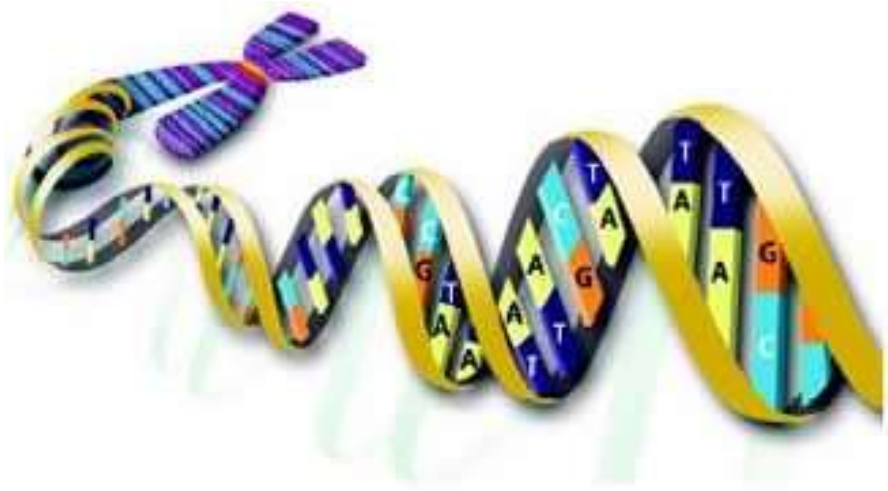
Gene sequencing



SEQUENCING

“Sequencing” means finding the order of nucleotides on a piece of DNA

- **Gene**
- Distinct portion of the DNA that codes for a type of protein or for an RNA chain.

- **DNA**

  A nucleic acid, made up of four similar chemicals called bases - A, T, C, and G that are repeated over and over in pairs.

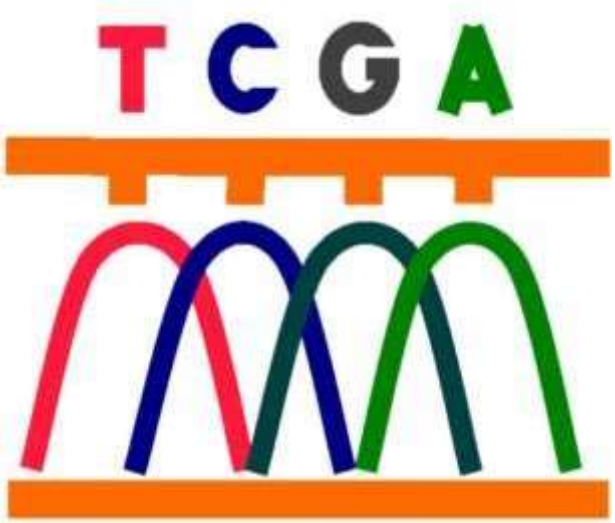
- **Gene sequencing**

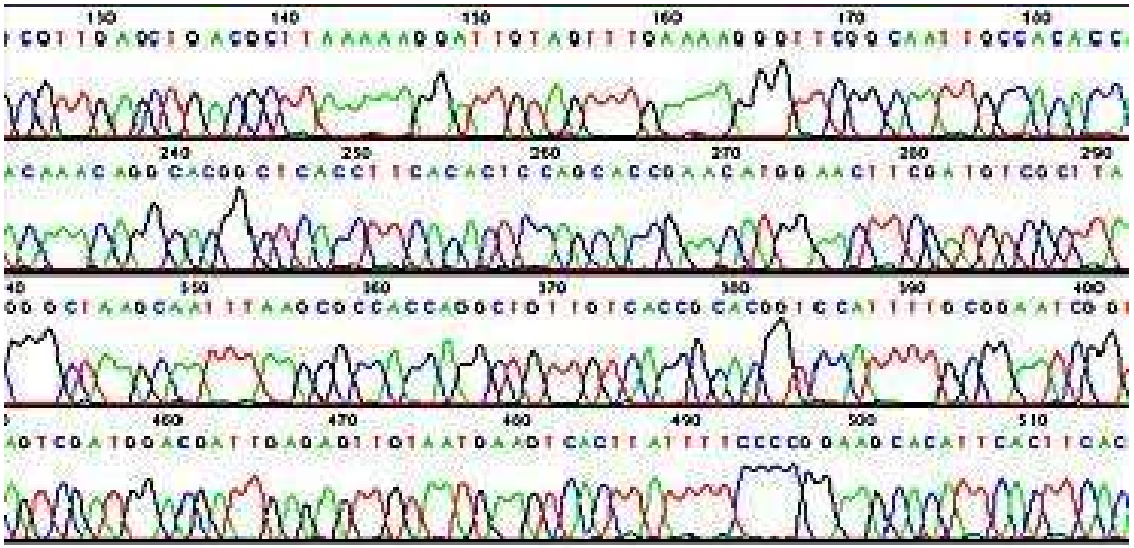
  Gene sequencing is a process in which the individual base nucleotides in an organism's DNA are identified.

- Gene sequencing = DNA sequencing = nucleotide sequencing = base sequencing

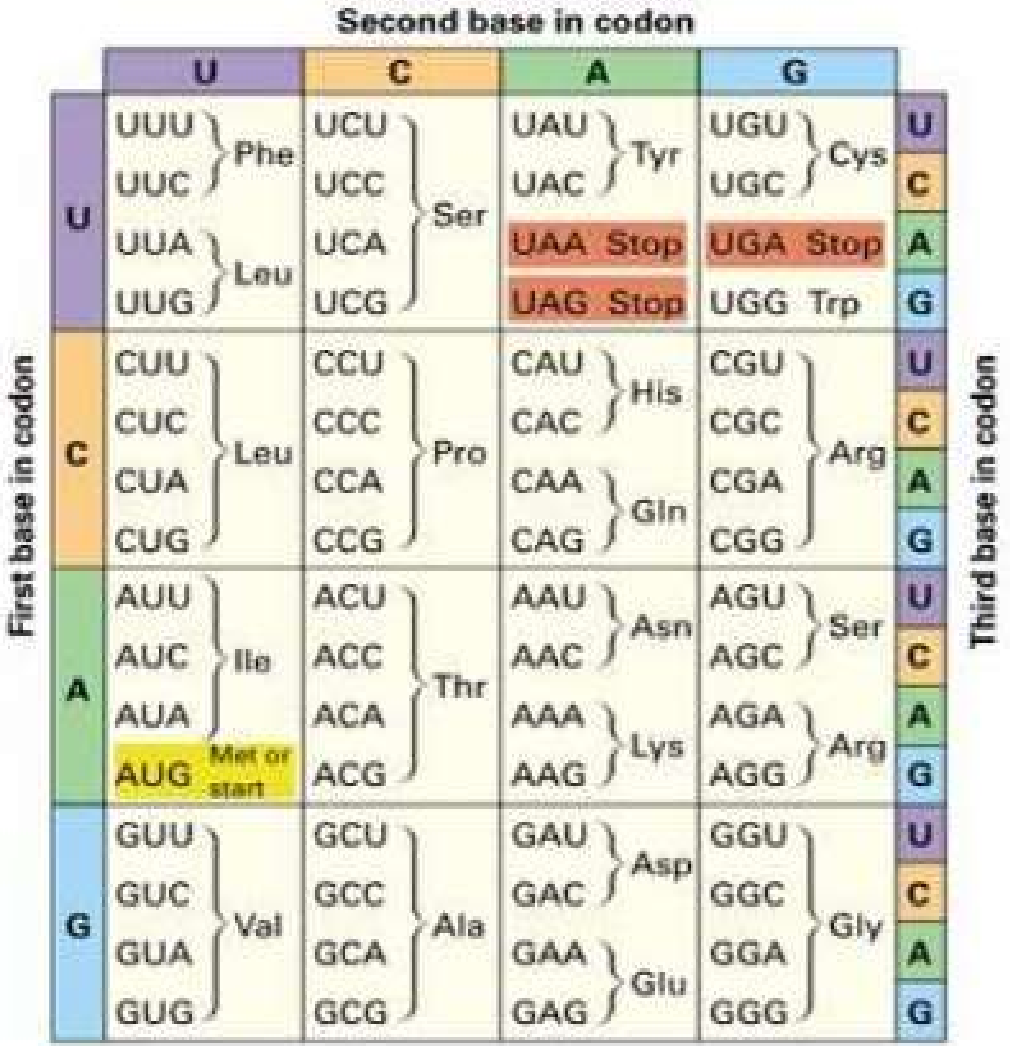
- The process of determining the order of nucleotides adenine (A), thymine (T), cytosine (C), and guanine (G) along a DNA strand.

DNA Sequencing





- **Nucleotide order determines Amino acid order, and by extension, protein structure and function**



Need of sequencing

- Understanding a particular DNA sequence can shed light on a genetic condition and offer hope for the eventual development of treatment.

- An alteration in a DNA sequence can lead to an altered or non functional protein, and hence to a harmful effect in a plant or animal.

- Simple point mutations can cause altered protein shape and function.

DNA Sequence variation can change the Protein produced by a particular gene

Gene A from person 1	Codon results in particular Amino Acid (AA) sequence	GCA	AGA	GAT	AAT	TGT
		Ala	Arg	Asp	Asn	Cys
Gene A from person 2	Codon change makes no difference in AA (redundant code)	GCG	AGA	GAT	AAT	TGT
		Ala	Arg	Asp	Asn	Cys
Gene A from person 3	Codon change results in different AA sequence	GCA	AAA	GAT	AAT	TGT
		Ala	[illegible]	Asp	Asn	Cys

- Simple point mutations such as this can cause altered protein shape and function.
- Diseases such as Sickle Cell Anaemia and Cystic Fibrosis are caused by point mutations

SEQUENCING METHODS

To determine the order of the nucleotide bases adenine, guanine, cytosine, and thymine in a molecule of DNA two methods were used

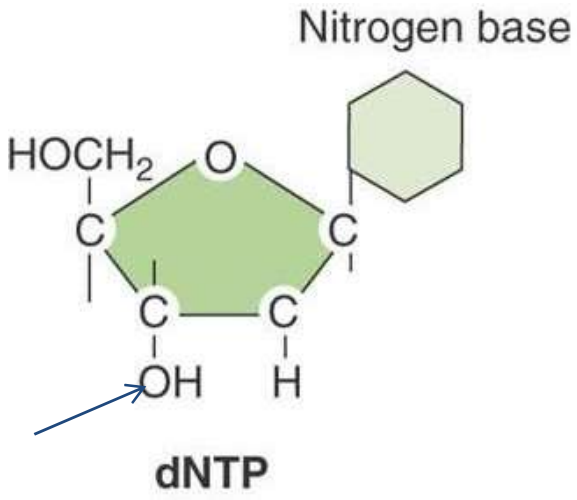
1. Sanger; Chain Termination Sequencing method
2. Maxam and Gilbert; Chemical Sequencing method

These two methods are most popular conventional methods

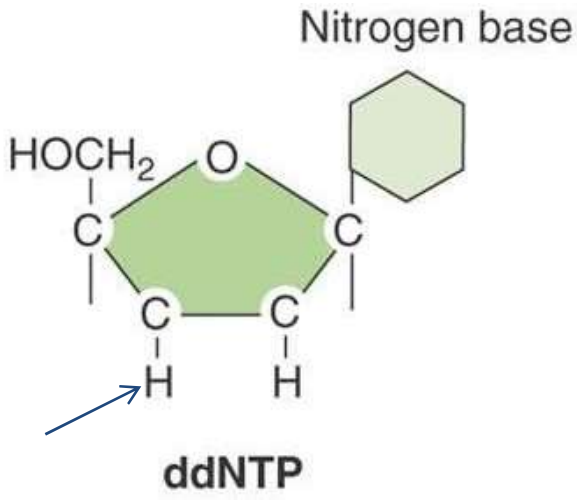
- Robotics and automated sequencing are based on these methods

Sanger's- Chain Termination Sequencing

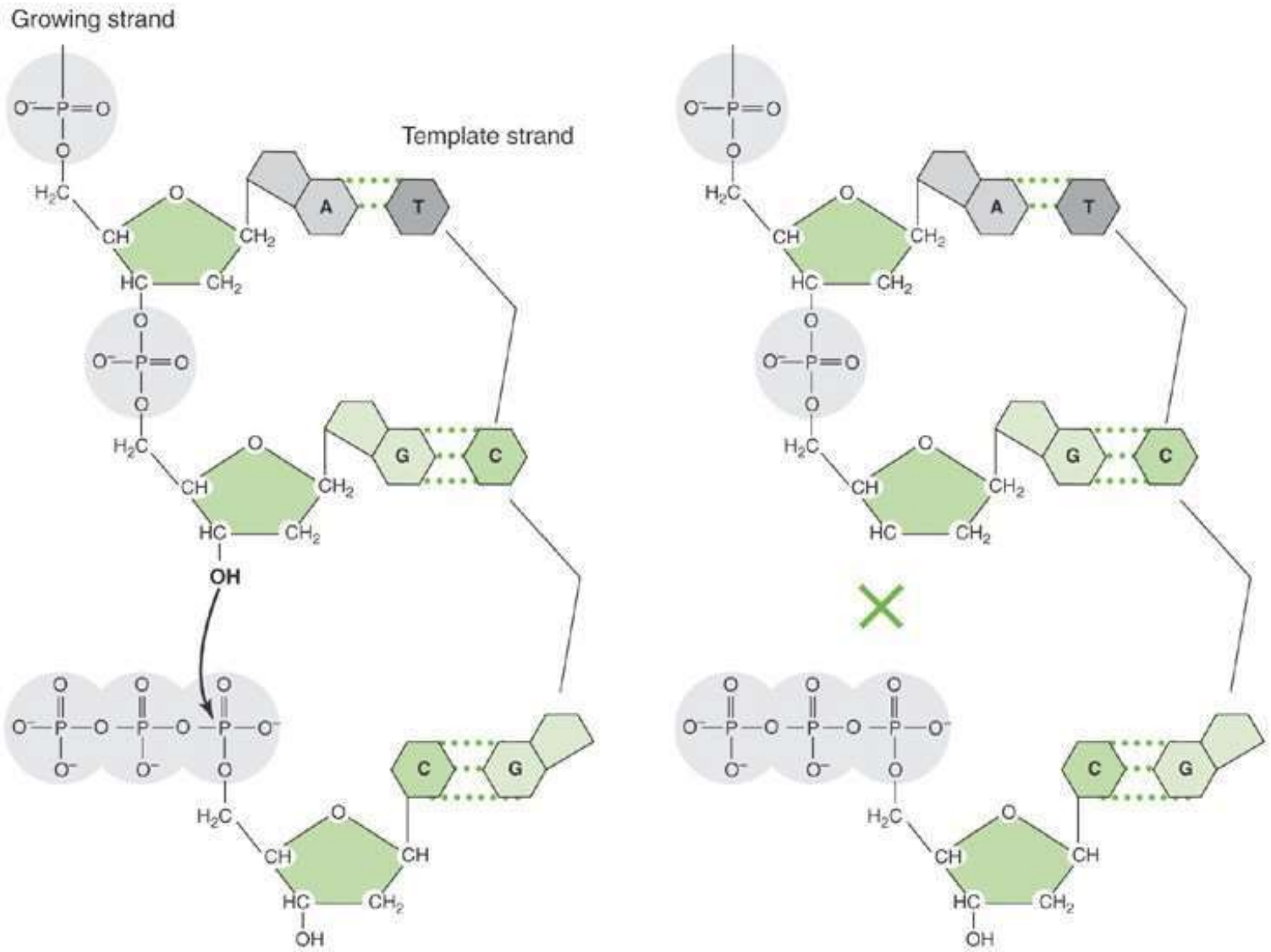
- It is PCR based method
- A modified DNA replication reaction
- Growing chains are terminated by dideoxynucleotides



dNTP



ddNTP



The 3’-OH group necessary for formation of the phosphodiester bond is missing in ddNTPs.

Sanger Sequencing: Process

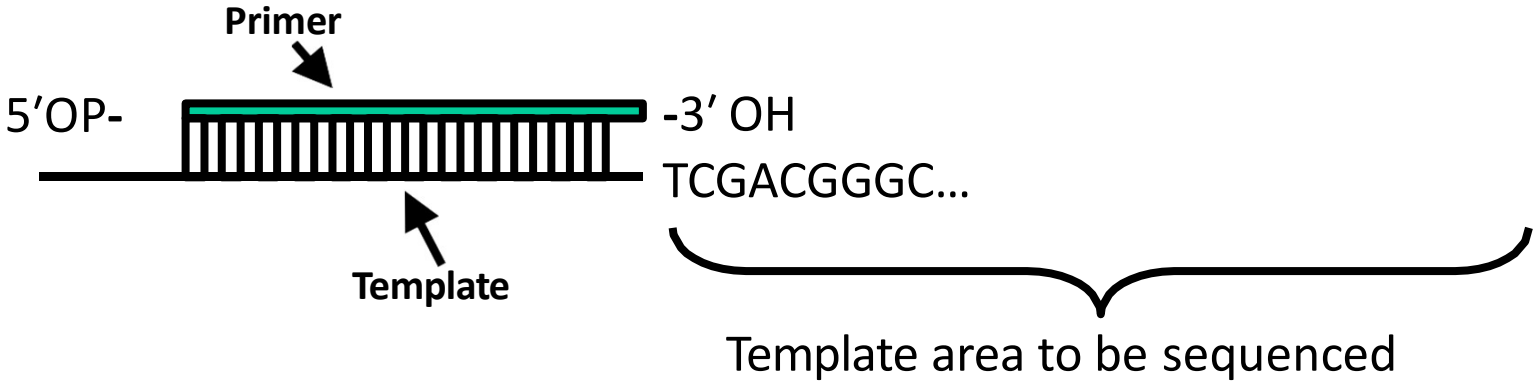
1. Get enough quantity of DNA (Run PCR)

2. Aliqot DNA sample into four different tubes

3. Prepare PCR reaction mix as below:

- Primer, Taq polymerase, template(ssDNA), dNTPS (All) and ddNTPs(ddATP, ddGTP,ddCTP & ddTTP respectively)
  - **To each reaction is added only one of the four dideoxynucleotides (ddATP, ddGTP, ddCTP, or ddTTP)**

- A sequencing reaction mix includes labeled primer and template.

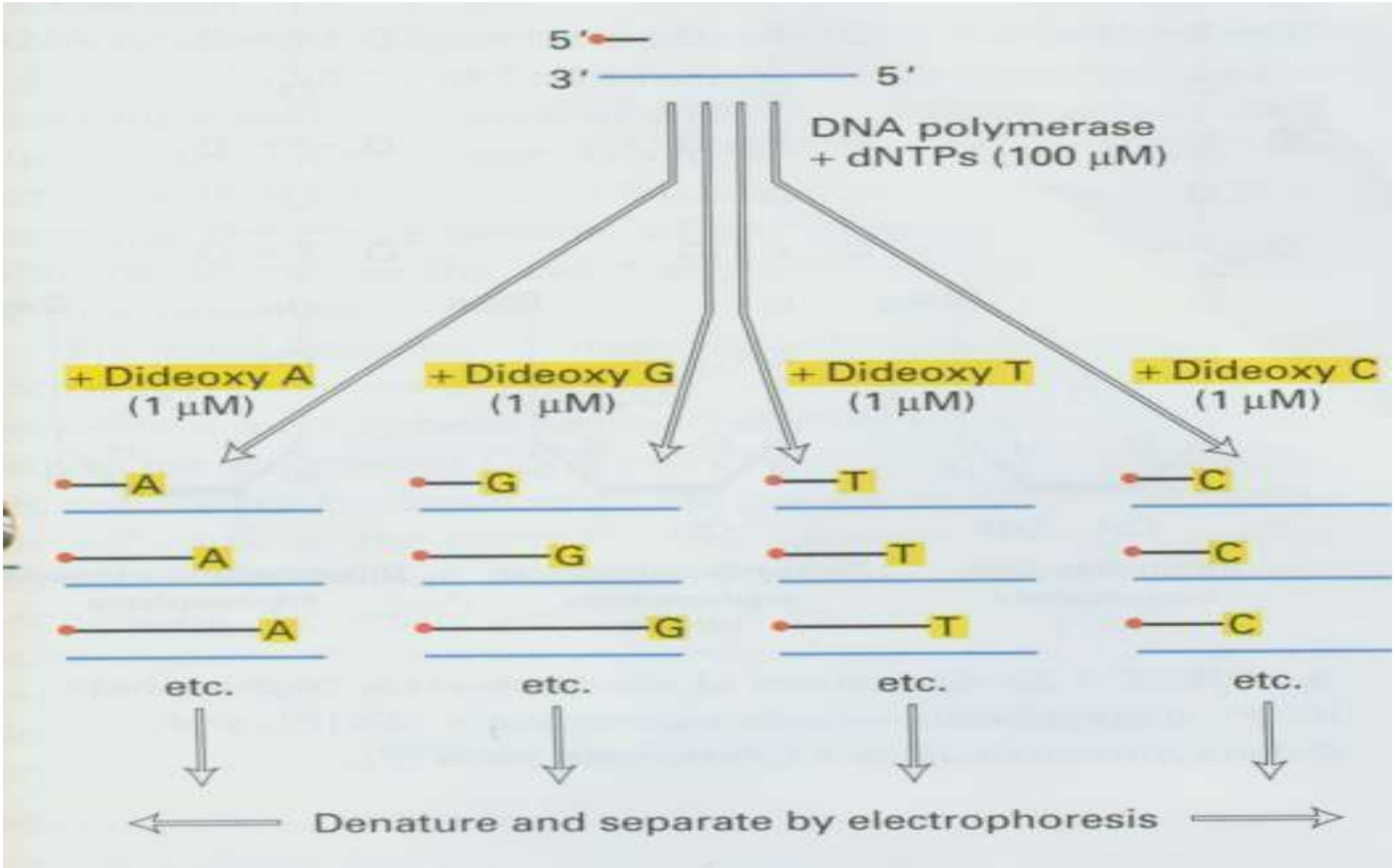


- Dideoxynucleotides are added separately to each of the four tubes.

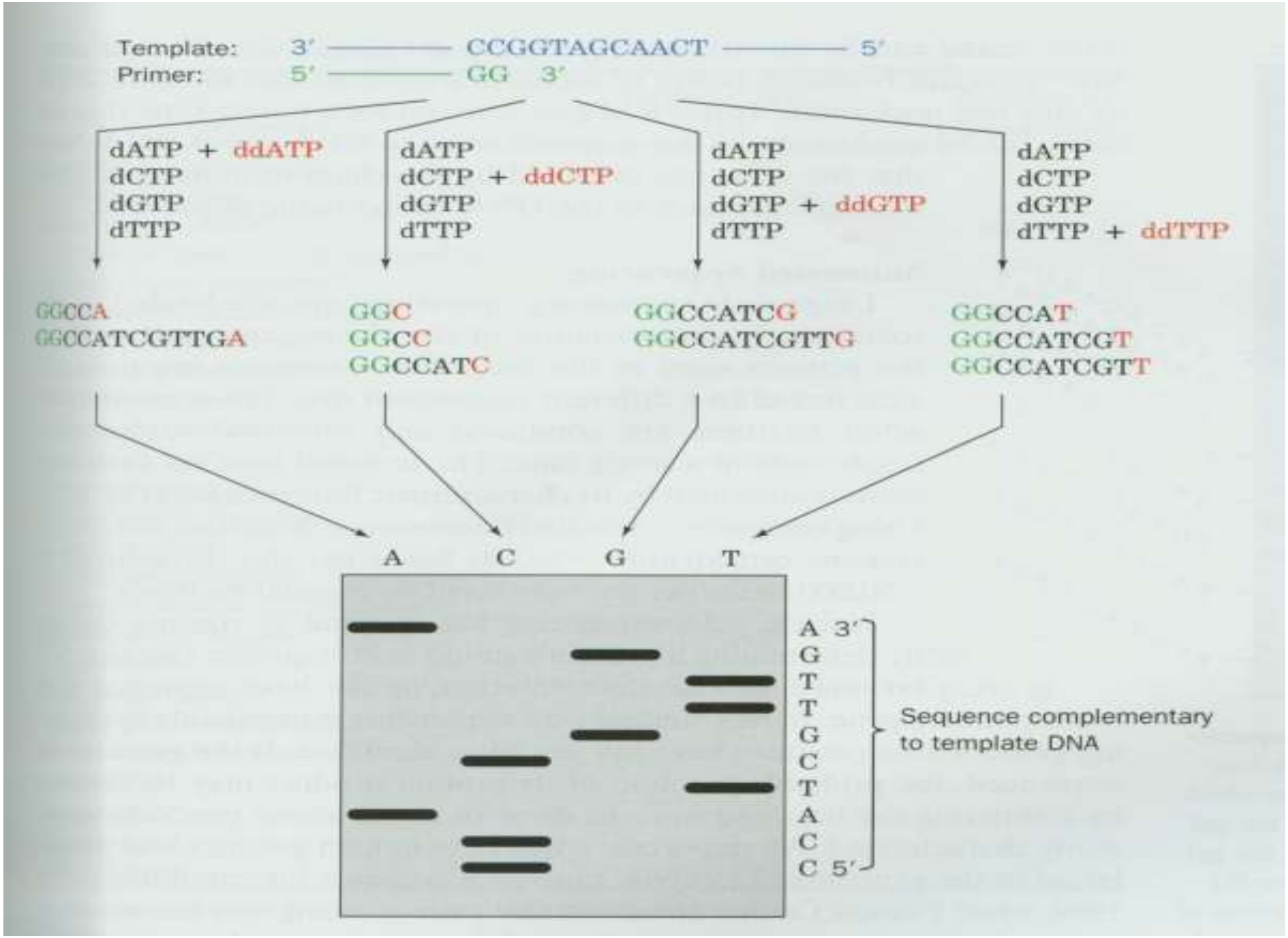
Sanger Sequencing: Process

1. Run PCR

- With addition of enzyme (DNA polymerase), the primer is extended until a ddNTP is encountered.
- The chain will end with the incorporation of the ddNTP.
- With the proper dNTP:ddNTP ratio, the chain will terminate throughout the length of the template.
- All terminated chains will end in the ddNTP added to that reaction.


5'
3'
5'
DNA polymerase
+ dNTPs (100 μM)
+ Dideoxy A
(1 μM)
+ Dideoxy G
(1 μM)
+ Dideoxy T
(1 μM)
+ Dideoxy C
(1 μM)
A
A
A
G
G
G
T
T
T
C
C
C
etc.
etc.
etc.
etc.
Denature and separate by electrophoresis

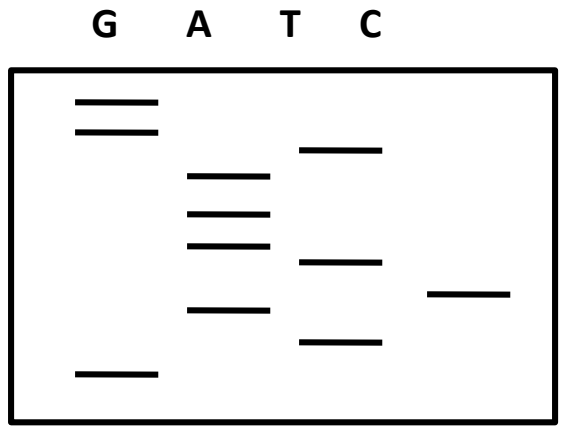
- The collection of fragments is a sequencing ladder.
- The resulting terminated chains are resolved by electrophoresis.
- Fragments from each of the four tubes are placed in four separate gel lanes.


Template: 3′ CCGGTAGCAACT 5′
Primer: 5′ GG 3′
dATP + ddATP
dCTP
dGTP
dTTP
dATP
dCTP + ddCTP
dGTP
dTTP
dATP
dCTP
dGTP + ddGTP
dTTP
dATP
dCTP
dGTP
dTTP + ddTTP
GGCCA
GGCCATCGTTGA
GGC
GGCC
GGCCATC
GGCCATCG
GGCCATCGTTG
GGCCAT
GGCCATCGT
GGCCATCGTT
A
C
G
T
A 3′
G
T
T
G
C
T
A
C
C 5′
Sequence complementary to template DNA

Chain Termination (Sanger) Sequencing

Longer fragments

Shorter fragments



3′
G
G
T
A
A
A
T
C
A
T
G
5′

Sequencing gels are read from bottom to top (5′ to 3′).

Advantages

- **Most popular method.**
- **Simpler and quicker allowing large output.**
- **Within an hour the primer-annealing and sequencing reactions can be completed.**

Disadvantages

- **Yielding of poor results owing to secondary structure in the DNA as sometimes DNA polymerases terminate chain elongation prematurely.**
- **The sequence is obtained not from the original DNA molecule but from an enzymatic copy. So, there is a chance of incorporation of wrong bases.**

Maxam–Gilbert sequencing method

Stages

1. The double-stranded fragment to be sequenced is isolated and radioactively labeled at the 5'-ends with ^{32}P.

2. The fragment is then cut with restriction enzyme and thus the label is removed from one end.

3. The fragment of DNA with one end labeled is denatured.

4. Four identical samples of these end-labeled DNA restriction fragments are subjected to chemical cleavage at different chemical nucleotides.

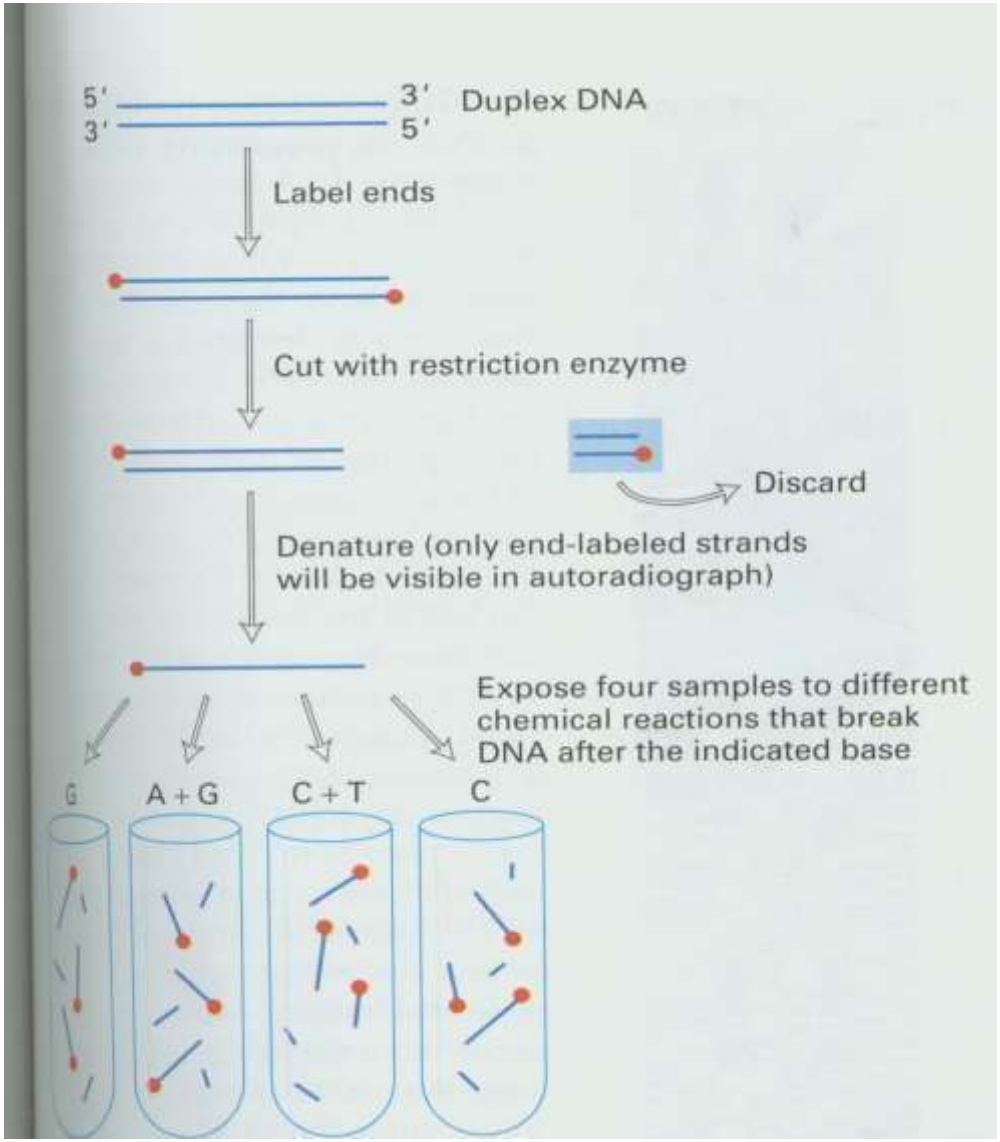


Figure: Maxam-Gilbert method (continued)

Lodish, H.;Berk, A. *et. al.* (4th ed); *Mol. Cell Biol.*; W. H. Freeman and Co. (2000) p: 233

Stages

There are four specific sets of chemical reactions that selectively

cut the DNA backbone at G, A+G, C+T, or C residues.

- **G only: Dimethyl sulphate (DMS) and piperidine**
- **A+G : DMS, piperidine**
- **C+T : Hydrazine, piperidine**
- **C only : Hydrazine, alkali, piperidine**

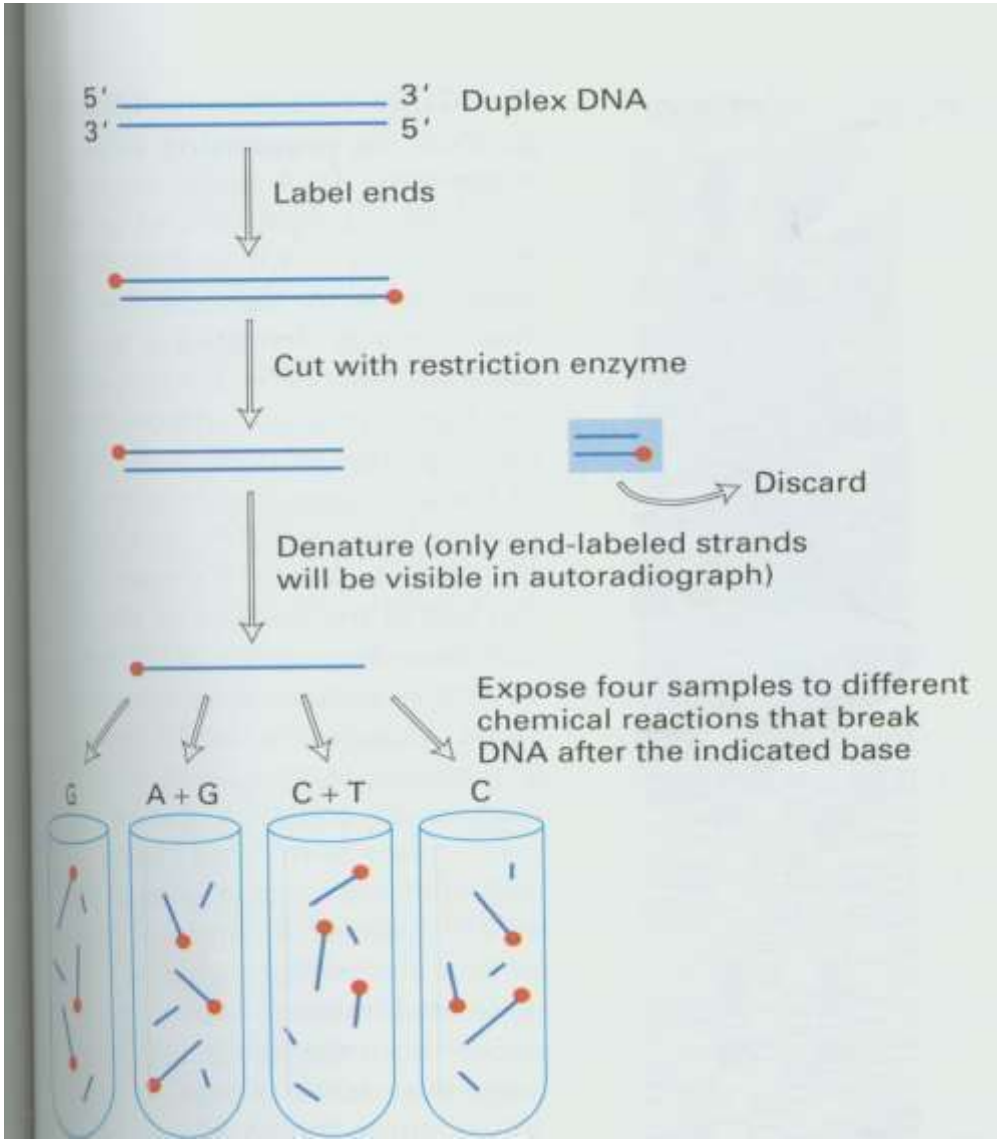
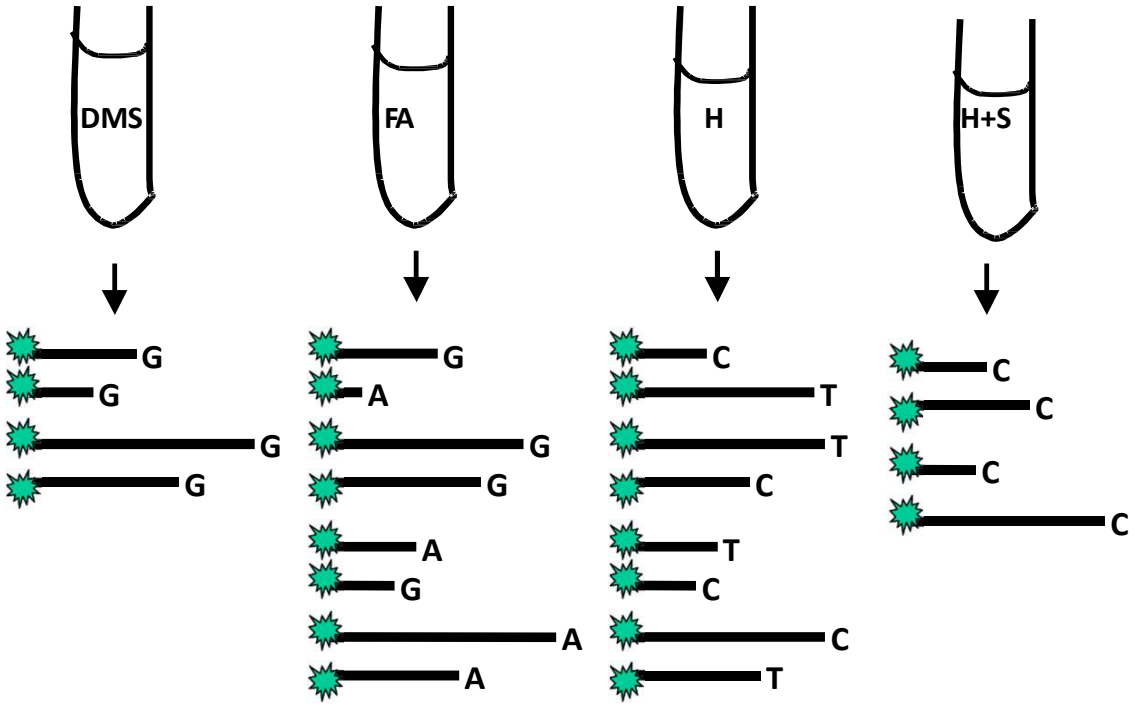


Figure: Maxam-Gilbert method (continued)

Lodish, H.;Berk, A. *et. al.* (4th ed); *Mol. Cell Biol.*; W. H. Freeman and Co. (2000) p: 233

Maxam-Gilbert Sequencing



Maxam-Gilbert sequencing is performed by chain breakage at specific nucleotides.

Contd…

6. For each labeled chain to be broken only once, the reactions are controlled.

7. The labeled sub-fragments created by the four reactions have the ^{32}P label at one end and the chemical cleavage point at the other end.

8. The reaction products are separated by polyacrylamide gel electrophoresis which is based on size. Smallest fragment goes fastest.

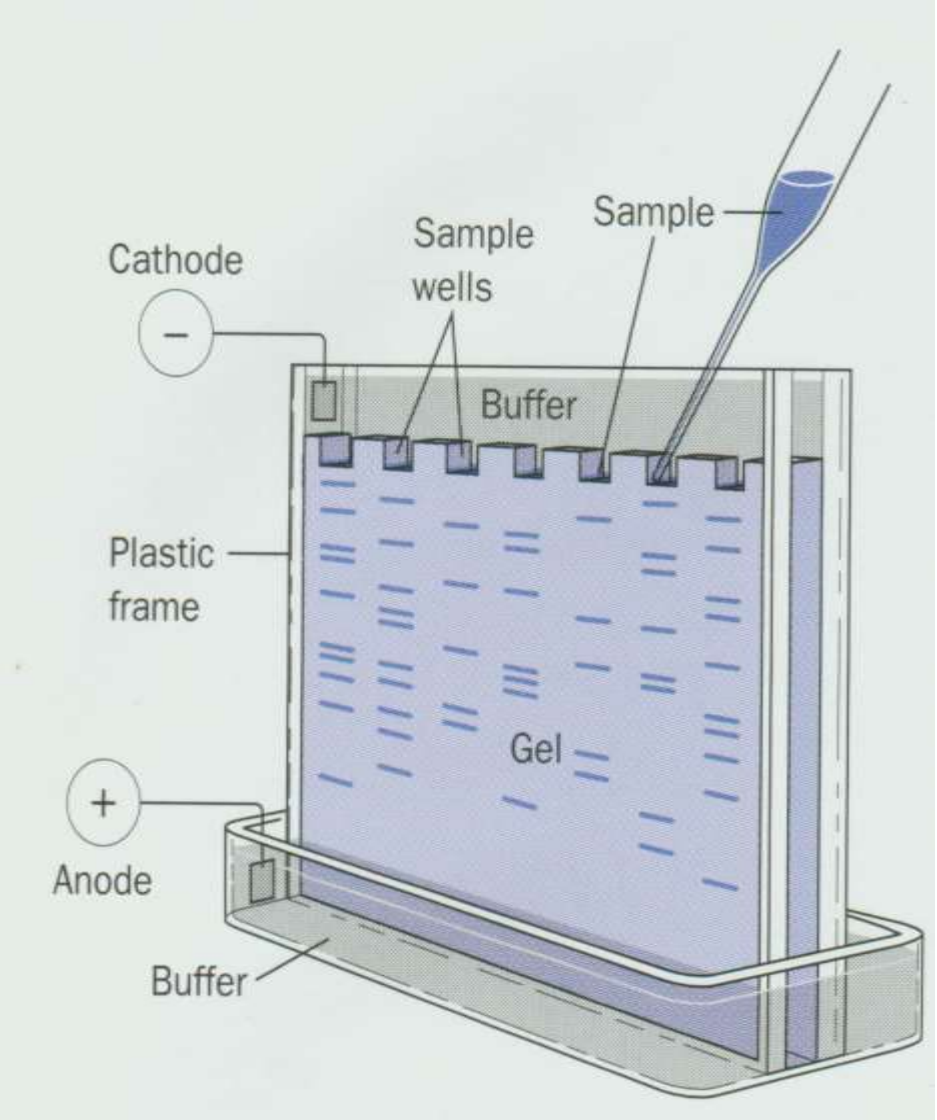


Figure: Apparatus for gel electrophoresis

Voet, D.; Voet, J. and Pratt, C. (upgrade ed) *Fundamentals of Biochemistry*; John Wiley and Sons, Inc (2002); p: 58

9. **The labeled fragments in the gel are visualized by autoradiography.**
10. **The sequence is read from bottom to top of the gel.**

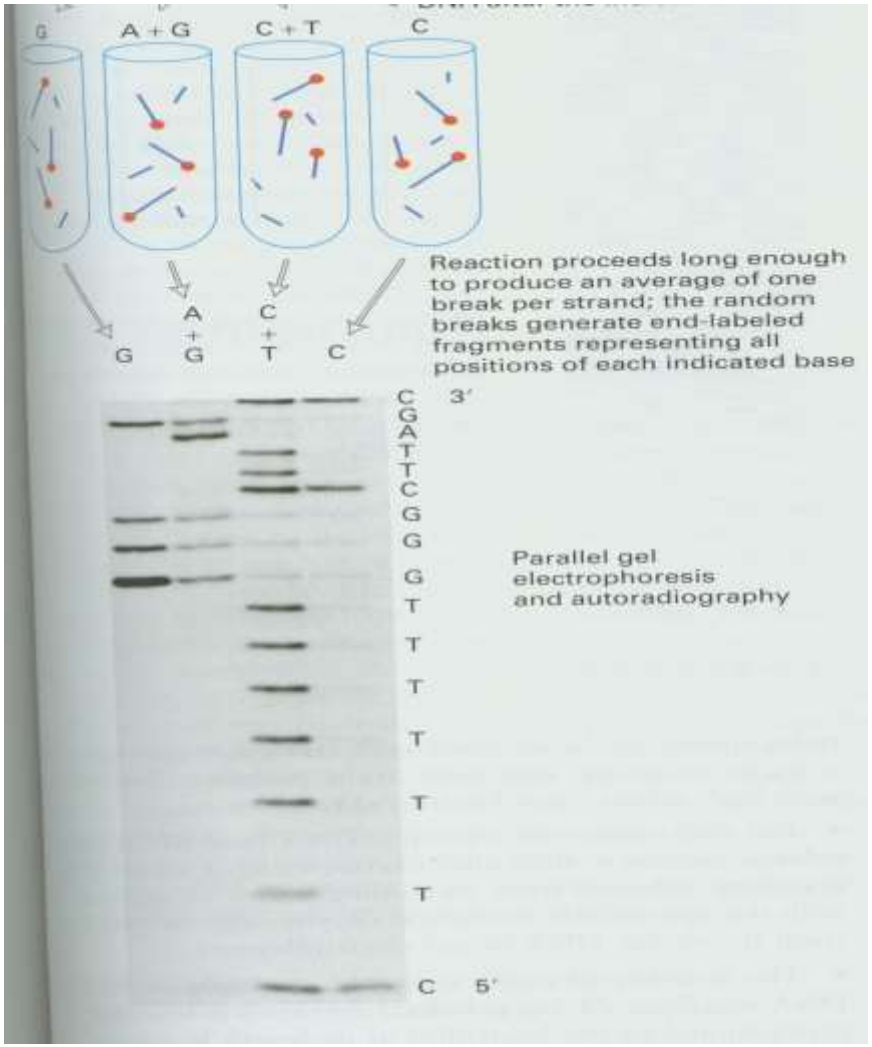


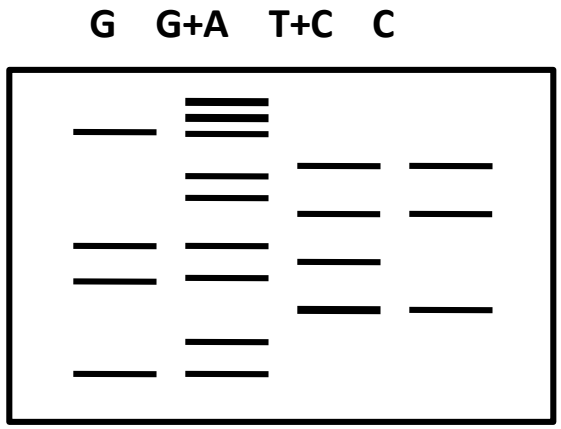
Figure: Maxam-Gilbert method

Lodish, H.;Berk, A. *et. al.* (4th ed); *Mol. Cell Biol*.; W. H. Freeman and Co. (2000) p: 233

Maxam-Gilbert Sequencing

Longer fragments

Shortest fragments



3′
A
A
G
C
A
A
C
G
T
G
C
A
G
5′

Sequencing gels are read from bottom to top (5′ to 3′).

Advantages

- No premature termination due to DNA sequencing. So, no problem with polymerase to synthesize DNA.
- Stretches of DNA can be sequenced which can not be done with enzymatic method.

Disadvantages

- Not widely used.
- Use of radioactivity and toxic chemicals.