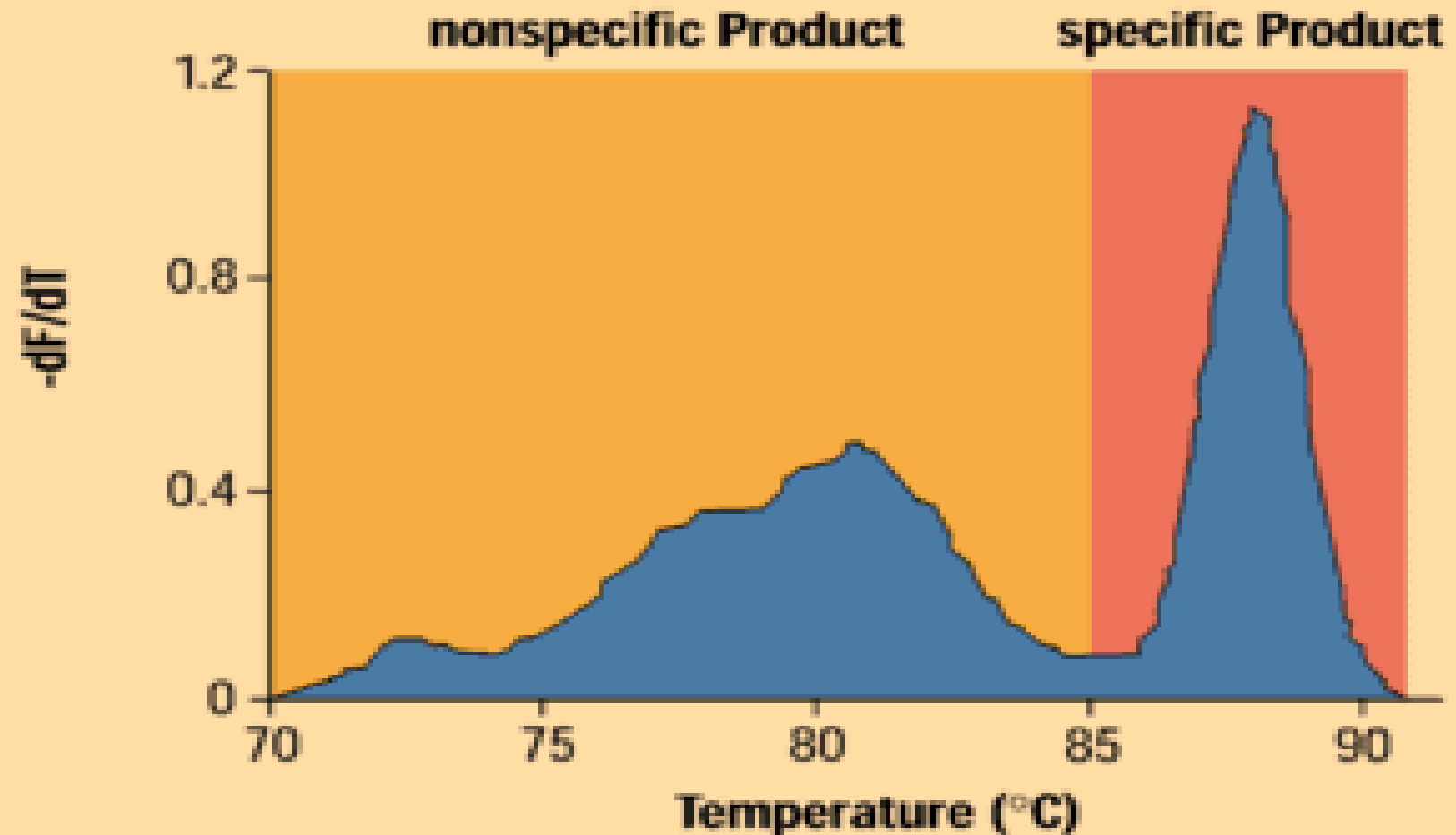


**Binding to dsDNA - enhanced emission of excited light SYBR Green**



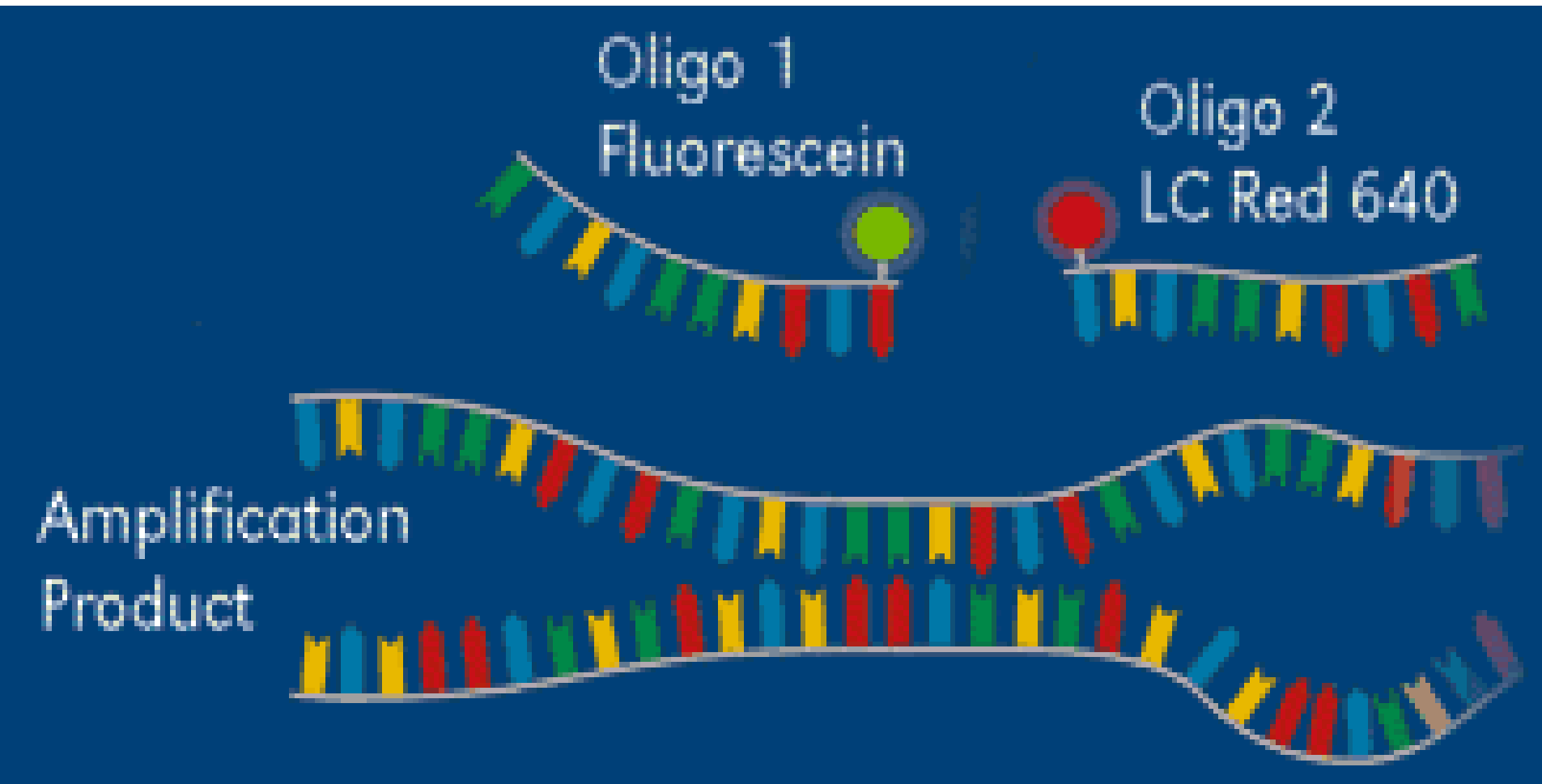
Polymerase

## ■ Melt primer dimer signal away



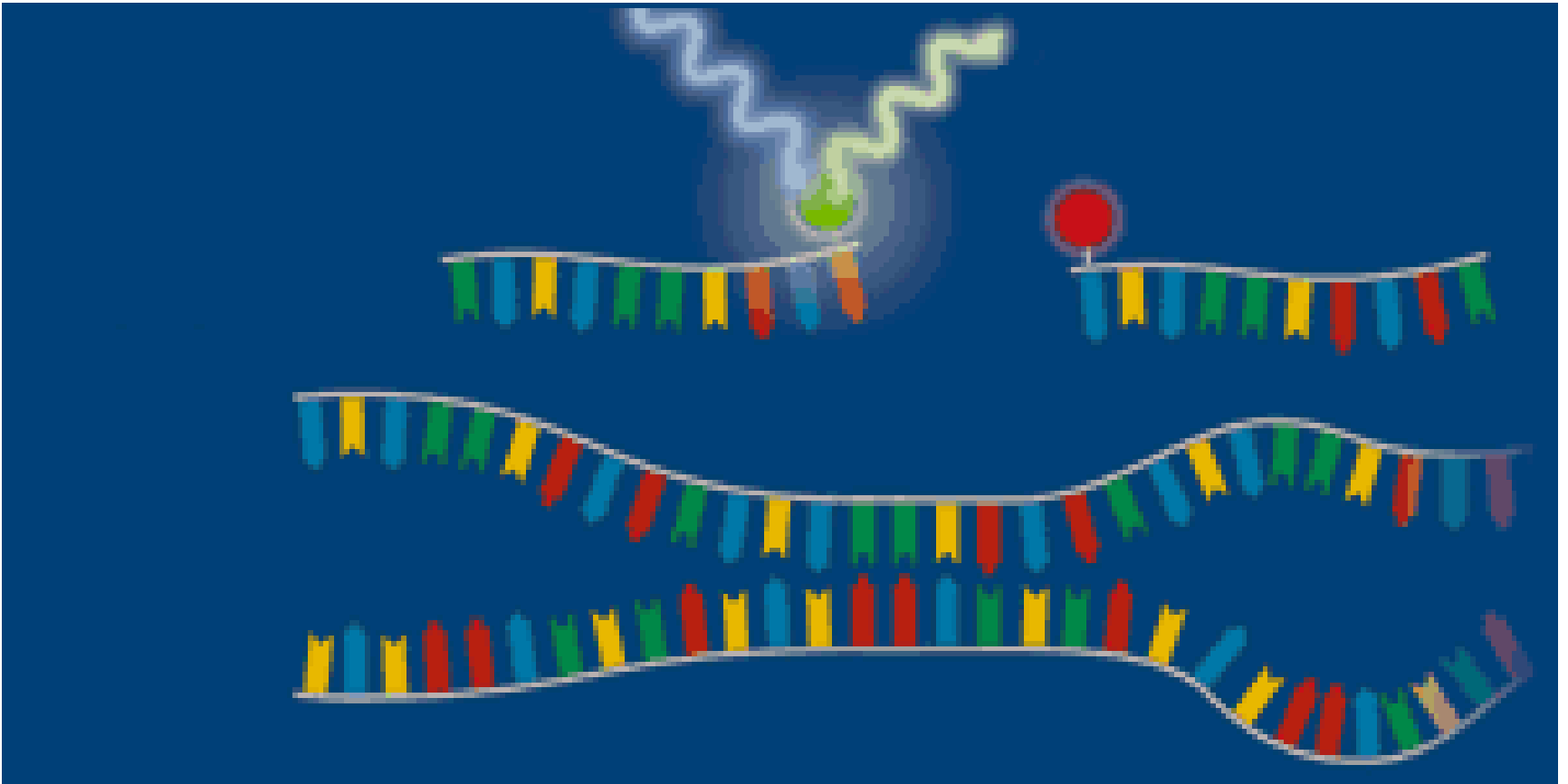
Distinguishing of specific signal – high temperature – close to  $T_m$

# FRET (Fluorescence Resonance Energy Transfer)

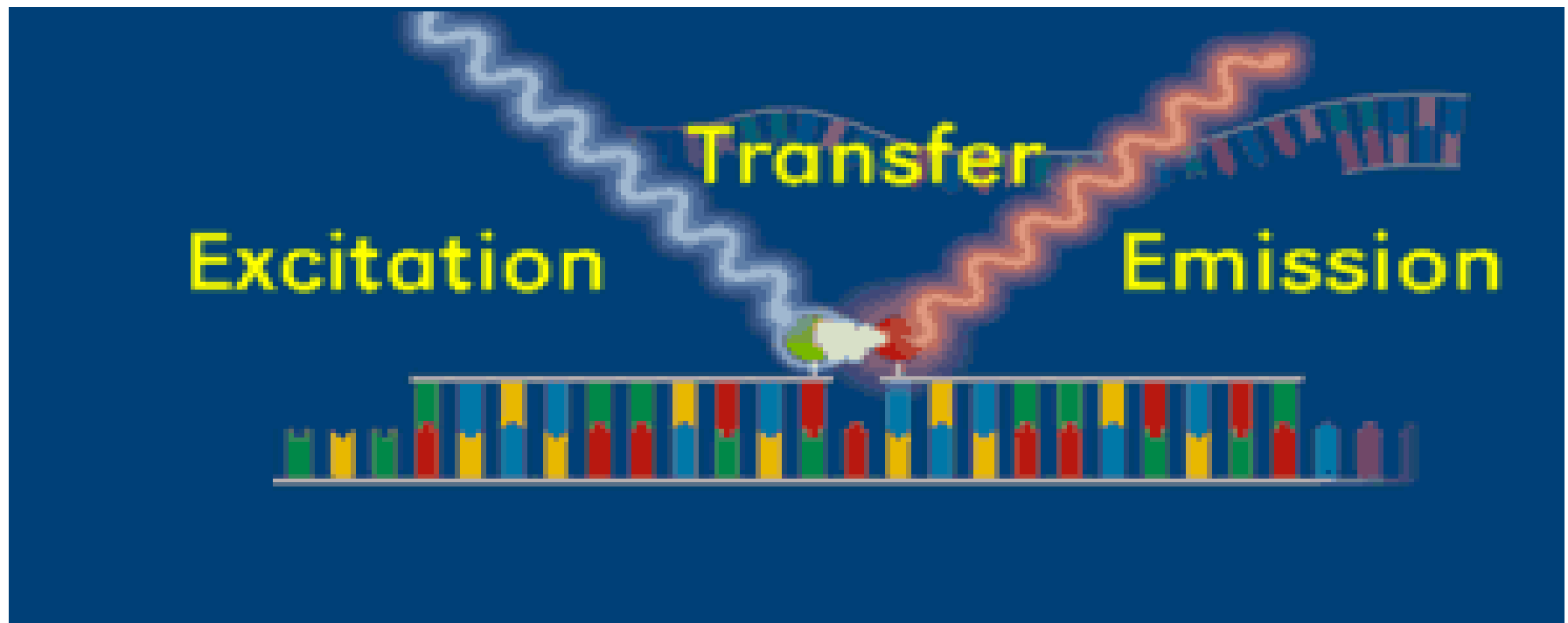


Oligo 1 with fluorescent probe (fluorescein) on 3'-end

Oligo 2 with other fluorescent probe (LC red 640) on 5'-end



**Oligonucleotides designed to hybridize with DNA fragments in „head to tail“ orientation  $\Rightarrow$  fluorescence labels in a close proximity**

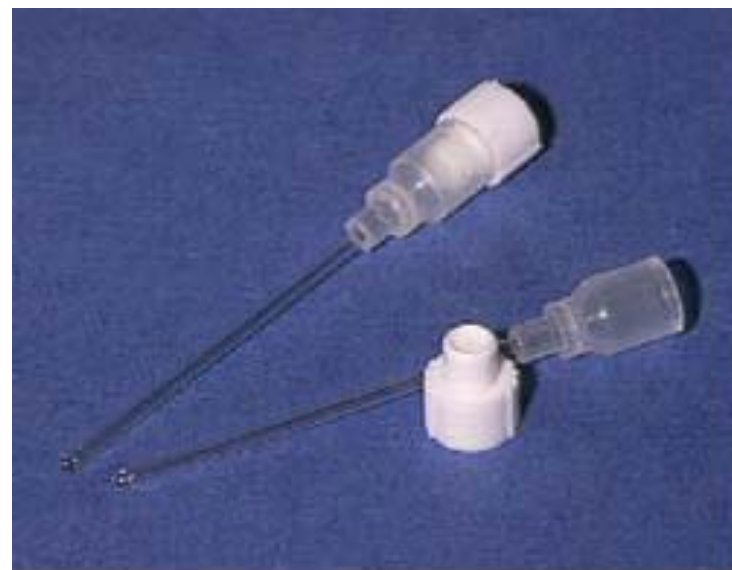
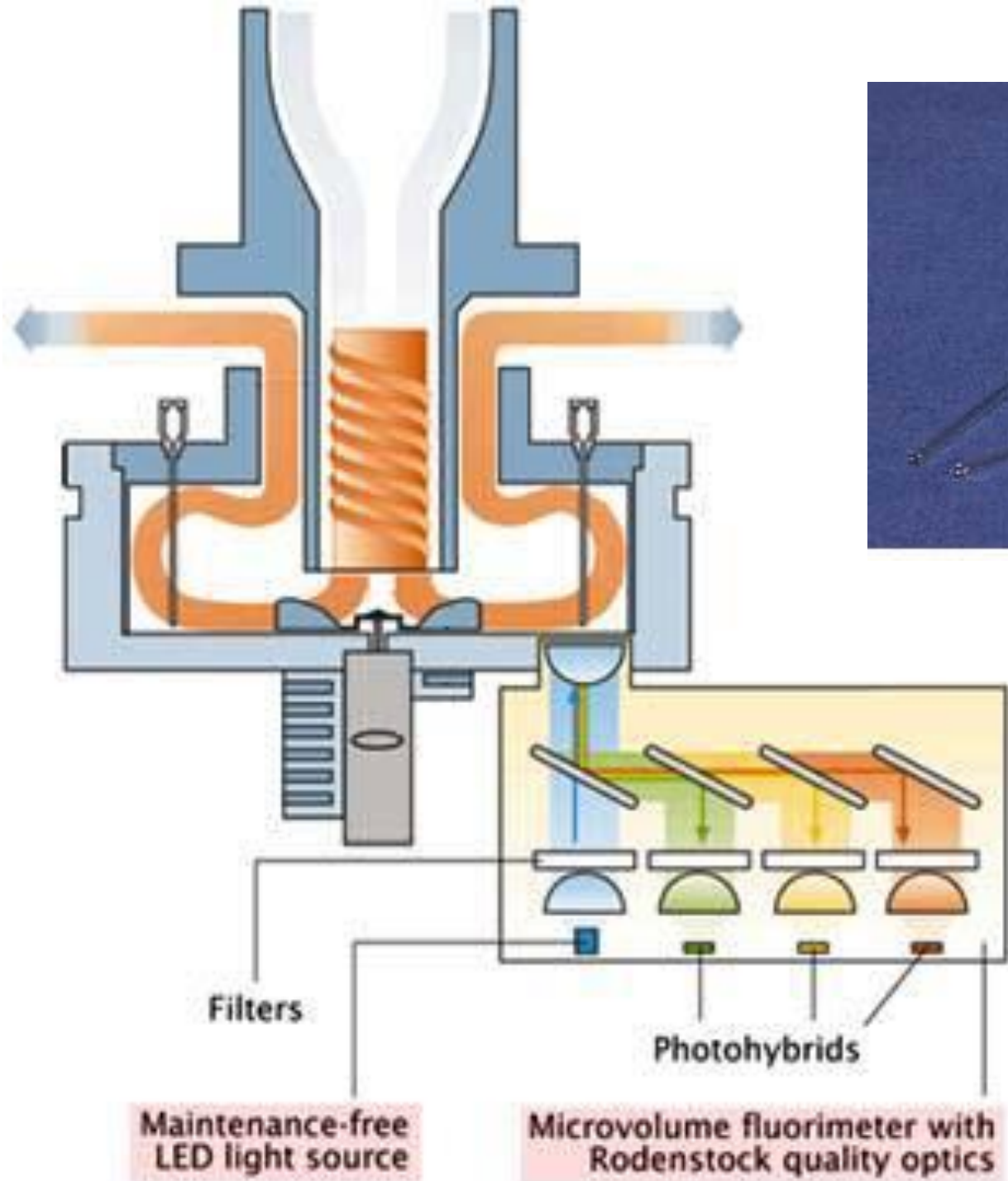


**Fluorescein - excitation by LED (Light Emitting Diode)**

**- emission - green fluorescence light – long wavelength**

**Neighboring dye - emitted energy - excitation LC Red 640 on the second probe – emission of red fluorescence light of higher wavelength**

**FRET – only in close proximity (1–5 nucleotides) – only at annealing temperature – the probe falls off during polymerization temperature**



# Real Time Instruments

Real Time Instruments combine thermal cyclers with fluorescent detection



ABI 7500 Real Time System



iCycler iQ®



Smartcycler®



LightCycler®

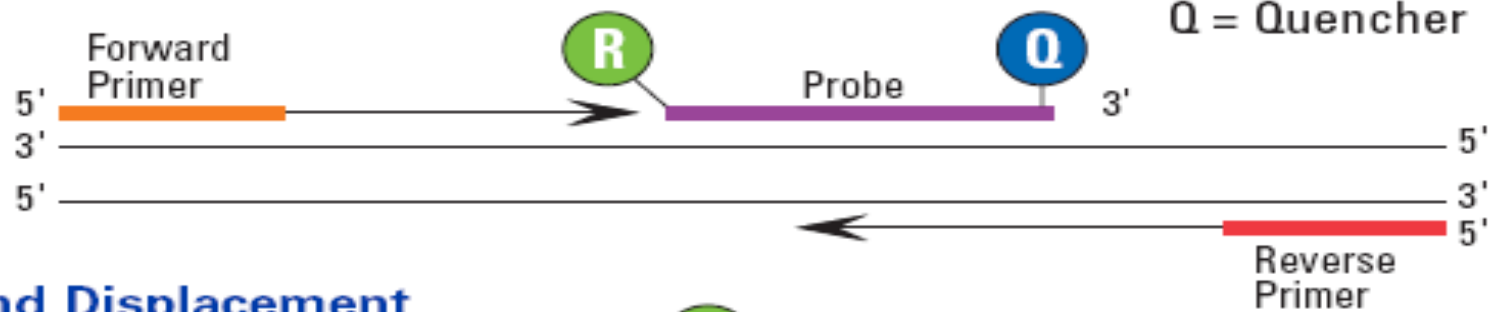


# TaqMan PCR

## Probe (TaqMan)

- specific for the sequence between both primers
- fluorescently labeled
- 5' reporter fluorochrome (e.g. 6-carboxyfluorescein [6-FAM]) and quenching fluorochrome (6-carboxy-tetramethyl-rhodamine [TAMRA]) on any T or at the 3' end
- probe  $> T_m$  than primers
- 100% hybridized during extension
- quenching – if both fluorochromes are part of the probe
- 5' - 3' nuclease activity of *Taq* degradation of the probe
  - release of reporter fluorochrome from the quencher

## Polymerization



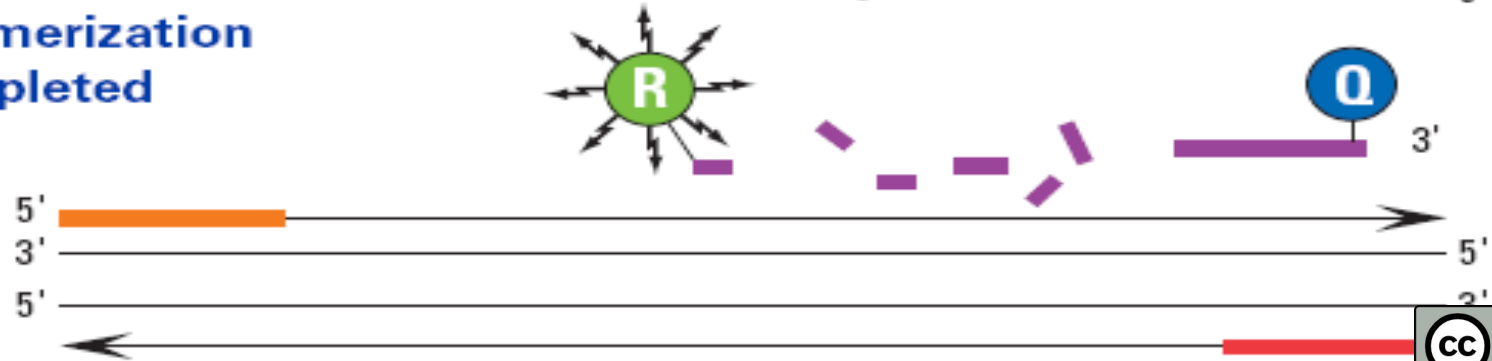
## Strand Displacement



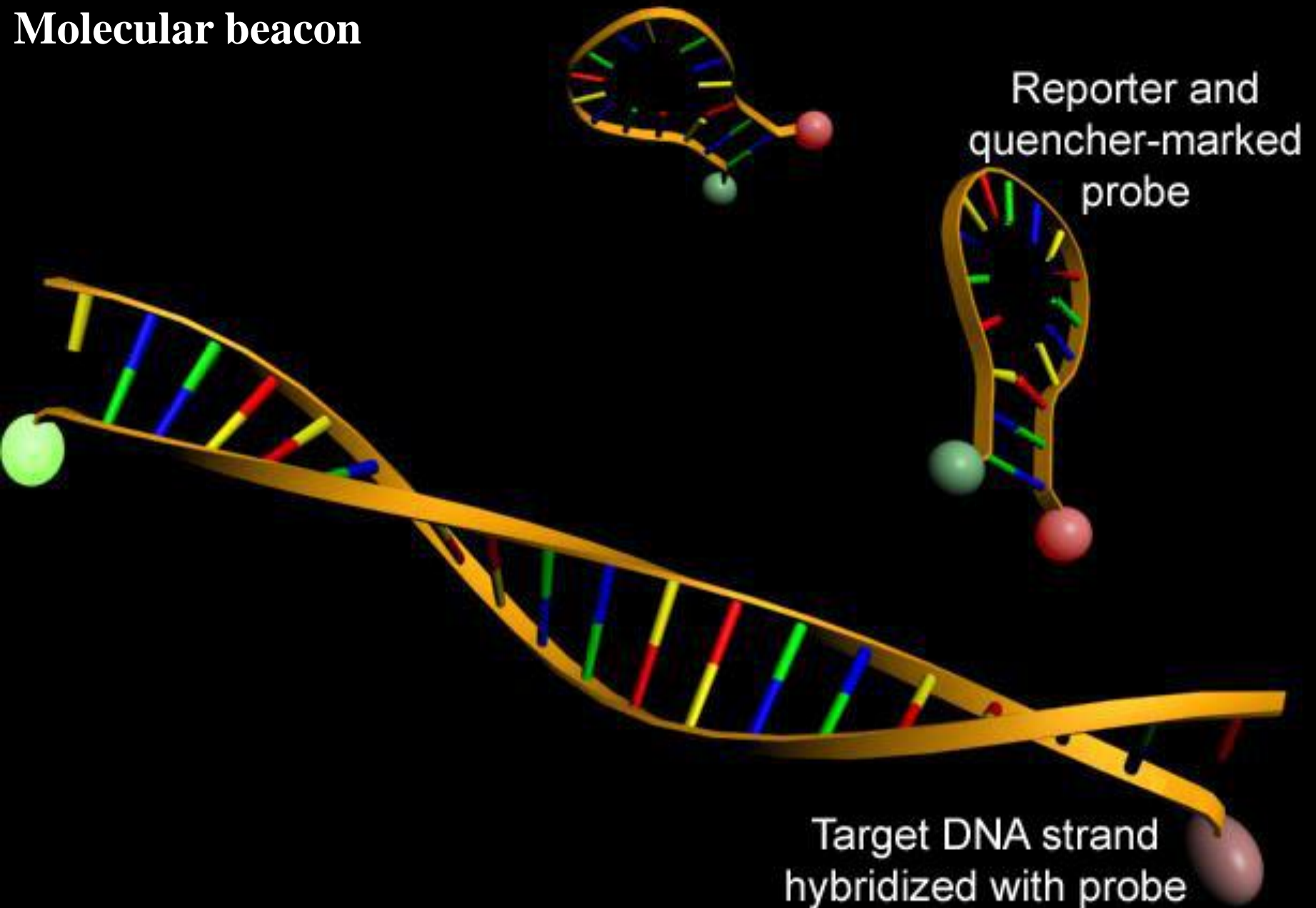
## Cleavage



## Polymerization Completed

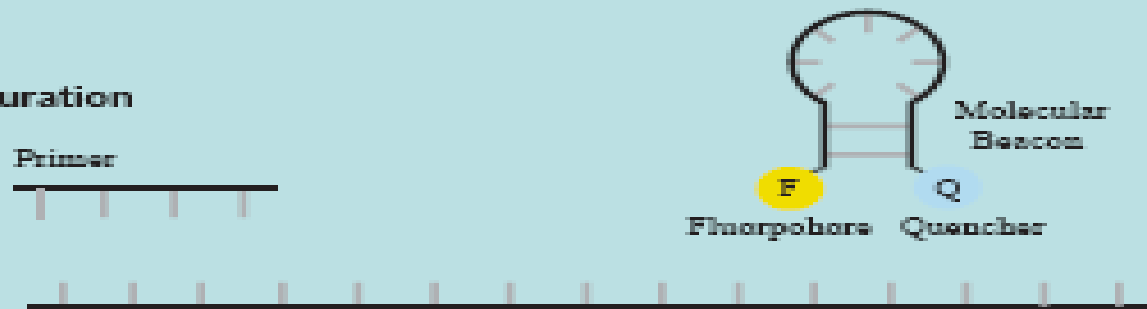


# Molecular beacon

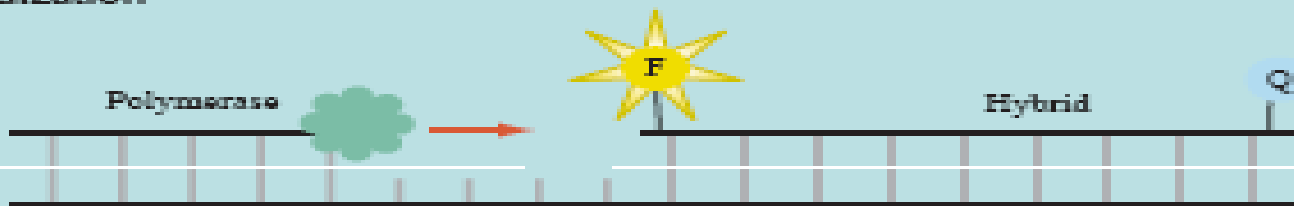


# Molecular Beacon

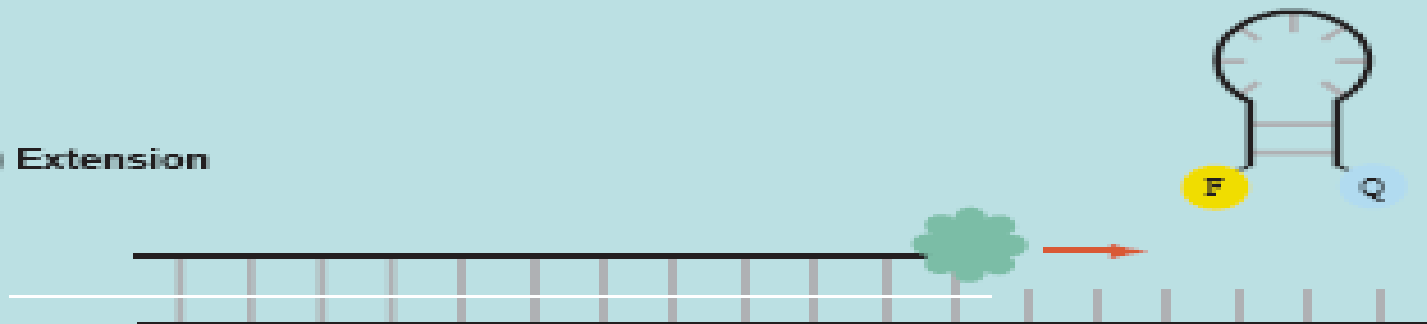
## 1) Denaturation



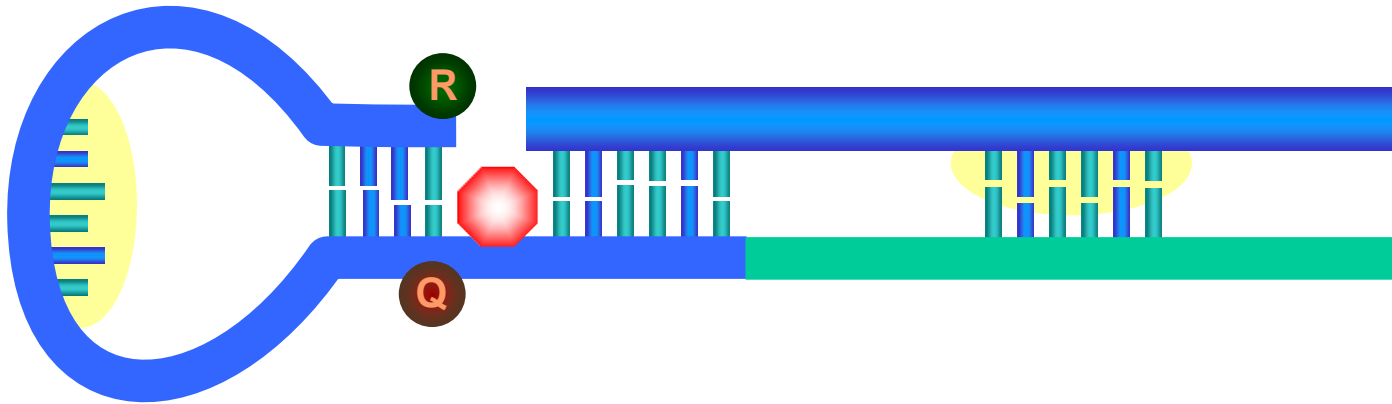
## 2) Hybridization



## 3) Extension

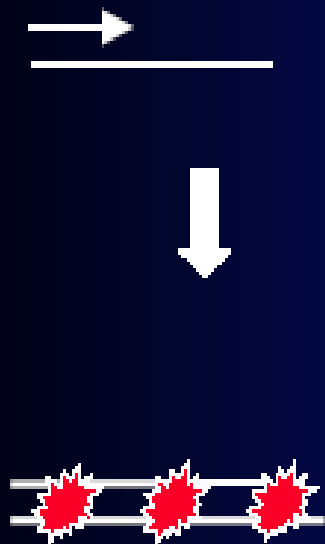


# SCORPIONS

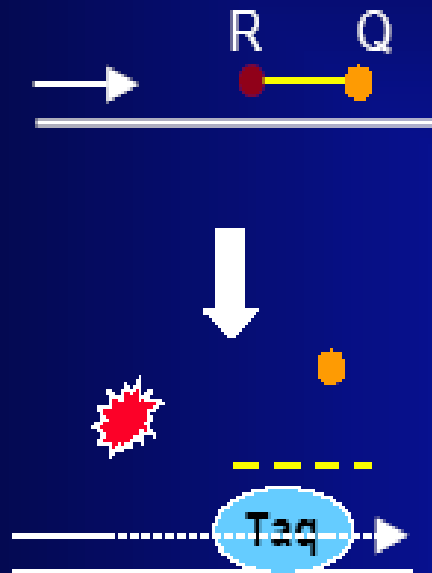


# Methods of fluorescence detection

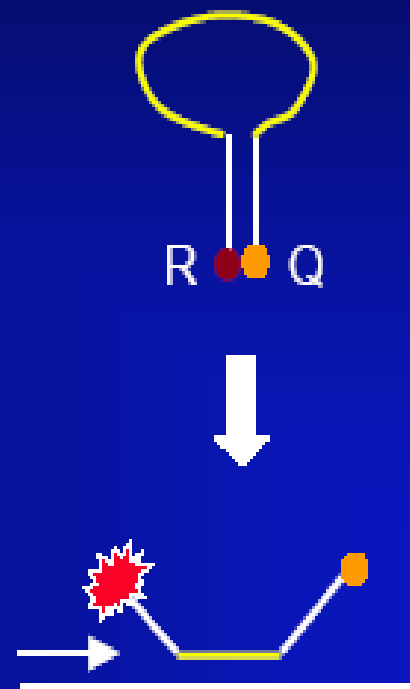
SYBR Green



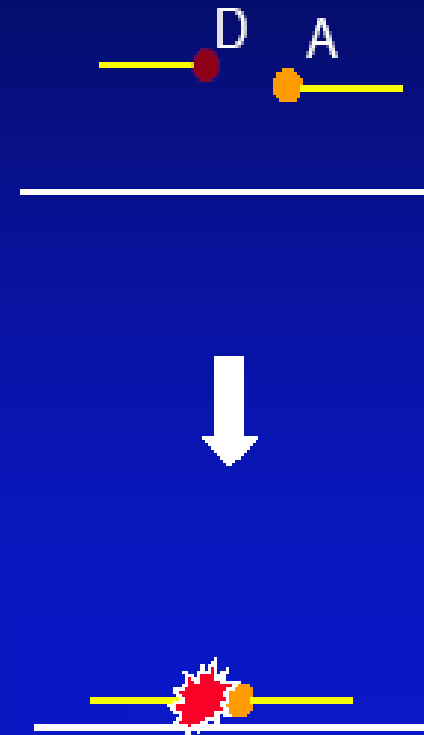
Taqman



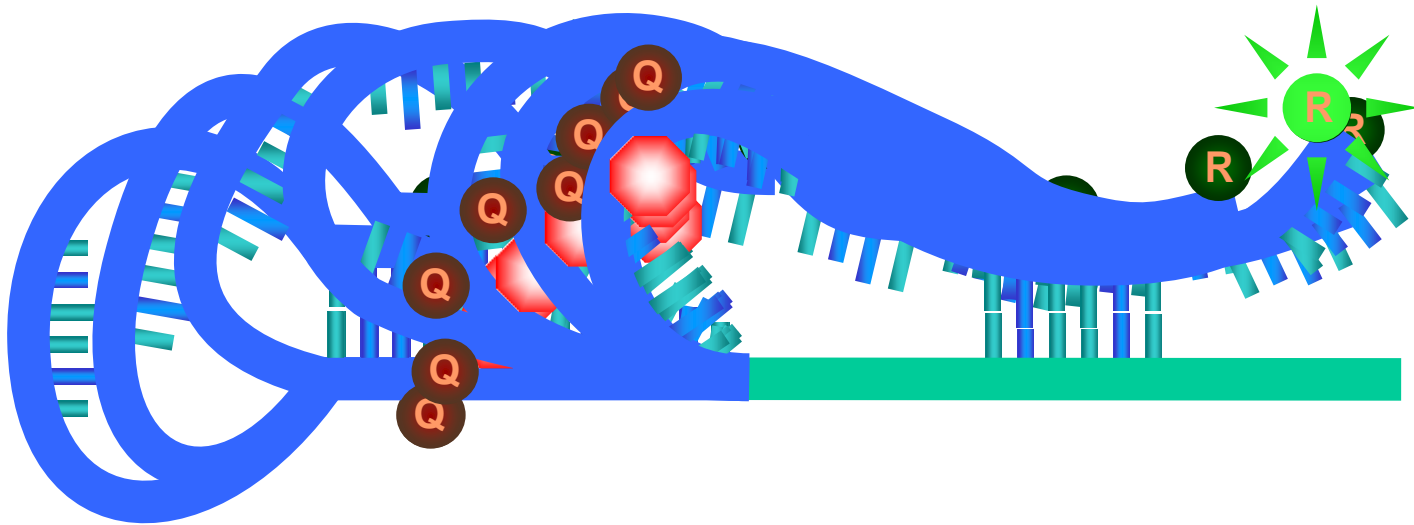
Molecular  
Beacons



Light  
Cycler



# SCORPIONS



Uveřejněné materiály jsou určeny studentům Vysoké školy chemicko-technologické v Praze

jako studijní materiál. Některá textová i obrazová data v nich obsažená jsou převzata

z veřejných zdrojů. V případě nedostatečných citací nebylo cílem autora/ů záměrně poškodit event. autora/y původního díla.

S eventuálními výhradami se prosím obračejte na autora/y konkrétního výukového materiálu, aby

bylo možné zjednat nápravu.

The published materials are intended for students of the University of Chemistry and Technology, Prague as a study material. Some text and image data contained therein are taken from public sources. In the case of insufficient quotations, the author's intention was not to intentionally infringe the possible author(s) rights to the original work.

If you have any reservations, please contact the author(s) of the specific teaching material in order to remedy the situation.



# Gene therapy



EUROPEAN UNION  
European Structural and Investing Funds  
Operational Programme Research,  
Development and Education



MINISTRY OF EDUCATION,  
YOUTH AND SPORTS

# Target of gene therapies

Treatment of inherited and acquired genetic diseases

- Cancer
- AIDS / HIV

## Good news

Some promising results - treatment of SCID.

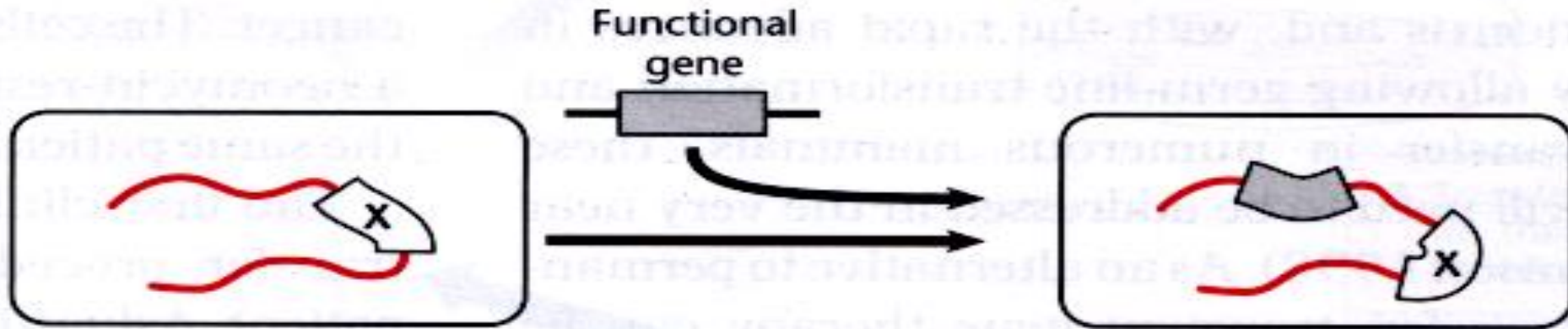
## Problems

Lack of sufficient knowledge of the potential risks

→ Complications

- **X** — Mutant (disease) gene – loss of function
- **D** — Mutant (disease) gene – dominant gain of function

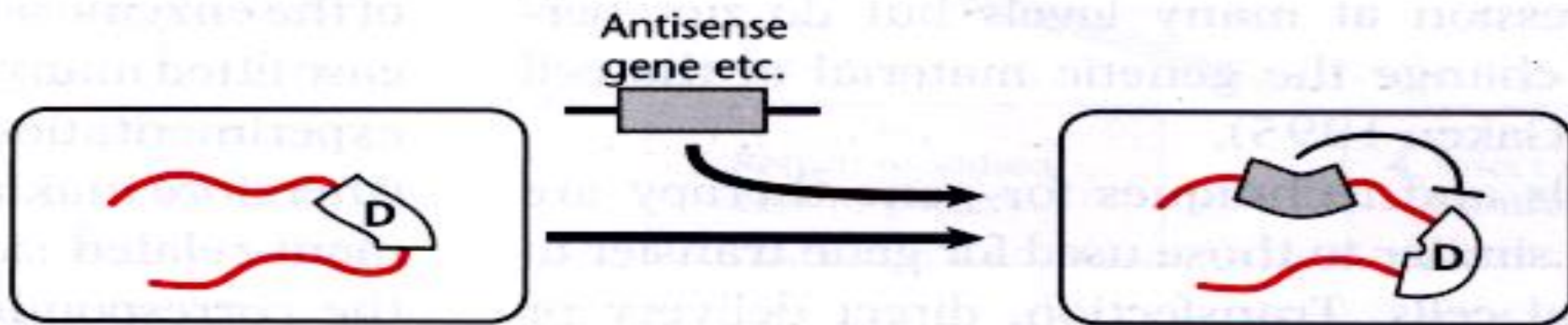
## 1 Gene augmentation therapy



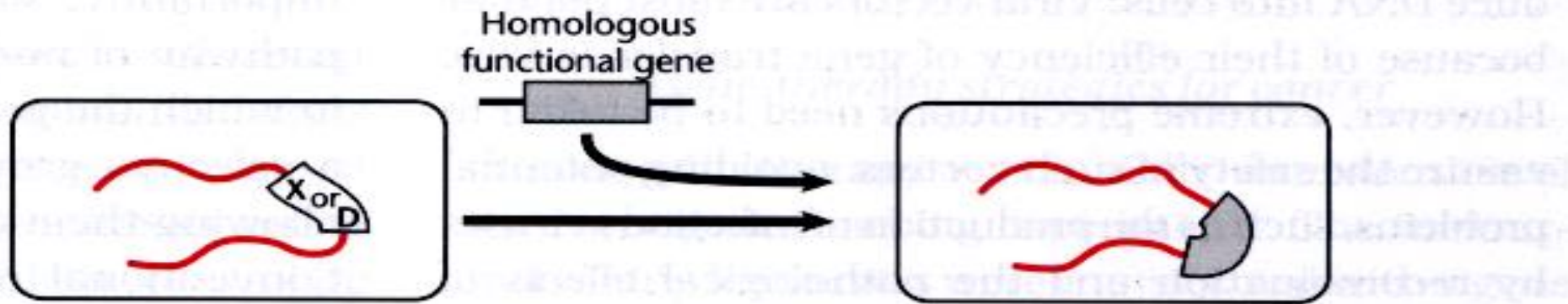
## Gene augmentation

- **X** — Mutant (disease) gene – loss of function
- **D** — Mutant (disease) gene – dominant gain of function

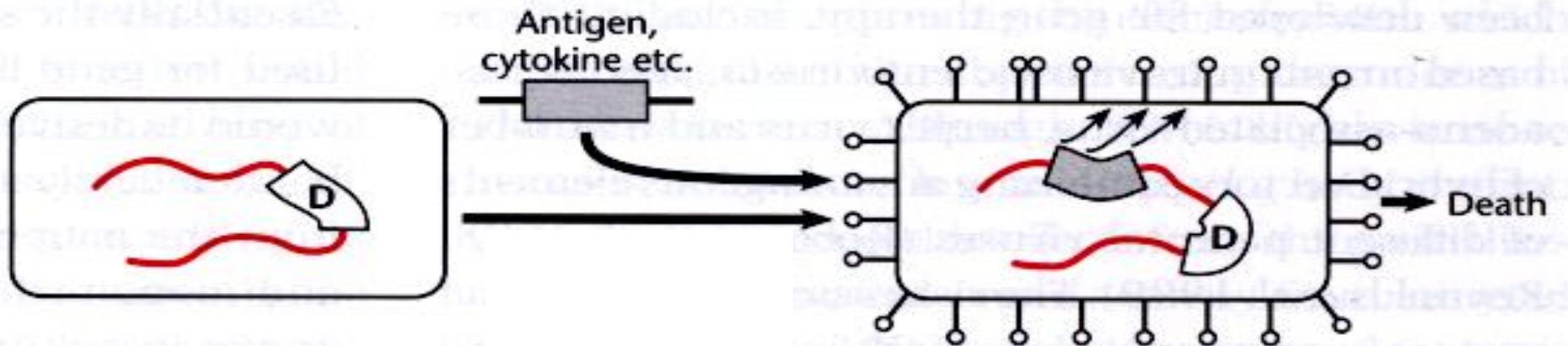
## Gene inhibition



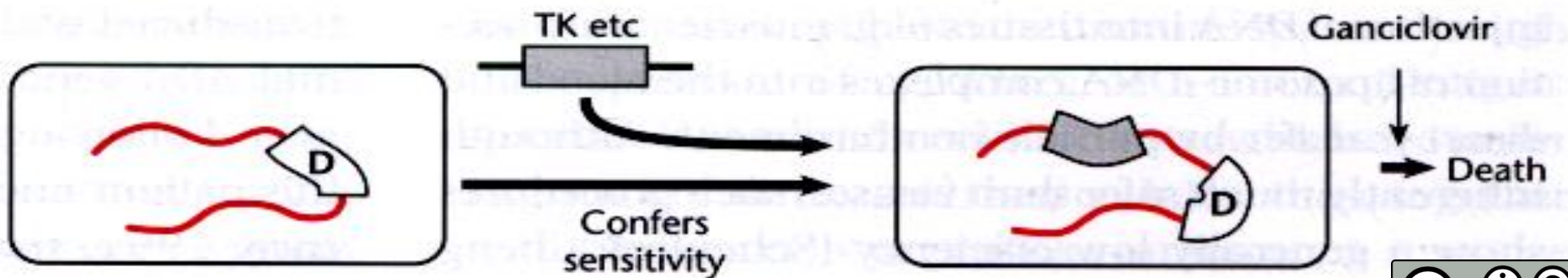
### 3 Gene targeting



### 4 Assisted killing

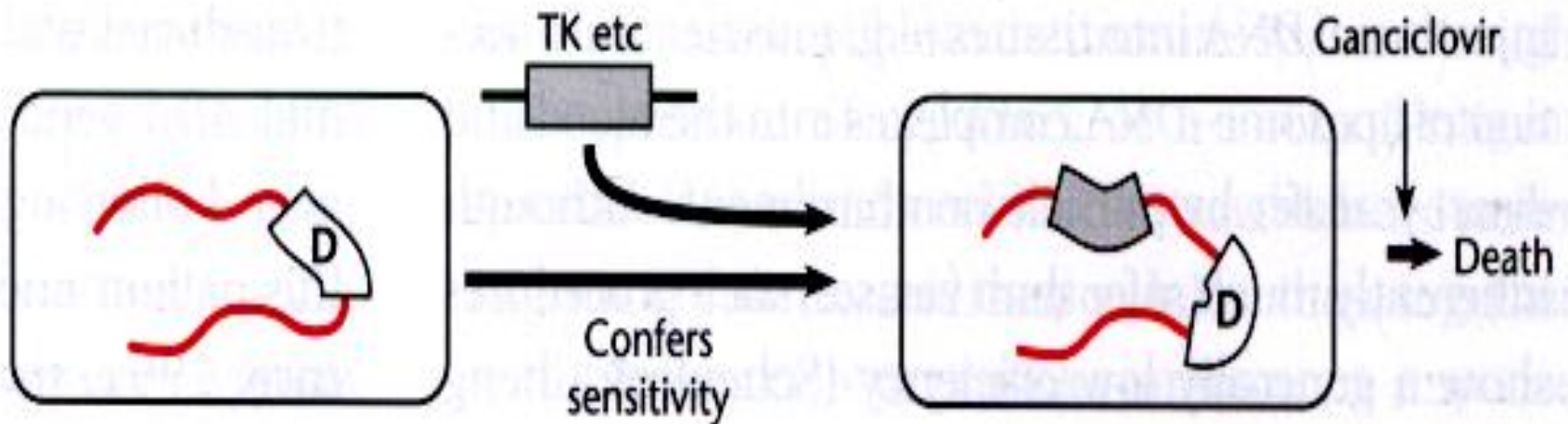


### 5 Prodrug therapy



## Assisted killing

## Prodrug - therapy



**Ganciclovir phosphorylation – GCV-triphosphate  
Incorporation to DNA and apoptosis induction**



# Genetic diseases

Type 1: A single locus (gene) - faulty - responsible for the disease, a hereditary 100%

Examples of sickle cell disease

hypercholesterolaemia

cystic fibrosis

Type 2: Polygenic <100% heritable, can depend on environmental factors and lifestyle

Examples

heart diseases

cancer

diabetes

alcoholism

schizophrenia

criminal behavior

# Single-gene diseases

Disease	Defect	Target cells
Severe combined immunodeficiency	Adenosin deaminase 4	Kostní dřeň nebo T-lymfocyty
Haemophilia	Factor VIII, Factor IX deficiency	Játra, svaly, fibroblasty
Cystic fibrosis	loss of CFTR gene	Lung
Haemoglobinopathy	Gene for $\alpha$ or $\beta$ globin	Bone marrow



# Polygenic diseases

<100% heritable

can depend on environmental factors and lifestyle

## **Examples:**

heart diseases

cancer

diabetes

alcoholism

schizophrenia

criminal behavior

**Cardiovascular diseases**

Atherosclerosis

endothelium

**Infectious diseases**

AIDS  
hepatitis B

T cells, macrophage  
Hepatocytes

Cirrhosis

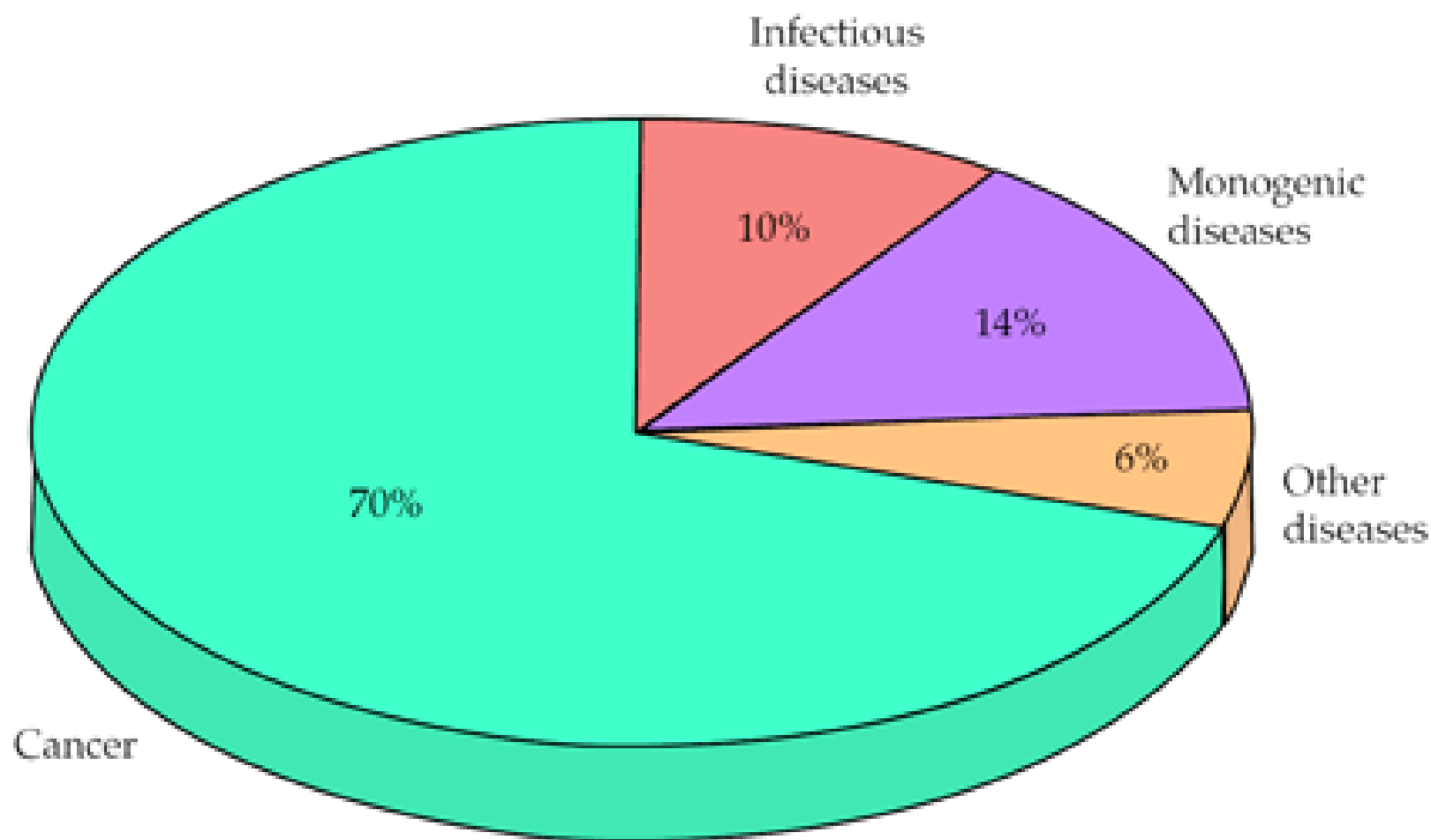
Fibrogenesis

Hepatic stellate cells

Autoimmune diseases

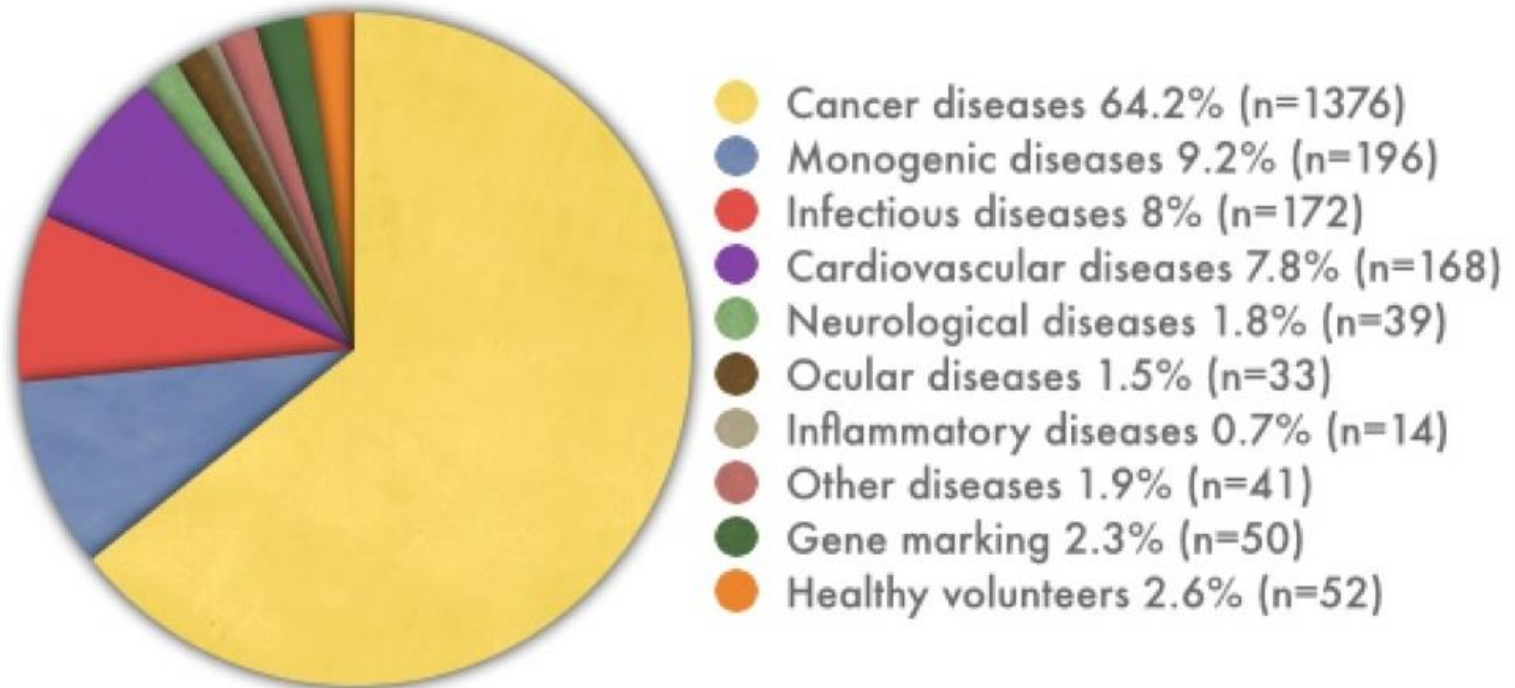
Lupus, diabetes

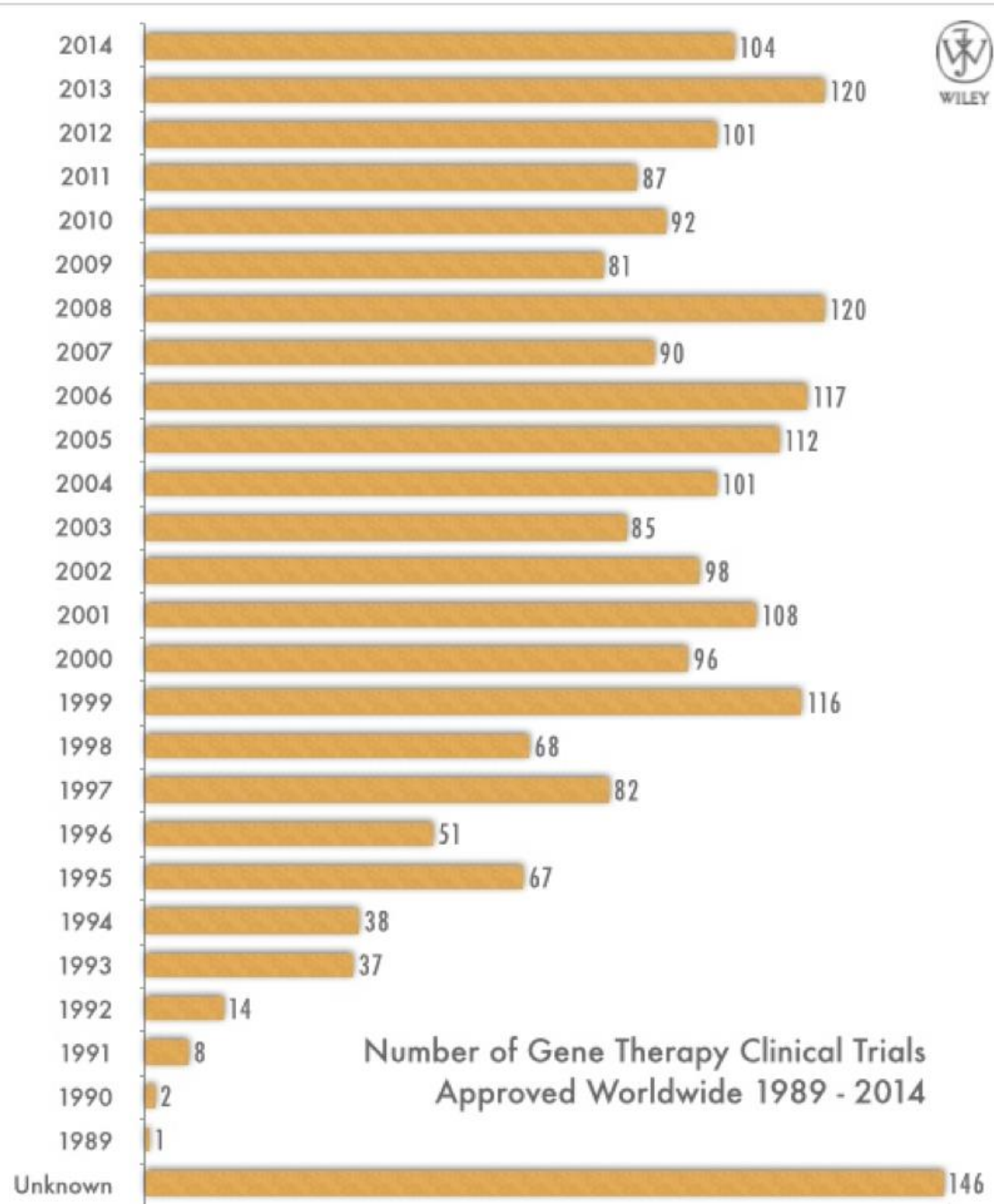
MHC



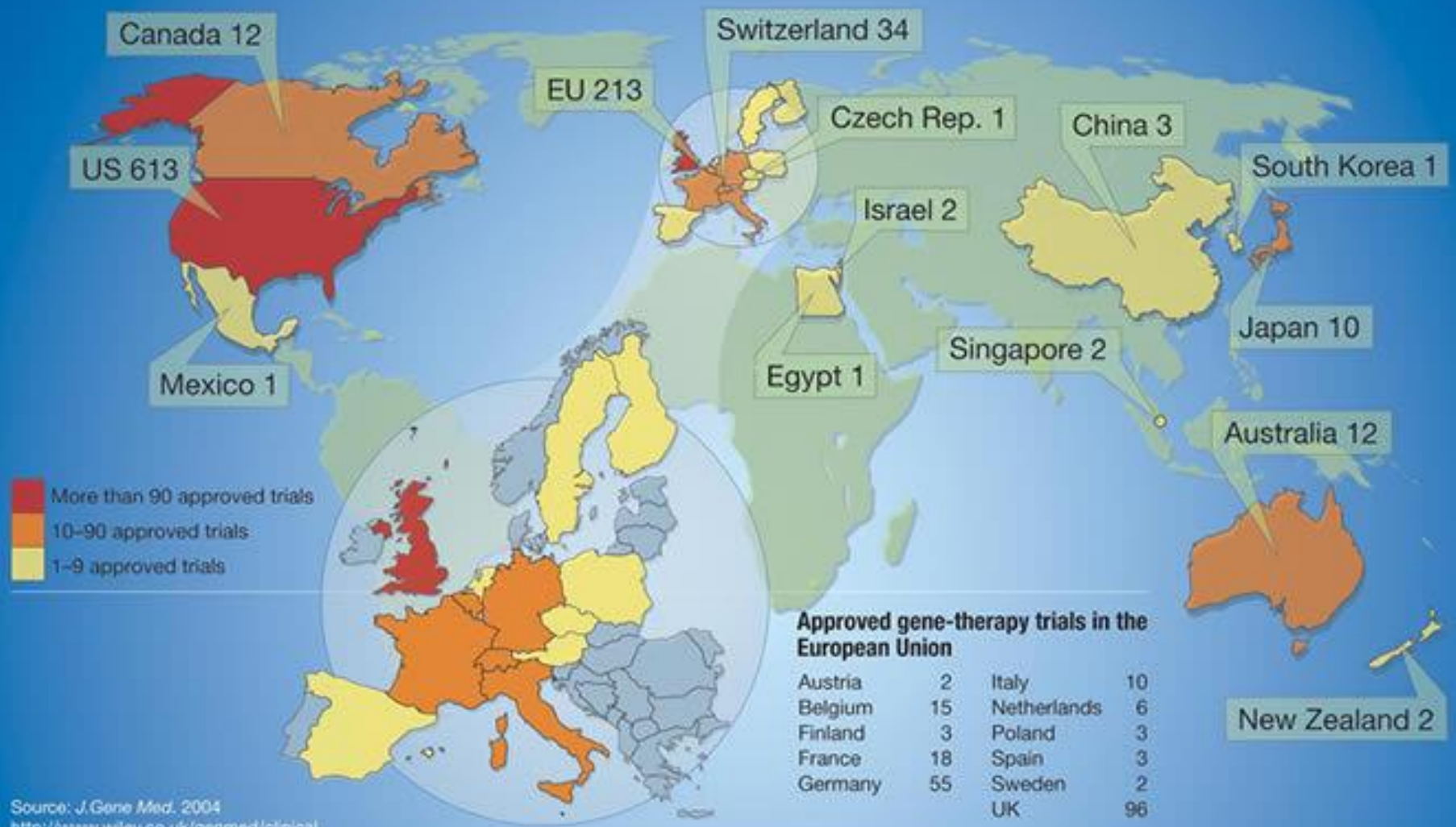
**Fig. 1 : Proportion of protocol for human gene therapy trials relating to various types of diseases<sup>52</sup>**

## Indications Addressed by Gene Therapy Clinical Trials





# Number of approved gene-therapy trials



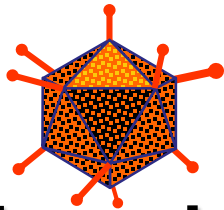
Source: J Gene Med, 2004  
<http://www.wiley.co.uk/genmed/clinical>

# Gene therapy principles

Exogenous DNA  
+ vector (viral)



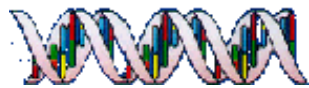
**AAV**



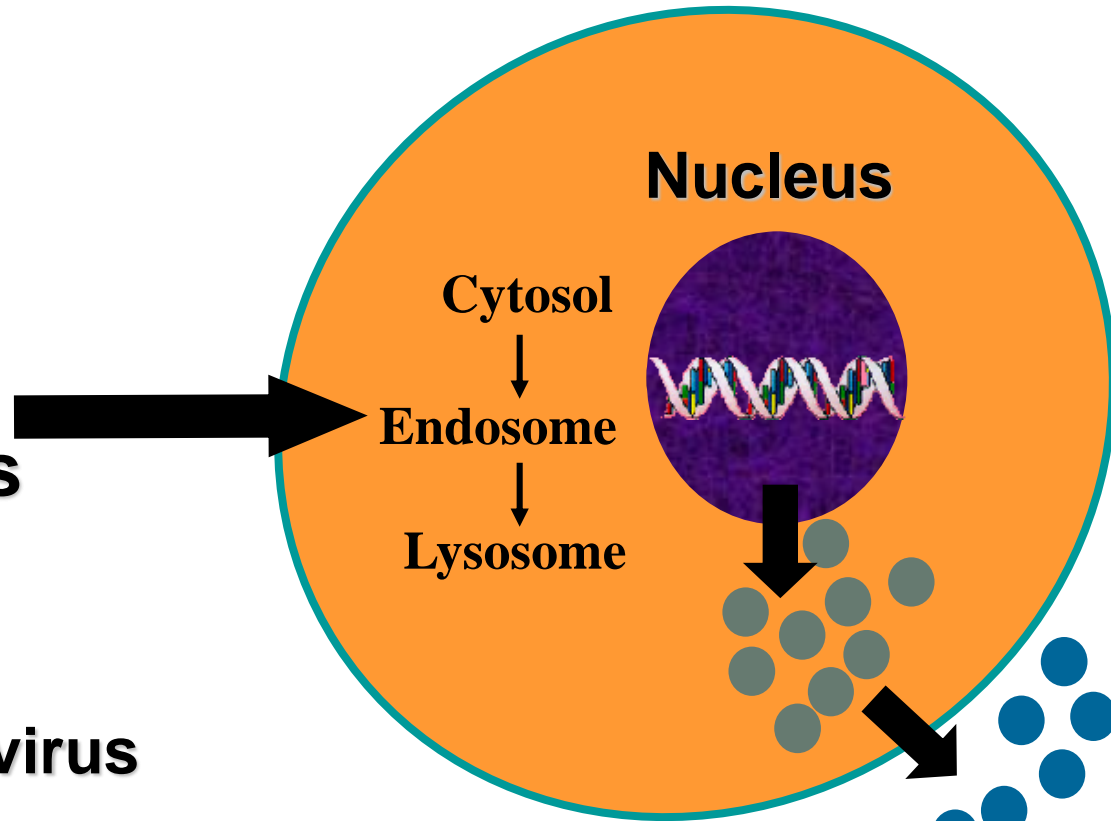
**Adenovirus**



**Retrovirus/Lentivirus**



**Isolated DNA**



**Target  
cell**

***Therapeutic  
protein***

# **Gene therapy hazard**

**Mutagenesis**

**Retroviral life cycle**

**– possible insertion to gene**

**Inactivation of tumor suppressor genes**

**Activation of protooncogene**



# Gene therapy – restoration of the immune system

## Severe Combined Immune Deficiency (SCID)

Fatal dysfunction of the immune system



Bubble boy, David



Commonly referred to a "bubble boy" disease, SCID, gets its moniker from David Vetter, a boy who lived out his 12 years of life in a germ-free plastic bubble.

# Severe Combined Immunodeficiency (SCID)

Recurrent infections diseases of bacterial, viral and fungal persistent including severe sepsis, meningitis, opportunistic pathogens - e.g. pneumonia

T lymphocytes reduced number of missing or weak proliferation in vitro

B lymphocytes absent or non-functional low levels of Ig after the disappearance of maternal IgG - lack of specific antibody response

Fatal without immune reconstitution of the immune system

# Severe Combined Immunodeficiency (SCID)

$\gamma$  chain ( $\gamma$ c) cytokine receptor deficiency (IL-2)  
blocks the differentiation of T and NK lymphocytes; no B- and T-cell functions

Autosomal recessive - homozygous for the defective ADA gene in T cells

Bubble boy, David, lived until removed from isolation after a bone marrow transplant to restore his immune system.

# SCID

**Linked to X chromosome (IL-2) - the most common form of SCID (males)**

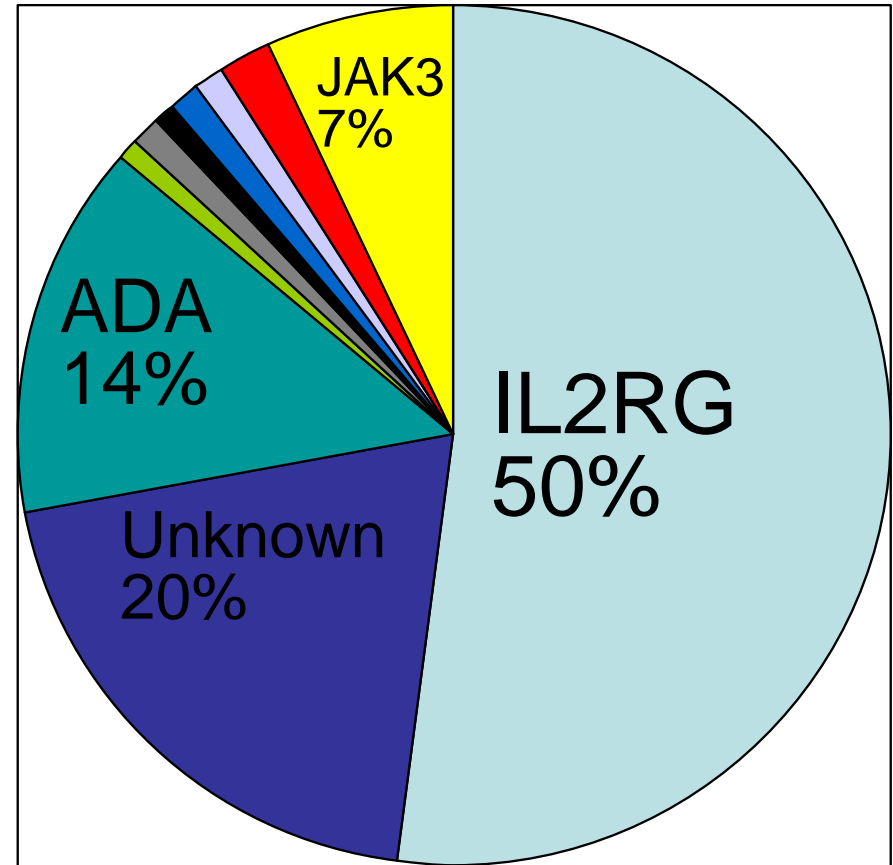
**Specific genetic defect = 80% of cases**

**Clinical application based on:**

**prenatal diagnosis**

**prediction of BMT response**

**gene therapy**

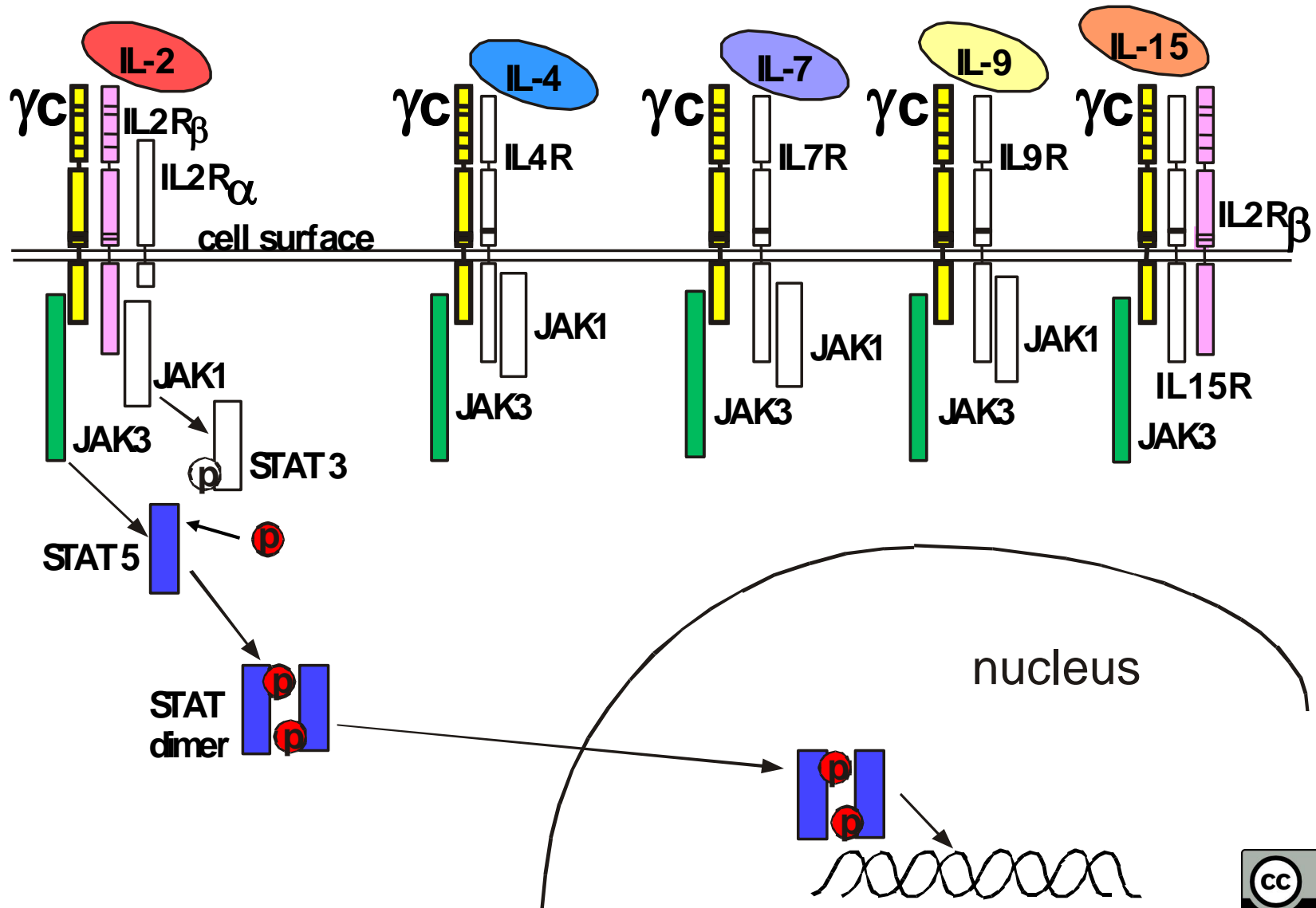


**Gene analysis 2002**

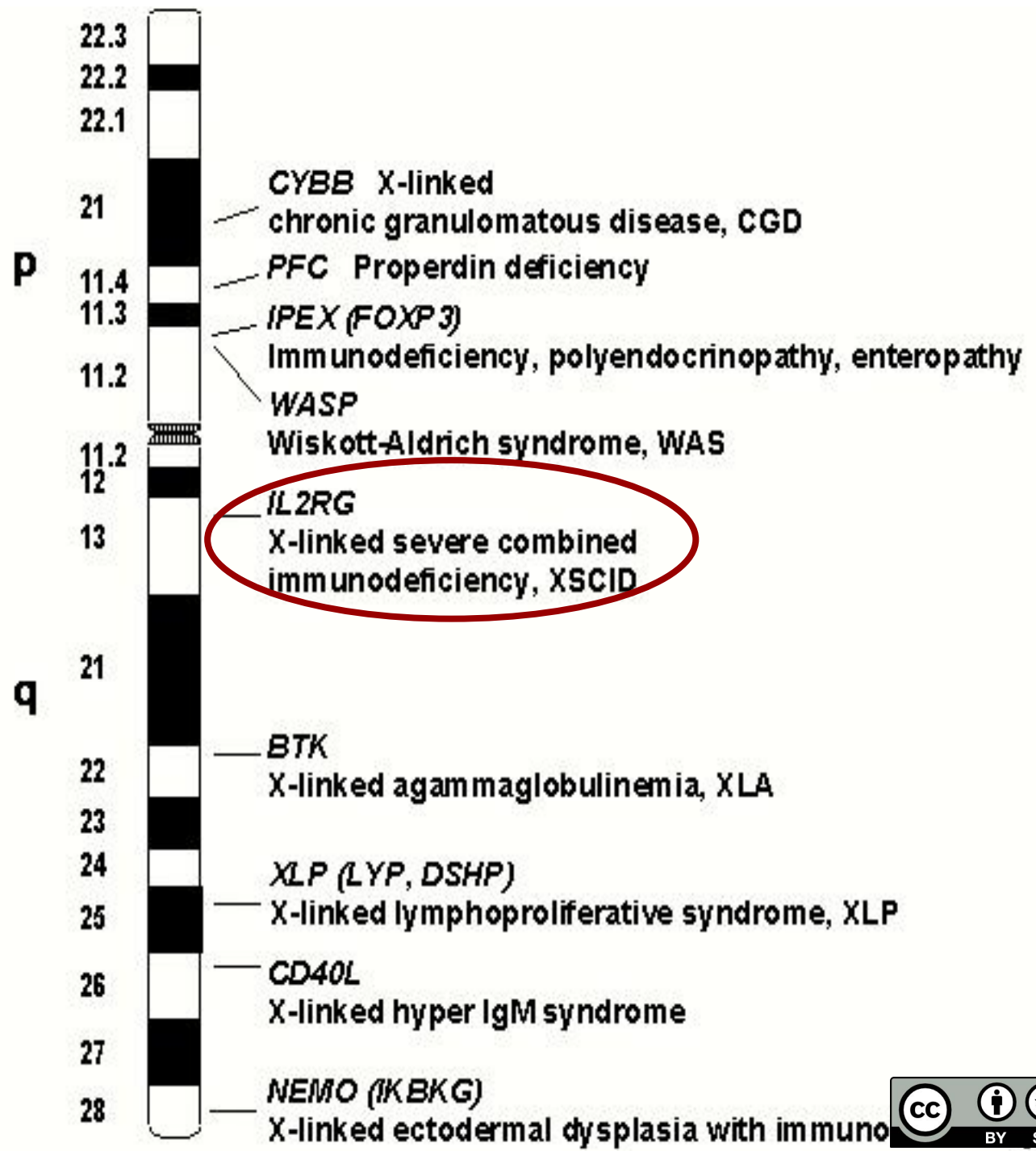
# SCID GENES

<b>IL2RG</b>	<b>50%</b>	<b>T-</b>	<b>B+</b>	<b>NK-</b>
<b>ADA</b>	<b>14%</b>	<b>T-</b>	<b>B-</b>	<b>NK-</b>
<b>JAK3</b>	<b>7%</b>	<b>T-</b>	<b>B+</b>	<b>NK-</b>
<b>IL7R<math>\alpha</math></b>	<b>7%</b>	<b>T-</b>	<b>B+</b>	<b>NK+</b>
<b>RAG1</b>	<b>&lt;5%</b>	<b>T-</b>	<b>B-</b>	<b>NK+</b>
<b>RAG2</b>	<b>&lt;5%</b>	<b>T-</b>	<b>B-</b>	<b>NK+</b>
<b>ARTEMIS</b>	<b>&lt;5%</b>	<b>T-</b>	<b>B-</b>	<b>NK</b>
<b>CD45</b>	<b>rare</b>	<b>T-</b>	<b>B+</b>	<b>NK+</b>

## *IL2RG* product: common $\gamma$ -chain ( $\gamma$ c) of cytokine receptor



# X-Linked Primary Immune Diseases



# **SCID - bone marrow transplantation (BMT)**

**Best results - BMT – 3rd month (even 1st) of age**

transplantation of hematopoietic stem cells to neonatal SCID patients – improved survival

**Best results - HLA-compatible twin**

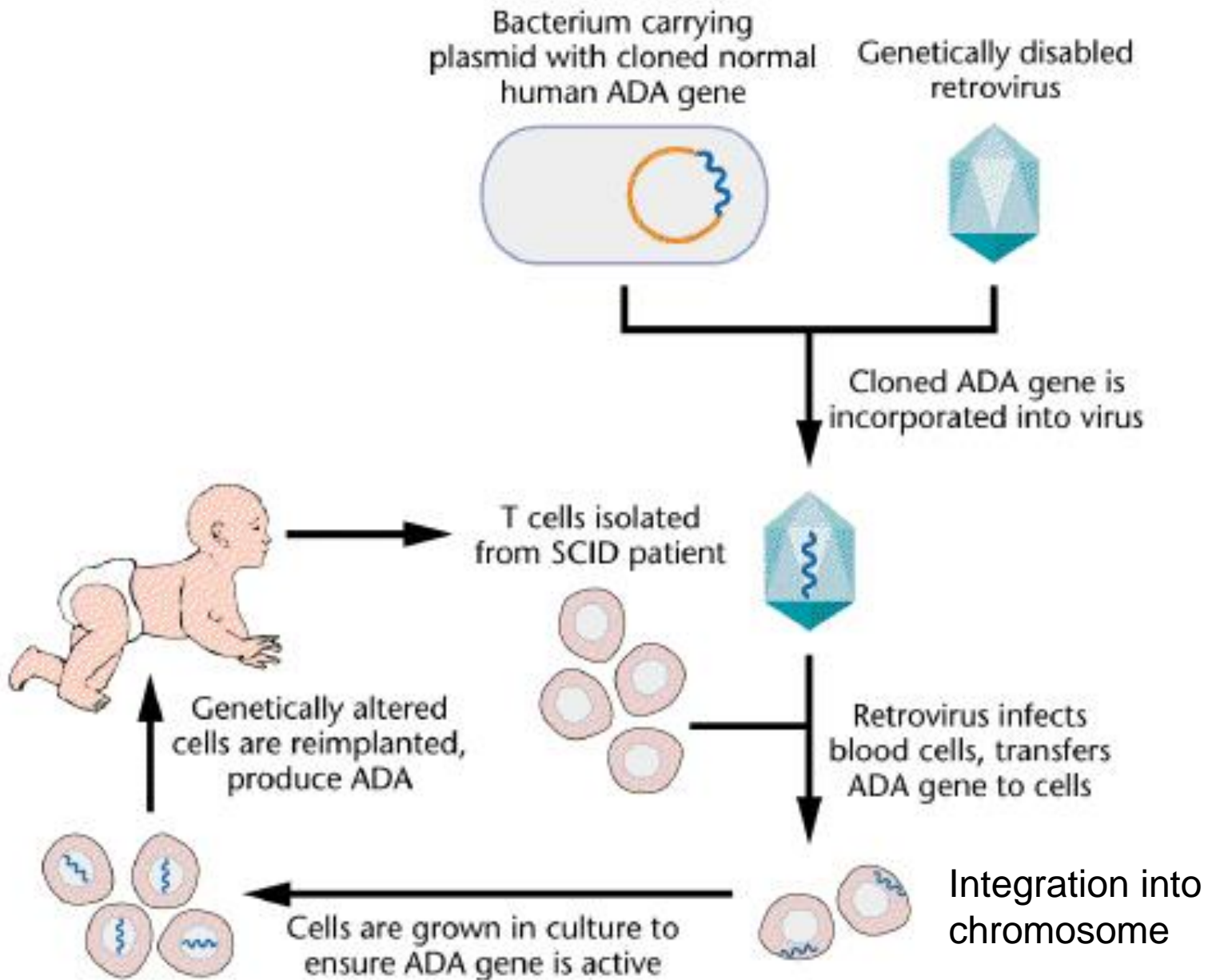
**Often only a partial resolution of disease**

Complete curation with BMT in some cases

the others – a continuous treatment with Ig



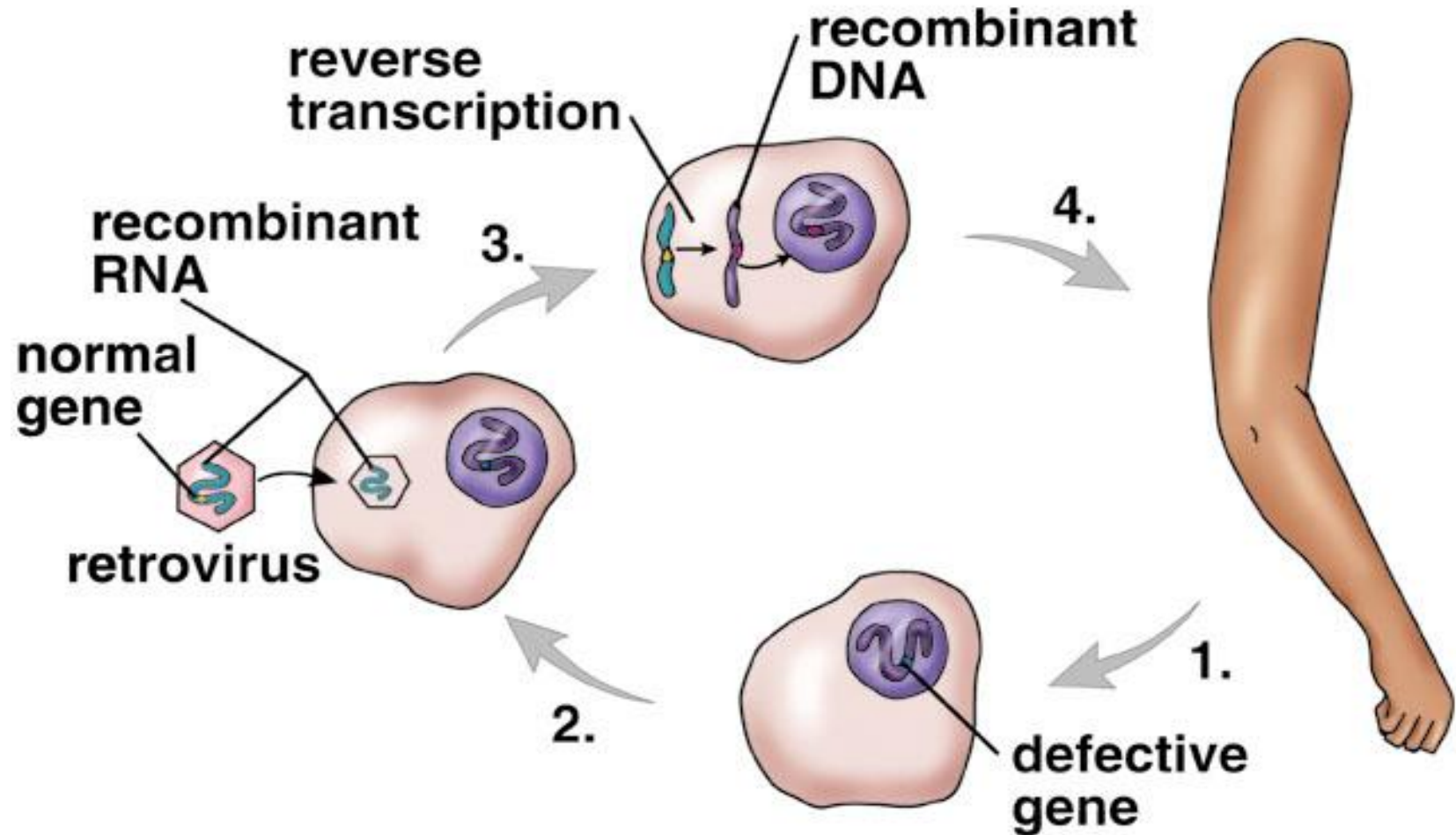
# Gene Therapy for SCID



# Gene Therapy – SCID

## (Severe Combined Immuno Deficiency)

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



**1st phase of clinical trial:**

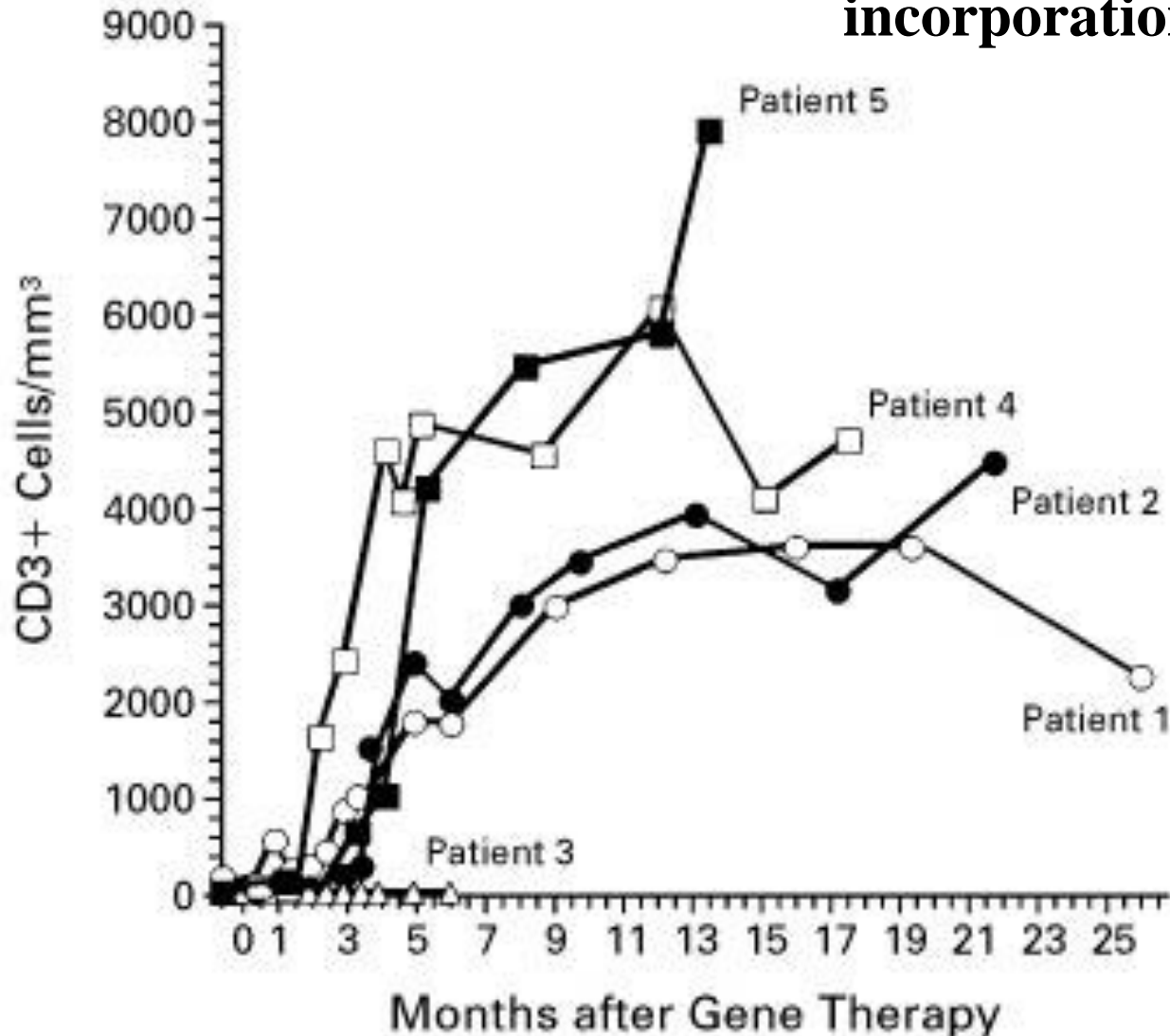
**Newborns: 14 - 26 milion CD<sup>34+</sup> cells/kg**

**~ 5 - 9 milion contained the therapeutic gene**

**Bone marrow reinfusion**

# X-Linked SCID – Clinical trial (11 patients)

Number of lymphocytes after  
incorporation of  $\gamma$ c gene



# **X-Linked SCID clinical trial**

**Follow up: 3 - 13 months  
of 10 from 11 patients**

**T cells  
NK cells  
B cells**



**Cell number  
normalization**

**Normal immune response to vaccination**

**Patients discharged**

# The first patient (SCID): Ashanti de Silva, treated at 4 (1990), is now 32



R. Michael Blaese, MD with Ashanthy DeSilva (left) and Cindy Kisik (both treated in 1990)  
at the IDF 2013 National Conference, June 29.



**Almost 3 years after GT, leukemia in 2 children**  
**Uncontrolled exponential proliferation of T cells**

**Gene insertion close to promotor of *LMO2* proto-oncogene**  
**- aberrant transcription and expression of *LMO2***

**Science 17<sup>th</sup> October 2003**

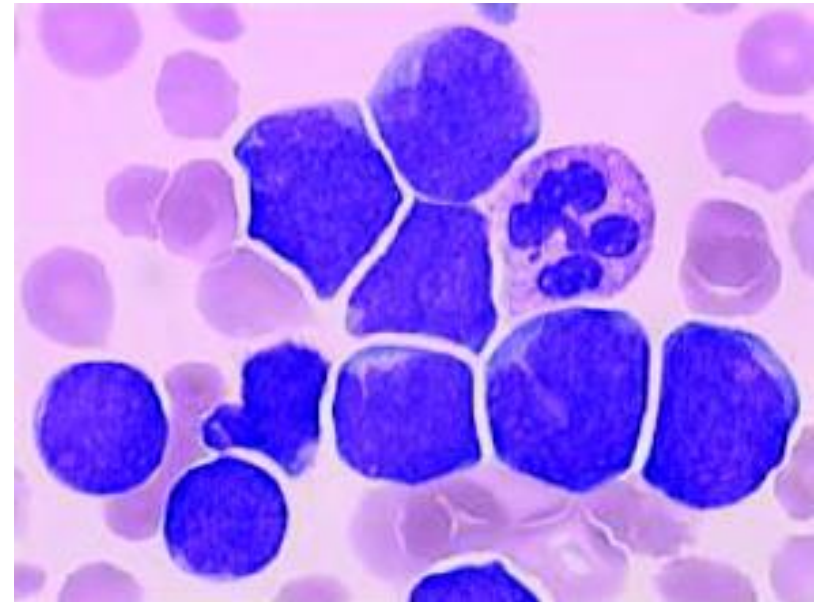
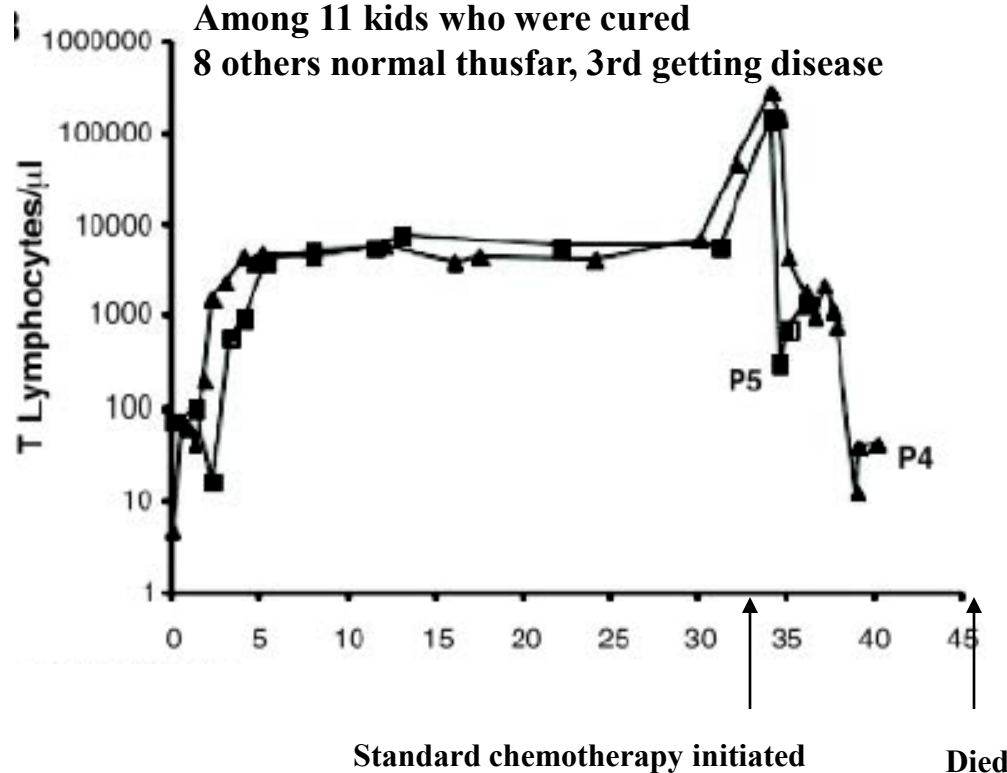
***LMO2*-Associated Clonal T Cell  
Proliferation in Two Patients  
after Gene Therapy for SCID-X1**

S. Hacein-Bey-Abina,<sup>1,2\*</sup> C. Von Kalle,<sup>6,7,8</sup> M. Schmidt,<sup>6,7</sup>  
M. P. McCormack,<sup>9</sup> N. Wulffraat,<sup>10</sup> P. Leboulch,<sup>11</sup> A. Lim,<sup>12</sup>  
C. S. Osborne,<sup>13</sup> R. Pawliuk,<sup>11</sup> E. Morillon,<sup>2</sup> R. Sorensen,<sup>19</sup>  
A. Forster,<sup>9</sup> P. Fraser,<sup>13</sup> J. I. Cohen,<sup>15</sup> G. de Saint Basile,<sup>1</sup>  
I. Alexander,<sup>16</sup> U. Wintergerst,<sup>17</sup> T. Frebourg,<sup>18</sup> A. Aurias,<sup>19</sup>  
D. Stoppa-Lyonnet,<sup>20</sup> S. Romana,<sup>3</sup> I. Radford-Weiss,<sup>3</sup> F. Gross,<sup>2</sup>  
F. Valensi,<sup>4</sup> E. Delabesse,<sup>4</sup> E. Macintyre,<sup>4</sup> F. Sigaux,<sup>20</sup> J. Soulier,<sup>21</sup>  
L. E. Leiva,<sup>14</sup> M. Wissler,<sup>6,7</sup> C. Prinz,<sup>6,7</sup> T. H. Rabbitts,<sup>9</sup>  
F. Le Deist,<sup>1</sup> A. Fischer,<sup>1,5†‡</sup> M. Cavazzana-Calvo<sup>1,2†</sup>

# *LMO2*-Associated Clonal T Cell Proliferation in Two Patients after Gene Therapy for SCID-X1

In addition, *Lmo2* transgenic mice were shown to develop T-ALL (28) within 10 months, despite the fact that the transgene expression was not restricted to T cells (29–32).

Among 11 kids who were cured  
8 others normal thusfar, 3rd getting disease





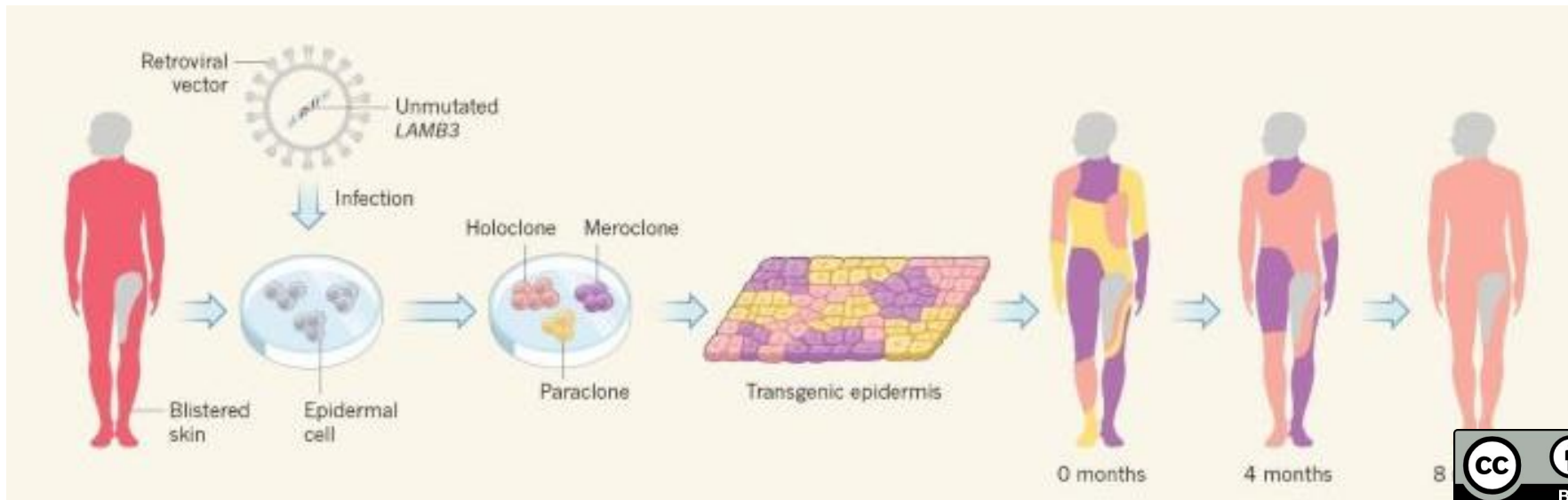
# Epidermolysis bullosa

Mutation: genes for keratine, plektine, laminin, collagen and integrin

2015 – repeated transplantation of skin, 80% of surface including the stomach

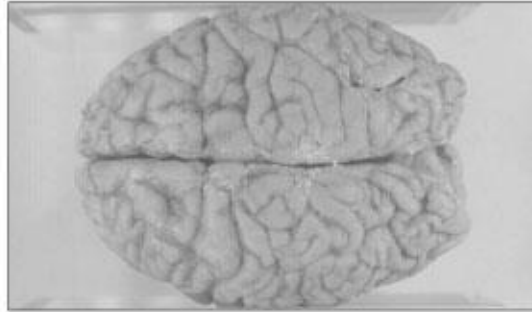
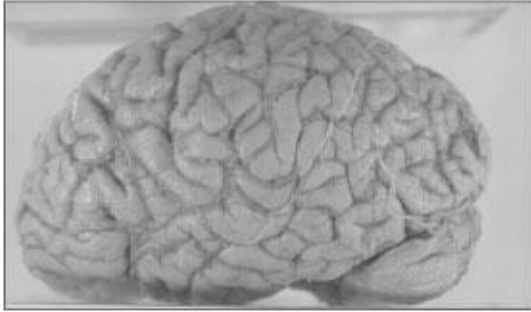
- modified keratinocytes from stem cells - producing laminin

Traditional treatment consists of bandages, ointments and pain medication



# Neurodegenerative diseases

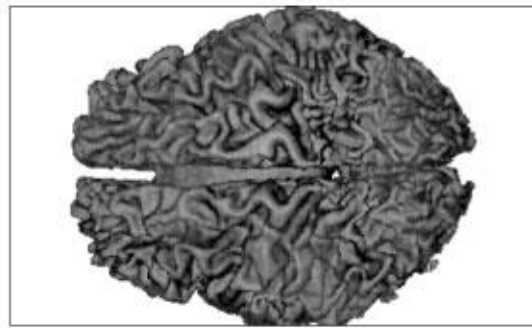
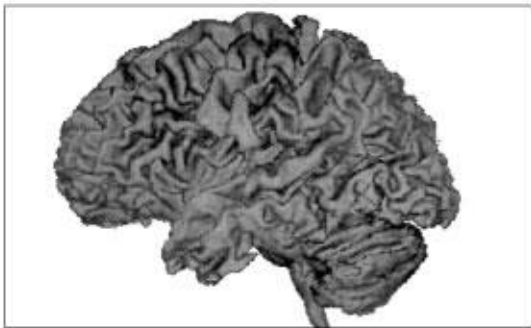
A. The brain of a normal elderly person



B. The brain of a person with Alzheimer's disease



C. The brain of a person with alcoholism



<http://www.niaaa.nih.gov/publications/arh25-4/254images/300.jpg>

[http://teachpol.tcnj.edu/amer\\_pol\\_hist/fi/000001ed.jpg](http://teachpol.tcnj.edu/amer_pol_hist/fi/000001ed.jpg)

# Parkinson disease

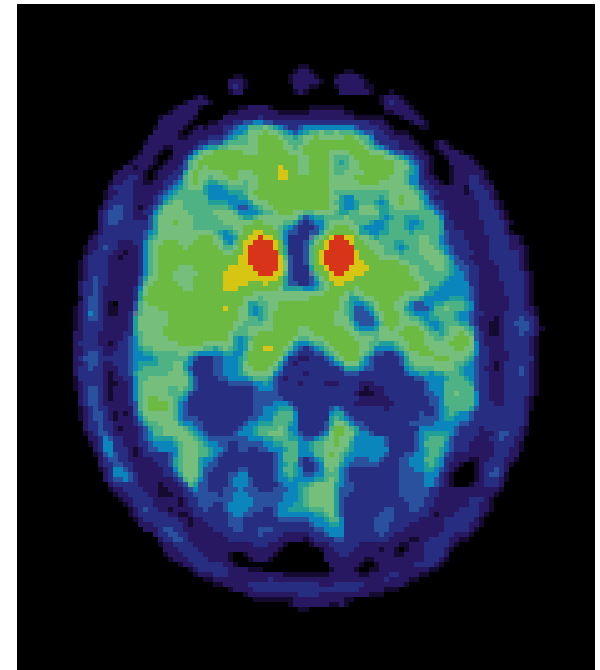
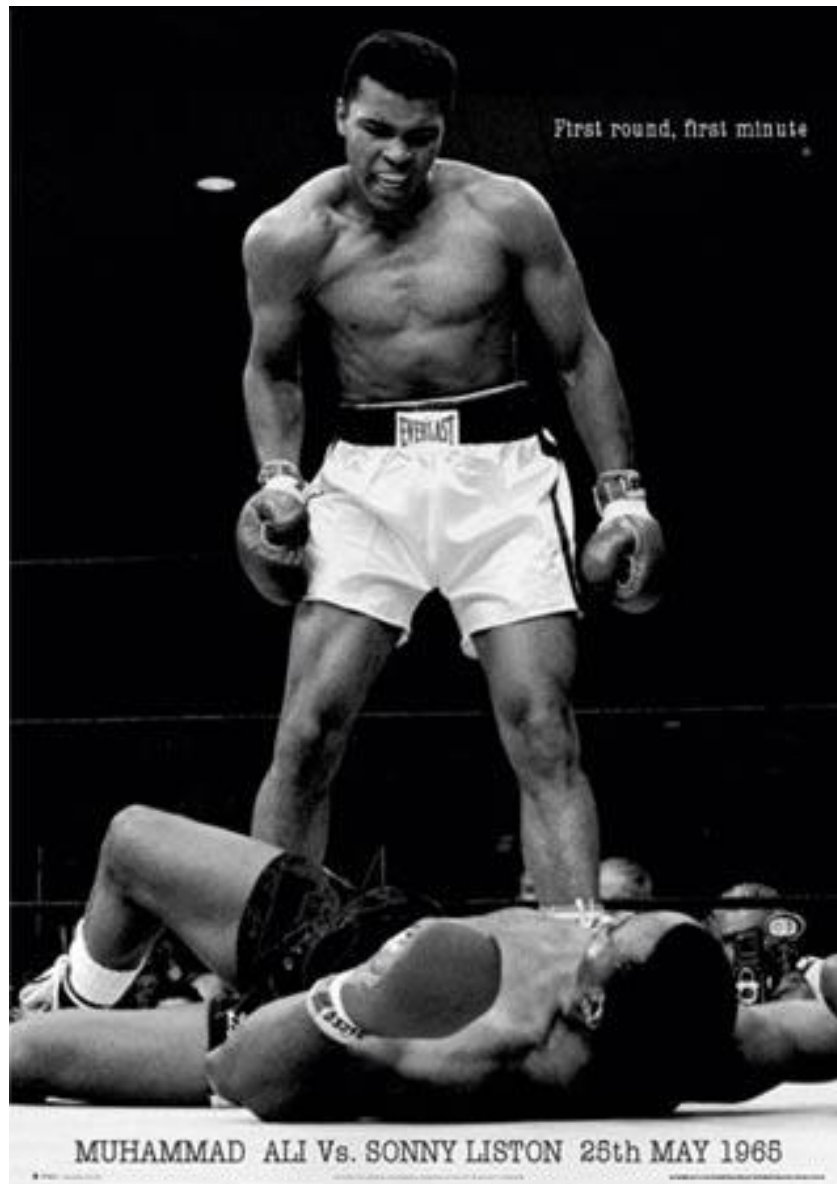
## – gene therapy



**1. Gene for a growth factor – neurogenesis**

**2. GAD (glutamic acid decarboxylase) gene insertion  
– responsible for the production of neurotransmitter GABA  
( $\gamma$ -amino butyrate) – control of movement**

# Parkinsonova choroba



**PET illuminates the  
Parkinsonian Brain**





## Parkinson disease

Loss of dopaminergic neurons

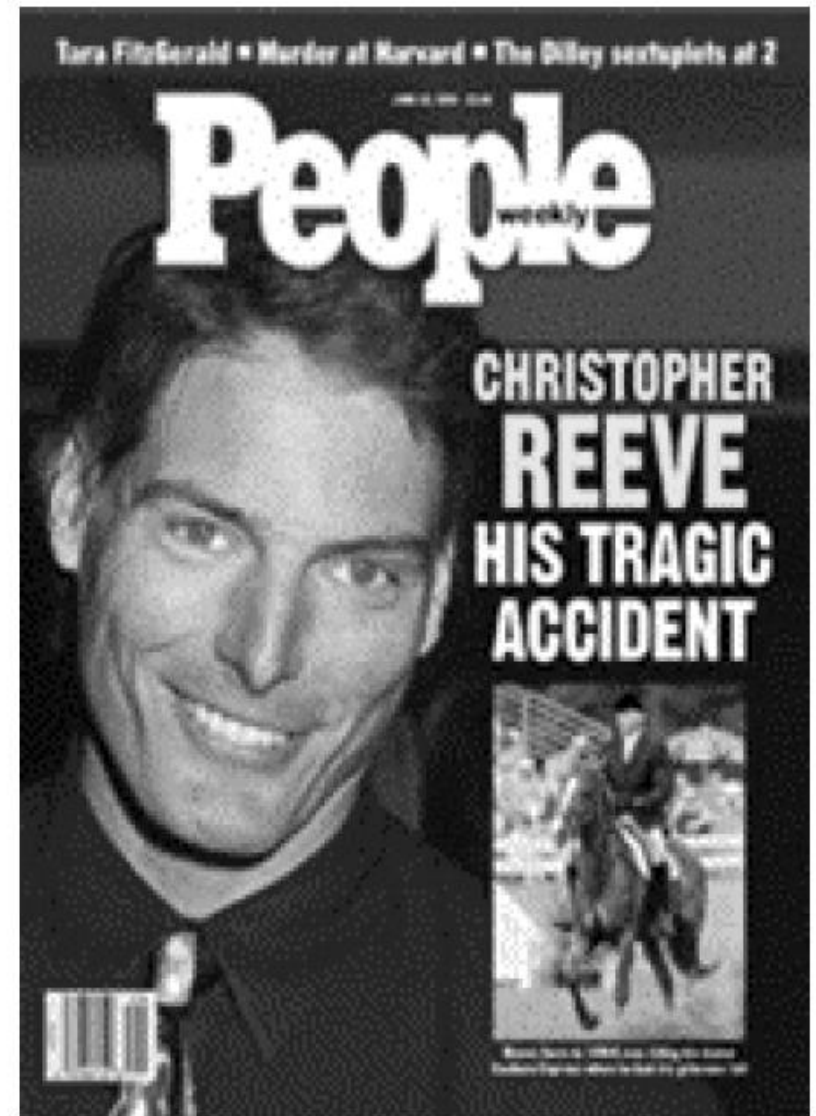
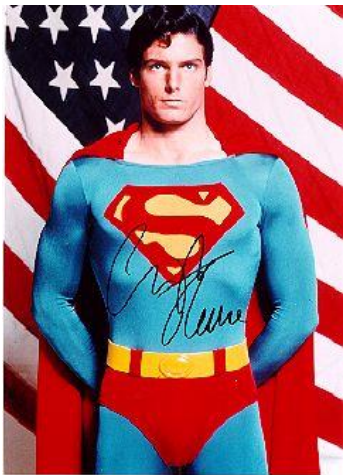
12 patients not responding to therapy for 5 years

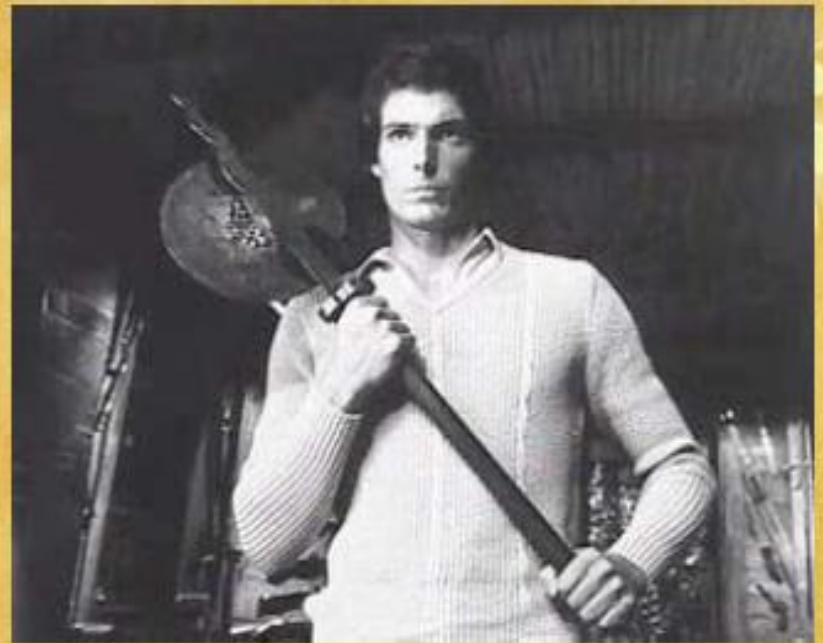
- insertion of gene for glutamate decarboxylase (GAD) to brain

Vector - adeno-associated virus (AAV)

After 1st year – improved motoric function by 27%

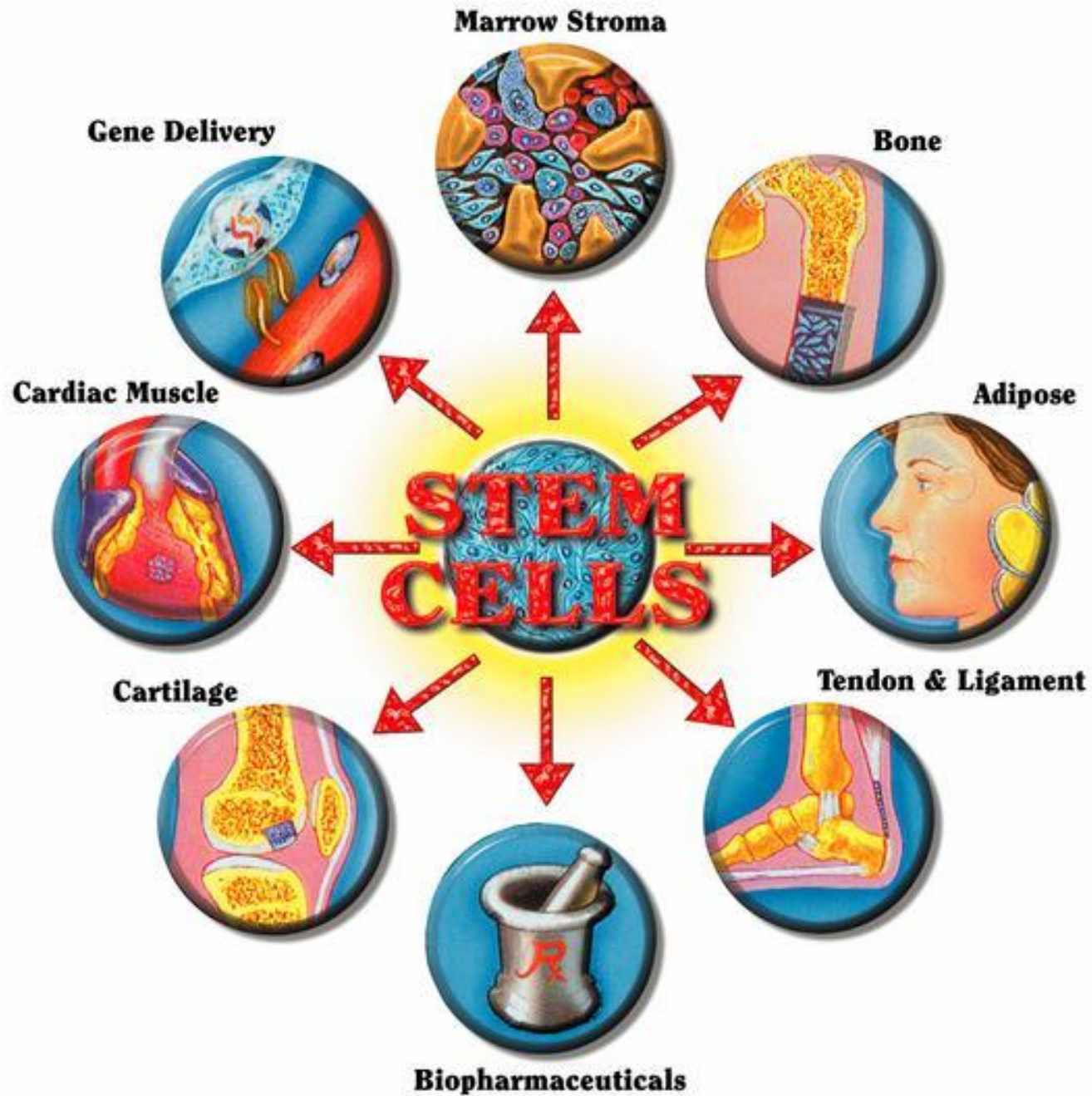
- ▶ 250,000 Americans with spinal cord injuries
- ▶ Approximately 11,000 new injuries occur each year in US.





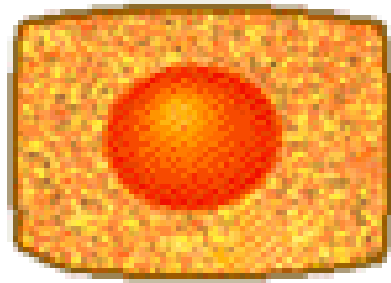
Stem cells





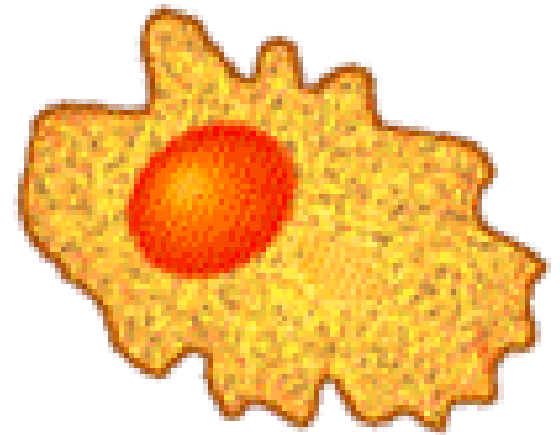


# Anticancer gene therapy



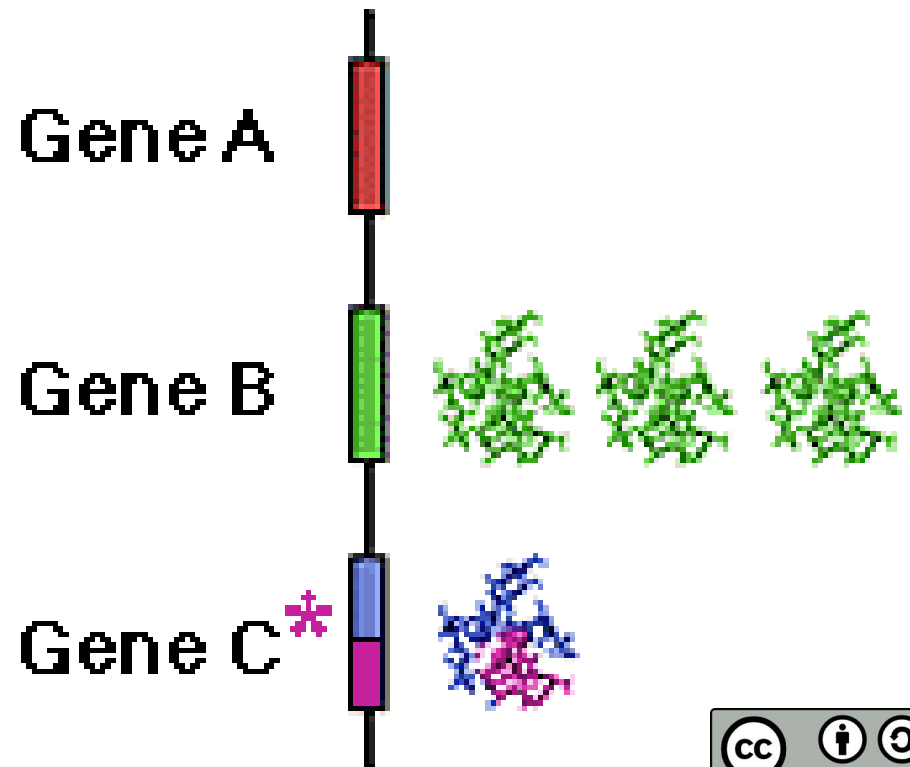
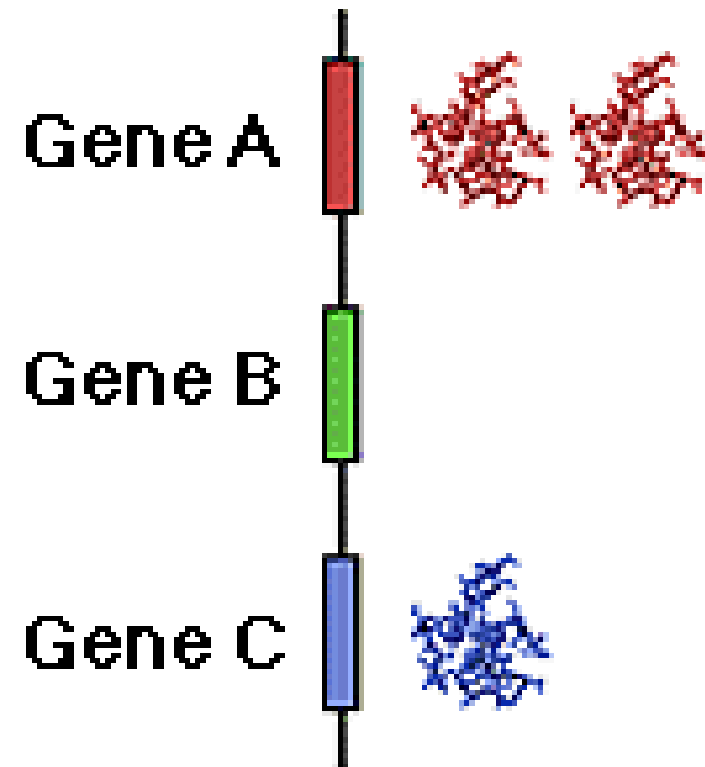
**Normal cell**

---



**Cancer cell**

---



**Which gene?**

**Which vector?**

**Which type of transfection?**

**Which target tissue?**

**Which model system?**

# Cancer inducing genes

## Oncogenes

Central regulators of the cell cycle

Oncogenic mutations establish a persistent growth signal

# DNA repair genes

**Maintain genetic stability**

**Mutations lead to hypermutability**

# Tumor suppressor genes

Negative regulators of cell growth

Loss of function - tumorigenesis

# Kontrolní body buněčného cyklu

**DNA damage**



**p53**



**Induction**

**p21**



Growth signal



cyclin/cdk



G1 phase



S phase

# Gene therapy of solid tumors

Cancer – polygenic disease

Change in a single gene can stop the tumor growth

- Gene replacement
- Gene knockout
- Suicide gene therapy
- Immunomodulatory gene therapy



# Strategies for anticancer gene therapy

**Decreased expression of oncogenes**

**Introducing tumor suppressor e.g. p53, Rb**

**Dominant negative proteins, growth factors / receptors**

**Induction of apoptosis / increased apoptosis: e.g. induction BAX inhibition of anti-apoptotic pathways: e.g. BCL-2**

**Sensitization to radiation and chemotherapy**

**Prodrug therapy**

**Others: increased frequency of DNA breaks, decreased DNA repair**

# **Dominant-negative proteins**

- Growth factors / receptors**

## **Induction of apoptosis**

- e.g. induction of BAX (pro-apoptotic factor), caspases**
- inhibition of anti-apoptotic pathways: e.g. BCL-2**

# Anti-angiogenic strategy

**> 40 natural anti-angiogenic molecules**

**e.g. endostatin, angiostatin**

**- expression of soluble receptors of angiogenic factors**

**- *partial elimination before arrival to its target***

**- antisense RNA**

**- VEGF - vascular endothelial growth factor**

**- FGF - fibroblast growth factor**

**Also the opposite attempts – treatment of cardiovascular diseases**

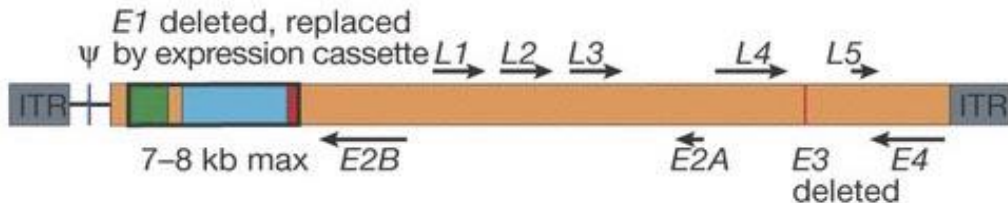
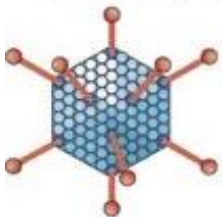
# Vector types

RNA viruses (retroviruses), Murine leukemia virus (MMLV),  
Human Immunodeficiency Virus (HIV) Human T-cell  
lymphotropic virus (HTLV)

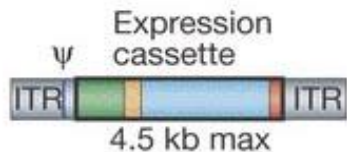
DNA viruses, adenovirus-adenovirus associated virus (AAV),  
herpes simplex virus (HSV), pox virus Foamy virus

Non-viral vectors Naked DNA Liposomes The liposome-  
polycation complexes Peptide Systems

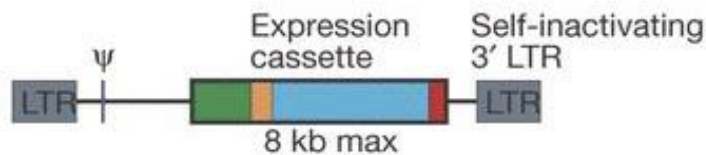
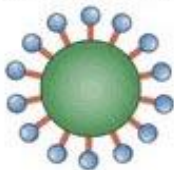
Adenovirus (~36 kb genome)



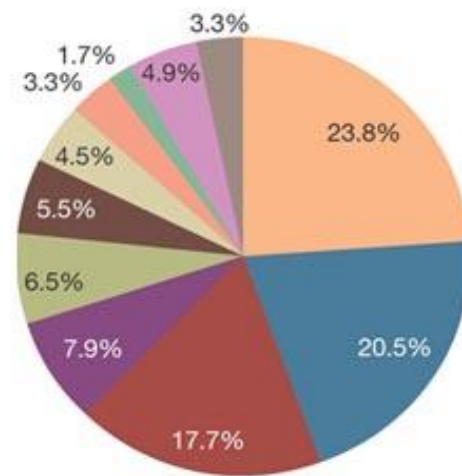
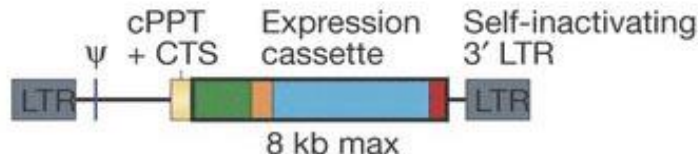
Adeno-associated virus (4.7 kb genome)



Retrovirus (7-10 kb genome)



Lentivirus (9-10 kb genome)

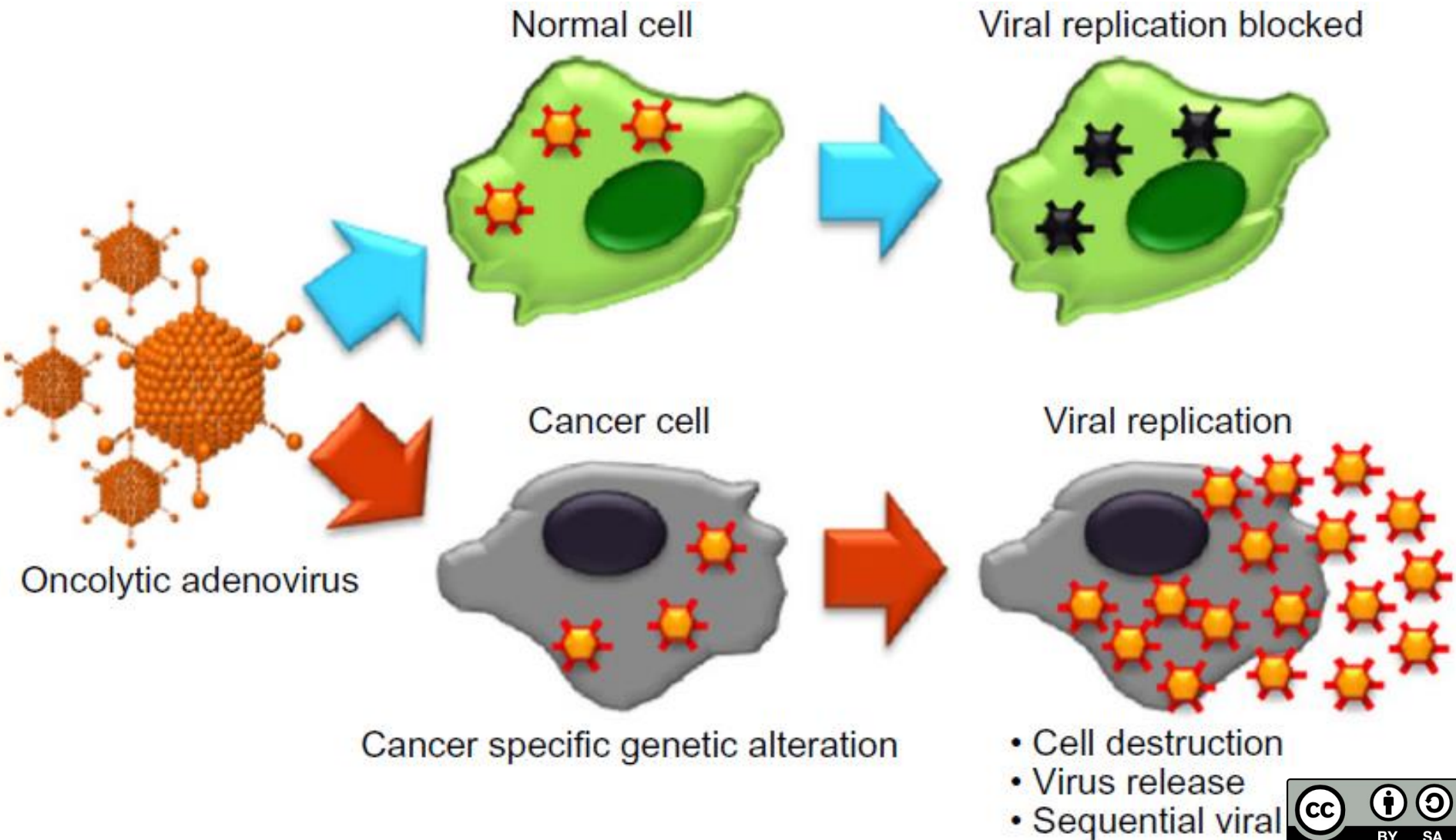


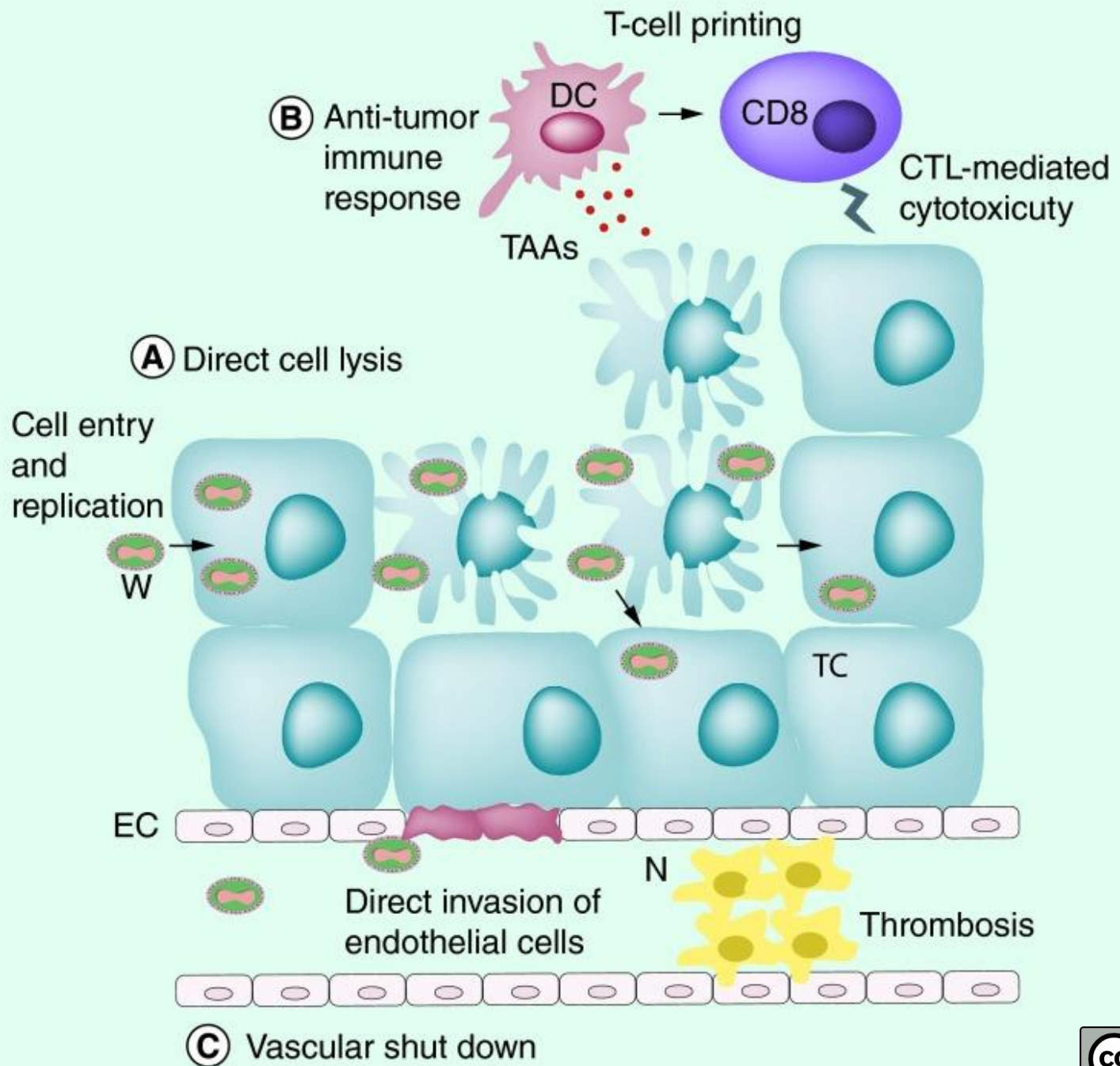
- Adenovirus (n = 400)
- Retrovirus (n = 344)
- Naked/plasmid DNA (n = 304)
- Vaccinia virus (n = 133)
- Lipofection (n = 109)
- Poxvirus (n = 93)
- Adeno-associated virus (n = 75)
- Herpes simplex virus (n = 56)
- Lentivirus (n = 29)
- Other categories (n = 82)
- Unknown (n = 55)

# Oncolytic viral therapy

## Oncolytic virus

- propagation more effective in cancer cells compared to normal cells
- modified viruses with defective genes







# Vaccinia virus (VV)

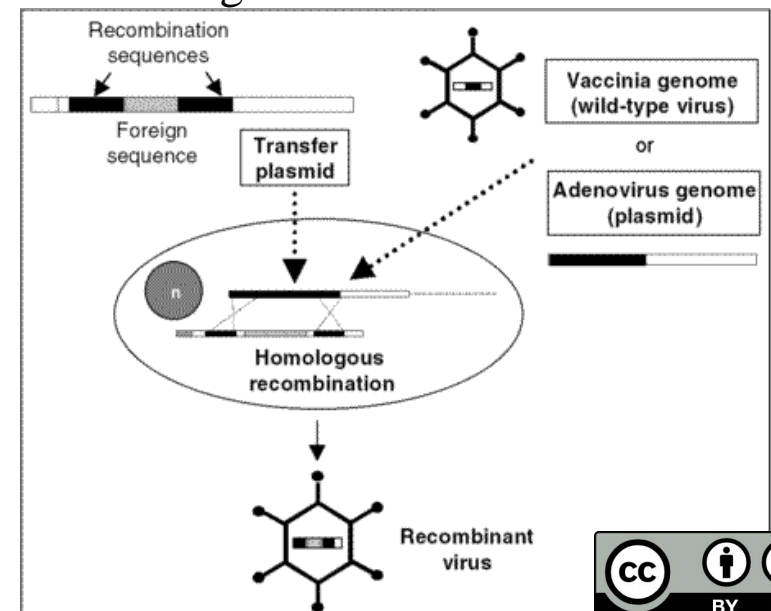
dsDNA virus -192 kbp – up to 25 kbp exogenous DNA

Short life cycle – 8 h in the cytoplasm – no risk of integration into genome. Replication 2 h after infection

own RNA polymerase – virus less dependent on the host apparatus

epidermal growth factor receptor (EGFR) and other proteins overproduced in tumor cells support replication of VV

viral antigens released into tumor – strong inflammatory response – overcoming of an immune escape of cancer. Moreover, a tumor cell lysis releases tumor antigens.





**a**

Virus



H-1PV  
naked, DNA,  
5 kb



Adenovirus  
naked, DNA,  
36–38 kb



HSV1  
enveloped, DNA,  
120–200 kb



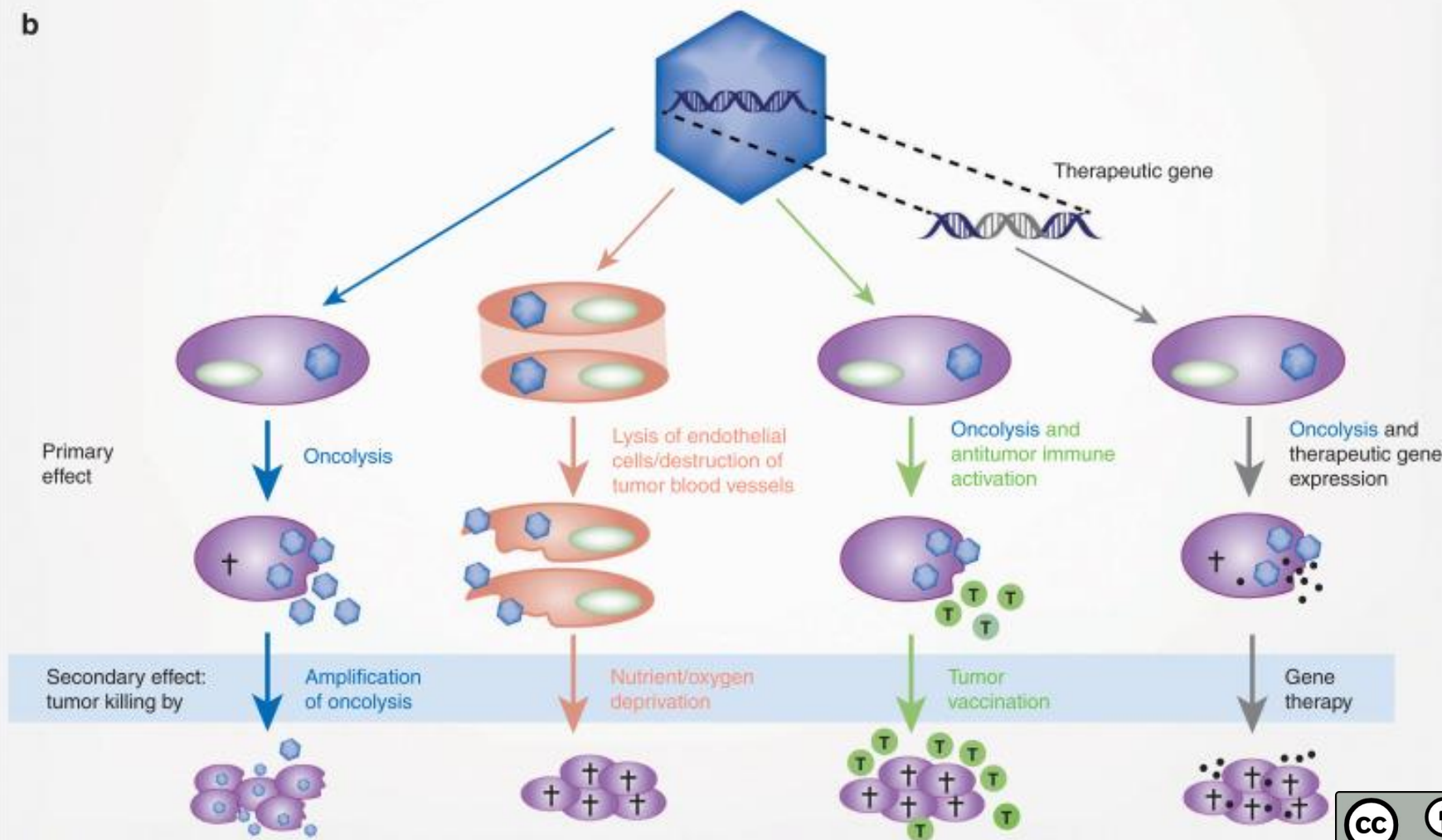
Vaccinia virus  
enveloped, DNA,  
130–280 kb



Reovirus  
naked, RNA,  
22–27 kb

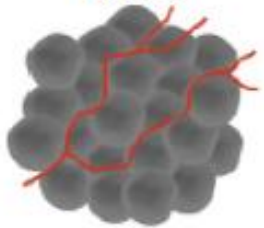


Measles virus  
enveloped, RNA,  
16–20 kb

**b**

# Optimalisation of therapeutic effect of oncolytic viruses

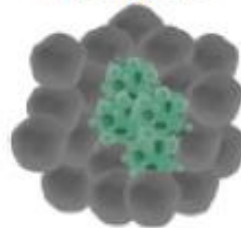
Delivery to the tumor



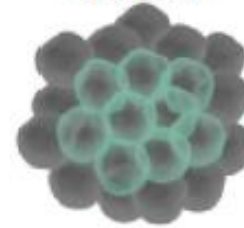
Virus infectivity



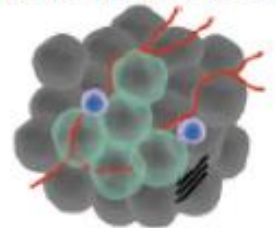
Virus oncolysis



Virus spread



Tumor microenvironment



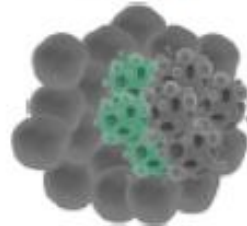
Tumor heterogeneity



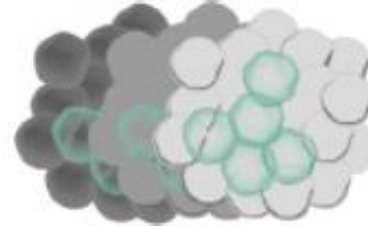
Virus productivity



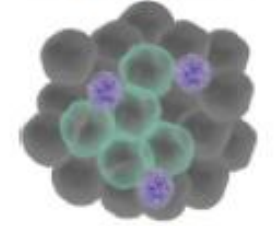
Bystander killing



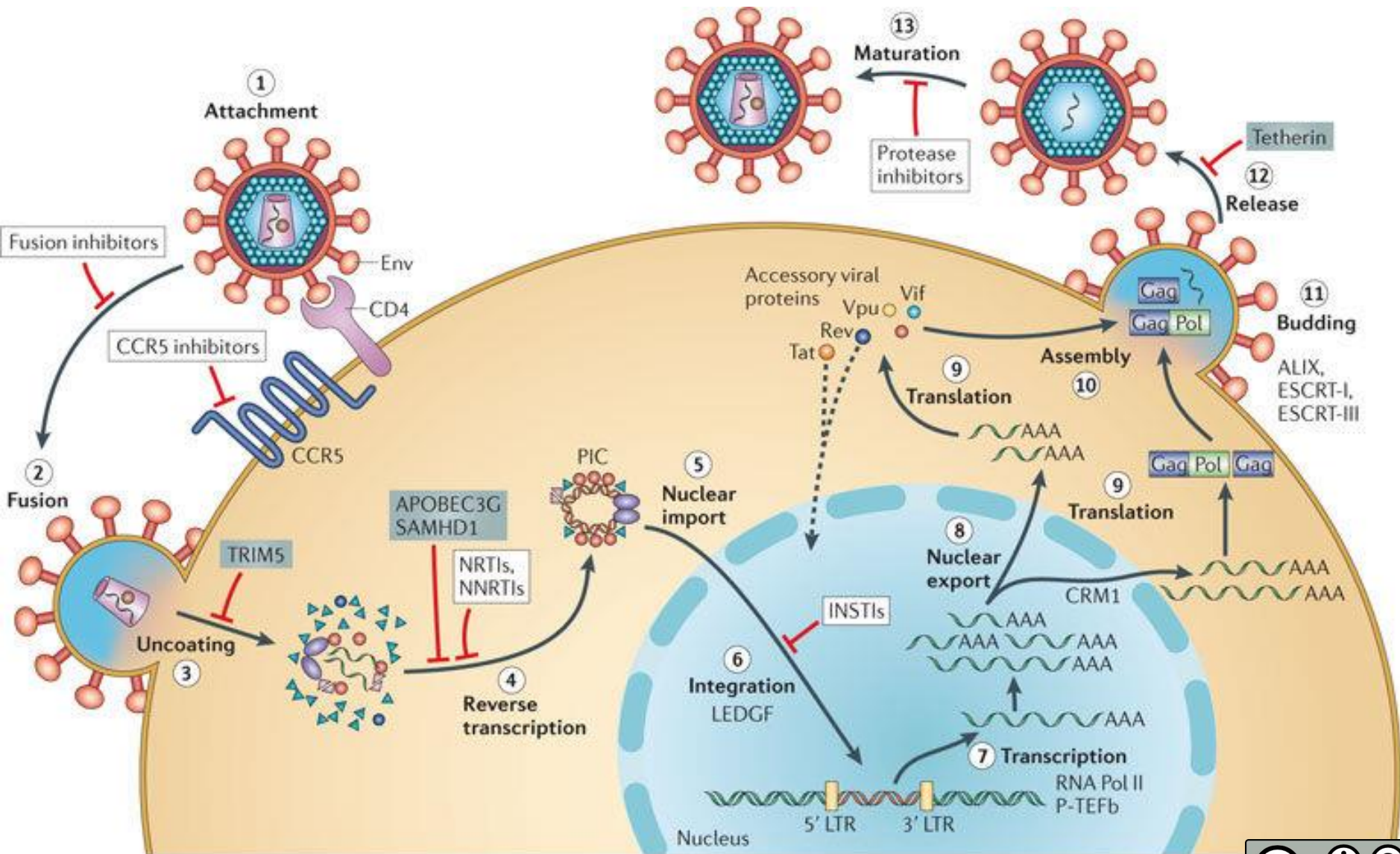
Virus persistence



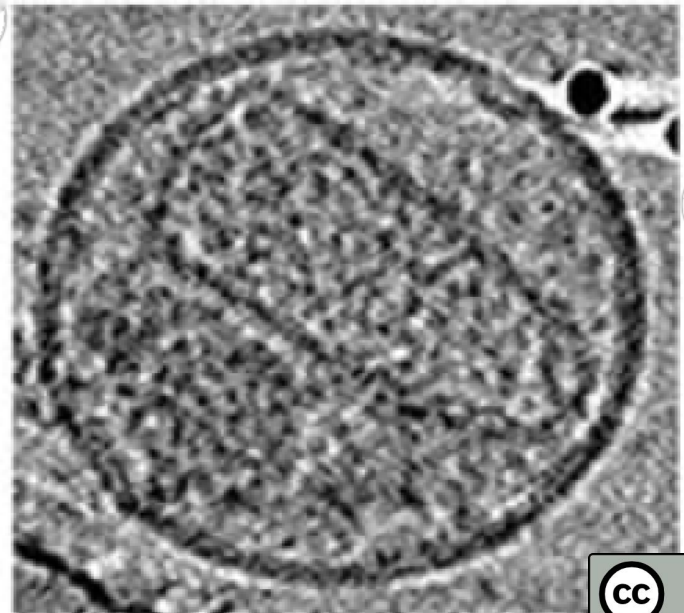
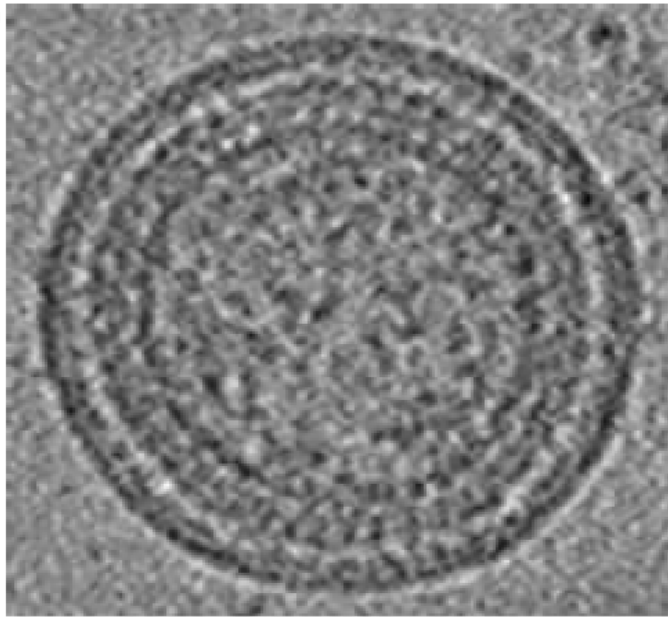
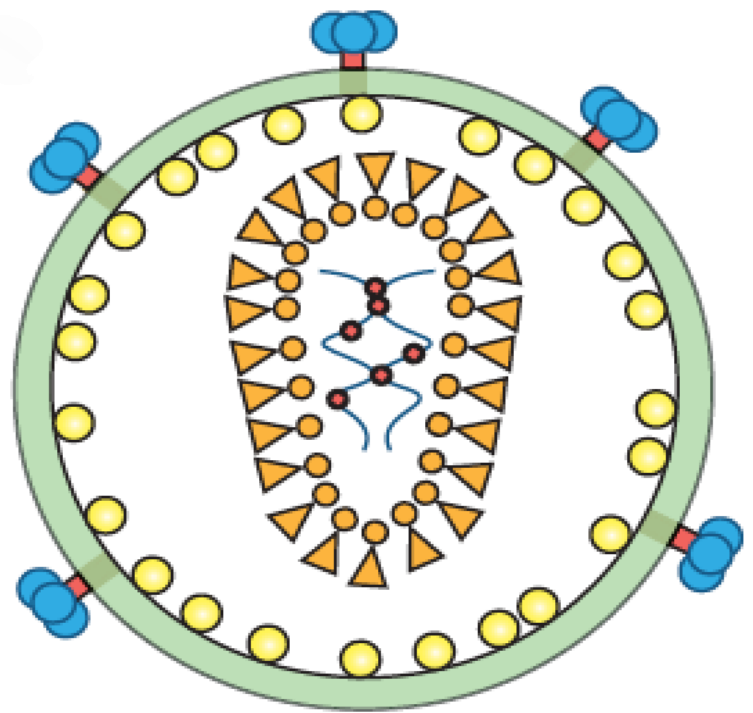
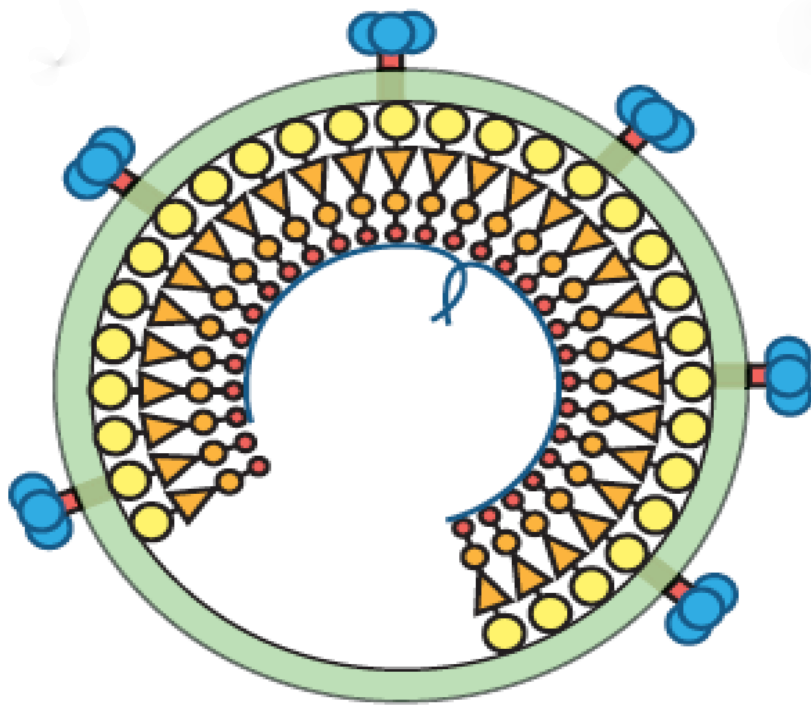
Immune engagement



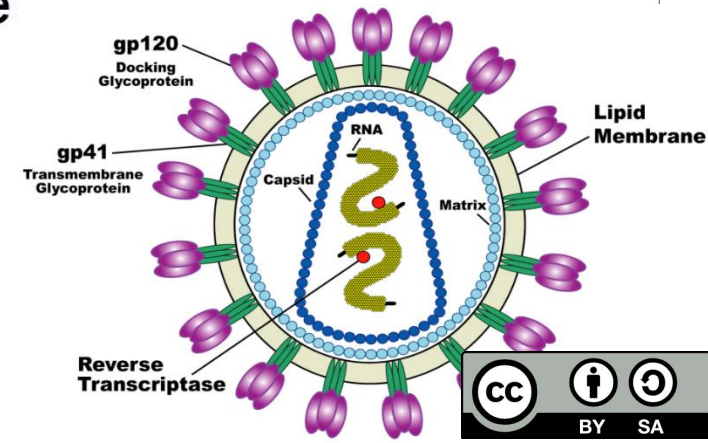
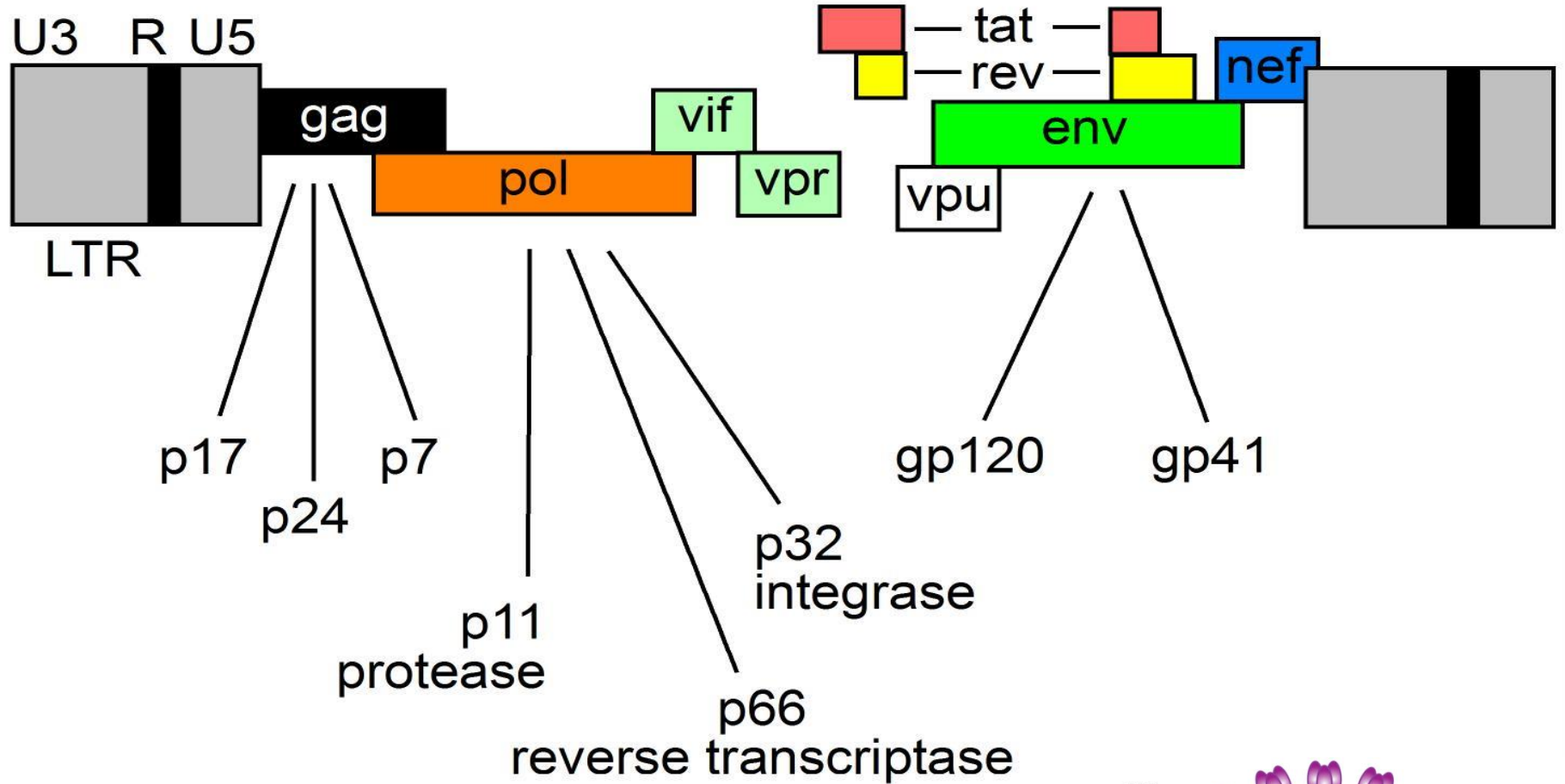
# Retrovirus life cycle



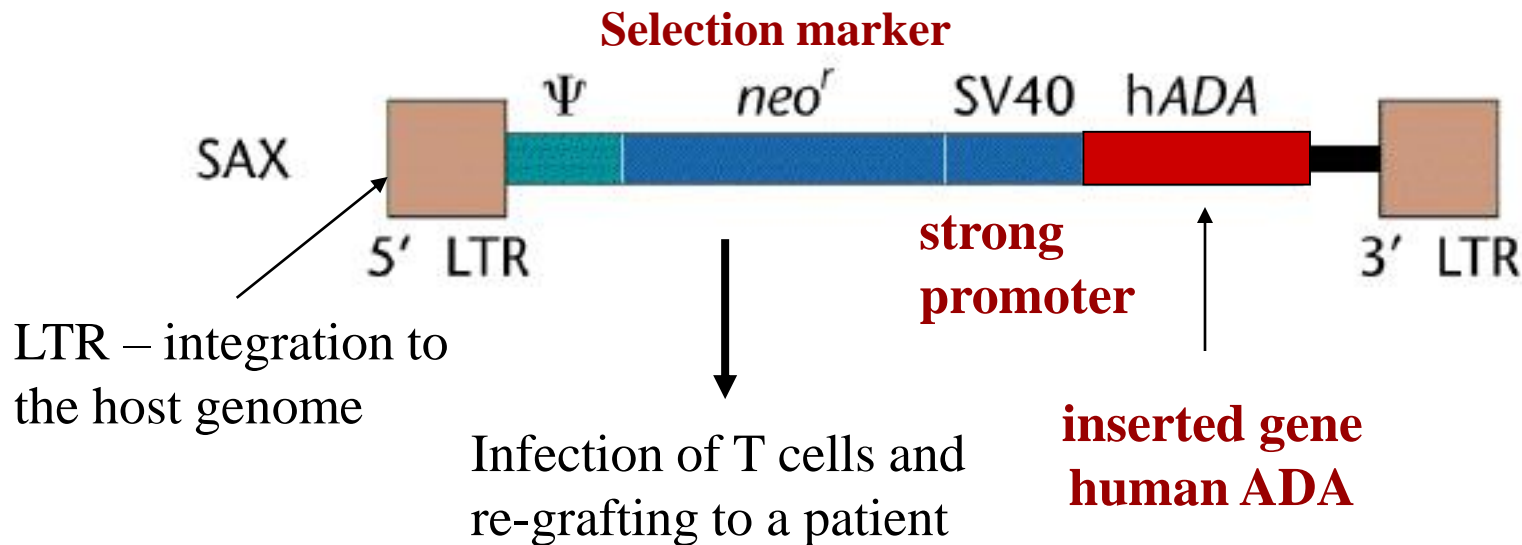
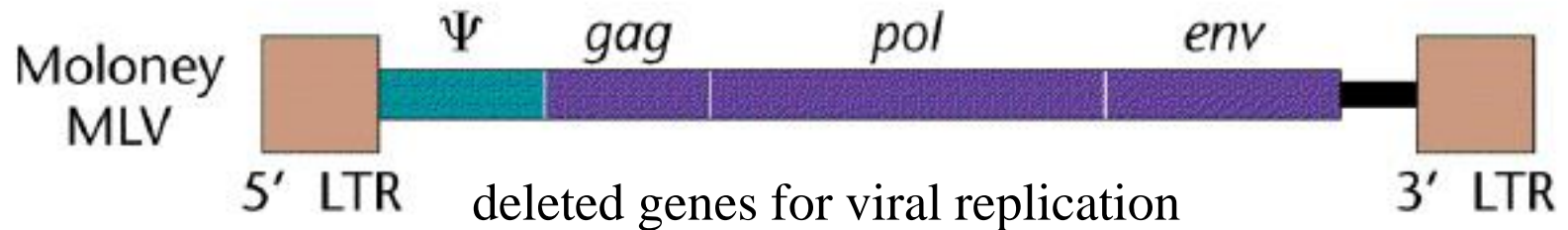




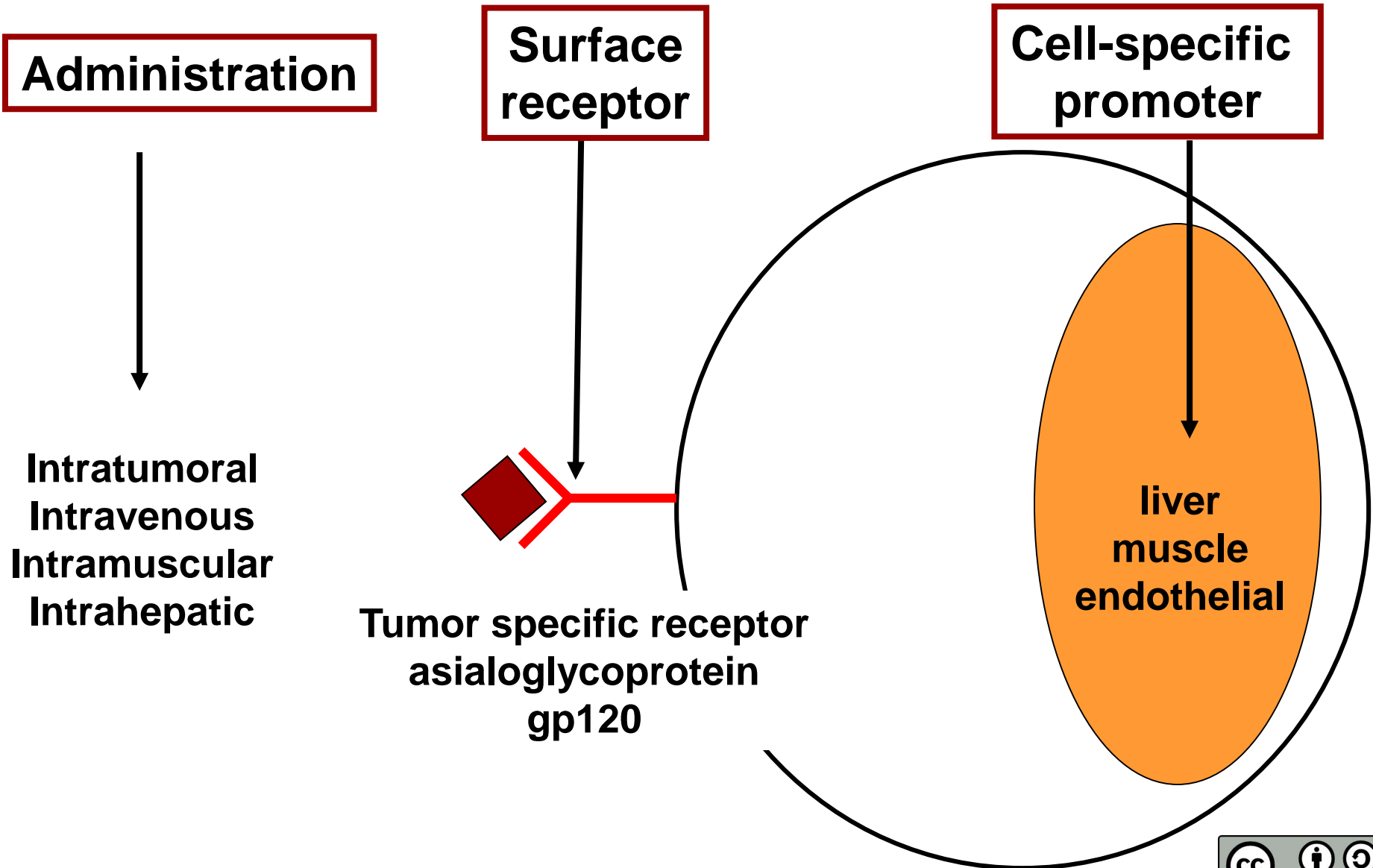
# Retroviral genome

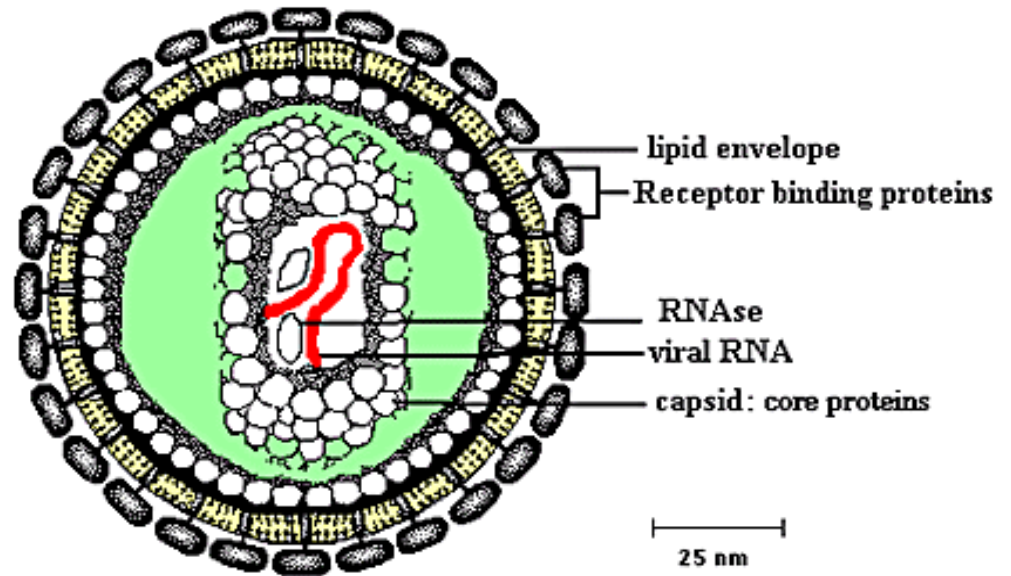


# Gene therapy - retroviral expression vector



# Cell targeting





**Diagram of a Retrovirus**

## **Benefits**

Random integration into the genome

Wide host range

Long-term transgene expression

## **Drawbacks**

Small capacity for therapeutic genes transfer

Infectivity limited to dividing cells

inactivation of complement

Safety



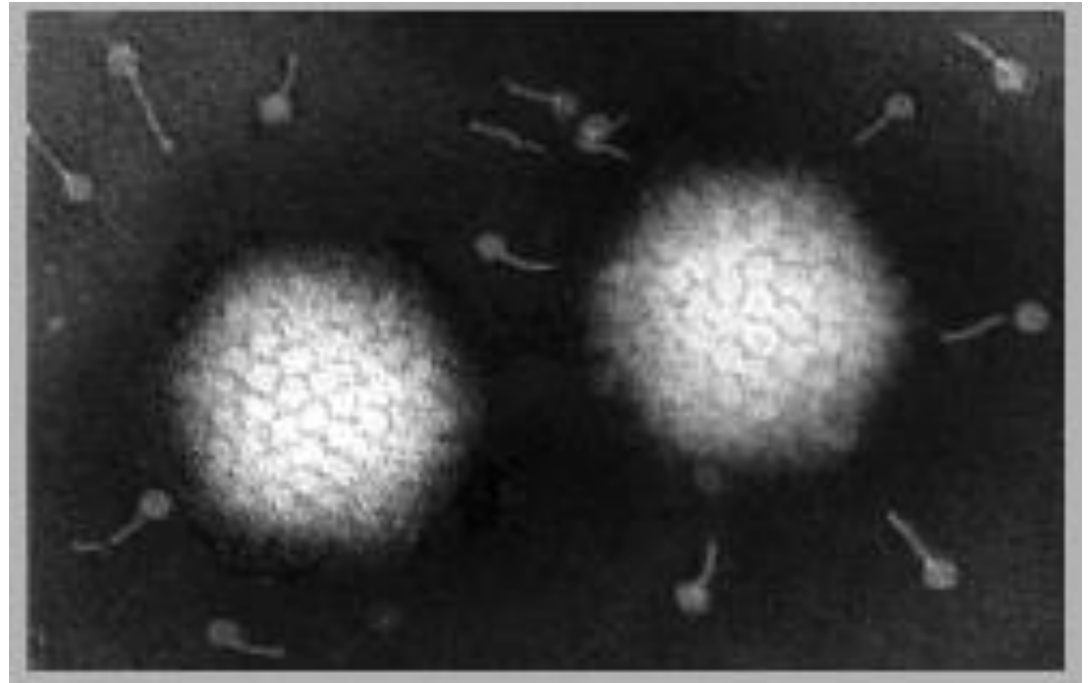
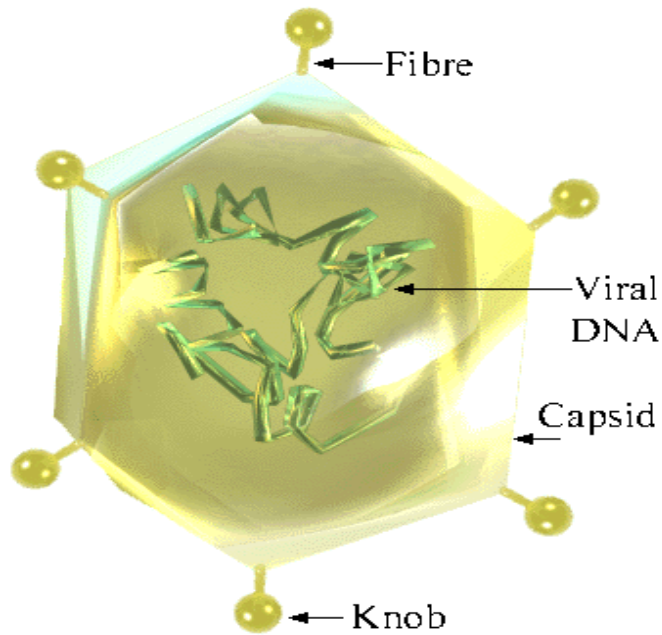
# Adenovirus

dsDNA virus

- respiratory, intestinal and ocular infections

**Onyx virus** – adenovirus with restricted replication

- especially in tumor cells

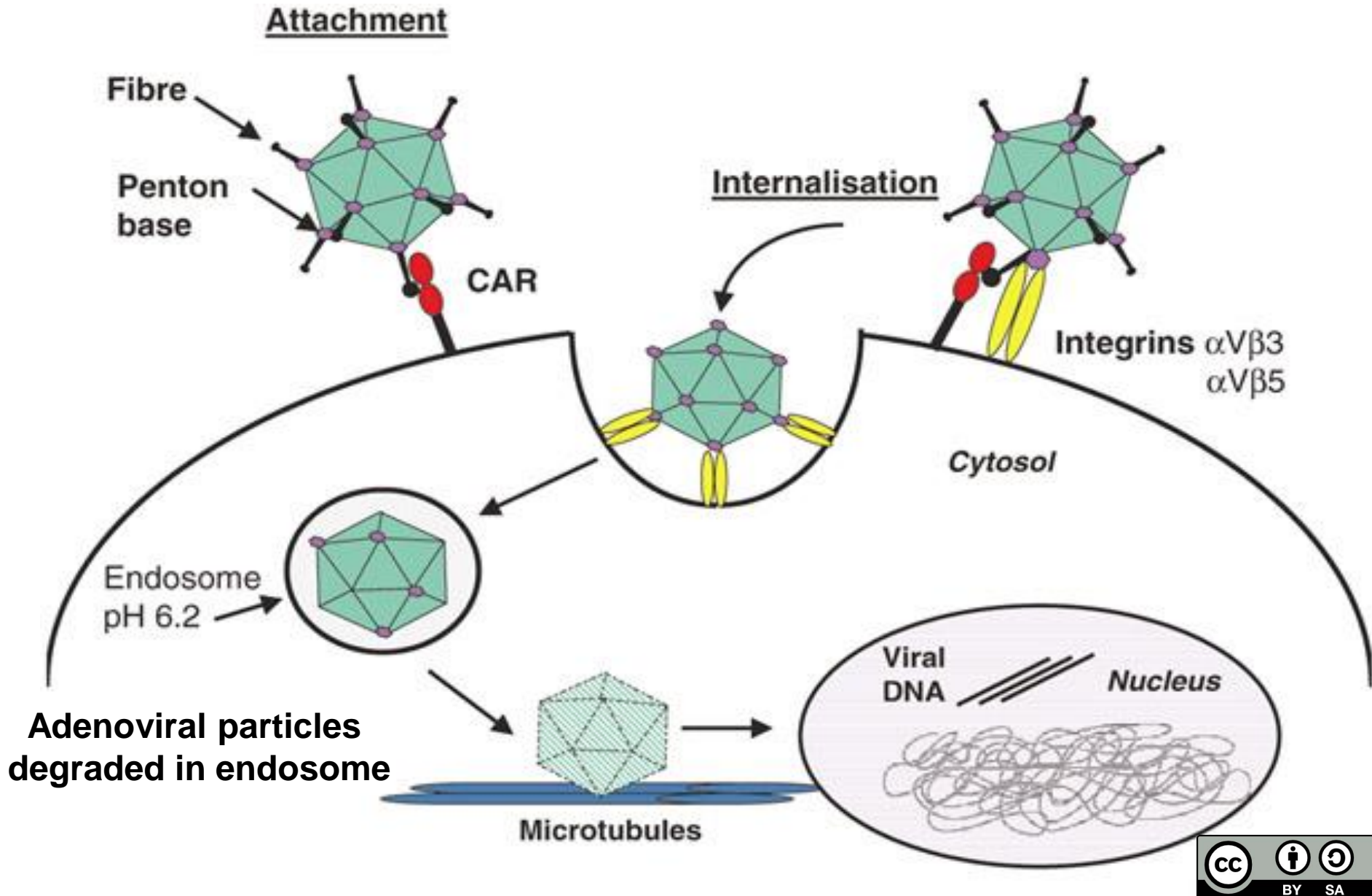


# Adenovirus

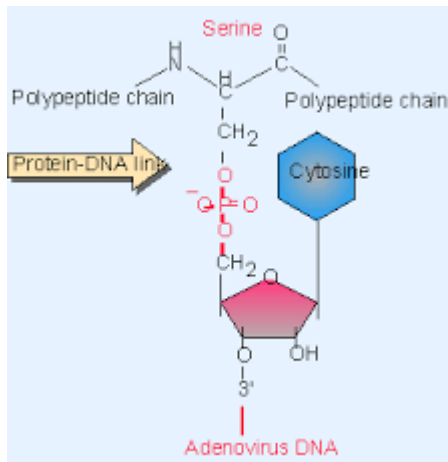
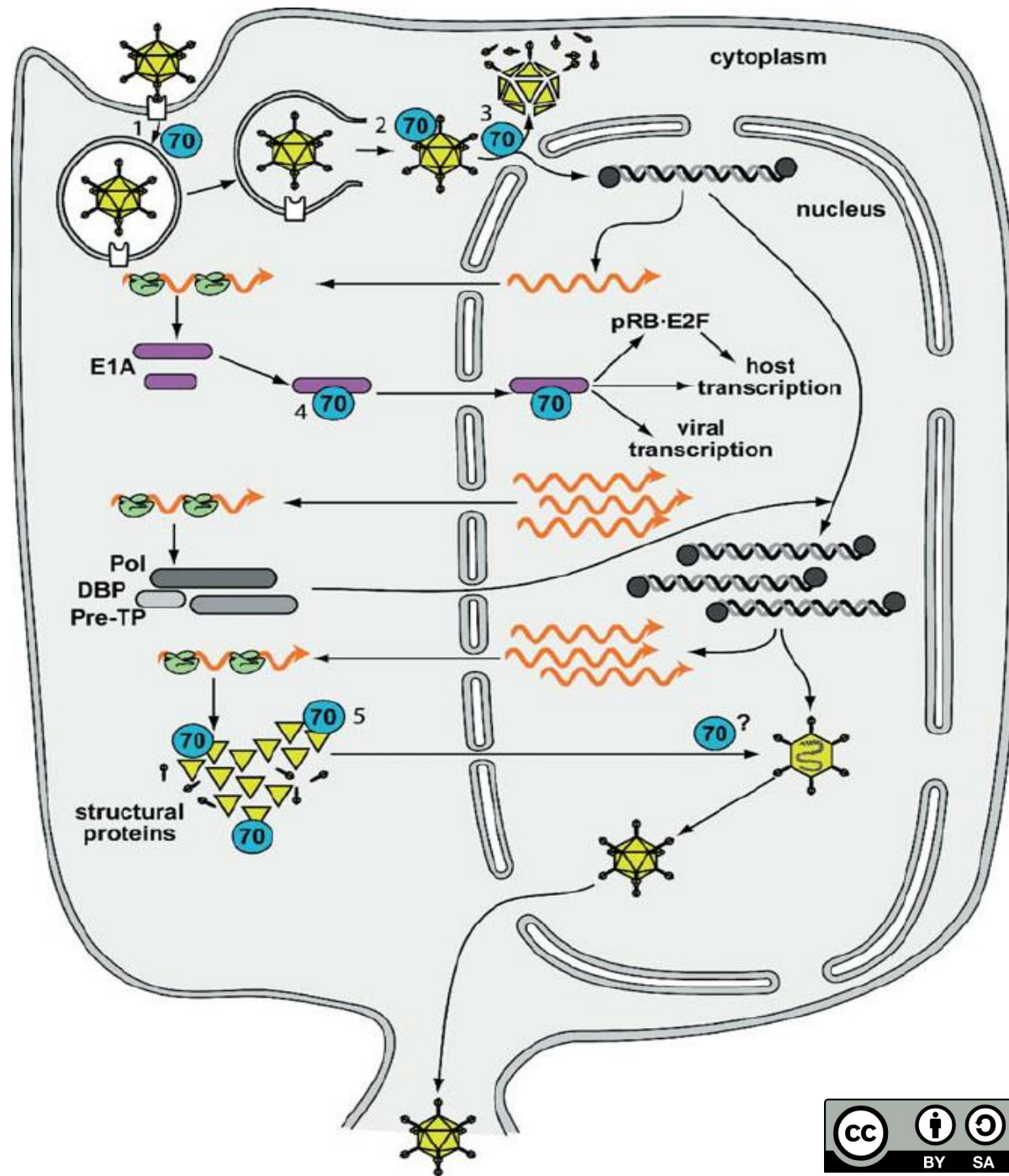
**36 kb dsDNA Genome**

**Entry - CAR (constitutive androstane receptor) and integrin co-receptor**

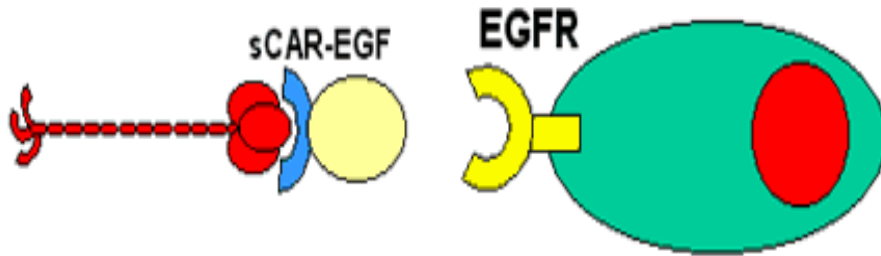
# Attachment



# Adenoviral life cycle



# Targeting of adenoviral vectors – tropism changes



Chemical conjugates

**Chemical conjugates**  
**heterogeneous,**  
**Difficult clinical certification**



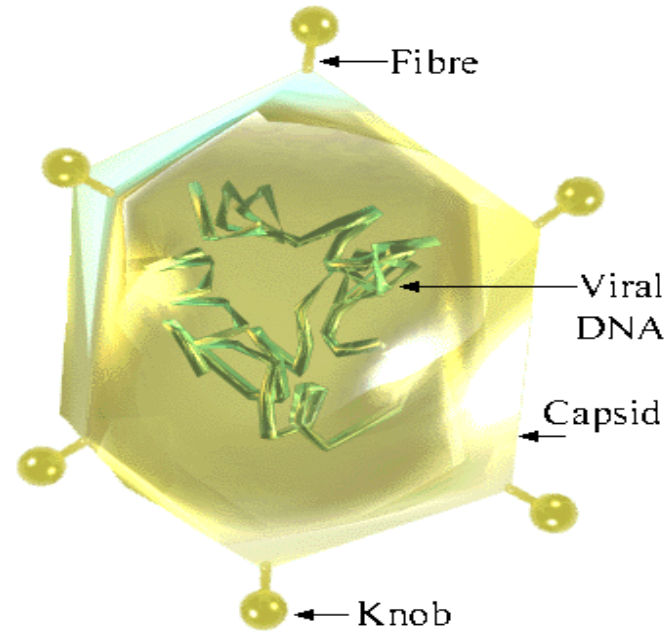
Bispecific single-chain  
antibodies (scFv)

**Recombinant genes -**  
**anti-knob + anti-receptor chain**  
**homogeneous**



Genetically altered  
fibre knob HI-loop

**Changes in adenoviral surface**  
**protein**



## Benefits

High efficiency transduction

High expression

Good capacity of incorporation of exogenous DNA

## Drawbacks

(Transient expression)

The problem of cell-specific targeting

Safety

# Other viral vectors

## **Adeno-associated virus** - ssDNA virus

Different cells (dividing and non-dividing)

The ability to integrate into the host genome

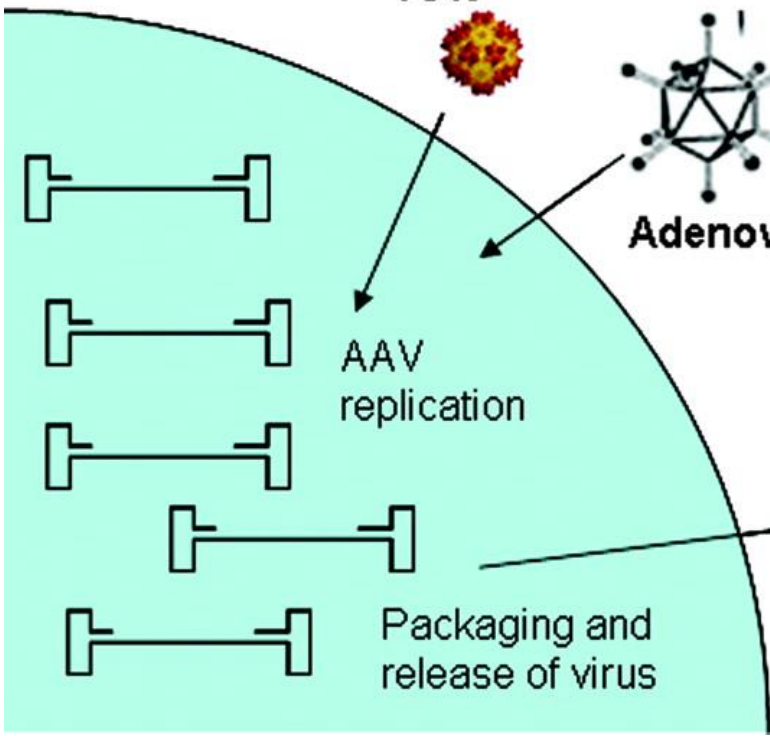
It is not associated with any disease

High efficiency transduction - a specific site on chromosome 19

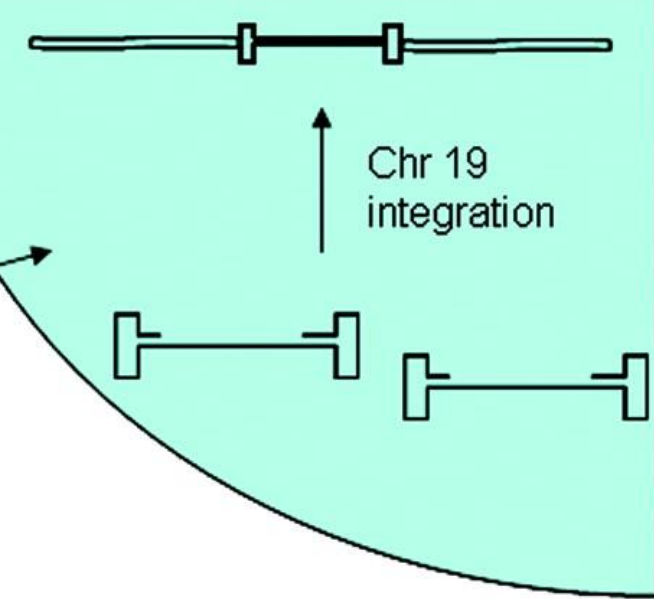
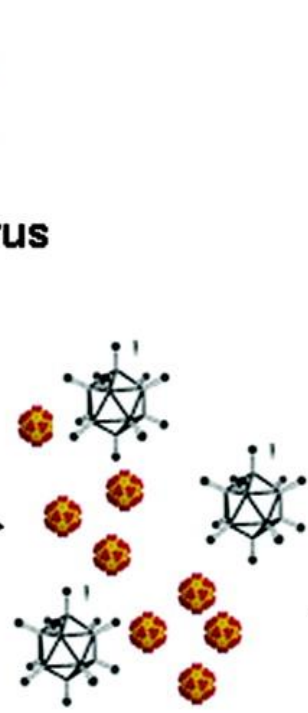


# AAV Life Cycle

## Helper Virus Co-infection

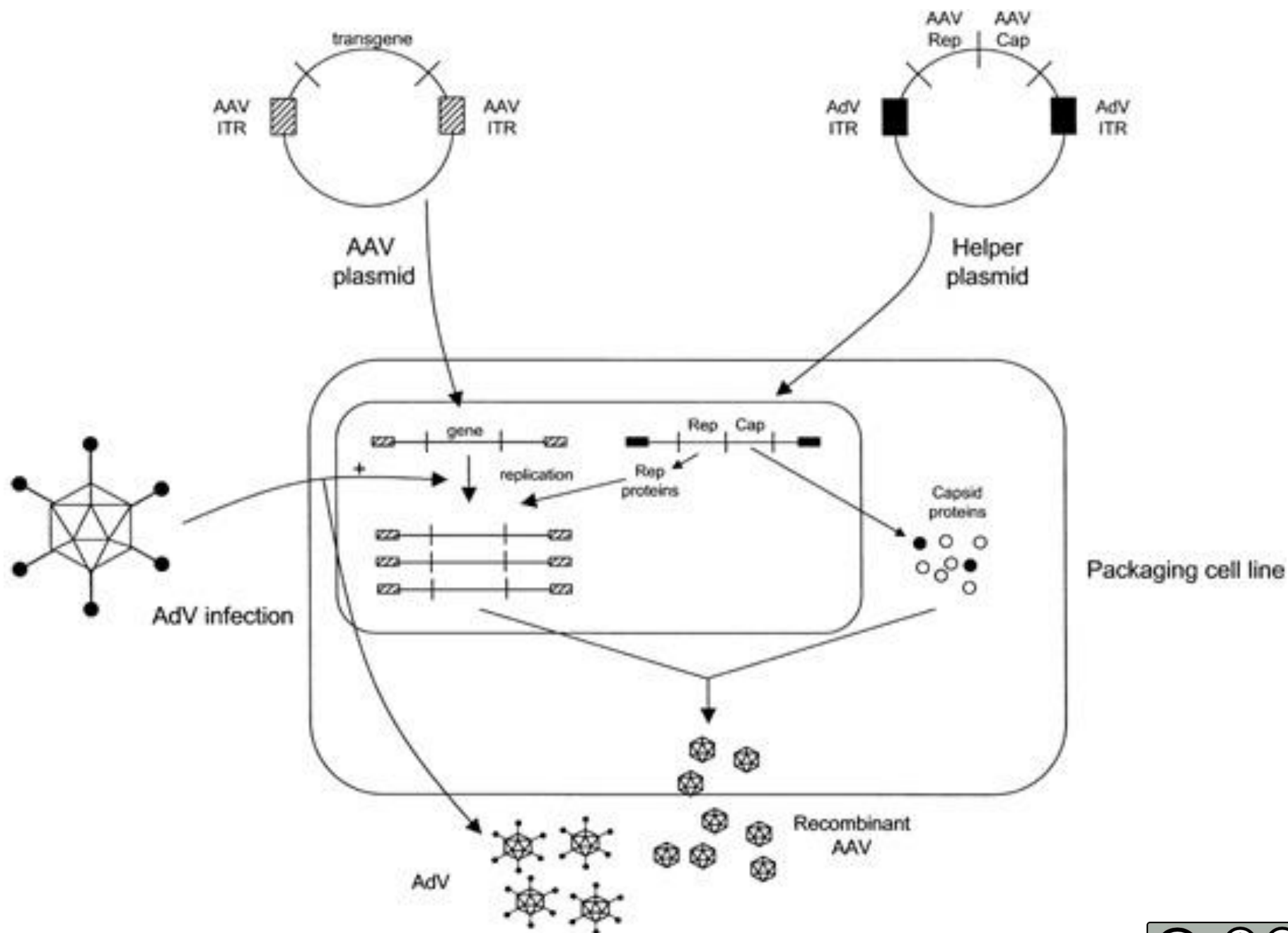


## AAV Infection



## Helper virus infection

Rescue of integrated AAV and replication



# **Herpes simplex virus**

dsDNA virus

infection of neurons

**Vaccinia virus**

**Syndbis virus**

**foamy viruses**

– yet imperfectly studied

# The ideal vector

Easy and reproducible production, high titre

Suitable size: insert - one or more genes

Targeting: limited to one type of cell

Without an immune response

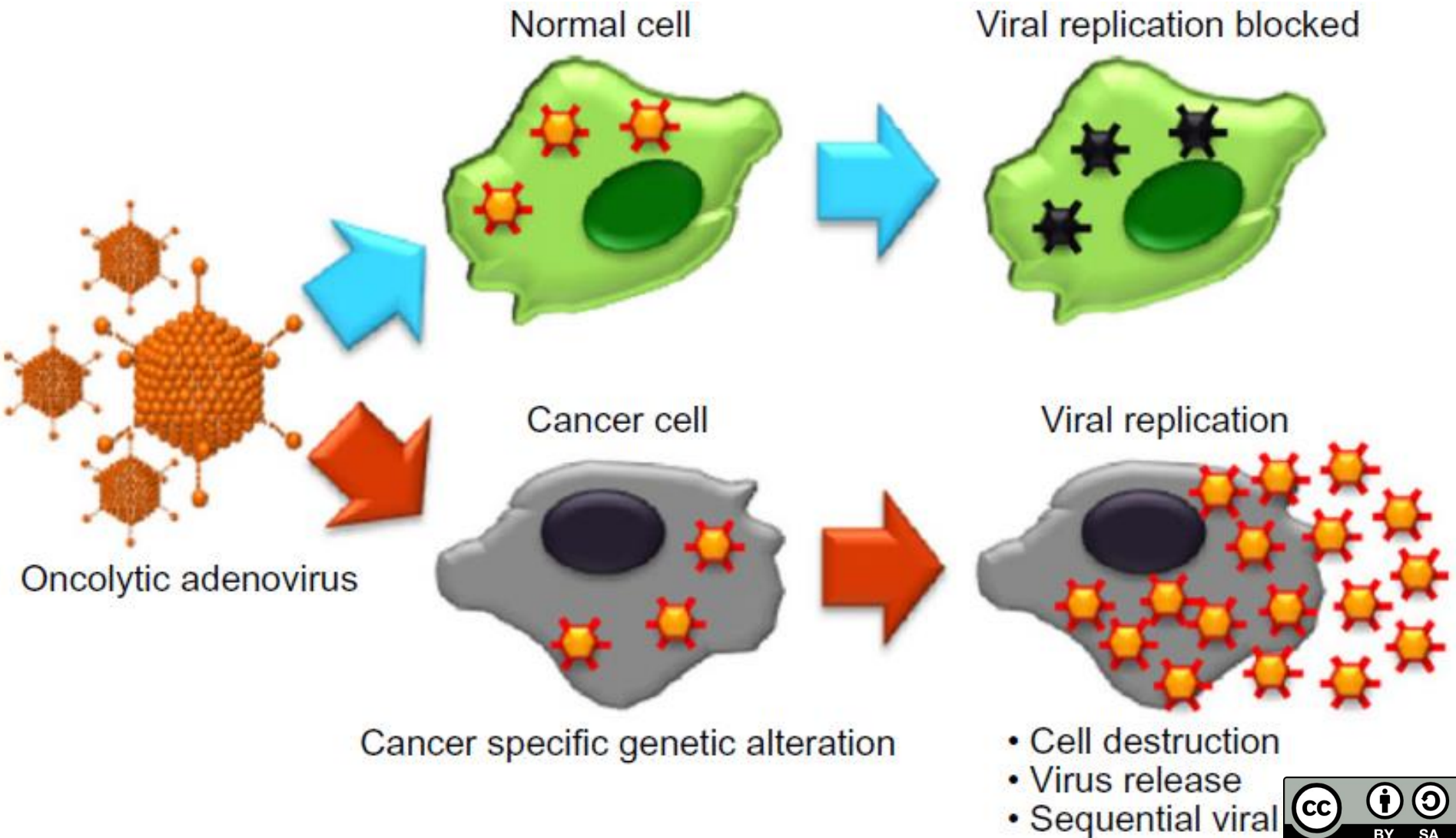
Stable - without mutations

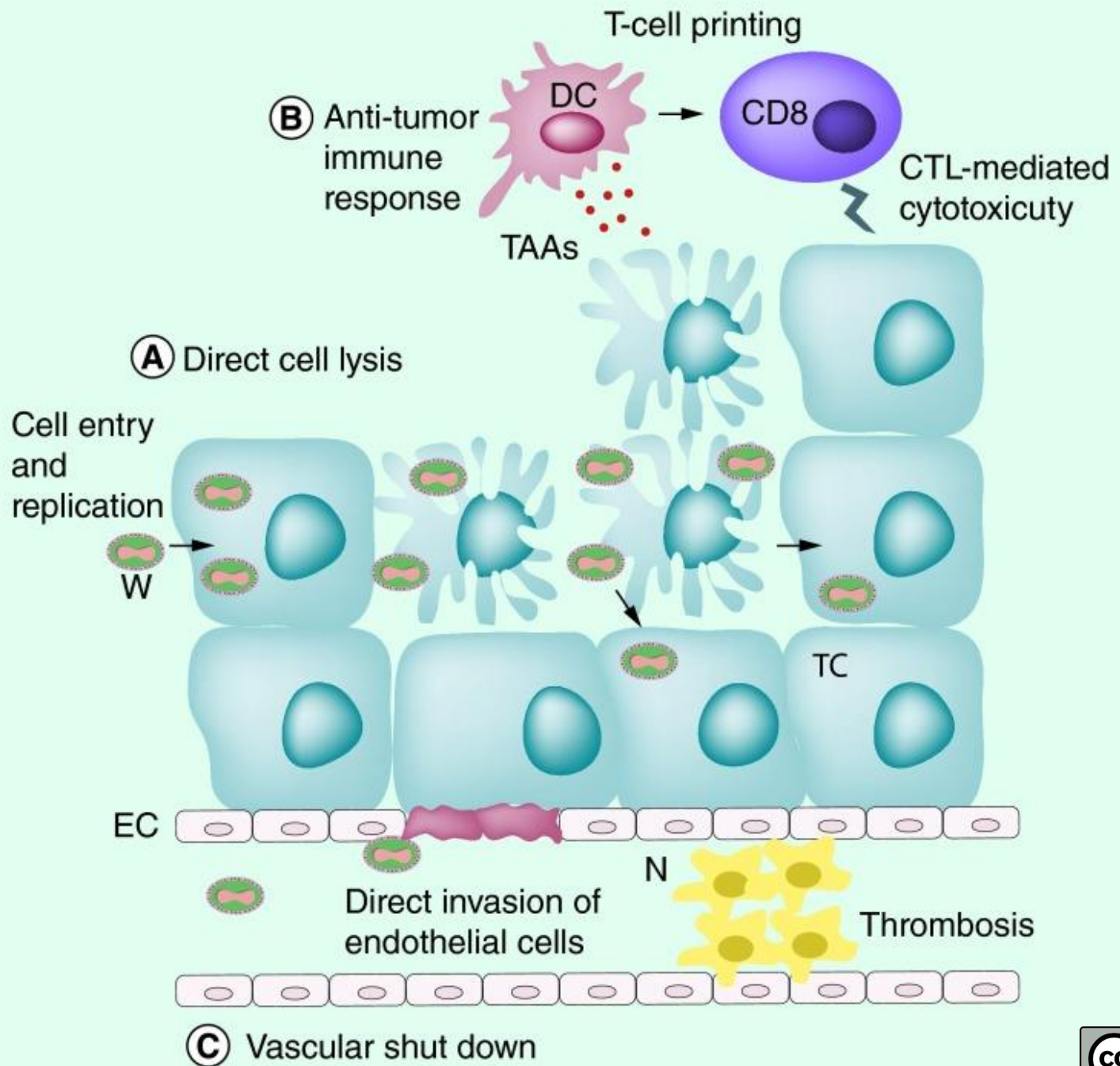
Inducible

# Oncolytic viral therapy

## Oncolytic virus

- propagation more effective in cancer cells compared to normal cells
- modified viruses with defective genes







# Vaccinia virus (VV)

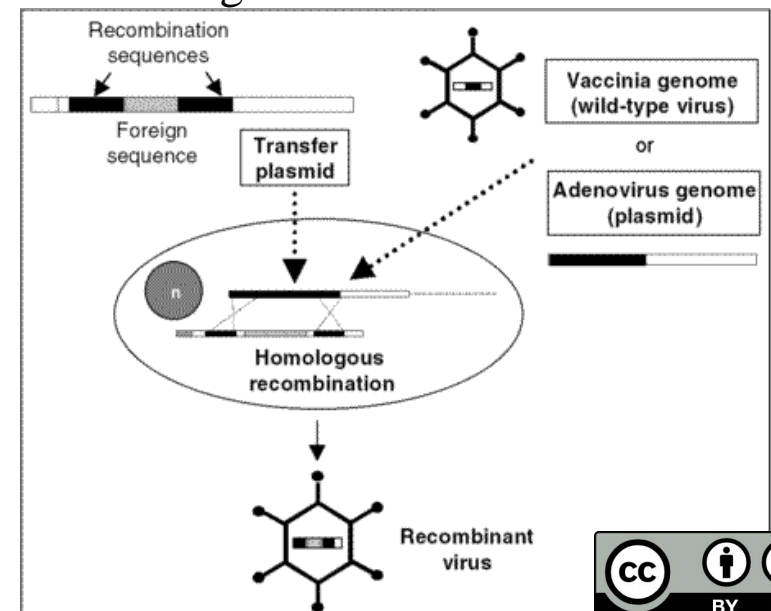
dsDNA virus -192 kbp – up to 25 kbp exogenous DNA

Short life cycle – 8 h in the cytoplasm – no risk of integration into genome. Replication 2 h after infection

own RNA polymerase – virus less dependent on the host apparatus

epidermal growth factor receptor (EGFR) and other proteins overproduced in tumor cells support replication of VV

viral antigens released into tumor – strong inflammatory response – overcoming of an immune escape of cancer. Moreover, a tumor cell lysis releases tumor antigens.





**a**

Virus



H-1PV  
naked, DNA,  
5 kb



Adenovirus  
naked, DNA,  
36–38 kb



HSV1  
enveloped, DNA,  
120–200 kb



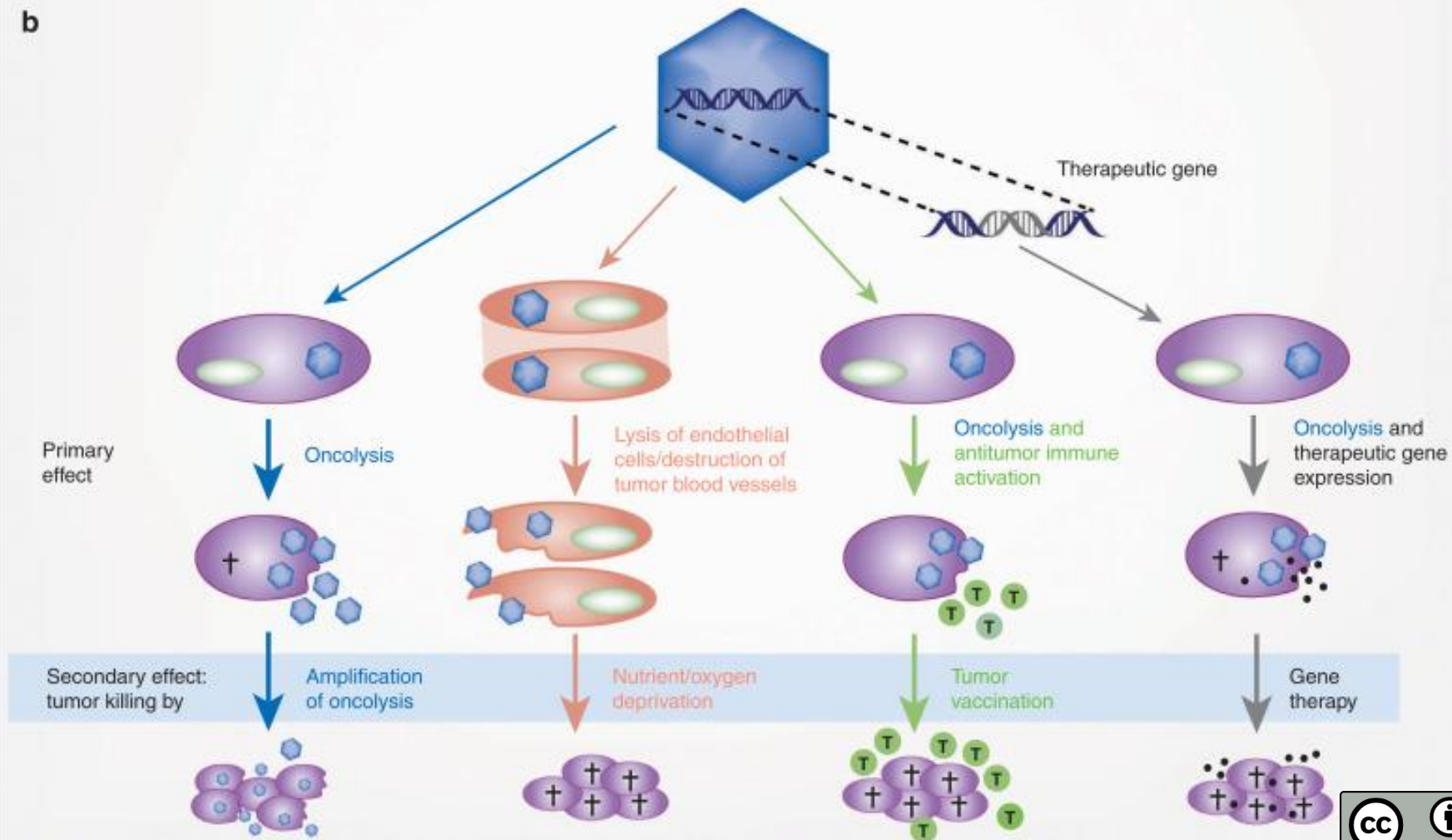
Vaccinia virus  
enveloped, DNA,  
130–280 kb



Reovirus  
naked, RNA,  
22–27 kb



Measles virus  
enveloped, RNA,  
16–20 kb

**b**

Uveřejněné materiály jsou určeny studentům Vysoké školy chemicko-technologické v Praze

jako studijní materiál. Některá textová i obrazová data v nich obsažená jsou převzata

z veřejných zdrojů. V případě nedostatečných citací nebylo cílem autora/ů záměrně poškodit event. autora/y původního díla.

S eventuálními výhradami se prosím obračejte na autora/y konkrétního výukového materiálu, aby

bylo možné zjednat nápravu.

The published materials are intended for students of the University of Chemistry and Technology, Prague as a study material. Some text and image data contained therein are taken from public sources. In the case of insufficient quotations, the author's intention was not to intentionally infringe the possible author(s) rights to the original work.

If you have any reservations, please contact the author(s) of the specific teaching material in order to remedy the situation.

# Gene therapy II



EUROPEAN UNION  
European Structural and Investing Funds  
Operational Programme Research,  
Development and Education



MINISTRY OF EDUCATION,  
YOUTH AND SPORTS

# **Non-viral DNA carriers**

**Liposomes**

**Naked DNA**

**Liposome-polycationic complexes**

**Peptide systems**

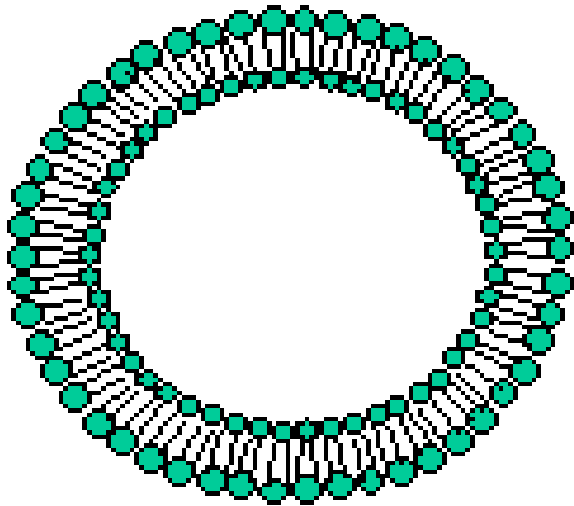
**Nanoparticles**

**Polymer**

**Metal**

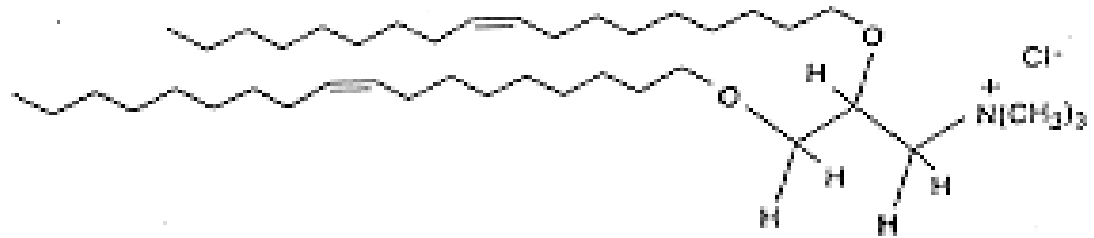
## Non-viral DNA carriers

### 1. Liposomes



### 2. Cationic polymers

positively charged lipids interact with negatively charged DNA (lipid-DNA complex)



### 3. DNA

### 4. Peptide-mediated gene transfer

Gene transfer to the brain – liposomes covered with polyethyleneglycol  
(viral vectors too large to cross a blood-brain barrier)  
Potential application in Parkinson or Huntington diseases.

## **Benefits:**

**Stable complex**

**Transfers long DNA chains**

**Specific targeting of cells**

**No induction of immune response**

## **Drawbacks:**

**Low efficiency of transfection**

**Transient expression**

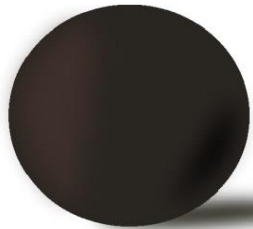
**Inhibition with serum**

**Toxicity to cell (low)**

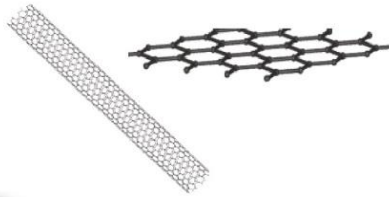
# Simple Nanocarriers

## Rigid Nanocarriers

Magnetic NPs



Carbon nanostructures



Gold NPs

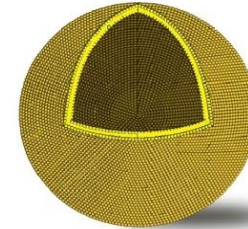


Mesoporous silica NPs

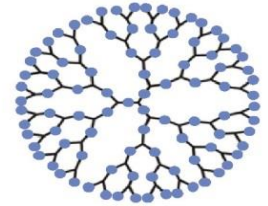


## Soft Nanocarriers

Liposomes



Polymeric NPs



**Stable, imaging potential**

**High biocompatibility**

**Magnetic control**

**High loading capacity**

**High biocompatibility**

**High loading capacity**

**High loading capacity**

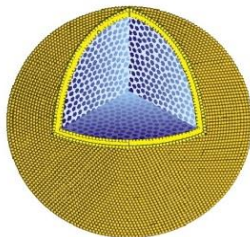
**Poor biodegradation**

**Rapid clearance**

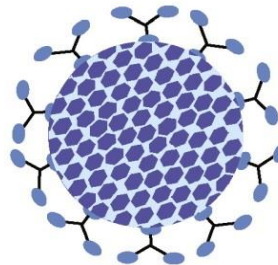
**Poor stability**

## Complex Nanocarriers

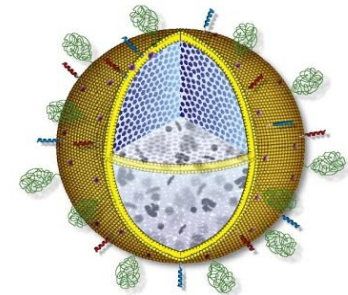
Lipid Coated Mesoporous Silica NPs



Polymer coated Mesoporous Silica NPs



Functionalized Protocells



**High loading capacity, stable, high biocompatibility, imaging potential, reduced clearance**

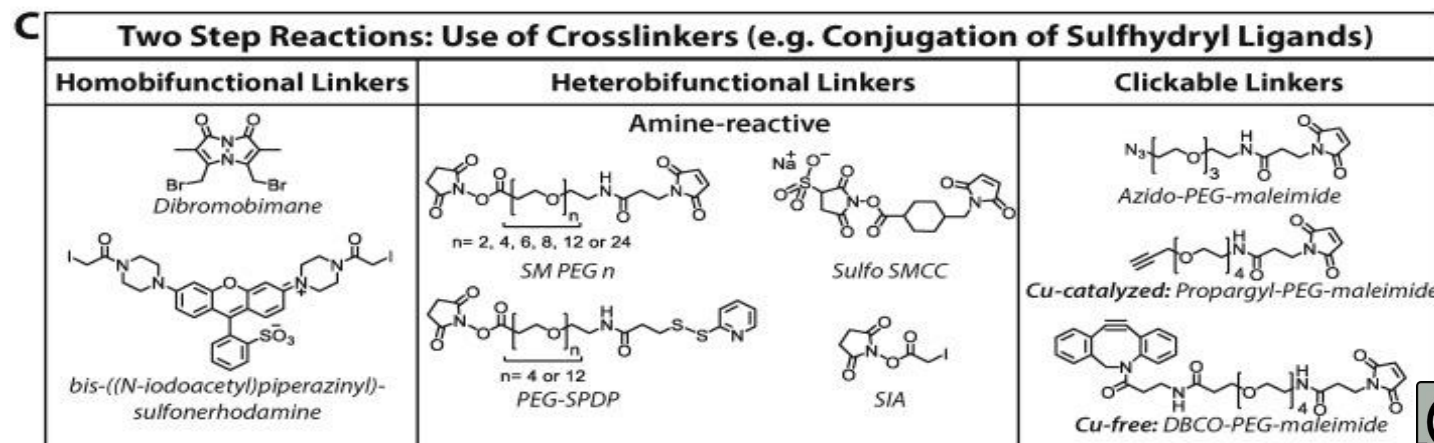
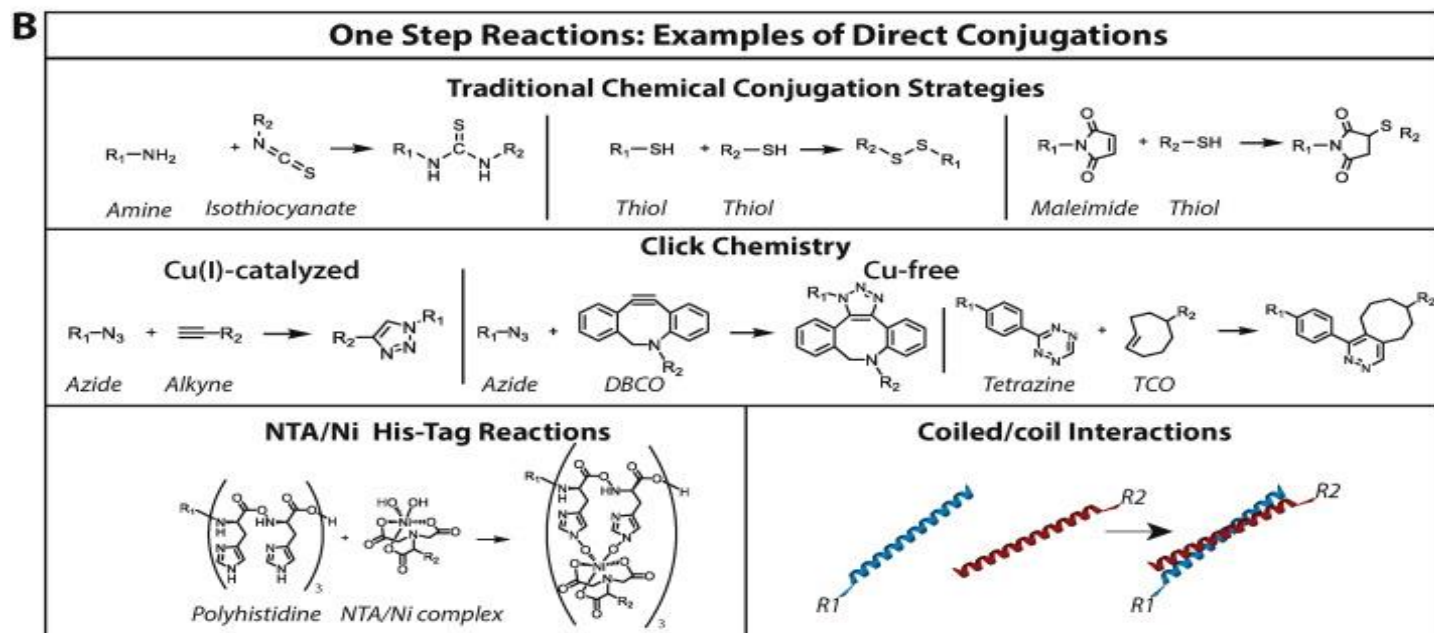
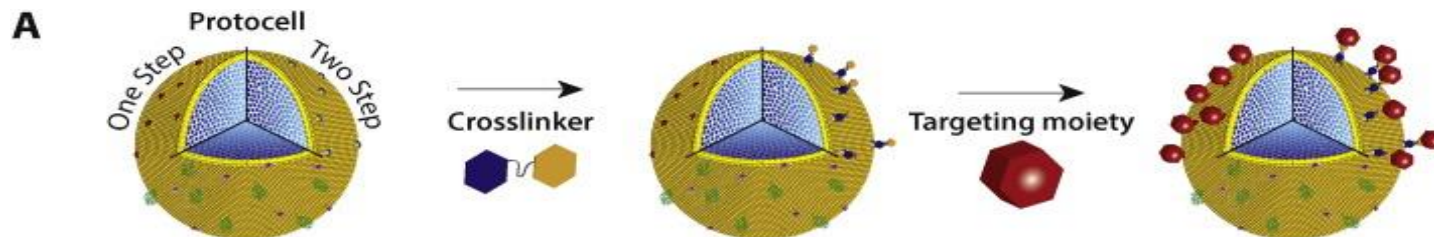
**Untargeted**

**Leaky**

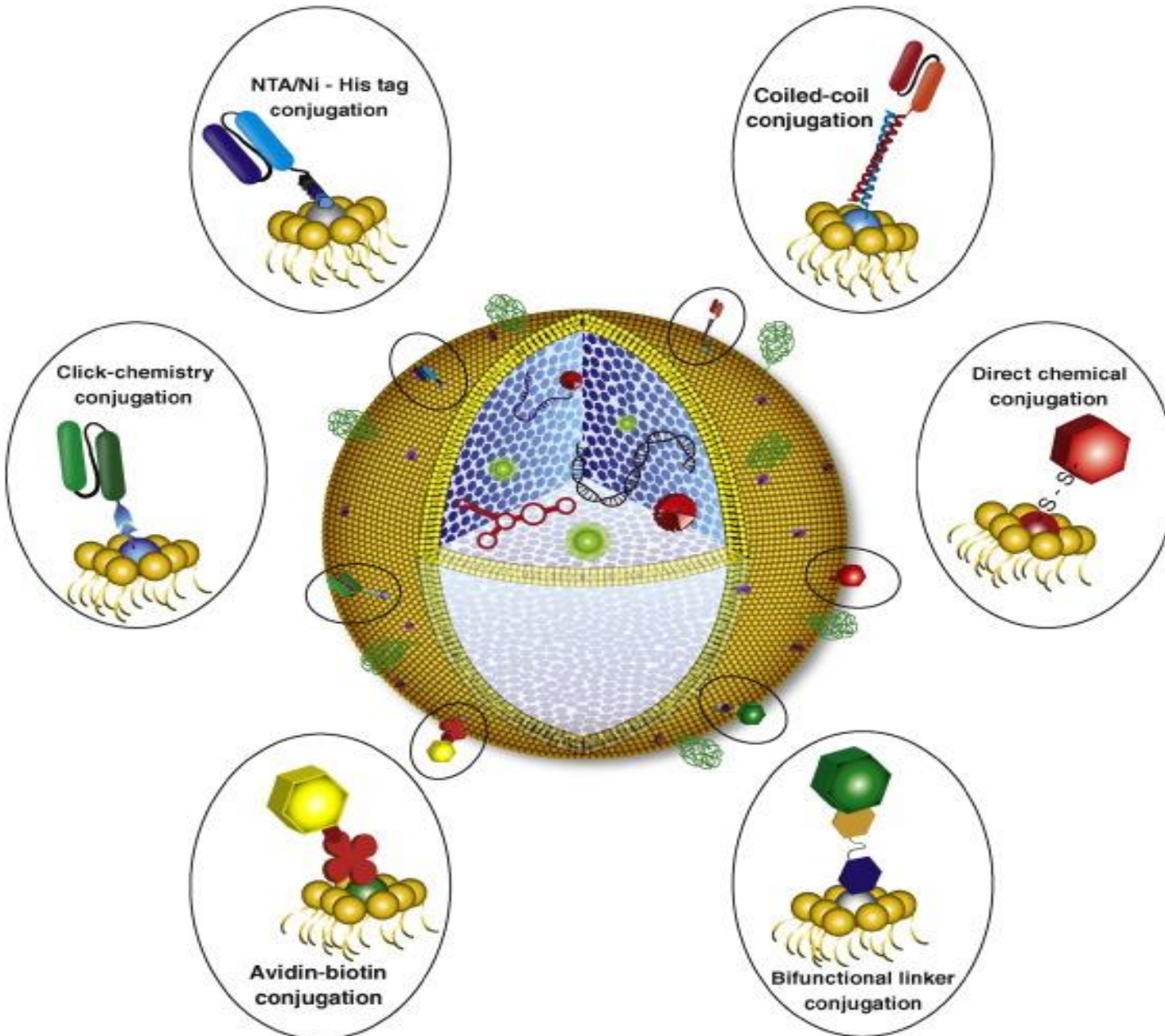
**Increased targeting**  
selective PEG or polymers







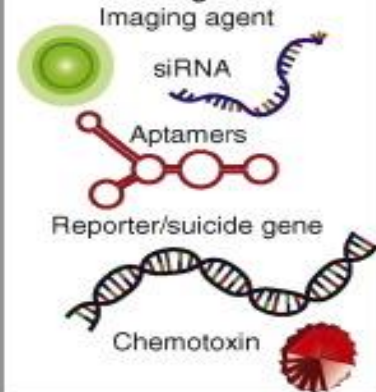
# Functionalised nanoparticle



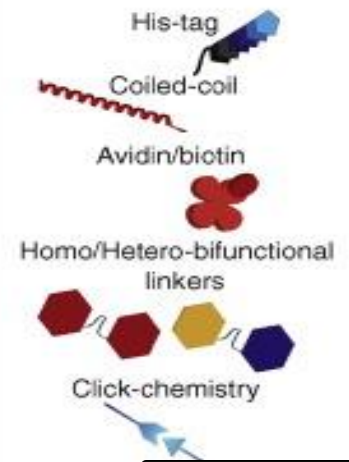
## Targeting moieties

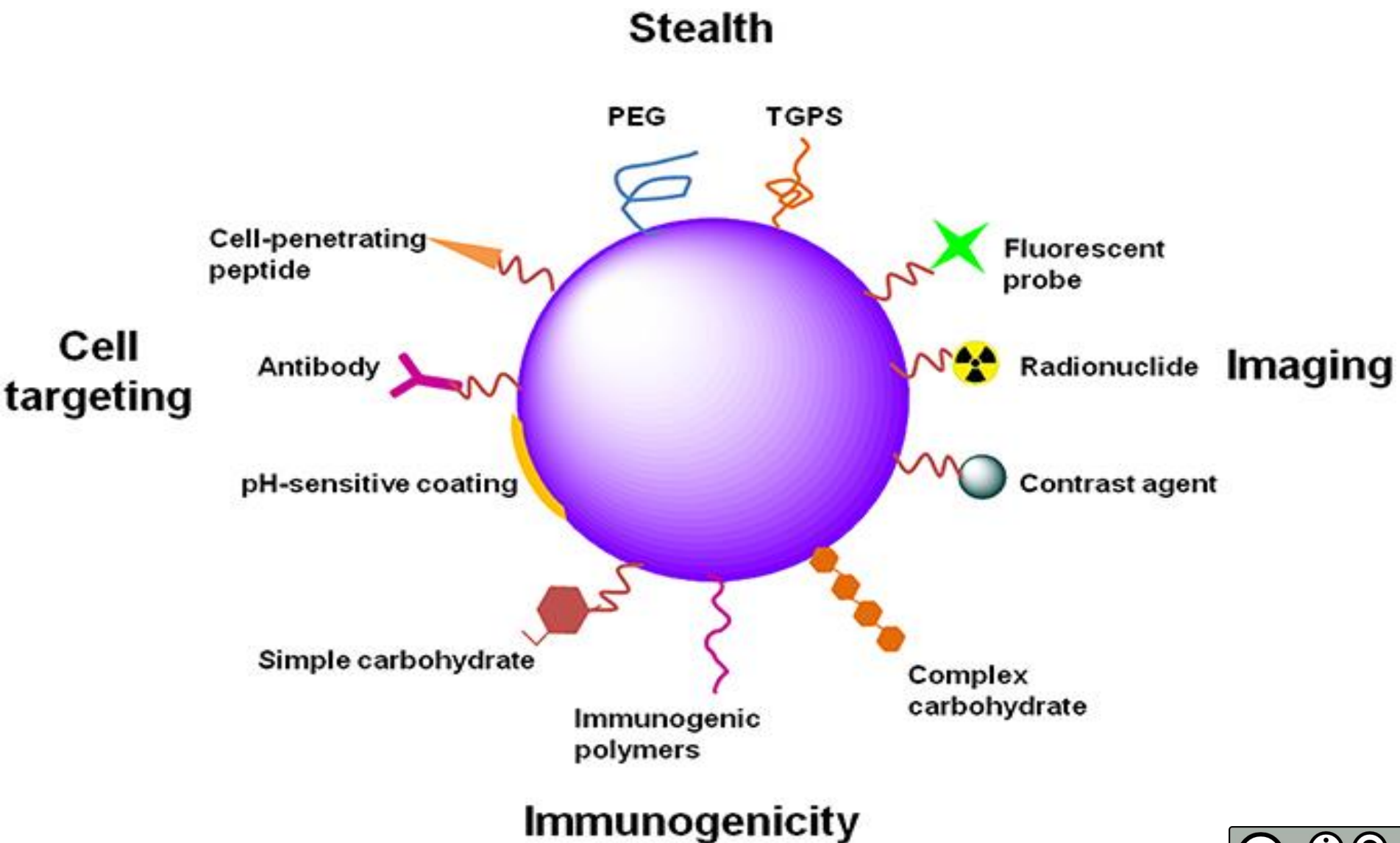


## Cargos



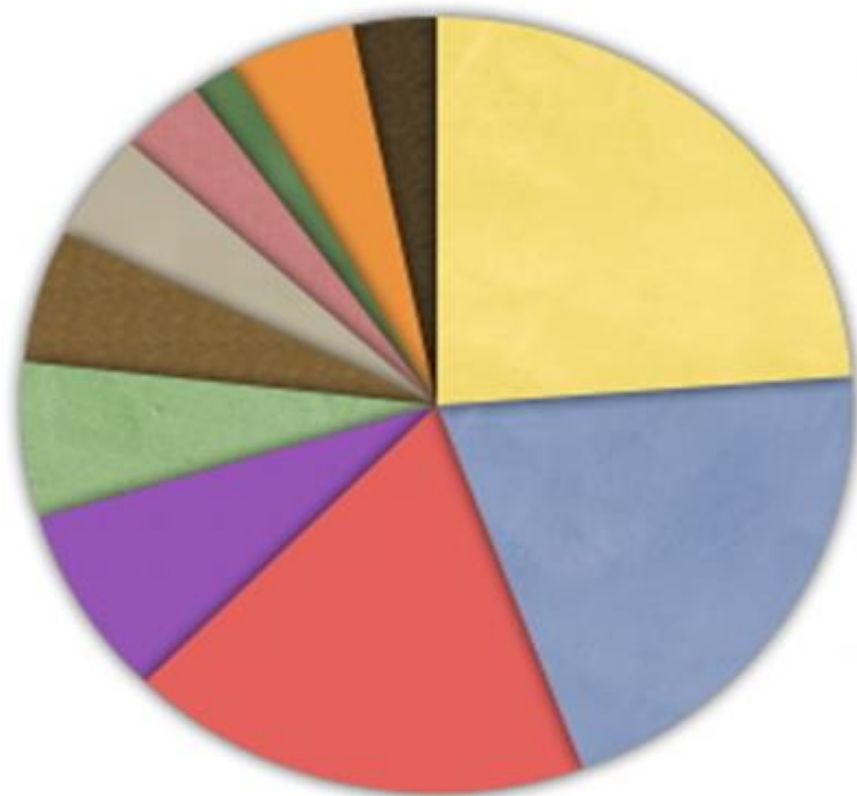
## Conjugation strategies







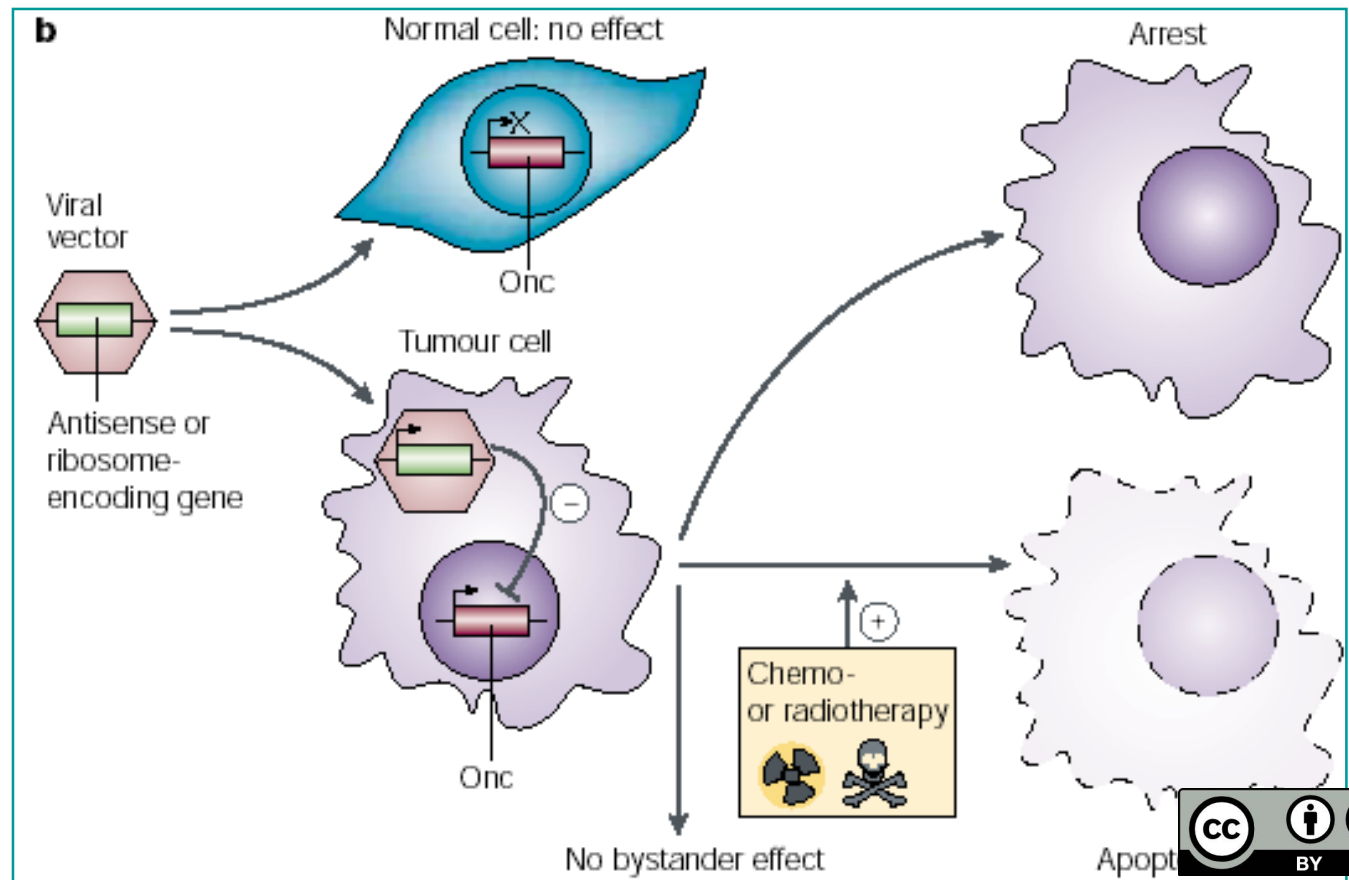
## Vectors Used in Gene Therapy Clinical Trials



- Adenovirus 23.8% (n=400)
- Retrovirus 20.5% (n=344)
- Naked/Plasmid DNA 17.7% (n=304)
- Vaccinia virus 7.9% (n=133)
- Lipofection 6.5% (n=109)
- Poxvirus 5.5% (n=93)
- Adeno-associated virus 4.5% (n=75)
- Herpes simplex virus 3.3% (n=56)
- Lentivirus 1.7% (n=29)
- Other categories 4.9% (n=82)
- Unknown 3.3% (n=55)

# Knockout gene therapy:

- inactivates oncogenes and reduces cellular proliferation.
- Approaches:
  - delivering a *dominant negative* mutant oncogene
  - delivering an interfering nucleic acid



## Meganucleases (microbial)

recognition sequence > 14bp specific →

effort to find another for different objectives

from microbial sources, mutagenesis, hybrid

## Hybrid Meganuclease



# CompoZr® Zinc Finger Nuclease Technology



## Zinc finger nucleases

Double-stranded breaks in DNA - without using a homologous NHEJ repair template -

Homologous recombination - DSB repair using the supplied template DNA

FokI (*Flavobacterium okeanoikoites*) IIS restriction endonuclease

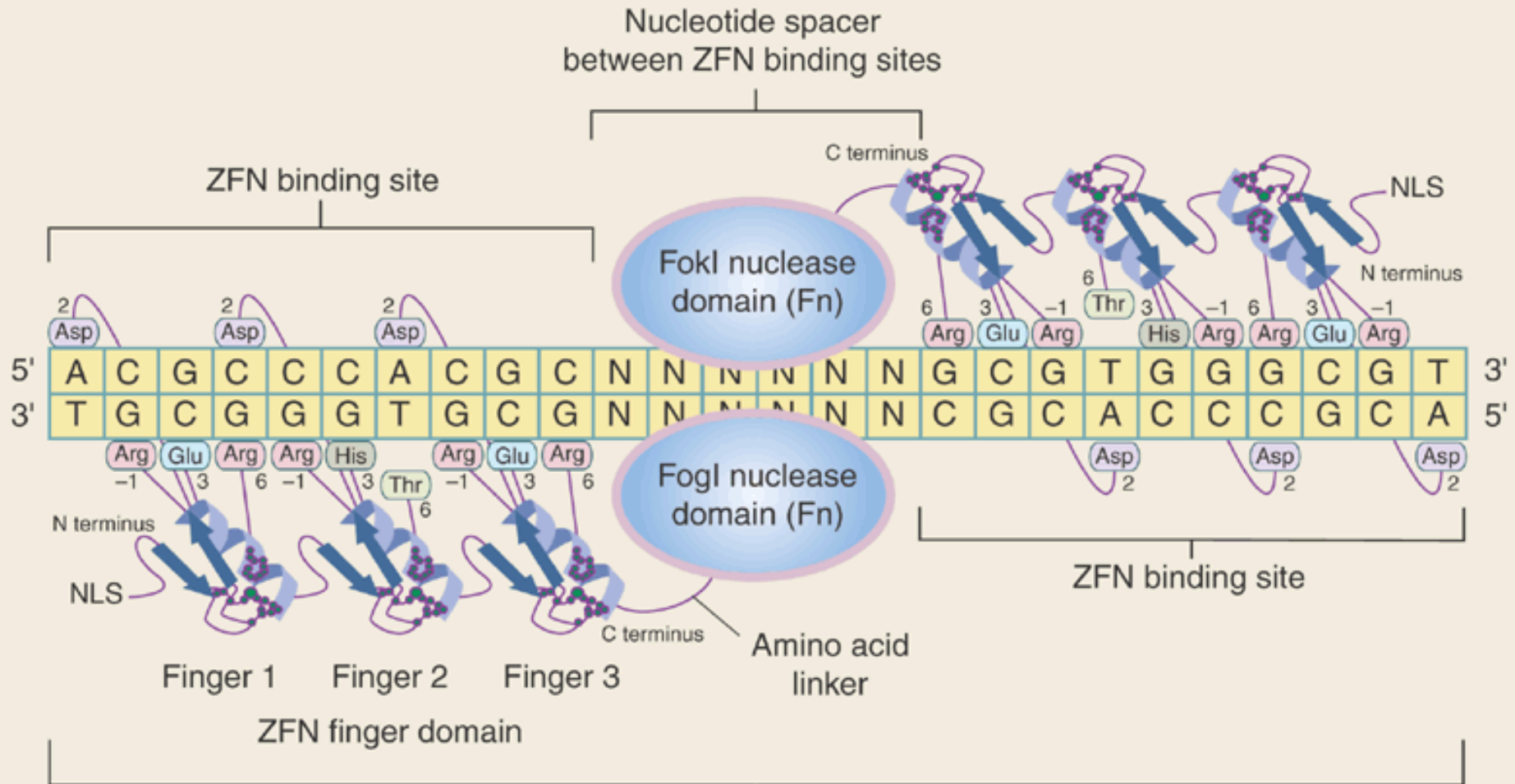
N-terminal binding domain and a C-terminal non-specific endonuclease domain

- Nuclease domain - required dimerization - enhanced specificity (both monomers DBD must recognize sequence)

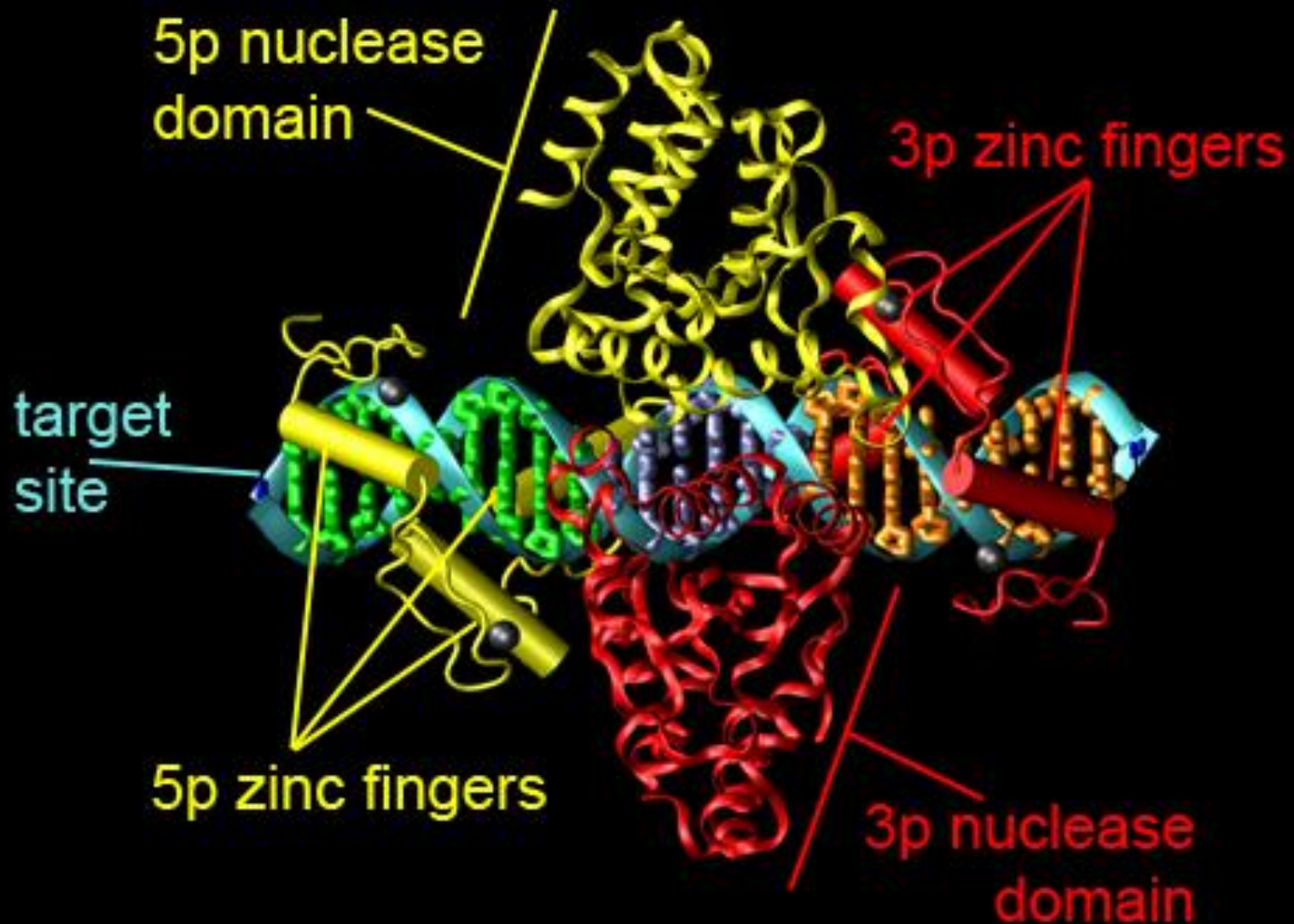


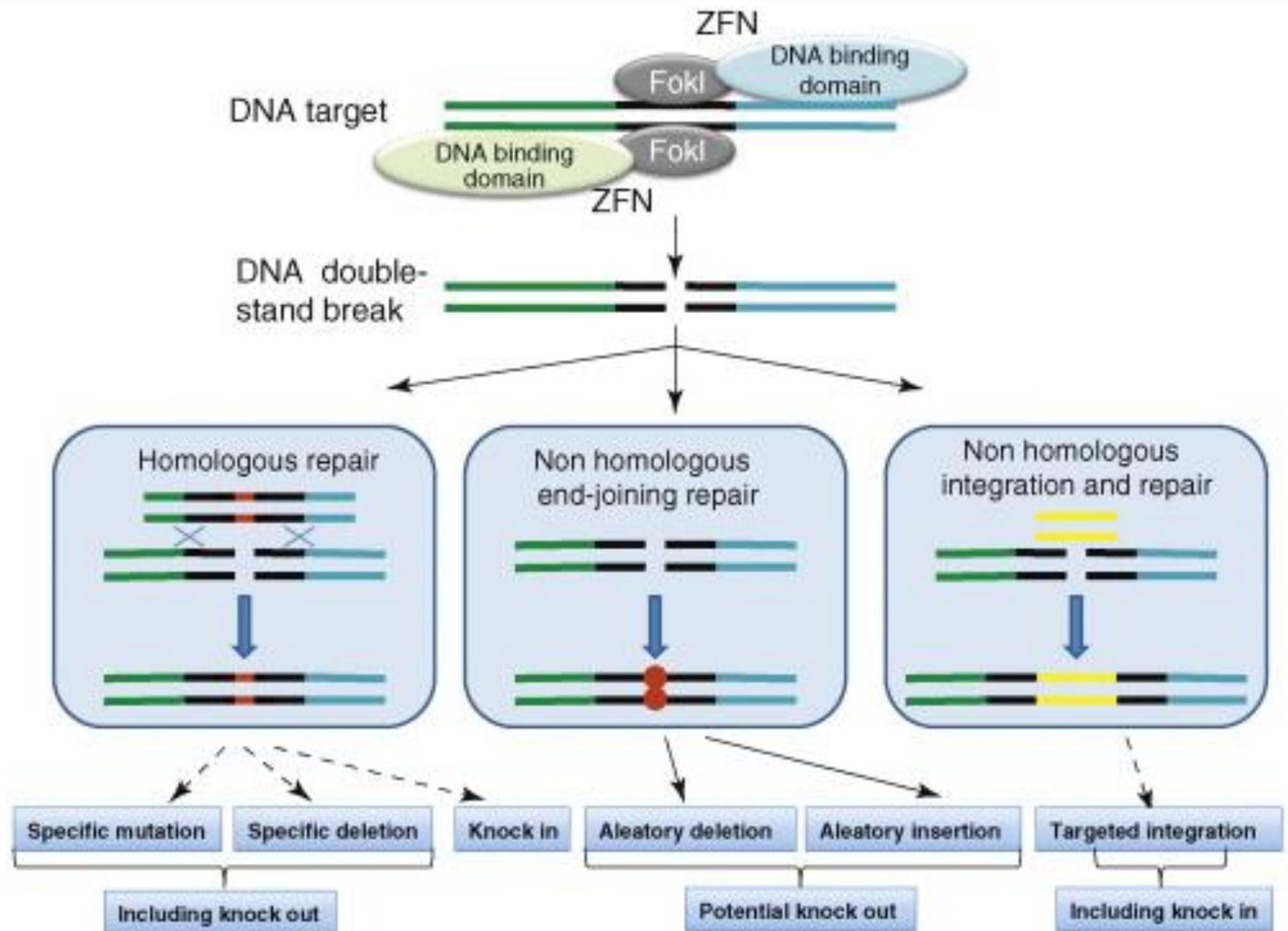


## Zinc finger nucleases

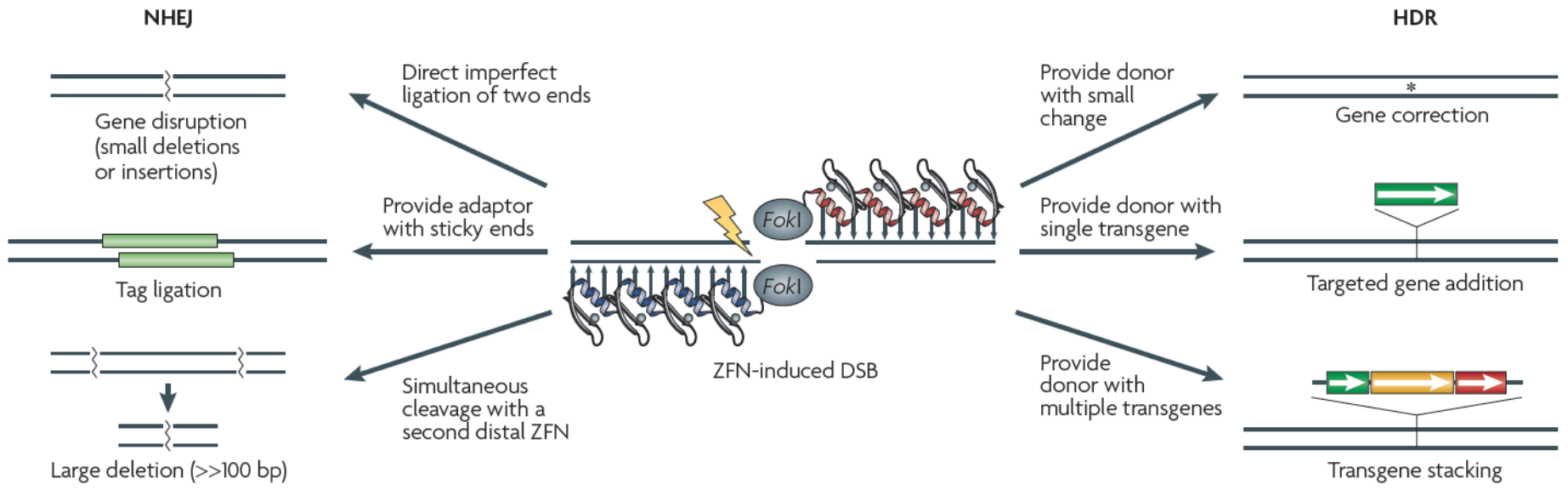


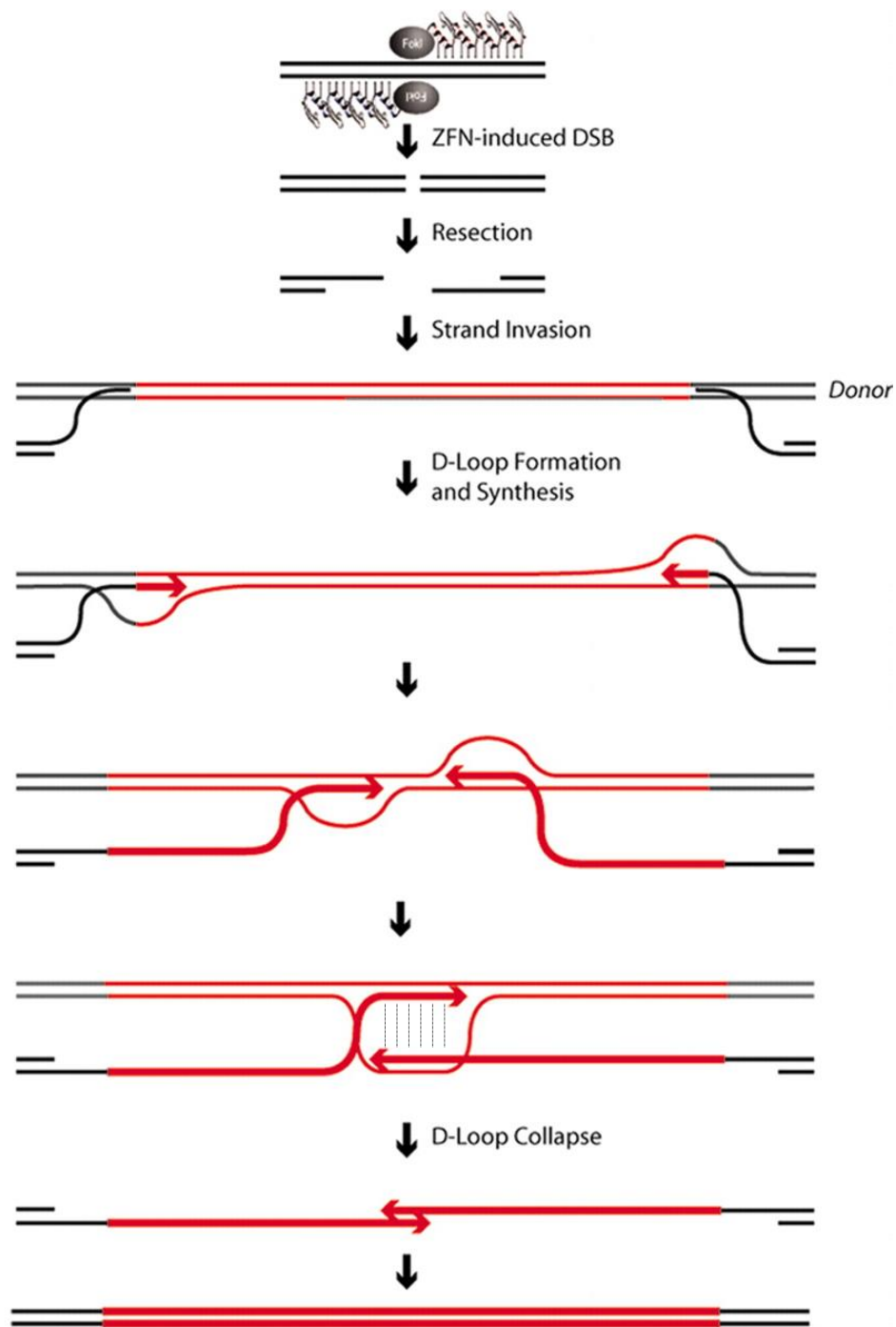
# ZFN





*TRENDS in Biotechnology*



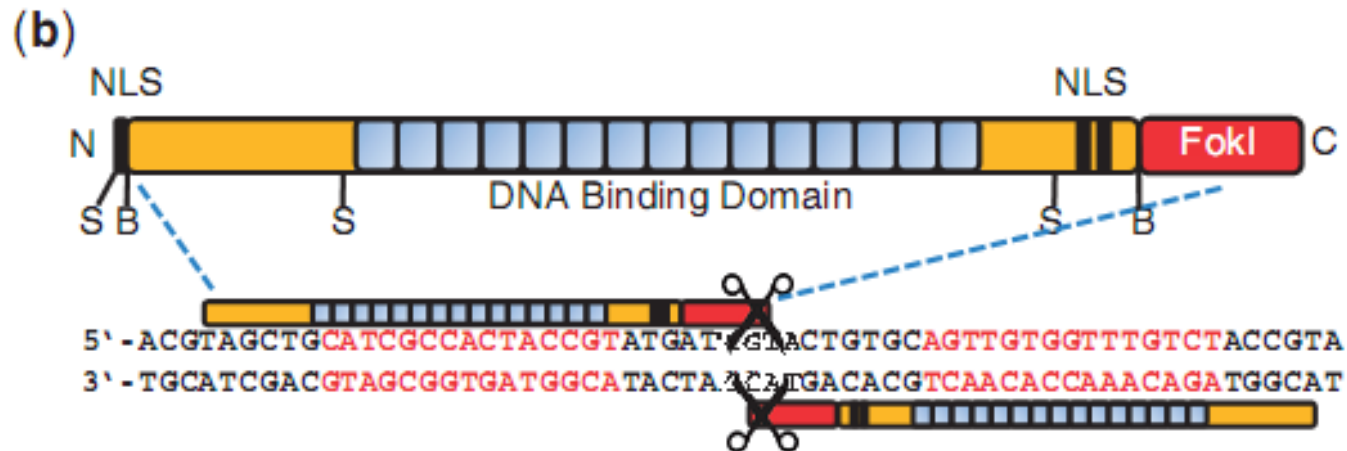
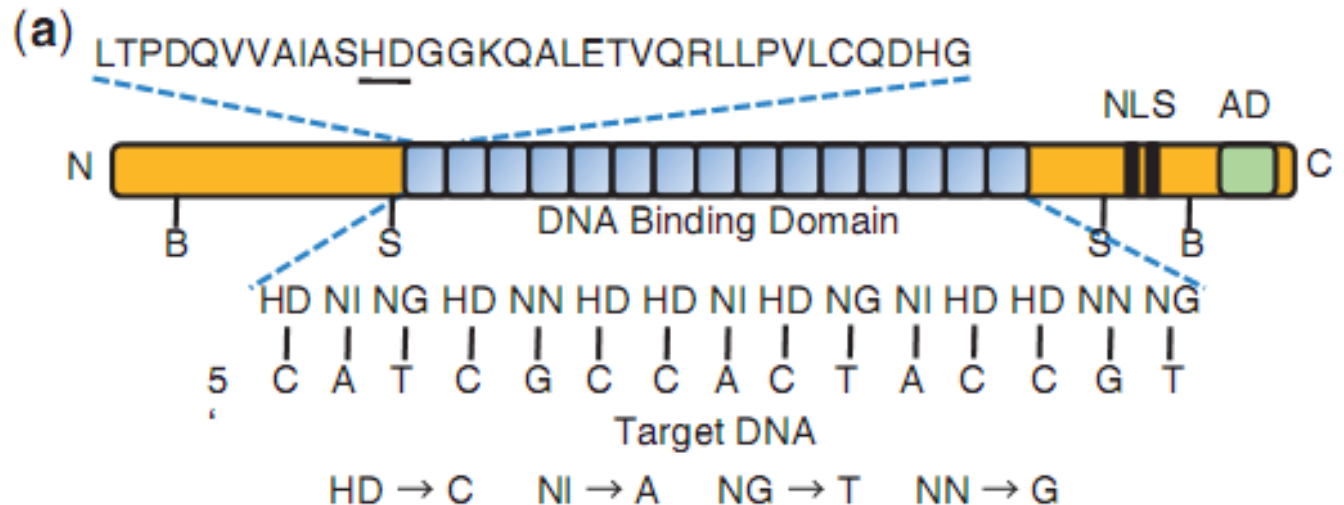


# Transcription activator-like effectors (TALEs)

## Transcription-activator like effector nucleases (TALENs)

DBD 33-34 AA conserved repeat; 12th and 13th AA

- Repeat Variable Diresidue - specific nucleotide recognition
- non-specific DNA cleavage domain - FokI endonuclease





# Hybrid Meganuclease



ZFN



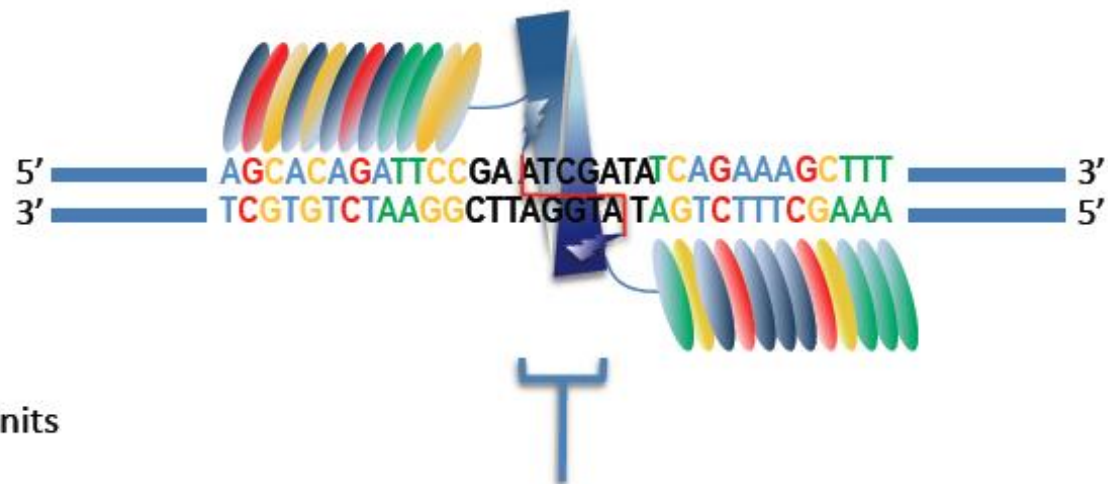
Zinc finger domains



TALEN



TALE subunits

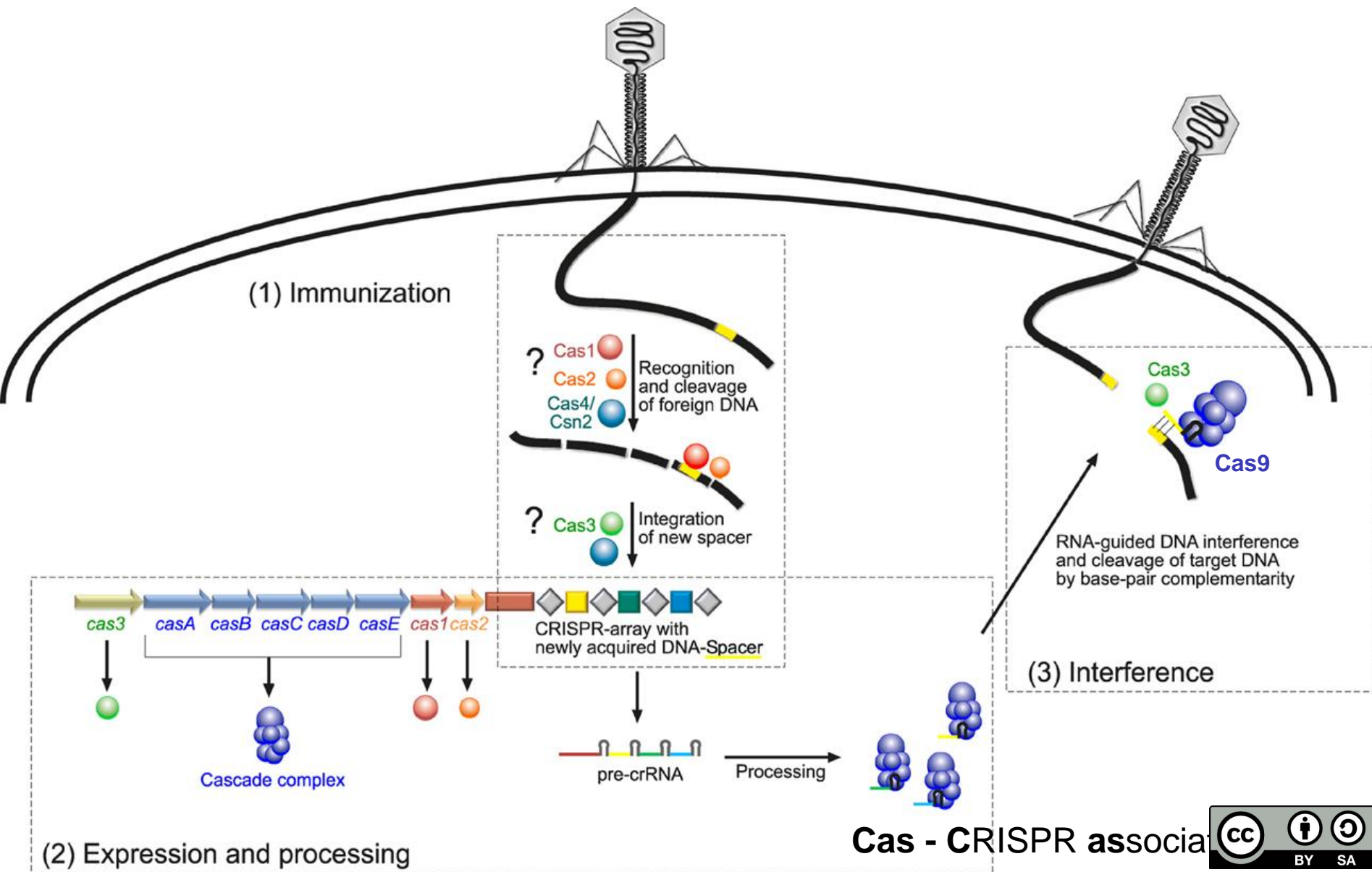


active FokI catalytic subunit heterodimer



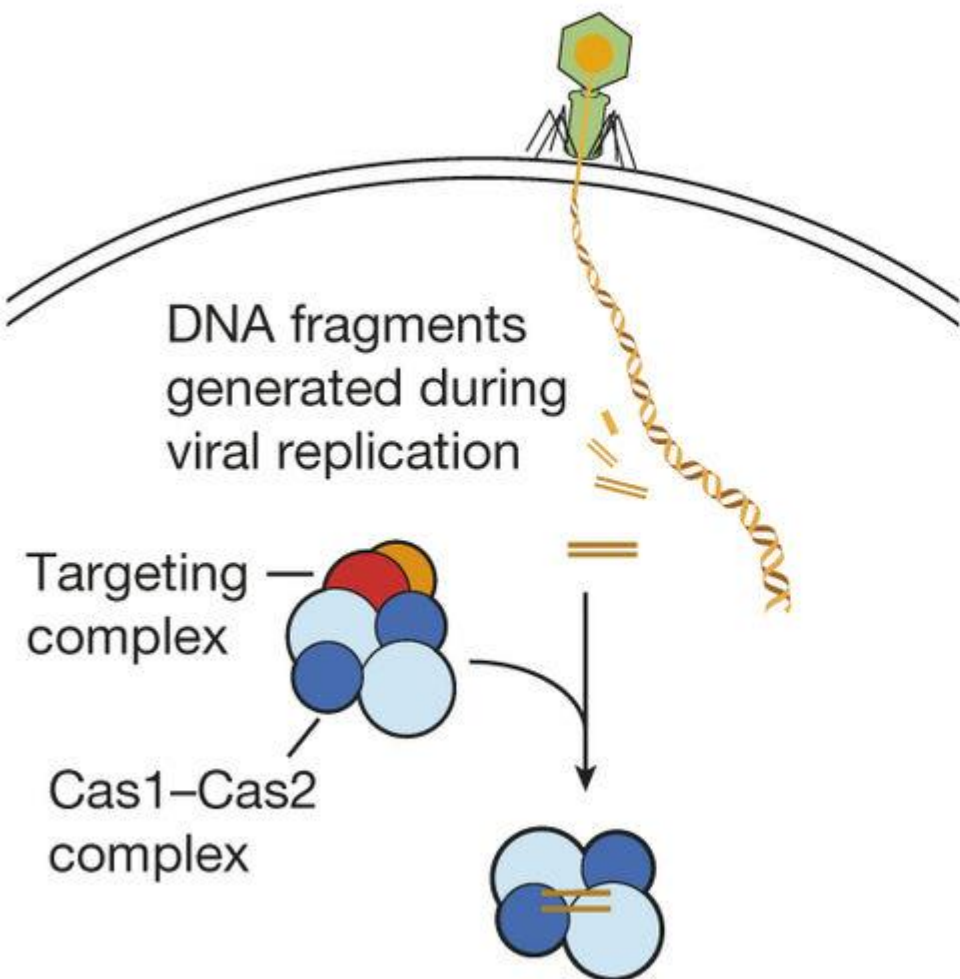
# CRISPRs

clustered regularly interspaced short palindromic repeats

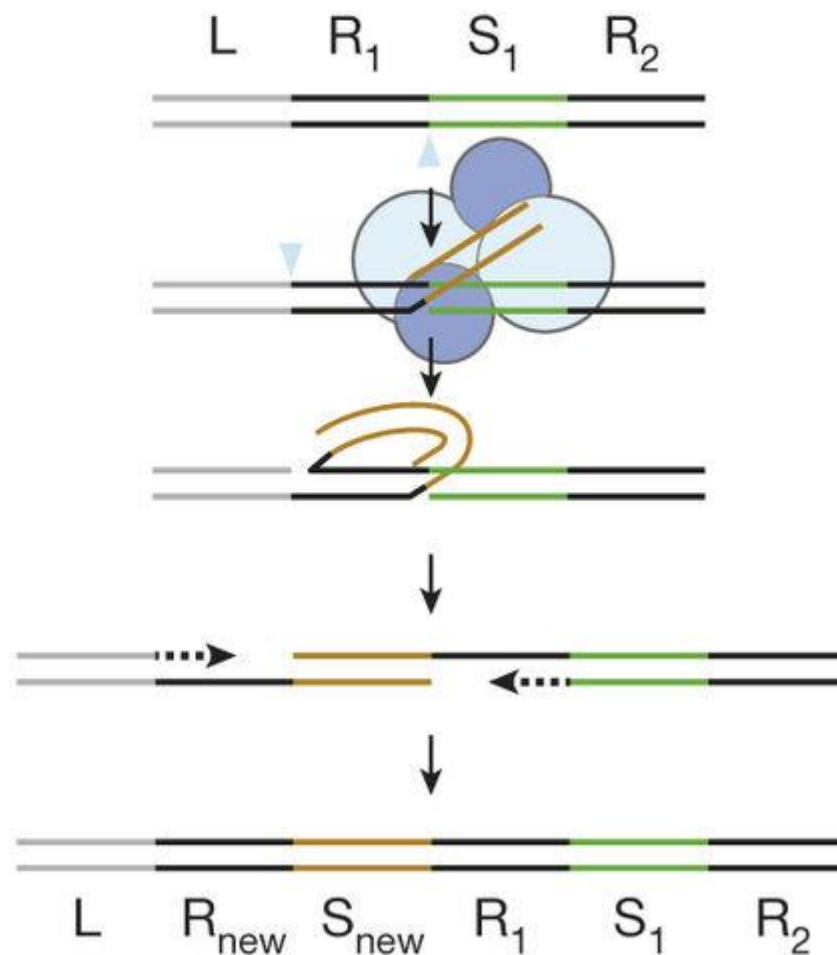


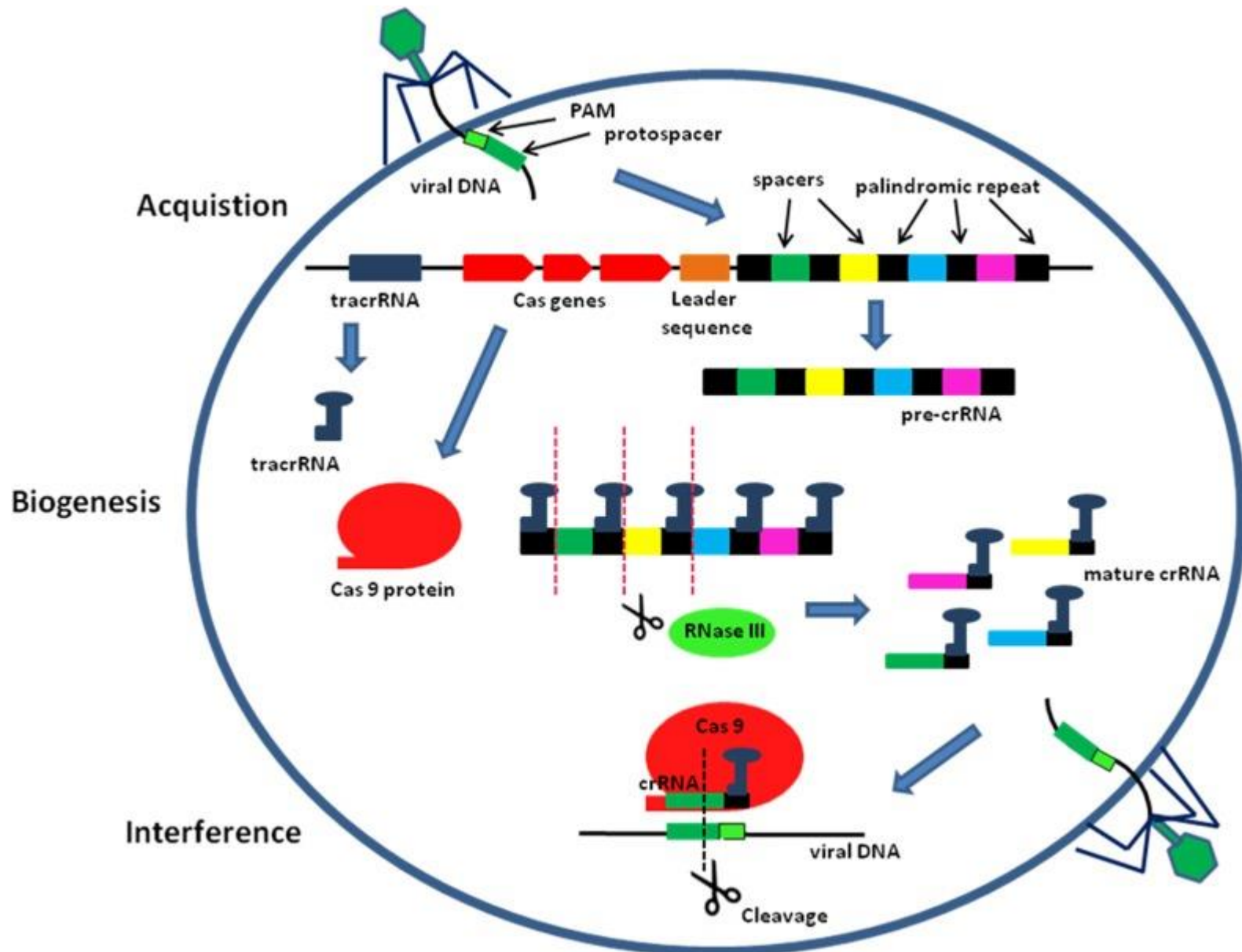
**a**

## Spacer sampling

**b**

## Spacer integration





# Cas9 - endonuclease - bacterial adaptive immunity (CRISPR associated protein 9)

## Multiple genes:

### Cas9 binds 2 small RNAs:

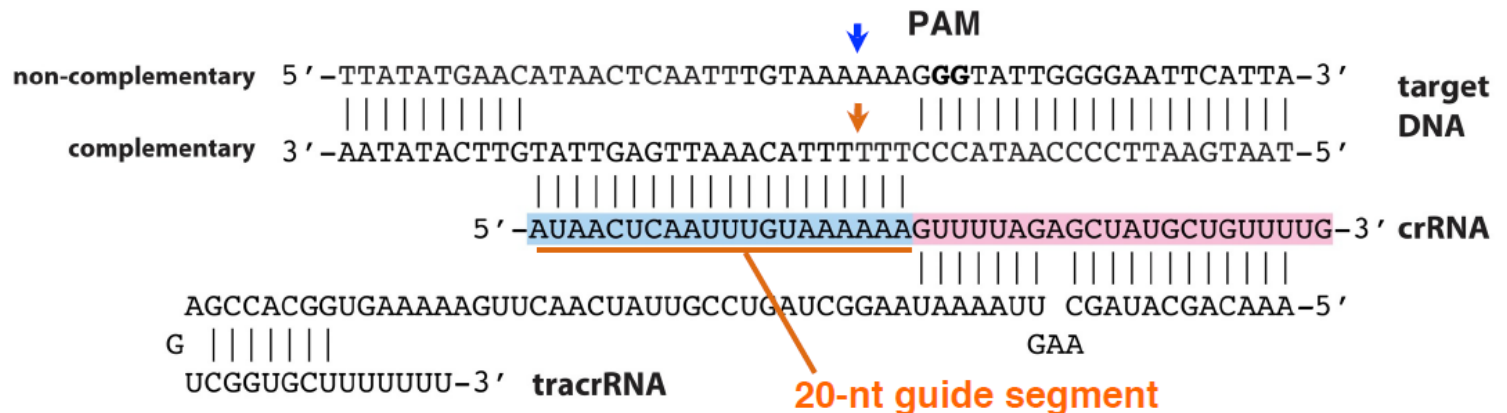
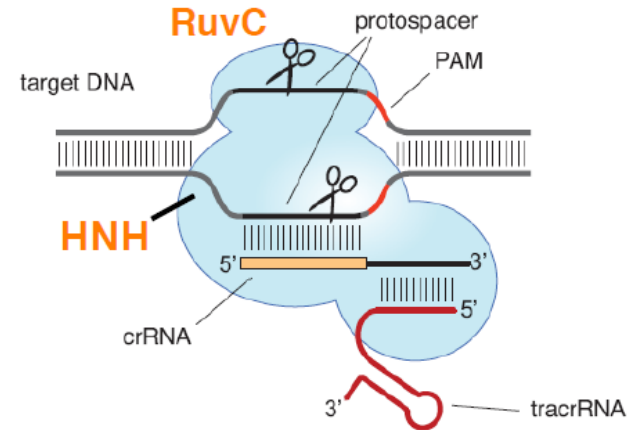
#### trans-activation RNA (tracrRNA)

- activation of Cas9 enzyme activity

#### CRISPR RNA (crRNA)

- substrate specificity - pairing with target DNA

#### protospacer adjacent motif (PAM) - recognized by Cas9



# Requirements for Cas9 mediated cleavage

## 8-12 nt "Seed" sequence in crRNA

near-perfect complementarity is critical for target binding

## Distal base-pairing

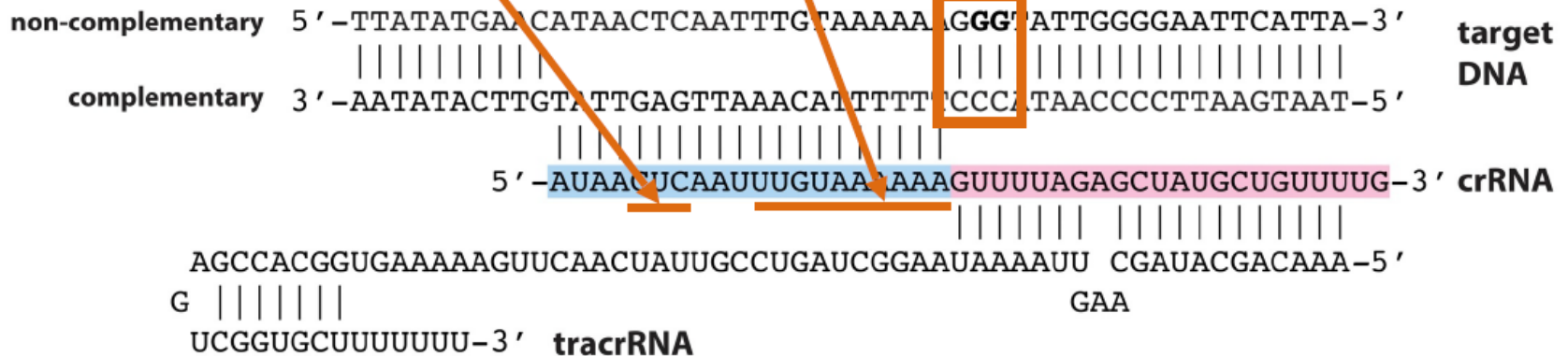
nt 16-17 base pairing is important for cleavage

## PAM motif in target DNA

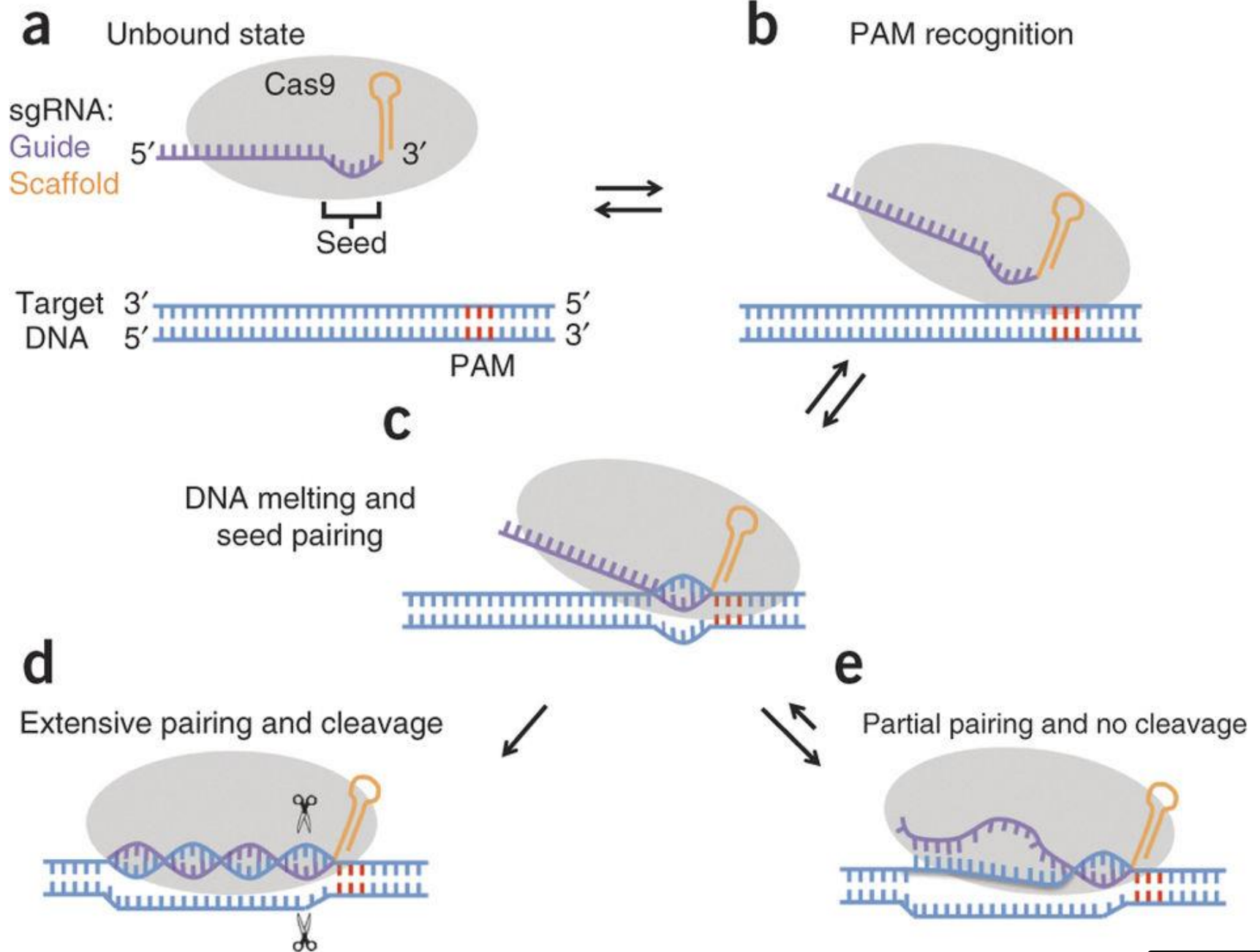
(**NGG** for *S. pyogenes* Cas9)

Cas9 does not recognise perfectly complementary target in the absence of PAM

**PAM**

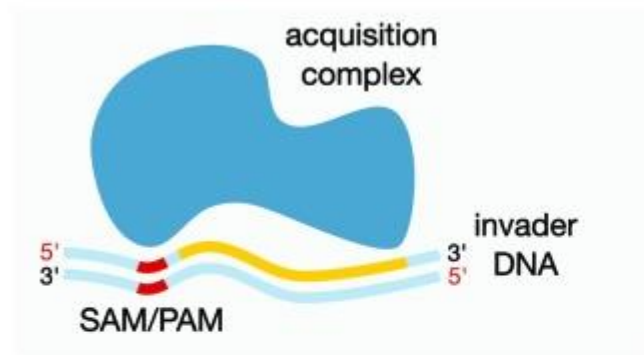






## A Type I

### SAM/PAM recognition

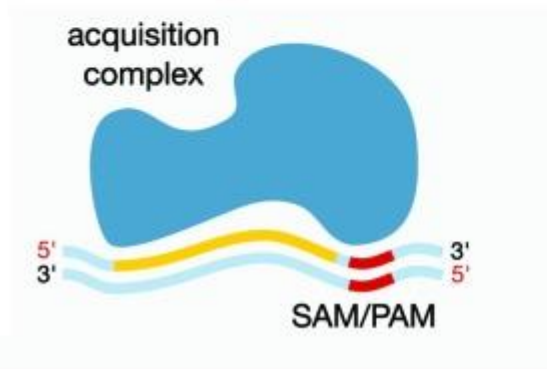


### TIM recognition

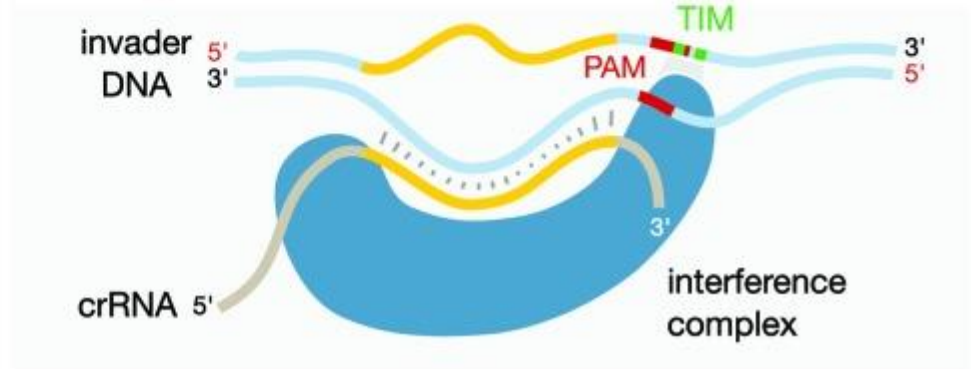


## B Type II

### SAM/PAM recognition

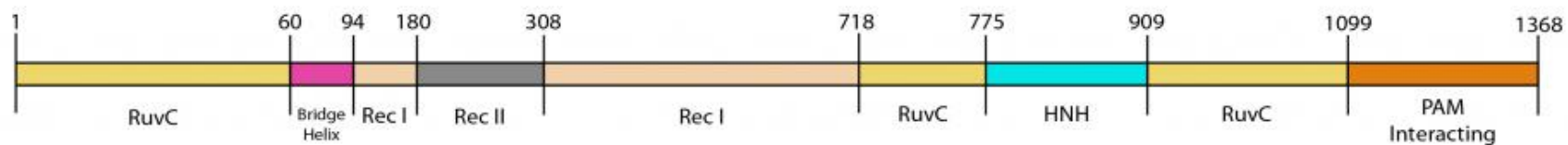
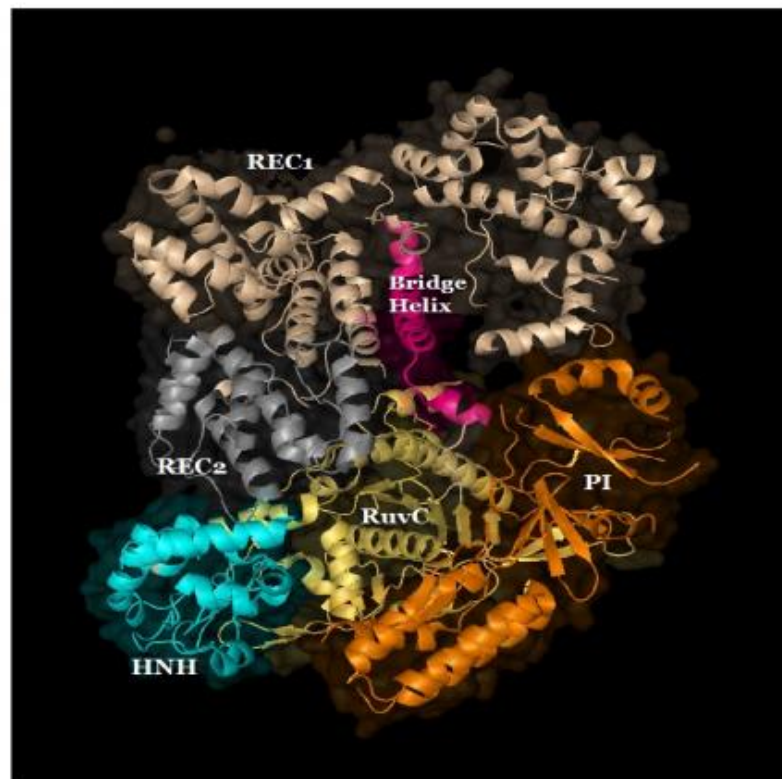
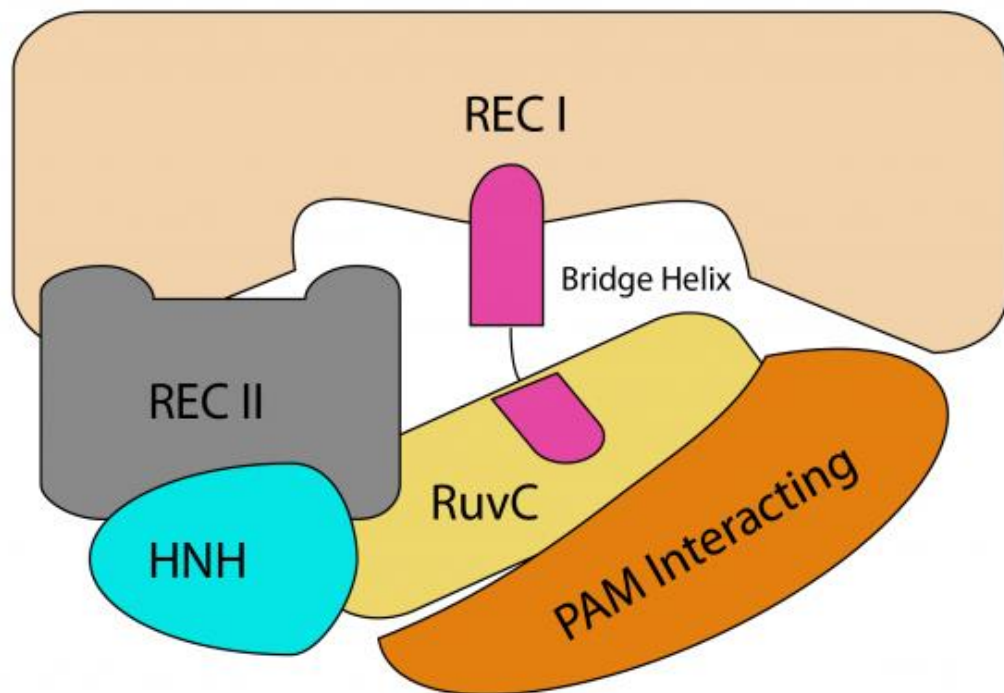


### TIM recognition



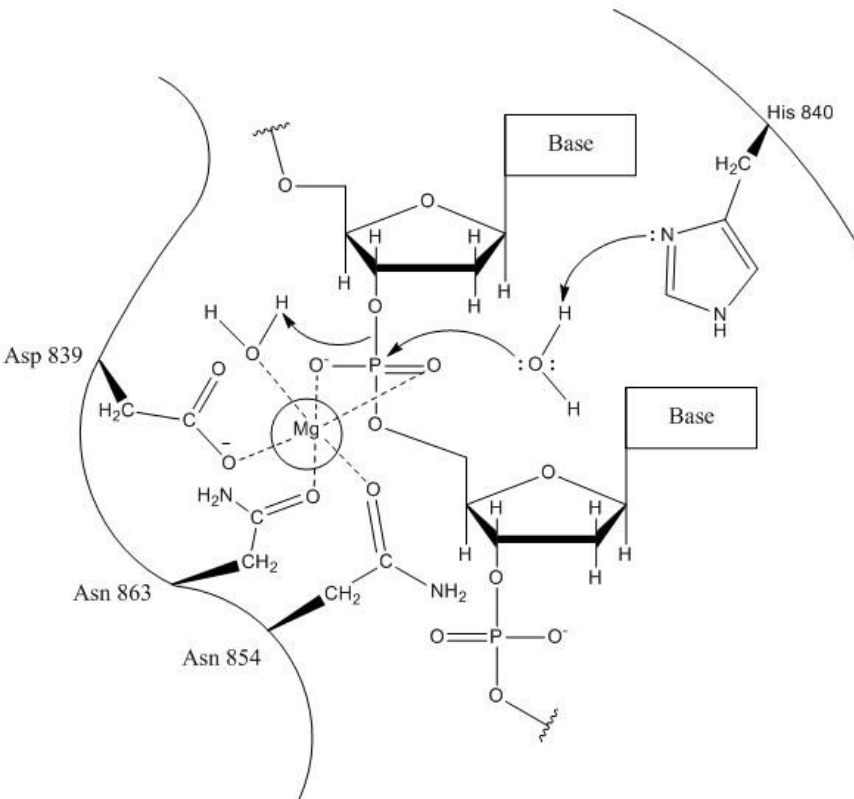
Overview of putative PAM, SAM and TIM interactions during acquisition and interference in type I and type II CRISPR systems. (A) The spacer acquisition motif (SAM) is recognized on the invader DNA by the Cas protein acquisition complex, which leads to the protospacer being excised by a putative ruler mechanism<sup>14</sup> and reinserted into a CRISPR locus by another putative ruler mechanism.<sup>42</sup> During interference by type I systems the target interference motif (TIM), on the crRNA-complementary DNA strand, is recognized by the Cas protein-crRNA complex where both TIM recognition and crRNA annealing are required for successful invader cleavage.<sup>29</sup> (B) In type II systems, the SAM/PAM motif is inferred to be recognized by a mechanism related to the type I systems but inverted on the dsDNA whereas TIM recognition occurs on the non-complementary DNA strand to the crRNA.



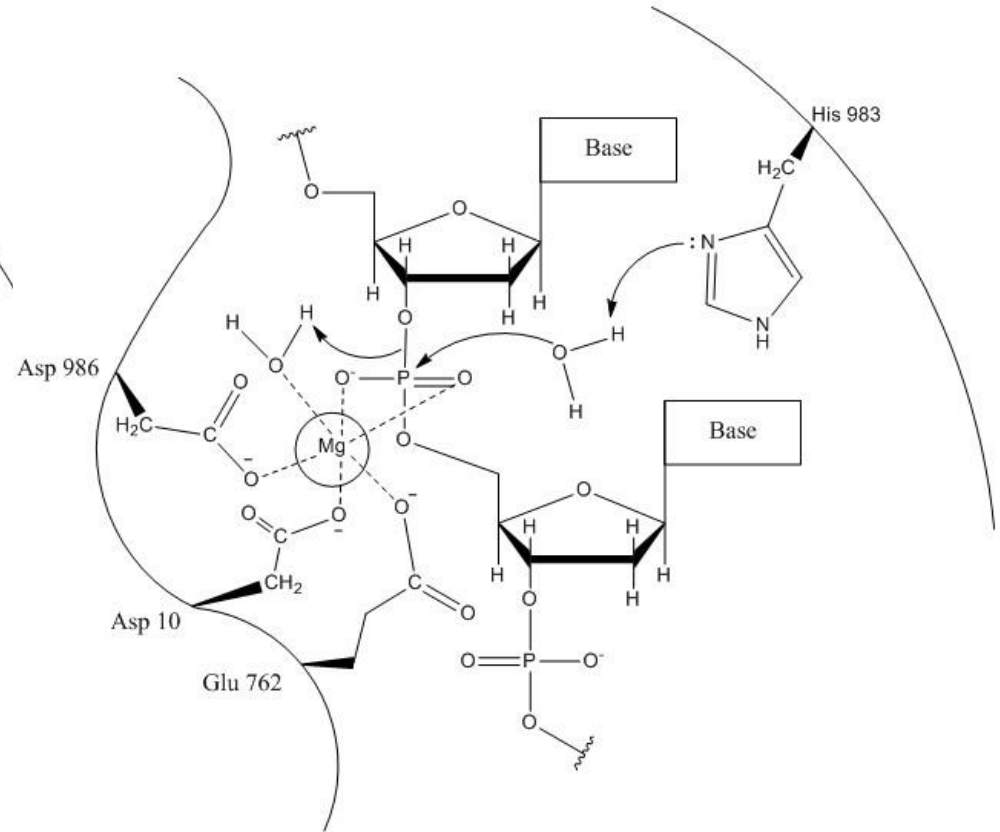


# ss endonucleases HNH a RuvC

## HNH



## RuvC



His activates a water molecule for attack of the scissile phosphate, which is made more electrophilic by coordination by a magnesium ion. Cleavage of the 3' – 5' phosphate bond

HNH named for characteristic histidine and asparagine residues

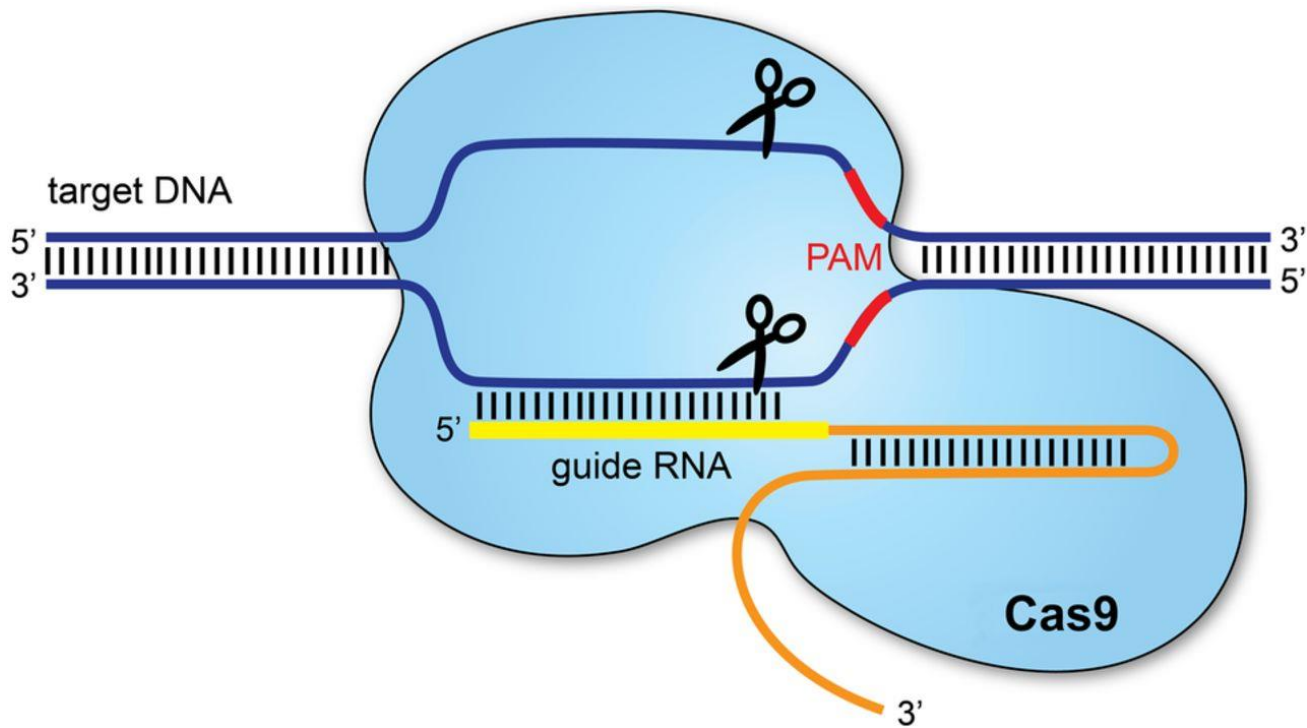
tracrRNA and crRNA can be combined into chimeric **single guide RNA (sgRNA)**: specific cleavage of target DNA

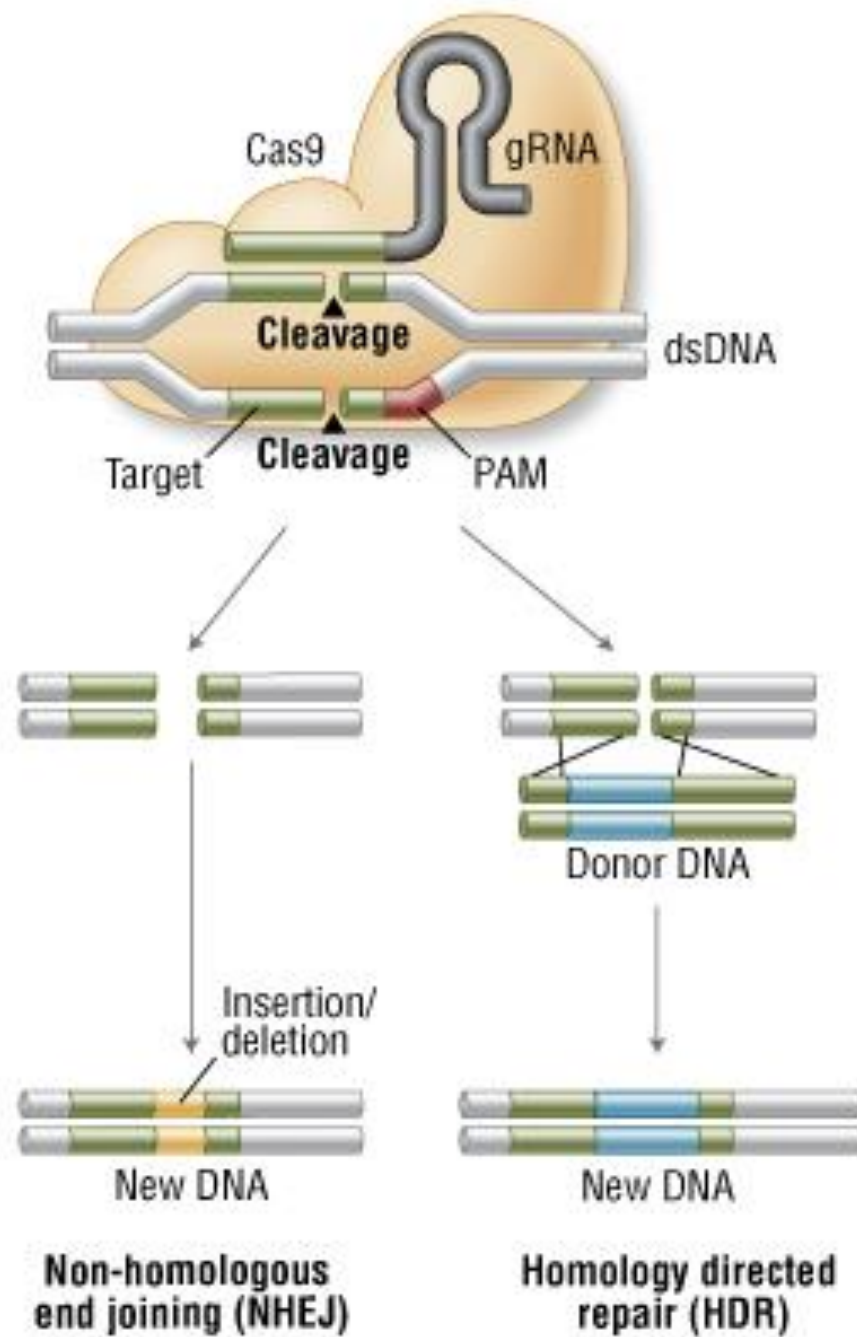
**2 conditions:**

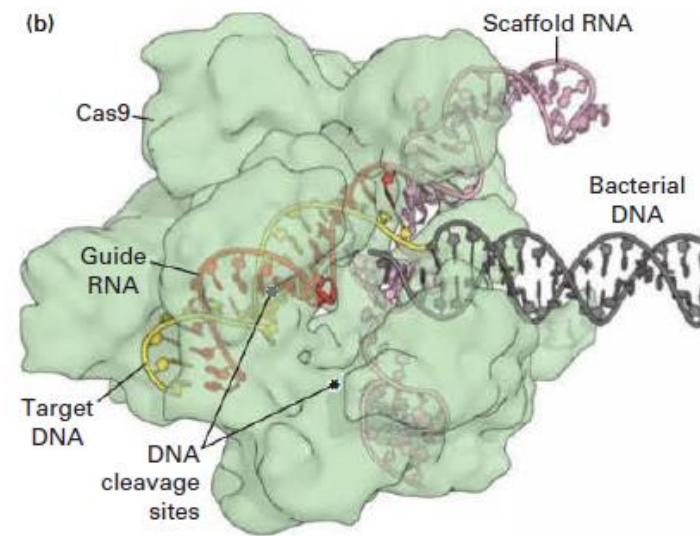
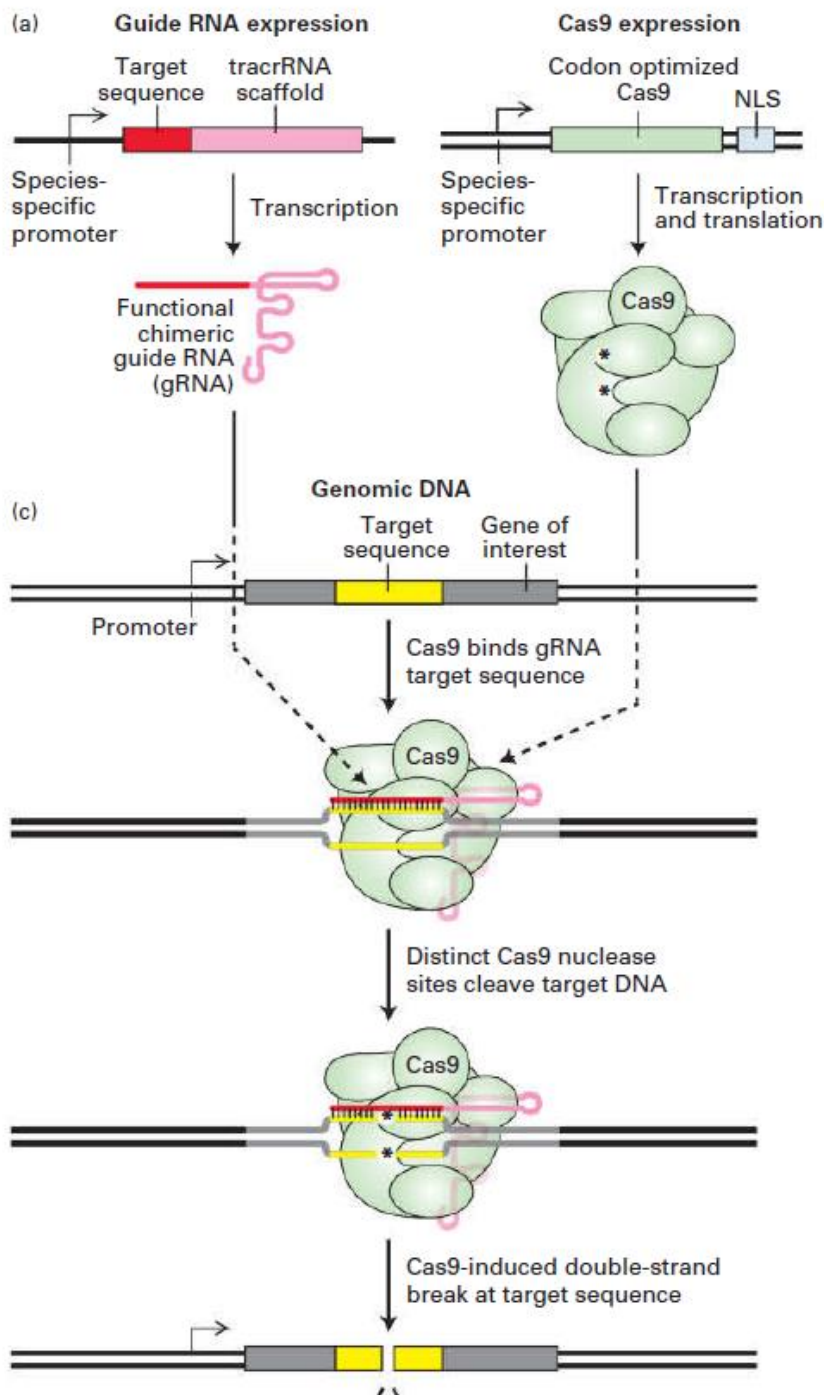
- 1) pairing of 5' terminal part of sgRNA with target DNA
- 2) NGG motif (PAM) after target sequence

**Sequence at 5' end of sgRNA determines the specific targeting of Cas9**

Cas9 employed for modification of genomes in human cells, rodents, fruit-flies, yeasts and bacteria.



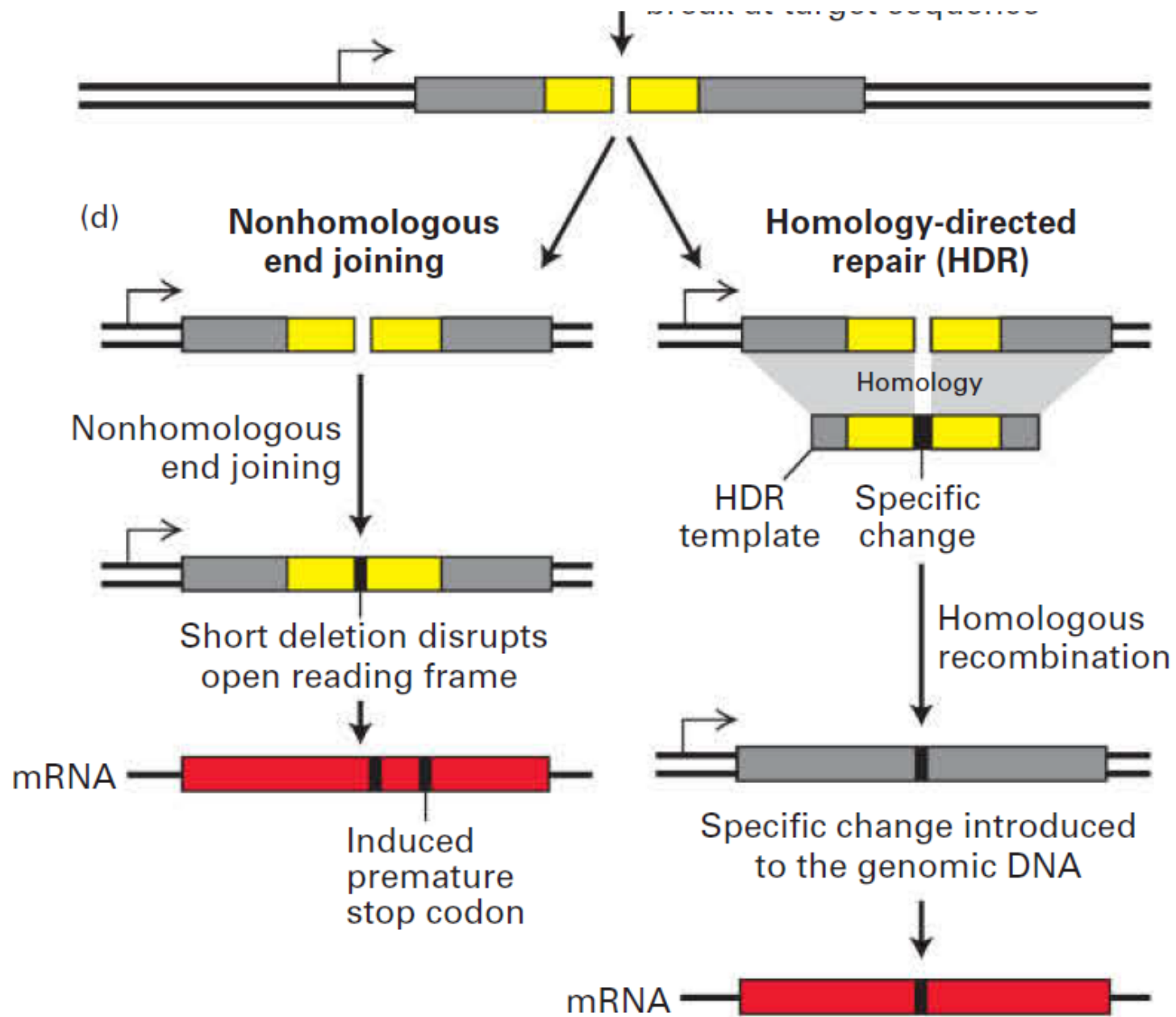




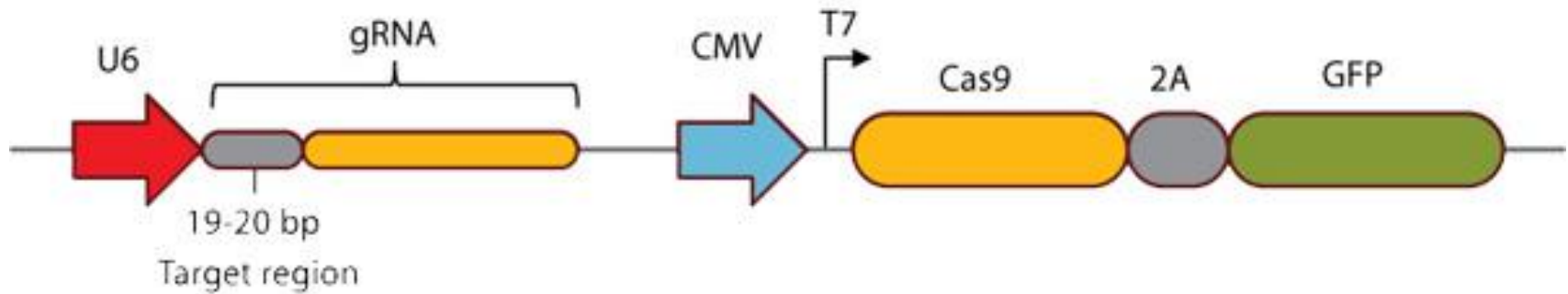
**Molecular Cell Biology**, Eighth Edition by Harvey Lodish et al. 2016, W.H. Freeman N.Y

**FIGURE 6-43 Single-nucleotide mutations can be introduced into the genome using an engineered CRISPR-Cas9 system.** (a) The genome of a target cell can be modified by expression of the double-stranded DNA endonuclease Cas9 and a guide RNA. Expression of these components can be achieved by transfection with plasmids carrying genes for Cas9 and the guide RNA or by direct injection of Cas9 mRNA and guide RNA. The guide RNA is composed of two parts: a sequence that folds into a hairpin scaffold structure that binds to Cas9, and a sequence of approximately 20 nt corresponding to the targeted site in the genome. Expression of these components can be achieved by transfection with plasmids carrying genes for Cas9 and the guide RNA or by direct injection of Cas9 mRNA and guide RNA. (b) A complex of guide RNA bound to Cas9 is





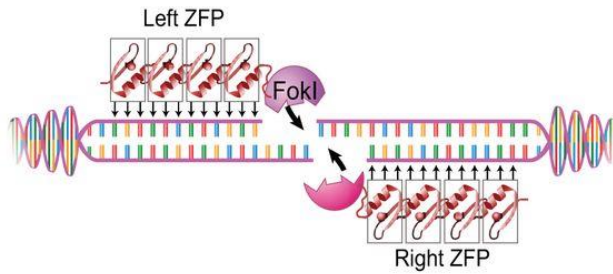
**Vector encoding Cas9, gRNA with selected target sequence  
and GFP to detect transfection efficiency**



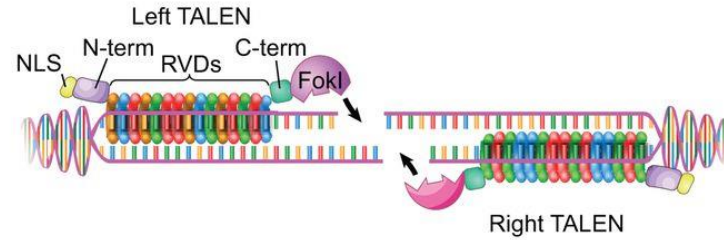
CRISPR/Cas Single Plasmid Format



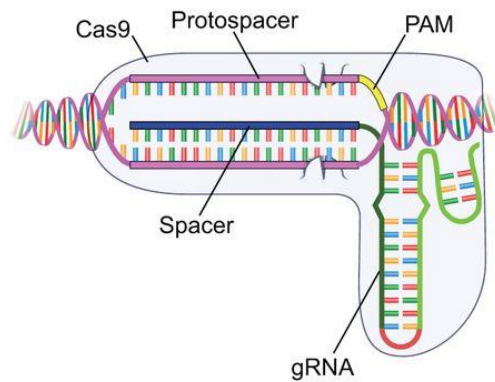
**A ZFN**



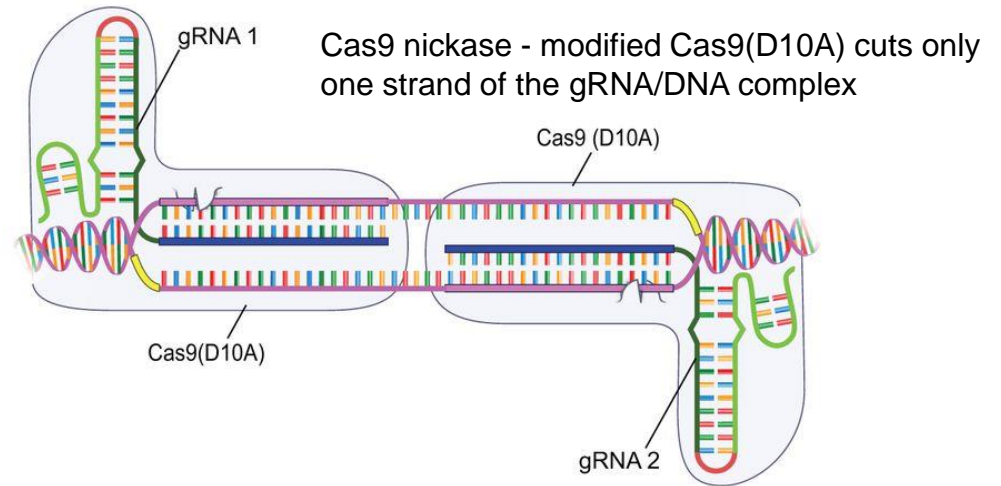
**B TALEN**



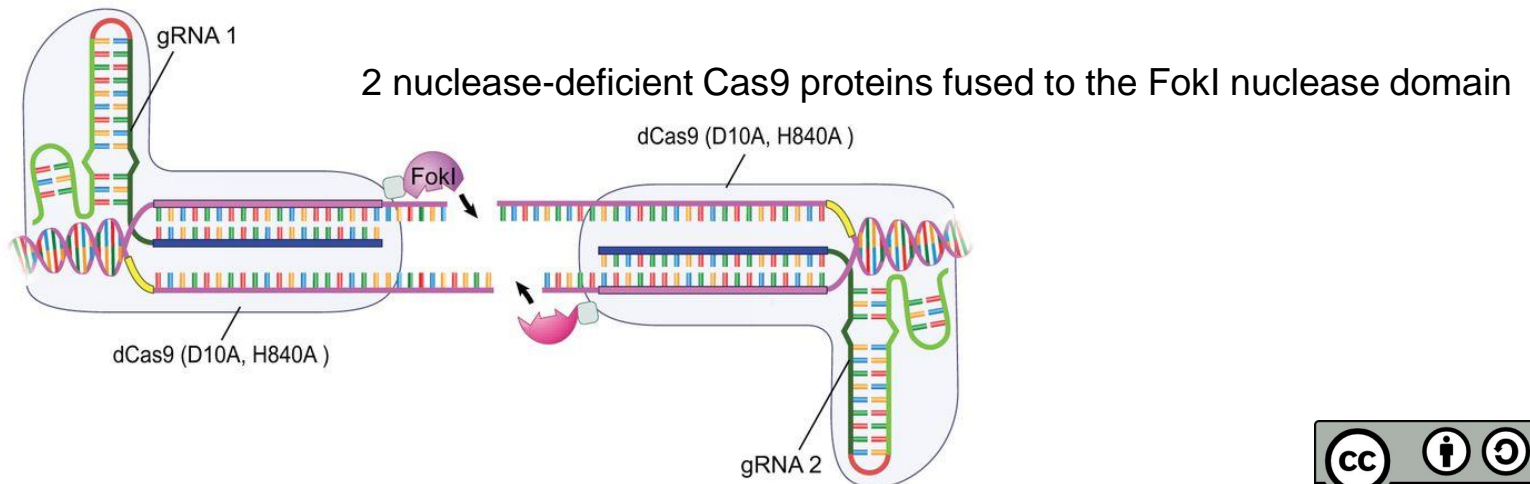
**C CRISPR/Cas9**



**D Cas9 nickase**



**E FokI-dCas9**



# **LoxP-Cre recombination system**

knock out of genes in specific cell types / organs

Phage P1 - protein Cre

Replication P1 - long multimers are created

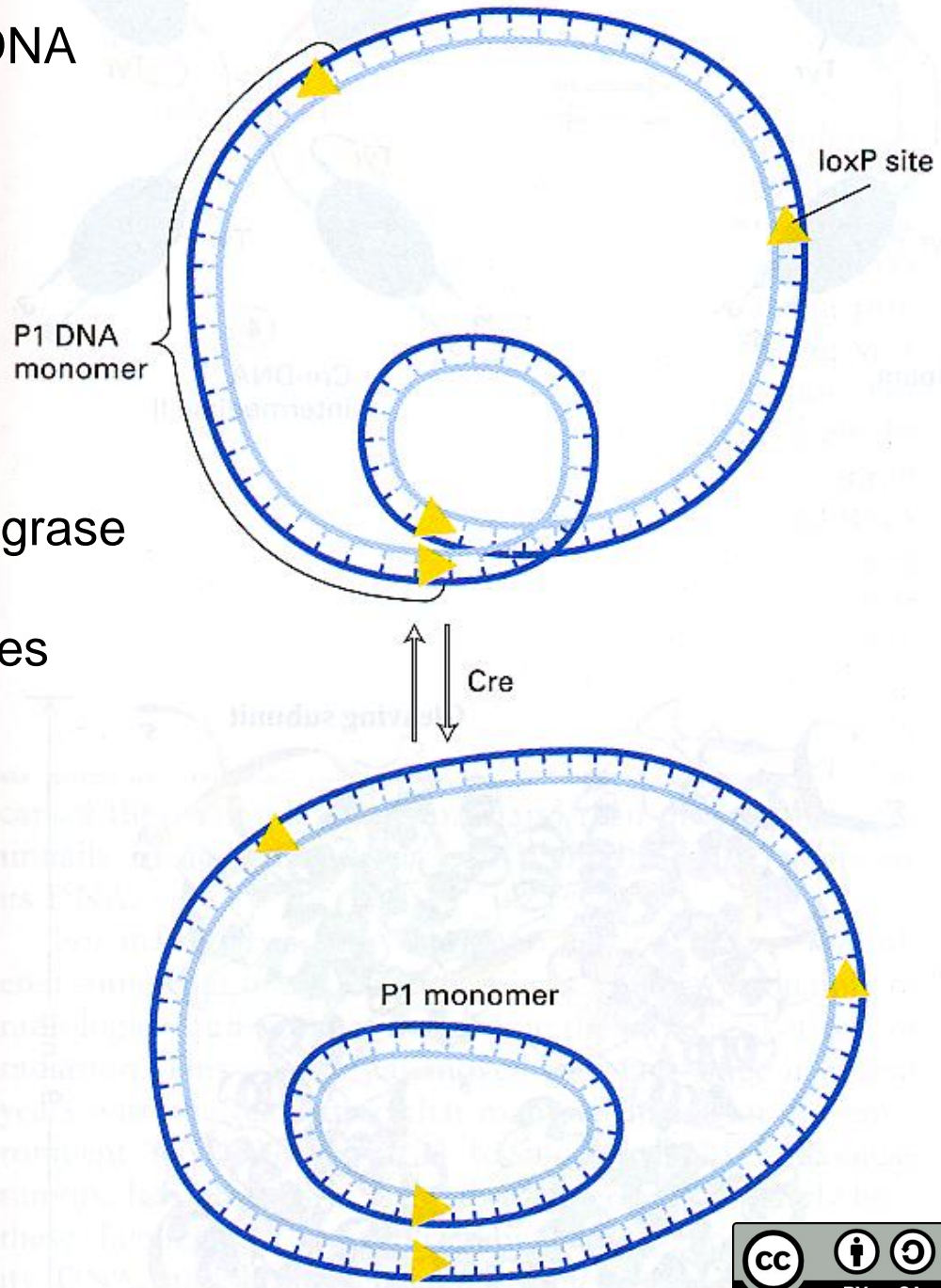
LoxP sites are responsible for creating monomers

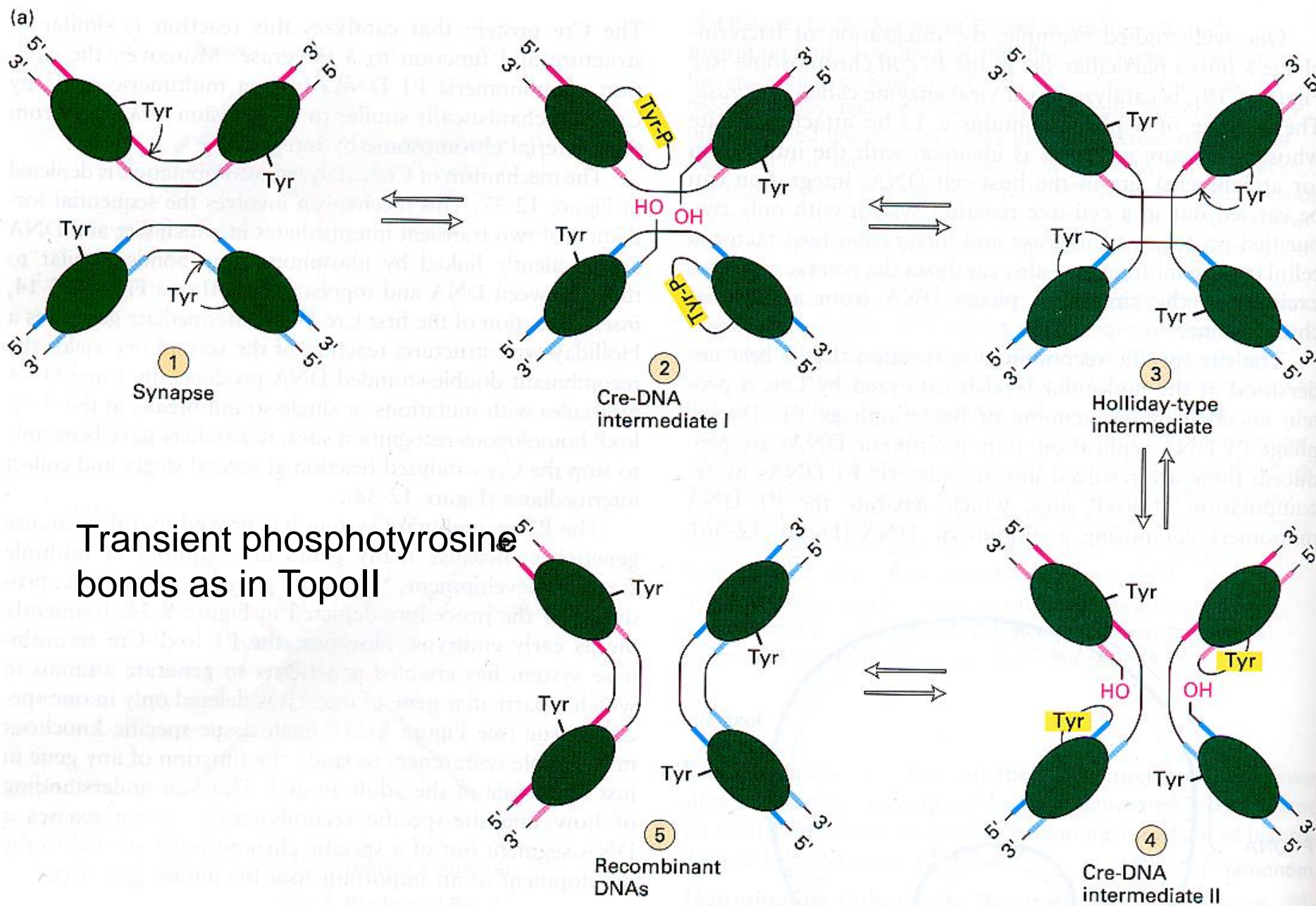
- a mechanism similar to integration
- intermediate Cre-DNA (phosphotyrosine linkage)

Phage P1 – replication - multimeric DNA

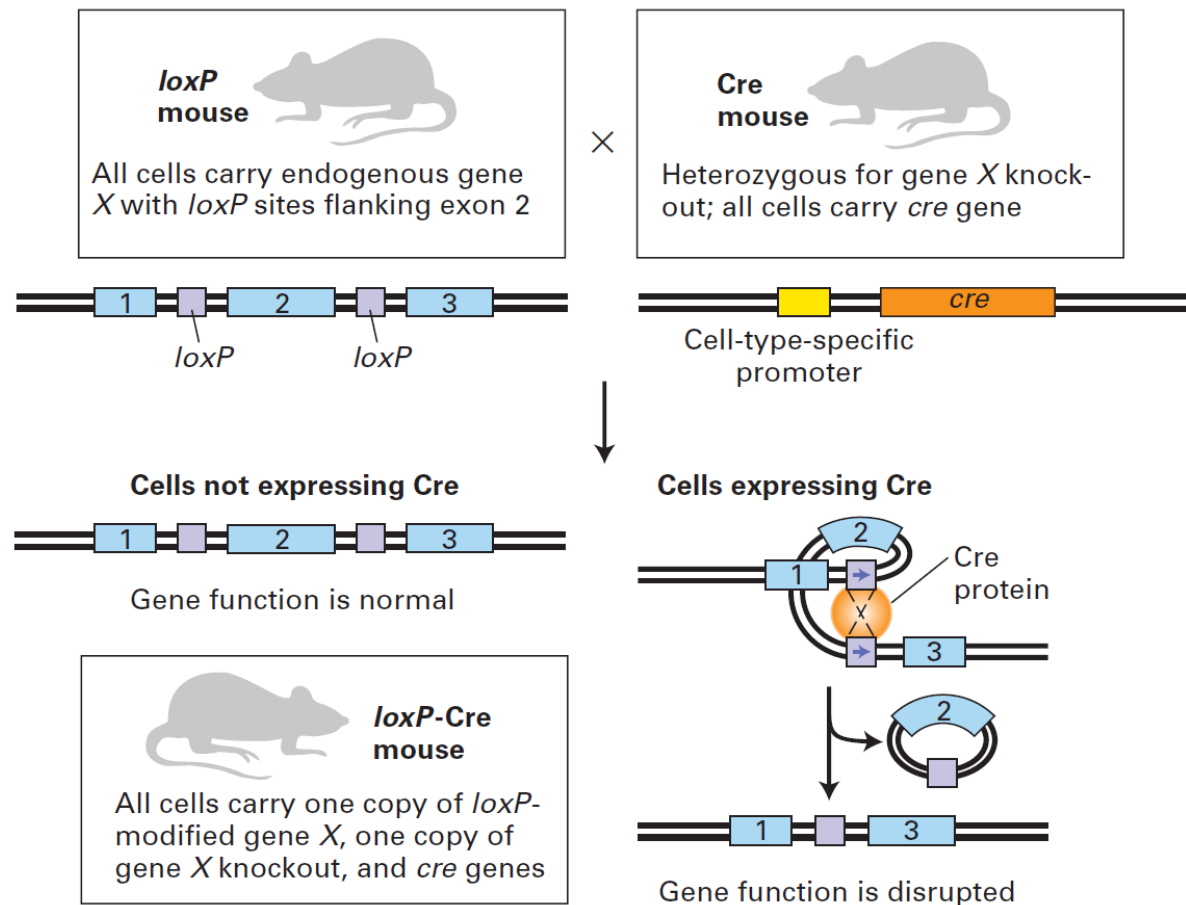
- Cre protein
- structure and function similar to integrase of  $\lambda$  phage
  - site-specific recombination in lox sites

Recombination – incorporation





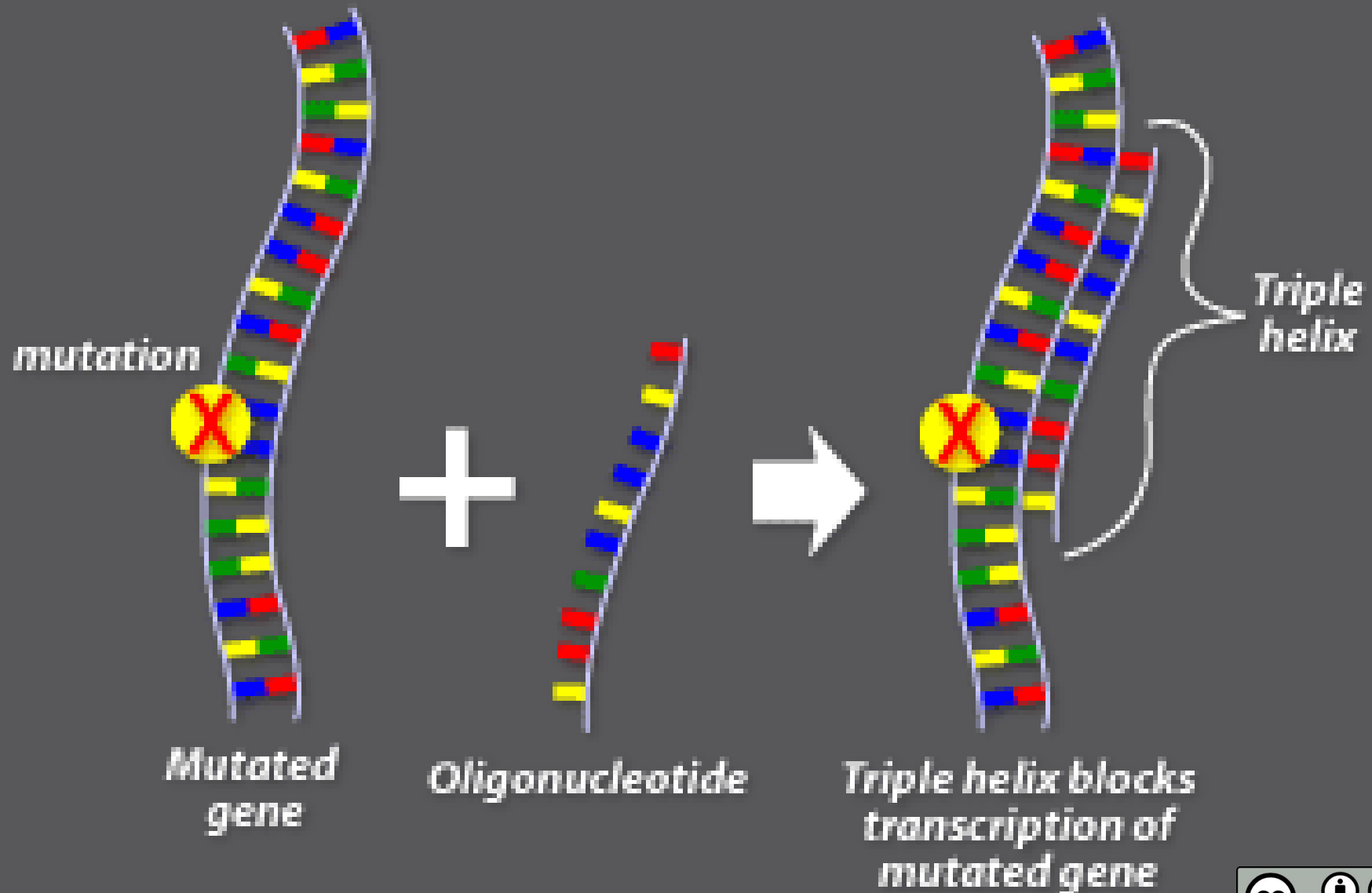




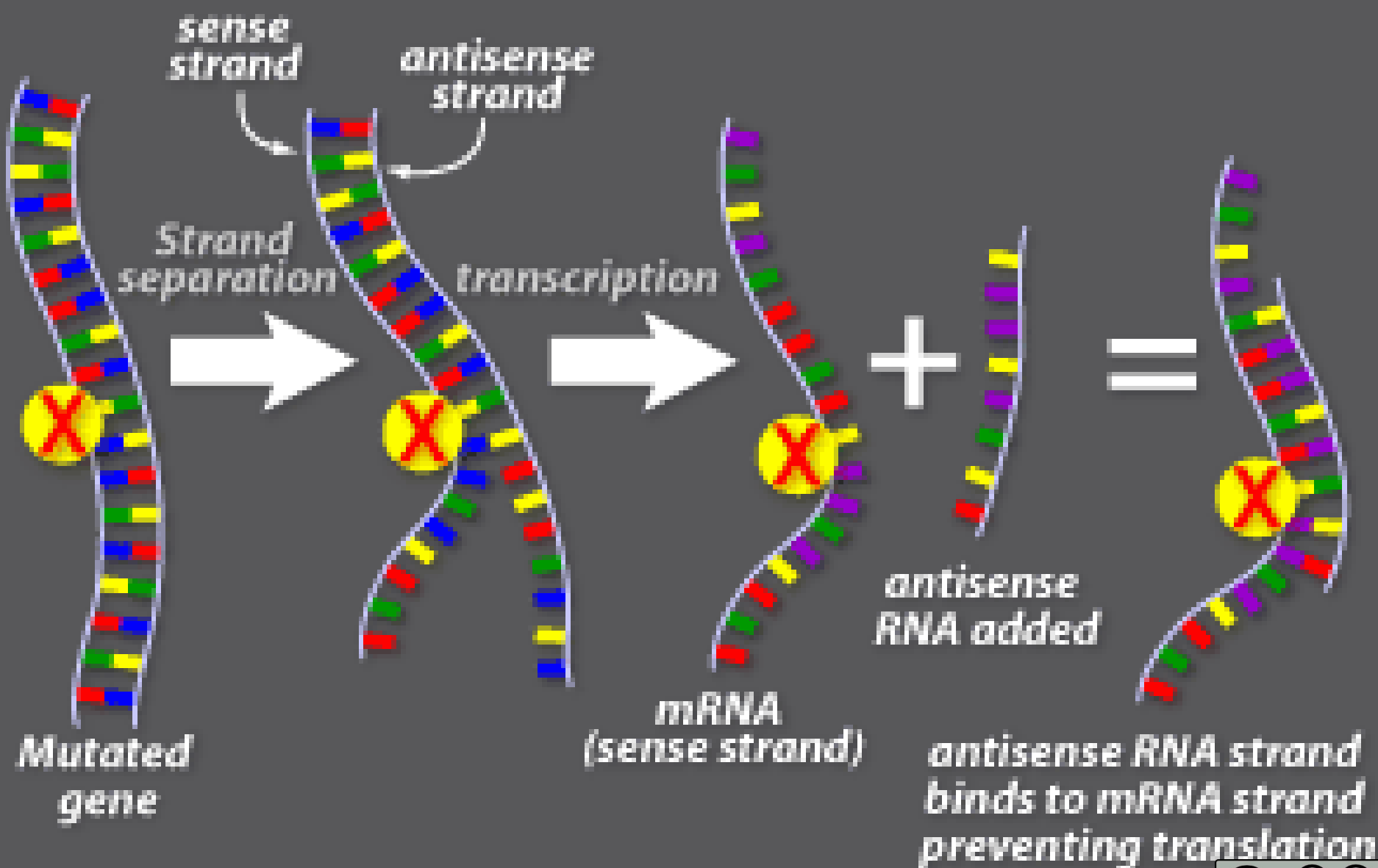
A *loxP* site (purple) is inserted on each side of an essential exon (exon 2) of the target gene (gene *X*; blue) by homologous recombination, producing a *loxP* mouse. *loxP* - in introns: do not disrupt the function of *X*.

A Cre mouse carries one gene *X* knockout allele and an introduced *cre* gene (orange) from bacteriophage P1 linked to a cell-type-specific promoter (yellow). The incorporated *cre* gene does not affect the function of other genes. In the *loxP*-Cre mice that result from crossing these two types of mice, Cre protein is produced only in those cells in which the promoter is active. Thus these are the only cells in which recombination between the *loxP* sites catalyzed by Cre occurs, leading to deletion of exon 2. Since the other allele is a constitutive gene *X* knockout, deletion between the *loxP* sites results in complete loss of function of gene *X* in all cells expressing Cre. By using different promoters, researchers can study the effects of knocking out gene *X* in various types of cells.

# Preventing Transcription of a Mutated Gene Using Triple-helix-forming Oligonucleotides



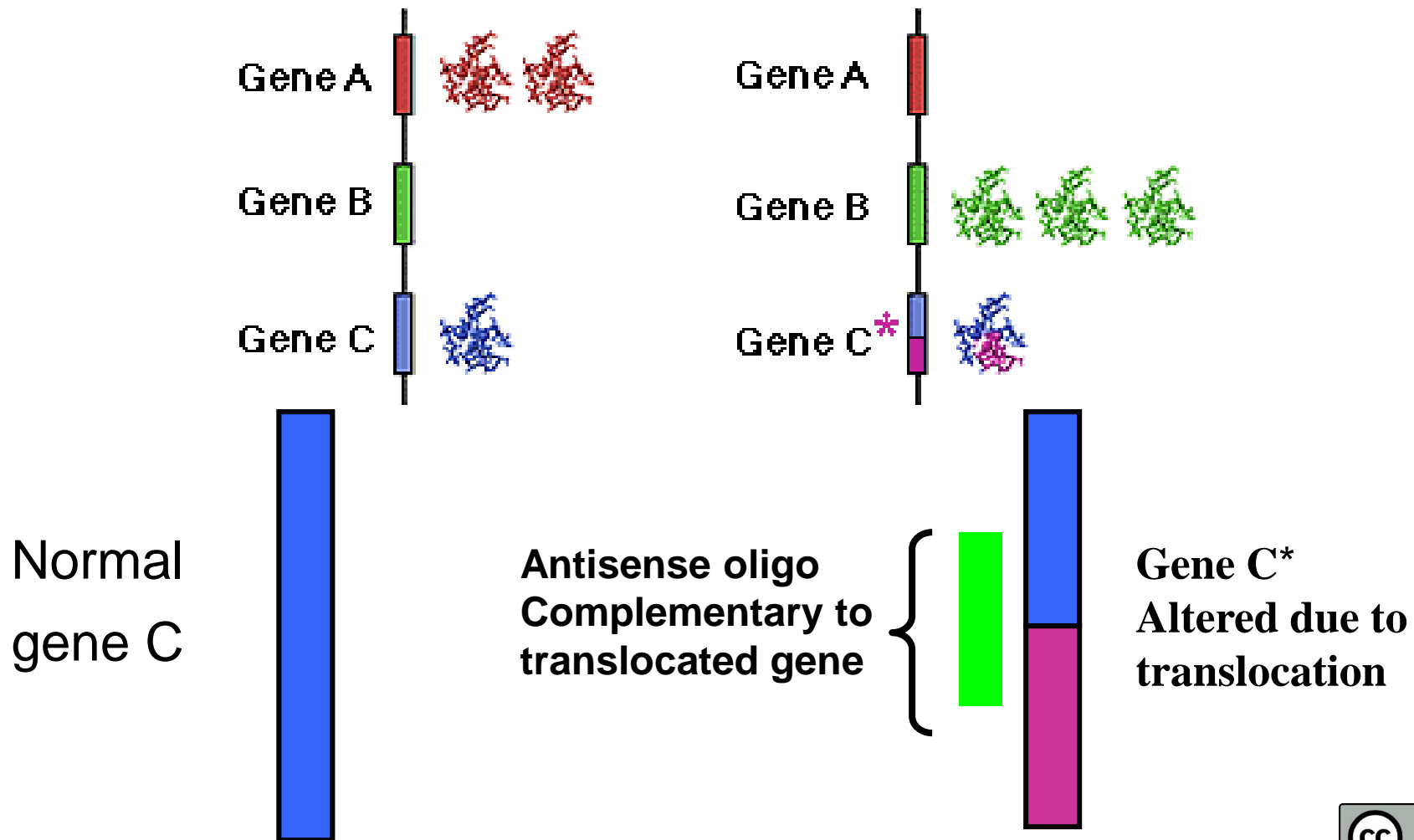
# Preventing Translation Using Antisense Technology





## Target sequences

- mRNA of gene overexpressed in cancer cell compared to normal cell (gene B)
- mRNA of mutant gene occurring only in cancer cells (Gen C\*)



# Benefits of antisense oligonucleotides

Antisense therapeutics block translation of mRNA – intervention preceding the synthesis of specific protein

*(Traditional medication – intervention following the synthesis of specific protein causing the disease)*

Antisense therapeutics block translation of only single specific mRNA  
– higher effectivity and less side effects

Antisense oligonucleotides – usually 16 nucleotides

# Vitravene®(fomivirsen): The World's First Antisense Drug (Novartis AG)



Antisense inhibitor CMVR: Vitravene

Virucidal – contrast to virostatic

Intravitreal application

Most common side-effect: ocular inflammation

# Limitations of antisense oligonucleotides

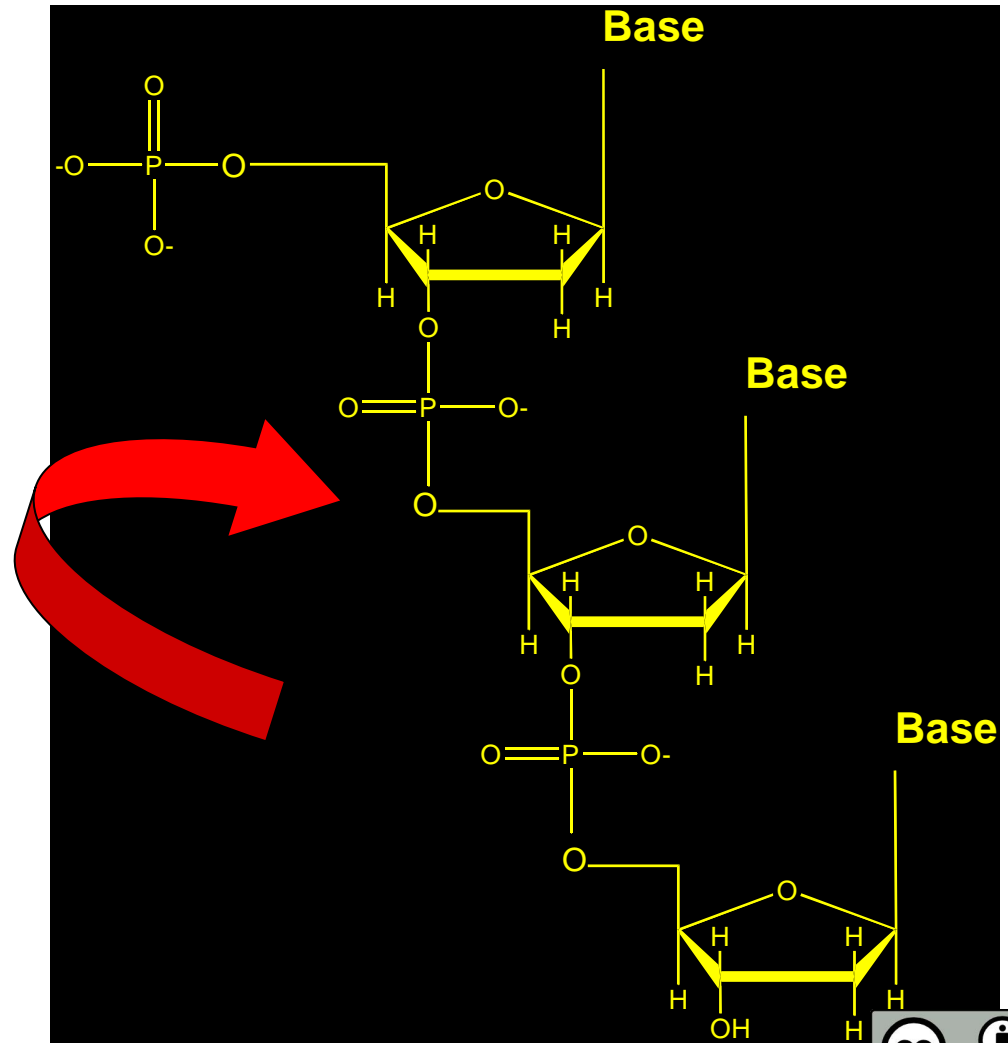
low stability - half-life:  $T_{1/2} = 1$  h in serum

low entry into the cell

the occurrence of "non-antisense" effects

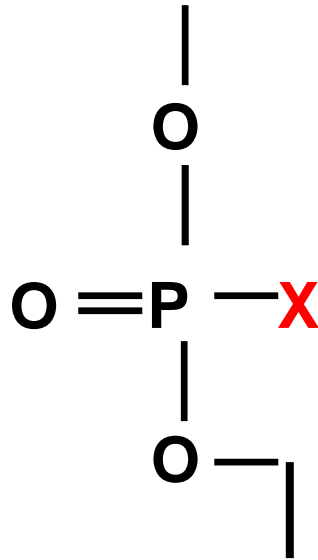
# Phosphodiester bonds – prone to degradation

Nuclease cleavage



# Oligonucleotide backbone

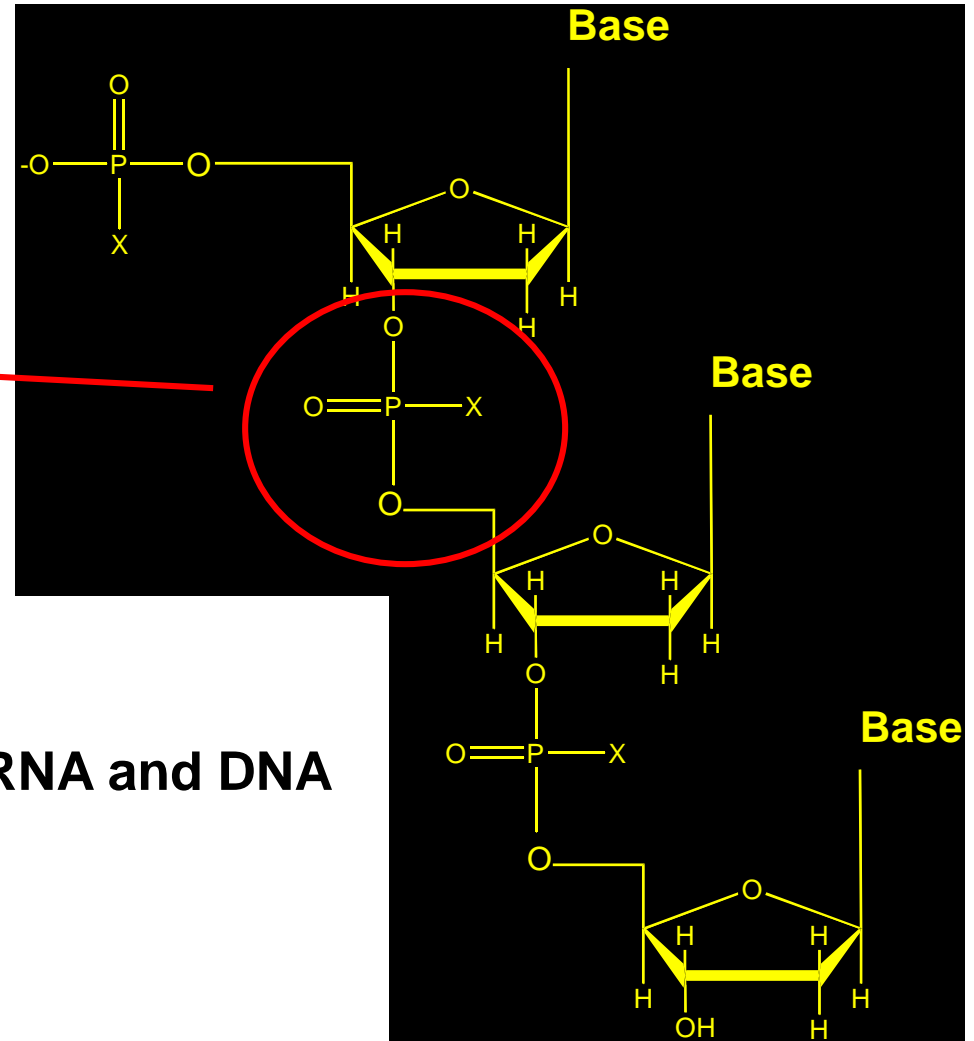
## - Modification – improved stability



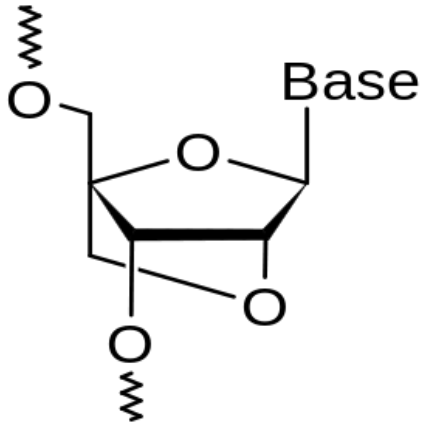
**X = O<sup>-</sup>**, Phosphodiester bond in RNA and DNA

**X = S<sup>-</sup>**, Phosphorothioate

**X = CH<sub>3</sub>**, methylphosphonate



# Locked nucleic acid (LNA)



2' O and 4' C bridge: "locks" ribose  
in the 3'-*endo* conformation

stable



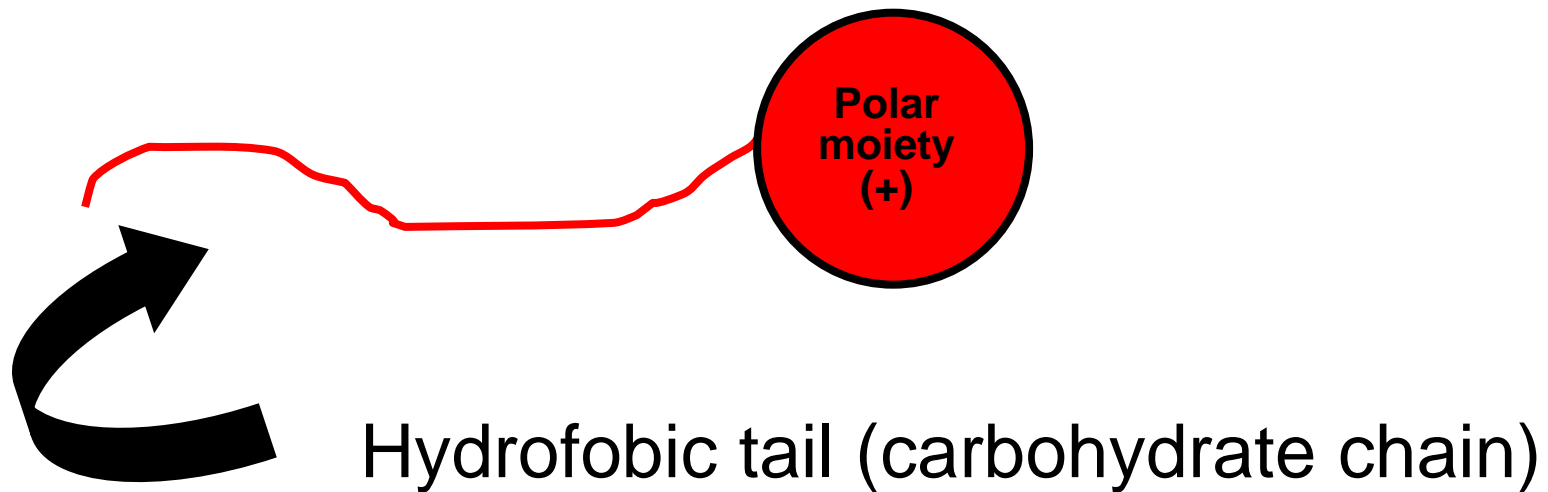
# Modification of ASO for the better uptake to cell

Conjugation of ASO with cationic surfactants

- Improved cell uptake
- liposomes

Microinjections

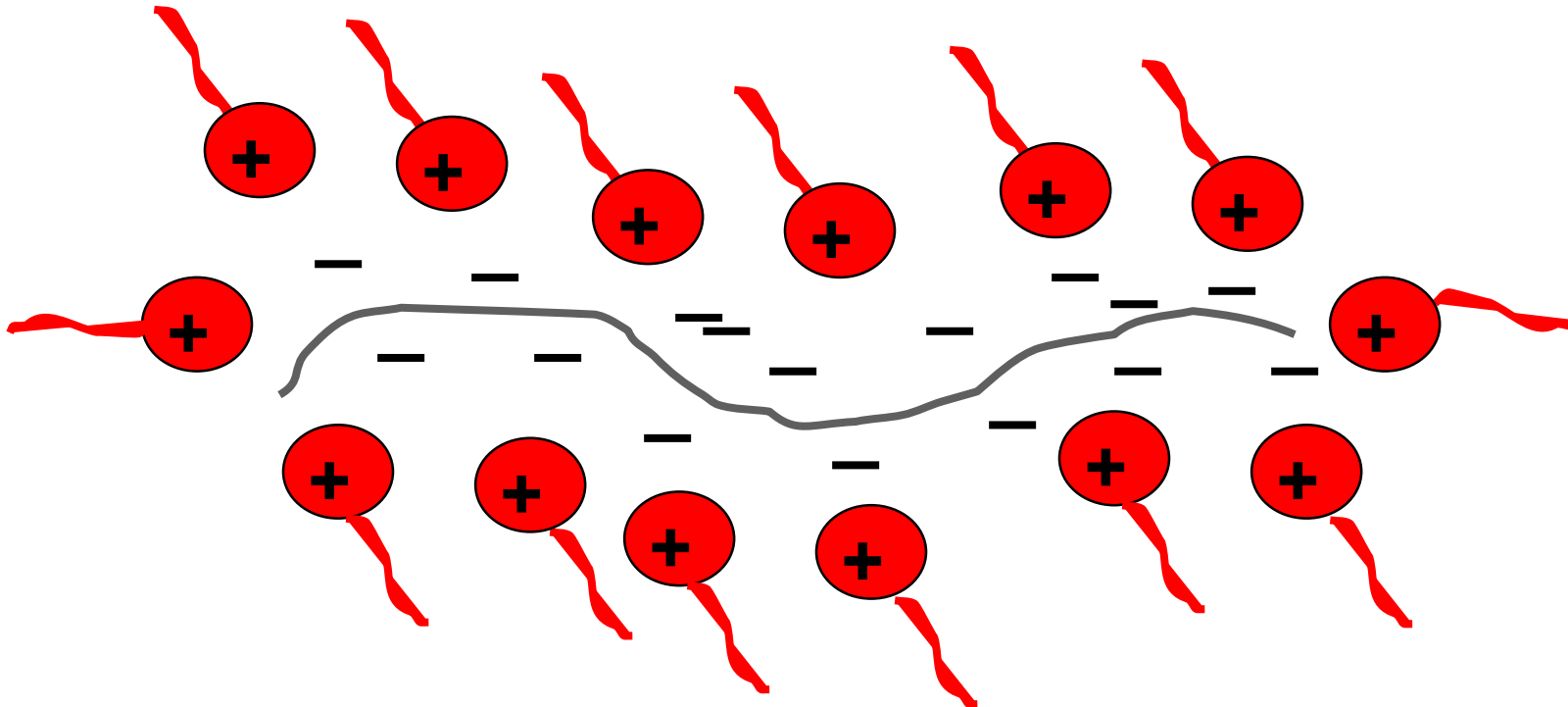
# Cationic surfactants



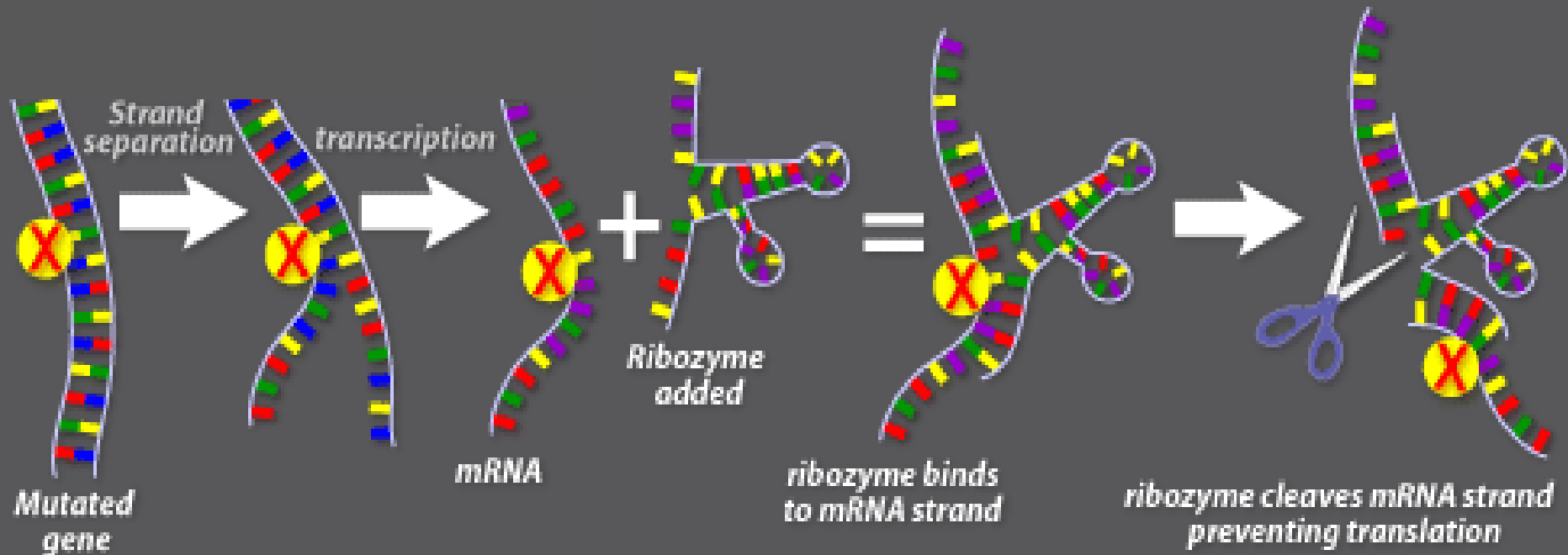
# Liposomes

Polar domains – neutralization of negative charge of antisense oligo

carbohydrate chain - penetration through cell membrane (lipid bilayer).



## Preventing Translation Using Ribozyme Technology

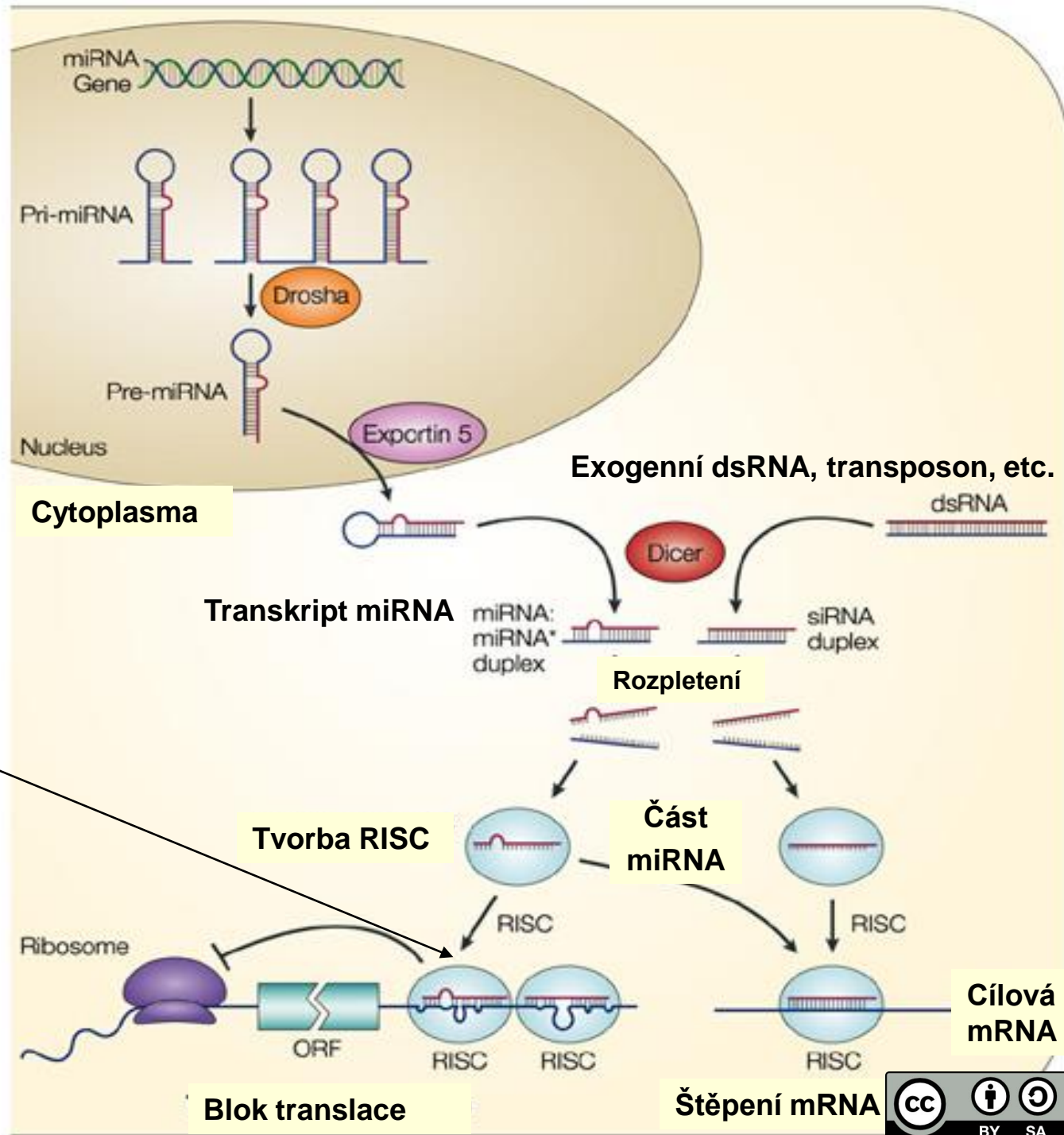


## miRNA x siRNA

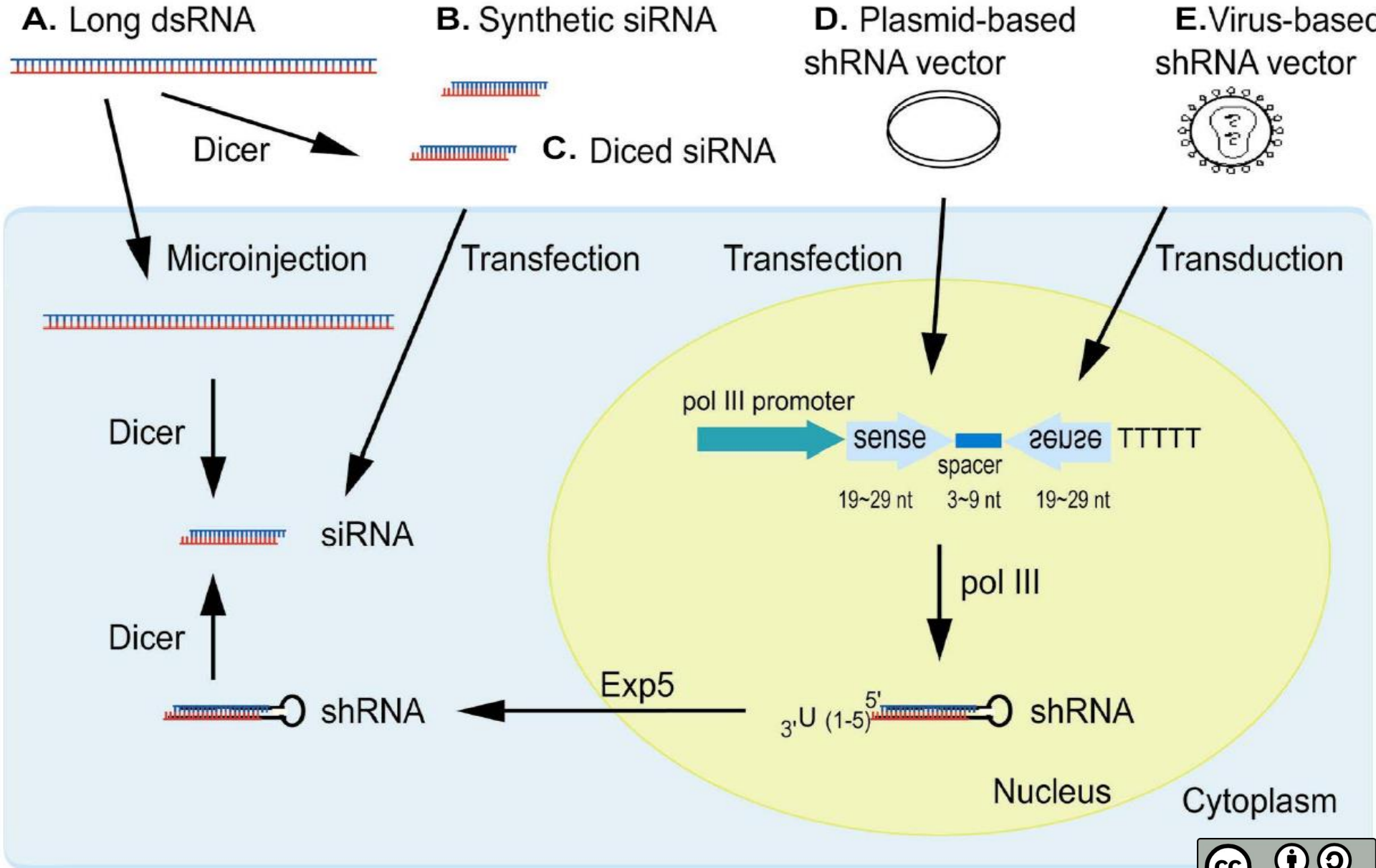
## Cleavage of mRNA x translation repression

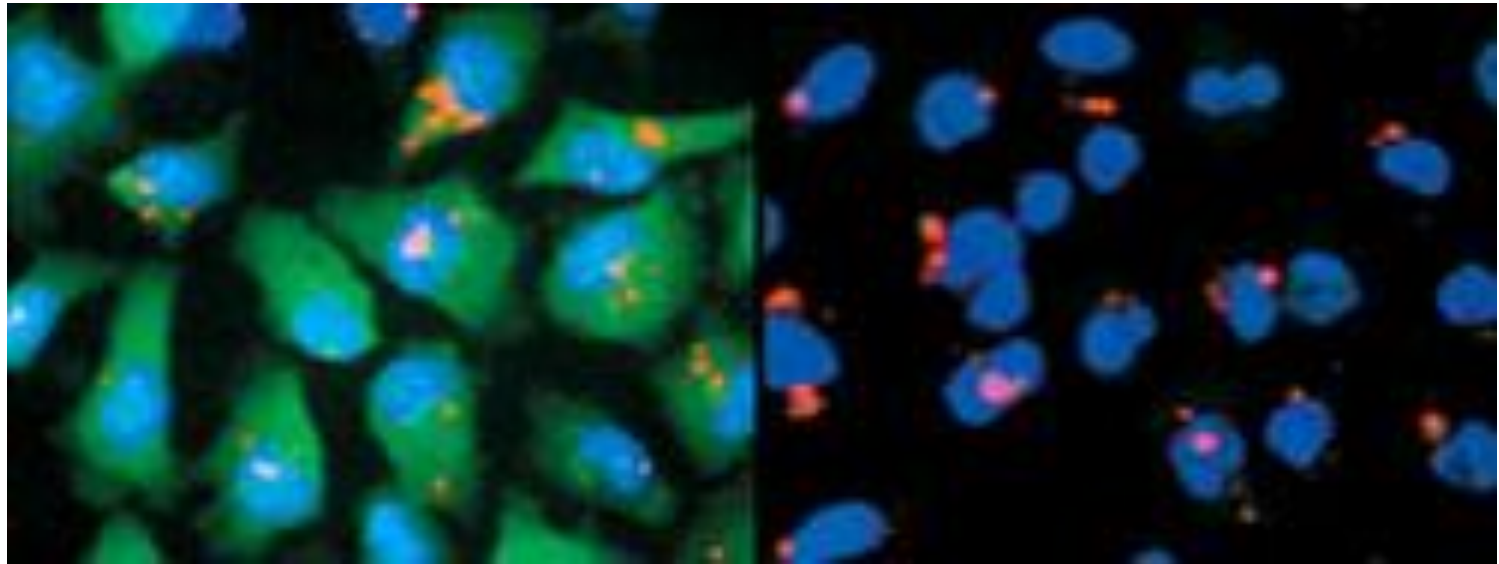
Imperfect pairing

animation



# siRNA





HeLa cells transfected with siRNA against GAPDH 48 h post- transfection

Red - siRNA

Green - GAPDH (Glyceraldehyde 3-phosphate dehydrogenase)

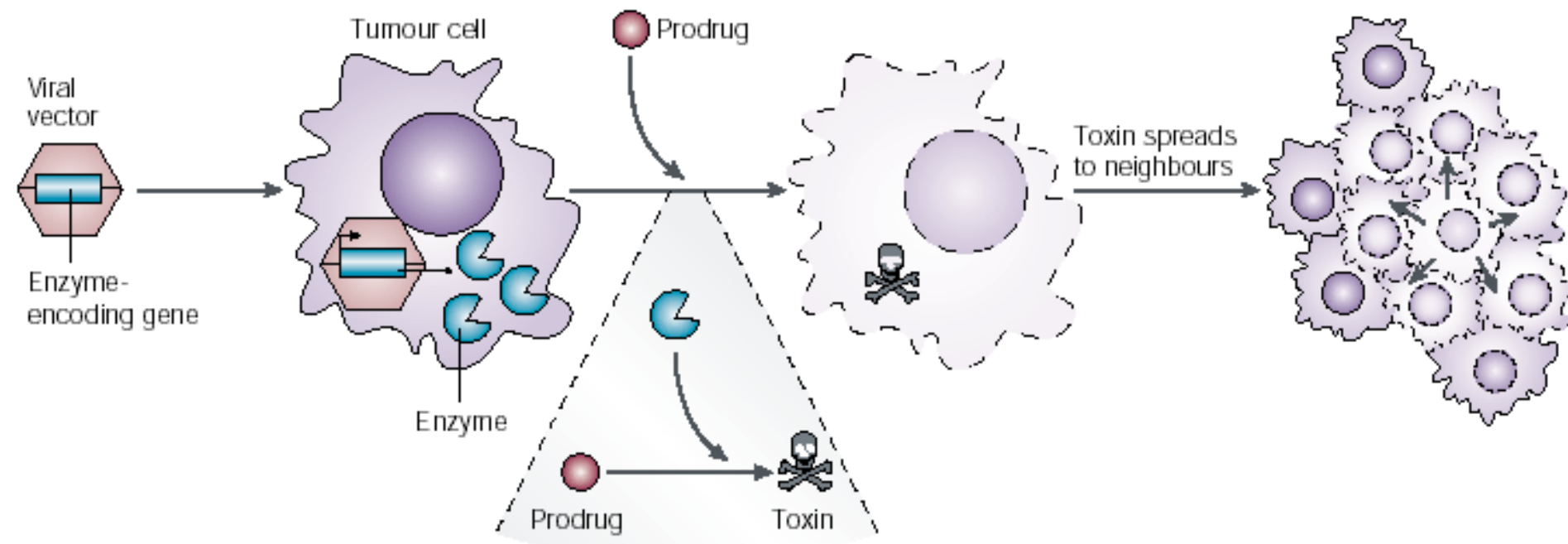
Non-functional (control) siRNA (left) - GAPDH expression

siRNA against GAPDH (right) – block of GAPDH protein expression



# Suicide gene therapy

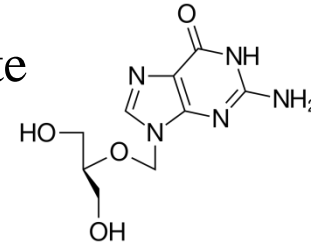
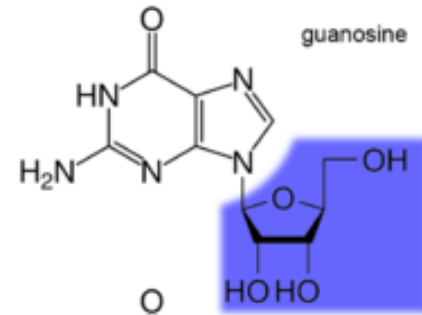
~ transduction of a gene that transforms a nontoxic „pro-drug“ into toxic substance.



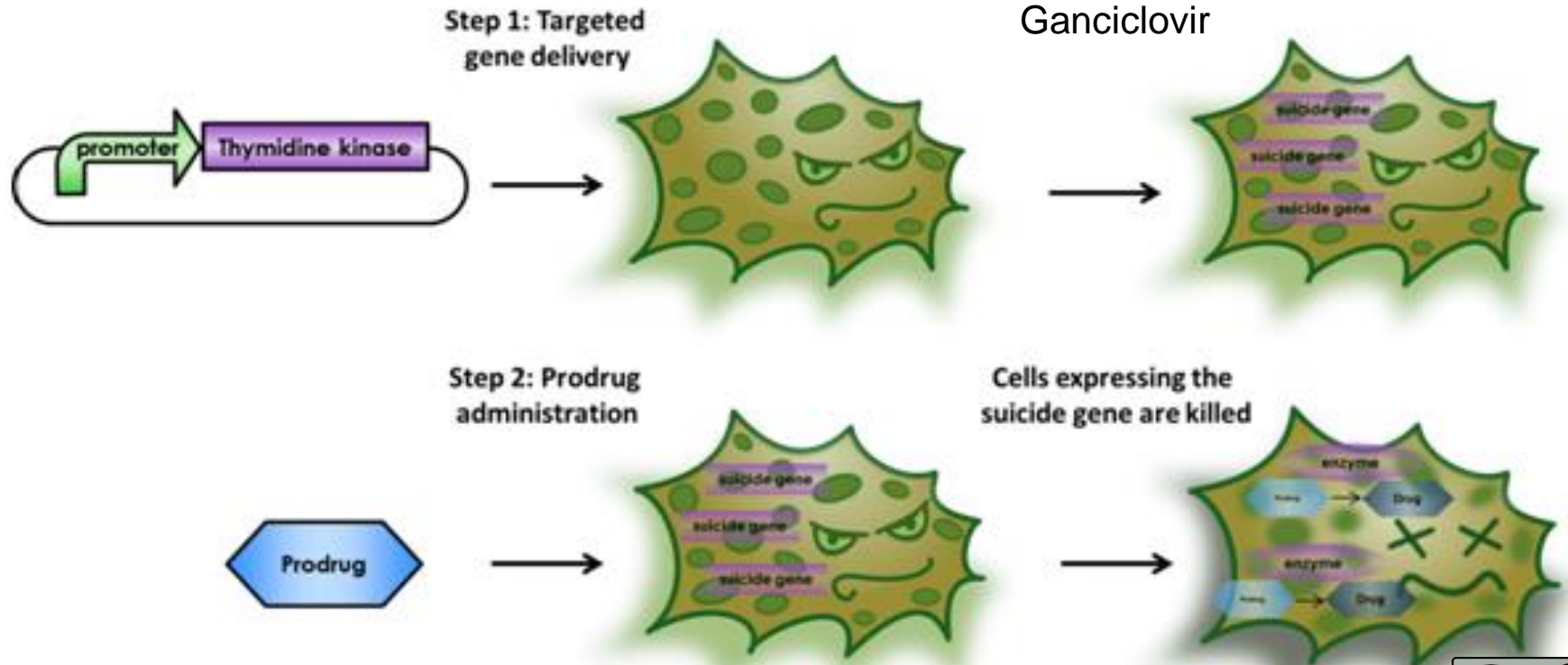
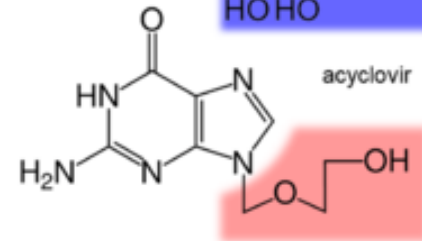
Gene for *E. coli* **cytosine deaminase** (CD) + 5-fluorocytosine  
Transformation of 5-fluorocytosine to 5-fluorouracil (more toxic)

Gene for **thymidine kinase** (TK) of herpes simplex virus (HSV-*tk*)  
+ analogs of nucleosides (acyclovir, ganciclovir)

TK phosphorylates ACV or GCV to monophosphate  
– inhibition of polymerases



Ganciclovir



## **Candidate genes**

- IL1, IL2, IL4, IL6, IL7, IL12 a JE/MCP1 (SCYA2)

## **Alternative approach**

- tumor cell antigens "presented" to immune system
- Attempts to modify a surface antigen by insertion of „co-stimulating" molecules

# DNA vaccines

Transfection with plasmid containing gene for tumor antigen

## Benefits

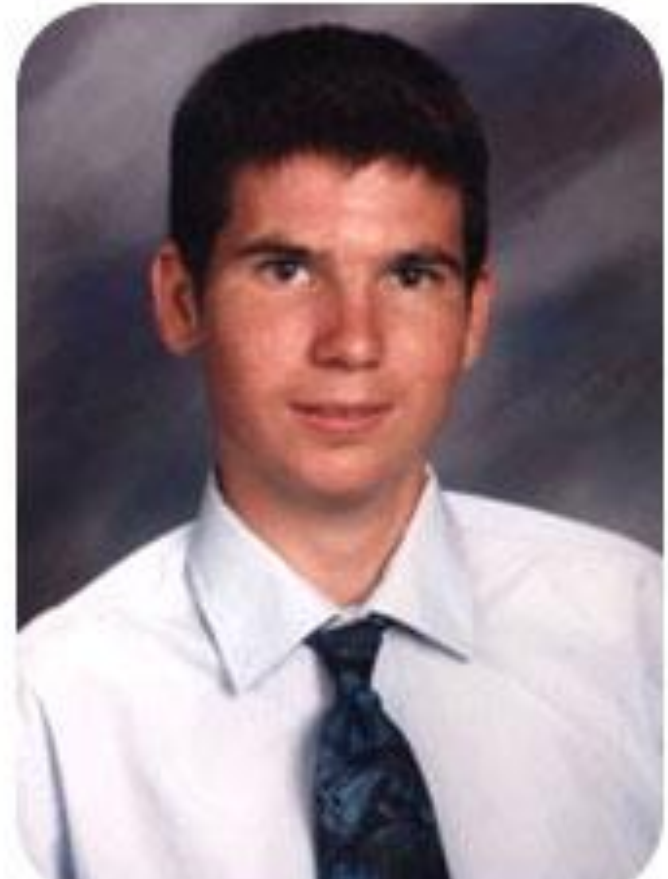
- strong immune response, stable, cheap, natural conformation

# Limitations of gene therapies

Several failures

Jesse Gelsinger

Rare liver disease (1/40 000)



**Ornithintranscarbamylase deficiency (OTC) – ammonia metabolism regulation**

X-linked (inherited from mother)

Diet – low consumption of proteins + 32 pills / day (still spasms)

Gelsinger - volunteer in 1999 gene therapy experiment (University of Pennsylvania) - adenovirus expressing OTC

Died after 4 days from multi-organ failure - excessive immune response to adenovirus

Discontinuation of gene therapies in US - discussion, regulation

(3 patients with SCID - France)

## **Result:**

**Vector – gene insertion also to other tissues than liver**  
**- systemic inflammatory response**  
**fever, coma, dead**

## **Why?**

**Animal experiments indicated that the dose is OK (?).**  
**Adenoviral vectors – sometimes inflammatory response**  
**Patient was already weakened by the disease**



## **Problems with gene therapies**

### **Risks**

**Problematic targeting of** circulating cells e.g. lymphocytes – easier targeting of tissues

**Short-term effect** – problems with stable integration. Patients must go through several cycles of treatment.

**Immune response** – insertion of foreign body, stimulation of immune system – risk of failure

**Problems with viral vectors** - toxicity, immune and inflammatory response. Possibility of mutation to virulent form.

Possibility of inactivation x induction of gene as a result of integration

Retroviral integration to dividing cells

**Multigenic diseases** – many pathologies (cardiovascular diseases, hypertension, Alzheimer disease, arthritis, diabetes – combined effect of many genes.

# **Ethical aspects of GT**

**What is norm and what is illness, who will judge it?**

**Is disability an illness?**

**Should it be cured or prevented?**

**Pathway to upgrade – intelligence, „genetic doping“**

**In case of somatic GT – will be repeated again in progeny  
(GT in germ cells - heritable)**

# **Etical aspects of GT**

**Is somatic gene therapy (of informed adult) more or less ethic than therapy of embryo (or even egg / sperm as a prevention of transmission from parent to progeny)?**

**GTs are extremely expensive. Who will be selected for?**

**Who will pay it?**

**Somatic therapy targets only the cells of the patient  
– no transmission to descendants**

Uveřejněné materiály jsou určeny studentům Vysoké školy chemicko-technologické v Praze

jako studijní materiál. Některá textová i obrazová data v nich obsažená jsou převzata

z veřejných zdrojů. V případě nedostatečných citací nebylo cílem autora/ů záměrně poškodit event. autora/y původního díla.

S eventuálními výhradami se prosím obračejte na autora/y konkrétního výukového materiálu, aby

bylo možné zjednat nápravu.

The published materials are intended for students of the University of Chemistry and Technology, Prague as a study material. Some text and image data contained therein are taken from public sources. In the case of insufficient quotations, the author's intention was not to intentionally infringe the possible author(s) rights to the original work.

If you have any reservations, please contact the author(s) of the specific teaching material in order to remedy the situation.

# Introduction to genetic engineering



EUROPEAN UNION  
European Structural and Investing Funds  
Operational Programme Research,  
Development and Education



MINISTRY OF EDUCATION,  
YOUTH AND SPORTS

# Homepage of the Department of Biochemistry and Microbiology

Education -

.....are delivered in English....

→ Genetic engineering (**www**)

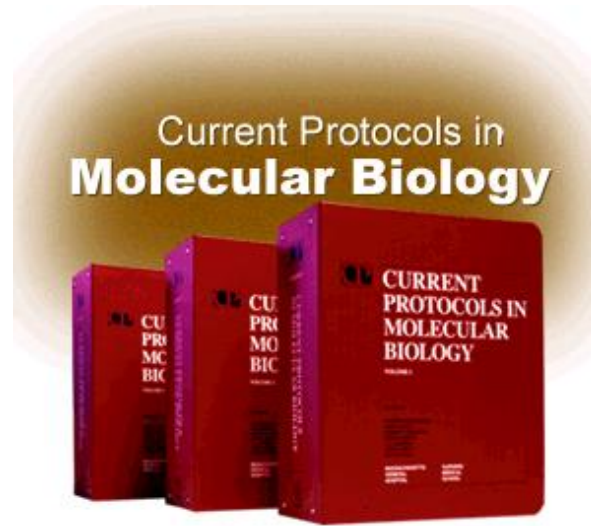
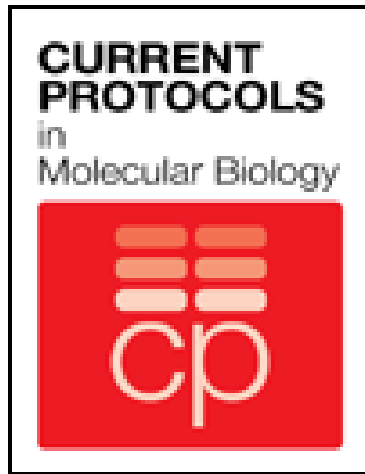
**Password: giruml**

Print \$1 550  
Online \$ 850

## Current Protocols in Molecular Biology

F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl

John Wiley & Sons



Citation example:

Medberry, S., Gallagher, S., and Moomaw, B. 2004.

Overview of digital electrophoresis analysis. *Curr. Protoc. Mol. Biol.* 66:10.5.1-10.5.11.

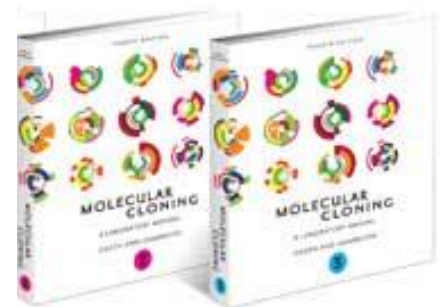
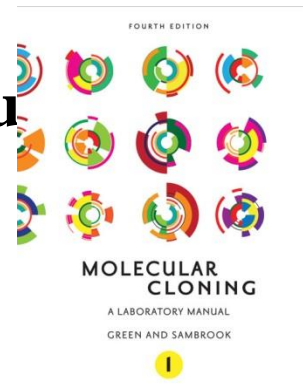
## Molecular Cloning: A Laboratory Manual (Fourth Edition)

J. Sambrook, M. R. Green

Cold Spring Harbor Laboratory Press 2012

(J. Sambrook, T. Maniatis)

\$ 365



Web





Production, isolation and characterization of DNA (plasmids and other vectors).

Transformation and transfection, cell strains.

Restriction endonucleases, ligases and other enzymes for cloning.

Identification of recombinant DNA

- DNA labeling, probes, hybridization, Southern blot.

DNA analysis - genome mapping, restriction analysis, sequencing.

Gene banks, mRNA isolation, PCR.

Mutations, controlled mutagenesis, phenotype analysis.

Gene expression - regulation and control: RNA and protein analysis, fusion proteins

Product isolation - affinity chromatography fusion

Gene therapy

# Methods of genetic engineering

Lab equipment - companies

Material

Safety (Chem. Biol., GMO)

Ethic questions

# **Cells**

## **Sources**

## **Genetically modified strains**

## **Storage, resuscitation, testing**

# *Escherichia coli*

## **DNA amplification**

DH1, DH5	recombinant deficient, for DNA amplification
DH5 $\alpha$	additional mutations: $\Delta 80$ lacZ $\Delta$ M15 for $\alpha$ -complementation
JM101	F' for M13 infection
JMN110	F', in addition methylation deficient

## **Protein expression**

BL21(DE3)	integrated RNA polymerase gene of bacteriophage T7 in $\lambda$ DE3, mutated protease genes
-----------	---

M13 is a filamentous E. coli bacteriophage specific for male (F factor-containing) cells

# Plasmids

**extrachromosomal elements**

**cca 1 000 - 200 000 base pairs**

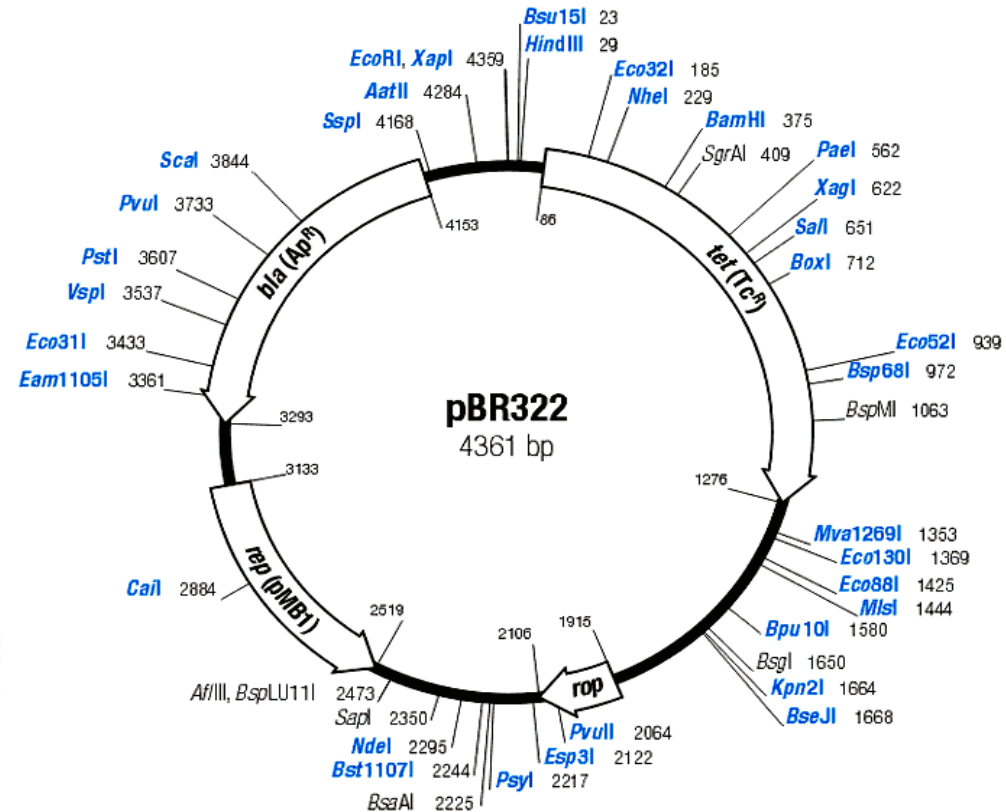
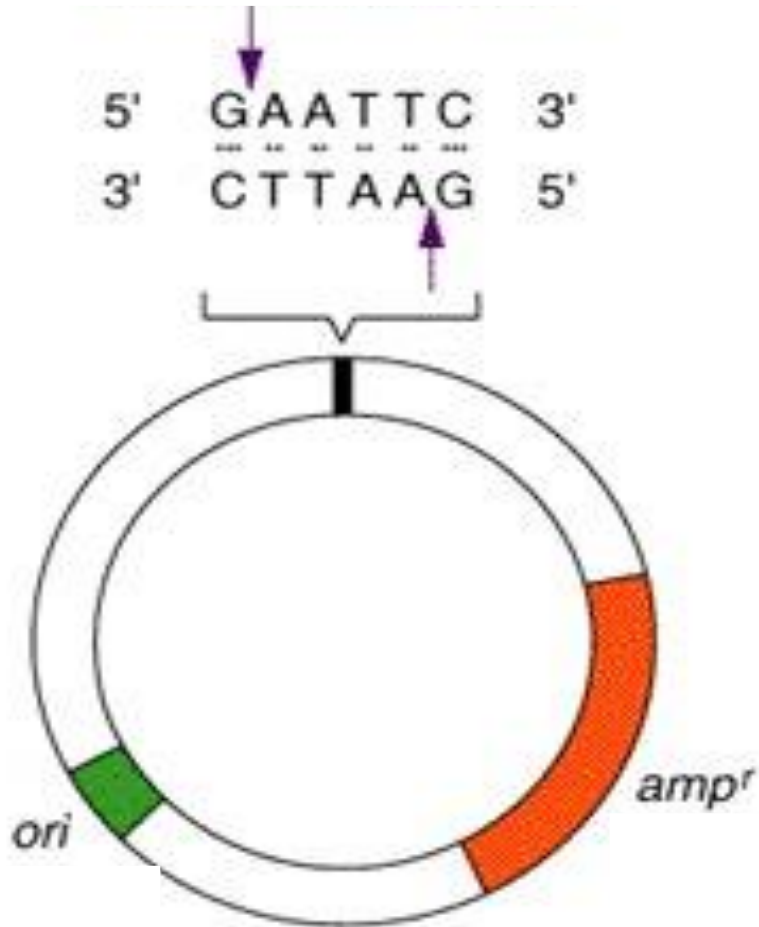
**ds, circular DNA**

**autonomous replication**

**synchronized x non-synchronized with chromosome replication**

**„relaxed“ plasmids - high copy number  $> 20/\text{cell}$**

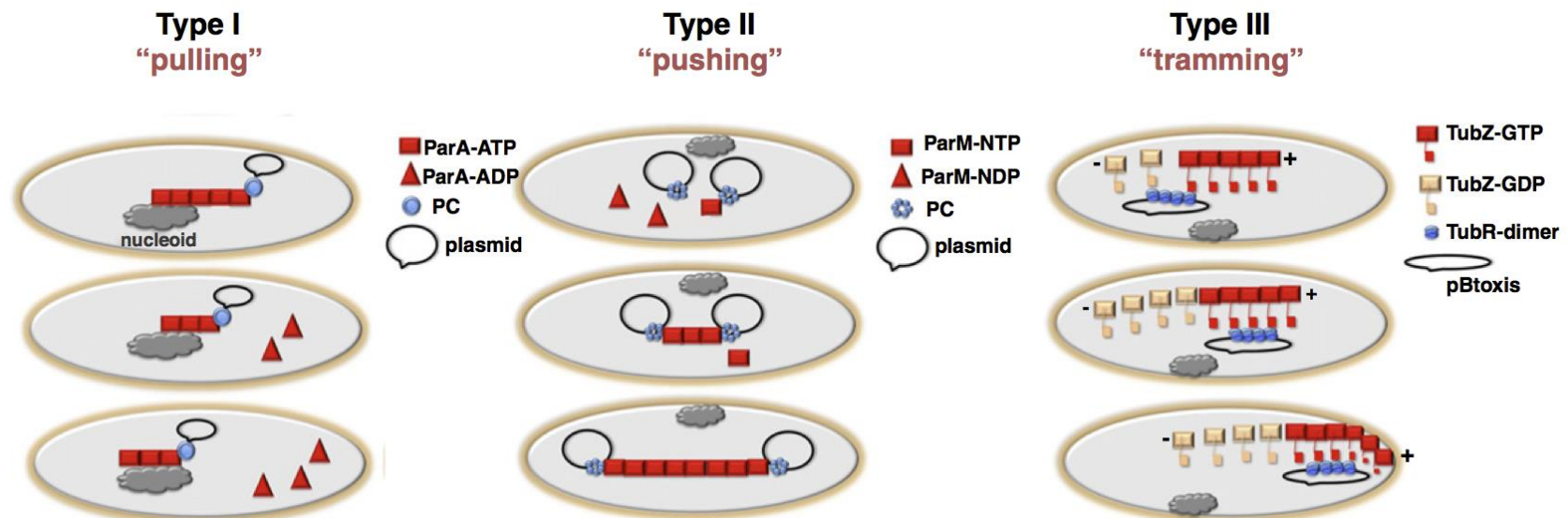
# Plasmid



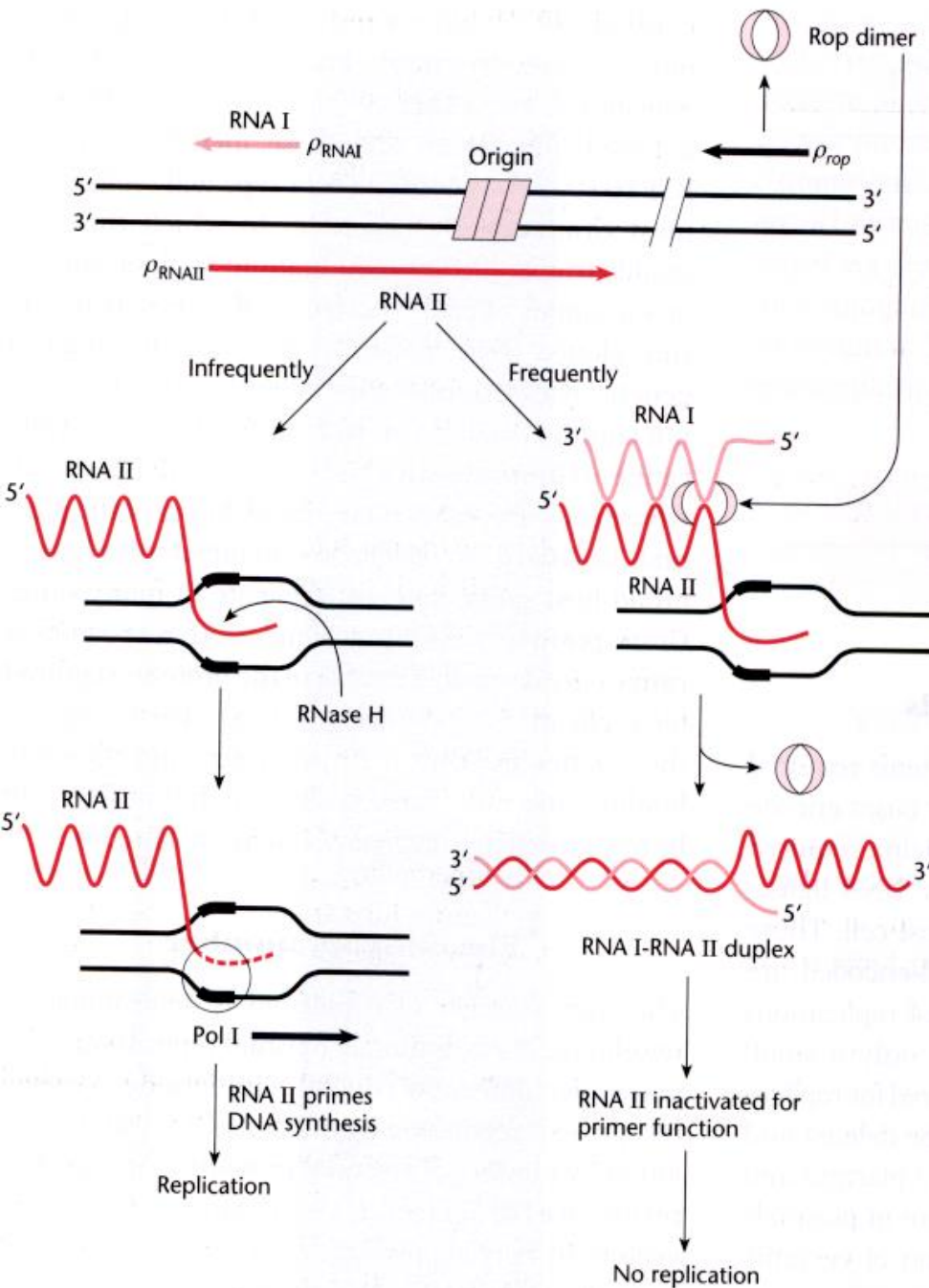
# Plasmid instability

Natural Plasmids - Partitioning Control - Partitions of the Area  
- Par protein - polymerization - segregation of plasmids - 3 types

Required for retaining low copy plasmids  
"Low copy number"







## CONTROL OF THE PLASMID COPY NUMBER

Rop protein (repressor of primer) - regulator

Binding of DNA and RNA II-RNase

→ 3' end - primer

RNA I binding - "kissing complex" →

inhibition of RNA II cleavage and inhibition of replication initiation

Rop dimer - increases affinity of RNA I to RNA

II - inhibition of cleavage

***The insert:***

**fewer copies of plasmids with large inserts or genes that create a toxic product.**

***The E. coli strain:***

***Growth conditions:***

**aeration, temperature, culture volume, antibiotic, and medium affect copy number. Some ORIs are temperature sensitive; some ORIs can be induced by the addition of chloramphenicol**

***The culture inoculum:***

**fresh**

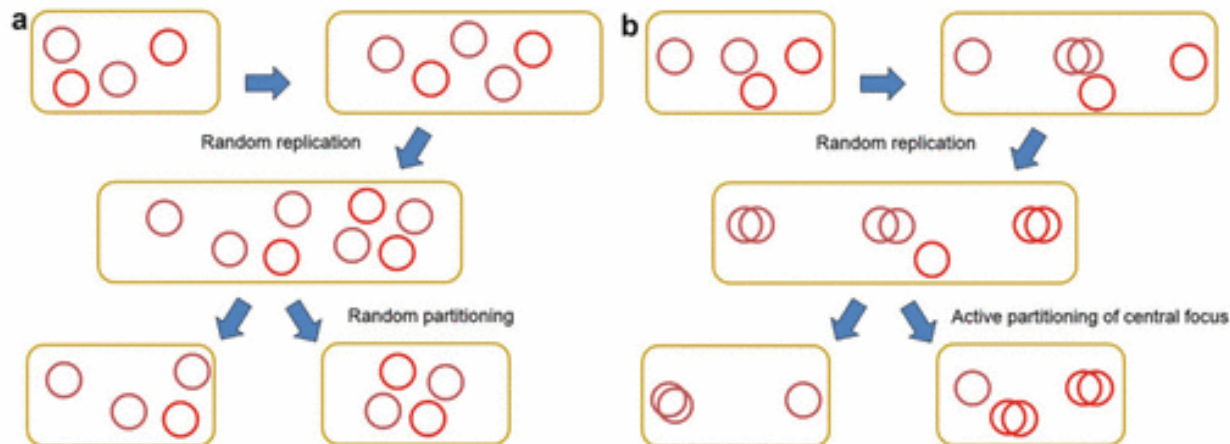
# Plasmids incompatibility

Related plasmids mutually "incompatible",

- The problem of coexistence in the same cell

They share the control mechanisms of replication - are competing for the number of copies  
total number of copies is unchanged

Fluctuations in plasmids copy numbers – random distribution



# Plasmids incompatibility

## Replicons of common cloning vectors

Common Vectors	Copy Number <sup>+</sup>	ORI	Incompatibility Group	Control
pUC	~500-700	pMB1 (derivative)	A	Relaxed
pBR322	~15-20	pMB1	A	Relaxed
pET	~15-20	pBR322	A	Relaxed
pGEX	~15-20	pBR322	A	Relaxed
pColE1	~15-20	ColE1	A	Relaxed
pR6K	~15-20	R6K*	C	Stringent
pACYC	~10	p15A	B	Relaxed
pSC101	~5	pSC101	C	Stringent
pBluescript	~300-500	ColE1 (derivative) and F1**	A	Relaxed
pGEM	~300-500	pUC and F1**	A	Relaxed

# **Introduction of plasmid into the cells**

**competent cells  $\text{CaCl}_2$  – temperature shock  
electroporation**

**Eukaryotic cells – shuttle vectors**

**Yeasts**

**LiCl  
electroporation**

**Tissue cultures**

**electroporation**

**$\text{Ca}_3(\text{PO}_4)_2$**

**DEAE dextrane**

**liposomes**

**gene gun**

# Selectable markers

Antibiotics	Concentration		Inhibition
	Stock solution (mg/ml)	Working solution ( $\mu$ g/ml)	
Ampicillin	100	50-100	blocks cell wall synthesis
Chloramphenicol*	34	25-150	50S ribosomal subunit
Kanamycin	40	30-50	inhibits ribosome translocation
Streptomycin	50	20-50	binds 30S subunit
Tetracycline**	10	10-50	inhibits tRNA / ribosome binding
Nalidixic acid	5	15	DNA synthesis – inhib. gyrase
Gentamycin	10	15	binds 50S subunit



- \* Stock solution in 100% ethanol
- \*\* Stock solution in 50% ethanol in dark

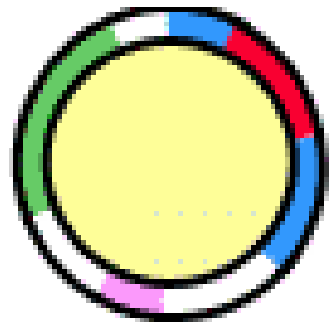
$\beta$ -lactamase – penicillin, ampicillin, cephalosporin

Chloramphenicolacetyltransferase

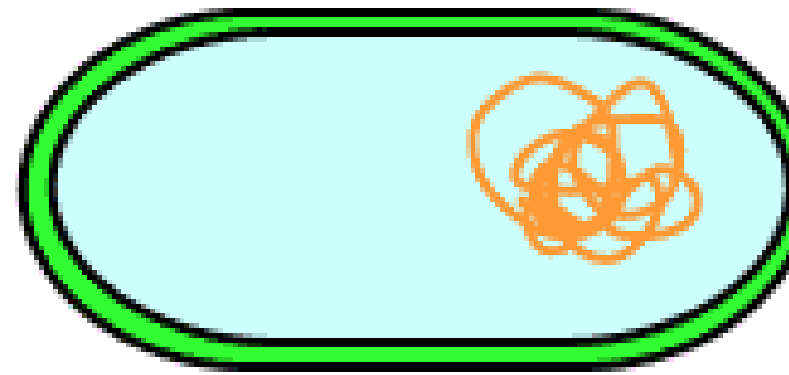
Neomycinphosphotransferase (also kanamycin)

Active export of  $Mg^{2+}$  tetracycline

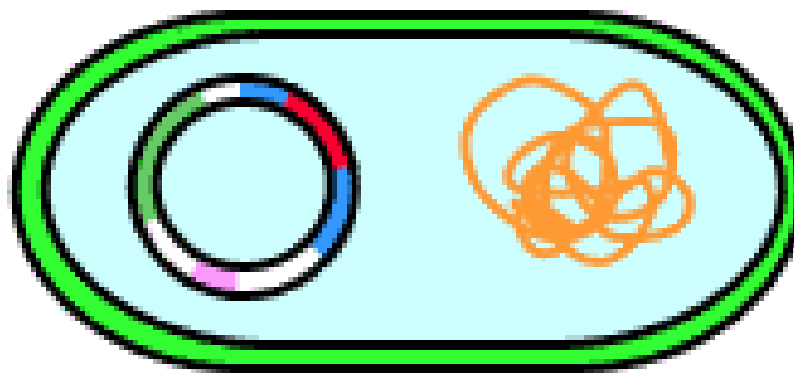




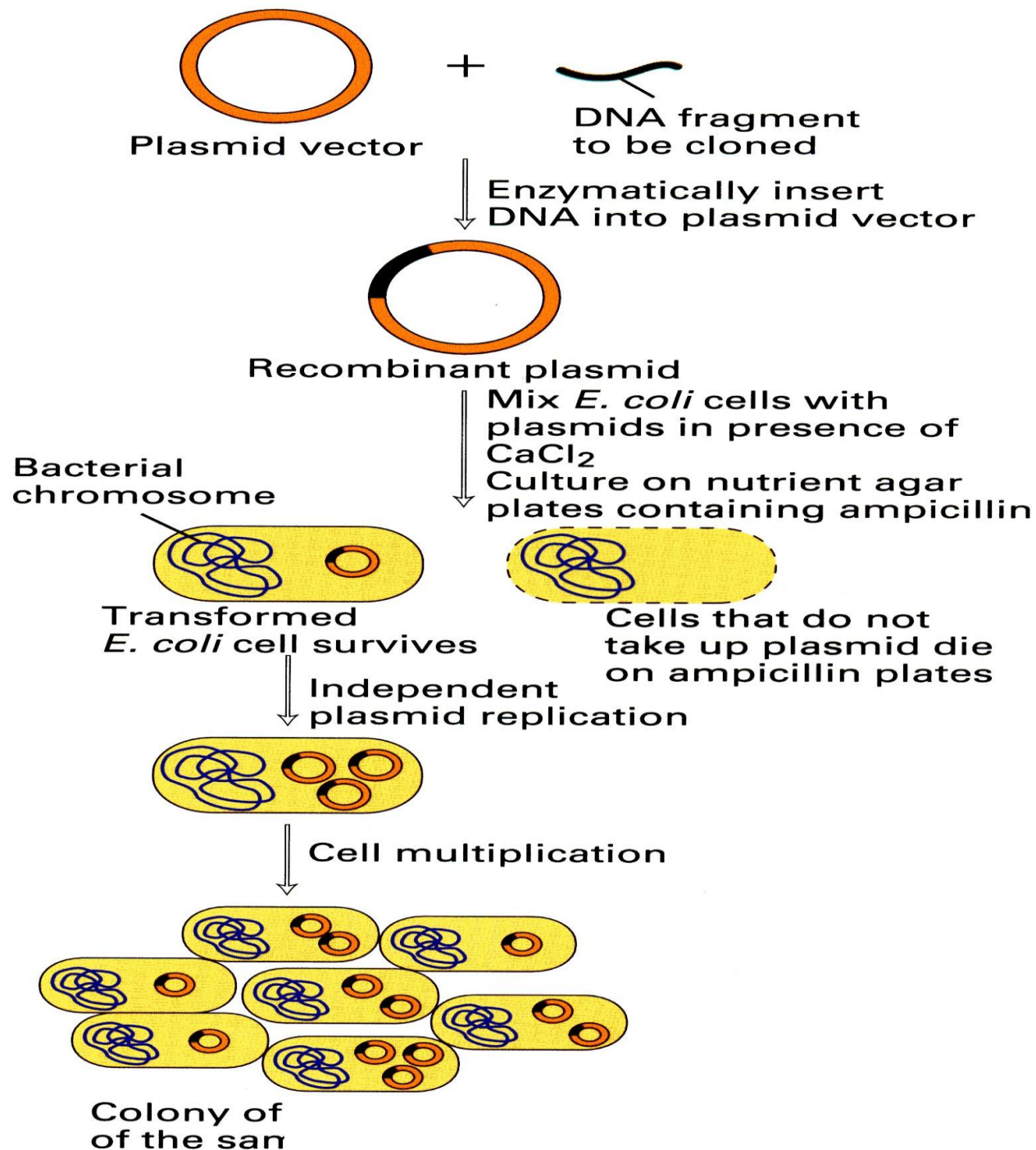
**Recombinant  
Plasmid**



*E. coli*



**Transformed  
Cell**



Some cells survive due to degradation of antibiotics around resistant colony - "satellite colony"

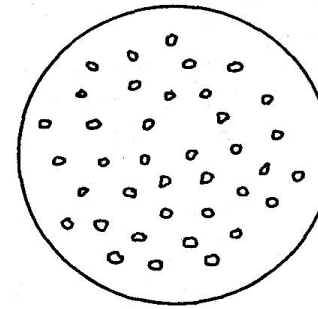
E.g. Due to  $\beta$ -lactamase production



# Identification of colonies with **recombinant** plasmid

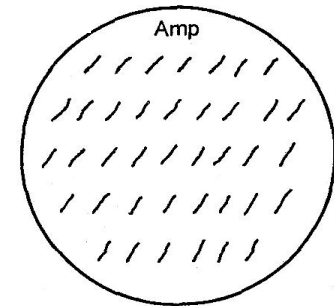
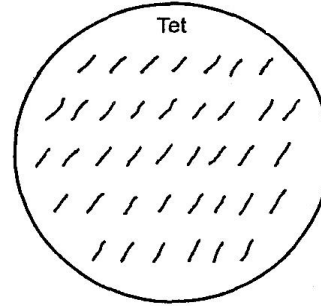
## Insertion inactivation

Amp



Transformed cells

Transfer by toothpick



Overnight incubation

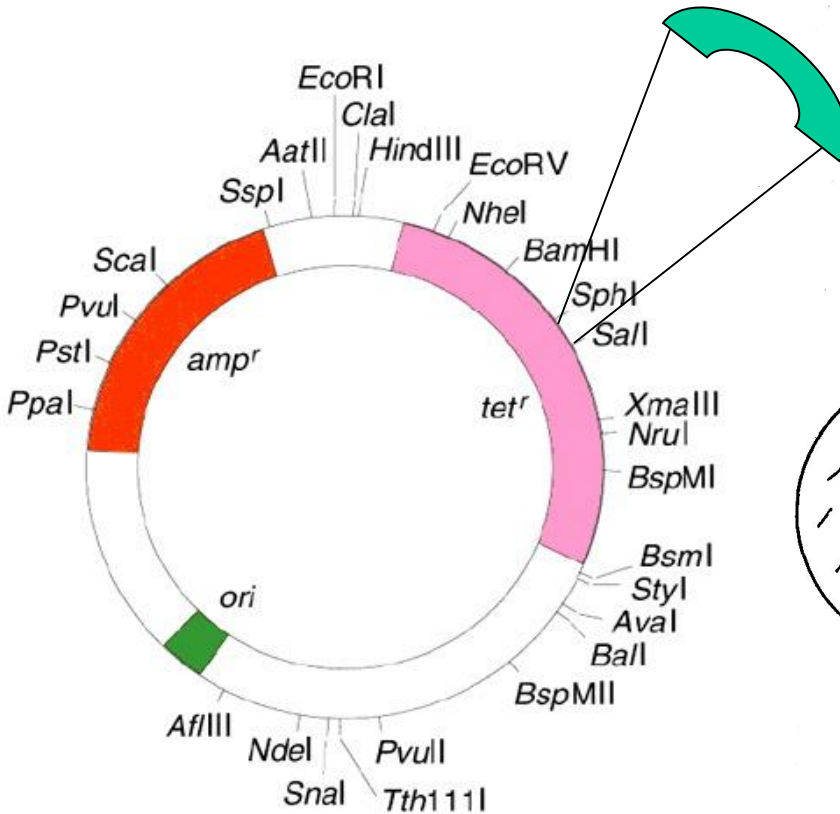
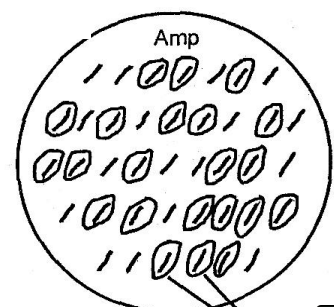
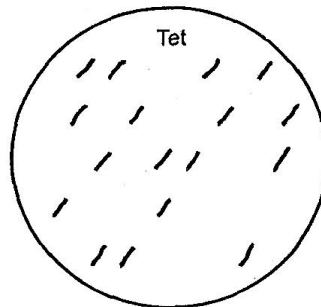
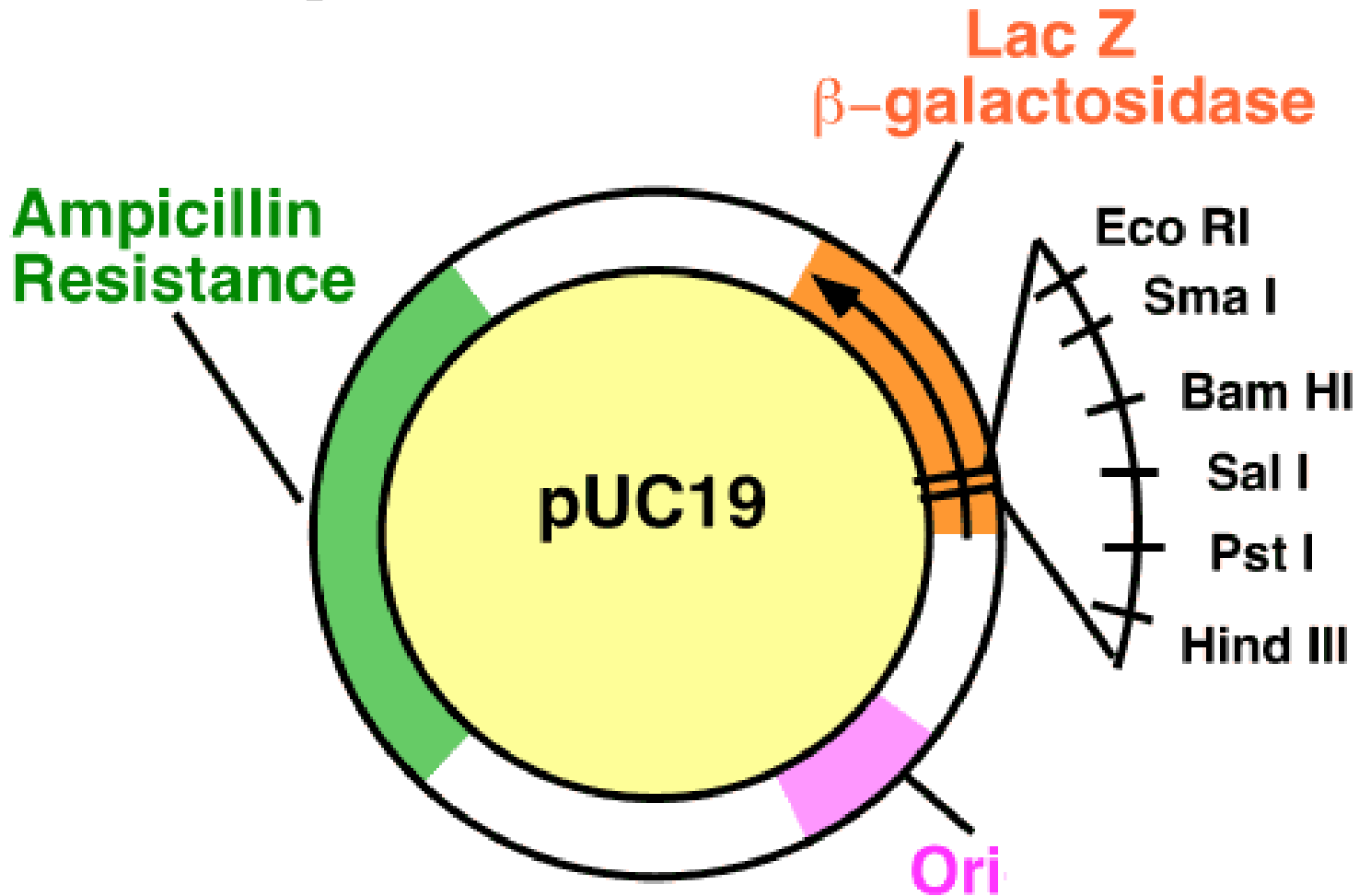
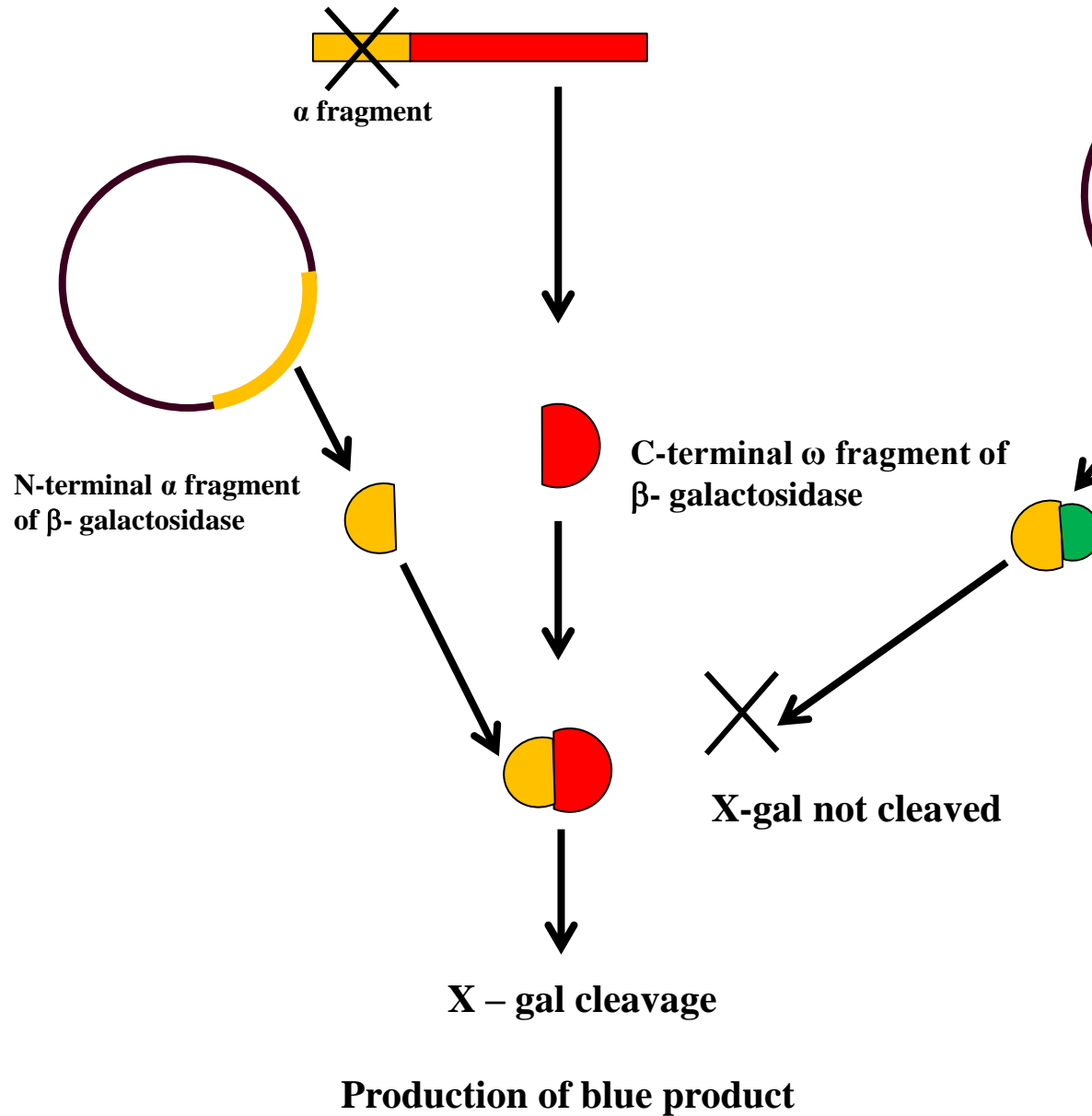


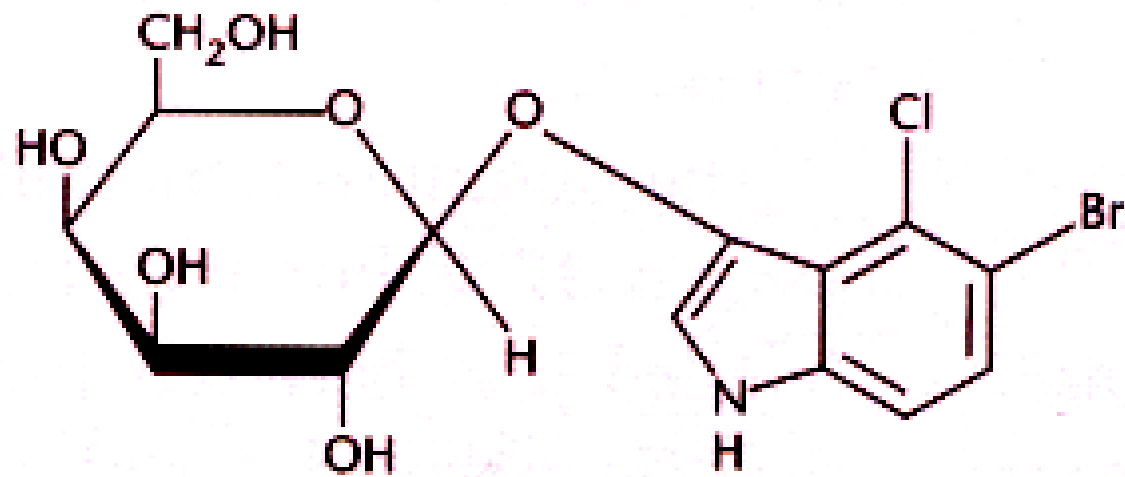
Figure 20.4 Structure of *E. coli* plasmid cloning vector pBR322, a circular DNA molecule

# Complementation

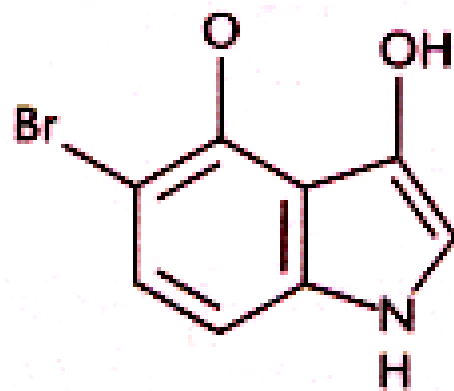
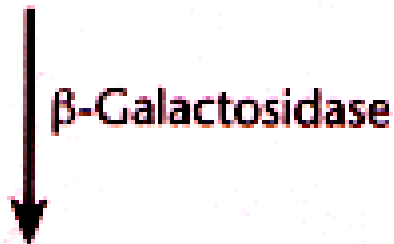


# Chromosome – $\beta$ -galactosidase mutation

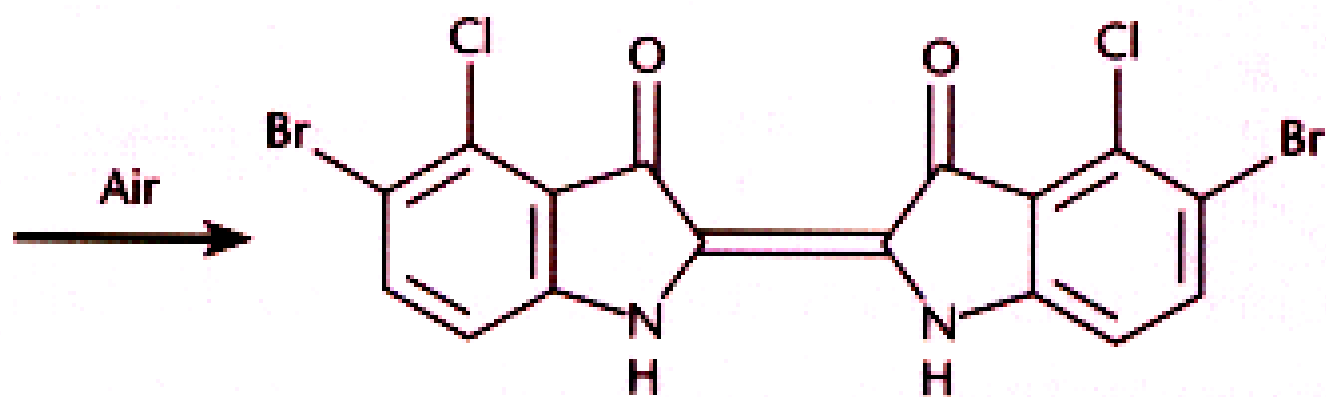




5-Bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside (Xgal)

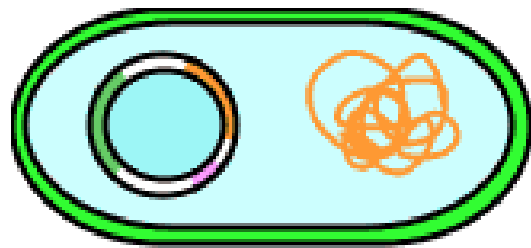


5-Bromo-4-chloroindoxyl



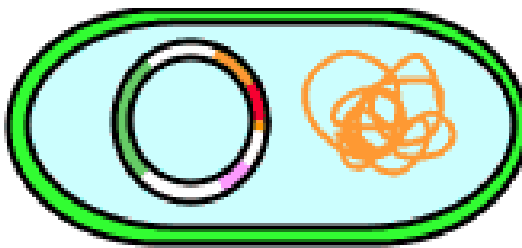
5,5'-Dibromo-4,4'-dichloroindigo





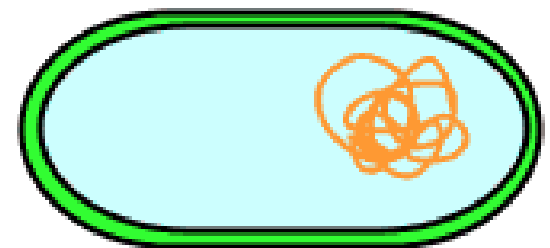
**Cell + pUC19**

+

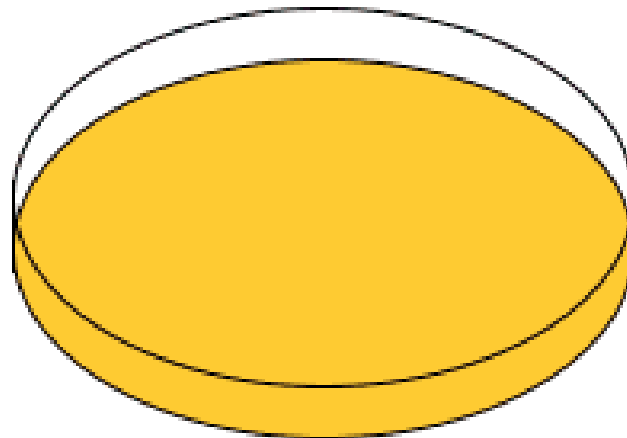


**Cell+pUC19  
+  
Insert**

+



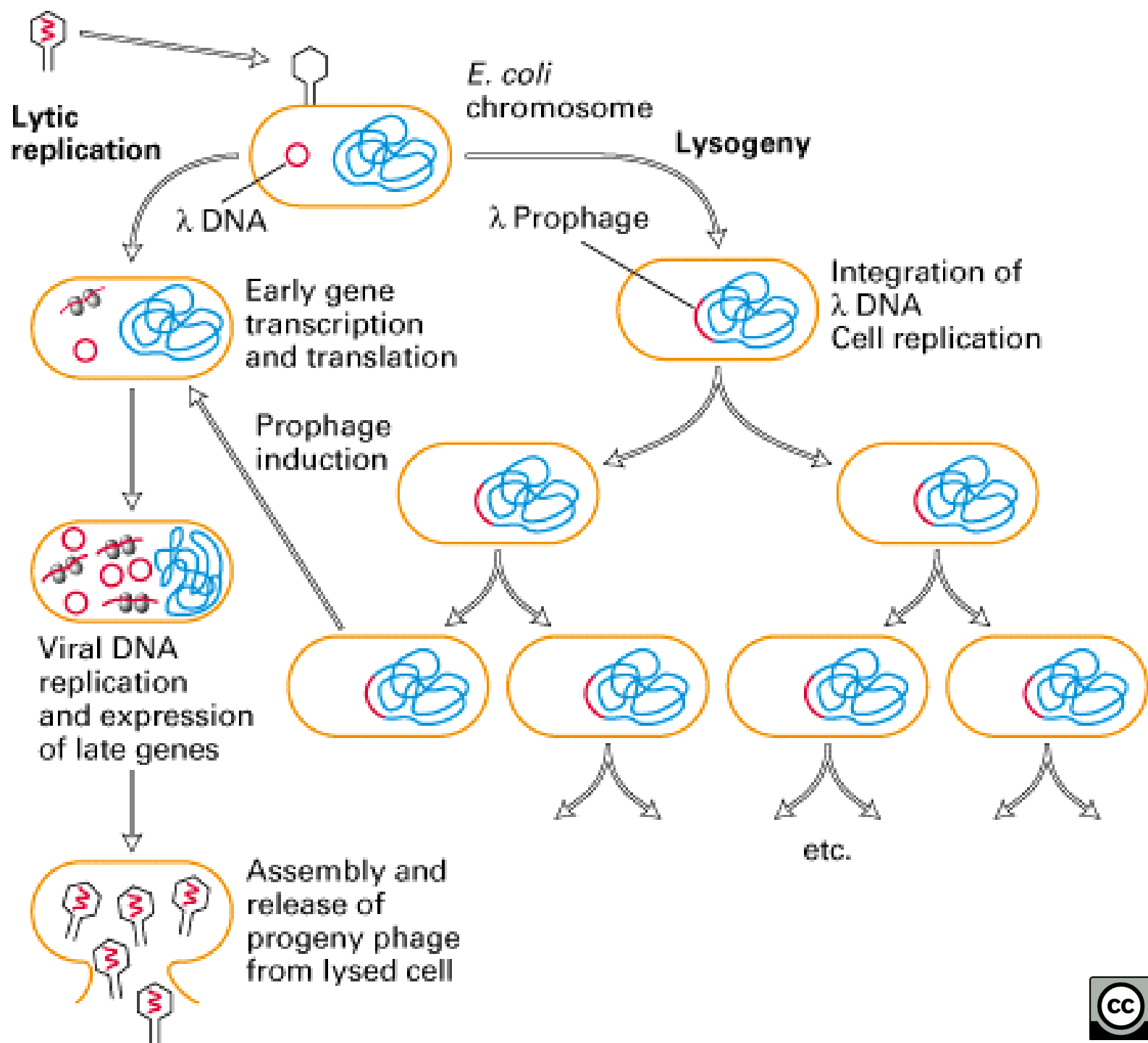
**Non-transformed  
Cell**

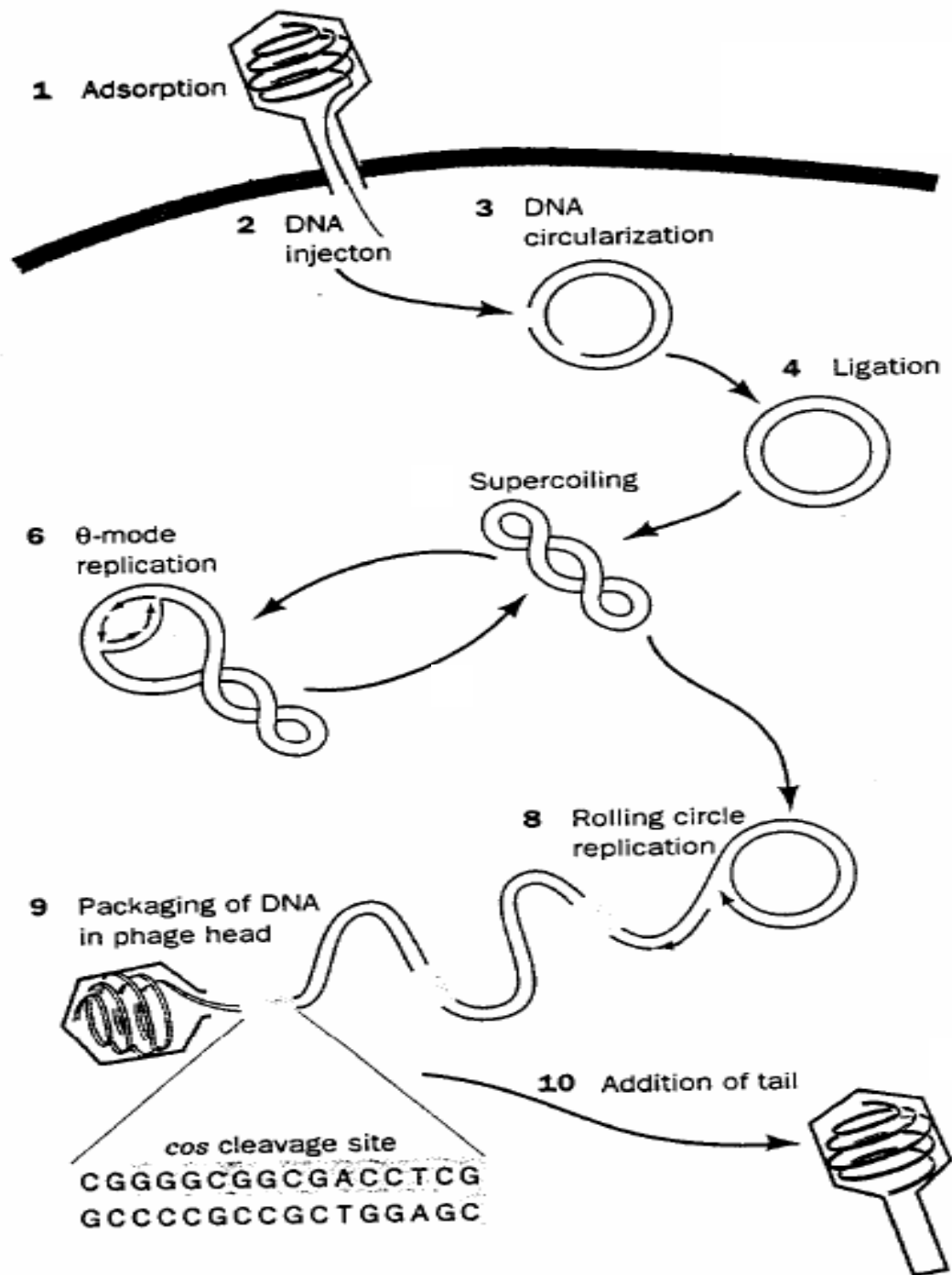


**Nutrient medium  
+  
ampicillin  
+  
X-gal**

# Bacteriophage lambda

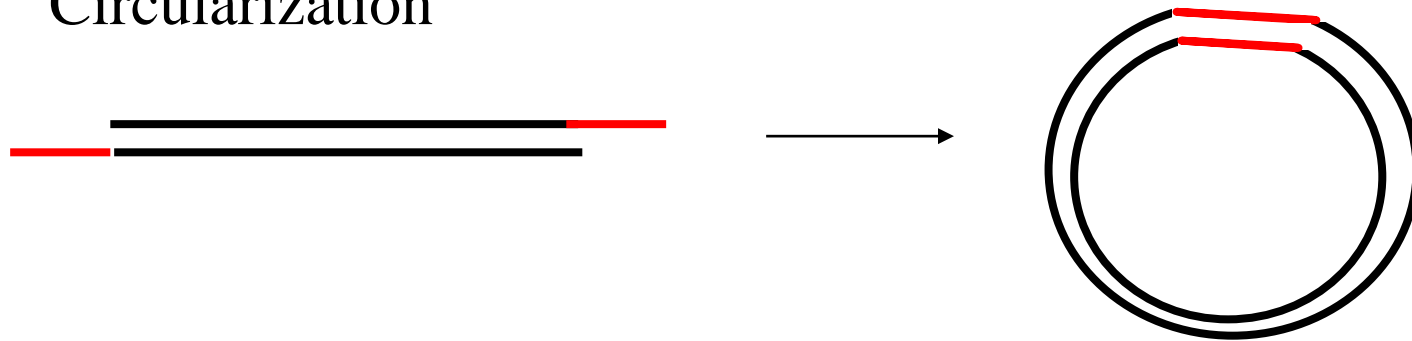




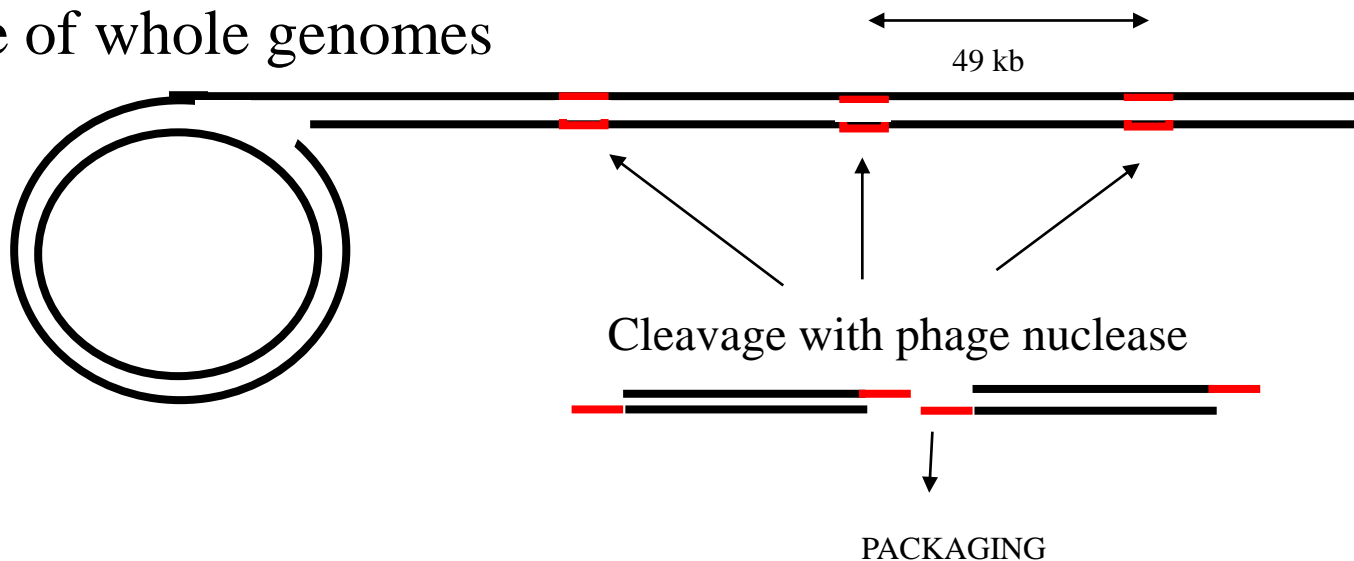


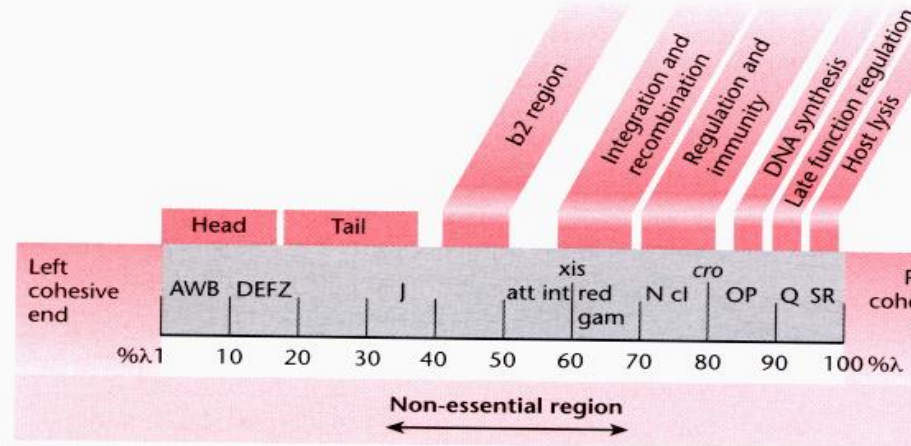
# Cos sequences

Circularization

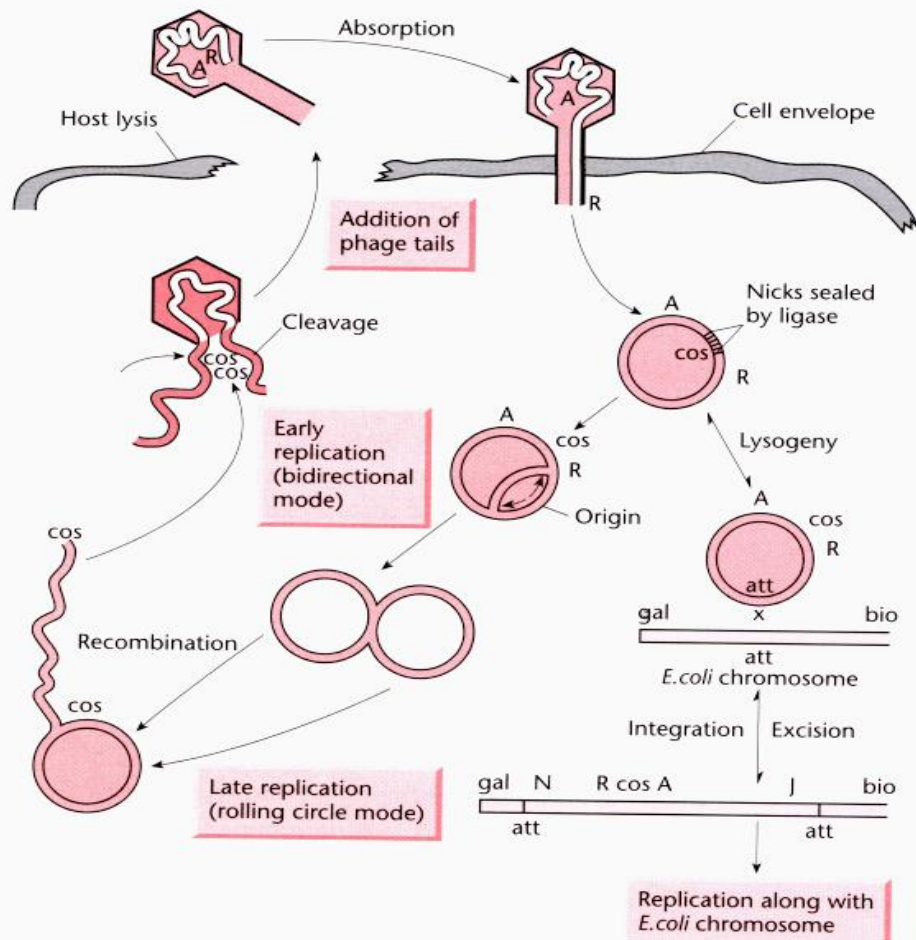


Cleavage of whole genomes

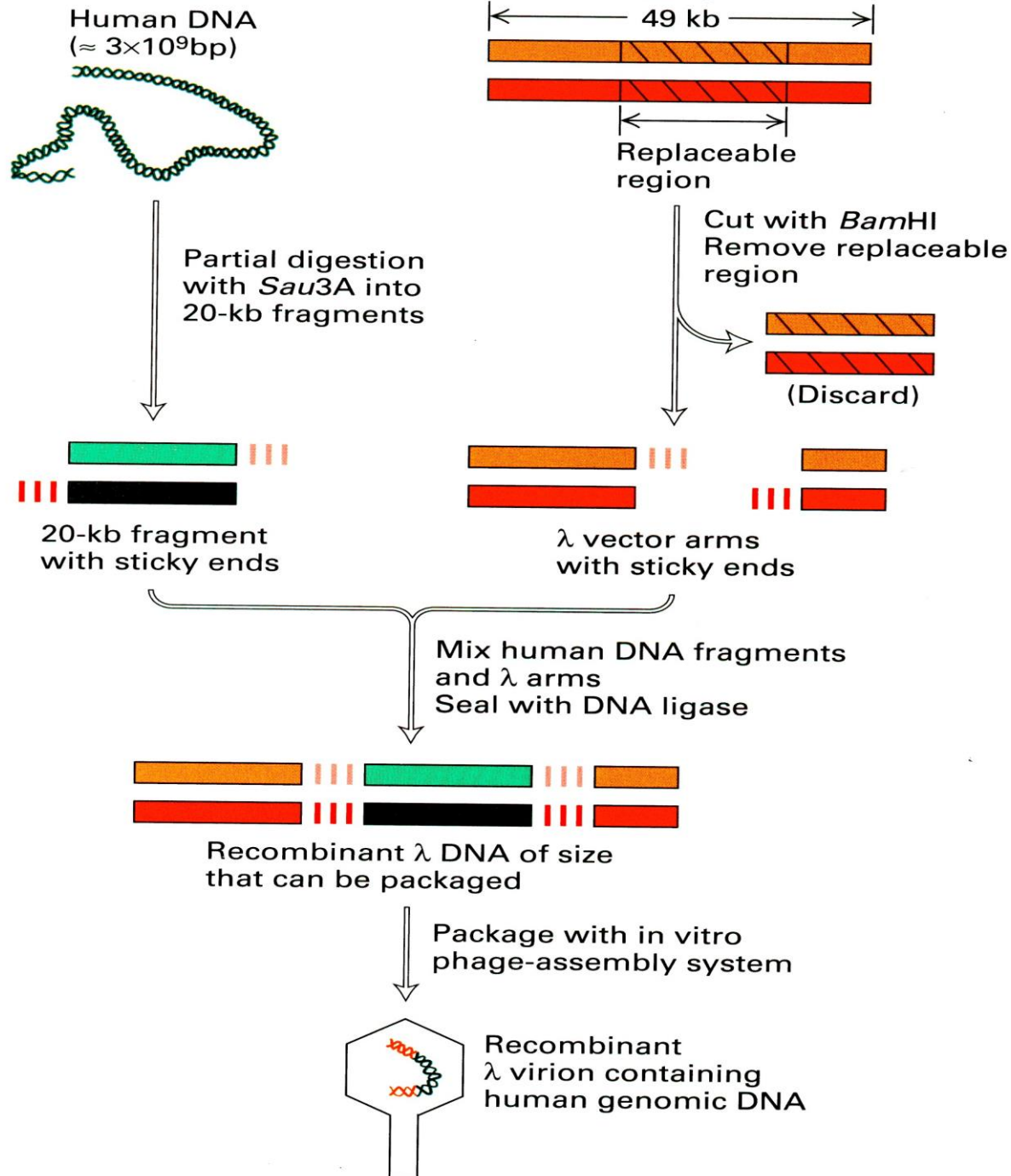




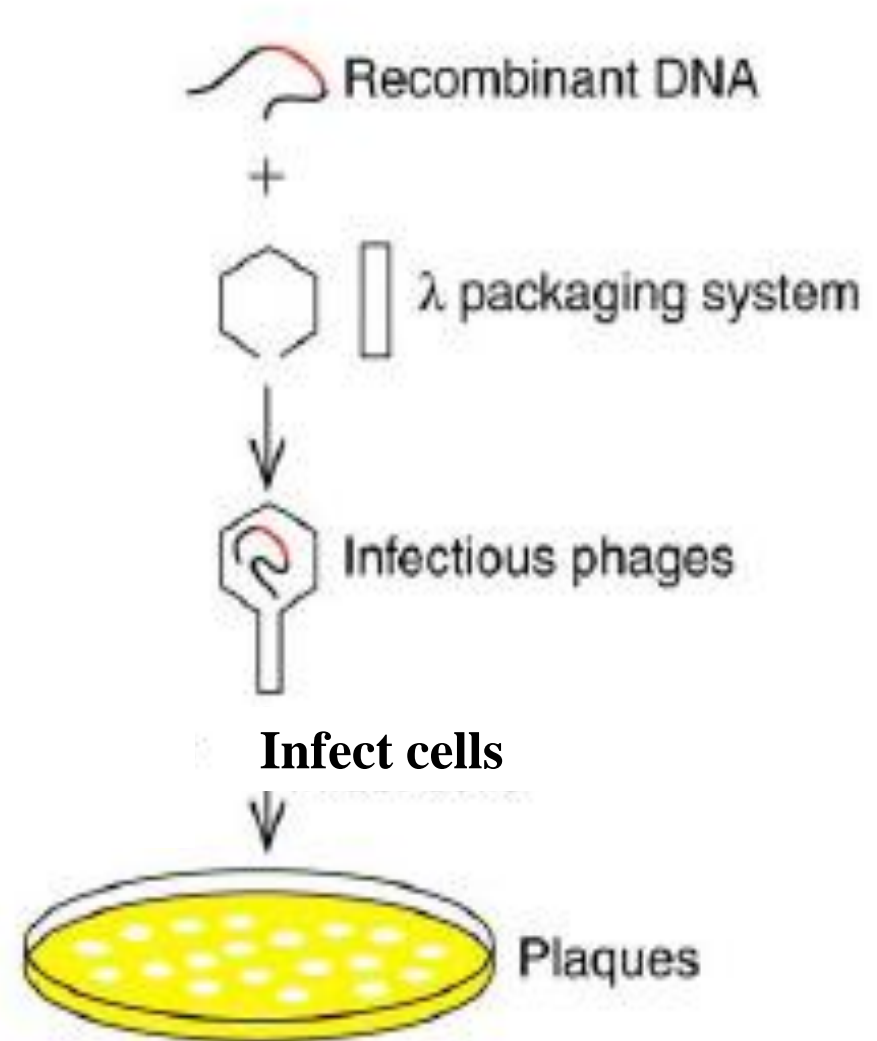
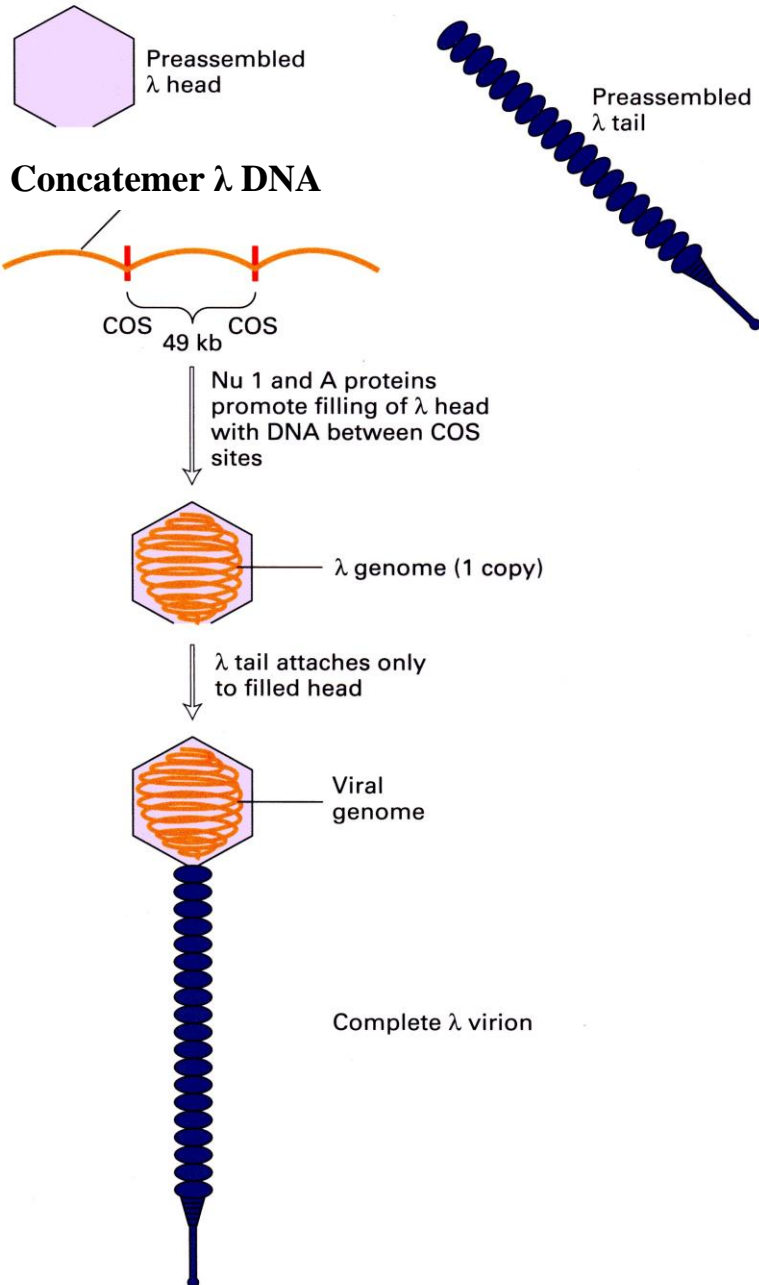
Lytic cycle

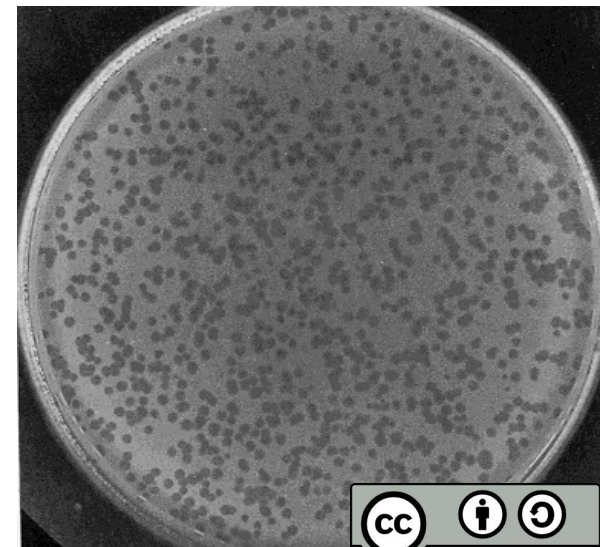
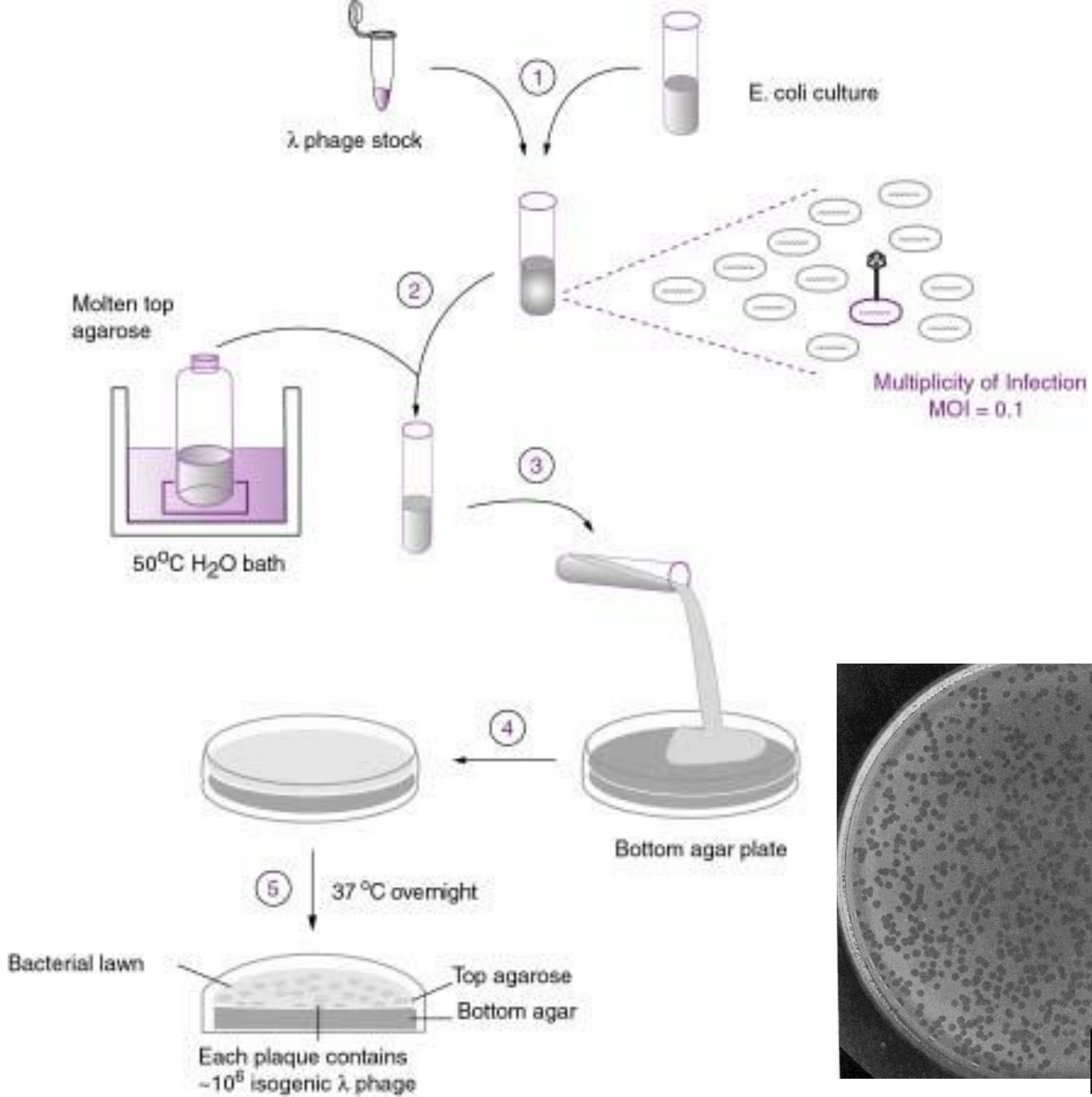


Lysogenic cycle









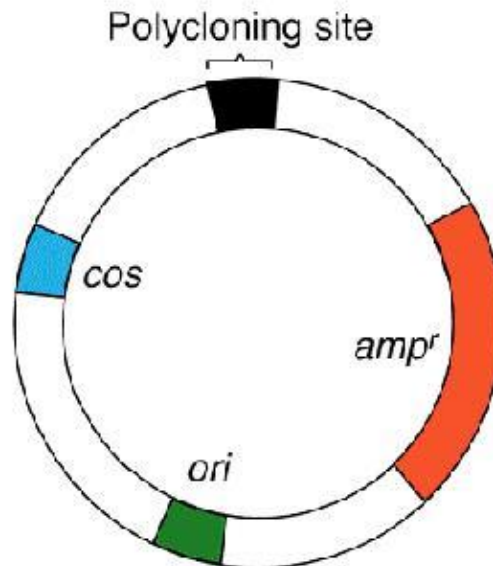
# Cosmids

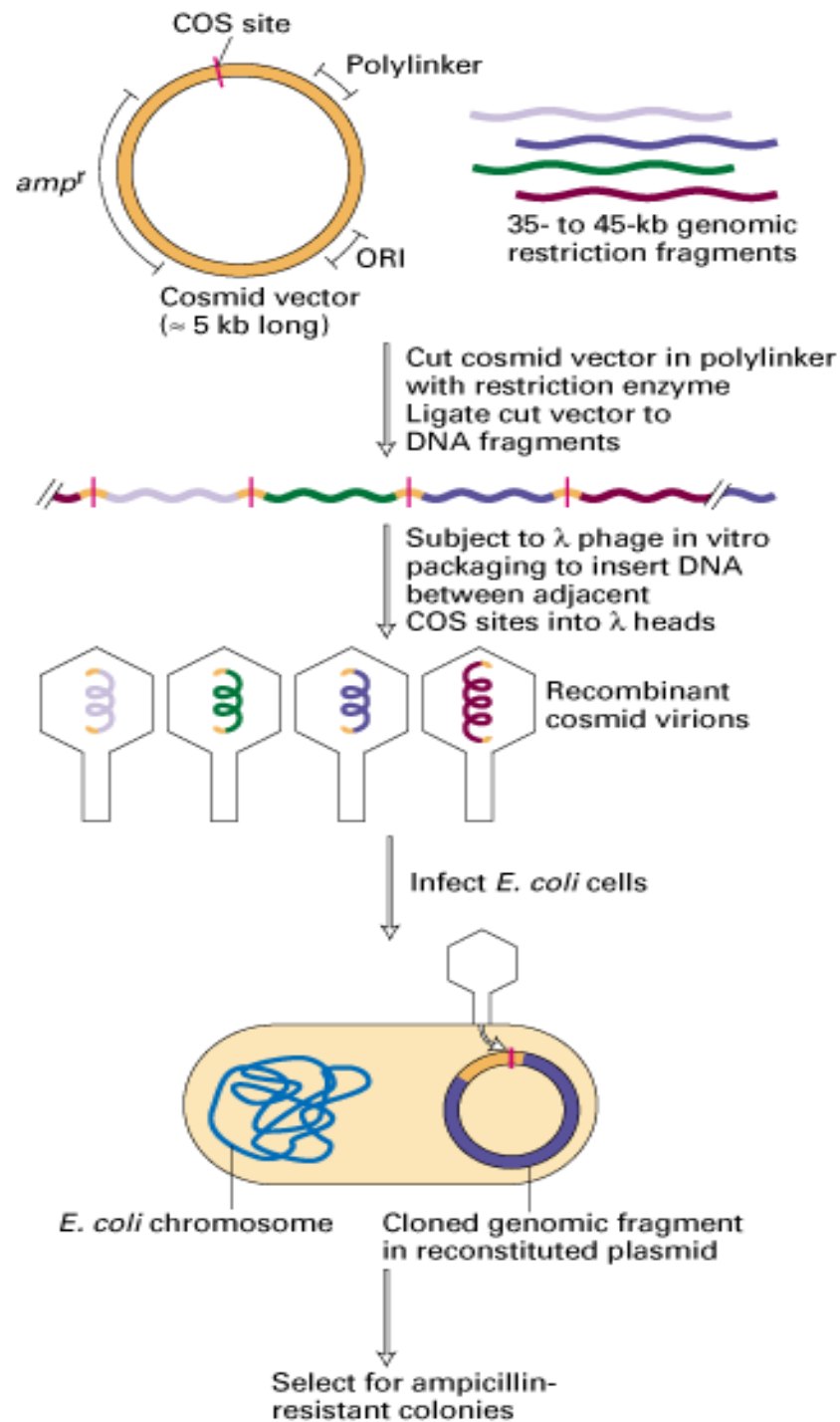
Similar to plasmids,

*cos* sequences from  $\lambda$  phage – efficient infection

- large DNA fragments (~ 40 kbp)

- Linearized vector with insert: co-ligated into concatamers and incorporated into phage heads (packaging extract)



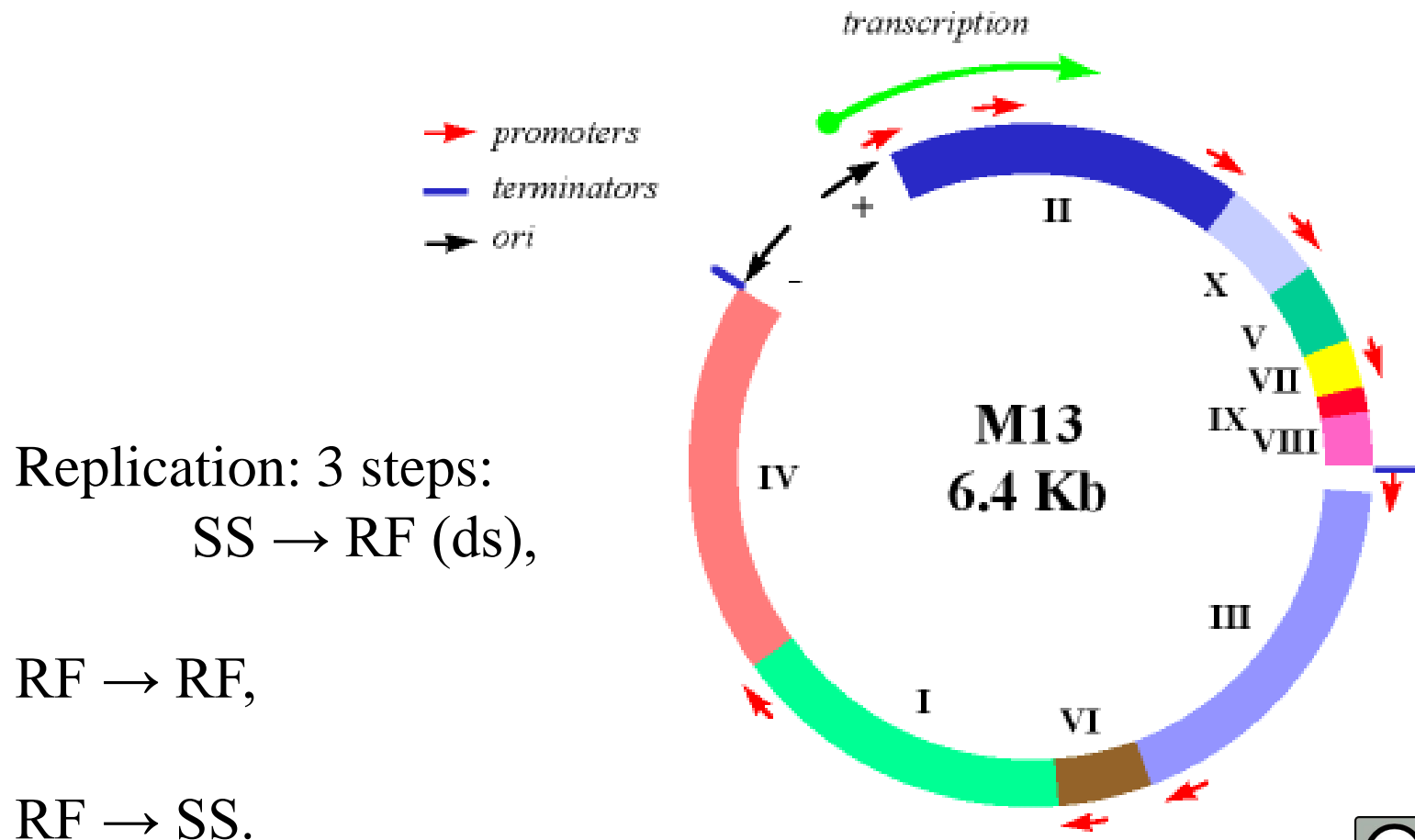


# Filamentous bacteriophages

M13, f1, fd - closed circular ss DNA

6400 bp

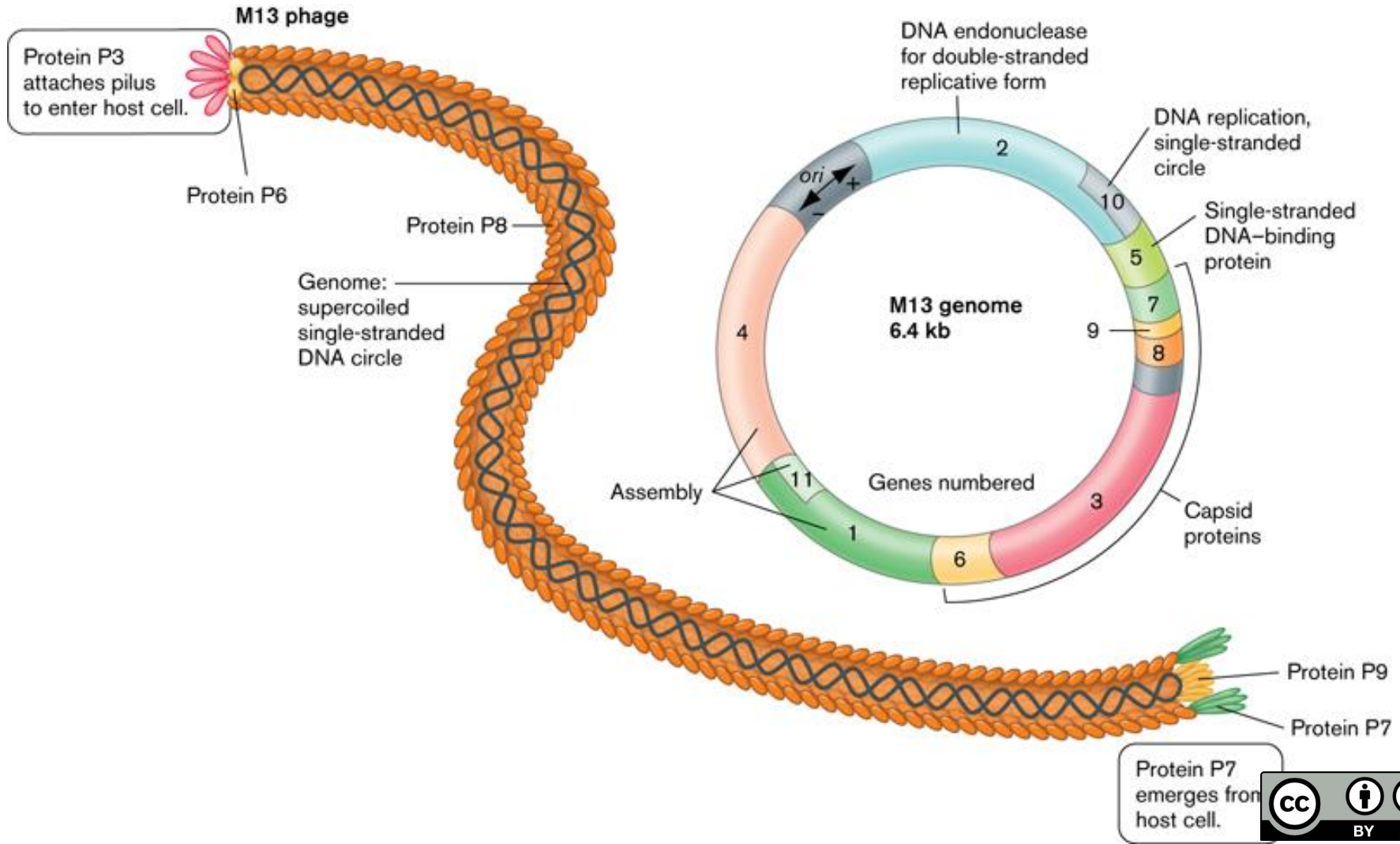
Production of ss templates for sequencing or *in vitro* mutagenesis



# Filamentous bacteriophages

M13, f1, fd - closed circular ss DNA 6400 bp

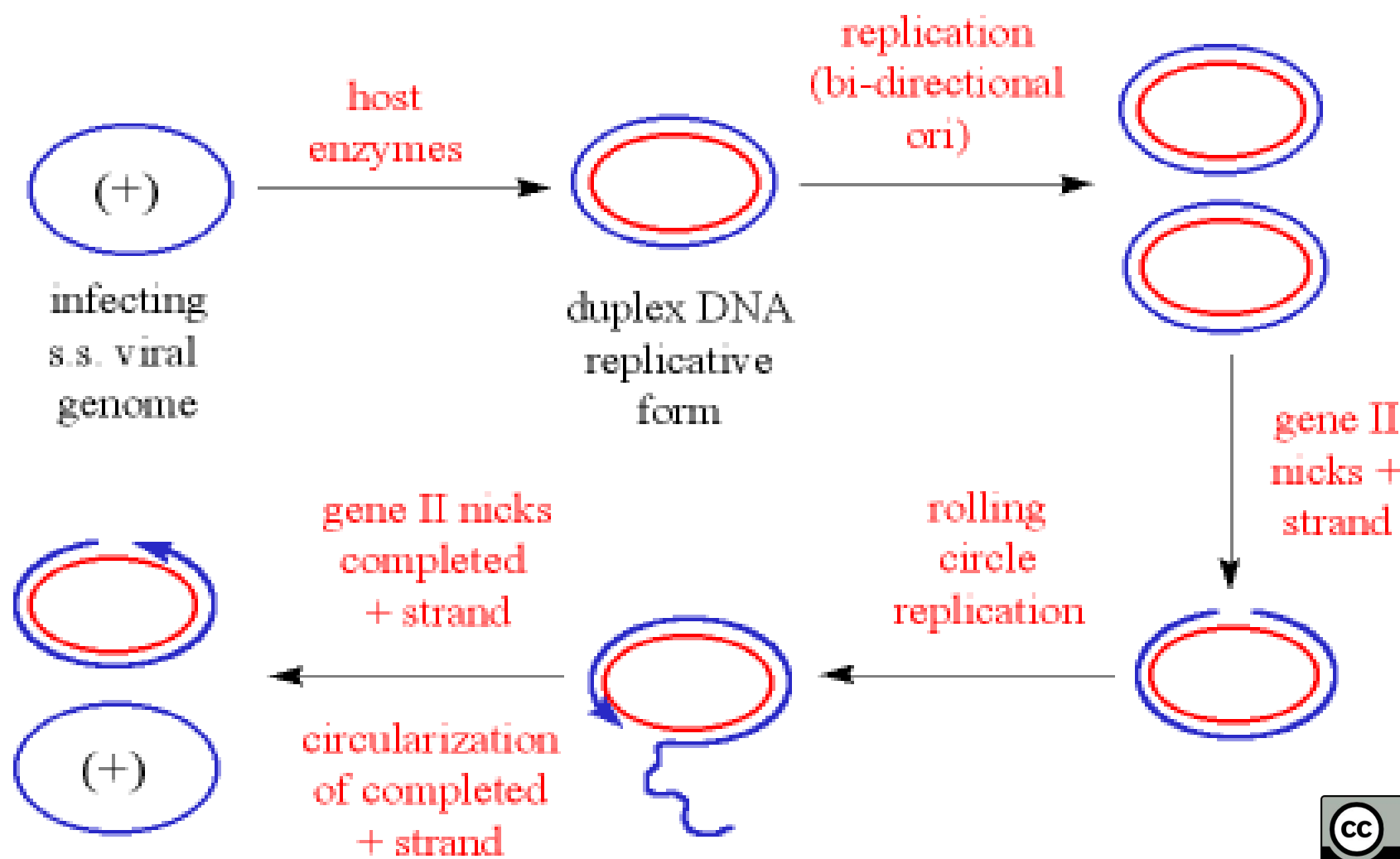
Production of ss templates for sequencing or *in vitro* mutagenesis



First 15-20 min. replication (+) chains circularized into ds (RF) form

Expression of gene V product - ss DNA binding protein

Prevents formation of RF – instead formation of circular ss (+) DNA (M13 genome)



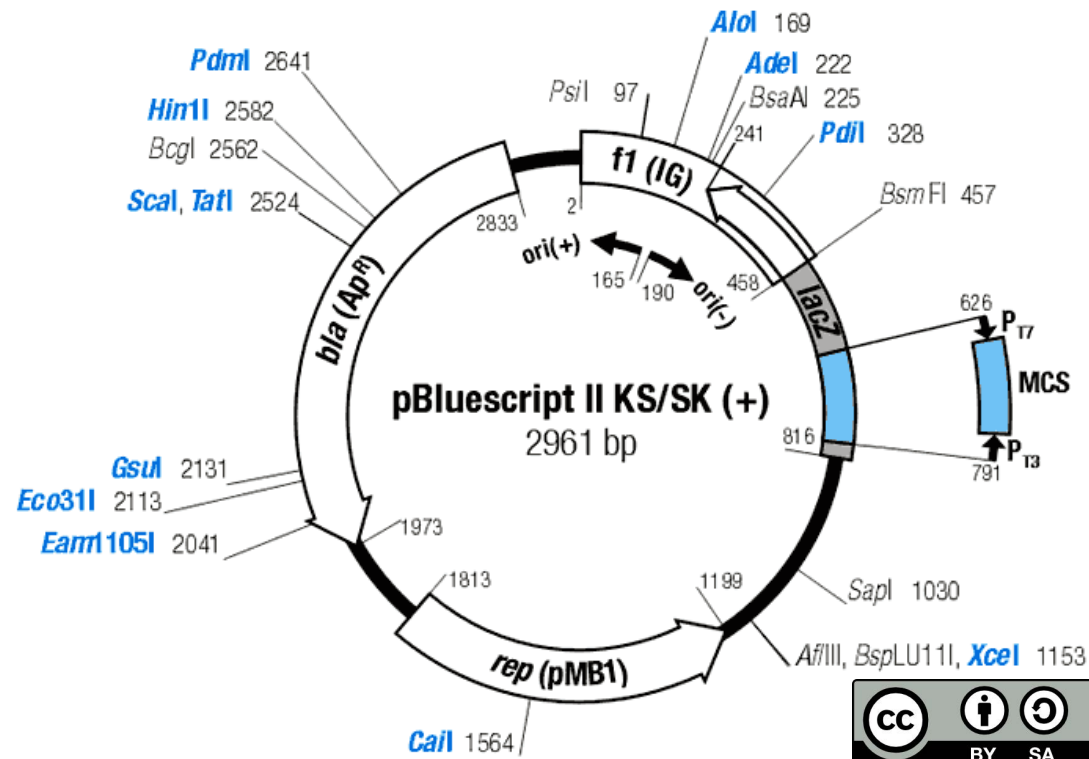


# PHAGEMID (phasmide)

Plasmid with origin of replication of filamentous phages - M13, f1 or fd

Production of ss templates for  
Sequencing or *in vitro*  
mutagenesis

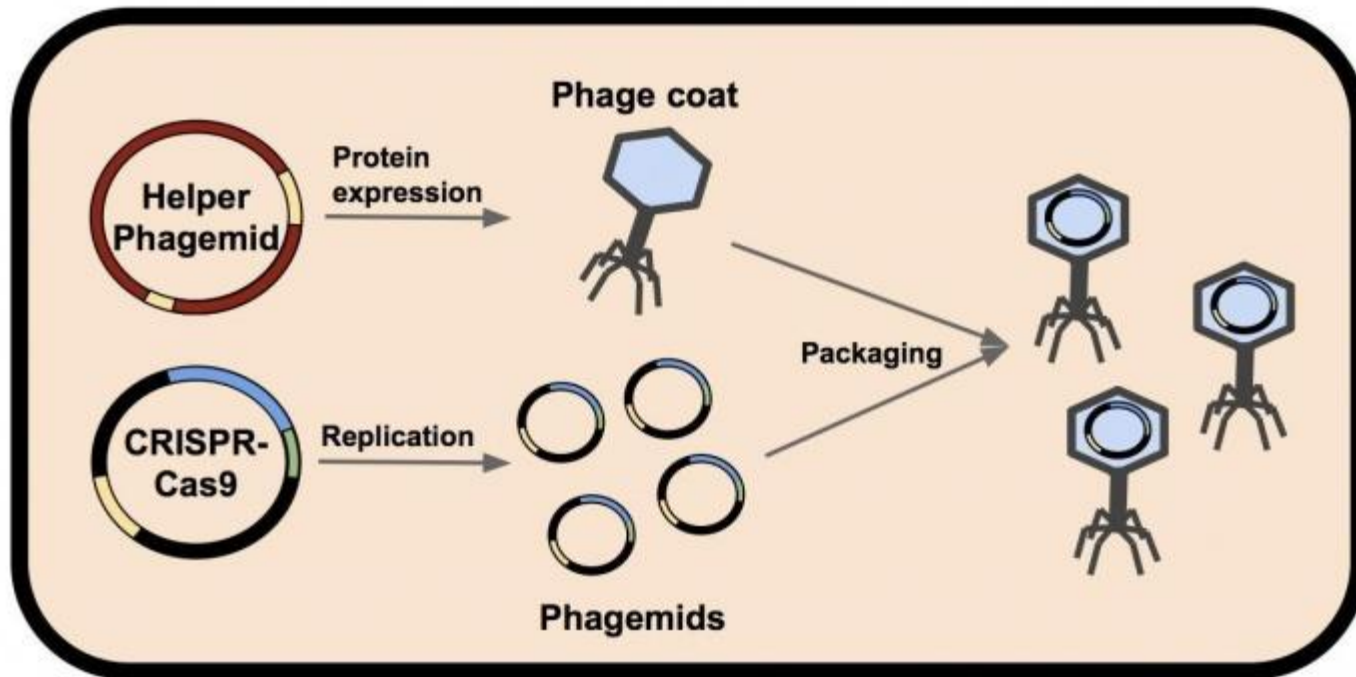
Polycloning site  
and „blue-white“ selection



**Replicates as filamentous phages**

**→ ssDNA in presence of helper virus**

**(defective virus encoding viral structural proteins and enzymes required for DNA packaging)**



## Plasmids

Easy manipulation

Short DNA fragments  
Up to 10 kb

Inefficient transformation

## Phages

Difficult manipulation

large DNA fragments  
~ 20-40 kb

Efficient transformation

ability to prepare ssDNA

# BACs bacterial artificial chromosomes

## Incorporate large fragments

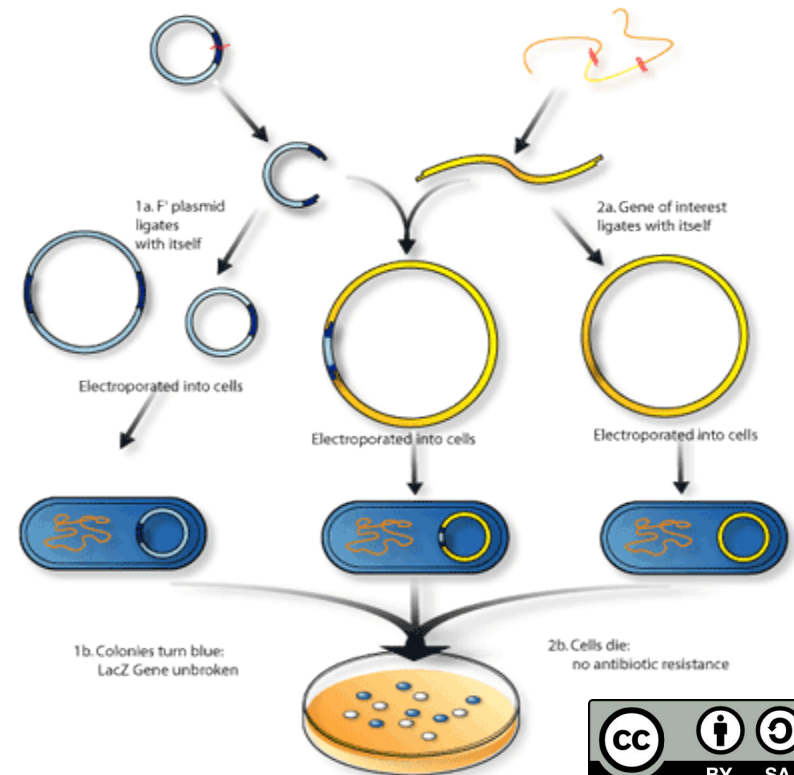
*oriS* – origin of replication,

*parA*, *parB*, a *parC* stable heritability

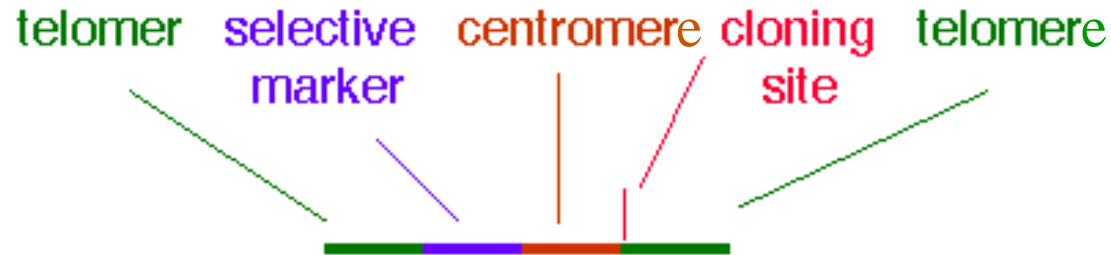
selection marker (chloramphenicol - resistance)

cloning sites

- Inserts up to hundreds of kbp, one copy, stable
- transformation – electroporation



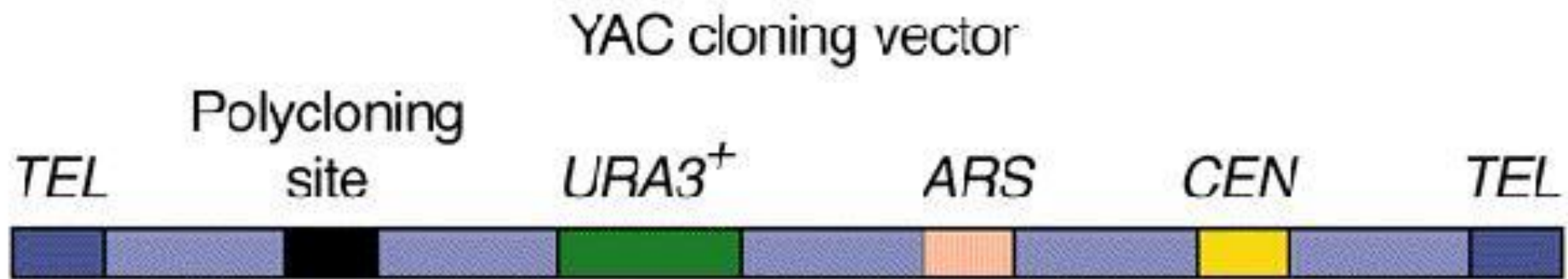
# YAC – yeast artificial chromosome



**Linear DNA vector – mimics yeast chromosome**

**Centromere, telomere..**

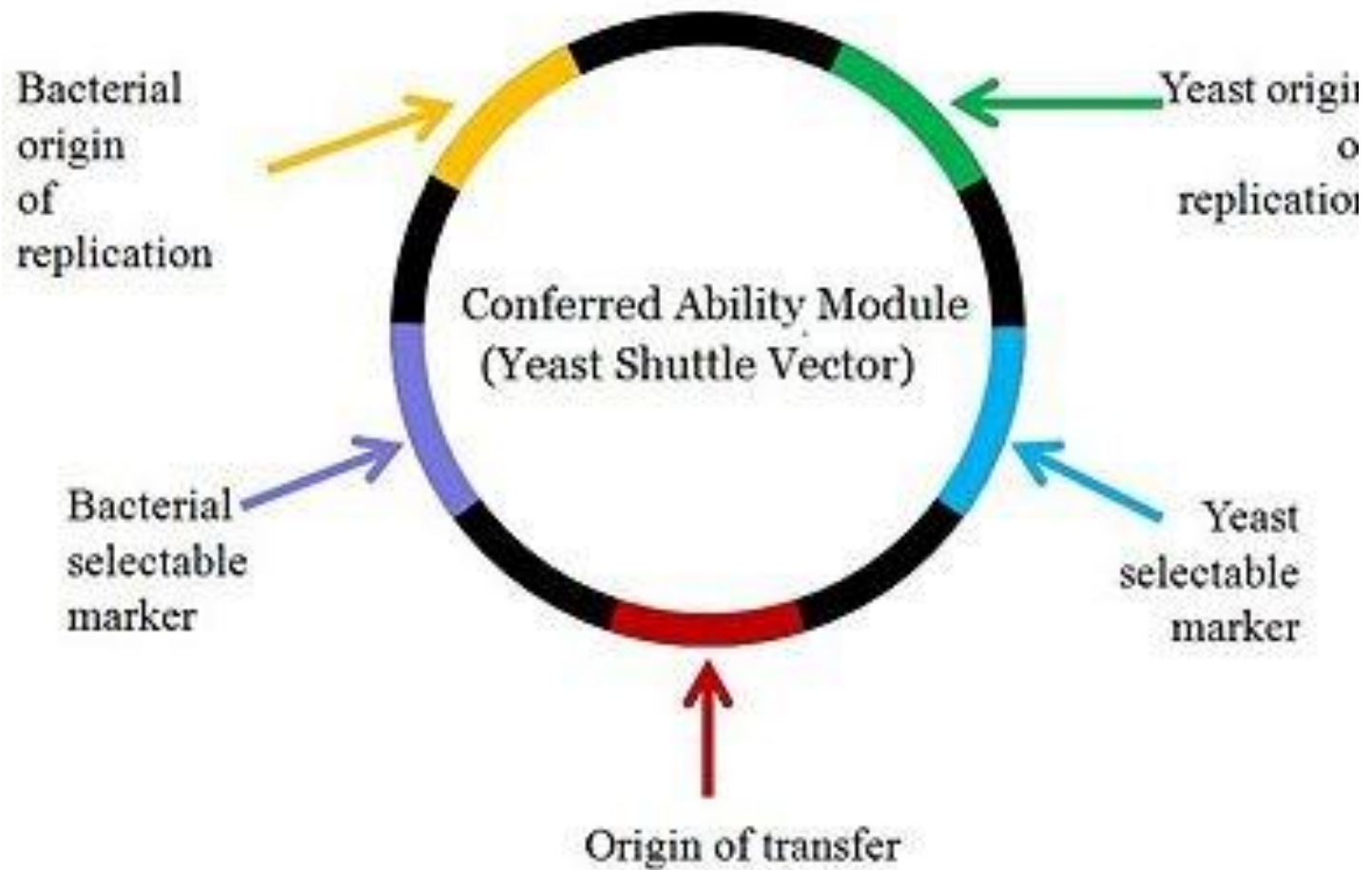
**Marker for selection, incorporation of thousands of bps**



<b>Vector</b>	<b>DNA Insert</b>	<b># clones complete human genomic library</b>
<b><i>E. coli</i> plasmid</b>	<b>&lt; 10 kb</b>	<b>330 000</b>
<b>λ phage</b>	<b>20 kb</b>	<b>170 000</b>
<b>cosmid</b>	<b>40 kb</b>	<b>83 000</b>
<b>BAC</b>	<b>200 kb</b>	<b>17 000</b>
<b>YAC</b>	<b>1000 kb</b>	<b>3 300</b>

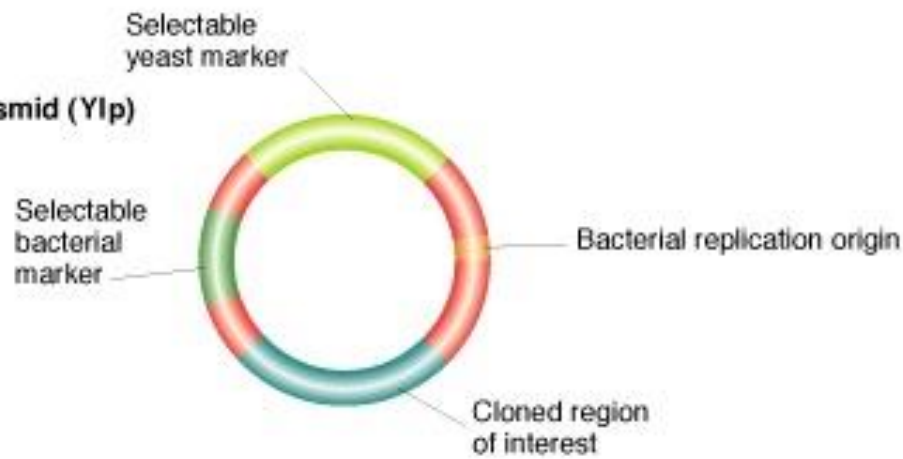
→ **BAC or YAC clones**

## Yeast vectors - „shuttle“





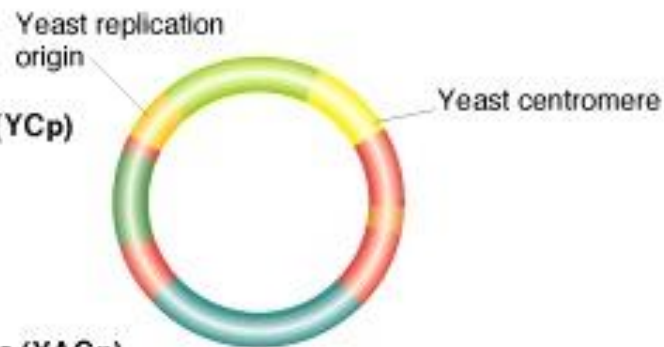
(a) Yeast integrative plasmid (YIp)



(b) Yeast episomal plasmid (YEp)



(c) Yeast centromere plasmid (YCp)



(d) Yeast artificial chromosome (YACp)

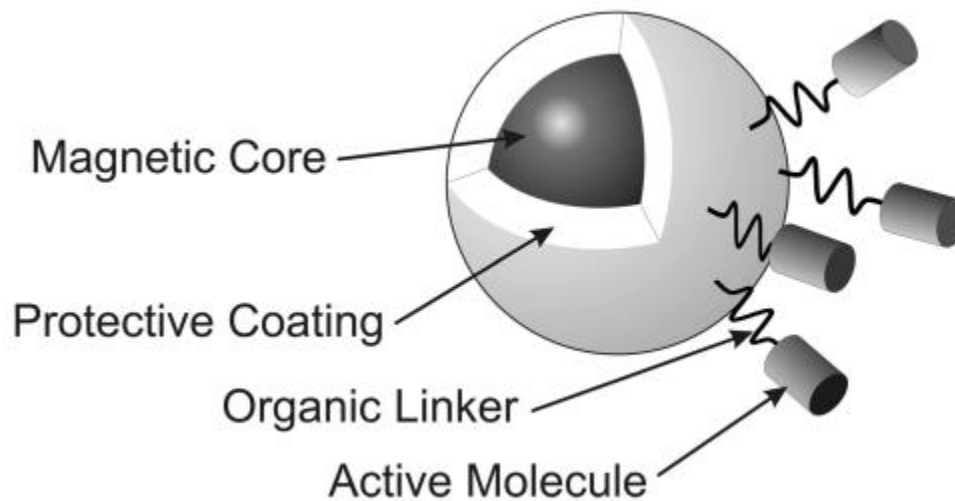


# Nanoparticles

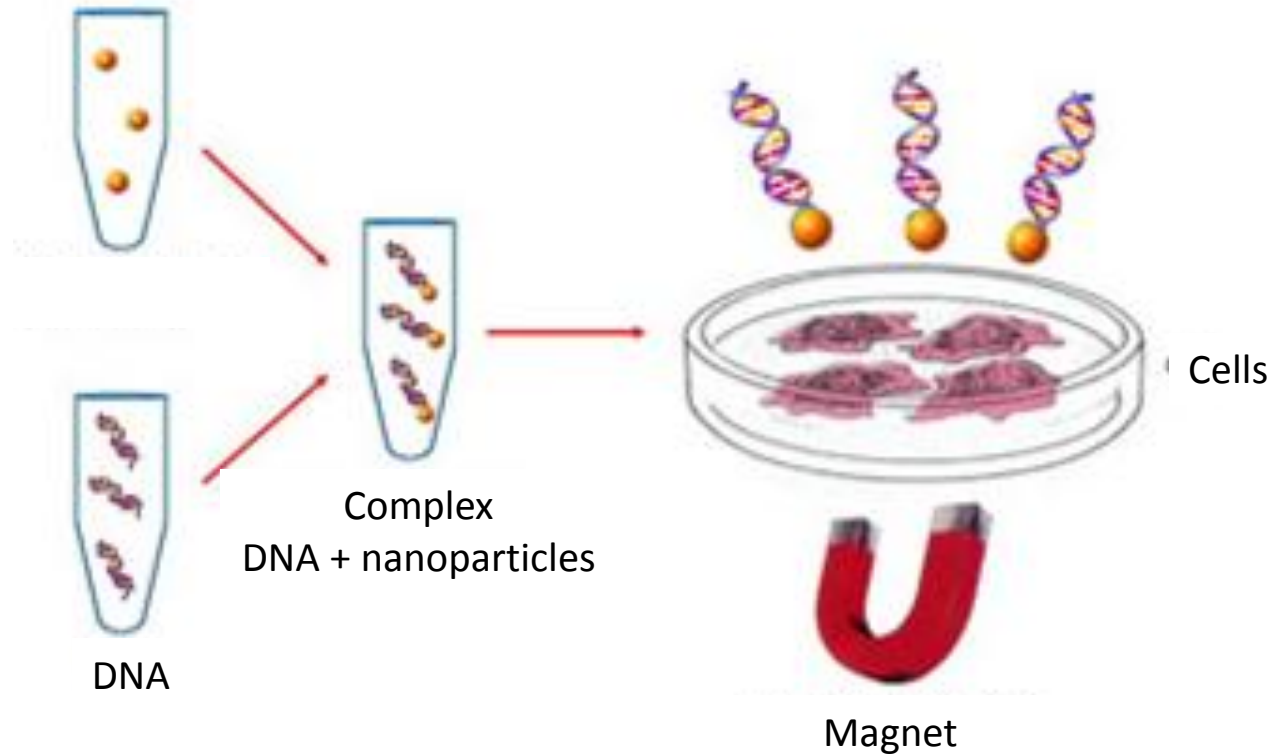
Ormosil – organic-modified silica particles

Non-viral vector

– successful DNA transfer into specific cells of organism



# Magnetofection



# Carbon nanofibers

- DNA-coated

Combination with magnetic material

- Allows oriented trafficking
- Universal (also for cells difficult to transfect)

# Viral vectors

**Almost 100% infection efficiency**

**Some viruses – integration into genome**

**Tropism – possible to modify**

**– change of surface glycoproteins**

**(e.g. specific lentiviral „coat envelope“**

**- replaced with promiscuous glycoprotein G of Vesicular stomatitis virus)**

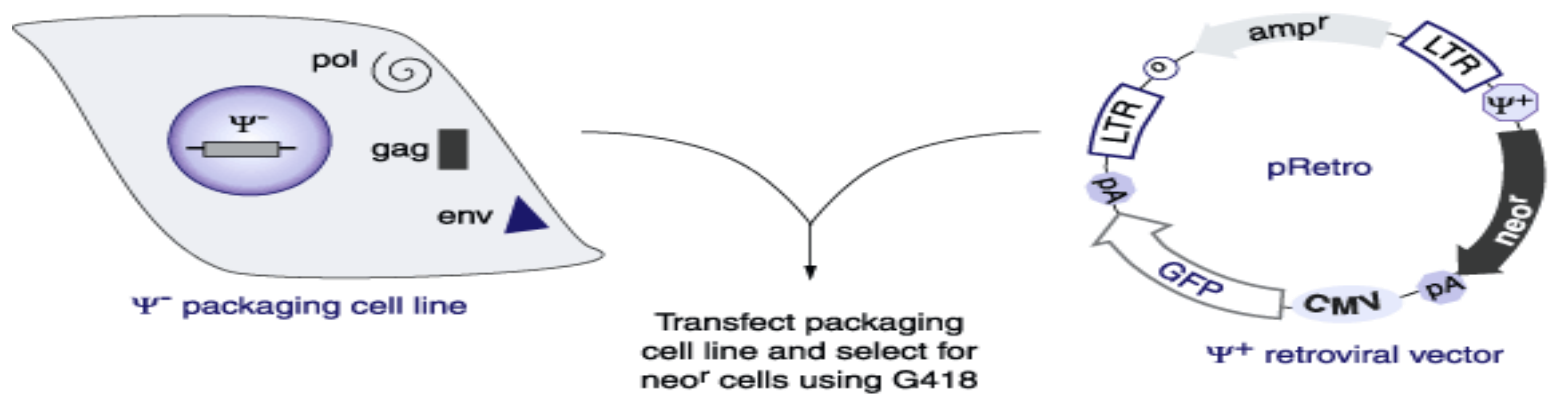
**- Sometimes advantageous specific tropism – targeted introducing of DNA**

**- vector – minimal x maximal effect on the cell physiology**

# Retroviral vectors

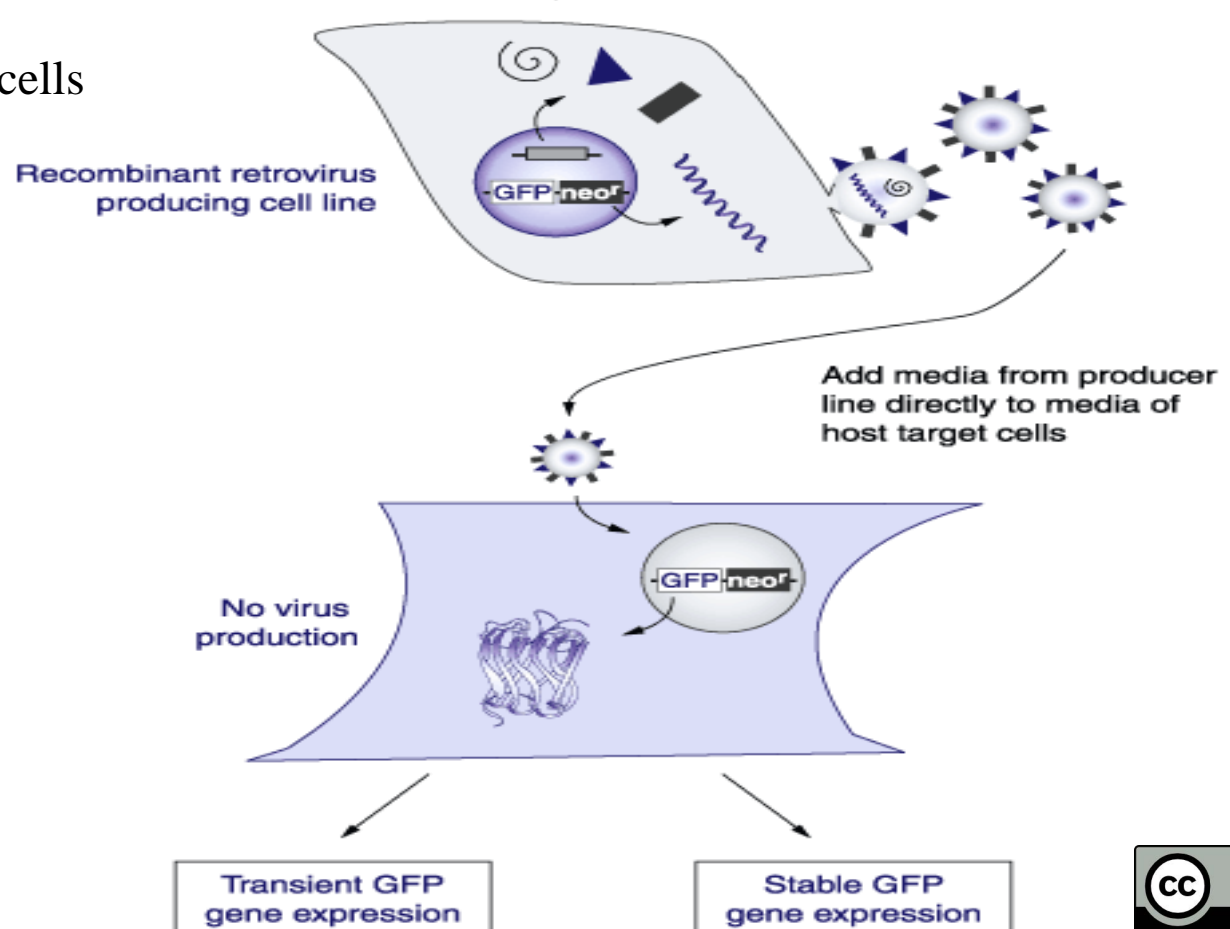
**RNA  $\rightarrow$  DNA - RT**

**Stable integration - IN**



# Retroviral expression vectors

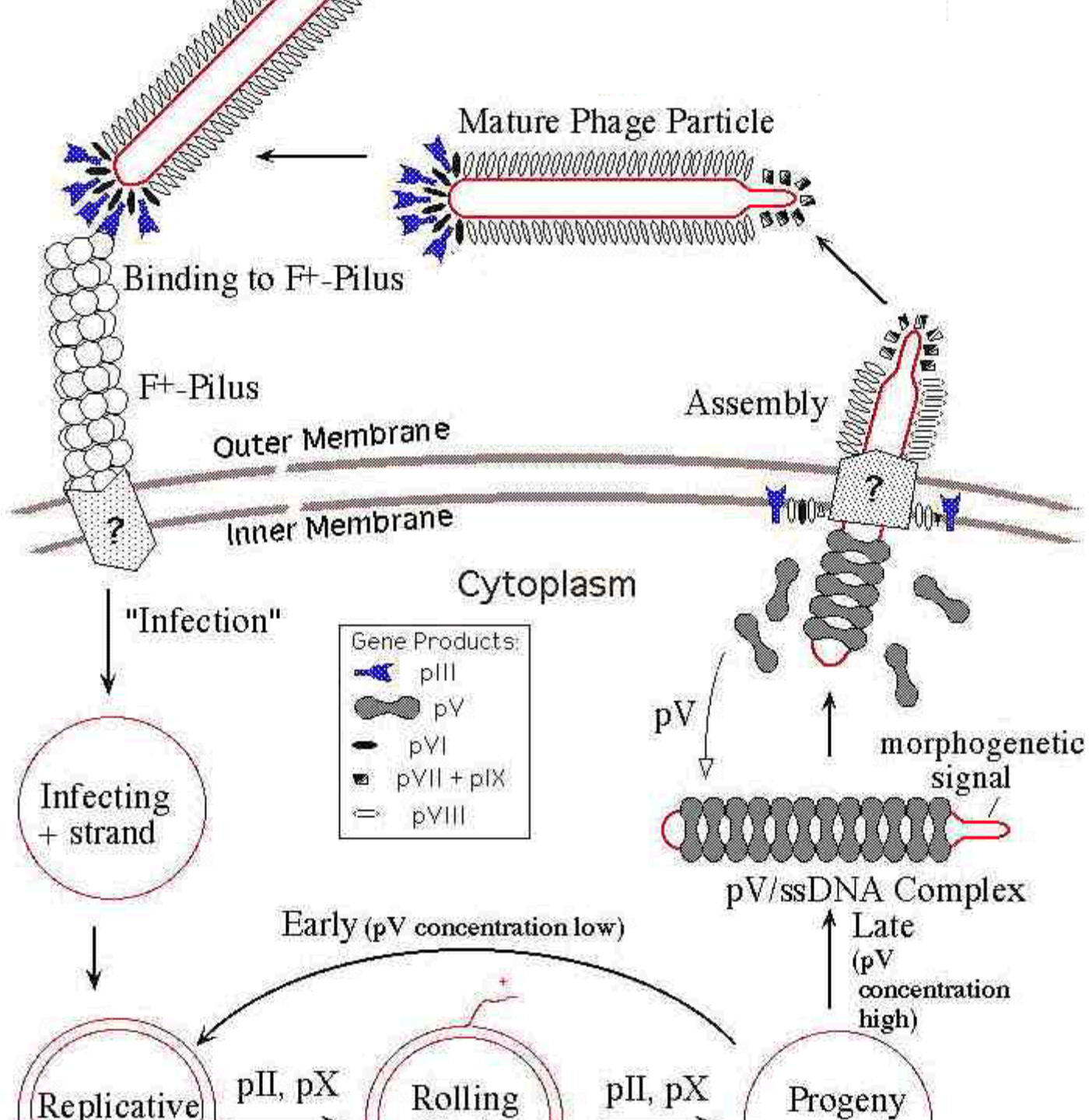
gene transfer – to selected cells

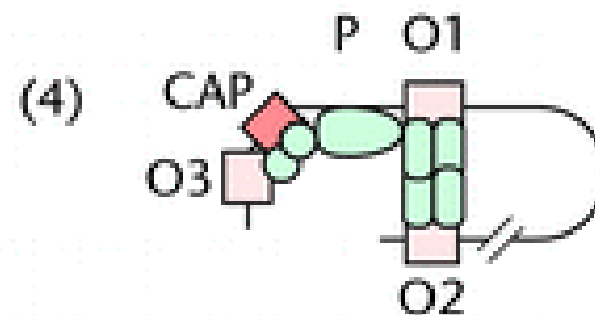
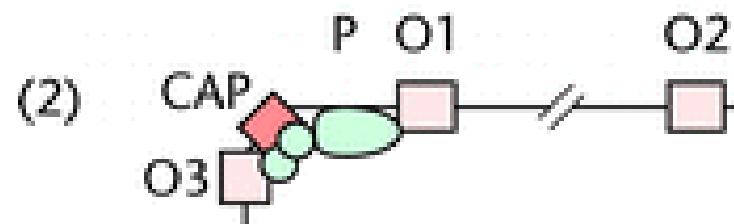
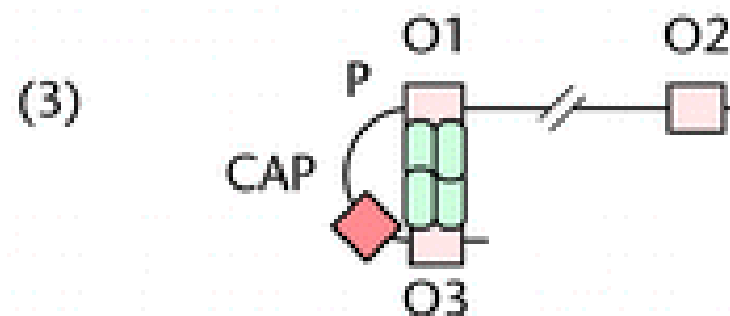
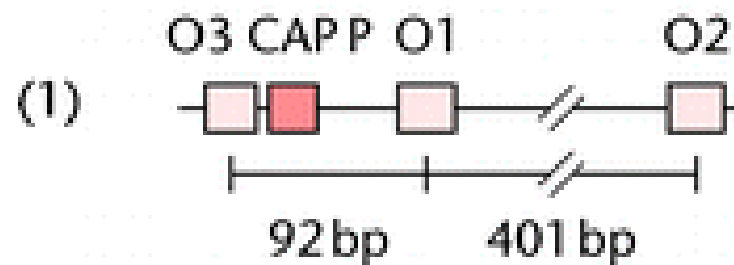
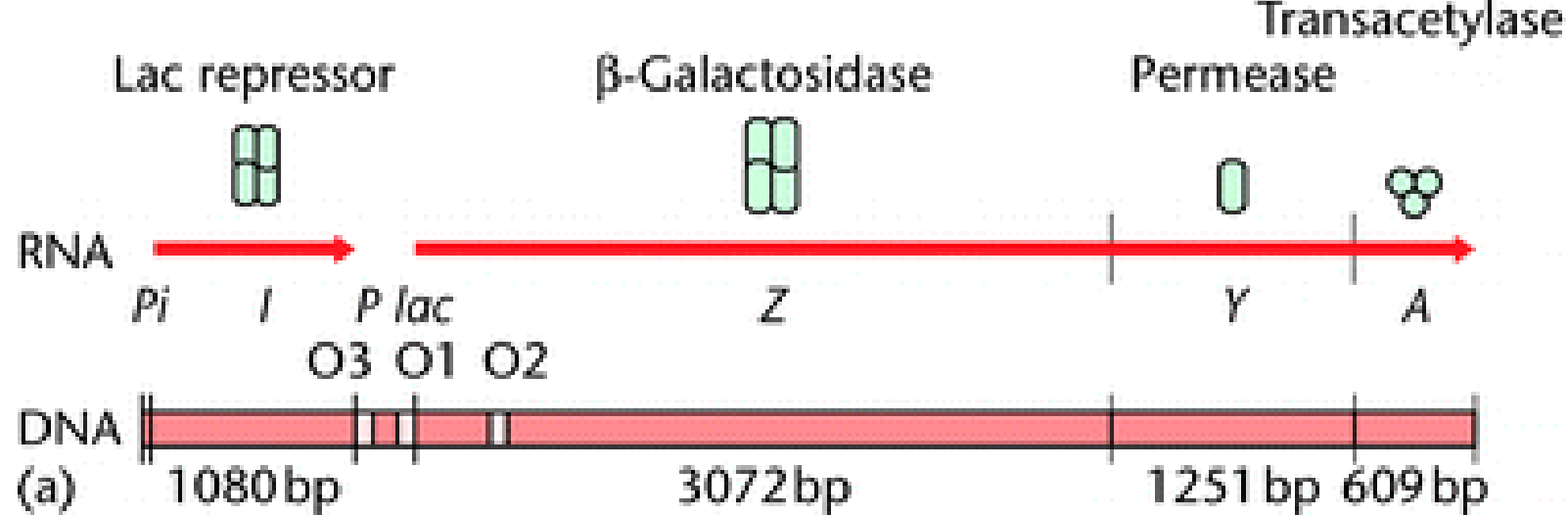




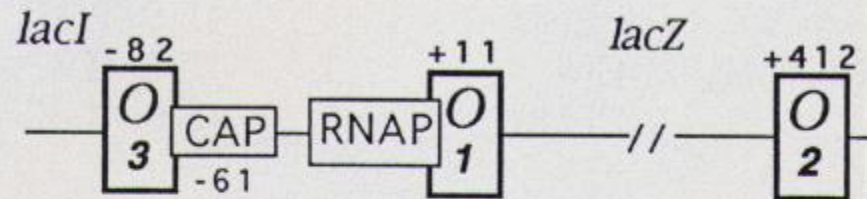








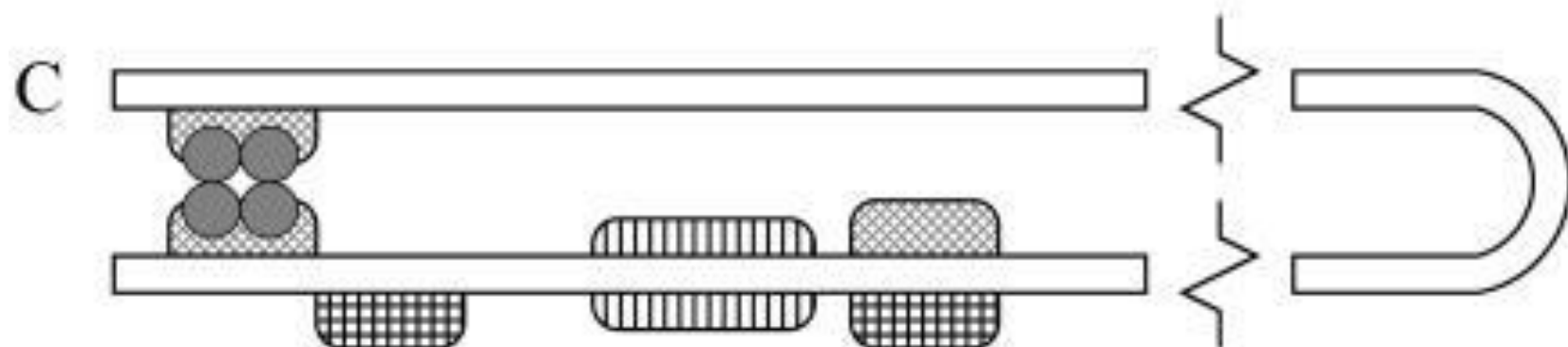
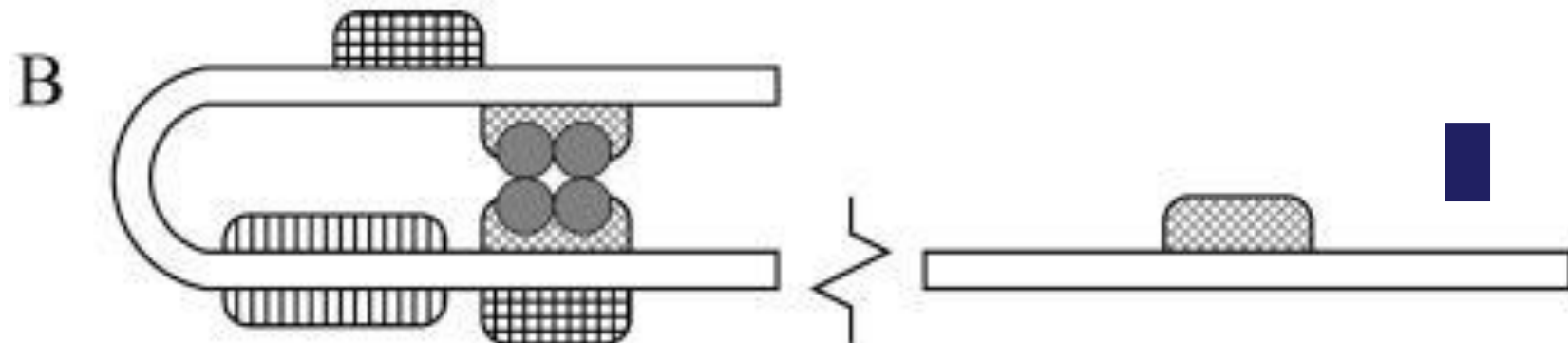
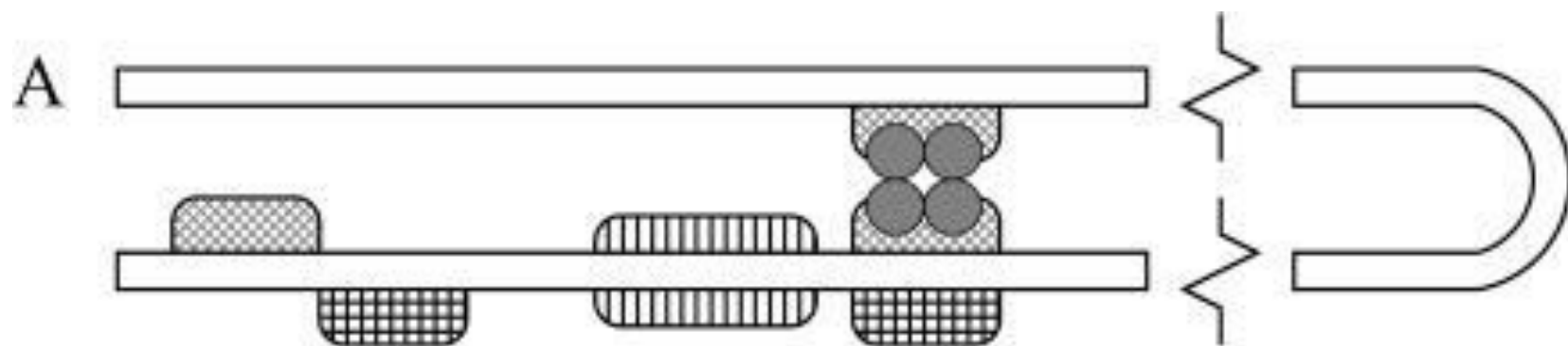
(b)

**A****B**

*O*<sub>1</sub> 5' AATTGTGAGCGGATAACAATT 3'  
*O*<sub>2</sub> 5' AAaTGTGAGCGagTAACAAcc 3'  
*O*<sub>3</sub> 5' ggcaGTGAGCGcAacgCAATT 3'

**C**

1                      10                      21  
 5' GAATTGTGAGC-GCTCACAATT 3'  
 3' TTAACACTCG-CGAGTGTTAAG 5'



gene VIII - protein – tubular assembly ca 2 700 identical subunits enveloping the viral genome

M13 phage  
s.s. DNA genome

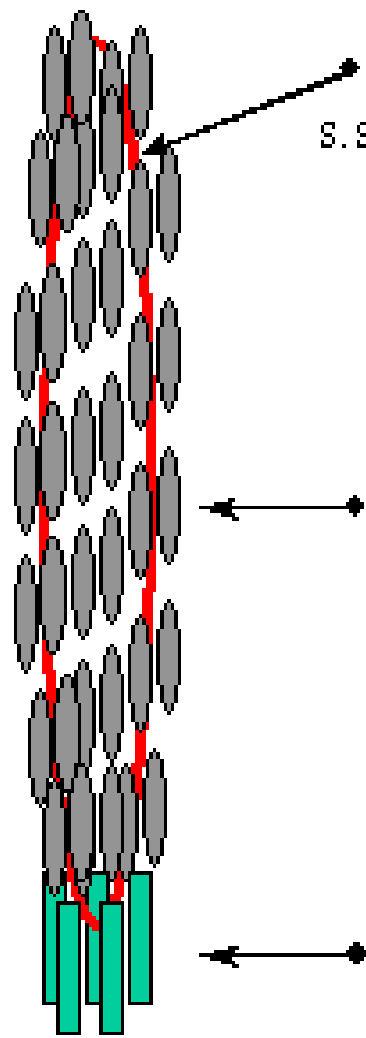
5 - 8 copies of the product of gene III at the end of the phage

-Binding to sex pilus

Pilus of *E. coli* contains the "F factor"  
(extrachromosomal element)

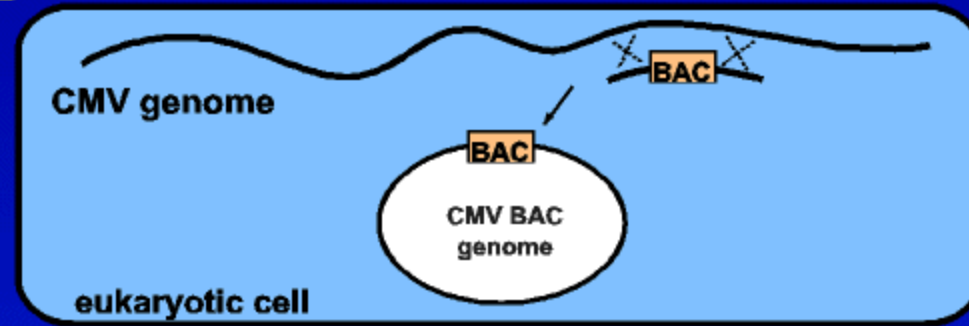
~2,700 copies  
of gene VIII  
protein

5-8 copies  
of gene III  
protein

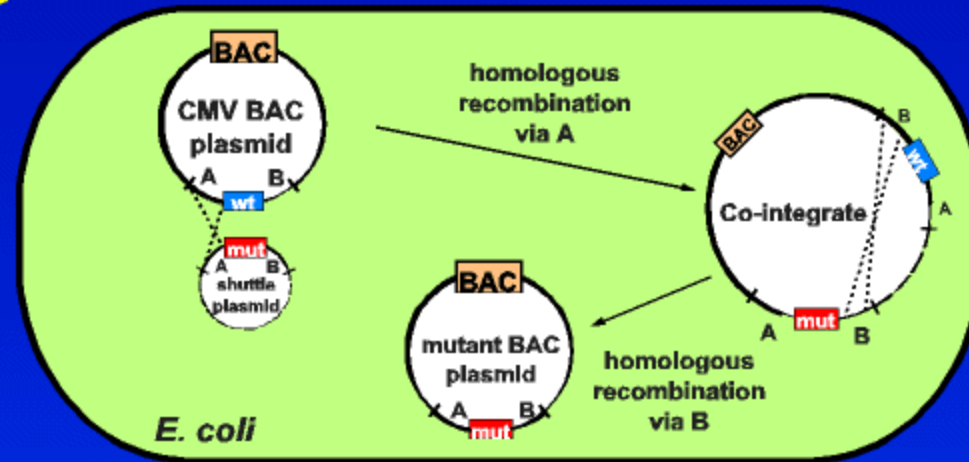


# as Infectious Bacterial Artificial Chromosomes in *E. coli*

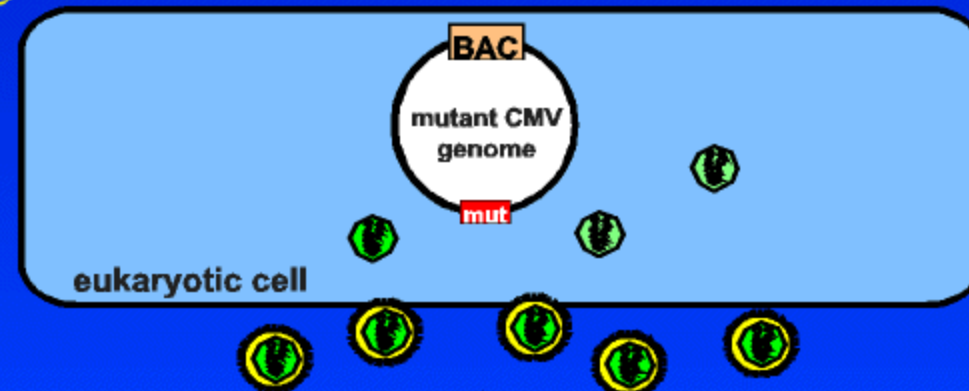
A



B



C





Phage  $\lambda$   
50 kb  
2 pages



*Escherichia coli*  
(bacterial)  
4.7 Mb  
200 pages



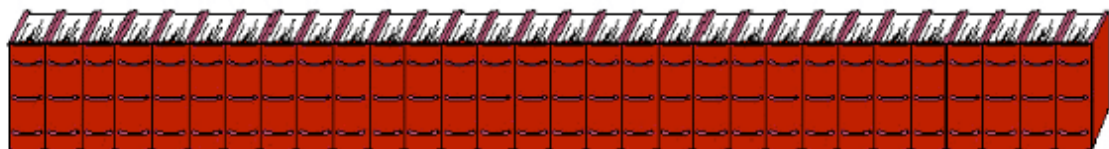
*Saccharomyces cerevisiae*  
(yeast)  
12.5 Mb  
500 pages



*Caenorhabditis elegans*  
(nematode)  
*Arabidopsis thaliana*  
(plant)  
100 Mb  
3 volumes



*Drosophila melanogaster*  
(fruit fly)  
165 Mb  
5 volumes



Human being  
3000 Mb  
80 volumes



25 kb per page  
1500 pages  
per volume  
(2 inches thick)

**Uveřejněné materiály jsou určeny studentům Vysoké školy chemicko-technologické v Praze**

**jako studijní materiál. Některá textová i obrazová data v nich obsažená jsou převzata**

**z veřejných zdrojů. V případě nedostatečných citací nebylo cílem autora/ů záměrně poškodit event. autora/y původního díla.**

**S eventuálními výhradami se prosím obračejte na autora/y konkrétního výukového materiálu, aby**

**bylo možné zjednat nápravu.**

**The published materials are intended for students of the University of Chemistry and Technology, Prague as a study material. Some text and image data contained therein are taken from public sources. In the case of insufficient quotations, the author's intention was not to intentionally infringe the possible author(s) rights to the original work.**

**If you have any reservations, please contact the author(s) of the specific teaching material in order to remedy the situation.**