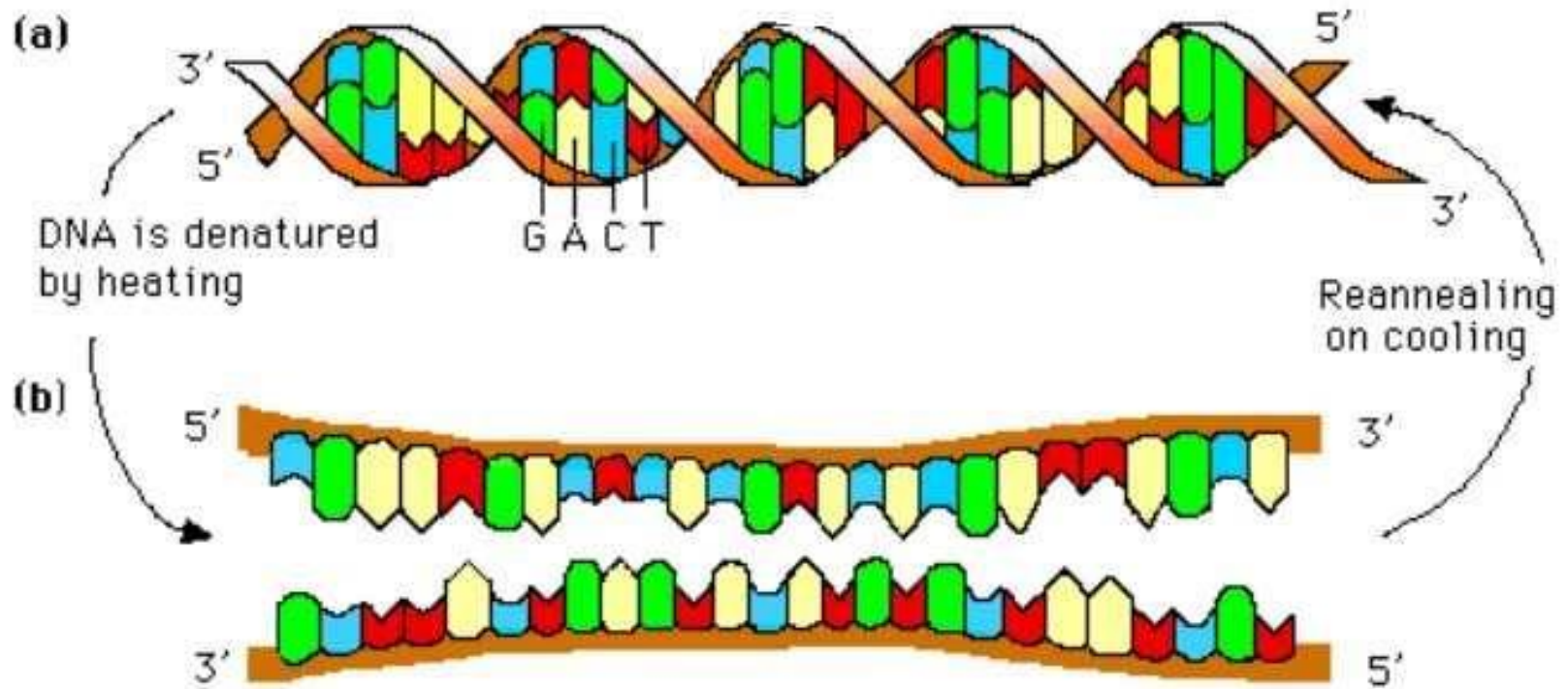
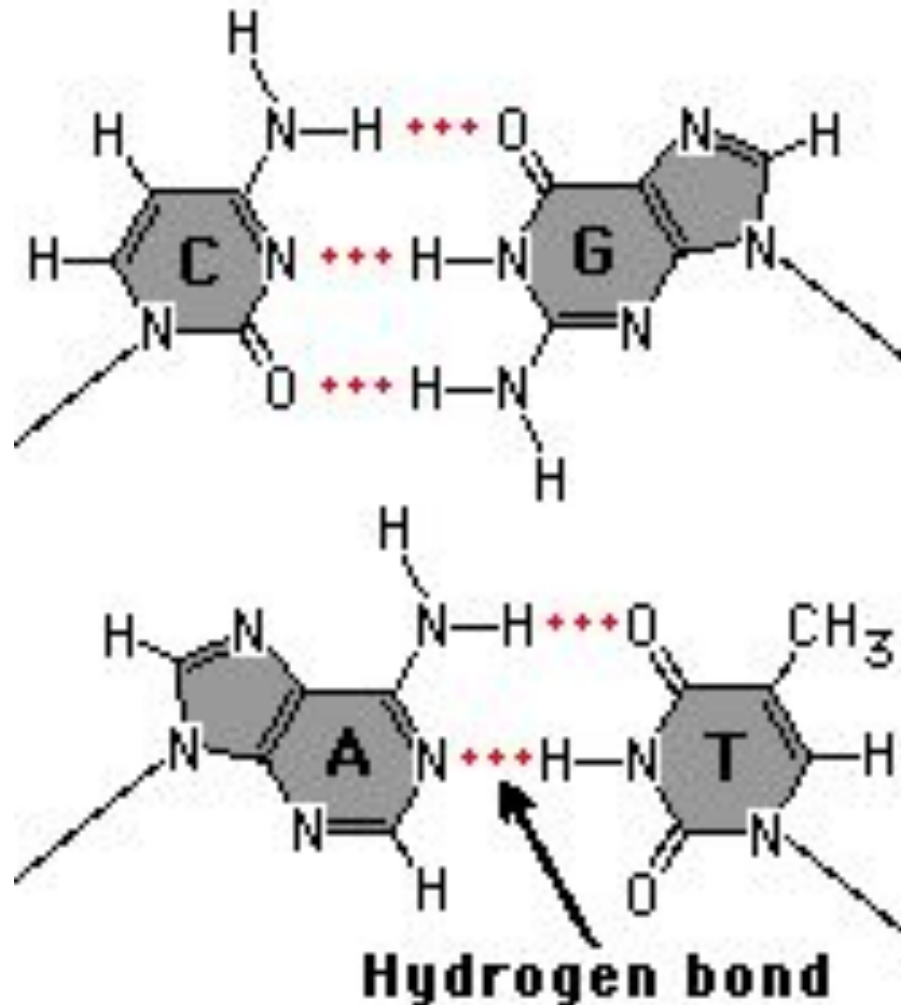


# 1. Hybridization (DNA-DNA or DNA-RNA)



# HYBRIDIZATION? –

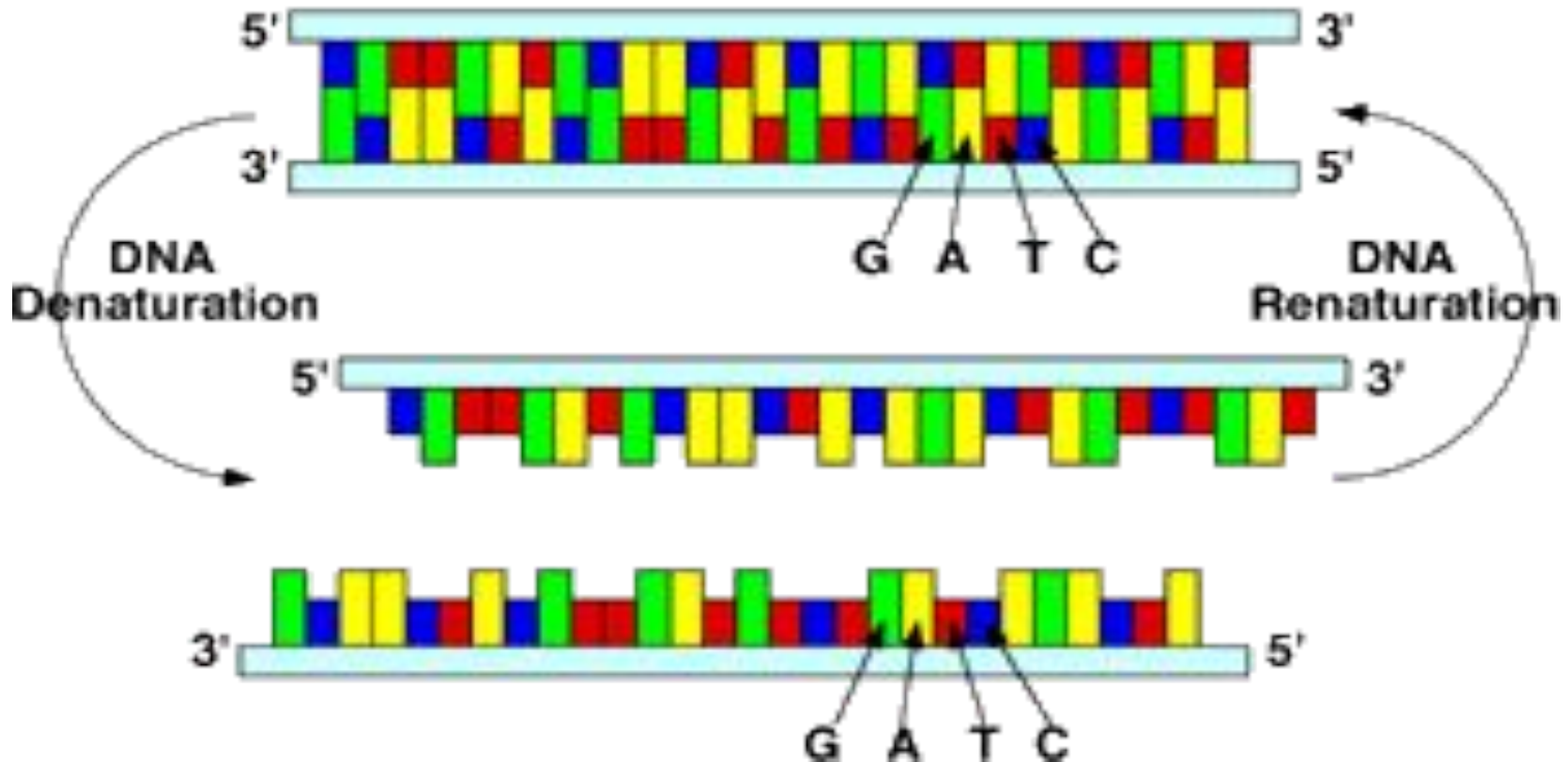
Yes, it is about this familiar picture



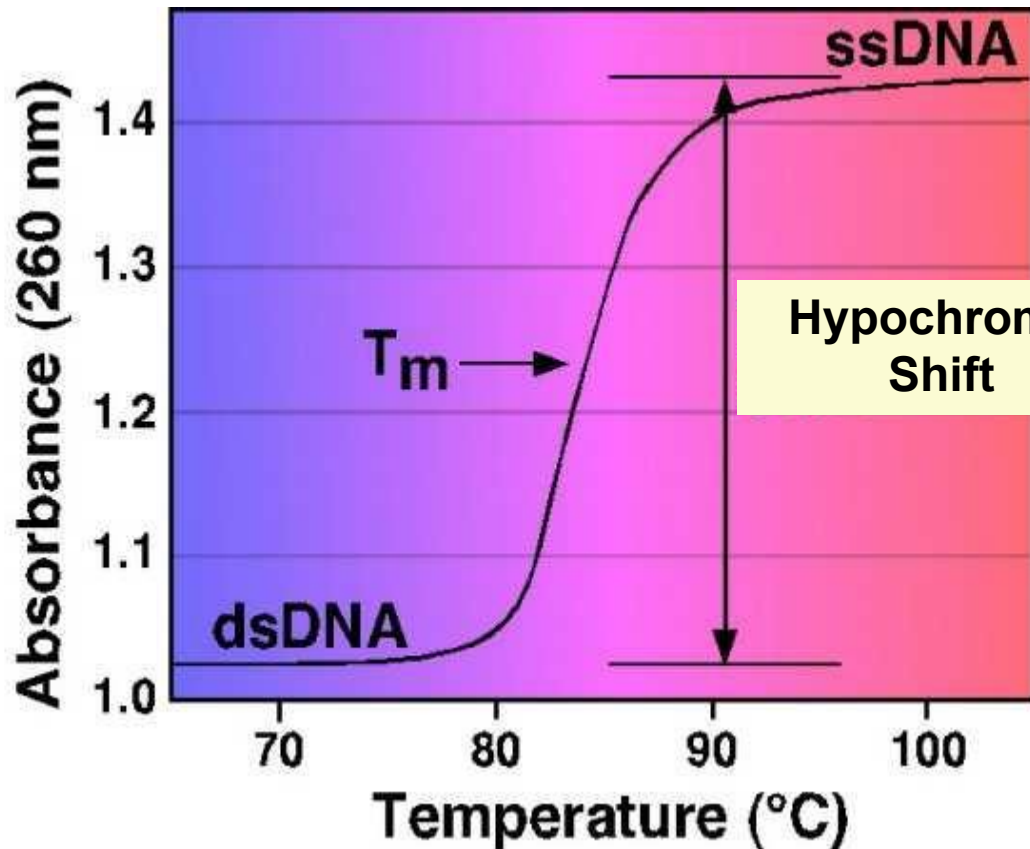
# We can denature and renature DNA by heating/cooling

## Denaturation/Renaturation

### DNA



As DNA is heated, it reaches a temperature where the strands separate (DNA melts).



**ssDNA**

H-bonds between basepairs are broken and the strand unwind.

**Melting curve.**

**Unstacked bases**  
(random orientation)  
**absorb more light**  
than neatly stacked  
(oriented) base-pairs

**$T_m$  (melting temp)**

The temperature at which DNA is half unfolded

**$T_m$  is a measure of the stability of dsDNA under a given set of conditions**

# T<sub>m</sub> of DNA is affected by:

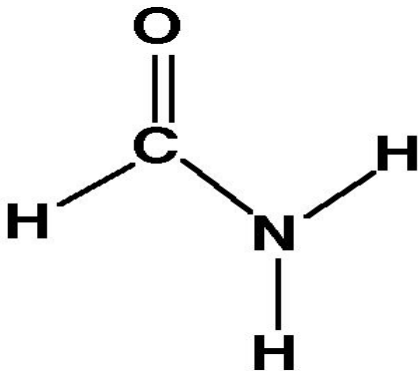
## 1. Base Composition :

higher the GC content, the higher the T<sub>m</sub>.

## 2. Ionic Strength :

as the ionic strength increases, so does T<sub>m</sub>.  
Double helical DNA is stabilized by cations.

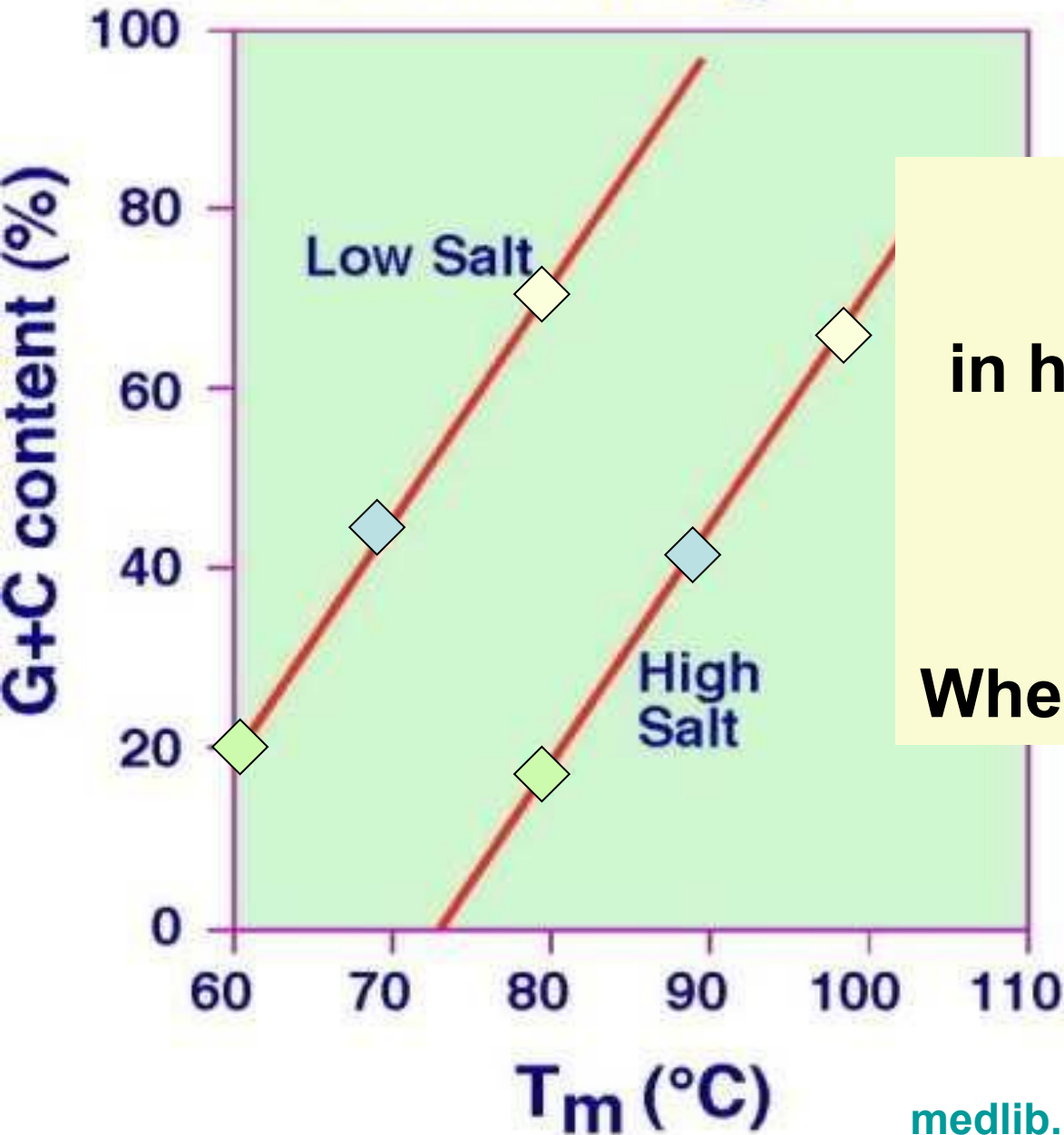
**Divalent cations (Mg<sup>2+</sup>) are more effective**  
than monovalent cations (Na<sup>+</sup> or K<sup>+</sup>).



## 3. Organic Solvents –

formamide for instance  
lowers the T<sub>m</sub> by weakening  
the hydrophobic interactions.

# $T_m$ depends on G+C content and ionic strength

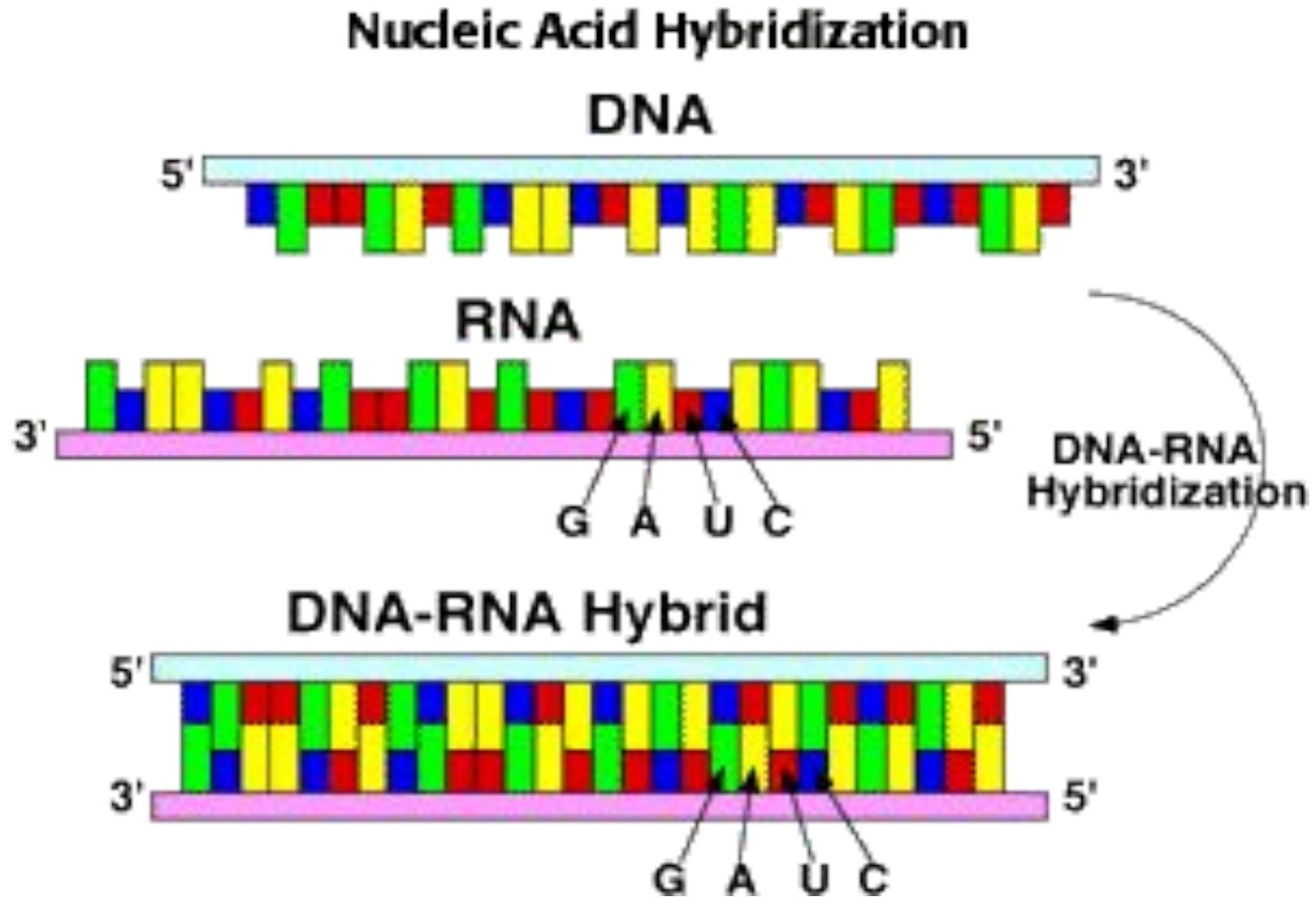


**DNA  
more STABLE  
in high-salt conditions.**

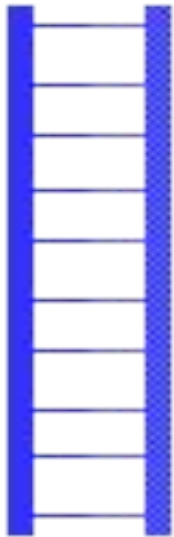
**DNA  
more STABLE  
When contains many GC**

# RNA can bind DNA

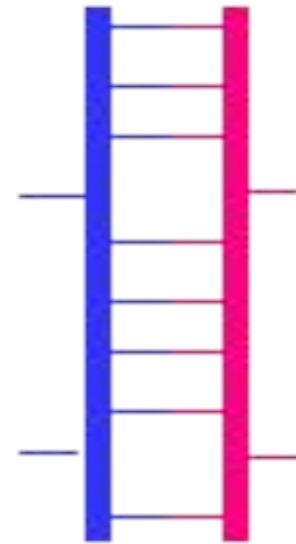
(U is equivalent of T in hybridization)



# Hybridization could be less than perfect

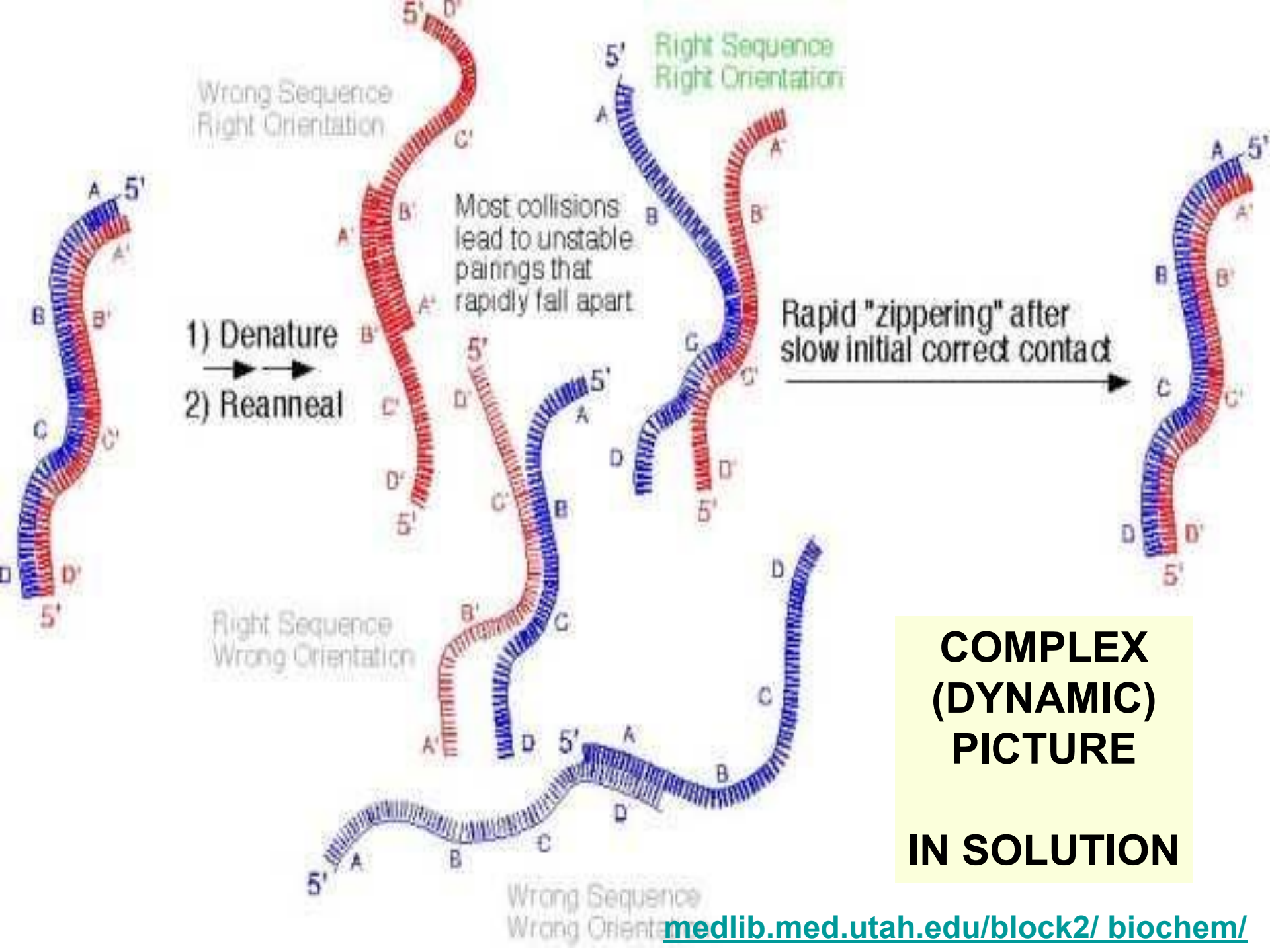


Homoduplex DNA,  
both strands from  
the same species



Heteroduplex DNA,  
strands from  
different species,  
with some bases  
unpaired



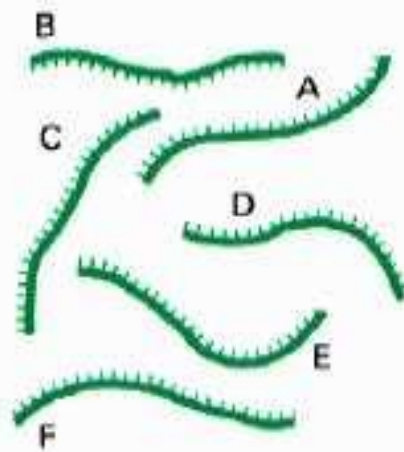


**COMPLEX  
(DYNAMIC)  
PICTURE  
IN SOLUTION**

Single-stranded DNA probes for gene A

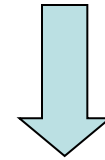


+



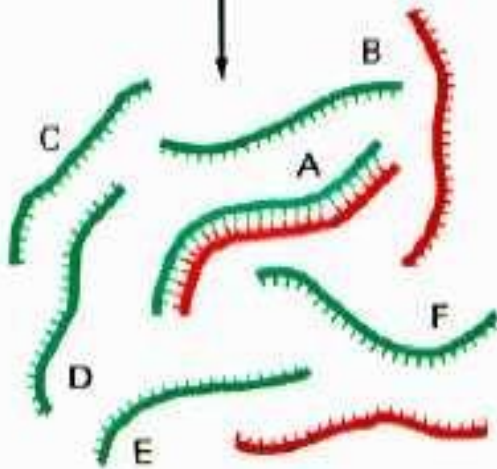
Mixture of single-stranded DNA molecules

**42 C**  
is more stringent  
condition  
than **35 C**



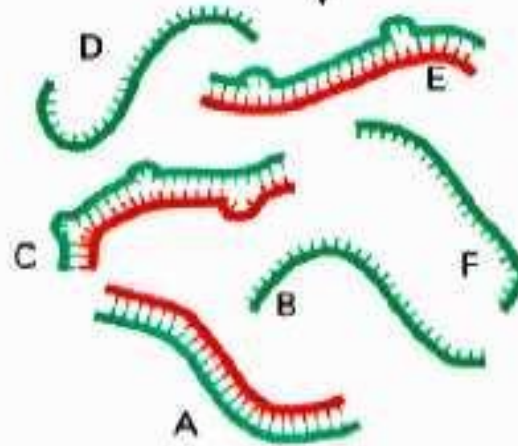
**(hybridization is more specific)**

Hybridization in 50% formamide at 42°C



Only A forms stable double helix

Hybridization in 50% formamide at 35°C



A, C, and E all form stable double helices

# So, hybridization is a most obvious phenomenon to use for **specific DNA detection**

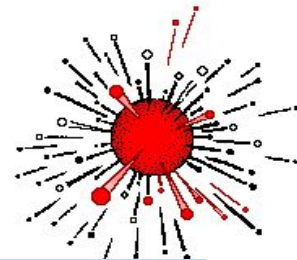
**Specific probe self-anneals to target DNA**

Only problem – DNA is invisible

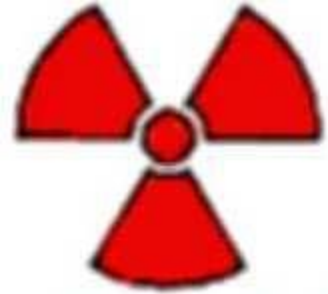
**How to visualize DNA?**



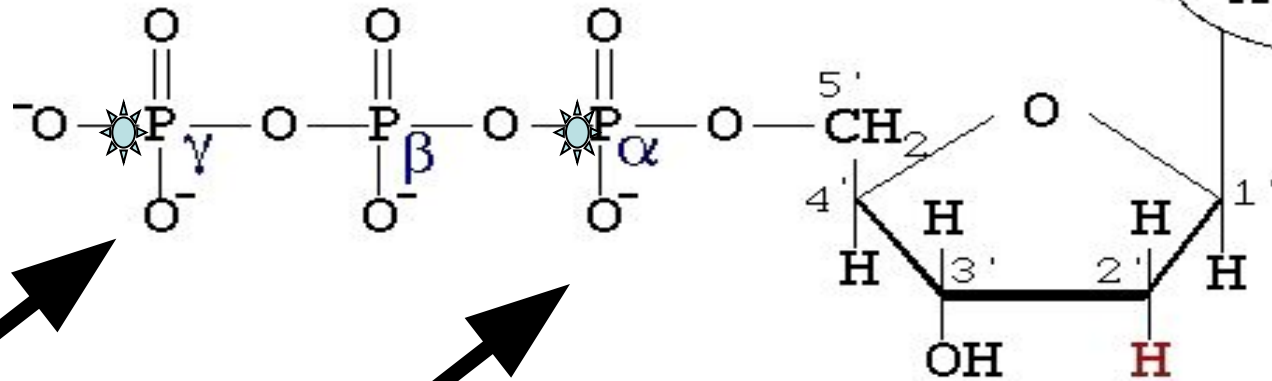
**Radioactively**



**Fluorescently**



# Polymerase labeling of DNA

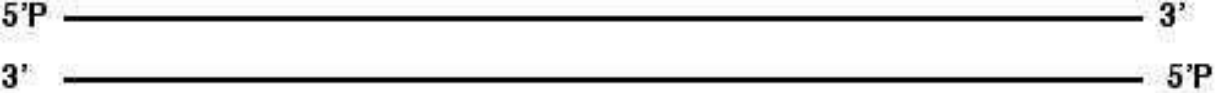


**Gamma-33-ATP**

**Alpha-32-ATP**

dNTP

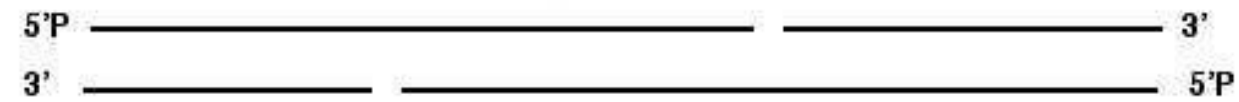




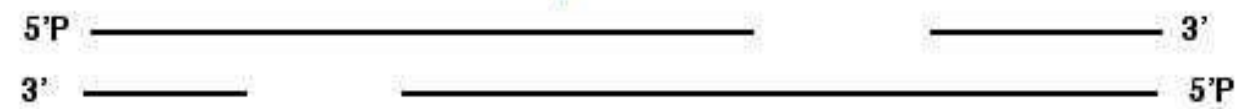
**DNase I**

**Labeling by  
NICK TRANSLATION**

Will work without DNase,  
as there are always nicks in DNA

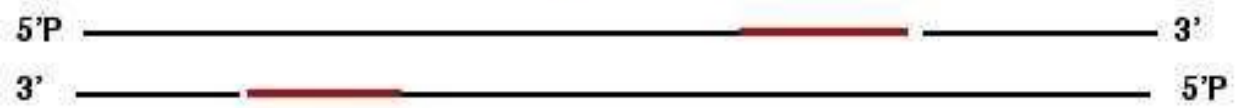


**Polymerase I  
(exonuclease activity)**



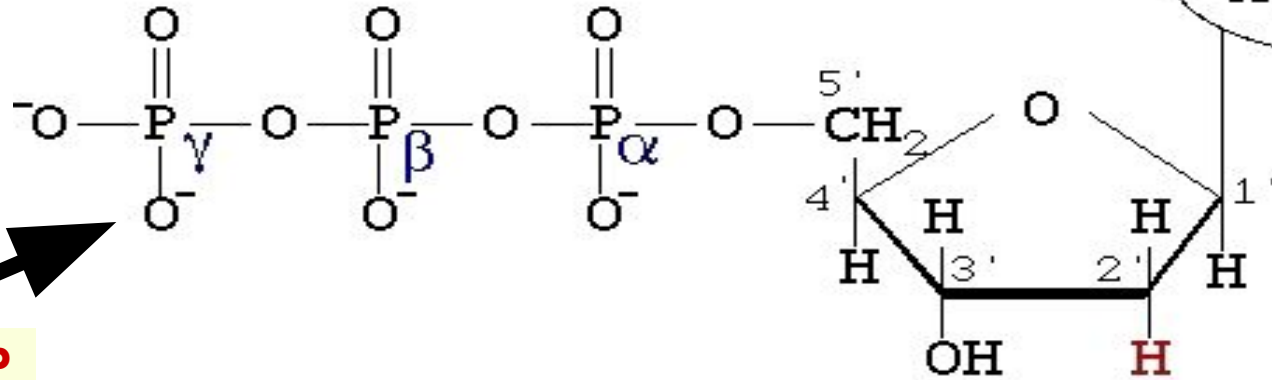
+ dTTP  
+ dCTP  
+ dGTP  
+ dAT<sup>32</sup>P

**Polymerase I  
(polymerase activity)**

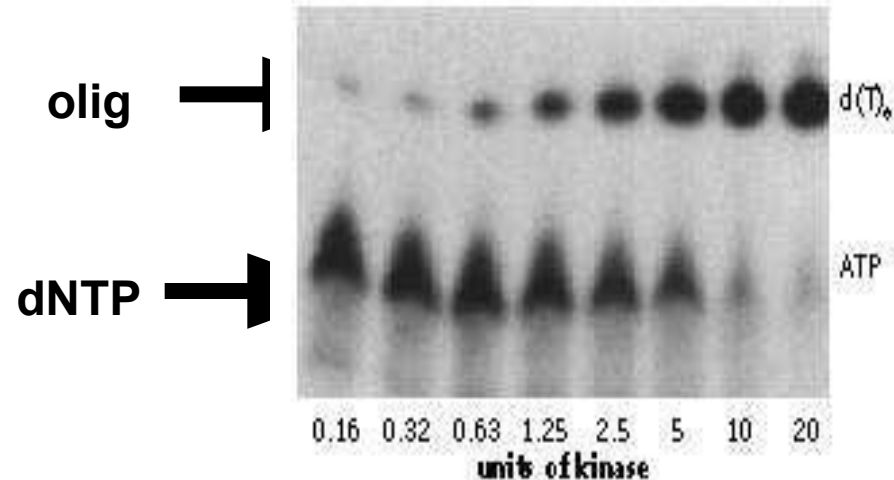
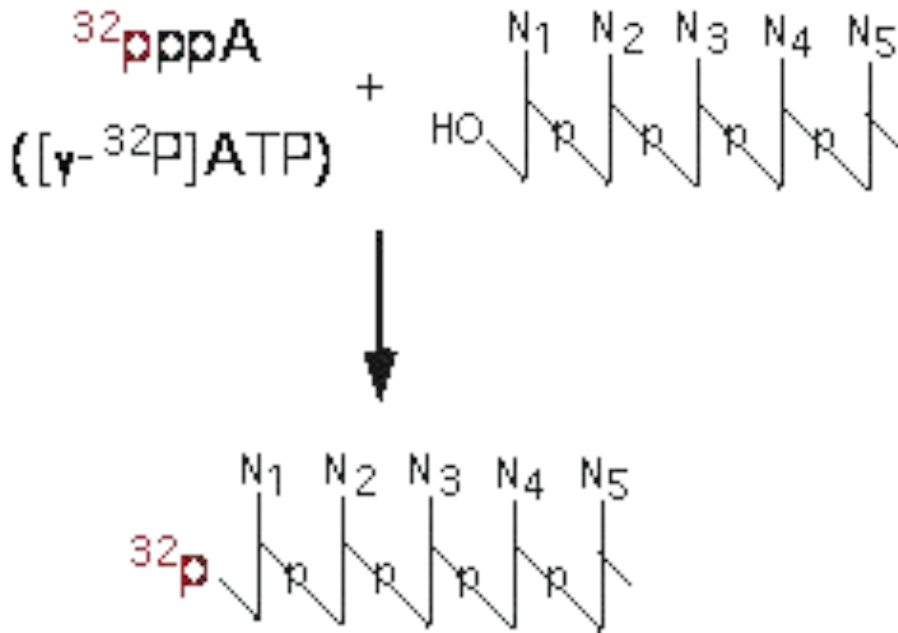


# T4 polynucleotide kinase labeling

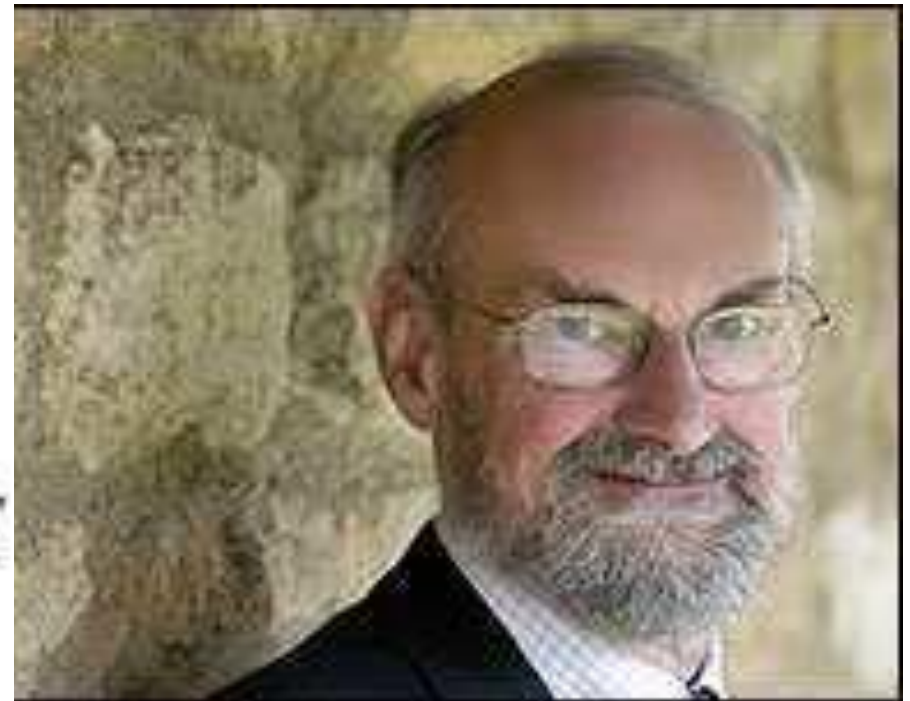
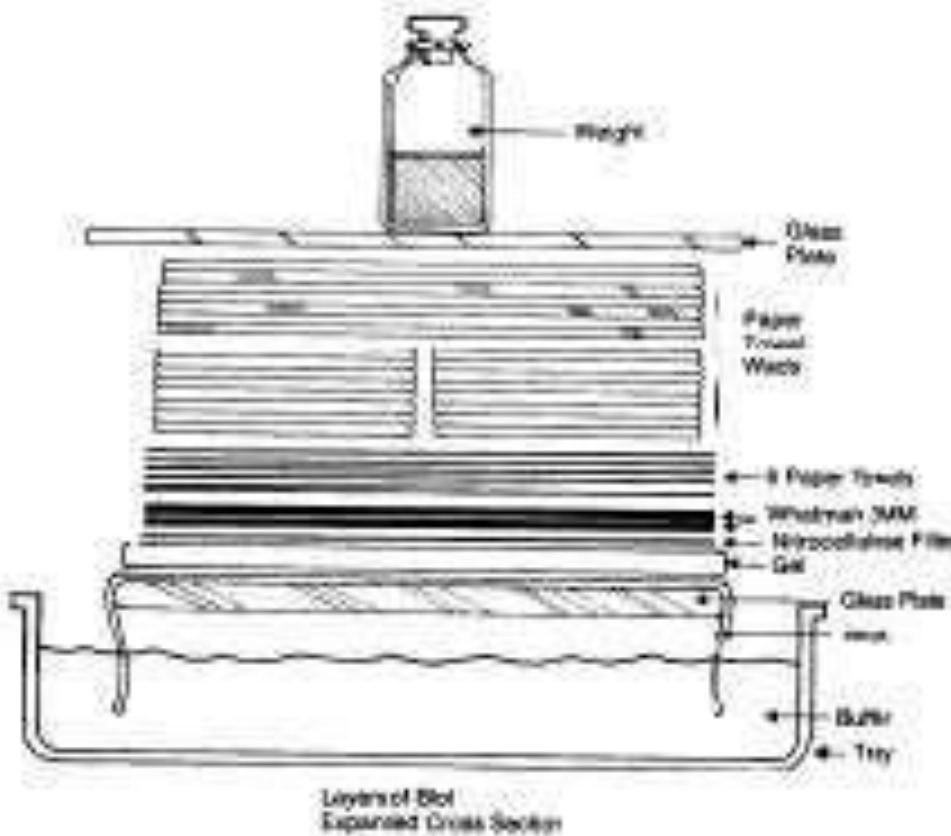
Used for oligonucleotide labeling

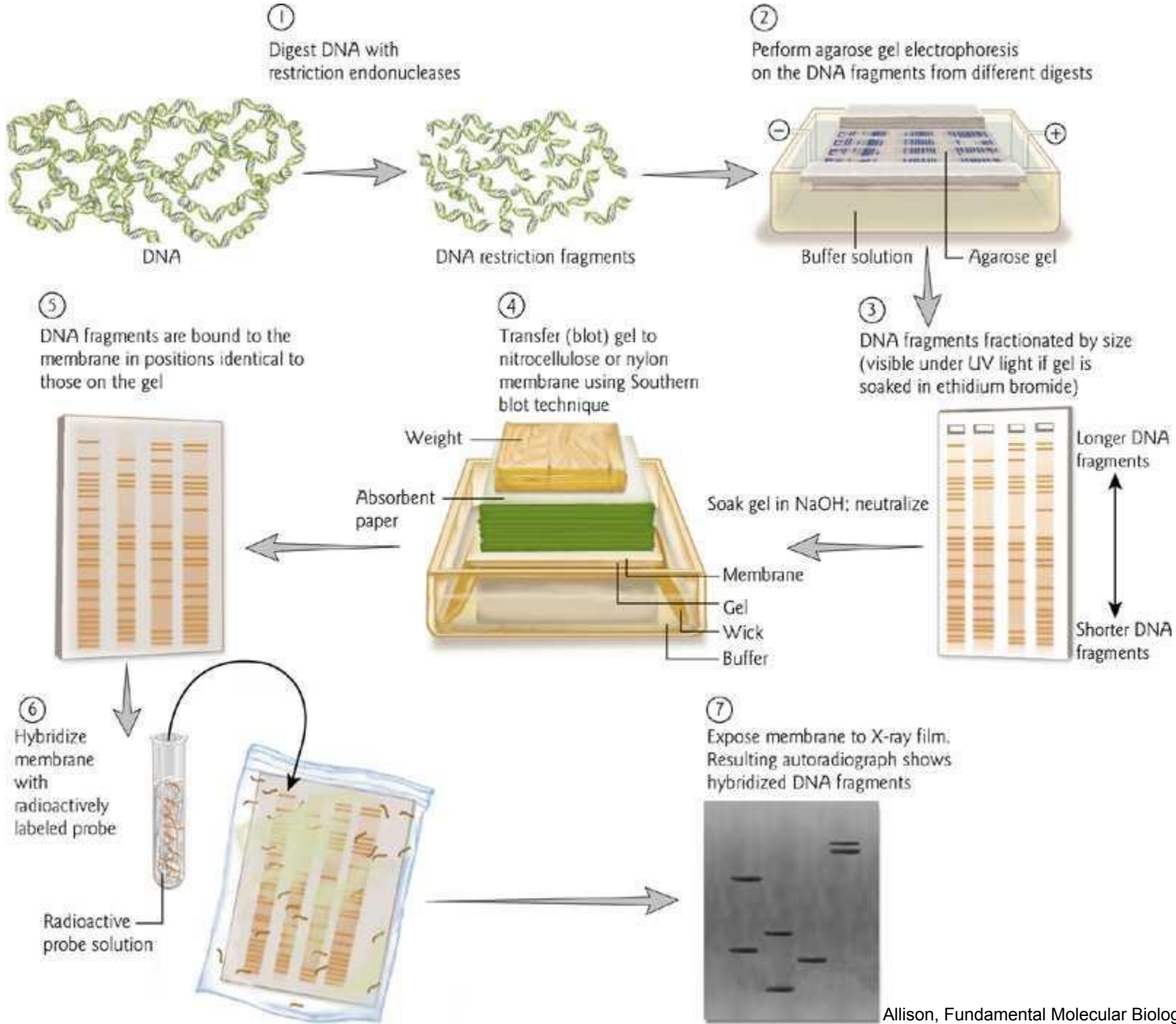


Gamma-33-ATP



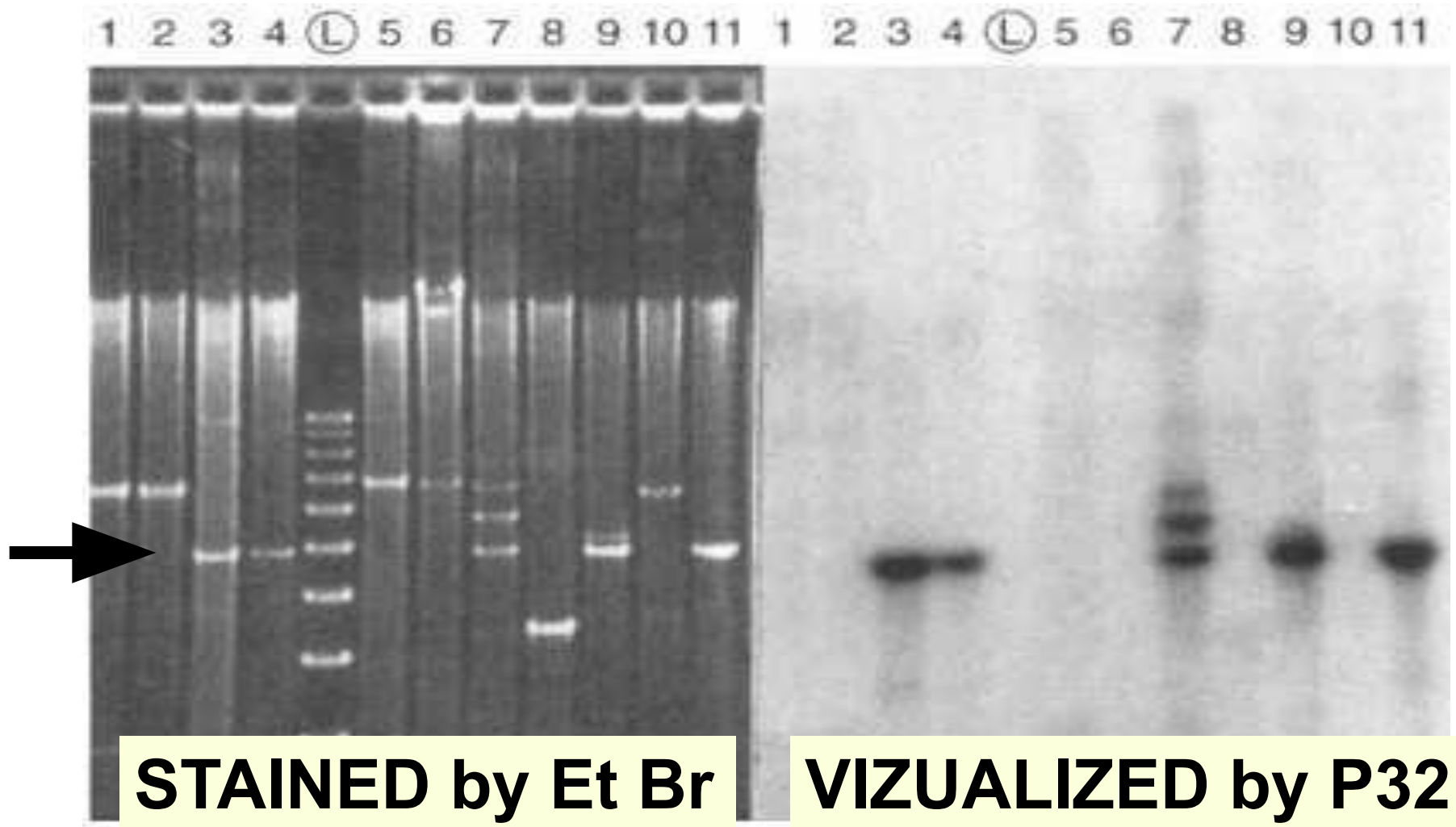
# Professor Sir Ed Southern, Whitley Professor of Biochemistry at the University of Oxford



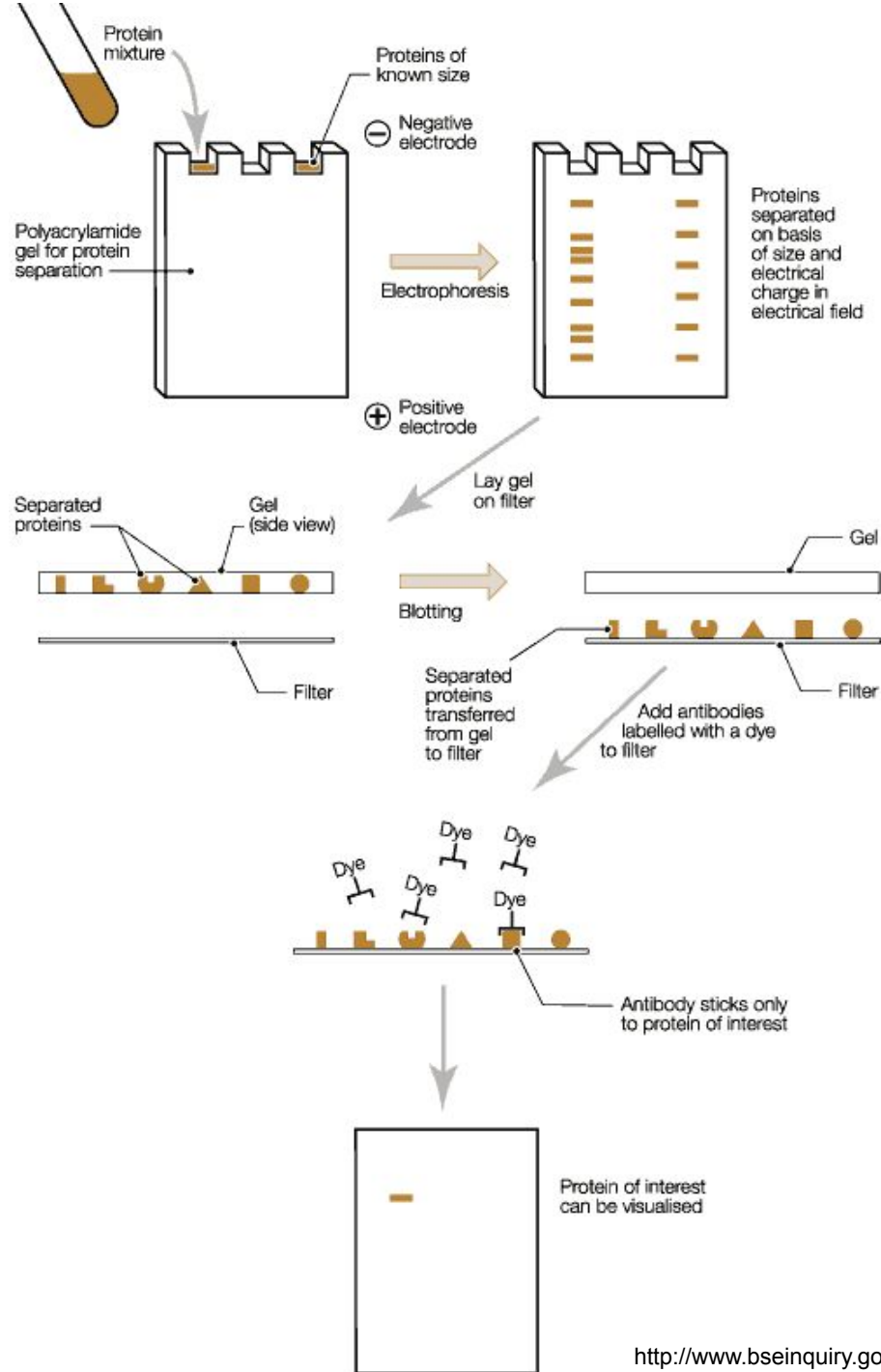




# “Real” Southern blot (DNA-DNA blot)



# Western Blot



# What different types of information can be provided by each different blot type?

Southern



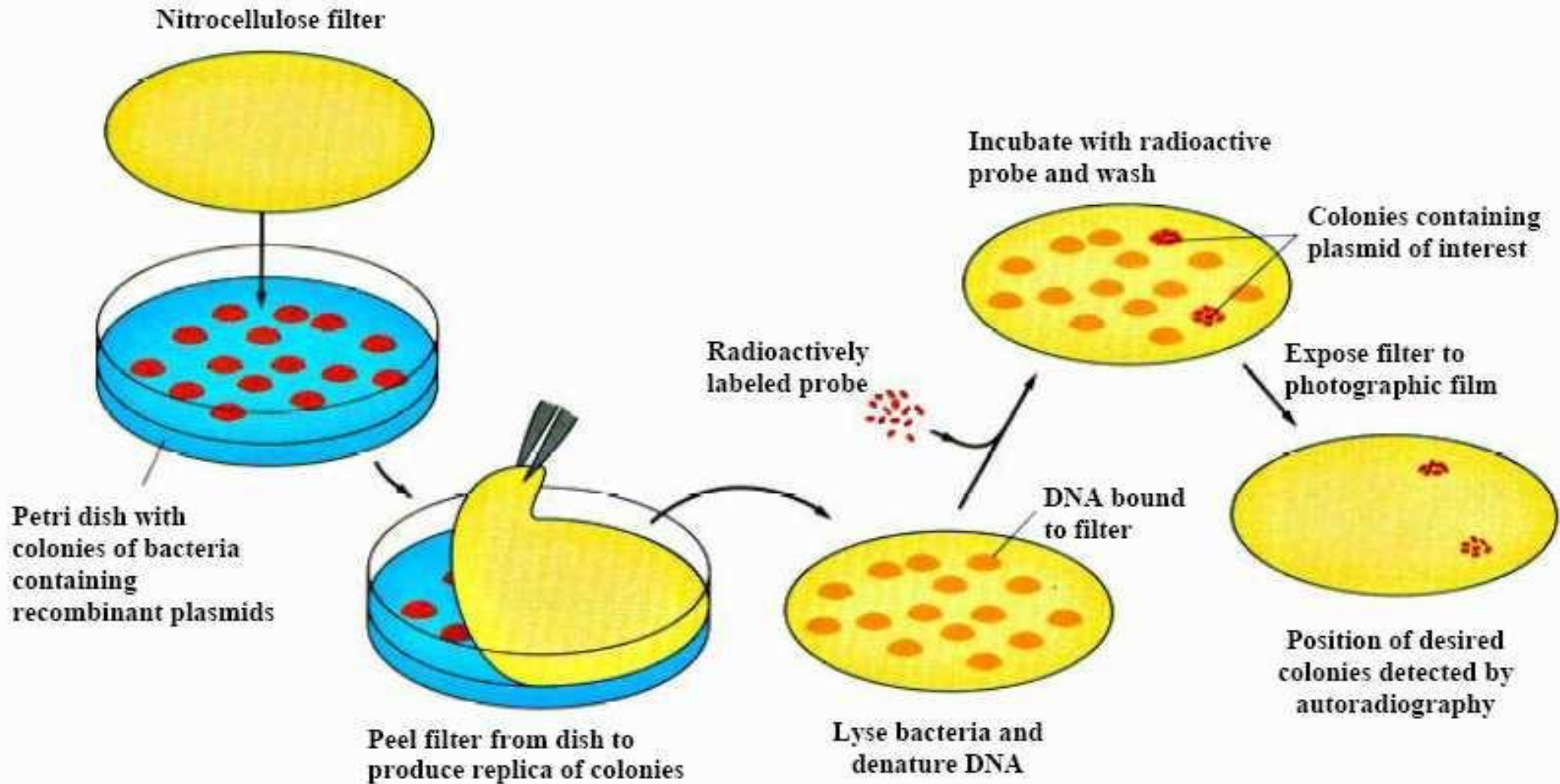
northern



western



# Colony hybridization assay for the identification of bacterial colonies carrying a particular DNA clone



## Same for arrayed clones

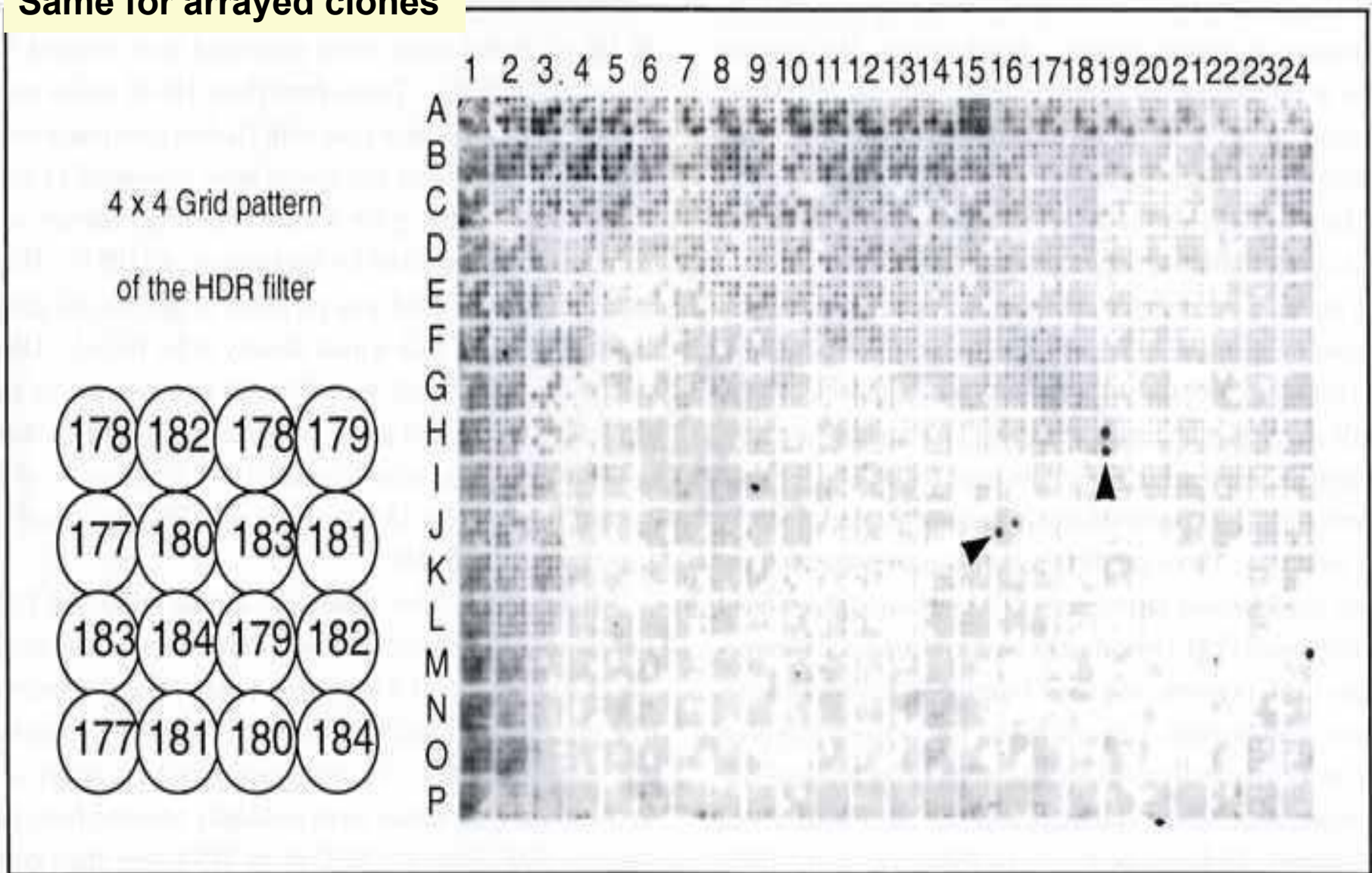
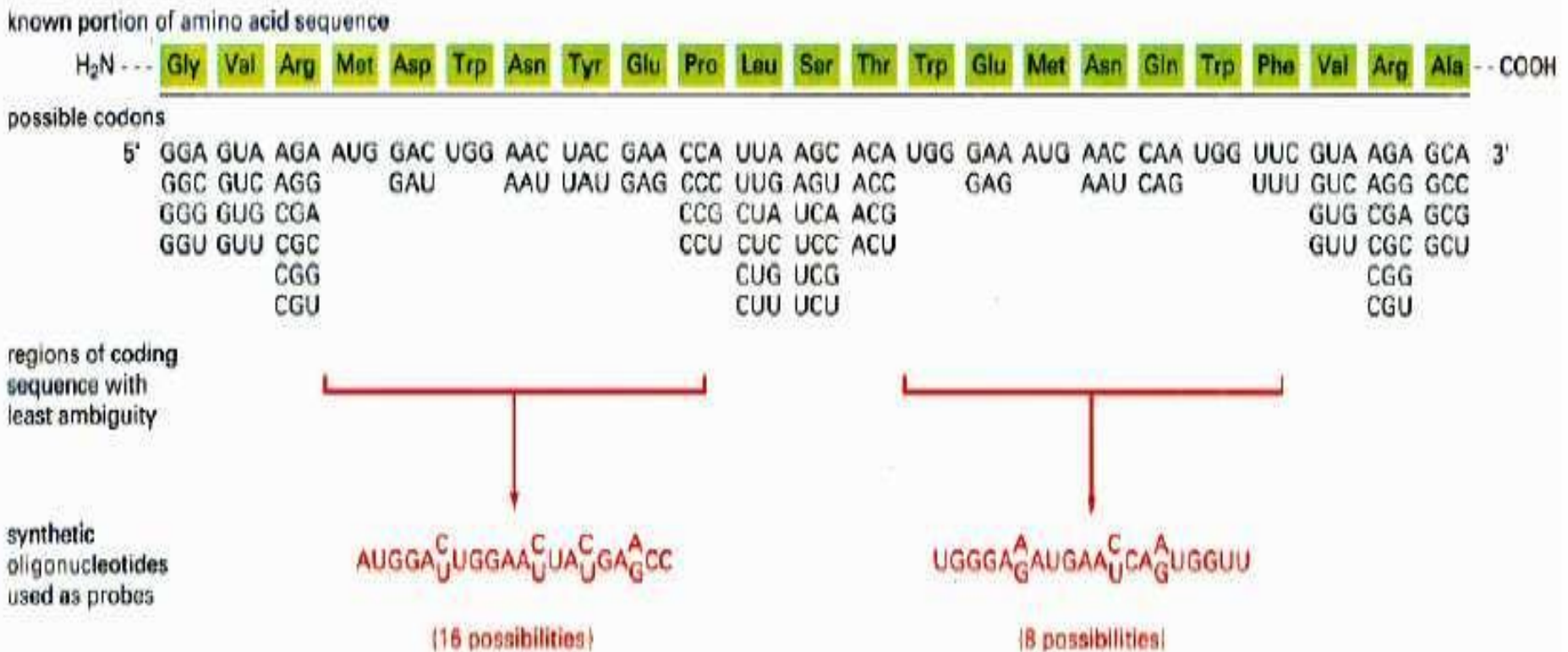


Fig. 2. Library screening by hybridization. Colony hybridization to a high density colony filter using the sequence of the single-copy gene *SLI* as a probe. 3072 BAC clones were gridded on a membrane with a 4 x 4 double offset pattern. Positive clones were detected as doublet signals.

# Design of degenerated synthetic hybridization probes (for already known proteins only)



The **degeneracy** of a primer is the number of unique sequences it corresponds to (5 in one of the examples below).

TTGATTCTTAAGA.....CCAGTGTAGC  
GGATTGTAGGCT.....ACTGAGGAGTTAT  
CTGGACTGTTGCCA.....CGGGCAGGAGTAC  
GAGATTCTGCATC.....CTAGTGTGGTAGC  
AGACTCTTAAGA.....AGAGAGGAGCTA

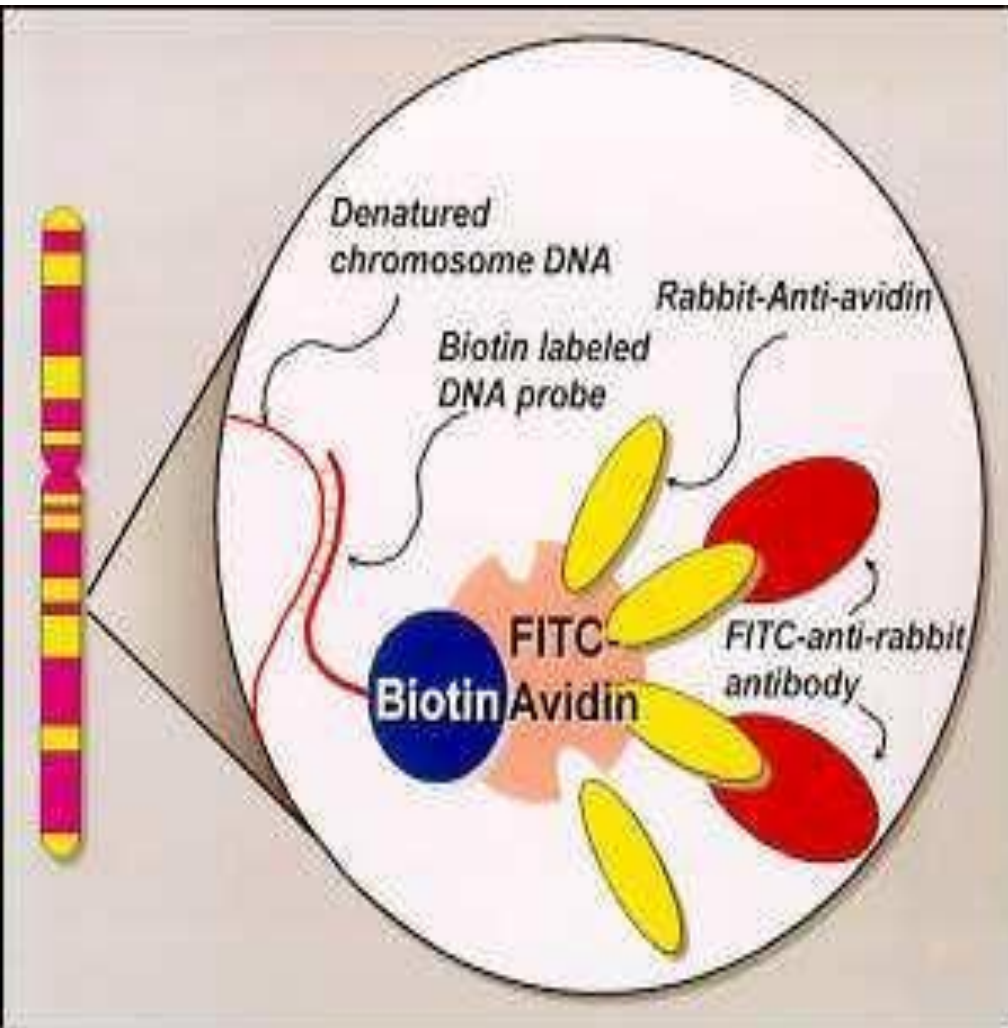
$$\begin{matrix} & C & & C \\ & | & & | \\ G & A & T & G \\ & | & & | \\ & T & & G \end{matrix}$$
  
5' primer

$$\begin{matrix} & A & & G \\ & | & & | \\ G & C & G & T \\ & | & & | \\ & T & & A \end{matrix}$$
  
3' primer

can be used when some of the related genomic sequences are unknown, or known only in a related species.

Up to  $10^{10}$  degeneracy is tolerated

# Fluorescent probe hybridization



A DNA probe, covalently bound to **biotin**, is hybridized to a denatured chromosome preparation.

An **avidin-bound fluorescent label (FITC)** is layered on top of the cells, and the **avidin-FITC binds the biotin**. The signal is amplified further by layering **rabbit anti-avidin antibody** (which binds the avidin-FITC), and then layering **FITC-labeled anti-rabbit antibody on top**.

Fluorescence will be detected only where the DNA probe has hybridized to the chromosome.



# How streptavidin/biotin binding is working?

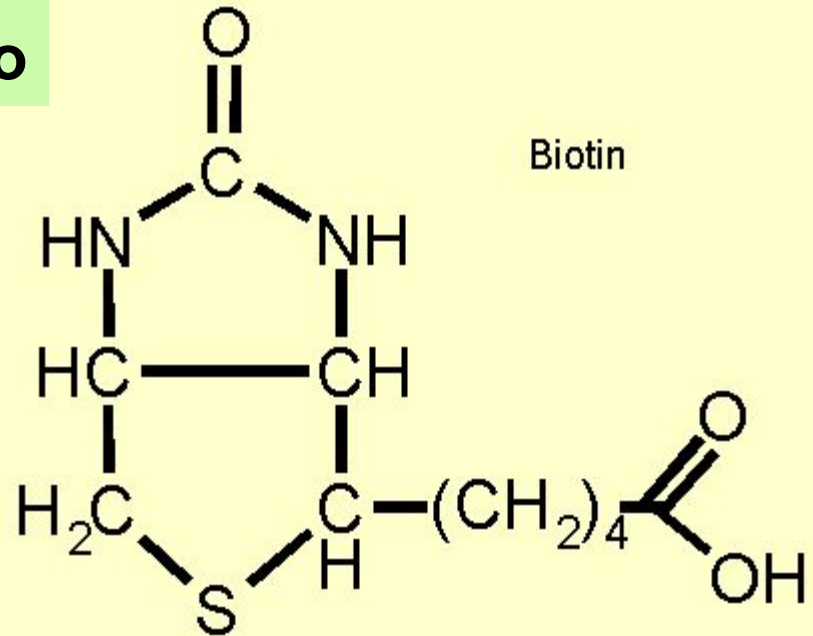
**Streptavidin is a protein.  
1 mole of SA binds 4 moles Bio**

(67 kD protein  
from *Streptococcus avidinii*)

**BIOTIN is a vitamin B  
(small thing)**

Biotin could be added to nucleotide  
and incorporated into the probe

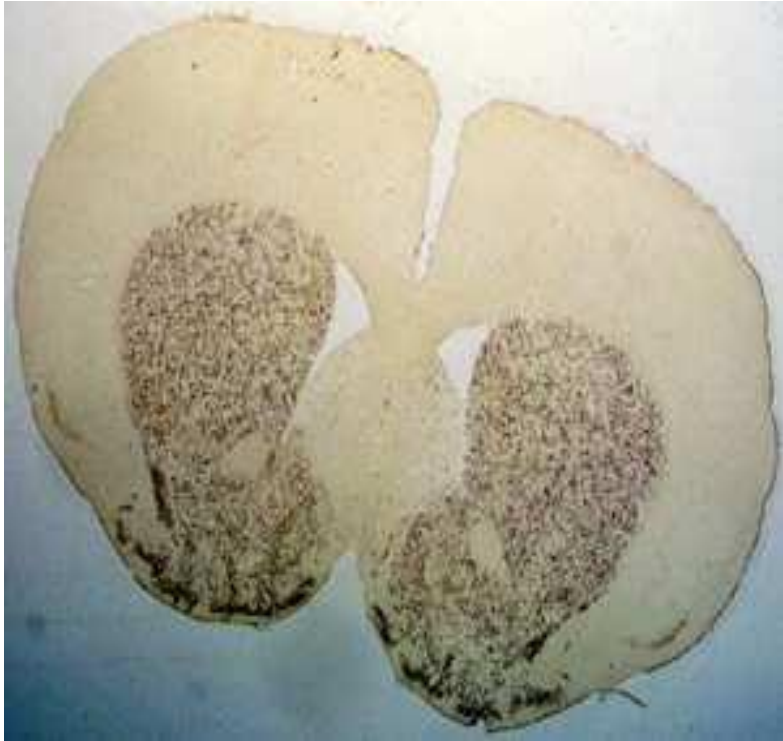
Avidin could be  
conjugated with fluorophore



Largest free energies of association  
of yet observed for noncovalent binding  
of a protein and small ligand  
in aqueous solution ( $K_{\text{assoc}} = 10^{14}$ ).  
Complex is extremely stable.

# ***IN SITU* HYBRIDIZATION**

is an imaging method to visualize mRNA expression in tissues and cells.

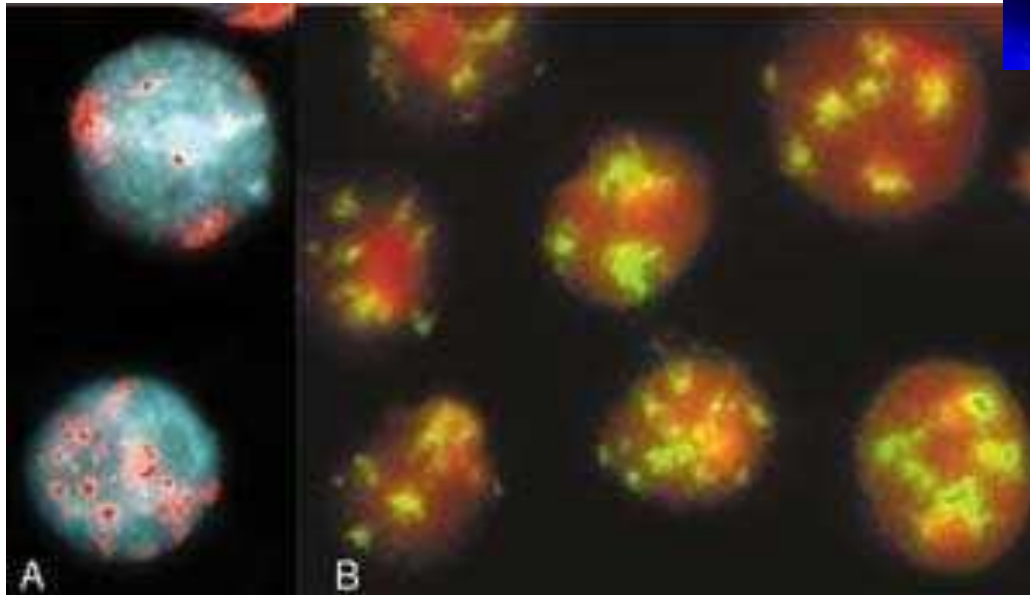
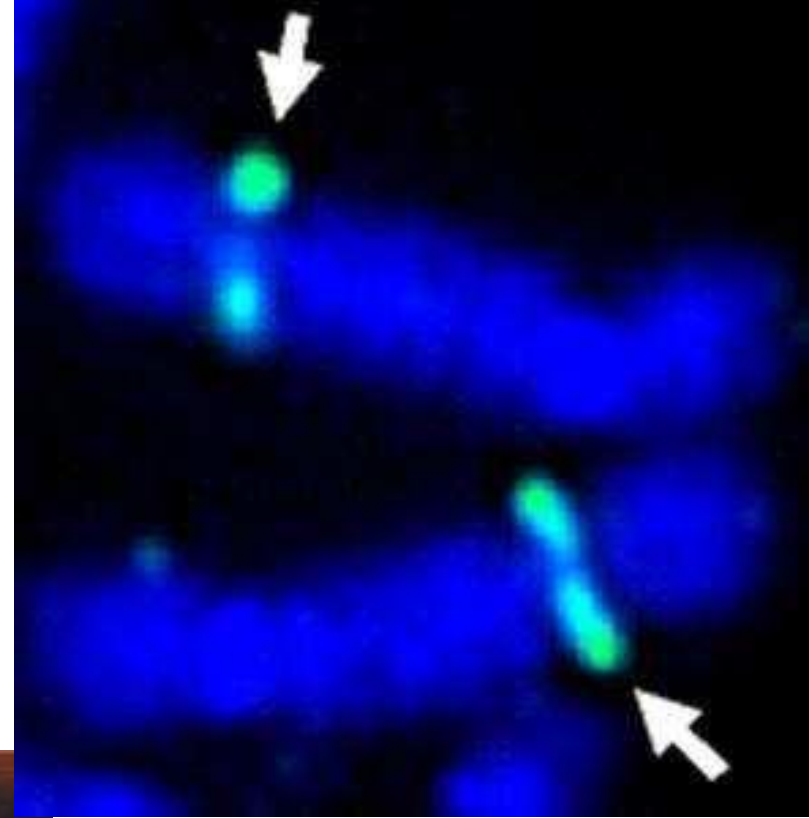


**Enkephalin gene expression  
in the mouse brain**



**The HuC transcript is expressed specifically  
in the nervous system of this E18 mouse**

**Metaphase FISH analysis  
using the BAC probe RP11-104M2  
labeled with FITC (green)  
hybridized to a normal metaphase cell  
confirms the chromosomal localization  
of the probe (gene) to 4q28.**

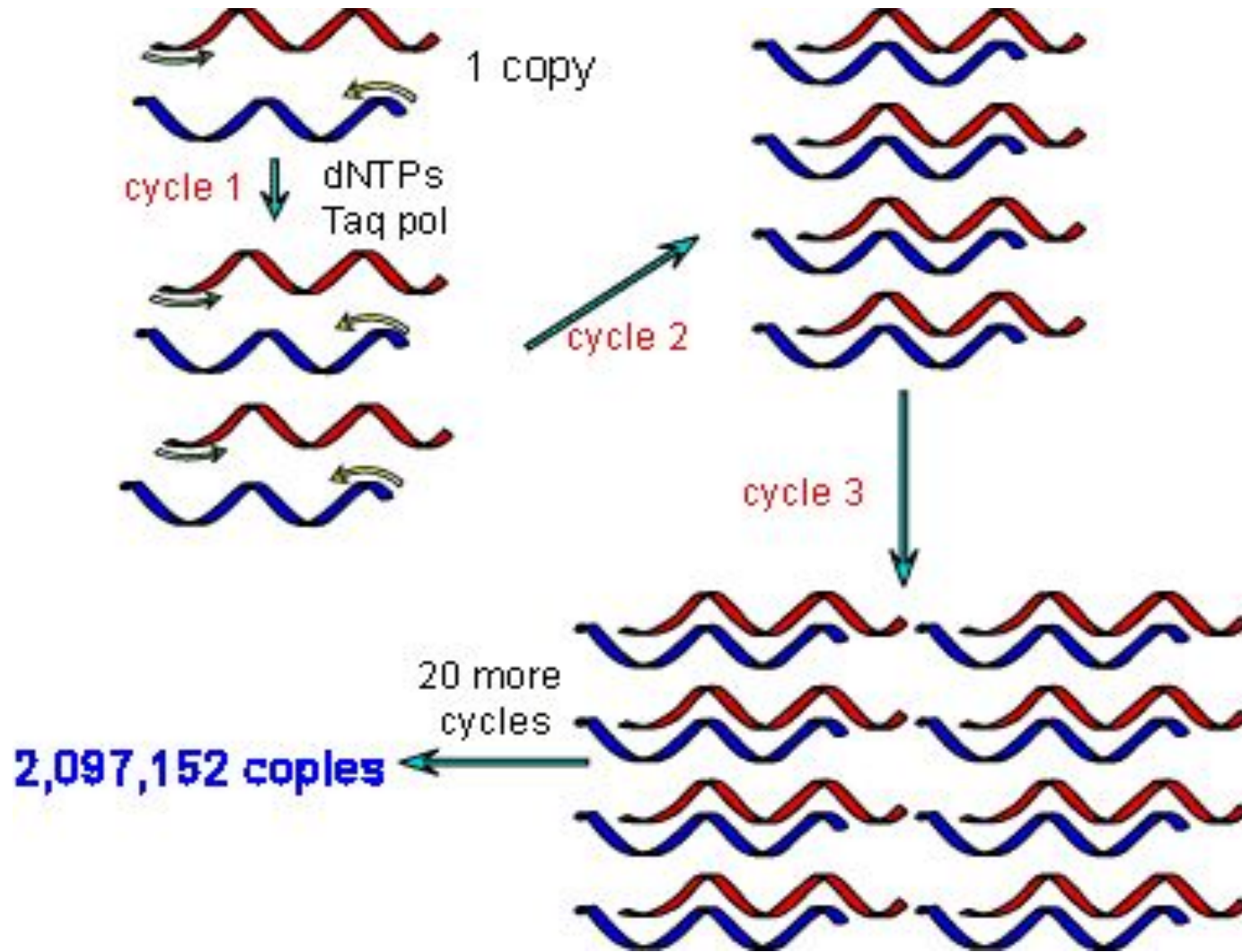


[www.infobiogen.fr/.../](http://www.infobiogen.fr/.../)

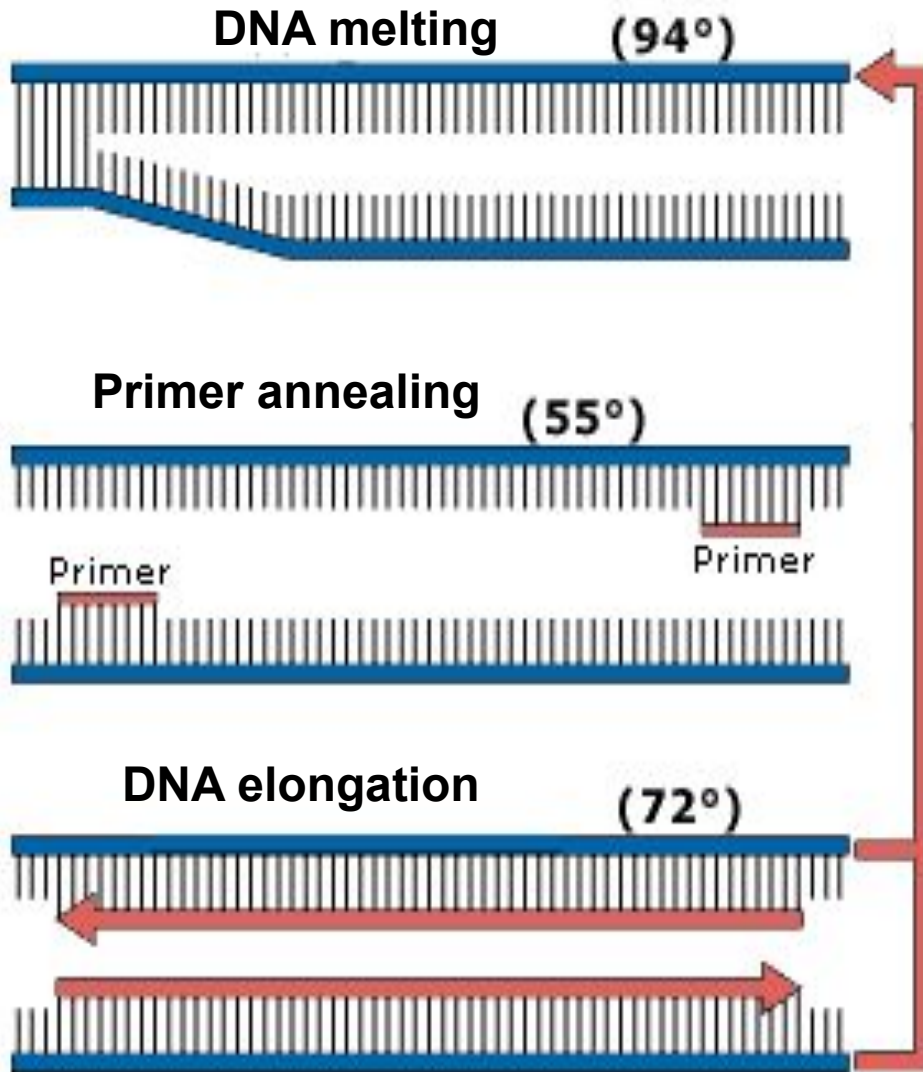
**FISH labeling of the centromeric highly repeated DNA**



# 2. Polymerase chain reaction (PCR)



# Polymerase Chain Reaction (PCR)



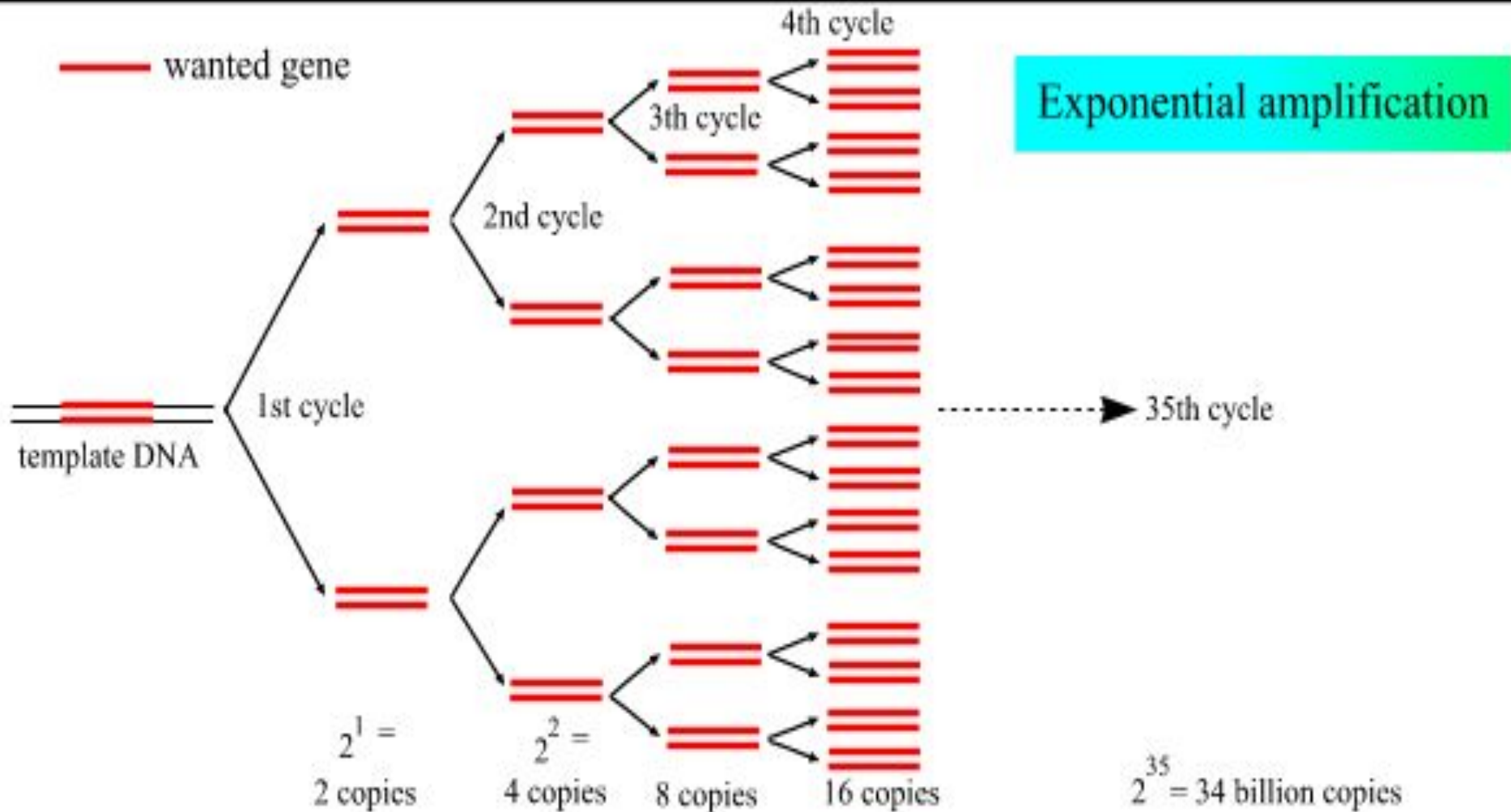
**Nobel Prize in Chemistry 1993,  
at age 48**

**Kary Mullis  
(invented PCR in 1983)**

**PhD**

***"The Cosmological Significance  
of Time Reversal,"***

# Exponential nature of PCR amplification



# PCR Reagent Components

---

- Template (DNA or RNA)
- Two primers
- dNTPs
- PCR buffer
  - KCl, Tris, MgCl<sub>2</sub>
- Thermostable DNA Polymerase (*Taq*)





# Basic Primer Design

- Usually 15 - 30 nucleotides in length
- Confirm uniqueness of primers with computer programs
- Optimize base pairing at 3' end is critical
- Random base distribution and average G+C content
  - Avoid long A+T and G+C-rich regions if possible
- Minimize internal secondary structure
- Try for equal T<sub>m</sub> for both primers

$$T_m = 69.3 + 0.41 * (\%GC) - 650/\text{length}$$

use OLIGONEW or PRIMER software

# Avoid primer dimer formation

5' CCTGCAACGTGAACTTGGCAC3'  
3' GCACGTTGCACTTGAACCGTGAAGTCT...5'

5' ...TTCATGTGTGCACTTGAACCAAGTTCAGG3'  
3' CGTGAACCTGGTTCAGTCC5'



Primer-Dimer Formation

5' CCTGCAACGTGAACTTGGCAC3' →  
← 3' CGTGAACCTGGTTCAGTCC5'

Marginally problematic primer

5' -GCTTTGTCCGATGAATCGT-3'

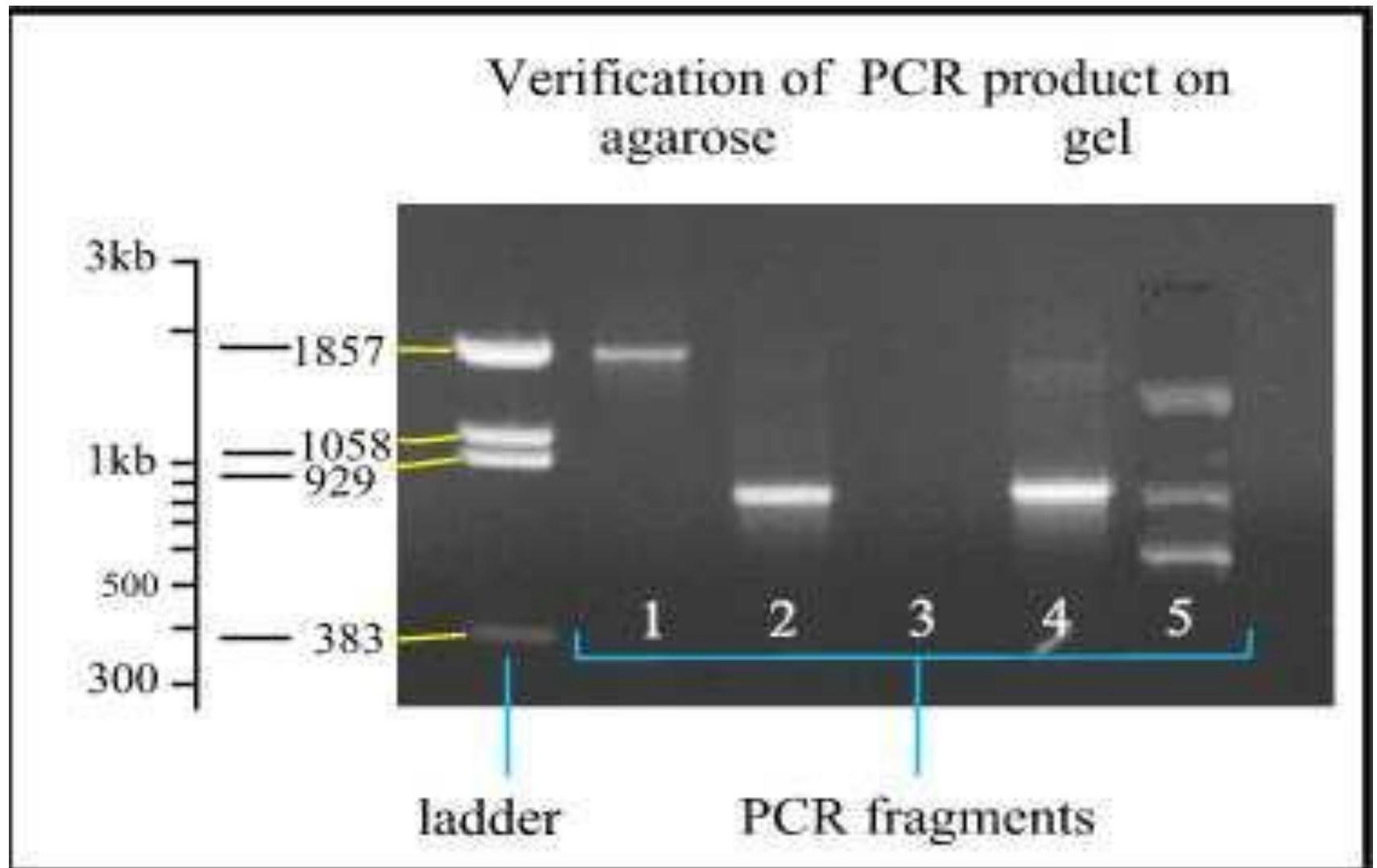
|||||  
3' -TGCTAAGTAGCCTGTTTCG-5'

# Use Software to avoid of such problems

## General Primer Picking Conditions

<u>Primer Size</u>	Min: <input type="text" value="18"/>	Opt: <input type="text" value="20"/>	Max: <input type="text" value="27"/>	
<u>Primer Tm</u>	Min: <input type="text" value="57.0"/>	Opt: <input type="text" value="60.0"/>	Max: <input type="text" value="63.0"/>	<u>Max Tm Difference:</u> <input type="text" value="100.0"/>
<u>Product Tm</u>	Min: <input type="text"/>	Opt: <input type="text"/>	Max: <input type="text"/>	
<u>Primer GC%</u>	Min: <input type="text" value="20.0"/>	Opt: <input type="text"/>	Max: <input type="text" value="80.0"/>	
<u>Max Self Complementarity:</u>	<input type="text" value="8.00"/>	<u>Max 3' Self Complementarity:</u>	<input type="text" value="3.00"/>	
<u>Max #N's:</u>	<input type="text" value="0"/>	<u>Max Poly-X:</u>	<input type="text" value="5"/>	
<u>Inside Target Penalty:</u>	<input type="text"/>	<u>Outside Target Penalty:</u>	<input type="text" value="0"/>	<u>Set Inside Target F</u>
<u>First Base Index:</u>	<input type="text" value="1"/>	<u>CG Clamp:</u>	<input type="text" value="0"/>	
<u>Salt Concentration:</u>	<input type="text" value="50.0"/>	<u>Annealing Oligo Concentration:</u>	<input type="text" value="50.0"/>	<u>(Not the concentra</u> <u>template.)</u>
<input checked="" type="checkbox"/> <u>Liberal Base</u>	<input type="checkbox"/> <u>Show Debugging Info</u>			
<input type="button" value="Pick Primers"/>	<input type="button" value="Reset Form"/>			

# Typical PCR gel (Every PCR should be gel-verified)



# Fidelity of PCR is often an issue

## Conditions Affecting Fidelity

	Increase Fidelity	Decrease Fidelity
[dNTP]	40 - 50 <b>mkM</b> all equimolar	1 - 2mM with different molarities
[Mg <sup>++</sup> ]	1:1 with dNTP	6 - 8mM Mg <sup>2+</sup> + 0.5mM Mn <sup>2+</sup>
Temp <sub>anneal</sub>	Increase Temp	Decrease Temp
[Enzyme]	Decrease [Enz]	Increase [Enz]
Extension time	Decrease Time	Increase Time
# cycles	Decrease Cycles	Increase Cycles

# Proof-reading activity enzymes are required for

## High Fidelity PCR

---

- Requires enzymes with 3'-5' nucleolytic properties
- Reaction conditions are
  - Designed to “slow down” the reaction
  - Limit reagents to favor correct base pair matches

**High Yield and High Fidelity  
are mutually exclusive**



## PCR Primer Concentrations

Normal Primer Range: 0.1 - 1.0 mM

	<u>Primer Concentration*</u>	
	<u>Lower</u>	<u>Higher</u>
Product Specificity:	Increases	Decreases
Product Yield:	Decreases	Increases

\* Compared to optimum

# Input Copy Number, Number of Cycles, and Product Mass

Human Genomic 			Plasmid 6kb 		
Copies <sup>a</sup>	Cycles	Yield <sup>b</sup>	Copies <sup>a</sup>	Cycles	Yield <sup>b</sup>
100	33	160 pg	100	33	660 ag
20,000	25	32 ng	20,000	25	130 fg
600,000	20	960 µg	600,000	20	4 pg
			1.5x10 <sup>11</sup>	2	1 mg

<sup>a</sup> Input copies

<sup>b</sup> Product mass

**If complete copies is amplified**



Cofactor required for activation of Taq polymerase

- dNTPs bind free Mg<sup>2+</sup> in a 1:1 molar ratio
  - Increases or decreases in dNTPs must be matched with equivalent changes in [Mg<sup>2+</sup>]
- Each PCR reaction has an optimum concentration of free [Mg<sup>2+</sup>]
- Different Enzymes require different [Mg<sup>2+</sup>]



## Mg<sup>++</sup> Optimization

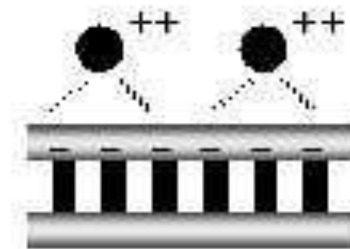
Suboptimal  
Mg<sup>2+</sup>



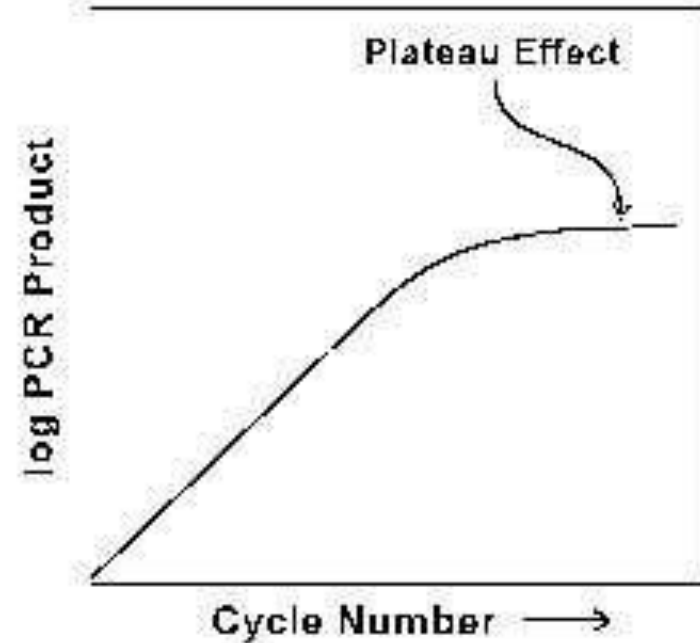
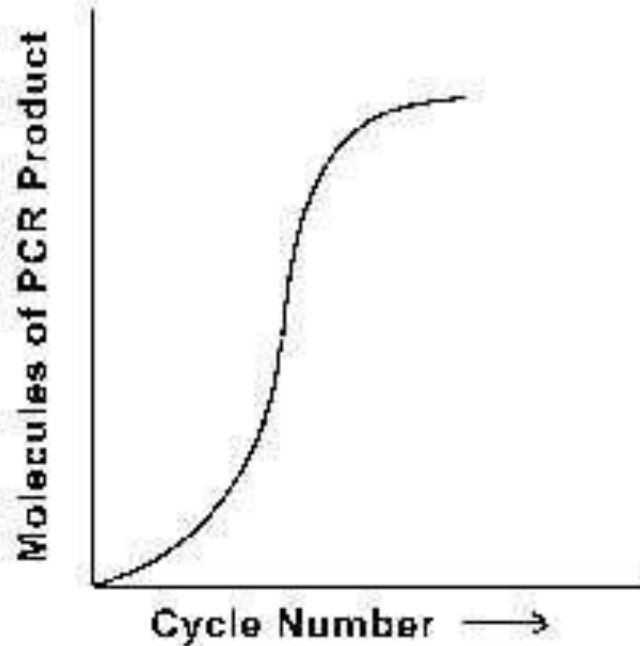
Optimal  
Mg<sup>2+</sup>



Excess  
Mg<sup>2+</sup>



# Plateau effect in PCR reaction



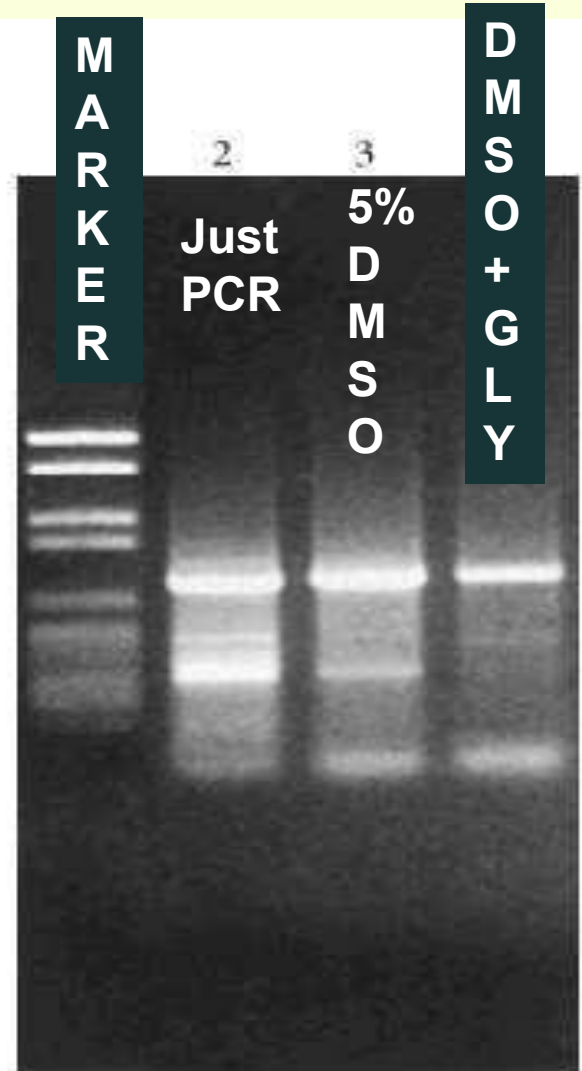
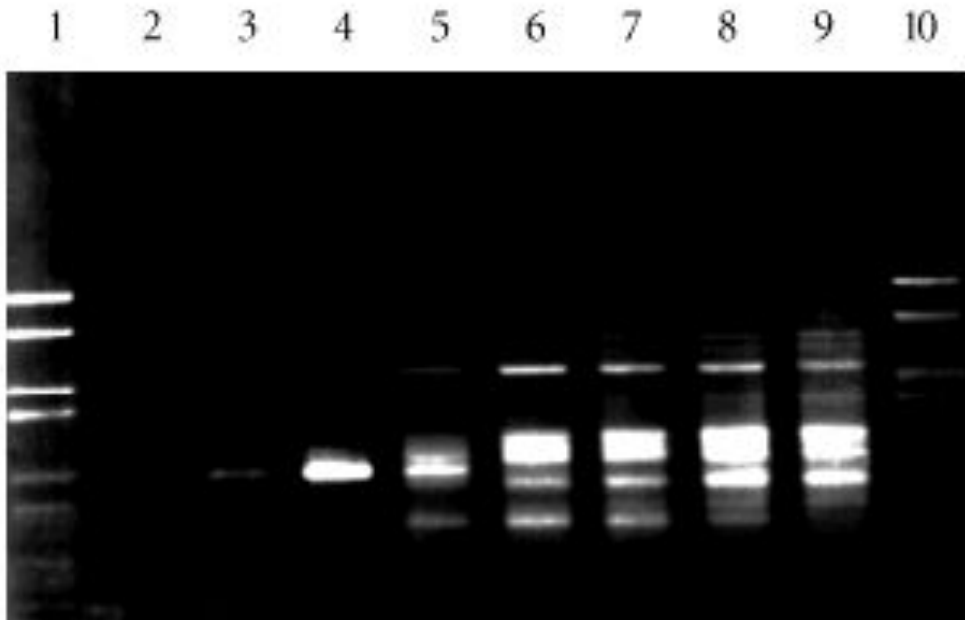
Plateau effect: The point at which product accumulation is no longer exponential (saturation), typically  $10^8$  to  $10^{12}$  molecules of product.

# Plateau effect in PCR reaction

- The point in a PCR where running more cycles does not result in a net gain of specific PCR
- Amplicons not in plateau - including nonspecific fragments - may continue to grow exponentially
- Hence, running PCRs into plateau may result in high background or smearing

# Non-specific PCR and how to improve it

Decrease in Mg concentraton



# PCR enzymes

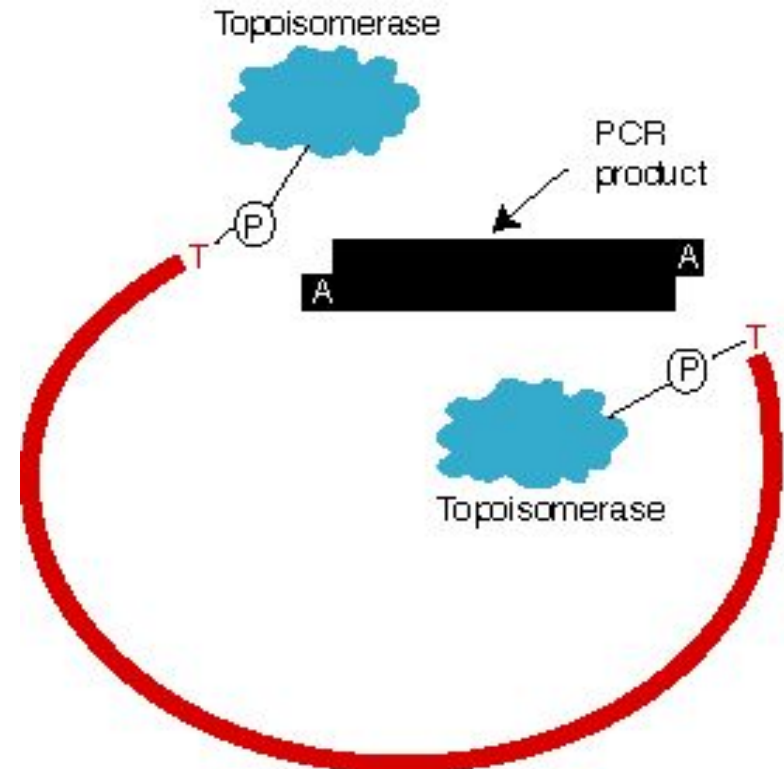
*Taq* DNA polymerase, the first enzyme used for PCR, is still the most popular.

- high processivity
- is the least expensive choice

**Half-life at 95C is 1.6 hours**

-- generates PCR products with single A overhangs on the 3'-ends  
(Suitable for TOPO-cloning)

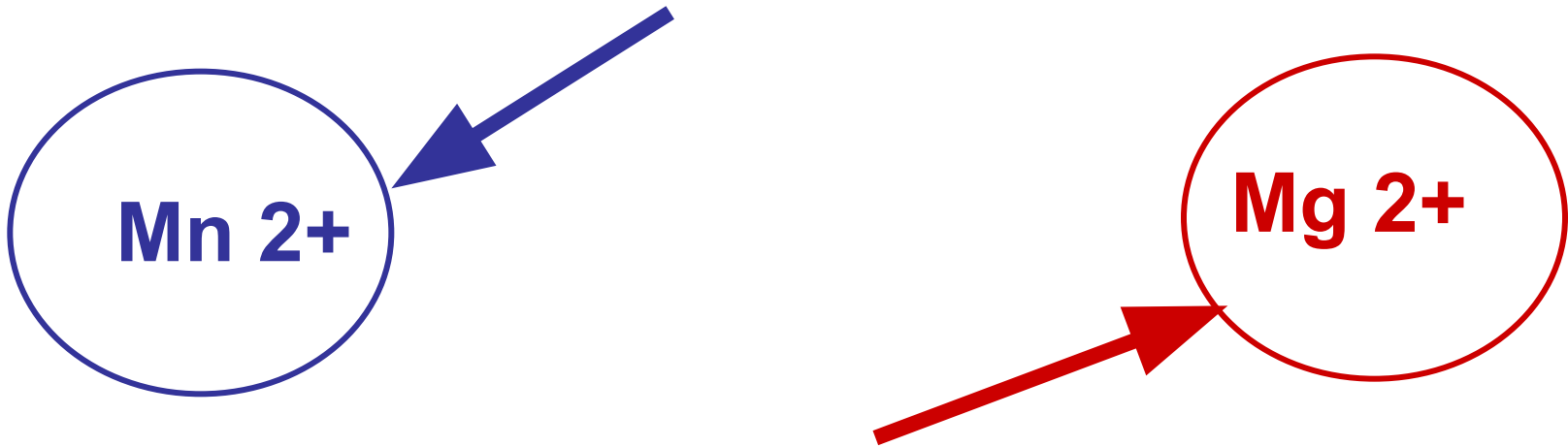
**“Topo”** cloning system (Invitrogen)



# Tth polymerase

*Thermus thermophilus* strain HB8.

RNA-dependent DNA-polymerase activity in the presence of  $Mn^{2+}$  ions.



DNA-dependent DNA-polymerase activity in the presence of  $Mg^{2+}$  ions.

**The fragment should be ideally smaller 1 kb.**

# Pfu polimerase

Proofreading or high fidelity DNA polymerases

(from *Pyrococcus furiosus*).

**approx. 1 / 2, 000,000 nucleotides before making an error.**

**In comparison Taq DNA polymerase makes an error in approx. every 1/ 10,000 nucleotides.**

**can tolerate temperatures exceeding 95°C, enabling it to PCR amplify GC-rich targets.**

**more expensive**



# Pol Vent (From *Thermococcus litoralis*)

also known as **Tli** polymerase

**Very thermostable: Half-life at 95 C is approximately 7 hours**

**3'->5' exonuclease activity presents**

**Vent error rate is intermediate between Taq and Pfu.**

**2-5 x 10<sup>-5</sup> errors/bp**

## **Other polymerases:**

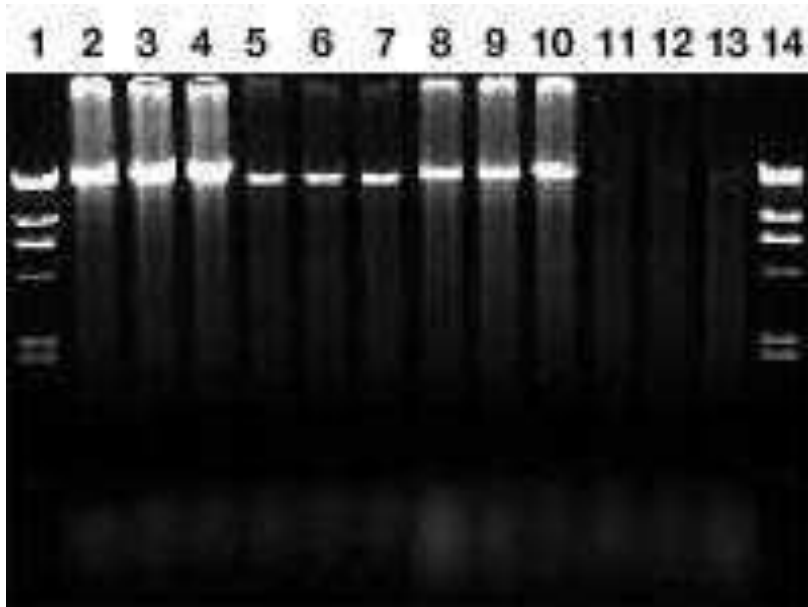
**Deep Vent** (Pyrococcus species GB-D) (New England Biolabs)  
New England Biolabs claims fidelity is equal to or greater than that of Vent.

**Replinase** (Thermus flavis) □ 1.03 x 10<sup>-4</sup> errors/base

# Long-Range PCR

## Use of two polymerases:

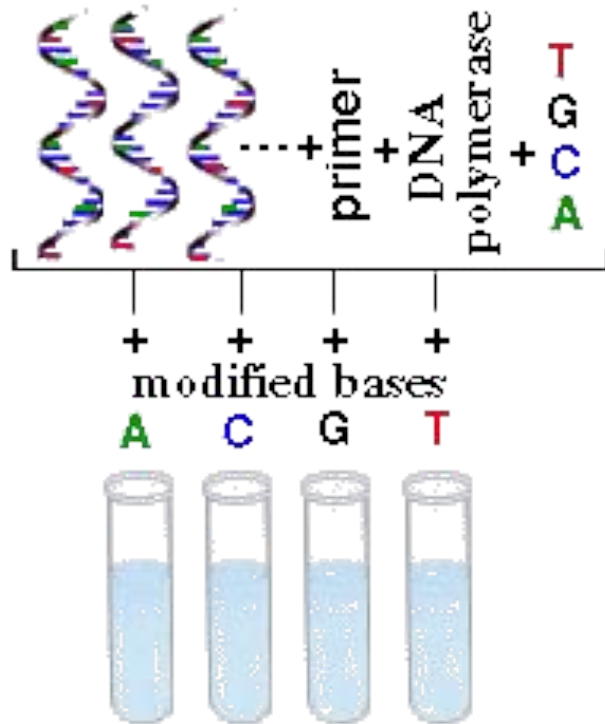
a non-proofreading polymerase **Taq** is the main polymerase in the reaction, a proofreading polymerase (3' to 5' exo) **Pwo** is present at a lower concentration.



22-24 kb PCR products are achieved on Qiagen and Eppendorf PCR mixes

**Taq+ Pwo**  
(*Pyrococcus woesei*) ;  
Pwo is very stable,  
2 hrs at 100 C

# 3. SEQUENCING: (Sanger method)

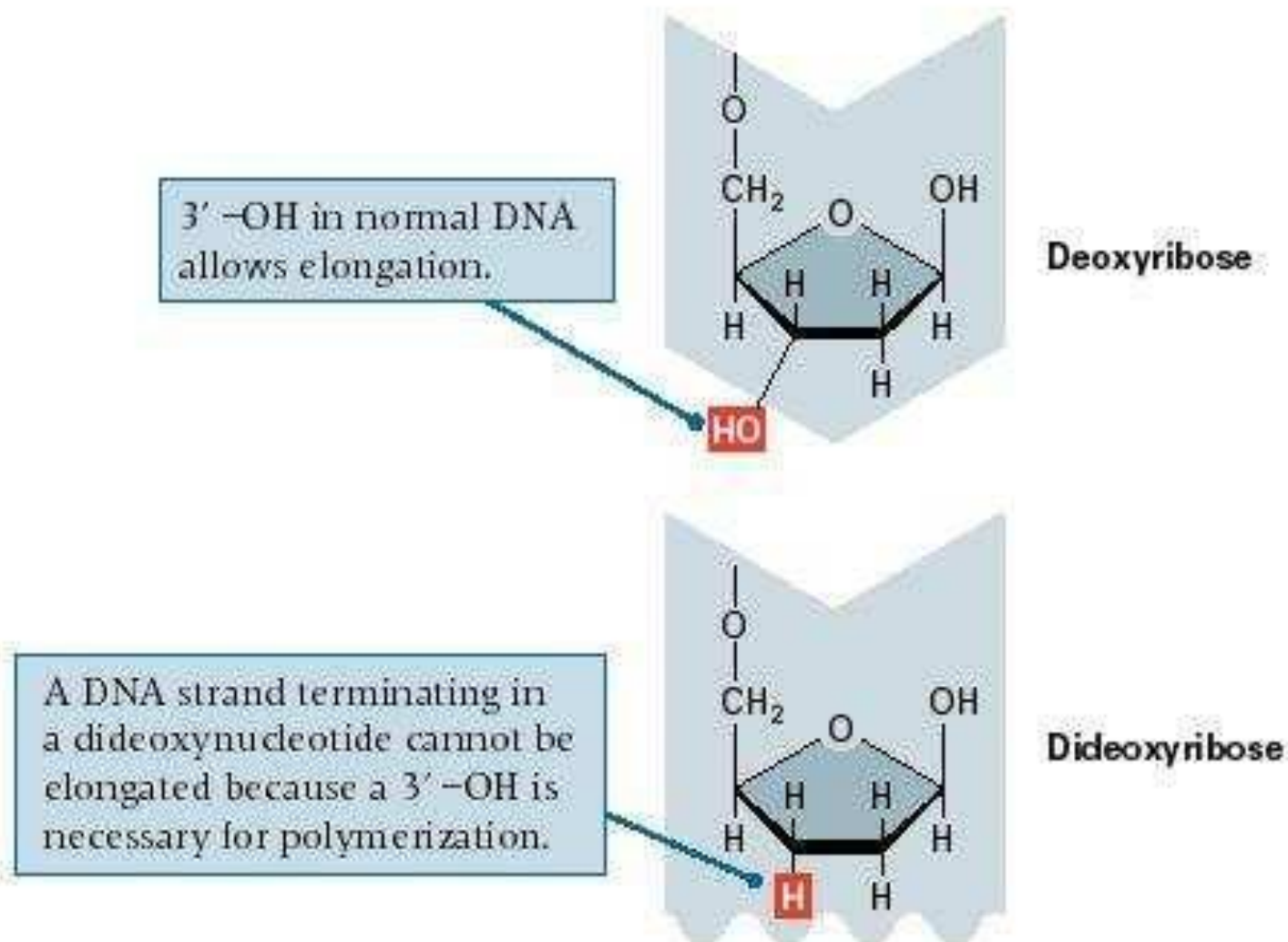


Sequencing of DNA  
by the Sanger Method

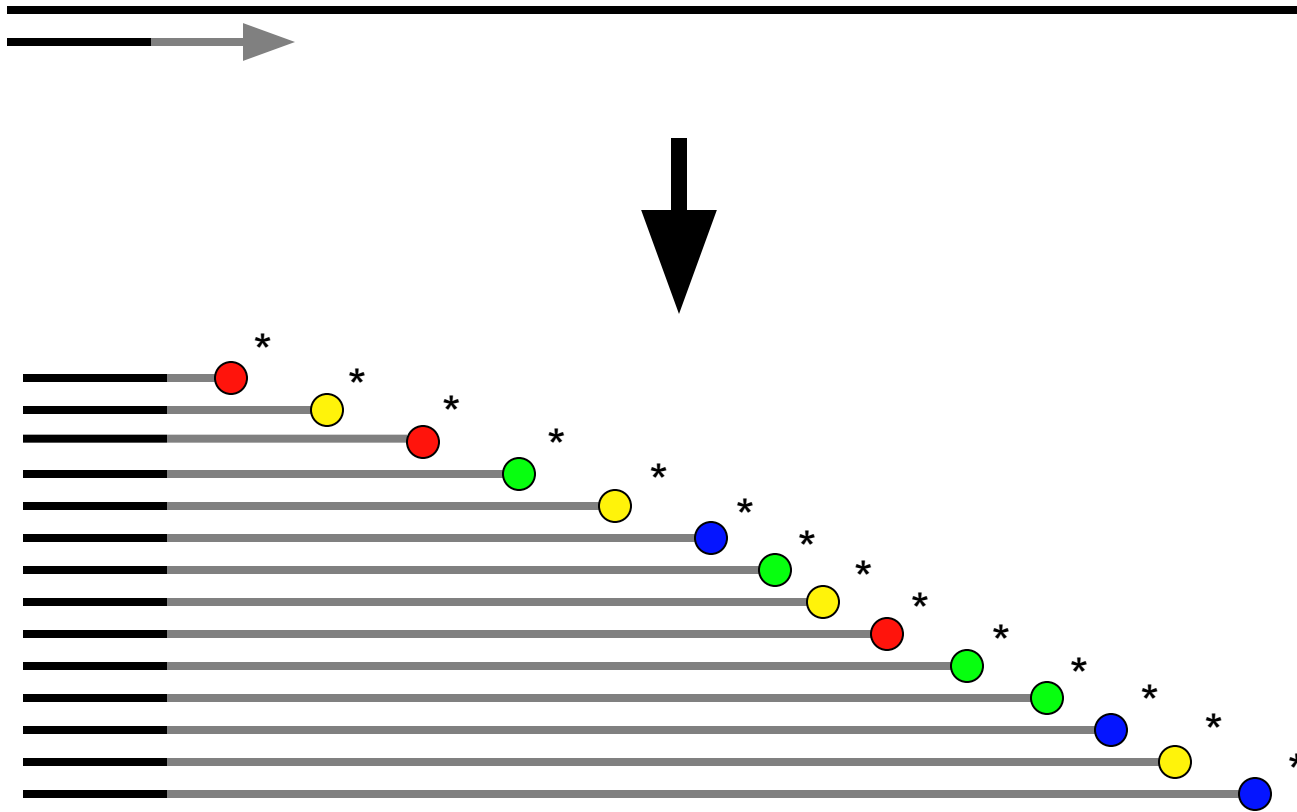


Frederick Sanger  
(Nobel prize 1980 with Paul Berg and Walter Gilbert)

# Dideoxynucleotide blocks chain elongation

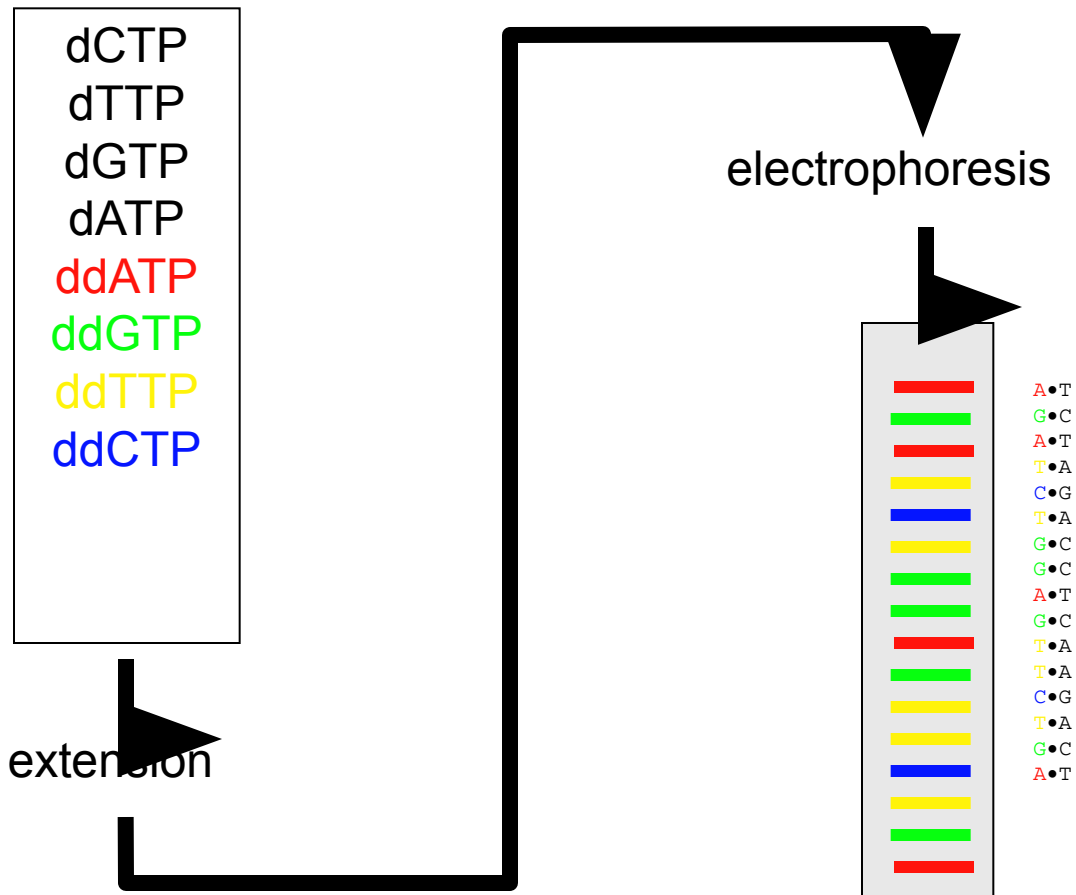


# DNA sequencing: chemistry, with terminator dyes

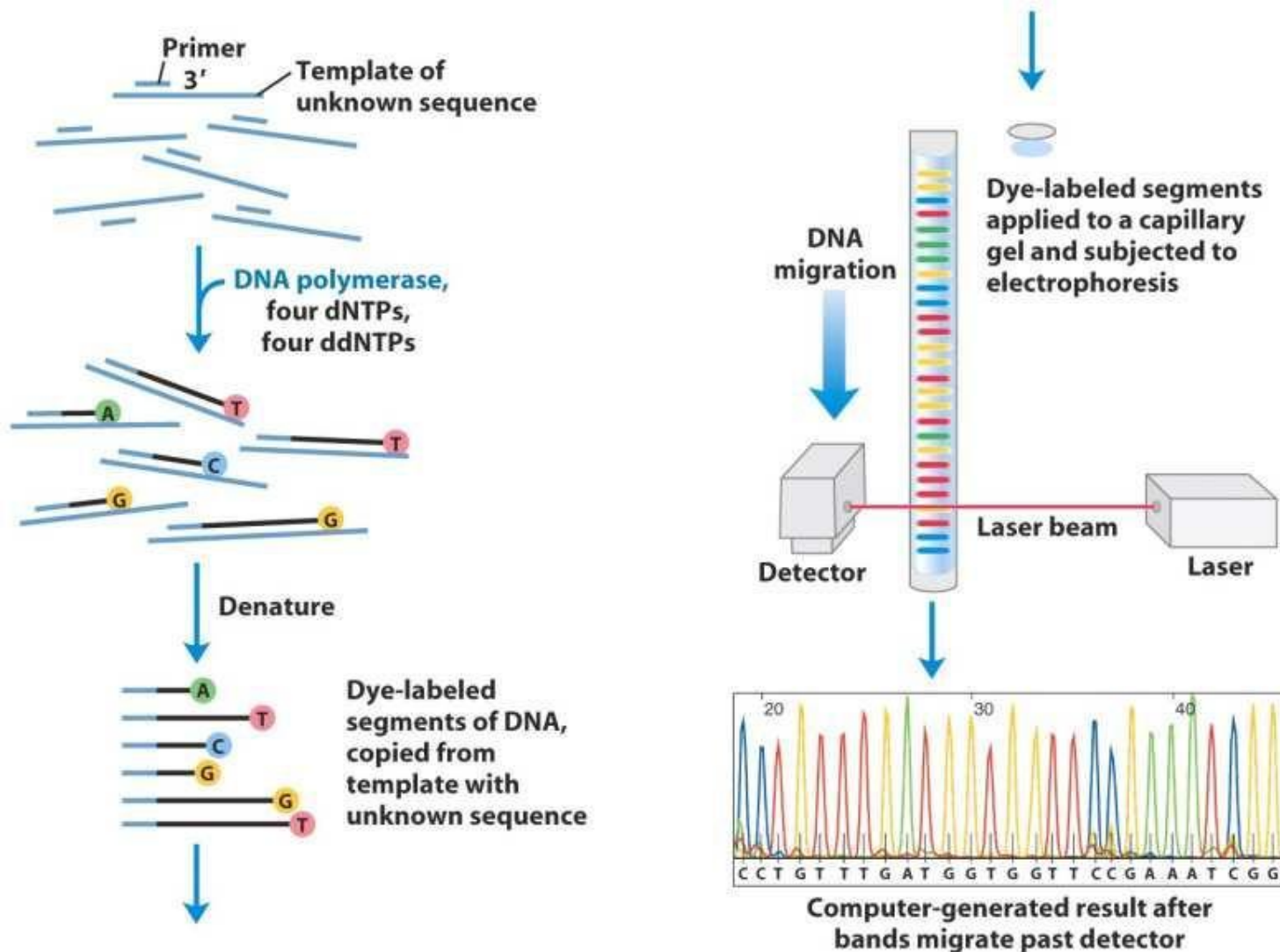


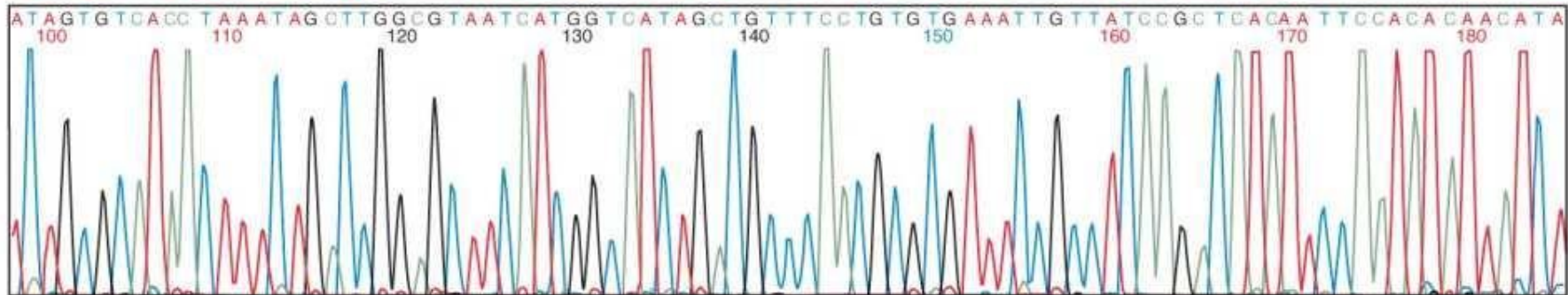
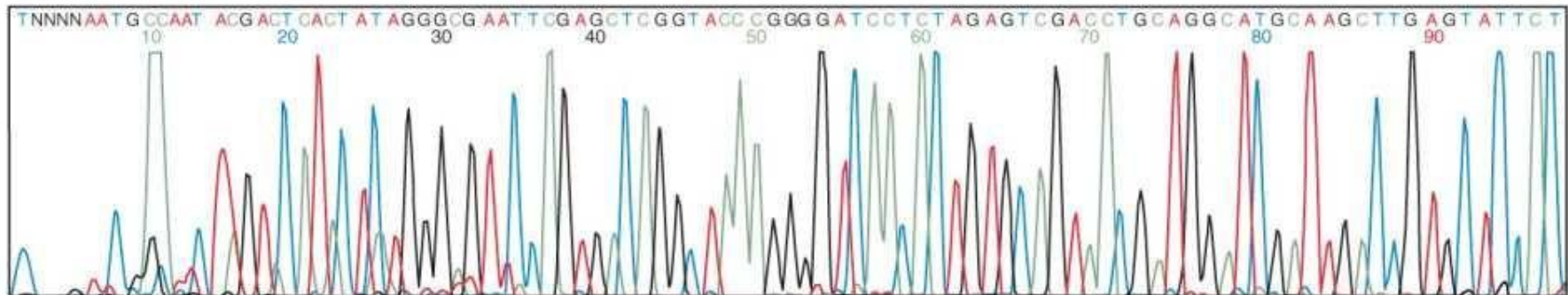
# Semi-automated fluorescent DNA sequencing:

template + polymerase +



# Cycle Sequencing - PCR





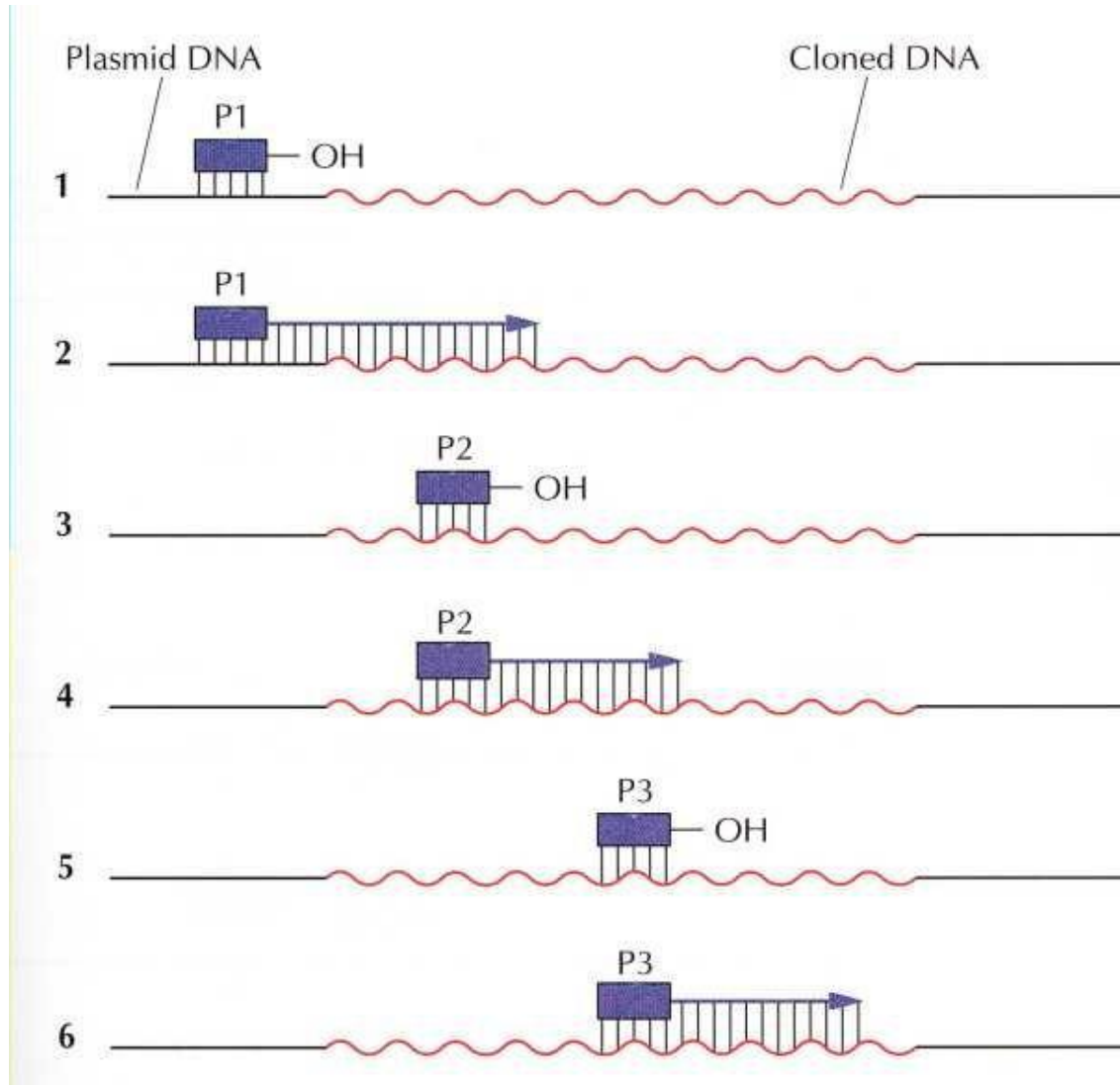


Applied Biosystems Inc., have designed an automated method that combines the PCR and actual sequencing

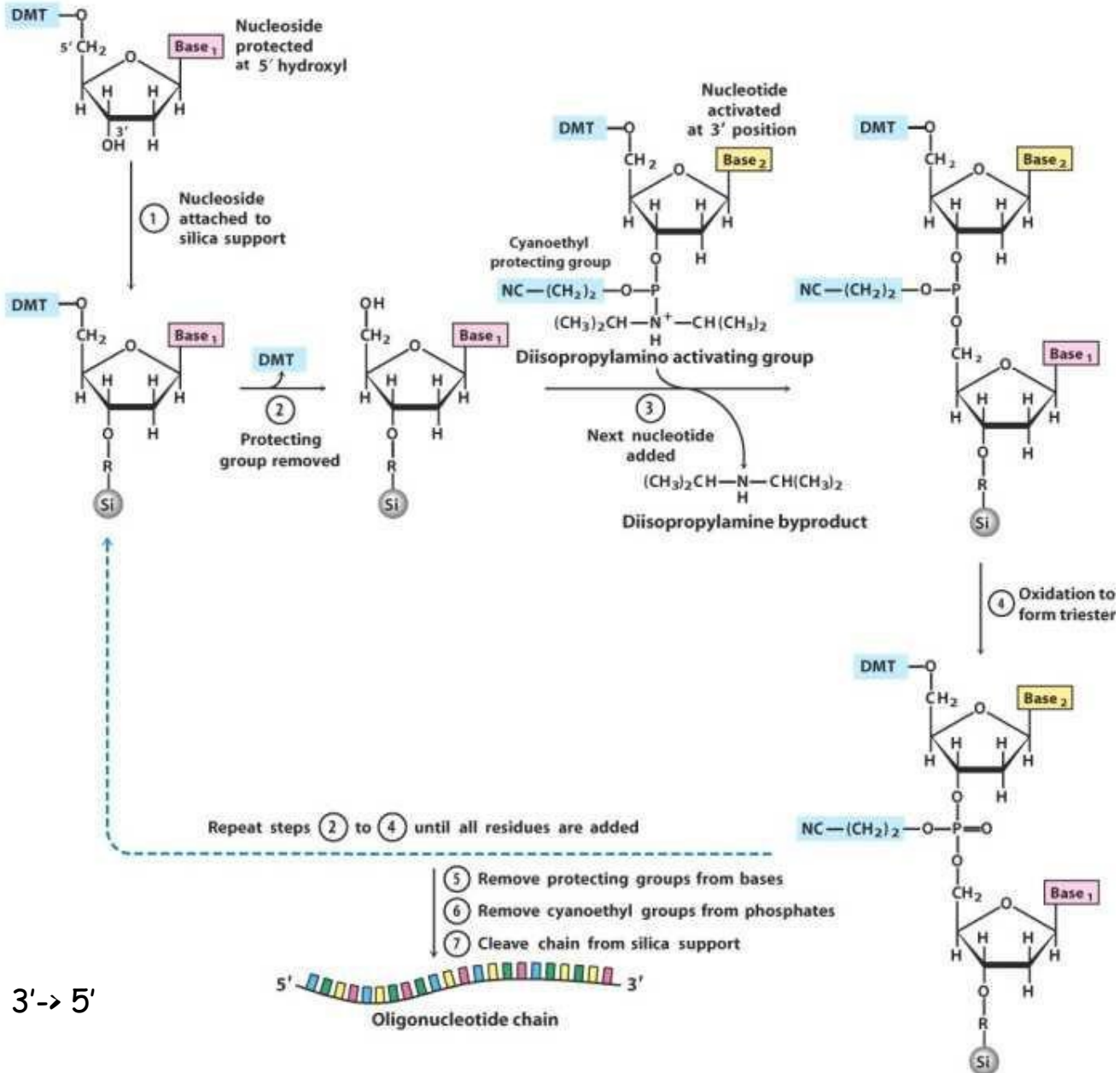


<<http://www.utmb.edu/proch/servo3.htm>>

# DNA sequencing by primer walking



# Chemical synthesis of DNA



Chain grows: 3' → 5'

# General consideration about Gene Expression

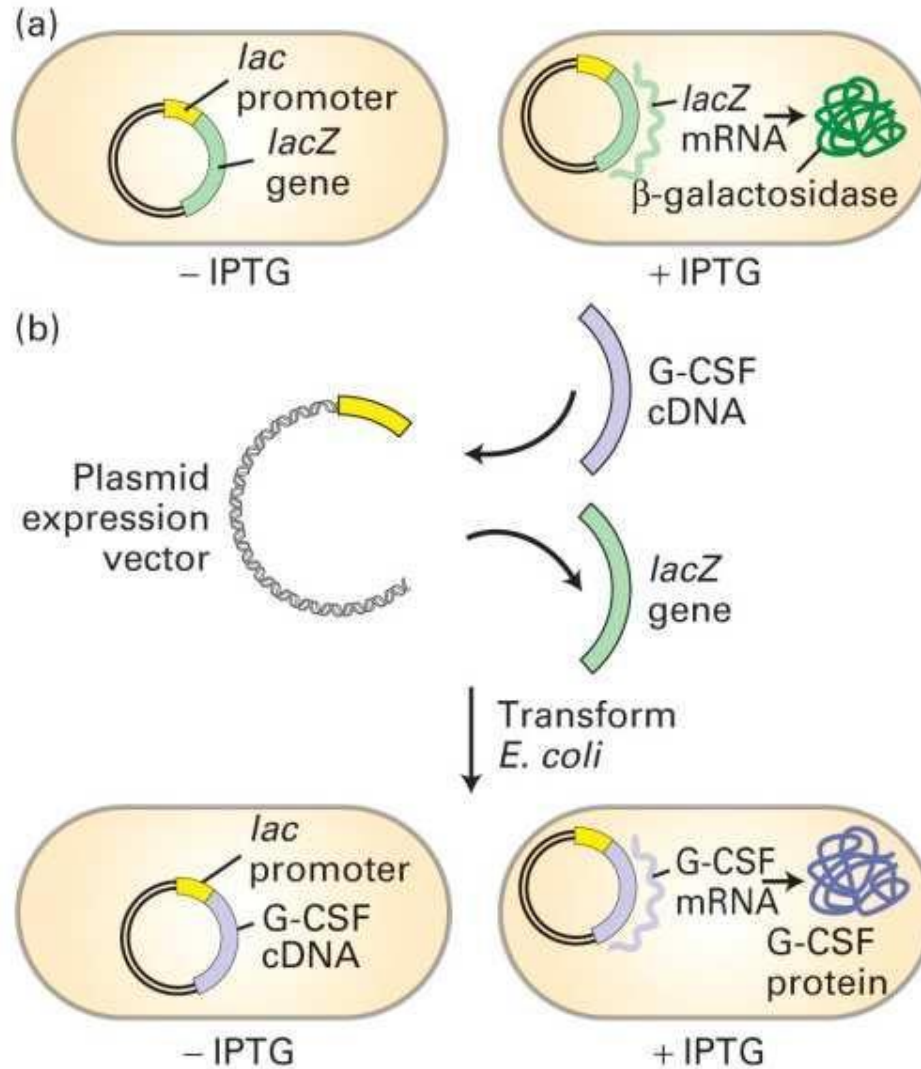
- Expression Host -> Expression System
- Promoter system -> expression vector
- Properties of product -> stability
- Production level

# Comparison of expression systems

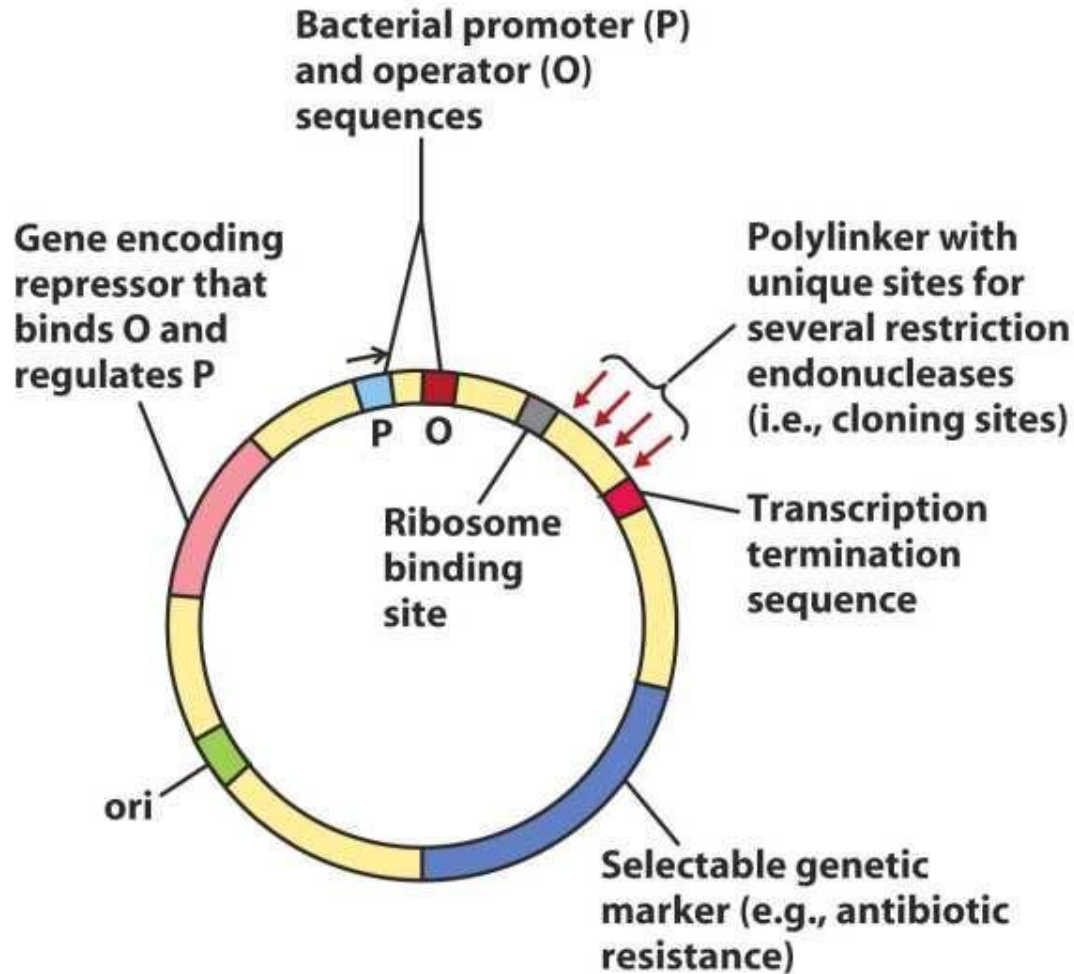
**Table 1**  
Comparison of Expression Systems

Desired characteristics	Expression system			
	Bacteria	Yeast	Insect	Mammalian cell culture
Cell growth	Rapid	Rapid	Slow	Slow
Complexity of growth medium	Minimum	Minimum	Complex	Complex
Cost of growth medium	Low	Low	High	High
Expression level	High	Low to high	Low to high	Low to moderate
Extracellular expression	Secretion to periplasm	Secretion to medium	Secretion to medium	Secretion to medium
Posttranslational modifications				
Protein folding	Refolding usually required	Refolding may be required	Proper folding	Proper folding
N-linked glycosylation	None	High mannose	Simple, no sialic acid	Complex
O-linked glycosylation	No	Yes	Yes	Yes
Phosphorylation	No	Yes	Yes	Yes
Acetylation	No	Yes	Yes	Yes
Acylation	No	Yes	Yes	Yes
$\gamma$ -Carboxylation	No	No	No	Yes

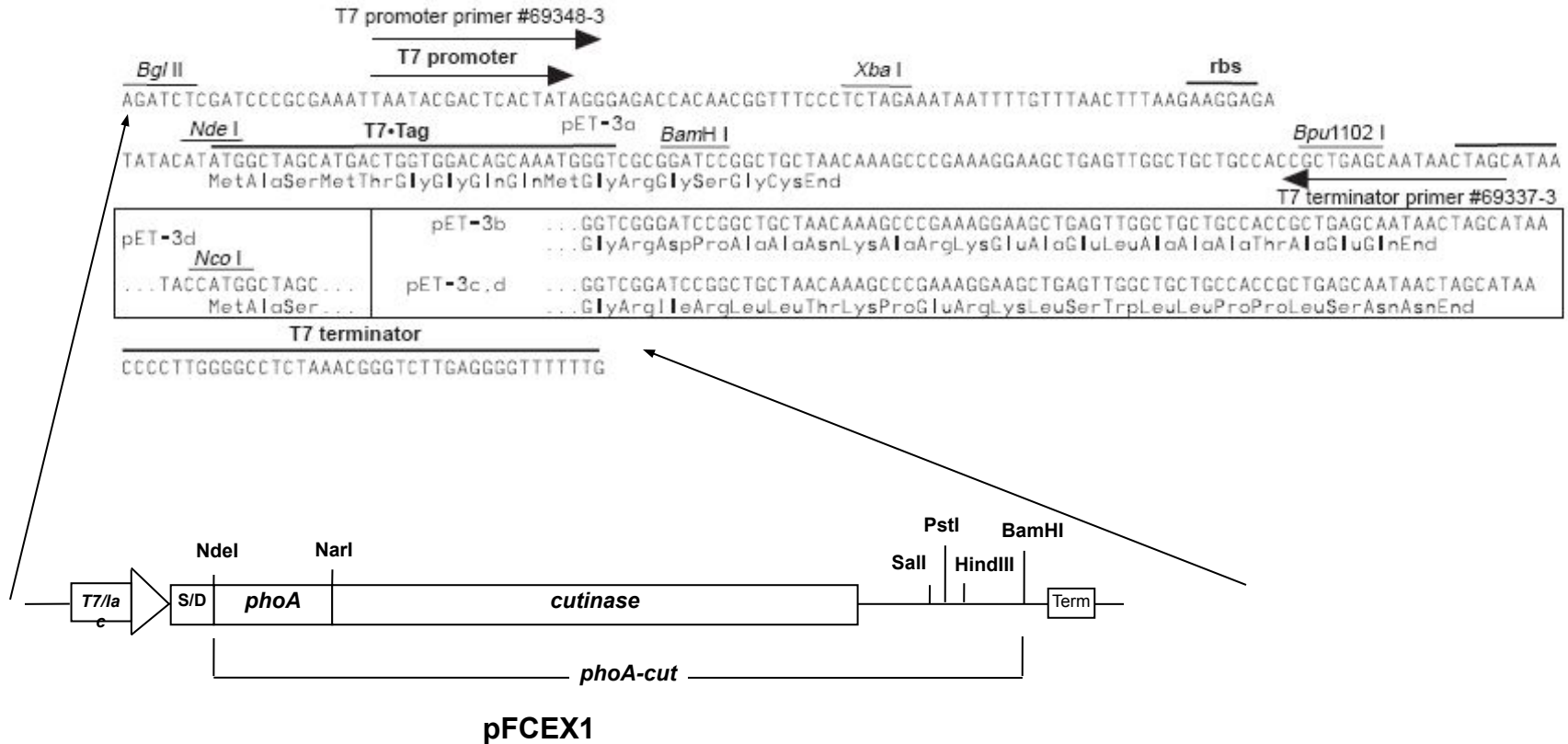
# Use of *Lac* promoter (*pLac*) for expression of foreign protein



# Prokaryotic Expression vector

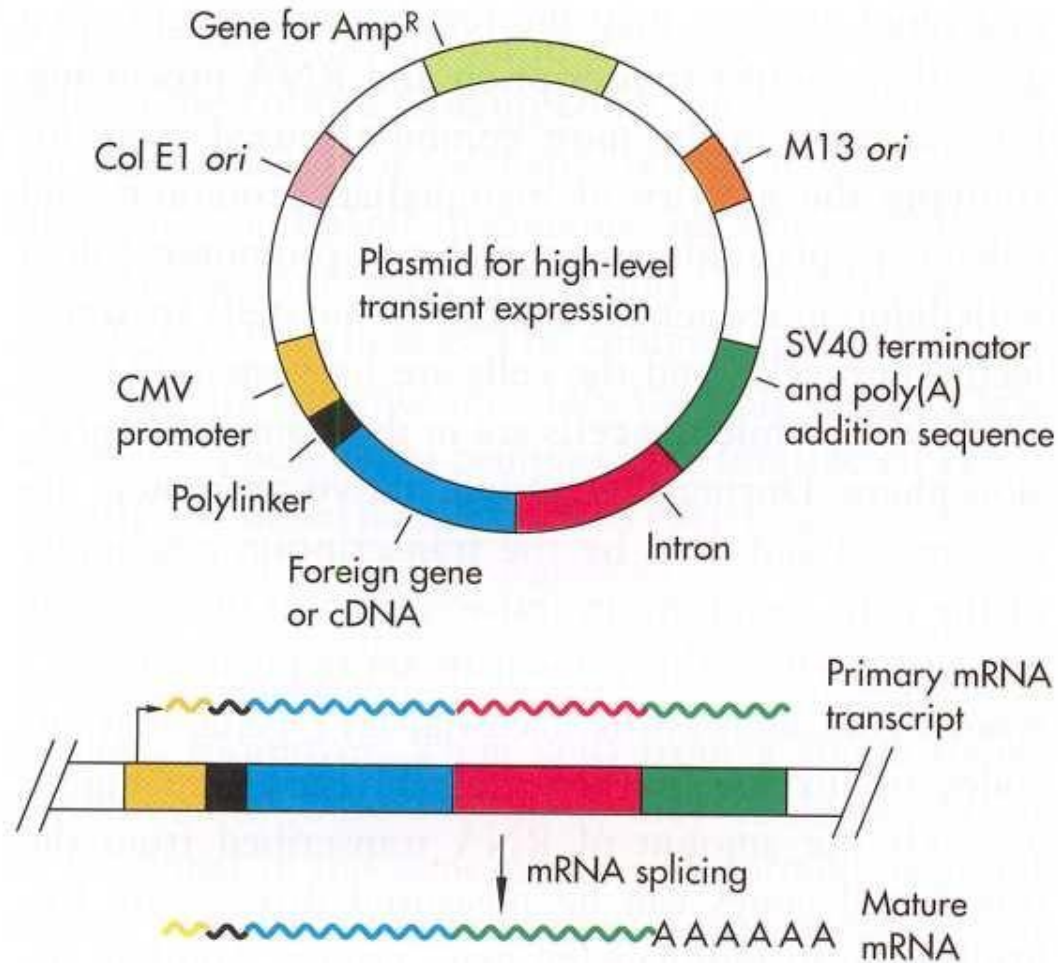


# Prokaryotic Expression vector

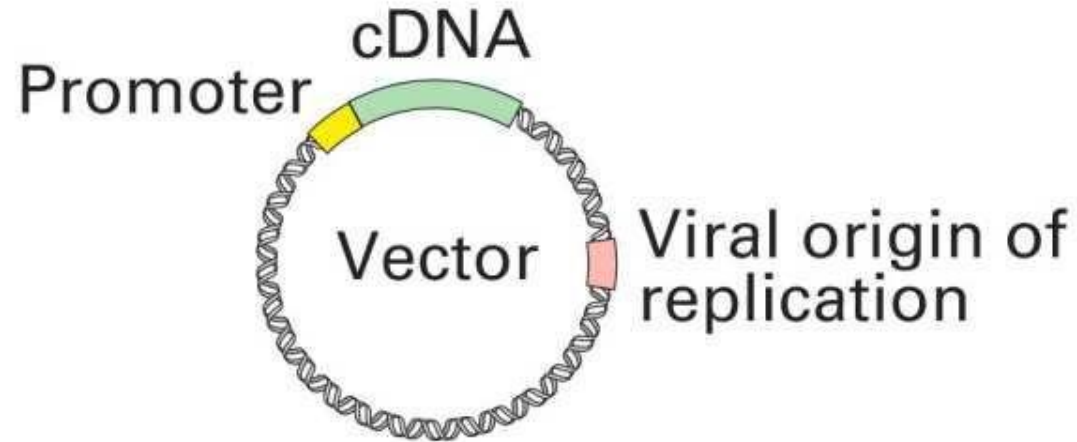




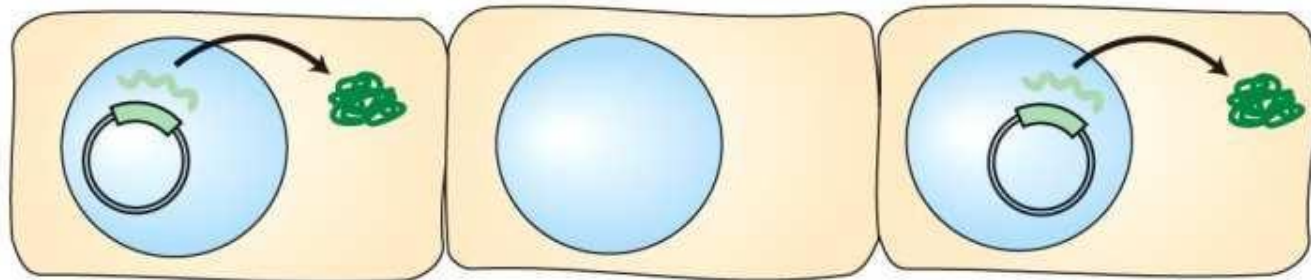
# Eukaryotic Expression vector



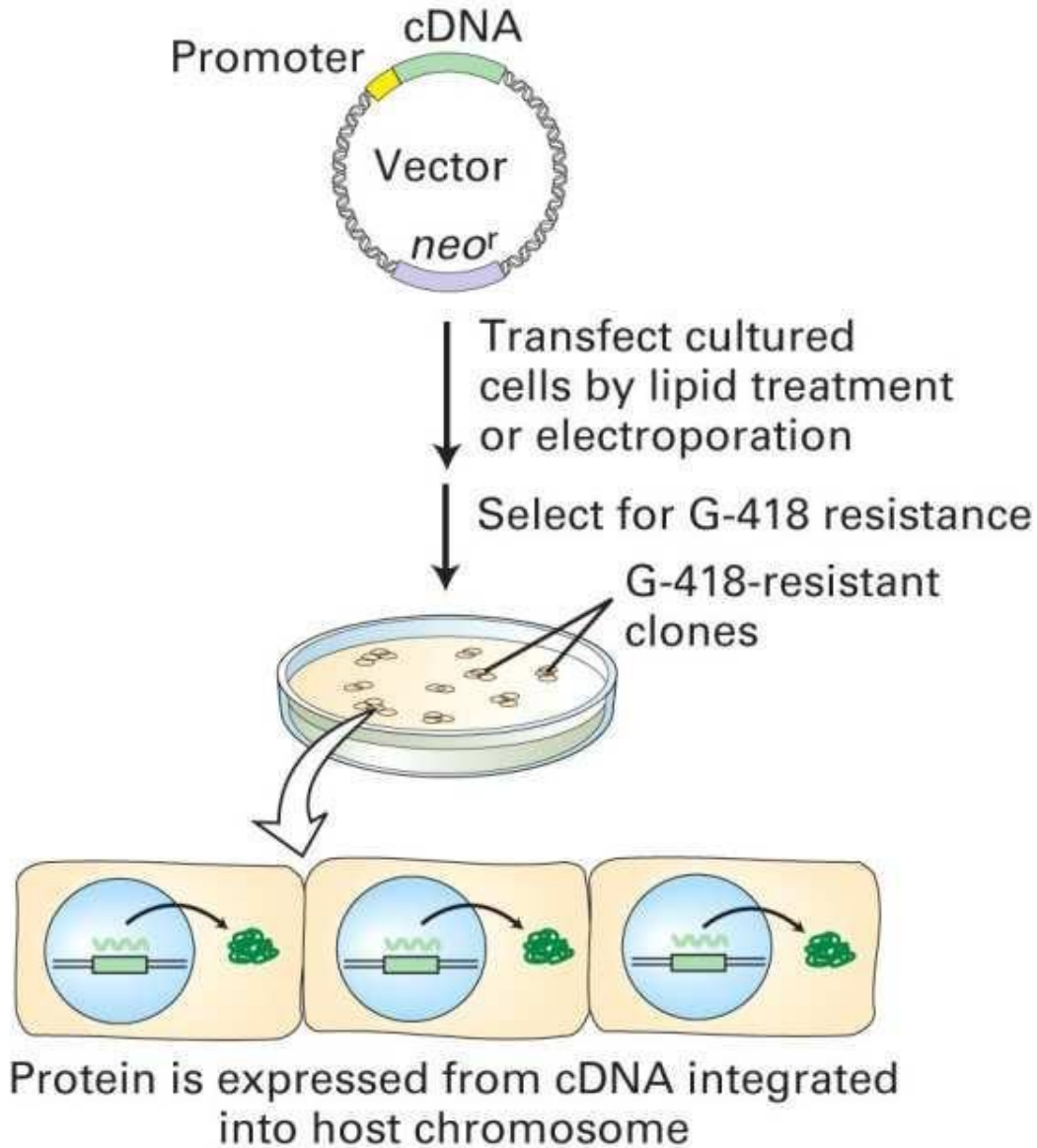
(a) Transient transfection



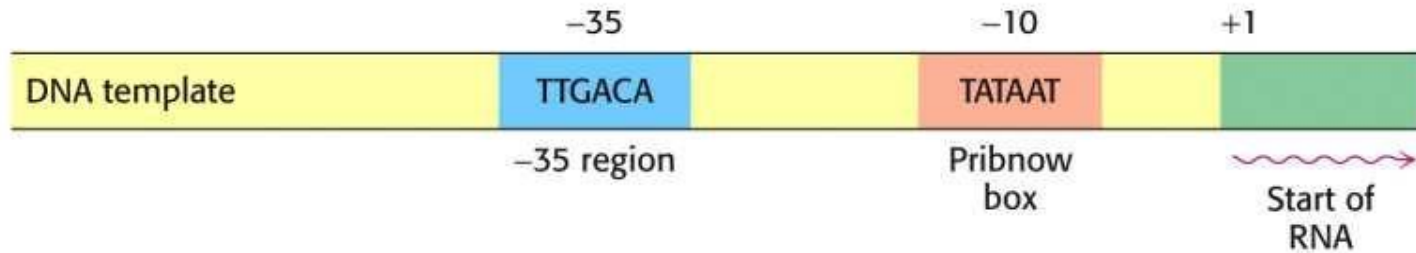
Transfect cultured cells by lipid treatment or electroporation



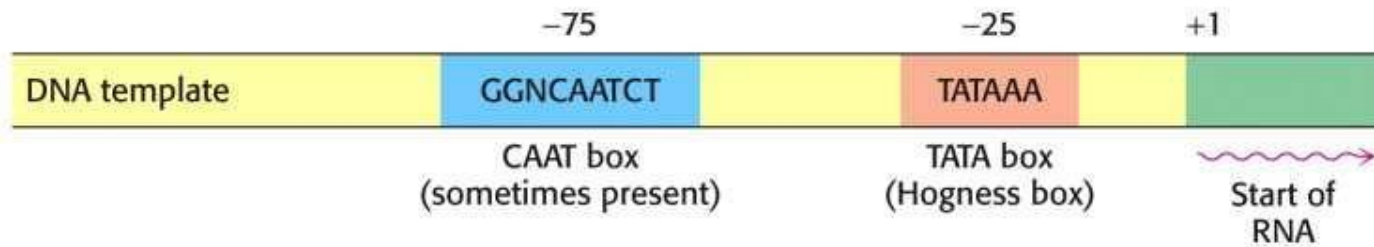
(b) Stable transfection (transformation)



# Promoters

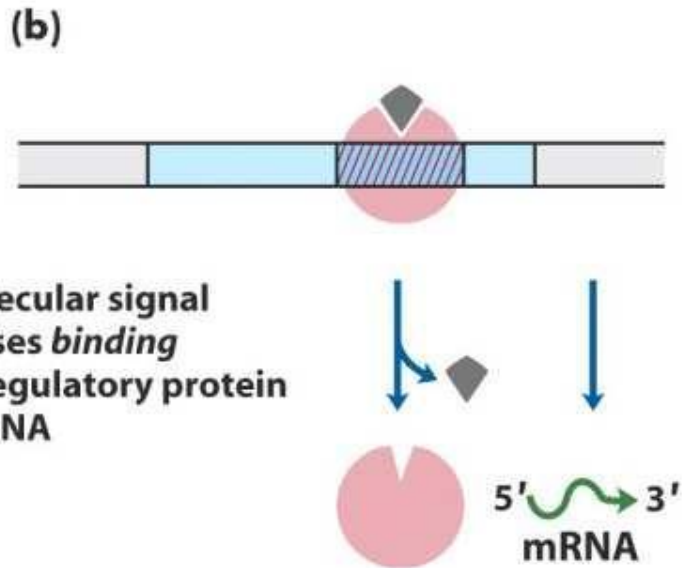
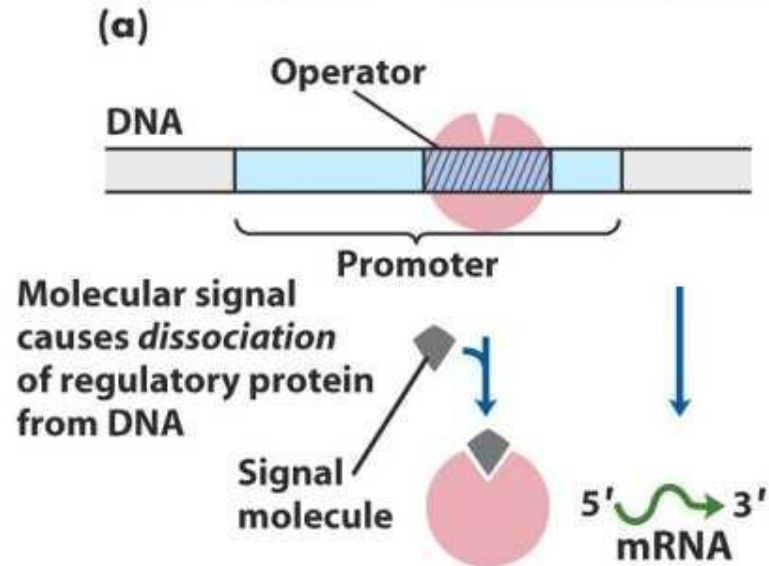


(A) **Prokaryotic promoter site**

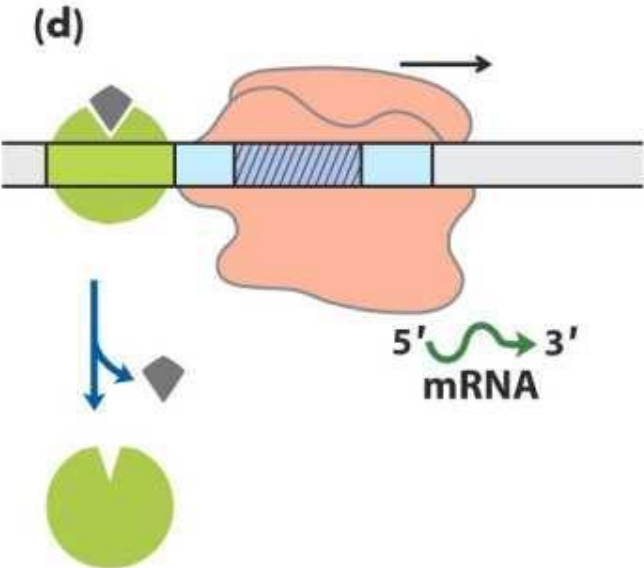
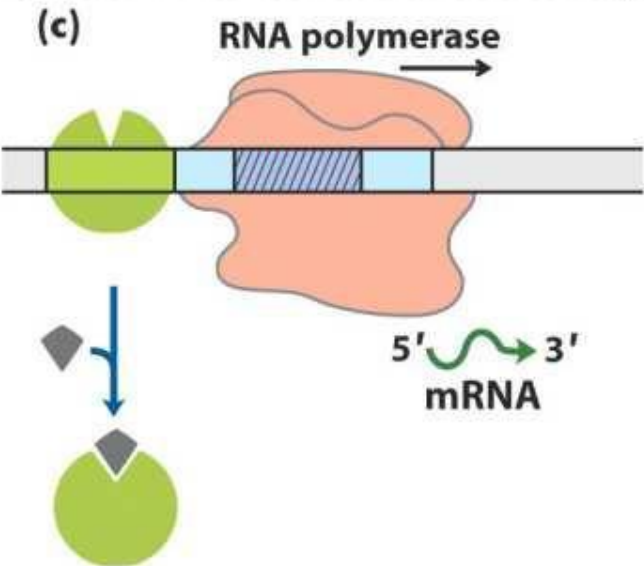


(B) **Eukaryotic promoter site**

**Negative regulation**  
(bound repressor inhibits transcription)

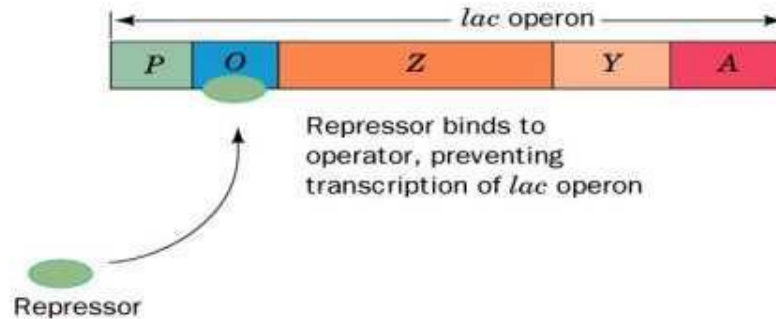


**Positive regulation**  
(bound activator facilitates transcription)

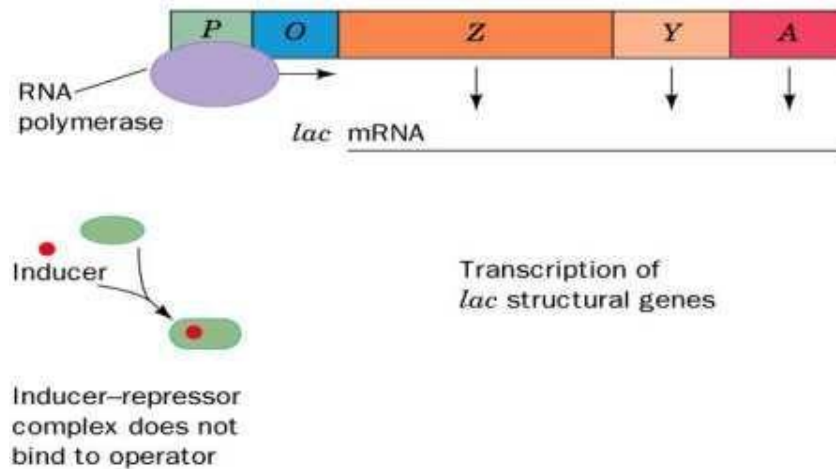


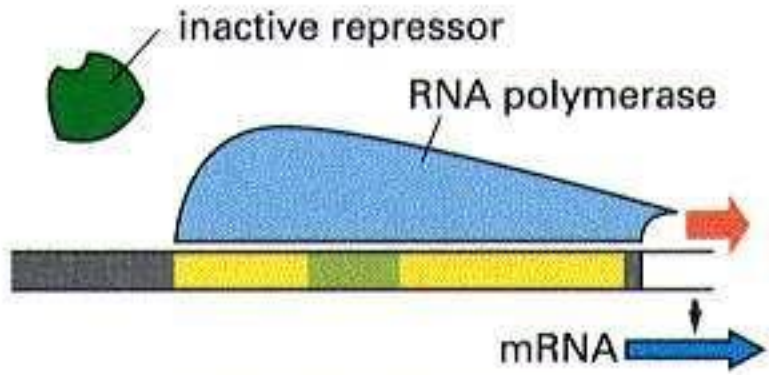
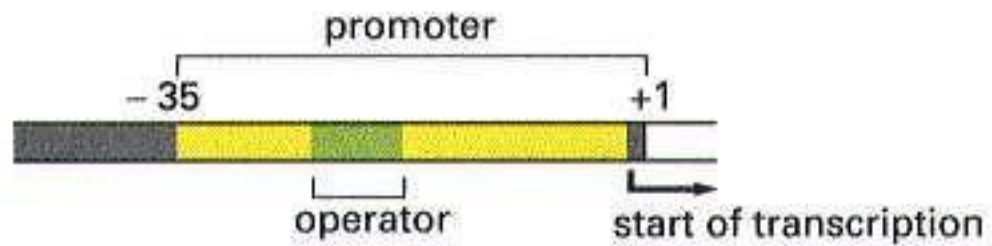
# Control of transcription of the *lac* operon.

(a) **Absence of inducer**

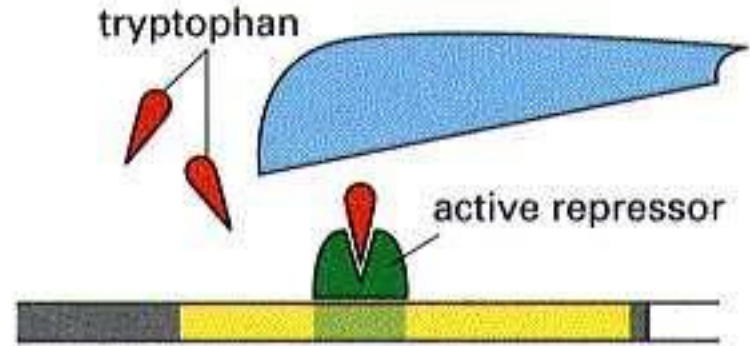


(b) **Presence of inducer**

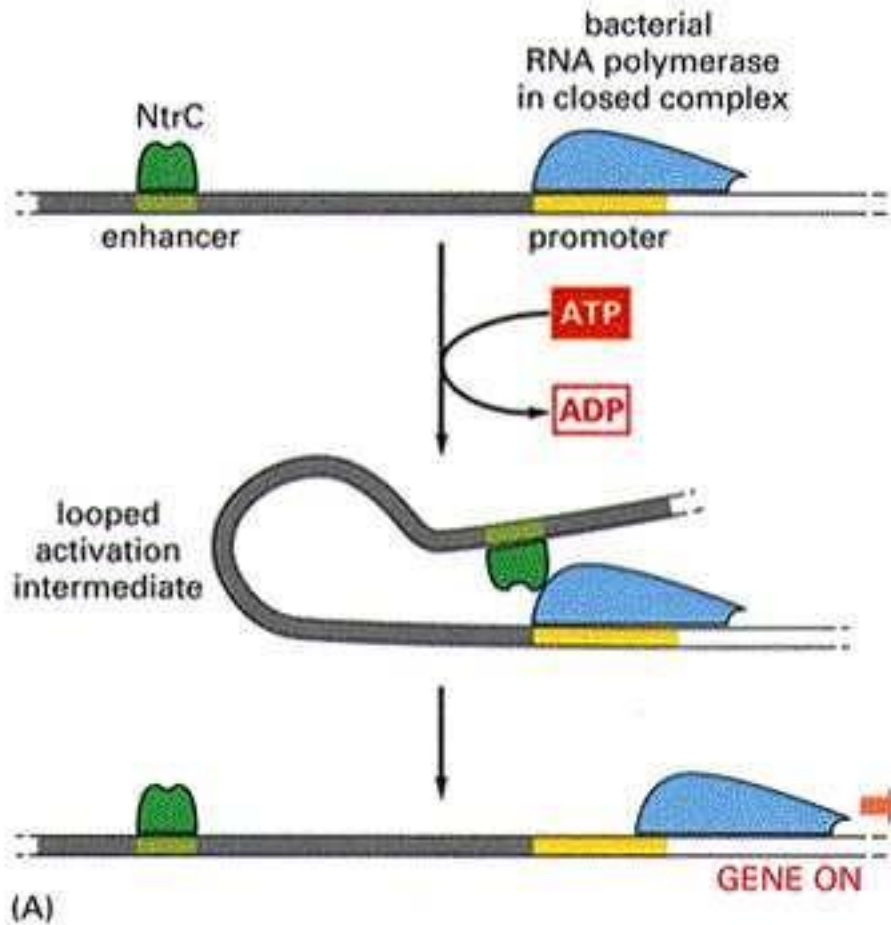




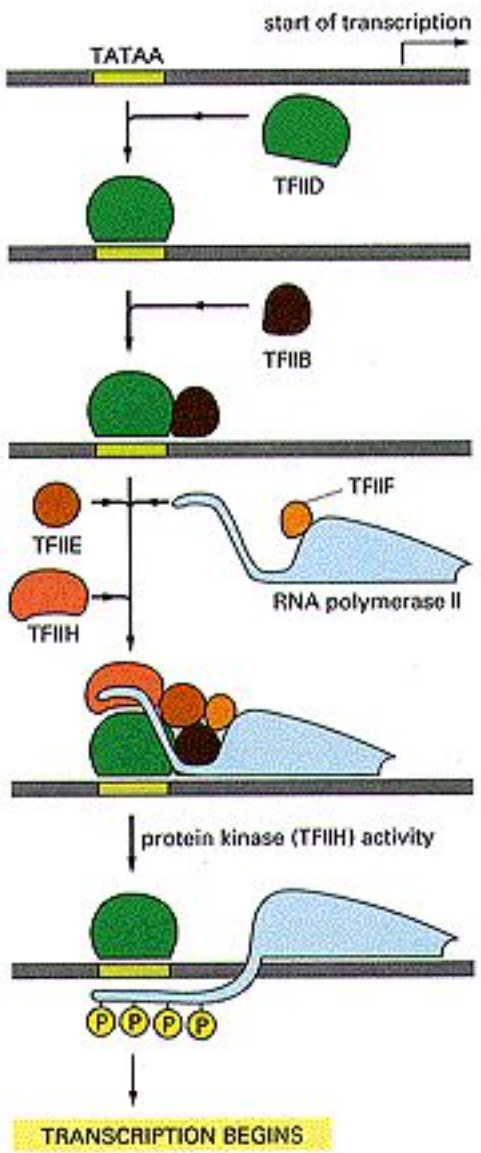
**GENES ARE ON**

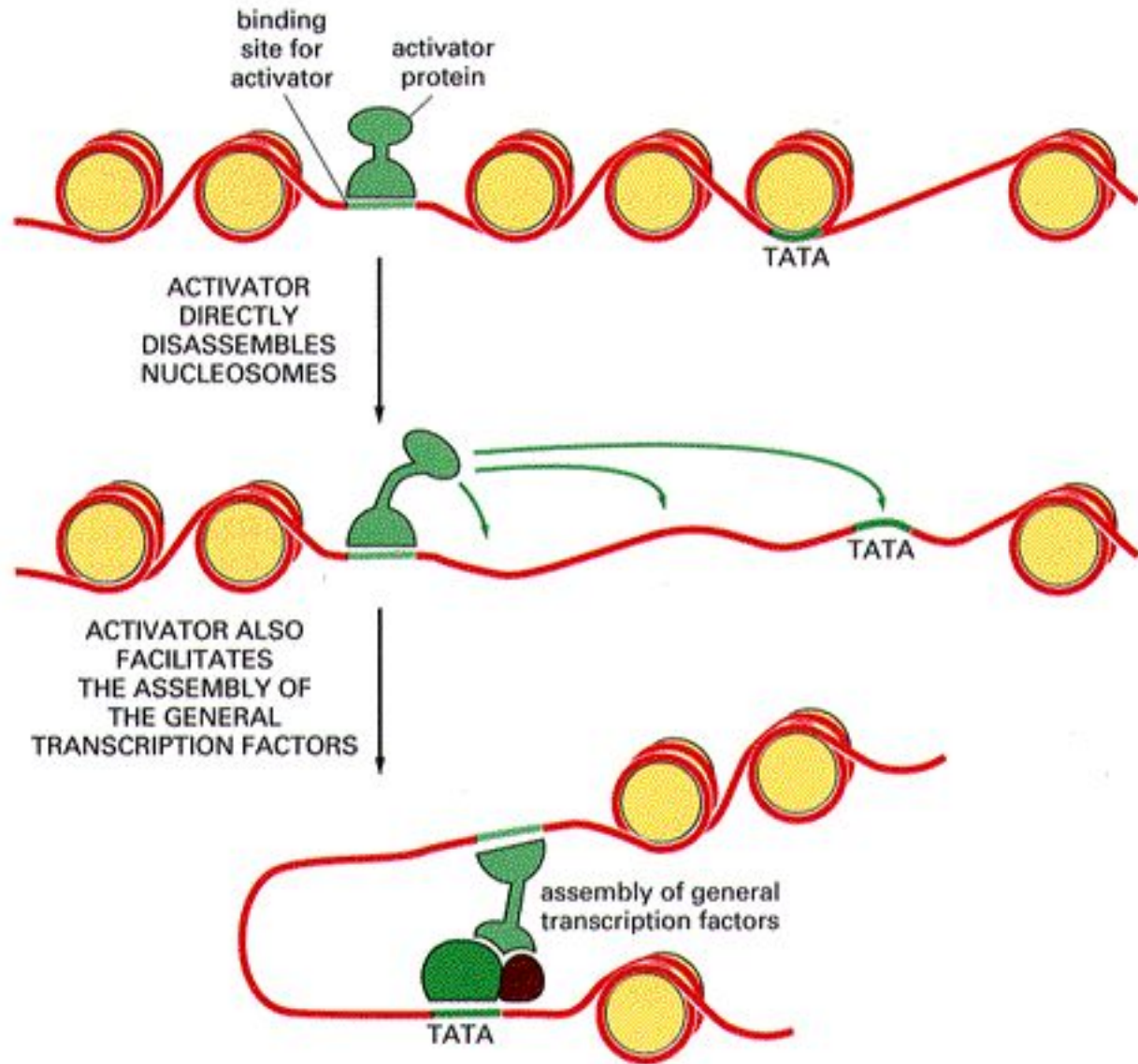


**GENES ARE OFF**

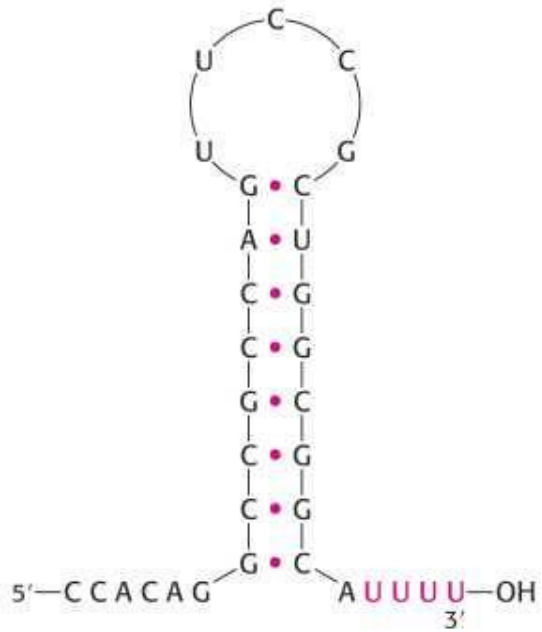




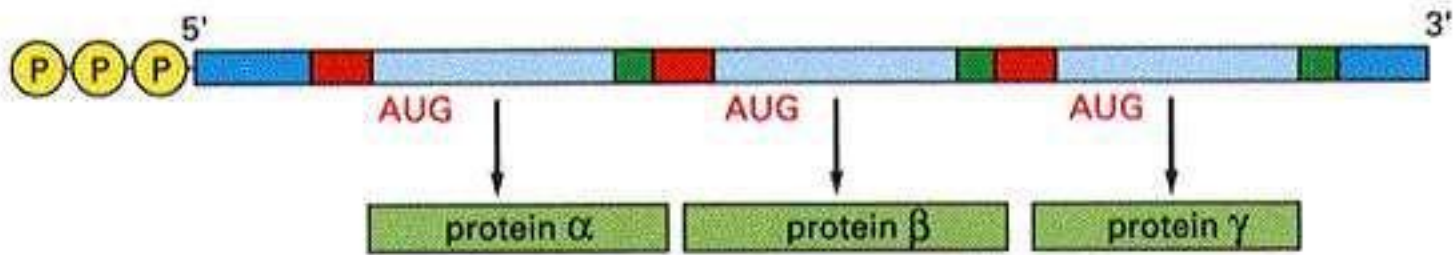




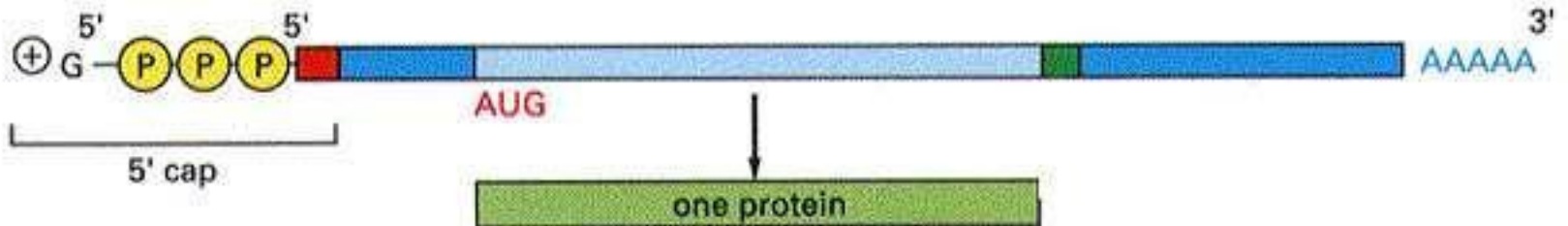
# Terminators




procaryotic mRNA





eucaryotic mRNA



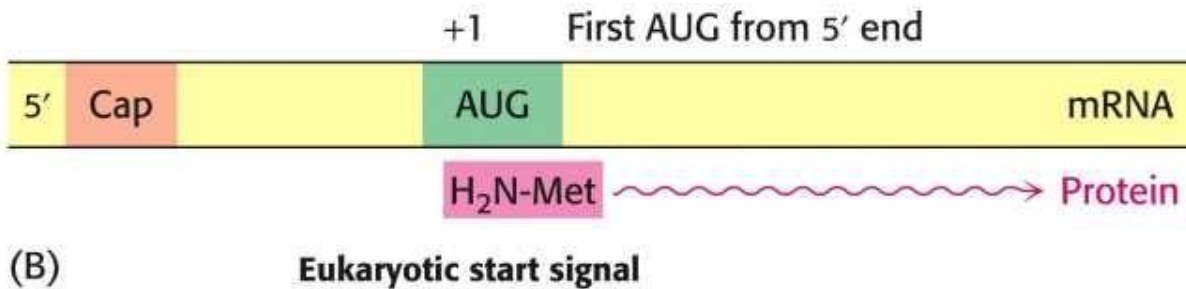
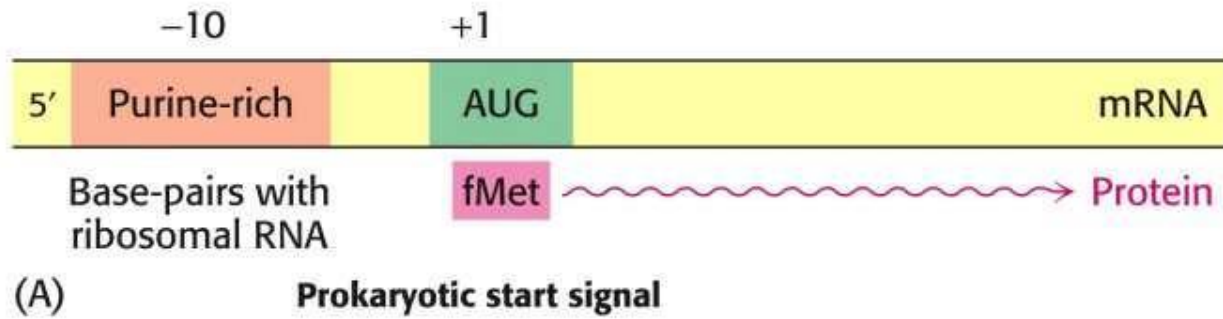
key:

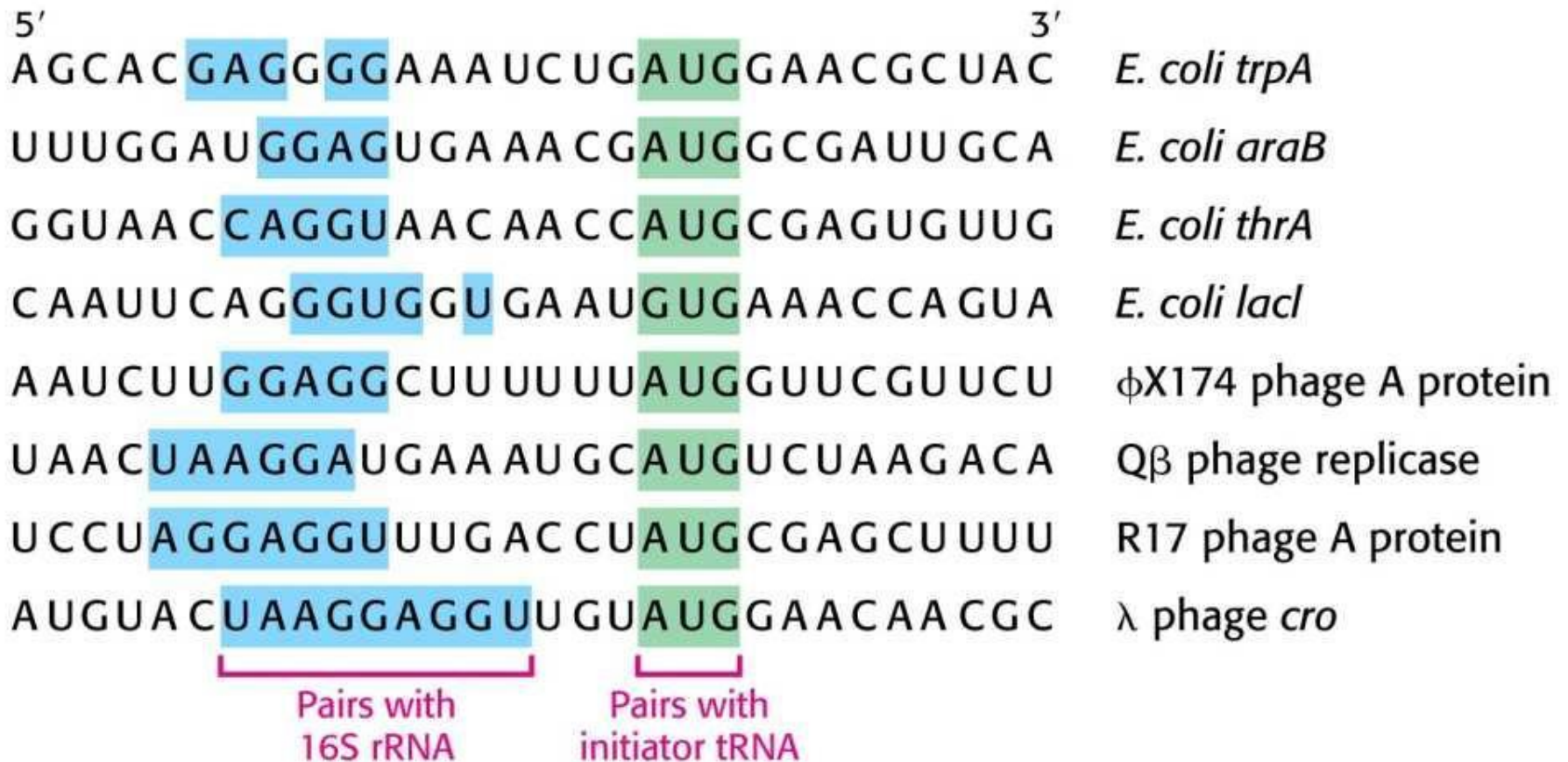
 ribosome-binding sites

 coding sequences

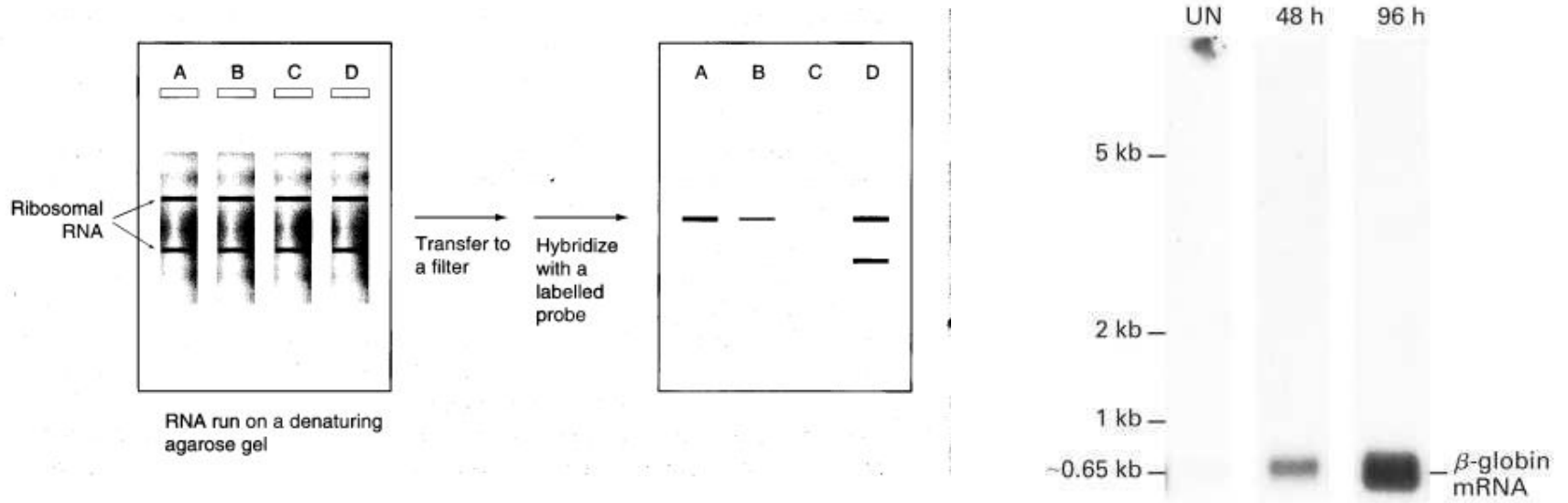
 noncoding sequences

 stop codons





# Northern Blot



-> to study transcription level

# Expression studies by microarray technique

