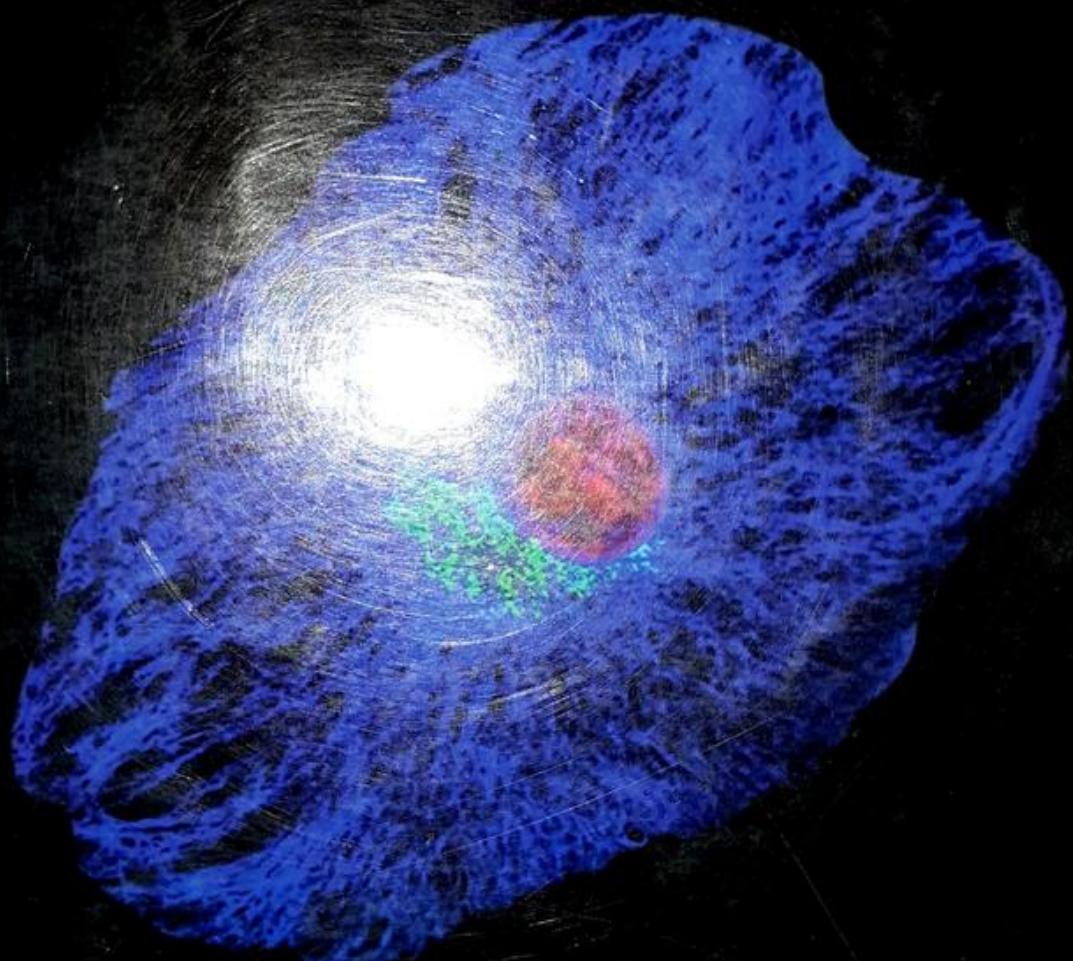


Fundamentals and Techniques of

Biophysics and Molecular Biology



Pranav Kumar

Contents

Biophysics

Chapter 01	Chromatography	1
Chapter 02	Electrophoresis	13
Chapter 03	Spectroscopy	26
Chapter 04	Mass spectrometry	47
Chapter 05	Centrifugation	53
Chapter 06	Microscopy	62
Chapter 07	Flow cytometry	78
Chapter 08	X-ray crystallography	89
Chapter 09	Patch clamp techniques	94
Chapter 10	Immunotechniques	99
Chapter 11	FRET and FRAP	114

Molecular Biology

Chapter 12	Molecular Biology Techniques	
12.1	Polymerase chain reaction	122
12.2	Nucleic acid hybridization	135
12.3	Labeling of nucleic acids	136
12.4	Blotting	146
12.5	Phage display	148
12.6	Yeast two-hybrid assay	150
12.7	Transcript analysis	151
12.8	DNA microarray	153
12.9	Electrophoretic mobility shift assay	156
12.10	Footprinting assay	157
12.11	Site-directed mutagenesis	160
12.12	DNA sequencing	167
12.13	Chromatin immunoprecipitation	175
12.14	Biosensors	176

Self Test

Index

Abbreviations

μm	micrometer
\AA	angstrom
A_{260}	absorbance at 260 nm
Ab	antibody
AD	activation domain
ADA	adenosine deaminase
ADP	adenosine 5'-diphosphate
AFLP	amplified fragment length polymorphism
AFM	atomic force microscopy
Ag	antigen
AMP	adenosine 5'-monophosphate
ATP	adenosine 5'-triphosphate
bis	bisacrylamide N,N'-methylenebisacrylamide
bp	base pair
BrdU	5-bromo-2-deoxyuridine
cccDNA	covalently closed circular DNA
CD	circular dichroism
cDNA	complementary DNA
CHEF	contour-clamped homogeneous electric field
CM	carboxymethyl
CNBr	cyanogen bromide
cpm	counts per minute
Da	dalton
dATP	deoxyadenosine triphosphate
DAPI	4',6-diamidino-2-phenylindole dihydrochloride
DBD	DNA binding domain
DBM	diazobenzyloxymethyl
ddNTP	dideoxynucleoside triphosphate
DEAE	diethylaminoethyl
DMS	dimethyl sulfate
dNTP	deoxynucleoside triphosphate
ELISA	enzyme-linked immunosorbent assay
EMSA	electrophoretic mobility shift assay
ER	electromagnetic radiation

EtBr	ethidium bromide
FACS	fluorescence activated cell sorter
FISH	fluorescence in situ hybridization
FITC	fluorescein isothiocyanate
FLIP	fluorescence loss in photobleaching
FRAP	fluorescence recovery after photobleaching
FRET	fluorescence (Förster) resonance energy transfer
FSC	forward scatter
GC	gas chromatography
GFP	green fluorescent protein
GLC	gas-liquid chromatography
GSC	gas-solid chromatography
HAT	hypoxanthine-aminopterin-thymidine
HGPRT	hypoxanthine-guanine phosphoribosyl transferase
HPLC	high performance liquid chromatography
IEF	isoelectric focusing
IFE	immunofixation electrophoresis
Ig	immunoglobulin
IR	infrared
kb	kilobase
kcal	kilocalorie
K _d	partition or distribution coefficient
kDa	kilodalton
LC	liquid chromatography
mAb	monoclonal antibody
MALDI	matrix-assisted laser desorption/ionization
Mb	megabase pair
MRI	magnetic resonance imaging
MS	mass spectrometry
NA	numerical aperture
nm	nanometer
NMR	nuclear magnetic resonance
ORD	optical rotatory dispersion
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PE	phycoerythrin
PFGE	pulsed-field gel electrophoresis

PI	propidium iodide
PMT	photomultiplier tube
RACE	rapid amplification of cDNA ends
RAPD	random amplification of polymorphic DNA
RCF	relative centrifugal field
R_f	relative front
RFLP	restriction fragment length polymorphism
RIA	radioimmunoassay
RPM	revolution per minute
RT-PCR	reverse transcription polymerase chain reaction
SDS	sodium dodecylsulfate
SEM	scanning electron microscope
SP	sulfopropyl
SPR	surface plasmon resonance
SSC	side scatter
STM	scanning tunneling electron microscopy
Taq	<i>Thermus aquaticus</i>
TdT	terminal deoxynucleotidyl transferase
TEM	transmission electron microscope
TEMED	N, N, N', N'-tetramethylethylenediamine
TLC	thin layer chromatography
T_m	melting temperature
TMS	tetramethylsilane
TOF	time-of-flight
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet
Vis	visible

Units, Conversion factors and Physical constants

Base units

Length	meter (m)
Mass	kilogram (kg)
Time	second (s)
Electric current	ampere (A)
Temperature	kelvin (K)
Amount of substance	mole (mol)
Luminous intensity	candela (cd)

Conversion factors

Multiplication factor	Prefix	Symbol
10^{12}	tera	T
10^9	giga	G
10^6	mega	M
10^3	kilo	k
10^2	hecto	h
10	deka	da
10^{-1}	deci	d
10^{-2}	centi	c
10^{-3}	milli	m
10^{-6}	micro	μ
10^{-9}	nano	n
10^{-12}	pico	p
10^{-15}	femto	f

Length units

Millimeters (mm)	Centimeters (cm)	Meters (m)	Kilometers (km)
1	0.1	0.001	0.000001
10	1	0.01	0.00001
1000	100	1	0.001
1000000	100000	1000	1

Volume units

Cubic centimeter (cm ³)	Cubic meter (m ³)	Liter (ltr)
1	0.000001	0.001
1000000	1	1000
1000	0.001	1

Conversion units

1 revolution = 2π radians = 360 degrees

1 degree = 60 minutes

π radians = 3.1416 radians = 180 degrees

1 Hz = 1 cycle/sec.

1 rpm (revolutions per minute) = 60 rps (revolutions per second) = 60Hz

1 Hz (Hertz) = 1 s⁻¹

1 N (Newton) = 1 kgm/s² = 100,000 dyne

1 dyne = 10^{-5} Newton

1 Pascal (Pa) = 1 N/m² = 6.895 kPa

1 atm (metric atmosphere) = 760 mm Hg at 0°C = 1.0132×10^5 N/m²

1 microbar = 0.1 N/m²

1 angstrom (\AA) = 10^{-10} m

Physical constants

Ideal gas law constant (R) = 1.987 cal/mole K

Boltzmann's constant (K) = 1.3×10^{-16} erg/K = 1.3×10^{-23} J/K

Planck's constant (h) = 6.62×10^{-27} erg-sec = 6.62×10^{-34} J.sec

Avogadro's number = 6.02×10^{23} mol⁻¹

Density of water = 1 g/cm³

Electron charge = 1.60×10^{-19} coulombs

Electron rest mass = 9.11×10^{-31} kg

Proton rest mass = 1.67×10^{-27} kg

Speed of light (c) = 3.00×10^8 m/sec

Gravitational constant (G) = 6.67×10^{-11} Nm²/kg²

Acceleration due to gravity (g) = 9.8 m/s²

Chapter 1

Chromatography

Chromatography is a physical method for separation of compounds. Tswet, Russian botanist (referred as the *father of chromatography*) is credited for the development of chromatography. He employed the technique to separate various plant pigments such as chlorophylls and xanthophylls by passing solutions of these compounds through a glass column packed with finely divided calcium carbonate. The separated species appeared as colored bands on the column, which accounts for the name he chose for the method (Greek *chroma* meaning 'color' and *graphein* meaning 'writing').

Chromatography is based on the fact that sample distributes or partitions itself to different extents in two different, immiscible phases, which is described by a partition or distribution coefficient, K_d . If we consider two immiscible phases A and B,

$$K_d = \frac{\text{Concentration of sample in phase A}}{\text{Concentration of sample in phase B}}$$

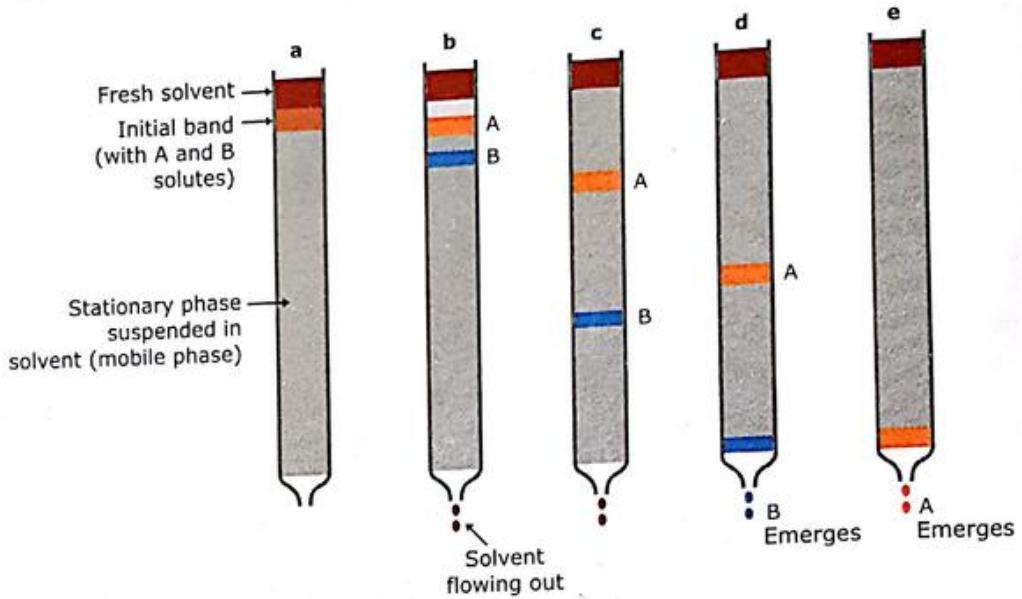
The two immiscible phases could be a solid and a liquid, or a gas and a liquid or a liquid and another liquid. One of the two phases is a stationary phase (a solid or a liquid supported on a solid) and does not move and the other is a mobile phase and moves with respect to first. The mobile phase may be a liquid (liquid chromatography) or a gas (gas chromatography). All chromatographic methods involve passing a mobile phase through a stationary (immobile) phase. The two phases are chosen so that the components of the sample distribute themselves between the mobile and stationary phases to varying degrees.

Classification of chromatographic methods

Chromatographic methods can be classified in two fundamental ways. The first classification is based on the physical means by which the stationary and mobile phases are brought into contact. On this basis, chromatography is classified into column and planar chromatography. In planar chromatography, the stationary phase is present as or on a plane. The plane can be a paper, serving as such or impregnated by a substrate as the stationary bed (paper chromatography, PC) or a layer of solid particles spread on a support e.g. a glass plate (thin layer chromatography, TLC). Planar chromatography is also termed open-bed chromatography. In thin layer chromatography, the stationary phase is a thin layer of silica gel or alumina on a glass, metal or plastic plate. Most commonly, silica gel is used as a stationary phase. In silica gel, the silicon atoms are joined via oxygen atoms in a giant covalent structure. The other commonly used stationary phase is alumina (aluminium oxide).



✓ In column chromatography, the stationary bed is within a tube. The particles of the solid stationary phase or support coated with a liquid stationary phase may fill the whole inside volume of the tube (*packed column*) or be concentrated on or along the inside tube wall leaving an open, unrestricted path for the mobile phase in the middle part of the tube (*open-tubular column*).



Chromatography represents a separation technique; whereas a chromatograph is a system for performing chromatography. The chart displaying the time dependent change in signal intensity as a result of the separation is called a chromatogram.

Figure 1.1 The figure shows separation of solutes A and B present in a mixture by chromatography. A continuous flow of solvent carries a mixture of solutes A and B. (a) The solvent carries the two solutes down the column. (b) After some time, solute B is moving at a much faster rate than A. (d) Finally, solute B emerges first, while solute A finally emerges in (e).

A more fundamental classification of chromatographic methods is based on the types of mobile and stationary phases and the mechanisms of retention (the manner in which the analyte interacts with the stationary phase).

Based on stationary phase, there are two subgroups of gas chromatography (GC) – gas-solid chromatography and gas-liquid chromatography. In similar fashion, liquid chromatography can be sub-grouped into liquid-solid chromatography and liquid-liquid chromatography.

The chromatographic techniques use several types of interactions to separate solutes. Based on types of interactions there are several chromatographic modes such as partition, adsorption, size exclusion and ion exchange.

1.1 Adsorption and partition chromatography

Chromatographic techniques can be classified into two main categories: partition and adsorption chromatography - depending on how the solute molecules interact with the stationary phase. In adsorption chromatography, the stationary phase is a solid material (known as adsorbent) on which the solute molecules are adsorbed. The adsorption process, which involves weak, non-ionic attractive van der Waals interaction and hydrogen bonding, occurs at specific adsorption sites. The adsorbents can be polar or non-polar molecules. Most commonly used adsorbents are polar, either acidic (e.g. silica) or basic (e.g. alumina) molecules. The adsorptive effects of the polar adsorbents are often due to the presence of hydroxyl groups and the formation of hydrogen bonds with the solute molecules. The strength of these bonds and

Analyte

An analyte (or solute) is the chemical entity being analyzed.

Detector

A detector is a device that monitors the concentration of the analyte.

Eluent

The eluent is an alternative term used for the mobile phase.

Effluent

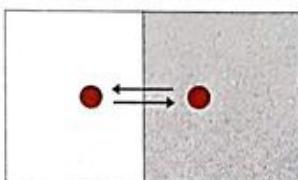
The effluent is the mobile phase that exits the column.

hence the degree of adsorption increases as the polarity of the solute molecule increases. Charcoal is a non-polar adsorbent that binds large and non-polar molecules.

In this chromatography, mobile phase is either a liquid (liquid-solid chromatography) or a gas (gas-solid chromatography). The separation mechanism depends upon differences in polarity between the different solute molecules. The more polar a molecule, the more strongly it will be adsorbed by a polar stationary phase. Similarly, the more non-polar a molecule, the more strongly it will be adsorbed by non-polar stationary phase. Less tightly bound compounds will be eluted out by the mobile phase earlier than the tightly bonded ones.

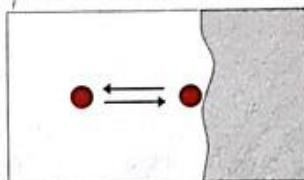
- ✓ In partition chromatography, the distribution of solutes between the two liquid phases is based primarily on solubility differences. It can be subdivided into liquid-liquid chromatography, in which the liquid stationary phase is attached to a supporting matrix by purely physical means or by covalent interactions.

Partition chromatography



Separation is based on
solute partitioning
between two liquid phases.
(Relative solubility)

Adsorption chromatography



The stationary phase is a solid.
Separation is due to a series of
adsorption/desorption steps.

old school

Partition chromatography

Martin and Synge in 1944 developed the methodology of partition chromatography and were honoured with Nobel Prize. In this chromatographic technique, separation is based on solute partitioning between stationary liquid phase and the mobile liquid phase. Substances which are more soluble in the mobile phase will pass rapidly through the system while those which favour the stationary phase will be retarded. It is of two types:

- ✓ Normal phase partition chromatography: Stationary phase is polar (usually water) and the mobile phase is relatively non-polar organic solvent such as hexane, heptane, ethyl-acetate. During elution least polar analyte (solute) is eluted first and the most polar last.
- ✓ Reverse phase partition chromatography: Stationary phase is non-polar and mobile phase is relatively polar. In this case, most polar solutes elute first and least polar elute last.

- ✓ Paper chromatography is a type of partition chromatography. In paper chromatography, the end of the paper is dipped into a solvent mixture consisting of aqueous and organic components. The solvent soaks into the paper by capillary action because of the fibrous nature of the paper. The aqueous component of the solvent binds to the cellulose of the paper and thereby forms a stationary phase with it. The organic component of the solvent continues migrating, thus forming the mobile phase. The rates of migration of the various substances being separated are governed by their relative solubilities in the polar stationary phase and the non-polar mobile phase. During the separation process, a given solute is distributed between the mobile and stationary phases according to its partition coefficient. The nonpolar molecules move faster than polar ones. The migration rate of a substance during paper chromatography is usually expressed as the dimensionless term R_f (Relative front), which is the ratio of the distance traveled by substance and solvent front.

R_f R_s S . front

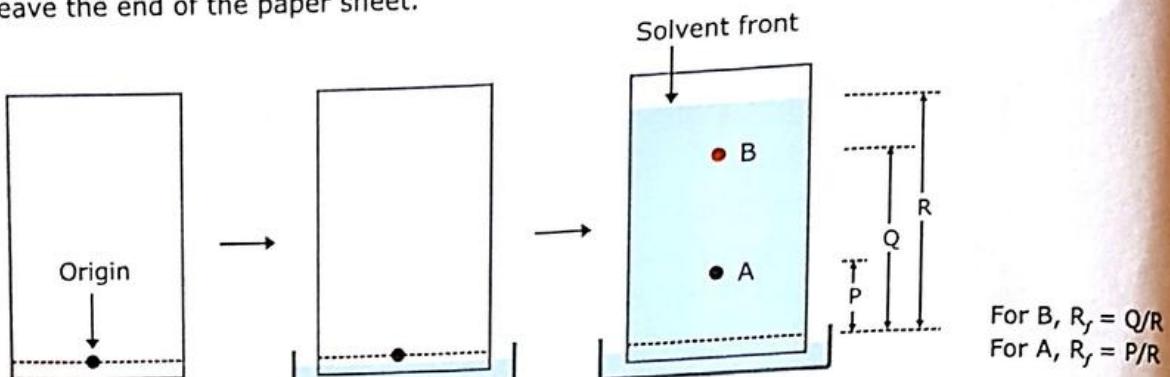
Chromatography

(3)

Relative front

$$R_f = \frac{\text{Distance traveled by substance}}{\text{Distance traveled by solvent front}}$$

Naturally the R_f can be calculated only in those instances when the solvent is not allowed to leave the end of the paper sheet.



Paper chromatograms can be developed either by ascending or descending solvent flow. There is little difference in the quality of the chromatograms and the choice is usually a matter of personal preference. Descending chromatography has two advantages: 1. it is faster because gravity aids the flow and 2. for quantitative separations of materials with very small R_f values, which therefore require long runs, the solvent can run off the paper.

1.2 Size exclusion chromatography

Size exclusion chromatography or molecular sieve chromatography separates molecules on the basis of size and shape. A column matrix filled with porous gel beads, made up of an insoluble and hydrated polymer such as polyacrylamide (Sephadex or BioGel P) or dextran (Sephadex) or agarose (Sephadex) acts as a stationary phase. Size exclusion chromatography includes: gel permeation chromatography and gel filtration chromatography. Gel permeation chromatography uses organic mobile solvent while gel filtration chromatography uses aqueous mobile solvent to separate and characterize molecules.

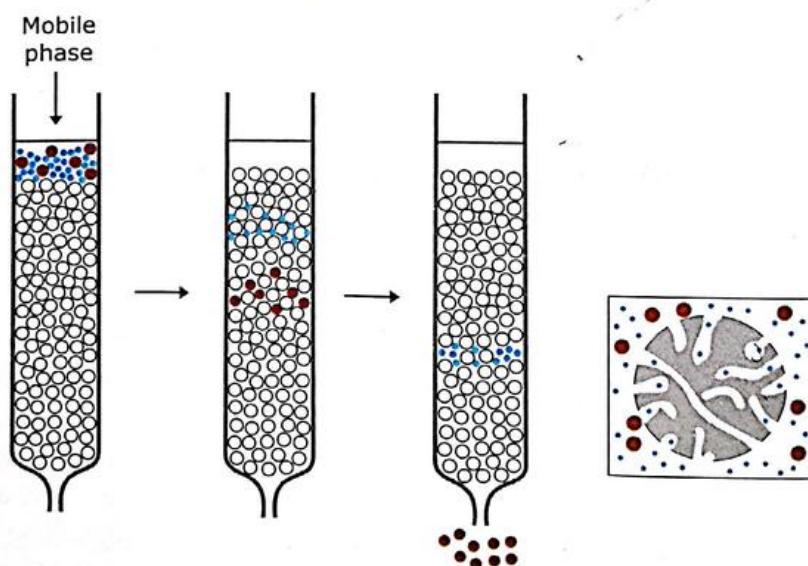


Figure 1.2 When the sample passes through the porous gel beads, small sample molecules can enter the pores, causing them to flow slower through the column. Large molecules which cannot enter the pores, pass through the column at a faster rate than the smaller ones. Correct pore sizes and solvents are crucial for a good separation.

The basis of size exclusion chromatography is very simple. If a solution containing molecules of various sizes is passed through the column, molecules smaller than the pores can enter the pores in the beads whereas larger molecules cannot. So larger molecules move faster and elute first. Since smaller molecules can enter the pores present in the beads, they have longer path and longer retention time than larger molecules that cannot enter the pores. Thus, if a mixture of proteins is applied to a column and then washed with an appropriate buffer, the first proteins to emerge from the column are those that are too large to enter the pores of the gel beads. Other proteins are eluted in decreasing order of their molecular size. The molecular mass of the smallest molecule unable to penetrate the pores of a given gel is said to be the gel's exclusion limit. For a given sample, the distribution coefficient K_d is dependent upon its size. If the molecule is completely excluded by the pore, then $K_d=0$; whereas if it enters into the porous beads and has accessibility to inner solvent, then $K_d=1$. For all other intermediate sizes, the K_d value will lie in the range of 0 to 1.

Table 1.1 Materials commonly used for porous gel beads in size exclusion chromatography

Material and trade name	Fractionation range* (Molecular mass in Da)
Dextran ✓	
Sephadex G-10	0 - 700
Sephadex G-25	1000 - 5000
Sephadex G-50	1500 - 30,000
Sephadex G-75	3000 - 70,000
Sephadex G-100	4000 - 150,000
Sephadex G-150	5000 - 300,000
Sephadex G-200	5000 - 800,000
Polyacrylamide ✓	
Bio-gel P-2	100 - 1800
Bio-gel P-6	1000 - 6000
Bio-gel P-60	3000 - 60,000
Bio-gel P-150	15,000 - 150,000
Bio-gel P-300	60,000 - 400,000
Agarose ✓	
Sepharose 2B	$2 \times 10^6 - 25 \times 10^6$
Sepharose 4B	$3 \times 10^5 - 3 \times 10^6$
Sepharose 6B	$10^4 - 20 \times 10^6$

*The molecular mass listed are for globular proteins. ✓

Size measurements by size exclusion chromatography

In order to obtain size information about a solute from a size exclusion chromatography experiment, the column must first be characterized in terms of the volumes accessible to analytes. The total volume (V_T) of a size exclusion chromatography column is divided into three parts:

1. The volume external to the packing material i.e. void volume, V_0 ;
2. The volume contained within the porous beads that is accessible to small molecules i.e. internal volume, V_i ;
3. The volume occupied by the packing material itself i.e. bed volume, V_g .

$$\text{Therefore, } V_T = V_0 + V_i + V_g$$

The values of the V_0 and V_i are determined experimentally by measuring the elution volumes of, respectively, a large solute that is totally excluded from the interior of the porous bead and a small solute that has access to all pores of the gel bead. The elution volume of a given solute, V_e , is the volume of solvent required to elute the solute from the column after it has first contacted the gel. The elution volume, V_e , of a solute that is partially included in the pores of the gel bead can be related to the void and internal volumes of a column by the following equation:

$$V_e = V_0 + \sigma V_i \quad \text{where, } \sigma \text{ is the partition coefficient of the solute.}$$

It is the partition coefficient, σ , which describes how much of the internal volume is available for the solute ($0 < \sigma < 1$). When σ is compared with the values measured for solutes of known size, it provides information about the molecular size of an unknown solute. If a series of solutes of known size is subjected to size exclusion chromatography, a linear relationship between partition coefficient and size is observed.

1.3 Ion exchange chromatography

Ion exchange chromatography is applicable for the separation of charged molecules. In this chromatographic technique, the stationary solid phase commonly consists of an insoluble matrix with covalently attached anions or cations (called ion exchanger). Solute ions of the opposite charge in the mobile liquid phase are attracted to the ion exchanger by electrostatic forces.

✓ Ion exchanger ✓

Ion exchangers are made up of two parts - an insoluble matrix and chemically bonded charged groups within and on the surface of the matrix. An ion exchanger is classified as cationic or anionic depending on whether it exchanges cations or anions.

Cation exchanger (also called acidic ion exchanger) : It is used for cation separation.

Anion exchanger (also called basic ion exchanger) : It is used for anion separation.

Each type of exchanger is also classified as strong or weak according to the ionizing strength of the functional group. An exchanger with a quaternary amino group is, therefore, a strongly basic anion exchanger, whereas primary or secondary aromatic or aliphatic amino groups would lead to a weakly basic anion exchanger. A strongly acidic cation exchanger contains the sulfonic acid group.

Table 1.2 Commonly used ion exchangers

Name	Type	Functional group
<i>Anion exchanger</i>		
DEAE-cellulose	Weakly basic	Diethylaminoethyl (DEAE)
QAE-Sephadex	Strongly basic	Quaternary aminoethyl (QAE)
Q-Sepharose	Strongly basic	Quaternary ammonium (Q)
<i>Cation exchanger</i>		
CM-cellulose	Weakly acidic	Carboxymethyl (CM)
SP-Sepharose	Strongly acidic	Sulfopropyl (SP)
SOURCE S	Strongly acidic	Methylsulphate (S)

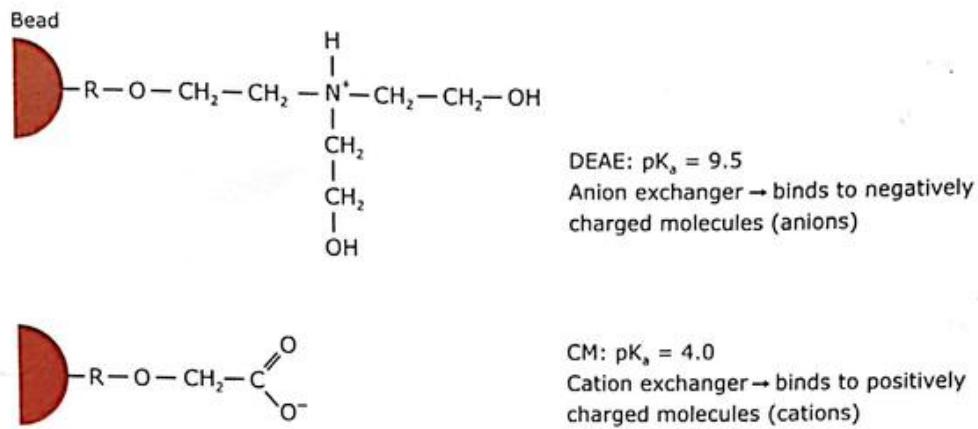


Figure 1.3 Ion exchangers – diethylaminoethyl (DEAE) and carboxymethyl (CM). The positive charge of DEAE attracts negatively charged molecules. CM is suitable for binding with positively charged molecules.

A molecular species in a given sample, which has the opposite charge to that of the charged group of the ion exchanger binds to the column. Separation of charged molecules occurs because different molecules have different degree of interaction with the ion-exchanger due to differences in their charges, charge density and distribution of charge on their surfaces. If a protein has a net positive charge at pH 7, it will usually bind to ion exchanger (cation exchanger) containing carboxylate groups, whereas a negatively charged protein will not. The bound molecules can be eluted by altering the pH of the eluting buffer or by increasing the salt concentration of the eluting buffer. A positively charged protein bound to cation exchanger can be eluted by increasing the salt concentration in the eluting buffer because cations present in the buffer compete with positively charged groups on the protein for binding to the ion exchanger. Proteins that have a low density of net positive charge will tend to emerge first, followed by those having a higher charge density.

Ion exchange chromatography can also be used to separate DNA from a cell extract. It is based on the interaction between anion exchanger (DEAE) and negatively charged phosphates of the DNA backbone. The anion-exchange resin consists of silica beads with a high charge density. When the cell extract passes through the column, all the negatively charged molecules bind to the resin and retained in the column. If the salt solution of gradually increasing concentration is passed through the column, the different types of molecule will elute in the sequence protein, RNA and finally DNA.

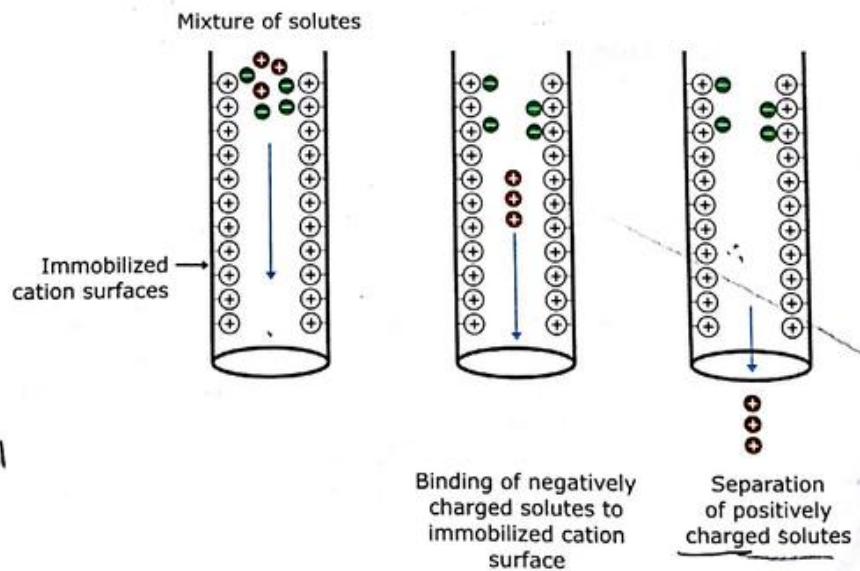


Figure 1.4 Column materials used for ion exchange chromatography contain charged groups covalently linked to the surface of an insoluble matrix. The charged groups of the matrix can be positively or negatively charged. Anion exchangers have positively charged groups that will attract negatively charged anions. When a mixture of solutes is loaded into the anion exchanger, negatively charged solutes bind to the exchanger.

anion exchanger + charged
cation exchanger - charged

cation exchanger
(cation exchanger)

The choice of ion exchanger for the purification of a protein largely depends on the isoelectric point, pI , of the protein. At a pH value above the pI of a protein, it will have a net negative charge and adsorb to an anion exchanger. Below the pI , the protein will adsorb to a cation exchanger.

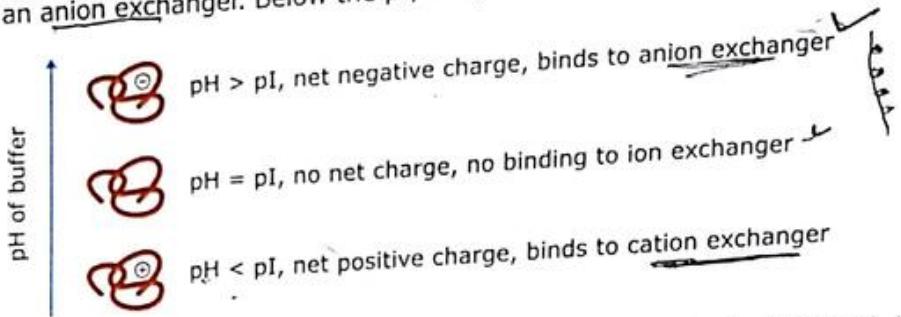


Figure 1.5 The net charge on a protein is influenced by the pH of its solvent. At $pH=pI$, the protein has zero net charge and, therefore, will not bind to a cation exchanger or an anion exchanger stationary phase. Adjusting the pH above or below the pI of the protein will lead to a net charge and protein binding to either an anion exchanger ($pH > pI$) or a cation exchanger ($pH < pI$) stationary phase.

For example, if the pI is 4 then, in most cases, it is advisable to choose an ion exchanger which binds to the protein at a $pH > 4$. Since at $pH > 4$ this protein is negatively charged, the ion exchanger has to be an anion exchanger, e.g. DEAE. One could also use a $pH < 4$ and a cation exchanger, but many proteins are not stable or aggregate under these conditions. If, in contrast, the protein we want to purify has a $pI = 10$, it is positively charged at a pH around 7. Thus, in general for this protein type we have to choose a cation exchanger, which is negatively charged at neutral pH.

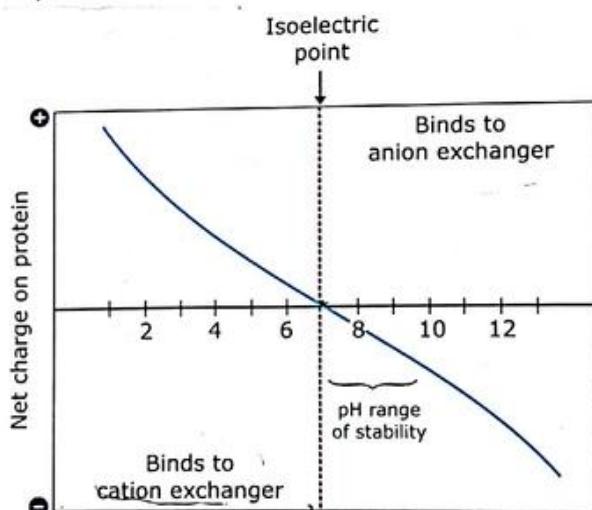


Figure 1.6 Diagram shows how the net charge of a hypothetical protein changes as a function of pH. Below the isoelectric point, the molecule has a net positive charge and would be bound to a cation exchanger. Above the isoelectric point, the net charge is negative, and the protein would bind to an anion exchanger. Superimposed on this graph is the pH range of stability for the hypothetical protein. The range of stability refers to the pH range in which the biomolecule is not denatured. Because it is stable in the range of pH 7.0-9.0, the ion exchanger of choice is an anionic exchanger. In most cases, the isoelectric point of the protein is not known. The type of ion exchanger must be chosen by trial and error.

1.4 Affinity chromatography

Affinity chromatography is a technique enabling purification of a biomolecule with respect to biological function or individual chemical structure. The substance to be purified is specifically and reversibly adsorbed to a ligand (binding substance), immobilized by a covalent bond to

a chromatographic bed material (matrix). Samples are applied under favourable conditions for their specific binding to the ligand. Substances of interest are consequently bound to the ligand while unbound substances are washed away. Recovery of molecules of interest can be achieved by changing experimental conditions to favour desorption.

A biospecific ligand that can be attached to a chromatography matrix covalently is one of the requirements for successful affinity purification. The binding between the ligand and target protein molecules must be reversible to allow the proteins to be removed in an active form. After washing away the contaminants, the coupled ligand must retain its specific binding affinity for the target proteins. Some examples of biological interactions that are usually used in affinity chromatography are listed in table 1.3.

Table 1.3 Typical biological interactions used in affinity chromatography.

<i>Types of ligand</i>	<i>Target molecules</i>
Enzyme ✓	Substrate analogue, inhibitor, cofactor
Antibody ✓	Antigen
Lectin ✓	Polysaccharide, glycoprotein, cell surface receptor, cell
Nucleic acid ✓	Complementary base sequence, nucleic acid binding protein
Hormone ✓	Receptor
Avidin ✓	Biotin
Calmodulin	Calmodulin-binding molecule
Poly(A)	RNA containing poly(U) sequences
Glutathione	Glutathione-S-transferase or GST fusion proteins
Proteins A and G	Immunoglobulins
Metal ions	Poly (His) fusion proteins, native proteins with histidine,

For example, the eukaryotic mRNA with poly (A) tail can be separated from other types of RNA molecules by oligo (dT)-cellulose affinity chromatography. Poly (A) tails form stable interaction with short chains of oligo (dT) that are attached to the support matrices. High salt is added to the chromatography buffer to stabilize the nucleic acid duplexes as only a few dT-A base pairs are formed. A low-salt buffer is used after non-polyadenylated RNAs have been washed from the matrix. This buffer helps to destabilize the double-stranded structures and elute the poly (A) RNAs from the resin.

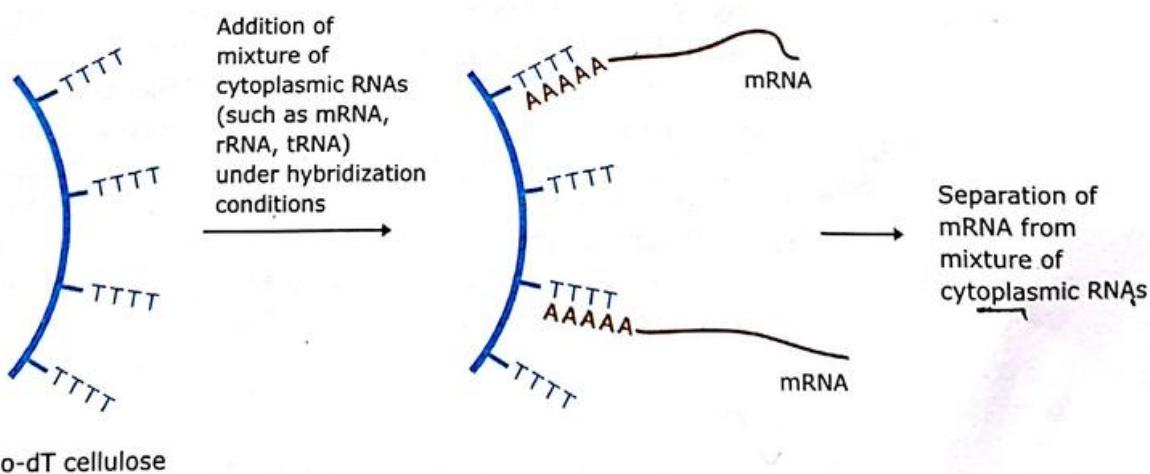


Figure 1.7 Isolation of mRNA by affinity chromatography. Messenger RNA is isolated from the cytoplasmic mixture of RNAs using the oligo (dT)-cellulose in a column. Cytoplasmic RNAs such as tRNA and rRNA which are not bound to the matrix beads can wash away and then mRNA can be eluted from the column using a low-salt buffer.

DNA affinity chromatography

DNA affinity chromatography facilitates the purification of sequence-specific DNA-binding proteins. In this method a double-stranded oligonucleotide of the correct sequence is chemically synthesized and linked to an insoluble matrix such as agarose. The matrix with the oligonucleotide attached is then used to construct a column that selectively binds proteins that recognize the particular DNA sequence. The whole process is explained in the following figure 1.8.

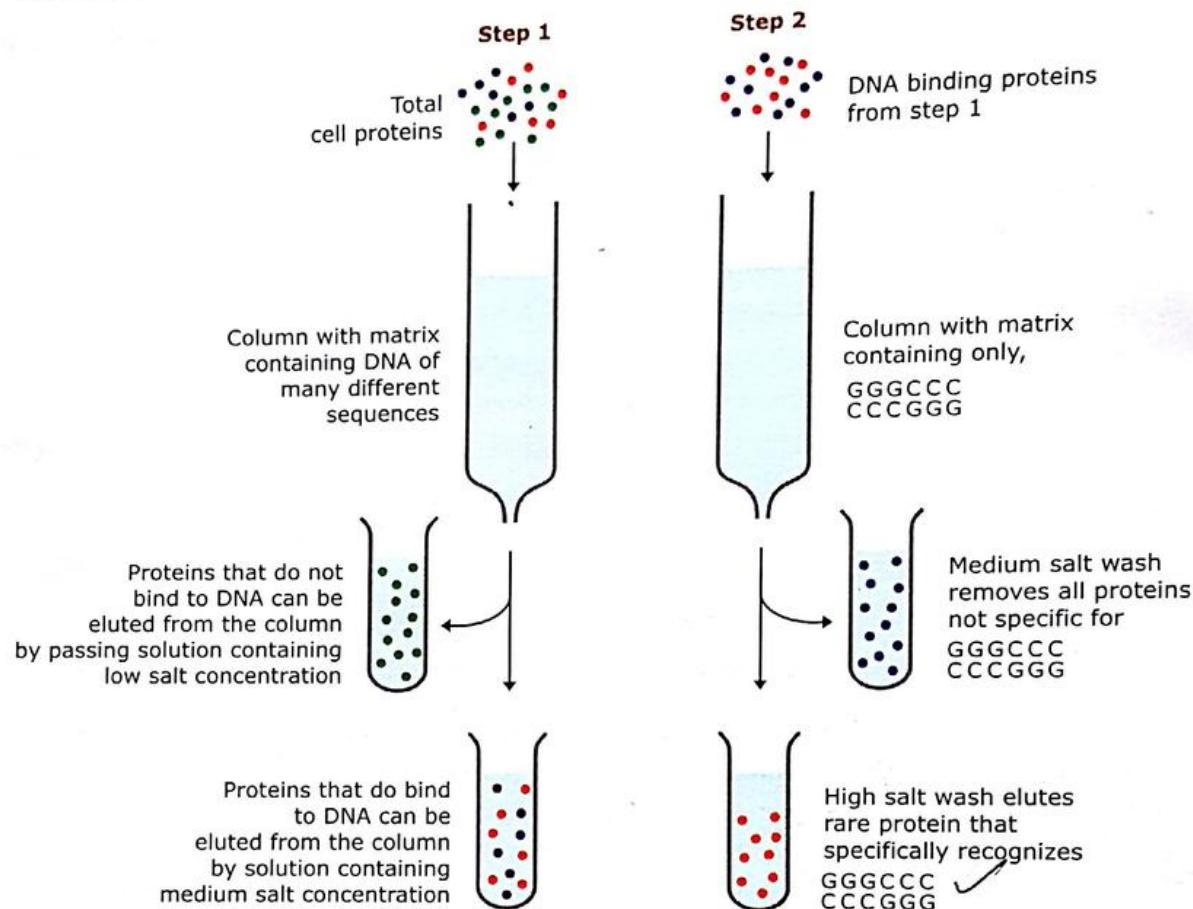


Figure 1.8 DNA affinity chromatography. In the first step, total cell proteins are passed through the column containing a huge number of different DNA sequences. In this process, the proteins that can bind DNA are separated from the remainder of the cellular proteins. Most sequence-specific DNA-binding proteins have a weak (nonspecific) affinity for bulk DNA and are, therefore, retained on the column. This affinity is largely due to ionic attractions, and the proteins can be washed off the DNA by a solution that contains a moderate concentration of salt. In the second step, the mixture of DNA-binding proteins is passed through a column that contains only DNA of a particular sequence. Typically, all the DNA-binding proteins will stick to the column, mostly by nonspecific interactions. These are again eluted by solutions of moderate salt concentration, leaving in the column only those proteins that bind specifically and, therefore, very tightly to the particular DNA sequence. These remaining proteins can be eluted from the column by solutions containing a very high concentration of salt.

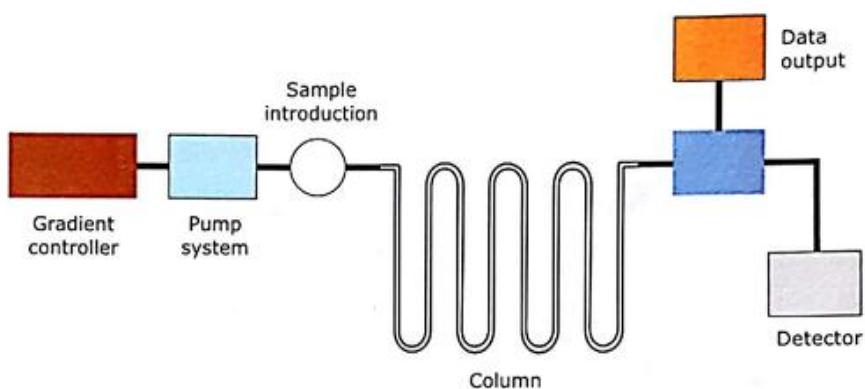
1.5

High performance liquid chromatography

High performance liquid chromatography (HPLC) is a column chromatography. It is liquid chromatographic technique i.e. mobile phase is liquid. Instead of a solvent (mobile phase) being allowed to drip through a column under gravity, it is forced through under high pressure.

It yields high performance and high speed compared with traditional column chromatography because the mobile phase is pumped with high pressure. The stationary phase may be solid or liquid. The HPLC can function in several chromatographic modes: adsorption, partition, ion exchange and size exclusion. Actually, some of these modes appear to cause separation in combination.

If the stationary phase is solid, the technique is called adsorption chromatography. If the stationary phase is a liquid, the sample is partitioned between the stationary and mobile liquid phase. This is called partition chromatography. There are two variants of partition chromatography depending on the relative polarity of the solvent and the stationary phase. In normal phase, the stationary phase is polar, and the mobile phase is non-polar. In reversed phase, the stationary phase is non-polar, and the mobile phase is polar. Reversed phase HPLC is the most commonly used forms of HPLC.



HPLC is basically a highly improved form of liquid column chromatography. Instead of a solvent being allowed to drip through a column under gravity, it is forced through under high pressures of up to 400 atmospheres. That makes it much faster. It also allows us to use very small size particles as column packing material which gives a much greater surface area for interactions between the stationary phase and the molecules flowing past it. This allows a much better separation of the components of the mixture.

The advantages of HPLC are the result of two major advances: 1. the development of stationary supports with very small particle sizes and large surface areas and 2. the improvement of elution rates by applying high pressure to the solvent flow.

Figure 1.9 A schematic diagram of a typical high-pressure liquid chromatograph. The basic components are a gradient controller, high-pressure pump, packed column, detector and recorder. A computer is used to control the process and to collect and analyze data. The purpose of the pump is to provide a constant, reproducible flow of solvent through the column. Two types of pumps are available – constant pressure and constant volume. HPLC columns are prepared from stainless steel or glass-Teflon tubing. Typical column inside diameters are 2.1, 3.2 or 4.5 mm for *analytical* separations and up to 30 mm for *preparative* applications. The length of the column can range from 5 to 100 cm, but 10 to 20 cm columns are common.

The parameters used to describe a HPLC column refer to the nature, type and size of its packing material and the dimensions of the column used. In all forms of chromatography, a measure of column efficiency is resolution. Resolution indicates how well solutes are separated. In HPLC, the increased resolution, as compared to classical column chromatography, is primarily the result of adsorbents of very small particle sizes and large surface areas. Since the smaller the particle size, the lower the flow rate; therefore, it is not feasible to use very small gel beads in liquid column chromatography as low flow rates lead to increased analysis time. In HPLC, increased flow rates are obtained by applying a pressure difference across the column. A combination of high pressure and adsorbents of small particle size leads to the high resolving power and short analysis time characteristic of HPLC.

1.6 Gas chromatography

In gas chromatography (GC), the mobile phase is a carrier gas. The stationary phase is either a solid adsorbent termed gas-solid chromatography (GSC), or a liquid on an inert support termed gas-liquid chromatography (GLC). This technique is used to analyze volatile

substances in the gas phase. Its use is, therefore, confined to analytes that are volatile, but thermally stable. It is the only form of chromatography that does not utilize the mobile phase for interacting with the analyte. GSC is based on a solid stationary phase in which retention of analytes occurs because of physical adsorption. In GLC, the analyte is partitioned between a gaseous mobile phase and a liquid phase immobilized on the surface of an inert solid packing or on the walls of a capillary tube.

The mobile-phase gas in GC is called the carrier gas. Different types of carrier gases are used in the GC. Carrier gas must be dry, free of oxygen and chemically inert. Helium, nitrogen, argon and hydrogen are generally used as carrier gases depending upon the desired performance and the detector being used. These gases are available in pressurized tanks. Pressure regulators, gauges, and flow meters are required to control the flow rate of the gas.

References

- Berg JM, Tymoczko JL and Stryer L (2007), *Biochemistry*, 6th ed. W.H. Freeman and Company.
- Chang R (2005), *Physical Chemistry for the Biosciences*, University Science Books.
- Dorsey JG, Foley JP, Cooper WT, Barford RA and Barth HG (1990), Liquid chromatography: theory and methodology. *Analytical Chemistry* 62:324R-356R.
- Freifelder D (1982), *Physical Biochemistry*, 2nd ed. W.H. Freeman and Company.
- Holler FJ and Crouch SR (2014), *Fundamentals of analytical chemistry*, 9th ed. Cengage Learning.
- Lodish H, Berk A, Kaiser CA et al (2008), *Molecular Cell Biology*, 6th ed. W.H. Freeman and Company
- Miller JJ (2005), *Chromatography: Concepts and Contrasts*, 2nd ed. John Wiley and Sons Inc.
- Skoog DA, Holler FJ and Crouch SR (2007), *Principles of instrumental analysis*, 6th ed. Thomson Publishing, USA.



Chapter 2

Electrophoresis

CElectrophoresis (*Electro* refers to the energy of electricity and *Phoresis*, from the Greek verb *phoros*, means to carry across) is a technique for separating or resolving charged molecules (such as amino acids, peptides, proteins, nucleotides, and nucleic acids) in a mixture under the influence of an applied electric field. Charged molecules in an electric field move or migrate, at a speed determined by their charge to mass ratio. According to the laws of electrostatics, an ion with charge 'Q' in an electric field of strength 'E' will experience an electric force, $F_{\text{electrical}}$

$$F_{\text{electrical}} = Q \cdot E$$

The resulting migration of the charged molecule through the solution is opposed by a frictional force

$$F_{\text{frictional}} = V \cdot f$$

where, V is the rate of migration of charged molecule and f is its frictional coefficient.

Frictional coefficient depends on the size and shape of the migrating molecule and the viscosity of the medium. In constant electric field, the force on charged molecule balances each other;

$$QE = Vf$$

so that each charged molecule moves with a constant characteristic velocity. The migration of the charged molecule in the electric field is generally expressed in terms of electrophoretic mobility (μ), which is the ratio of the migration rate of an ion to the applied electric field:

$$\mu = \frac{V}{E} = \frac{Q}{f}$$

So according to the equation, if two molecules have the same mass and shape, the one with the greater net charge will move faster towards an electrode.

Electrophoresis is of two types - moving boundary electrophoresis and zone electrophoresis. Moving boundary or free boundary electrophoresis is the electrophoresis in a free solution. It was developed by Tiselius in 1937. To separate the different charged molecules present in a mixture, the sample (dissolved in a buffer solution that serves as electrolyte and maintains the desired pH) is placed in glass tube connected to electrodes. When an electrical potential is applied across the tube, the charged molecules migrate toward one or the other electrode. Because different charged molecules migrate at different rates, a number of interfaces or boundaries are formed between the leading edge of each charged molecules and the remaining

The SI unit of
electrophoretic mobility
is $\text{m}^2 \text{ s}^{-1} \text{ V}^{-1}$



mixture. This electrophoretic technique is used for the analysis, but not for the fractionation of complex mixture.

In zone electrophoresis, sample is constrained to move in a solid support such as filter paper (paper electrophoresis) or a gel (gel electrophoresis). This technique can be used for both analytical and preparative purposes.

Gel electrophoresis

In gel electrophoresis, gel serves as molecular sieve. There are two basic types of materials used to make gels: agarose and polyacrylamide. Agarose is a natural colloid extracted from seaweed. Agarose gels have very large pore size and are used primarily to separate very large molecules with a molecular mass greater than 200 kDa. Agarose is a linear polysaccharide made up of the basic repeat unit agarobiose, which comprises alternating units of galactose and 3,6-anhydrogalactose. Agarose is usually used at concentrations between 1% and 3%. Agarose gels are used for the electrophoresis of both proteins and nucleic acids.

A polyacrylamide gel consists of chains of acrylamide monomers ($\text{CH}_2=\text{CH}-\text{CO}-\text{NH}_2$) cross-linked with N, N' -methylenebisacrylamide units ($\text{CH}_2=\text{CH}-\text{CO}-\text{NH}-\text{CH}_2-\text{NH}-\text{CO}-\text{CH}=\text{CH}_2$), the latter commonly called bis. The pore size of the gel is determined by both the total concentration of monomers (acrylamide + bis) and the ratio of acrylamide to bis. Polymerization of the acrylamide : bis solution is initiated by ammonium persulfate and catalyzed by TEMED (N, N, N', N' -tetramethylethylenediamine).

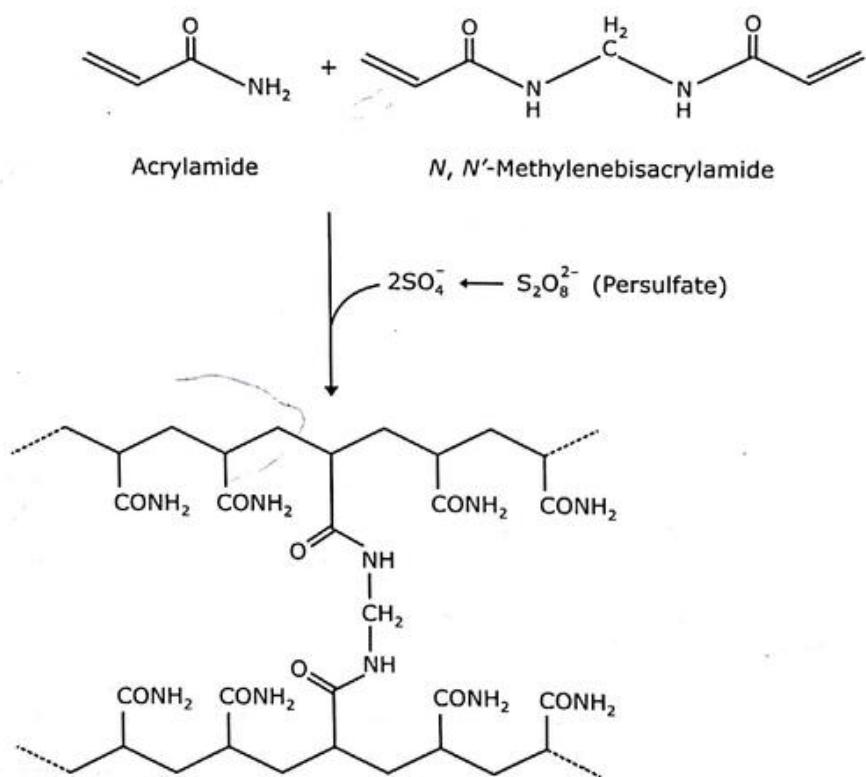


Figure 2.1 Formation of a polyacrylamide gel. In this reaction, acrylamide is used as the monomer and N, N' -methylenebisacrylamide is used as a cross-linking agent. The reaction of these two agents is begun by adding ammonium persulfate, where persulfate forms sulfate radicals that cause the acrylamide and N, N' -methylenebisacrylamide to combine. TEMED (not shown) is added to this mixture as a reagent that stabilizes the sulfate radicals. The size of the pores that is formed in the polyacrylamide gel will be related to how much N, N' -methylenebisacrylamide is used versus acrylamide. As the amount of N, N' -methylenebisacrylamide is increased, more cross-linking occurs and smaller pores are formed in the gel. As less N, N' -methylenebisacrylamide is used, larger pores are formed, but the gel also becomes less rigid.

Gel electrophoresis of proteins

SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Electrophoretic separation of proteins is most commonly performed in polyacrylamide gels. The relative movement of proteins through a polyacrylamide gel depends on the charge density (charge per unit of mass), mass (or size) and shape of the molecules. If two proteins have the same mass and shape, the one with the greater charge density will move faster through the gel. Similarly, if two proteins having same charge density and shape, the one of smaller mass or size will migrate faster than the large size protein. Shape is also a factor because compact globular proteins move more rapidly than elongated fibrous proteins of comparable mass. In SDS-PAGE, proteins are exposed to the negatively charged anionic detergent sodium dodecylsulfate (SDS) before and during gel electrophoresis. SDS binds to main chains at a ratio of about one SDS for every two amino acid residues, which imparts a large net negative charge on protein. The negative charge acquired by protein due to binding of SDS is usually much greater than the charge on native protein; this native charge thus becomes insignificant. If the protein itself has a very large positive or negative charge, this charge may not be negligible compared with the charge produced by the bound SDS. A specific example is histone H1 which has a molecular mass of about 21 kDa, but behaves in gel electrophoresis like a molecule of 30 kDa, because of its strong positive charge.

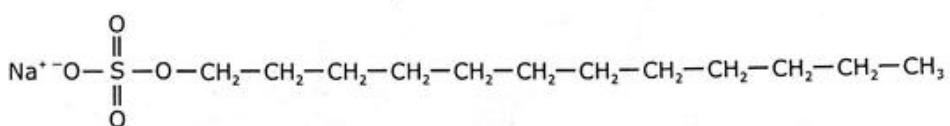


Figure 2.2 Structure of sodium dodecylsulfate (SDS).

SDS denatures proteins, causing multimeric proteins to dissociate into their subunits, and all polypeptide chains are forced into extended conformations. Along with SDS, protein is also treated with β -mercaptoethanol to break intrachain or interchain disulfide bonds. β -mercaptoethanol is a reducing agent that cleaves disulfide bonds.

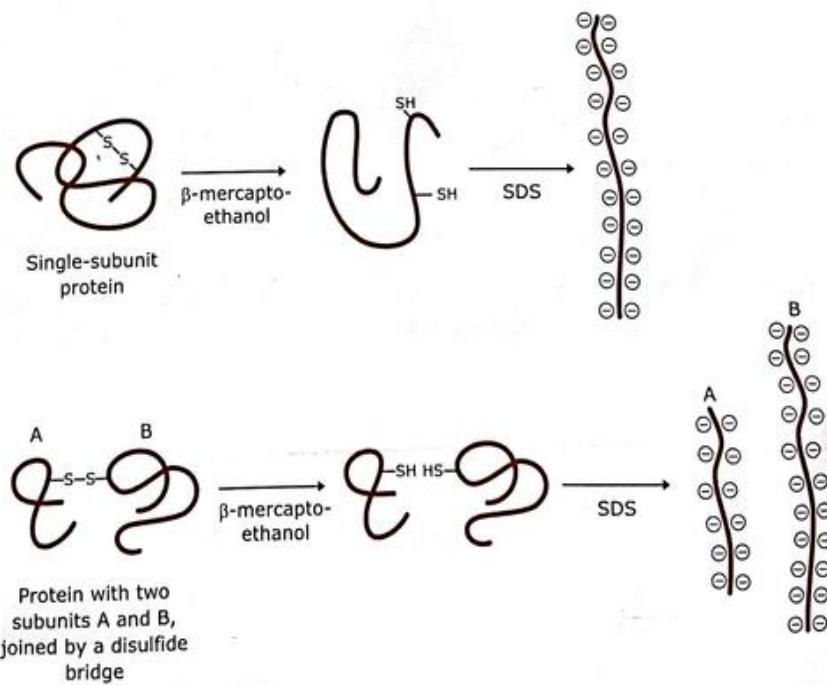


Figure 2.3 Schematic presentation of a protein in SDS. The protein is denatured and coated with a molecule with a uniform charge per unit length. The intrinsic charge of the protein is neutralised. Addition of β -mercaptoethanol reduces any disulfide bonds.

Surjan

Proteins treated with SDS have similar charge-to-mass ratio. This is because the amount of SDS bound per unit weight of protein is constant ~1.4 g of SDS per gram of protein. SDS treatment thus eliminates the effect of differences in shape and charge density, so that chain length, which reflects mass, is the sole determinant of the migration rate of proteins in SDS-PAGE.

Typically, the separating gel used is a 15% polyacrylamide gel. This gives a gel a certain pore size in which proteins of relative molecular mass 10 kDa move through relatively freely, whereas proteins of 100 kDa can only just enter the pores of the gel. This means that gels of 15% polyacrylamide are suited for separating proteins in the range of 10-100 kDa. For proteins with a molecular mass larger than 100 kDa, a larger-pored gel (e.g. 10% or even 7.5%) would be used.

Estimating the molecular weight of a protein

The molecular weight of a protein can be estimated by comparing the distance it migrates through a gel with the distances that proteins of known molecular weight migrate. The electrophoretic mobility of a protein on an SDS polyacrylamide gel is related to its molecular weight (M_r). Standard proteins of known molecular weight are subjected to electrophoresis. These proteins separate into a series of bands. A plot of $\log M_r$ of the marker proteins versus the distance migrated gives a straight line. Hence, a protein of unknown molecular weight is electrophoresed with two or more of known molecular weight, then the unknown can be calculated to an accuracy ranging between 5% to 10%.

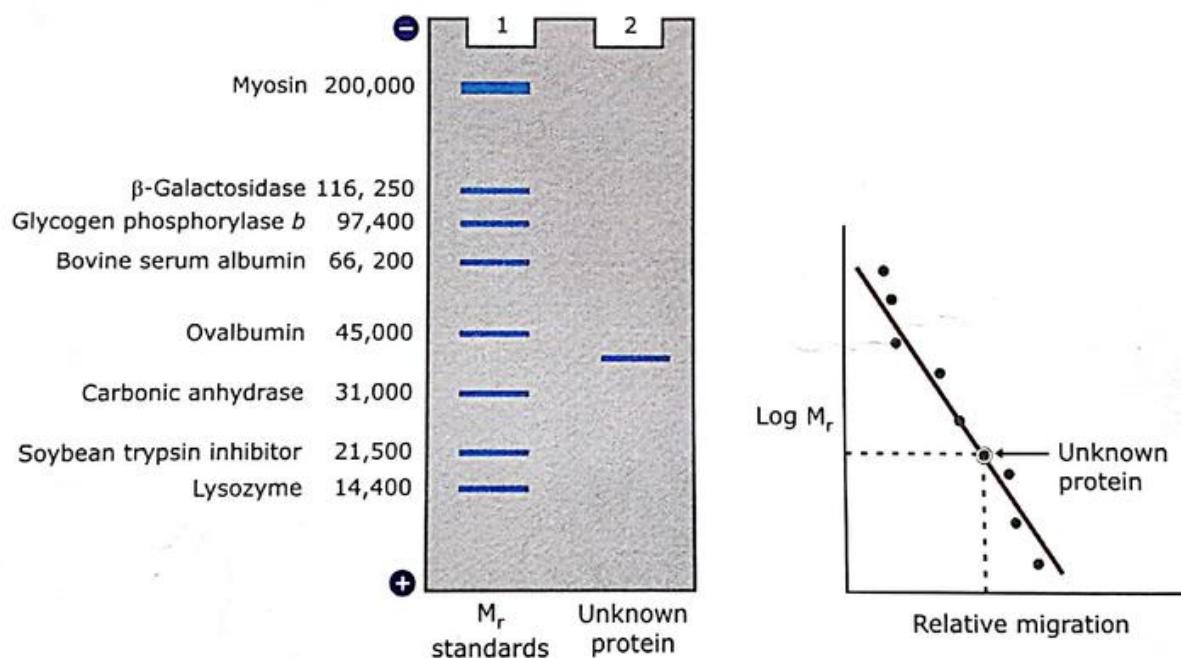


Figure 2.4 Electrophoresis can determine molecular weight. The electrophoretic mobility of many proteins in SDS-polyacrylamide gels is inversely proportional to the logarithm of their molecular weight.

✓ SDS-PAGE is rapid, sensitive and capable of a high degree of resolution. Bands resulting from electrophoretic separation can be located by a variety of techniques. Proteins are often visualized by staining. Coomassie brilliant blue is the most widely used dye which binds to proteins but not to the gel itself. Fluorescamine, a nonfluorescent molecule and silver stains are also used as alternative types of protein stain. In case of silver staining, silver ions are reduced to metallic silver on the protein, where the silver is deposited to give a black or brown band.

Discontinuous electrophoresis

In continuous gel system, separating media consists of a single gel with a uniform pH throughout. In discontinuous gel system, polyacrylamide gel is divided into three regions called the sample gel, stacking gel and separating gel. These gels can have different concentrations of the same support media, or may be completely different agents. The sample gel contains the mixture of proteins to be separated and is prepared using low concentrations of acrylamide so that pore sizes are large and do not influence the rates of migration of different size proteins. The stacking gel (or spacer gel) is similar to the sample gel. The proteins in the sample gel will concentrate into a small zone in the stacking gel before entering the separating gel.

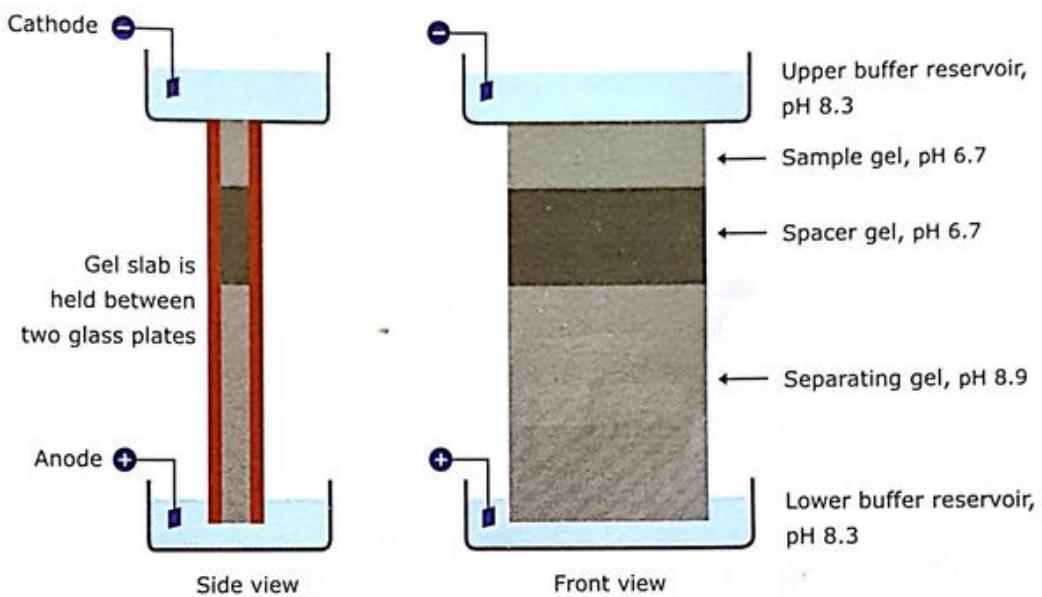


Figure 2.5 Essential components of Discontinuous (Disc) electrophoresis.

The separating gel (or running gel) differs from the other two regions due to greater concentrations of acrylamide; this results in smaller pore sizes and provides the sieving effect. The sample is usually dissolved in glycine-chloride buffer, pH 8 to 9, before loading on the gel. Glycine exists primarily in two forms at this pH, a zwitterion and an anion.

When the voltage is turned on, buffer ions (glycinate and chloride) and protein sample move into the stacking gel, which has a pH of 6.7. Upon entry into the upper gel, the concentration of glycine zwitterion increases and hence no electrophoretic mobility. Since most protein samples are still anionic at pH 6.7, they replace glycinate as mobile ions. The ion having a greater charge will move faster and is thus the leading ion, while the ion with the lesser charge will be the trailing ion. Therefore, the relative ion mobilities in the stacking gel are chloride > protein sample > glycinate. The sample will tend to accumulate and form a thin, concentrated band sandwiched between the chloride and glycinate as they move through the upper gel. Now, when the ionic front reaches the separating gel with pH 8.9 buffer, the glycinate concentration increases and anionic glycine and chloride carry most of the current. The protein sample molecules, now in a narrow band, encounter both an increase in pH and a decrease in pore size. The increase in pH would, of course, tend to increase electrophoretic mobility, but the smaller pores decrease mobility. The relative rate of movement of anions in the separating gel is chloride > glycinate > protein sample.

The separation of sample components in the separating gel occurs as described in an earlier section on gel electrophoresis. Discontinuous electrophoresis solves two problems of protein electrophoresis. It prevents aggregation and precipitation of protein during the entry from liquid sample to gel matrix and promotes well-defined bands.

Two-dimensional gel electrophoresis

Electrophoresis of all cellular proteins through an SDS gel can separate proteins having a relatively large difference in molecular mass (e.g. a 41 kDa protein from a 42 kDa protein). To separate proteins of similar mass, another physical characteristic must be exploited. In two-dimensional electrophoresis, proteins are separated in two sequential steps: first by their charge and then by their mass. In the first step, a cell extract is fully denatured by high concentrations (8M) of urea and then layered on a glass tube filled with polyacrylamide that is saturated with a solution of **ampholytes**, a mixture of polyanionic and polycationic molecules. When placed in an electric field, the ampholytes will separate and form a continuous gradient based on their net charge. The most highly polyanionic ampholytes will collect at one end of the tube, and the most polycationic ampholytes will collect at the other end. This gradient of ampholytes establishes a pH gradient. Charged proteins will migrate through the gradient until they reach their isoelectric point (pI), the pH at which the net charge of the protein is zero. This technique, called **isoelectric focusing (IEF)**, can resolve proteins that differ by only one charge unit.

Commonly used electrophoresis buffers for

Resolution of proteins

Separating gel buffer

1.5 M Tris-HCl

0.4% SDS

pH 8.8

Stacking gel buffer

0.5 M Tris-HCl

0.4% SDS

pH 6.8

Tris-glycine running buffer

25 mM Tris-base

192 mM glycine

0.1% SDS

Resolution of nucleic acids

Tris-acetate

40 mM Tris acetate

1 mM EDTA

Tris-phosphate

90 mM Tris phosphate

2 mM EDTA

Tris-borate

45 mM Tris-borate

1 mM EDTA

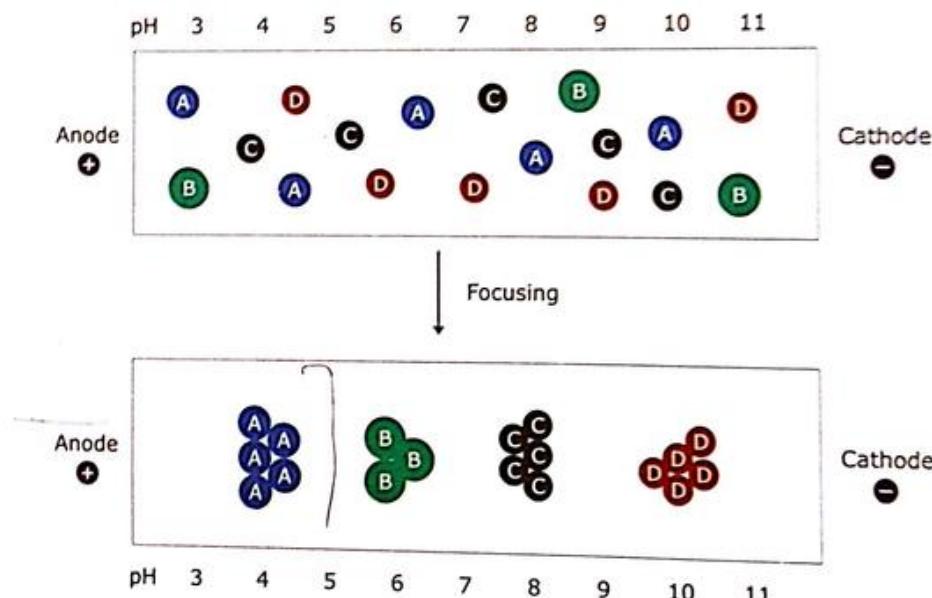


Figure 2.6 IEF. A mixture of proteins (A, B, C and D) is placed in a pH gradient in an electric field. The pH increases from acidic (pH 3) at the anode to basic (pH 10) at the cathode. When mixture of proteins is placed in a medium with a pH gradient and subjected to an electric field, they will initially move toward the electrode with the opposite charge. During migration through the pH gradient, the protein will either pick up or lose protons. As it migrates, its net charge will change. Eventually, the protein will arrive at the point where the pH gradient is equal to its pI . At this point because net charge is zero, it will stop migrating. In this way, proteins condense or are focused, into sharp bands in the pH gradient at their individual characteristic pI values. Focusing is a steady-state mechanism with regard to pH. Proteins approach their pI values at differing rates, but remain relatively fixed at those pH values for extended periods. By contrast, proteins in conventional electrophoresis continue to move through the medium until the electric field is removed.

Isoelectric focusing is carried out in an acrylamide gel. The gel contains ampholytes (for forming the pH gradient) together with 8M urea and a non-ionic detergent, both of which denature and maintain the solubility of the proteins being analysed. The denatured proteins, therefore, separate in this gel according to their isoelectric points.

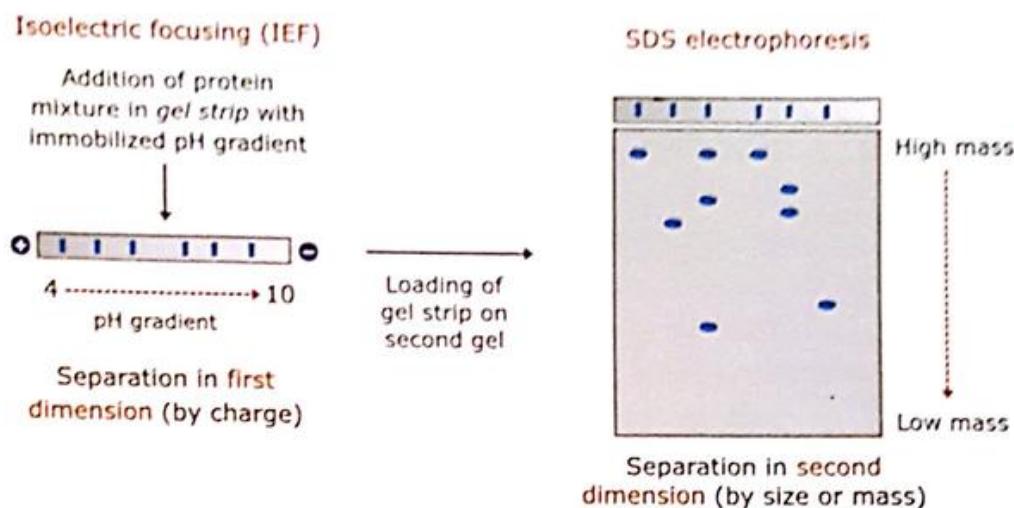


Figure 2.7 Two-dimensional gel electrophoresis. The first dimension in a 2-D gel electrophoresis experiment involves the separation of proteins according to their isoelectric point (pI) by isoelectric focusing (IEF). IEF works by applying an electric field to protein within a pH gradient. The proteins separate as they migrate through the pH gradient in response to the applied voltage. When a protein reaches a pH value that matches its pI, its net electrical charge becomes neutral and stops migrating. In this way, each protein in a sample becomes focused according to its pI. The resulting gel strip is applied to an SDS-polyacrylamide gel and the proteins are separated into bands on the basis of their masses (size).

Proteins that have been separated on an IEF gel can then be separated in a second dimension based on their size or mass. To accomplish this, the IEF gel is extruded from the tube and placed lengthwise on a second polyacrylamide gel, this time formed as a slab saturated with SDS. When an electric field is imposed, the proteins will migrate from the IEF gel into the SDS slab gel and then separate according to their mass. The sequential resolution of proteins by their charge and mass can achieve excellent separation of cellular proteins.

Isoelectric focusing is an electrophoretic method in which proteins are separated on the basis of their pIs. It makes use of the property of proteins that their net charges are determined by the pH of their local environments. Proteins carry positive, negative or zero net electrical charge, depending on the pH of their surroundings.

The net charge of any particular protein is the sum of all of its positive and negative charges. These are determined by the ionizable acidic and basic side chains of the constituent amino acids and prosthetic groups of the protein. If the number of acidic groups in a protein exceeds the number of basic groups, the pI of that protein will be at a low pH value and the protein is classified as being acidic. When the basic groups outnumber the acidic groups in a protein, the pI will be high with the protein classified as basic. Proteins show considerable variation in isoelectric points, but pI values usually fall in the range of pH 3-12 with mostly having pIs between pH 4 to 7.

Proteins are positively charged in solutions at pH values below their pI and negatively charged above their isoelectric points. Thus, at pH values below the pI of a particular protein, it will migrate toward the cathode during electrophoresis. At pH values above its pI, a protein will move toward the anode. A protein at its isoelectric point will not move in an electric field.

Blue Native PAGE

Blue native polyacrylamide gel electrophoresis (BN-PAGE) is a technique that allows separation of proteins present in multiprotein complexes in a native conformation with a higher resolution. By this method, proteins are separated according to their size and shape in a polyacrylamide gel. It has a higher resolution than gel filtration or sucrose density ultracentrifugation and can be used for protein complexes from 10 kDa to 10 MDa.

It is a native protein separation method that relies on the anionic dye *Coomassie Brilliant Blue* to confer negative charge. The anionic dye binds to proteins because of hydrophobic properties. Binding a large number of dye molecules imposes a net negative charge on the proteins that causes even basic proteins to migrate to the anode at pH 7.5 during BN-PAGE. This dye is also sufficiently soluble in water. Proteins are separated according to the size in acrylamide gradient gels. Protein migration gradually decelerates with running distance and with decreasing pore size of the gradient gel. Individual proteins must stop almost completely when they approach their size-dependent specific pore-size limit. Because negatively charged protein surfaces repel each other, the tendency of proteins to aggregate is considerably reduced. Furthermore, binding of dye makes proteins also water-soluble. This means that no detergent is required in the BN gels once Coomassie dye bind protein surfaces. Therefore, the risk of denaturation of detergent-sensitive proteins is minimized during BN-PAGE.

2.1 Immunoblotting

Separation of a mixture of proteins by electrophoretic techniques usually results in a complex pattern of protein bands or zones. Specific proteins can often be identified using an *immunoblotting* technique (also known as Western blotting). This technique requires an antibody against the test protein. After the initial separation by electrophoretic technique in a gel, the proteins are transferred (or blotted) from the gel to a membrane, usually nitrocellulose or polyvinylidene difluoride. During the transfer, the gel is at the negative electrode (cathode) side and the membrane at the positive electrode (anode) side. Proteins that are coated with negatively charged SDS will move from the negative side, the gel, to the positive side, the nitrocellulose.

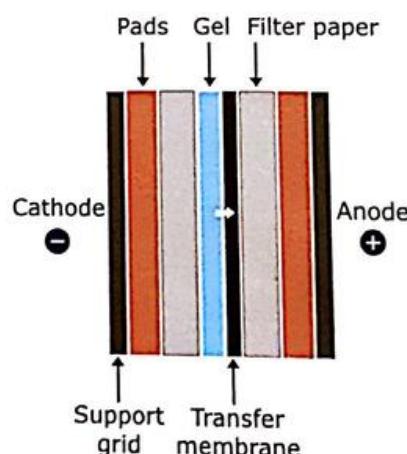


Figure 2.8 Schematic diagram showing the position of the gel, transfer membrane and direction of protein in relation to the electrode position.

Similar to ELISA, the nitrocellulose membrane that contains the resolved proteins will be first incubated with an unrelated protein to block nonspecific protein binding sites on the membrane, and then treated with a suitable primary antibody that is specific for the protein of interest.

Excess antibodies are then washed from the membrane and the bound antibody, which remains, is detected using a secondary antibody against the first. The secondary antibodies are conjugated with fluorescent or radioactive labels or enzymes that give a subsequent reaction with an applied reagent, leading to a coloring or emission of light, enabling detection.

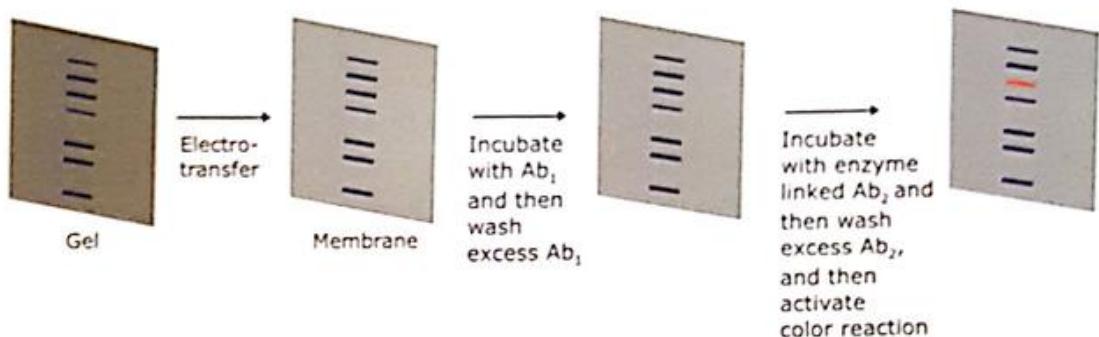


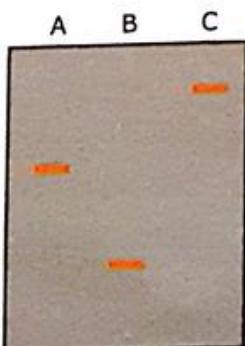
Figure 2.9 Immunoblotting using antibodies to detect specific proteins. Proteins are separated through a polyacrylamide gel, prior to blotting onto a nitrocellulose membrane. Non-specific protein binding sites are blocked on the membrane - using solubilized milk powder - before the primary antibody (Ab_1) is added. The primary antibody will specifically bind to the antigen against which it was raised. A labeled secondary antibody (Ab_2) is then added to detect the location of the primary antibody. The secondary antibody is often labeled with an enzyme whose activity, in the presence of appropriate substrates, results in a color change on the membrane.

Electrophoresis of DNA

DNA molecules are negatively charged, they migrate towards the anode in the presence of electric current. Because the charge-to-mass ratio is nearly the same for all DNA molecules, hence separation of DNA fragments occurs due to differences in mass or size and shape in gel electrophoresis. There are two common types of gel used for separation of DNA molecules – polyacrylamide and agarose. The gel acts as a sieve to selectively impede the migration of the DNA in proportion to its mass or size. Larger pieces of DNA move slowly, while smaller pieces of DNA move through the gel more quickly.

Electrophoresis separates DNA molecules, not only according to their mass or size, but also according to their shape and topological properties. DNA topoisomers have the same length, but different linking numbers. Even though topoisomers have the same mass, they can be separated from each other by gel electrophoresis. The basis of this separation is that the greater the supercoiling, the more compact the shape of a covalently closed circular DNA (cccDNA). The more compact the DNA, the more easily it is able to migrate through the gel. Thus, a relaxed cccDNA migrates more slowly than a highly supercoiled form of same circular DNA.

Rate of migration of DNA through agarose gel depends on
 1. Size of DNA
 2. Conformation of DNA
 3. Agarose concentration
 4. Nature of buffer
 5. Applied voltage



Electrophoretic separation of DNA topoisomers. Band on lane C represents relaxed circular DNA, lane B highly supercoiled and lane A less supercoiled cccDNA.

DNA molecules are invisible to the naked eye, but can be seen in gels by staining them with a solution of planar aromatic cations such as *ethidium bromide*, *acridine orange* or *proflavin*. Ethidium bromide is a molecule that becomes intercalated between the stacked bases of the DNA molecule. Soaking a DNA containing gel in ethidium bromide will result in the concentration of the chemical within the DNA. Illumination of the soaked gel with ultraviolet light (260–300 nm) results in fluorescence of ethidium bromide, and the DNA shows up on the gel as a band of fluorescence (orange).

The molecular weights of the DNA of each band can also be determined from the distance of the band from the origin. The equation that describes the relation between the molecular weights (M) and the distance traveled (D) is:

$$D = a - b (\log M)$$

where, a and b are constants that are a function of the time of electrophoresis, the buffer and the gel concentration, respectively.

The molecular weights can be obtained from a plot of D versus $\log M$, as long as at least two values of M are known.

Tracking dyes

The most common means of monitoring the progress of an electrophoretic separation is by following the migration of *tracking dyes* that are incorporated into the loading buffer. Two widely used dyes displaying different electrophoretic mobilities are *bromophenol blue* and *xylene cyanol*. The dyes migrate at specific speeds in a given gel concentration and usually run ahead of the smallest fragments of DNA. They are not associated with the sample DNA, and thus they do not affect its separation. The movement of the tracking dye is monitored, and when the dye approaches the end of the gel, or the desired distance, electrophoresis is terminated.

Pulsed-field gel electrophoresis

The sizes of the DNAs that can be separated by conventional gel electrophoresis are limited to ~50 kb in size. Very large DNA fragments are unable to penetrate the pores in an agarose gel and thus cannot readily be resolved. However, larger DNA fragments can be resolved from one another by regular changes in the orientation of the electric field with respect to the gel. With each change in the electric-field orientation, the DNA must realign its axis prior to migrating in the new direction. This technique is known as *pulsed-field gel electrophoresis* (PFGE), invented by Schwartz and Cantor in 1984. With this technique DNA fragments up to 10 Mb can be separated. The method basically involves electrophoresis in agarose gel, where two electric fields are applied alternately at different angles for defined time periods. When the electrical field is applied to the gel, the DNA molecules elongate in the direction of the electrical field. The first electrical field is then switched to the second field. The DNA must change conformation and reorient before it can migrate in the direction of this field. As long as the alternating fields are equal with respect to the voltage and pulse duration, the DNA will migrate in a straight path down the gel.

There are several variants of PFGE, all of which share the essential characteristic that the orientation of the electric field is periodically changed in a manner that results in net migration of DNA. Currently available advanced PFGE instrumentation utilizes multiple electrodes arranged in a hexagonal array. Contour-clamped homogeneous electric fields (CHEF) is one of the advanced variant of PFGE. In CHEF, some of the electrodes are clamped, or held to intermediate potentials, giving homogeneous electric fields necessary for straight, distortion-

free lanes. This instrumentation allows the manipulation of pulse time, field strength and pulse angle, all of which influence the migration rate of DNA through an agarose gel and the resolution of the separation.

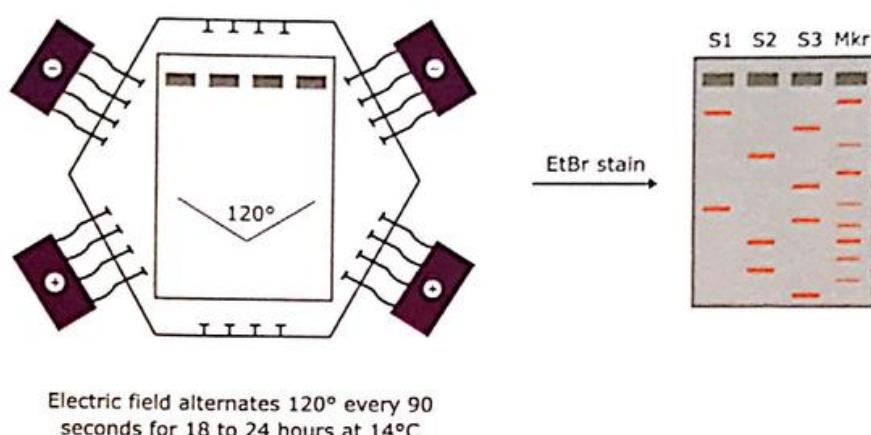


Figure 2.10 CHEF (Contoured clamped homogeneous electric field) systems use a hexagonal gel box that alters the angle of the fields relative to the agarose gel. After PFGE, DNA fragments are visualized by staining with ethidium bromide.

Electrophoresis of RNA

RNA molecules can be analyzed on both native or denaturing agarose and polyacrylamide gels. For most applications involving RNAs of less than or equal to 600 nucleotides, denaturing acrylamide gels are most appropriate. In contrast, agarose gels are generally used to analyze RNAs of more than 600 nucleotides. In non-denaturing RNA electrophoresis, due to intramolecular base pairing, RNA molecules can form extensive double-stranded structures. As a result, accurate sizing of RNA molecules is not always possible under non-denaturing conditions. Native RNA electrophoresis is, therefore, typically used to assess the overall quality of total RNA. Denaturing electrophoresis is recommended to precisely determine the size because, in this case, the migration rate of each molecule is dependent entirely on its length (size), and is not influenced by the intramolecular base pairs. The following are the two methods most commonly used today to separate denatured RNAs:

- Electrophoresis of RNA denatured with glyoxal/formamide through agarose gels.
- Pretreatment of RNA with formaldehyde and dimethylsulfoxide, followed by electrophoresis through gels containing up to 2.2 M formaldehyde.

Formaldehyde forms unstable Schiff bases with the imino-groups of guanine residues. This maintains RNA in the denatured state by intrastrand Watson-Crick base pairing. The denaturant is added to the sample before it is loaded onto the gel. Because the RNA is unable to form stable secondary structures, it migrates through agarose gels at a rate that is approximately proportional to the \log_{10} of its size. The Schiff bases are unstable and easily removed by dilution. RNA can be maintained in the denatured state only when formaldehyde is present in the buffer or gel. In case of *glyoxal*, the aldehyde groups react under slightly acidic conditions with the imino groups of guanine residues to form a cyclic compound that prevents formation of intrachain Watson-Crick base pairing.

2.3 Capillary electrophoresis

Capillary electrophoresis employs narrow-bore capillaries (typically 20-100 μm in internal diameter) to perform high efficiency separations of both large and small molecules. These separations are facilitated by the use of high voltages, which may generate an electroosmotic and electrophoretic flow of buffer solutions and ionic species, respectively, within the capillary. The properties of the separation have characteristics resemblance between polyacrylamide gel electrophoresis and high performance liquid chromatography. One advantage of using capillaries is that they reduce problems resulting from heating effects. Because of the small diameter of the tubing, there is a large surface-to-volume ratio, which gives enhanced heat dissipation. This helps to eliminate both convection currents and zone broadening owing to increased diffusion caused by heating.

The basic instrumental configuration for capillary electrophoresis is relatively simple. It includes a fused-silica capillary, two electrode assemblies, and two buffer reservoirs. The ends of the capillary are placed in the buffer reservoirs. After filling the capillary with buffer, the sample can be introduced by dipping the end of the capillary into the sample solution. In most frequently used injection mode, the tip of capillary is dip into the sample vial. The solution in vial is then pressurized to forced into the capillary.

One of the fundamental processes that drive capillary electrophoresis is *electroosmosis*. This phenomenon is a consequence of the surface charge on the wall of the capillary. The capillary surface in contact with a buffer solution is electrically charged. The fused silica capillaries that are typically used for separations have ionizable silanol groups in contact with the buffer contained within the capillary.

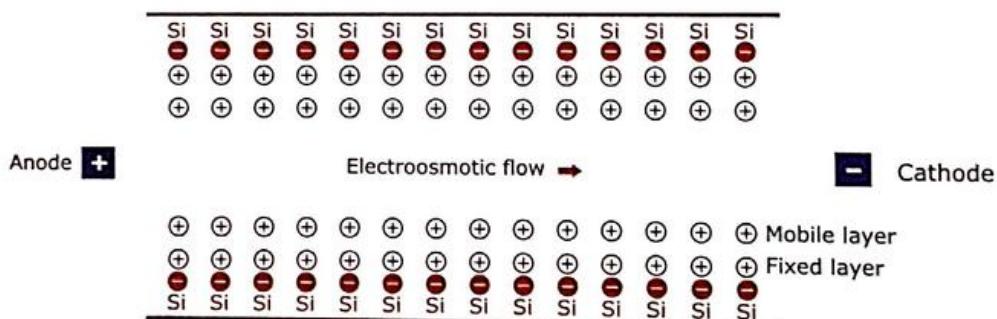


Figure 2.11 The production and effects of electroosmosis. This particular diagram shows a support that has a negatively charged interior. The interior wall of this fused-silica gel capillary has silanol groups at its surface, which can act as weak acids and form a conjugate base with a negative charge. The extent of electroosmosis, in this case, will depend on the pH of the running buffer, because this will affect the relative amount of the silanol groups that are present in their neutral acid form or charged conjugate base form.

The negatively-charged wall attracts positively-charged ions from the buffer, creating two inner layers of cations (called *diffuse double layers*) on the capillary wall. The first layer is referred to as the *fixed layer* because it is held tightly to the silanol groups. The outer, called the *mobile layer*, is farther from the silanol groups. When voltage is applied to the circuit, one electrode becomes net positive and the other net negative. The mobile cation layer is pulled in the direction of the negatively charged cathode. Since these cations are solvated, the bulk buffer solution migrates with the mobile layer, causing the *electroosmotic flow* of buffer solution.

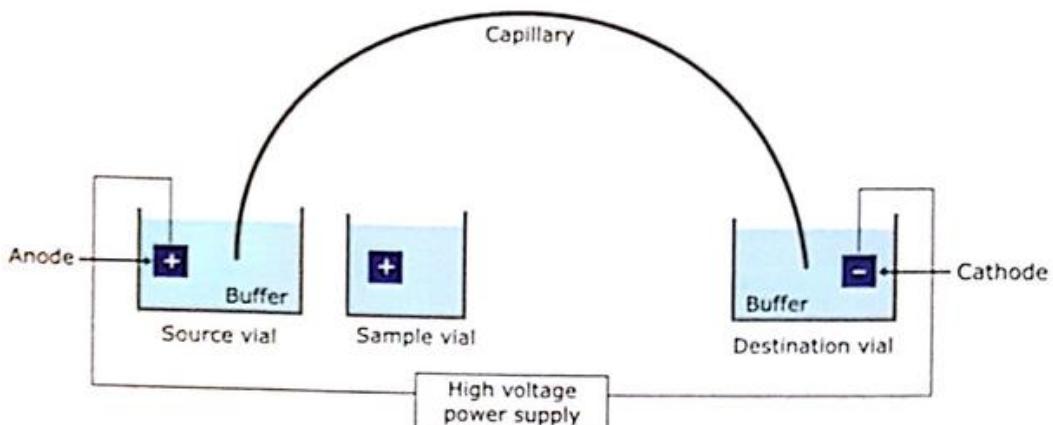


Figure 2.12 Diagram of capillary electrophoresis system.

Capillary electrophoresis generally refers to capillary zone electrophoresis, but this electro-phoretic techniques also include capillary gel electrophoresis, capillary isoelectric focusing and other classes also. Capillary zone electrophoresis, also known as free solution capillary electrophoresis, is the most commonly used technique. The separation is based on the differences in electrophoretic mobility, which is directly proportional to the charge on the molecule, and inversely proportional to the viscosity of the solvent and radius of the atom. The velocity at which the ion moves is directly proportional to the electrophoretic mobility and the magnitude of the electric field.

References

- Alberts B et al (2008), *Molecular Biology of the Cell*, 5th ed. Garland Science Publishing.
- Berg JM, Tymoczko JL and Stryer L (2007), *Biochemistry*, 6th ed. W.H. Freeman and Company.
- Brown TA (2007), *Genomes*, 3rd ed. Garland Science Publishing.
- Chang R (2005), *Physical Chemistry for the Biosciences*, University Science Books.
- Freifelder D (1982), *Physical Biochemistry*, 2nd ed. W.H. Freeman and Company.
- Garfin DE (2003), Two-Dimensional Gel Electrophoresis: An Overview. *Trends in Analytical Chemistry* 22 263-272.
- Hames BD (1998), *Gel Electrophoresis of Proteins: A Practical Approach*, 3rd ed. Oxford University Press, New York.
- Holler FJ, Crouch SR (2014), *Fundamentals of analytical chemistry*, 9th ed. Cengage Learning.
- Jorgenson JW (1986), Electrophoresis. *Analytical Chemistry* 58:743A-760A.
- Lodish H, Berk A, Kaiser CA et al (2008), *Molecular Cell Biology*, 6th ed. W.H. Freeman and Company.
- Manz A, Pamme N and Dimitri L (2004), *Bioanalytical Chemistry*, 3rd ed. Imperial College Press.

Chapter 3

Spectroscopy

Spectroscopy is the study of the interaction between electromagnetic radiation and matter. The matter can be atoms, molecules or ions. The nature of the interaction between radiation and matter may include – absorption, emission or scattering. It is the absorption, emission or scattering of radiation by matter that is used to quantitatively or qualitatively study the matter or a physical process. A study of the radiation absorbed or emitted by an atom or a molecule will give information about its identity and this technique is known as *qualitative spectroscopy*. Measurement of the total amount of radiation will give information about the number of absorbing or emitting atoms or molecules and is called *quantitative spectroscopy*.

3.1 Electromagnetic radiation

Electromagnetic radiation is a form of energy and has both electrical and magnetic characteristics. A representation of electromagnetic radiation with electric field (E) and the magnetic field (B) – at right angle to the direction of the wave – is depicted in the figure 3.1. The electric and magnetic fields in an electromagnetic wave oscillate along directions perpendicular to the propagation direction of the wave.

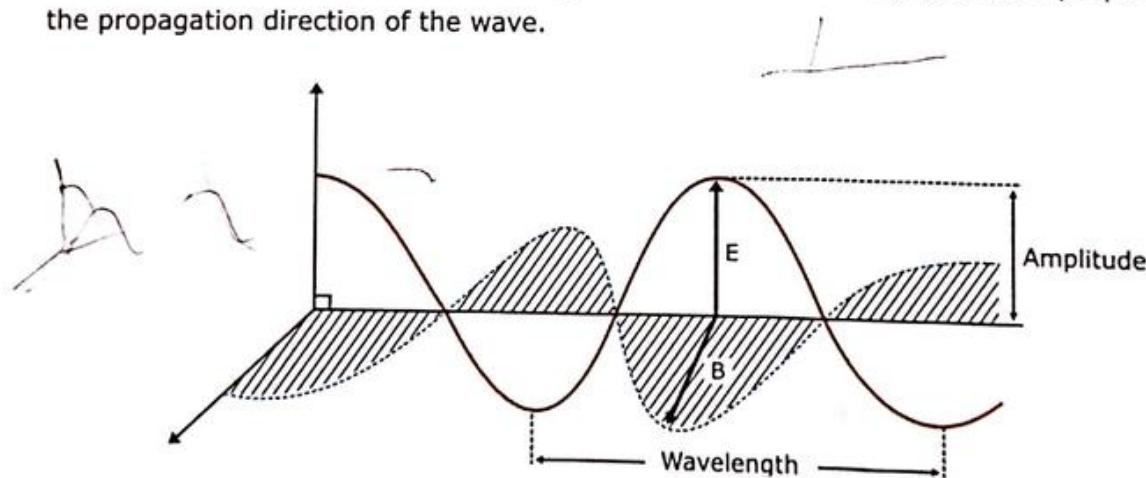


Figure 3.1 Electromagnetic radiation. A representation of electromagnetic radiation with the electric field (E) and the magnetic field (B) at right angles to the direction of the wave movement. Both fields oscillate at the same frequency.

Electromagnetic spectrum ranges from very short wavelengths (such as gamma rays) to very long wavelengths (radio waves). The visible region of the spectrum extends approximately

over the wavelength range 400–700 nm, the shorter wavelengths being the blue end of the spectrum and the longer wavelengths the red. The wavelengths between 400 and 200 nm make up the near ultraviolet region of the spectrum and wavelengths above 700 nm to approximately 2000 nm (2 μm) the infrared region.

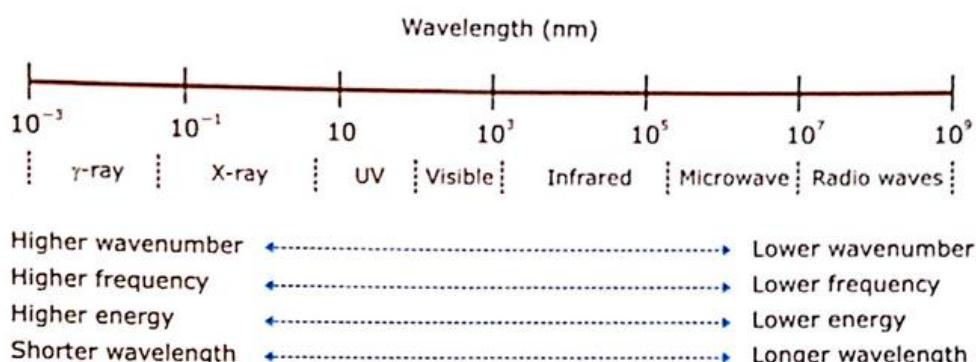


Figure 3.2 Range of electromagnetic radiation.

The energy associated with a given segment of the spectrum is related to its frequency and wavelength. Frequency (ν) is the number of wave cycles that pass through a point in one second. It is measured in Hz, where 1 Hz = 1 cycle/sec. The wavelength (λ) is the length of one complete wave cycle. It is often measured in centimeters. It is inversely proportional to the frequency (ν) and is governed by the relationship;

$$\nu = \frac{c}{\lambda} \quad \text{where, } c = \text{speed of light.}$$

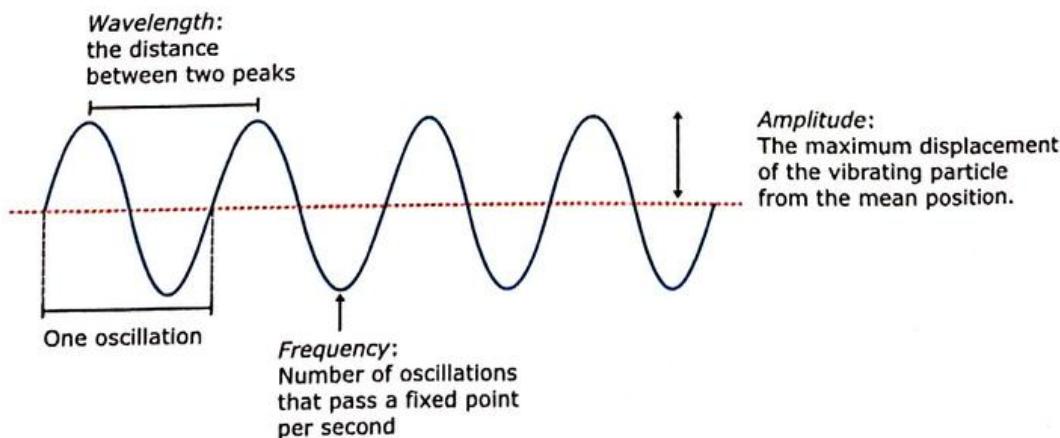
The energy is directly proportional to frequency and inversely proportional to wavelength. It is related to wavelength and frequency by the following equation:

$$E = h\nu = \frac{hc}{\lambda} \quad \text{where, Planck's constant } (h) = 6.6 \times 10^{-34} \text{ joules-sec.}$$

The radiation in the infrared region of the electromagnetic spectrum is also expressed in terms of wavenumber, rather than wavelength. It is the reciprocal of the wavelength. It is denoted by $\bar{\nu}$. Wavenumber is expressed in units of per centimeter (cm^{-1}).

$$\bar{\nu} = \frac{1}{\lambda} \quad \text{where, } \bar{\nu} \text{ is in units of } \text{cm}^{-1} \text{ and } \lambda \text{ is in units of cm.}$$

The main reason chemists prefer to use wavenumbers as units is that they are directly proportional to energy (*a higher wavenumber corresponds to a higher energy*).



and X-rays → ionization
Visible → vibration
Radio → rotation

The effect of electromagnetic radiation on interaction with matter depends on energy associated with the radiation. Very energetic radiations (such as UV and X-ray region of the spectrum) may cause an electron to be ejected from the molecule (ionization). Radiations in the infrared region of the spectrum have much less energy than radiations in the visible or UV regions of the electromagnetic spectrum. They can cause vibrations in molecules. Microwave radiation is even less energetic than infrared radiation. It can neither induce electronic transition in molecules, nor can it cause vibrations; it can only cause molecules to rotate.

Region of spectrum	Types of energy transitions
Ultraviolet visible	Electronic transition
Infrared	Molecular vibration
Microwave	Molecular rotation
Radio frequencies	Nuclear spin (in case of nuclear magnetic resonance) Electronic spin (in case of electron spin resonance)

V - ①
- V.
- R.
- N.

3.2 Types of spectroscopy

When radiation meets matter, the radiation is either scattered, emitted or absorbed. This gives rise to three principal branches of spectroscopy:

Absorption spectroscopy

Absorption spectroscopy studies radiation *absorbed* at various wavelengths. When a beam of electromagnetic radiation passes through a sample; much of the radiation passes through the sample without a loss in intensity. At selected wavelengths, however, the radiation's intensity is attenuated (decrease in number of photons). This process of attenuation is called absorption. Absorption spectroscopy can give both qualitative and quantitative information about the sample.

In absorption spectroscopy, an electromagnetic radiation is absorbed by an atom or molecule, which undergoes a transition from a lower-energy state to a higher energy or excited state. Absorption occurs only when the energy of radiation matches the difference in energy between two energy levels. The type of transition depends on the energy of electromagnetic radiation. For example, absorption of a radiation in ultraviolet and visible region promotes molecule's valence electrons to a higher-energy level. When a molecule absorbs infrared radiation, on the other hand, one of its chemical bonds experiences a change in vibrational energy.

Table 3.1 Major types of absorption spectroscopy

Electromagnetic radiation	Spectroscopic type
X-ray	X-ray absorption spectroscopy
UV/Vis	UV/Vis absorption spectroscopy
IR	Infrared absorption spectroscopy
Microwave	Microwave absorption spectroscopy
Radio wave	Electron spin resonance spectroscopy Nuclear magnetic resonance spectroscopy

Scattering spectroscopy

Scattering spectroscopy measures certain physical properties by measuring the amount of light that a substance scatters at certain wavelengths. One of the most useful applications of light scattering spectroscopy is Raman spectroscopy.

Emission spectrum
When the light emitted by a substance is passed through a prism and examined directly with a spectroscope, the spectra obtained are referred to as emission spectrum.

An absorption spectrum is a plot of the amount of light absorbed by a sample versus the wavelength of the light. The amount of light absorbed is called the absorbance.

Emission spectroscopy

Atoms or molecules that are excited to high energy levels can decay to lower levels by emitting radiation. The substance first absorbs energy and then emits this energy as light. Emission can be induced by sources of energy such as flames or electromagnetic radiations. For atoms excited by high-temperature, the light emission is commonly called atomic emission (*emission spectroscopy*); and for atoms excited with electromagnetic radiation, the light emission is called atomic fluorescence (*fluorescence spectroscopy*).

Fluorescence is an emission phenomenon. Fluorescence emission of light is observed when, after excitation by the absorption of a photon, an electron returns from the first excited state to the ground state. In the excited state, some energy is always lost by non-radiative processes (such as transitions between vibrational states). Therefore, the energy of the emitted light is always less than that of the absorbed light, and the fluorescence of a chromophore, thus, always occurs at greater wavelengths than its absorption. The difference between the absorption (excitation) and emission wavelengths is called the *Stokes shift*.

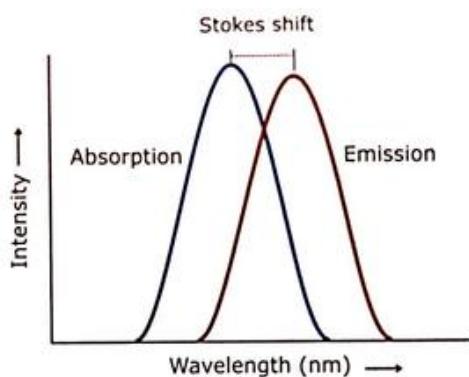


Figure 3.3 Stokes shift. It is the difference in wavelength or energy between the excitation and emission spectra of the same electronic transition. In practice, it is the difference between the excitation and emission maxima.

Principles of absorption spectroscopy

When electromagnetic radiation passes through a material, a portion of the electromagnetic radiation may be absorbed. If that occurs, the remaining radiation, when it is passed through a prism, yields a spectrum with a gap in it, called an *absorption spectrum*. The absorption spectrum is characteristic of a particular element or compound, and does not change with varying concentration. At a given wavelength, the measured absorbance has been shown to be proportional to the molar concentration of the absorbing species and the thickness of the sample the light passes through. This is known as the *Beer-Lambert law*.

Beer-Lambert law

When radiation falls on homogeneous medium, a portion of incident light is reflected, a portion is absorbed and the remainder is transmitted. The two laws governing the absorption of radiation are known as Lambert's law and Beer's law. In the combined form, they are referred to as *Beer-Lambert law*.

Lambert's law

It states that when monochromatic light passes through a transparent medium, the intensity of transmitted light decreases exponentially as the thickness of absorbing material increases.

Beer's law

It states that the intensity of transmitted monochromatic light decreases exponentially as the concentration of the absorbing substance increases.

Mathematical expression of Beer-Lambert law:

The relationship between concentration, length of the light path, and the light absorbed by a particular substance is expressed mathematically as shown below:

$$A = \log \frac{I_0}{I} = \epsilon cl$$

where, A = Absorbance

I_0 = Intensity of incident light

I = Intensity of light transmitted through the sample

ϵ = Extinction coefficient or absorption coefficient for an absorbing compound

c = Concentration of absorbing material in the sample

l = Pathlength (cm)

If the concentration is expressed in molarity, ϵ is termed as the *molar absorption coefficient*. If the concentration is expressed in g/liter, ϵ becomes the *specific absorption coefficient*. The absorbance is a dimensionless quantity. Theoretically, absorbance (A) can have any positive value; in practice, for UV and visible spectrometers, 'A' normally varies between zero and one. The ratio of the intensity of the transmitted light (I) to the intensity of the incident light (I_0) is called *transmittance* (T).

$$T = \frac{I}{I_0}$$

It measures the amount of light transmitted after passing through the medium. The smaller the transmittance, the greater the absorption of light. Transmittance is always a numerical value between zero (all light absorbed) and one (no light absorbed). It is common to convert transmittance into percent transmittance (%T).

$$\%T = \frac{I}{I_0} \times 100\% = T \times 100\%$$

Percent transmittance varies between 100% (all light transmitted) and 0% (no light transmitted). The transmittance and absorbance are inversely related. Moreover, the inverse relationship between transmittance and absorbance is not linear, it is logarithmic.

$$\begin{aligned} A &= \log_{10} \frac{1}{T} \\ &= -\log_{10} T \\ &= -\log_{10} \frac{\%T}{100} \\ &= -\log_{10} \%T + \log_{10} 100 \\ &= -\log_{10} \%T + \log_{10} 10^2 \\ &= 2 - \log_{10} \%T \end{aligned}$$

A spectrophotometer is an instrument that measures the intensity of a light beam as a function of its color, or wavelength. In absorption experiments, a spectrophotometer provides an accurate quantitative measurement of the fraction of light that passes through a defined pathlength of solution.

An instrument used to measure the absorbance by measuring the amount of light of a given wavelength that is transmitted by a sample is termed spectrophotometer. A spectrophotometer

Electrons in the vast majority of molecules fall into one of the three classes: σ electrons, π electrons and non-bonding electrons (n electrons). In chemical terms, a single bond between atoms, such as C-H, O-H only σ electrons, a multiple bond, C=C, C=N, contains π electrons in addition, while atoms such as nitrogen and oxygen possess n electrons. In general, the σ electrons are most firmly bound to the nuclei and hence require a great deal of energy to undergo transitions, while the π and n electrons require less energy, the n electrons usually requiring less than the π .

consists of a light source, a monochromator, sample holder (cuvette) and a light detector. The spectroscopic technique used to assess the concentration or amount of a given species is termed spectrometry.

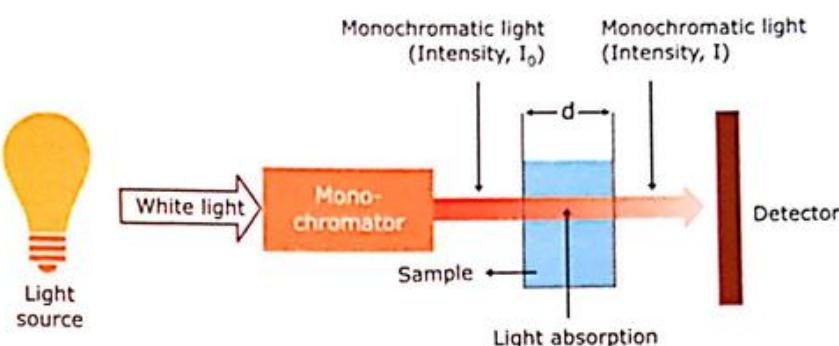


Figure 3.4 A spectrophotometer. Light from a light source passes through a monochromator for wavelength selection. Sample is contained in cuvette in a cuvette holder. Light passes through a cuvette and detected by detector.

3.3 UV/VIS absorption spectroscopy

When electromagnetic radiation passes through a transparent material, a portion of radiation may be absorbed. As a result of energy absorption, atoms or molecules pass from a state of low energy (ground state) to a state of higher energy (excited state). The electromagnetic radiation that is absorbed has energy exactly equal to the energy difference between the excited and ground states. UV/Vis absorption spectroscopy is based on the transitions of electrons from one molecular orbital to another due to the absorption of electromagnetic radiation of UV and visible region. As a molecule absorbs energy, an electron is promoted from an occupied orbital to an unoccupied orbital of greater potential energy. Generally, the transition of electrons occurs from the *highest occupied molecular orbital* to the *lowest unoccupied molecular orbital*.

Molecular orbitals with lowest energy are the σ -orbitals. The π -orbitals lie at high energy levels and non-bonding orbitals lie at even higher energies. The non-bonding orbitals contain a lone pair of electrons and they are stable, filled orbitals. Anti-bonding orbitals are normally empty and have higher energy than bonding or non-bonding orbitals.

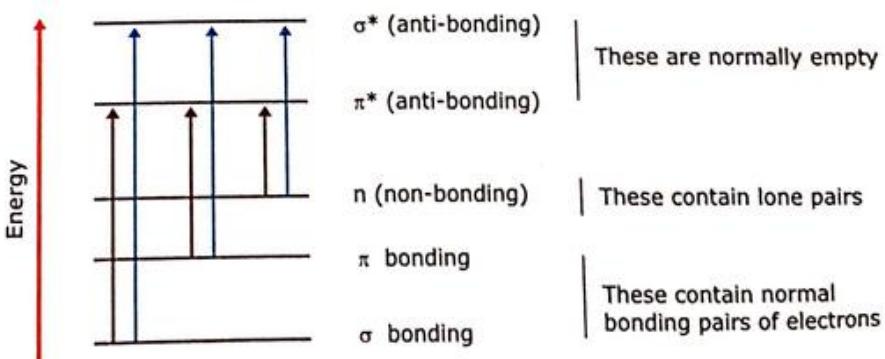


Figure 3.5 The possible electron jumps that light might cause are shown here. In each possible case, an electron is excited from a full orbital into an empty anti-bonding orbital. Each jump takes energy from the light, and a big jump obviously needs more energy than a small one. Each wavelength of light has a particular energy associated with it. If that particular amount of energy is just right for making one of these energy jumps, then that wavelength will be absorbed – its energy will be used in promoting an electron.

In an electronic transition, an electron moves from one orbital to another. Transition occurs between atomic orbitals in atoms and between molecular orbitals in molecules.

Chromophores are unsaturated organic functional groups that absorb in the UV or visible region.

When electromagnetic radiation (light) passes through a compound, energy from the radiation is used to promote an electron from a bonding or non-bonding orbital into one of the empty anti-bonding orbitals. The possible electron jumps that electromagnetic radiation might cause are shown in figure 3.5.

In each possible case, an electron is excited from a full orbital into an empty anti-bonding orbital. The energy gaps between these levels determine the wavelength of the electromagnetic radiation absorbed, and these gaps will be different in different compounds. The larger the gap between the energy levels, the greater the energy required to promote the electron to the higher energy level; resulting in light of higher frequency, and therefore shorter wavelength, being absorbed. Each wavelength of electromagnetic radiation has a particular energy associated with it. If that particular amount of energy is just right for making one of these energy jumps, then that wavelength will be absorbed - its energy will have been used in promoting an electron. An absorption of electromagnetic radiation in UV/Vis region (200 to 700 nm) cause only a limited number of the possible electron jumps. These jumps are from π bonding orbitals to π anti-bonding orbitals (π to π^*) and from non-bonding orbitals, n , to π anti-bonding orbitals (n to π^*). This means that in order to absorb electromagnetic radiation in the UV/Vis region, the molecule must contain either π bonds or atoms with non-bonding orbitals. Both n to π^* and π to π^* transitions require the presence of an unsaturated functional group to provide the p-orbitals. Molecules containing such functional groups and capable of absorbing UV/Vis radiation are called *chromophores*.

Molecules that show increasing degrees of conjugation require less energy for excitation and as a result absorb radiation of longer wavelengths. Molecules that contain conjugated systems, i.e. alternating single and double bonds, will have their electrons *delocalized* due to overlap of the p-orbitals in the double bonds. As the amount of delocalization in the molecule increases, the energy gap between the π bonding orbitals and π anti-bonding orbitals gets smaller; and therefore light of lower energy, and longer wavelength, is absorbed.

3.4 IR absorption spectroscopy

Each frequency of light has a certain energy. If a particular frequency is being absorbed as it passes through the compound being investigated, it must mean that its energy is being transferred to the compound. As with other types of energy absorption, molecules are excited to a higher energy state when they absorb infrared radiation (IR). The term *infrared* covers the range of the electromagnetic spectrum between 1 micrometer to 100 micrometers. It is commonly divided into three sub regions – near-IR, mid-IR and far-IR. A molecule absorbs only selected frequencies of IR radiation. The absorption of IR radiation corresponds to energy changes in the order of 8 to 40 kJ/mole.

IR spectroscopy is one type of *vibrational spectroscopy*. IR radiation does not have enough energy to induce electronic transitions as seen with UV and visible radiations. The energies of IR radiation correspond to the energies involved in bond vibrations. At temperatures above absolute zero, all the atoms in molecules are in continuous vibration with respect to each other.

Degree of freedom

A molecule consisting of n atoms has a total of $3n$ degrees of freedom. In a *nonlinear* molecule, 3 of these degrees are rotational and 3 are translational and the remaining ($3n-6$) correspond to fundamental vibrational freedom. In a *linear* molecule, 2 degrees are rotational and 3 are translational and $3n-5$ are vibrational degrees of freedom.

Table 3.2 Fundamental vibrational degrees of freedom for nonlinear and linear molecules

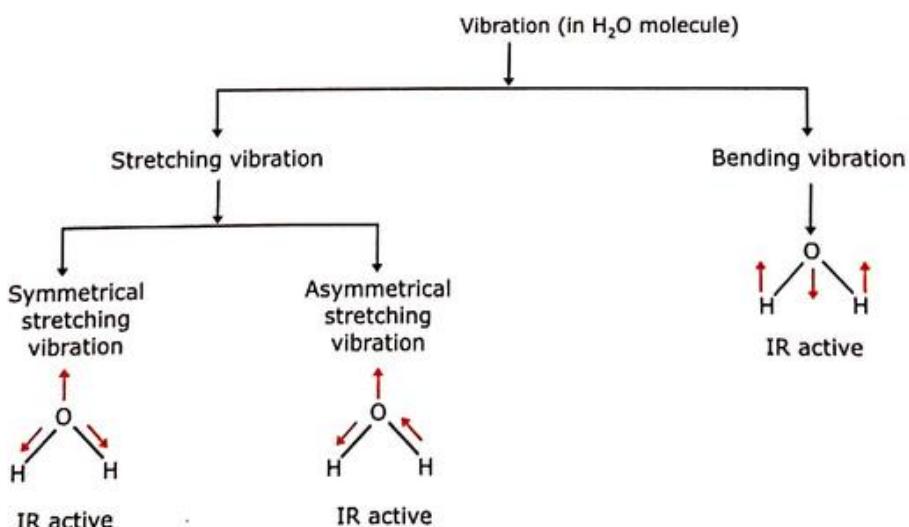
Molecule	Degrees of freedom			Total
	Translational	Rotational	Vibrational	
Nonlinear	3	3	$3n - 6$	$3n$
Linear	3	2	$3n - 5$	$3n$

The water is an example for a nonlinear molecule and has 3 fundamental vibrational degrees of freedom ($3 \times 3 - 6$). The CO_2 molecule is an example of a linear molecule and has 4 fundamental vibrational degrees of freedom ($3 \times 3 - 5$). Fundamental vibrational motion falls into the two main categories – *stretching* and *bending*.



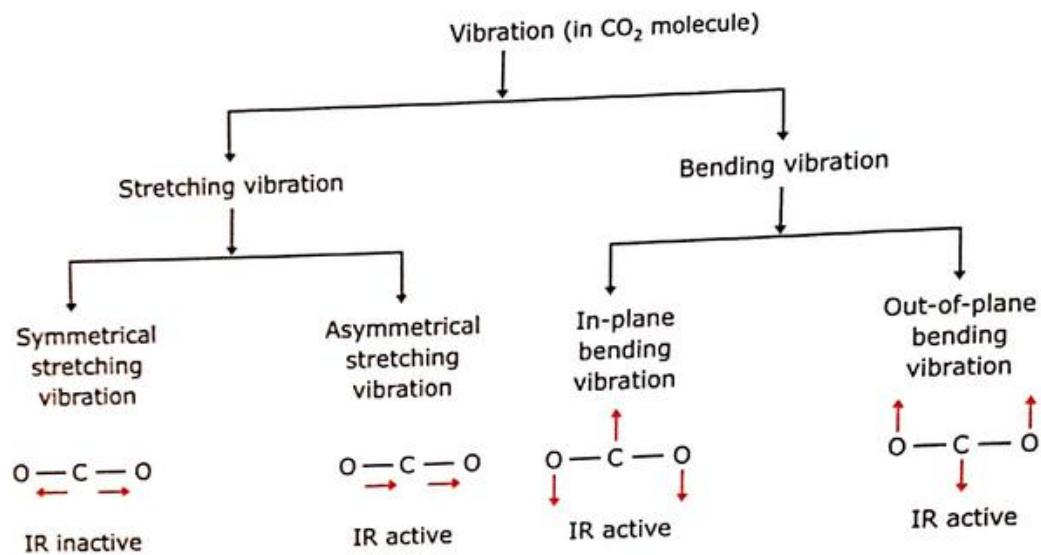
Stretching is a vibrational motion in which the bond length alters and the two nuclei move harmonically relative to each other. For a compound involving two atoms, there is only one bond that is capable of being stretched; and in general terms, it can be said that for a compound containing n atoms, there are $(n-1)$ possible stretching vibrations. Bending vibrations are characterized by a change in the angle between two bonds and are of four types: *scissoring*, *rocking*, *wagging* and *twisting*. Bending vibrations present many more possible variations, the total number being defined as $(2n-4)$ for linear molecules and $(2n-5)$ for nonlinear molecules.

Stretching and bending vibrations in water molecule: Water molecule has nonlinear structure. It has three fundamental vibrational modes – symmetrical stretching vibration, asymmetrical stretching vibration and bending vibration.



A stretching vibration involves changes in the length of an interatomic bond along its axis. Bending vibrations are characterised by a change in the angle between two bonds.

Stretching and bending vibrations in carbon dioxide molecule: Carbon dioxide molecule has linear structure. It has four fundamental vibrational modes : symmetric stretching vibration, asymmetric stretching vibration and two bending vibrations – *in plane* bending vibration and *out of plane* bending vibration.



IR active and inactive molecule

When the natural vibrational frequency of a molecule is equal to the frequency of the IR radiation directed on the molecule, the molecule absorbs the radiation. But not all bonds in a molecule are capable of absorbing infrared energy, even if the frequency of the radiation exactly matches that of the bond motion. To absorb IR radiation, the molecule must have a dipole moment and that must change during vibration. Thus, a bond must present an electrical dipole that is changing at the same frequency as the incoming radiation for energy to be transferred.

Homonuclear diatomic molecules such as H₂, N₂ and O₂ do not absorb IR radiation, as these molecules have a zero dipole moment and stretching vibration of the bonds also do not produce one. However, all molecules having a permanent dipole moment are also not IR active. In order to be IR active, a vibration must cause a change in the dipole moment of the molecule.

In linear symmetric molecules, symmetric stretching vibrations do not absorb IR radiation since there is no change in dipole moment. In the case of the asymmetric stretch, a dipole moment will be periodically produced and destroyed resulting in a changing dipole moment and therefore IR active. For example, CO₂ is a linear molecule and symmetric in structure, its equilibrium dipole moment is zero. In symmetrical stretching, both bonds are shortened or elongated to the same extent. As there is no change in dipole moment, it is IR inactive. In asymmetrical stretching, one of the bonds is shortened and the other is elongated. Hence, there is change in bond length and dipole moment. So it is infrared active.

Dipole moment depends on the variation in the distribution of electrons along the bond, and also its length, which is why stretching a bond can change its dipole moment. For bonds between unlike atoms, the larger the difference in electronegativity, the greater the dipole moment, and the more it changes when stretched.

A normal mode of vibration that gives rise to an oscillating dipole in the IR spectral range is said to be IR-active.

Conversely, a normal mode of vibration that does not give rise to an oscillating dipole moment (e.g. the stretching of a homonuclear diatomic molecule) cannot lead to IR absorption and is said to be IR-inactive.

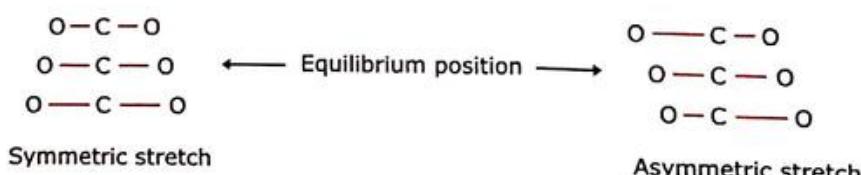


Figure 3.6 Symmetric and asymmetric stretching vibration of the carbon dioxide molecule.

Uses of the infrared spectrum

The energy involved in vibration depends on things like the length of the bond and the mass of the atoms at either end. This means that each different bond will vibrate in a different way, involving different amounts of energy. The amount of energy it needs to do this will vary from bond to bond, and so each different bond will absorb a different frequency (and hence energy) of IR radiation. In general,

1. Stretching frequencies are higher than corresponding bending frequencies. It is easier to bend a bond than to stretch or compress it.
2. Bonds to lighter atoms (such as hydrogen atom) have higher stretching frequencies than those to heavier atoms.
3. Triple bonds have higher stretching frequencies than corresponding double bonds, which in turn have higher frequencies than single bonds.

Values chiefly affected by mass of atoms: *Lighter atom, higher frequency*

C-H (3000 cm^{-1})

C-O (1100 cm^{-1})

C-Cl (700 cm^{-1})

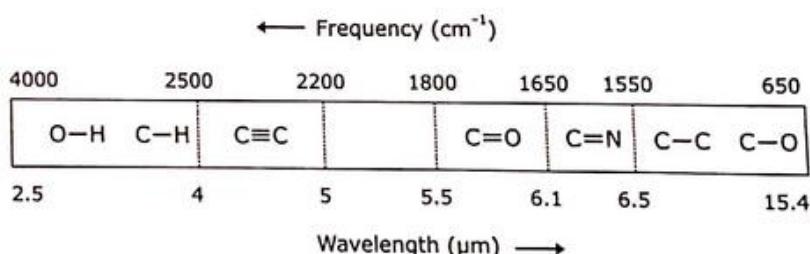
Values chiefly affected by bond strength: *Stronger bond, higher frequency*

C≡O (2143 cm^{-1})

C=O (1715 cm^{-1})

C-O (1100 cm^{-1})

An IR spectrum usually extends from radiation around 4000 cm^{-1} to 650 cm^{-1} and can be split into the *functional group region* and the *fingerprint region*. Many functional groups absorb infrared radiation at about the same frequency, regardless of the structure of the rest of the molecule. For example, C-H stretching vibrations usually appear between 3200 and 2800 cm^{-1} ; and carbonyl (C=O) stretching vibrations usually appear between 1800 and 1600 cm^{-1} . This makes these bands diagnostic markers for the presence of a functional group in a sample. These types of infrared bands are called *group frequencies* because they tell us about the presence or absence of specific functional groups in a sample.



In IR spectroscopy, spectroscopists normally use 'wavenumbers' instead of frequency. A wavenumber is the inverse of the wavelength. It has units of cm^{-1} . It is directly proportional to the frequency and the energy of the radiation. IR radiation with a high wavenumber has higher frequency and energy than IR radiation with a lower wavenumber. Because wavenumber and frequency are directly proportional to one another, it is common for the two terms to be used almost interchangeably.

Figure 3.7 Infrared spectrum. The approximate regions where various common types of bonds absorb for stretching vibrations only.

The region of the infrared spectrum from 1200 to 700 cm^{-1} is called the *fingerprint region*. This region usually contains a very complicated series of absorptions. Many different vibrations, including C—O, C—C and C—N single bond stretches, C—H bending vibrations, and some bands due to benzene rings are found in this region. This is called the *fingerprint region*. The fingerprint region is often the most complex and confusing region to interpret. The fingerprint region is different for each molecule just like a fingerprint is different for each person. Two different molecules may have similar functional group regions because they have similar functional groups, but they will always have a different fingerprint region. Since every type of bond has a different natural frequency of vibration, and since two of the same types of bonds in two different compounds are in two slightly different environments, no two molecules of different structure have exactly the same infrared absorption pattern, or infrared spectrum. By comparing the infrared spectra of two substances, one can establish whether they are identical or non-identical.

A second use of the infrared spectrum is to determine structural information about a molecule. The absorptions of each type of bond (N-H, C-H, O-H, C-O, C-C and other) are found only in certain small portions of the infrared region. Thus, a small range of absorption can be defined for each type of bond.

Analyzing an infrared spectrum

The instrument that determines the absorption spectrum for a compound is called an *infrared spectrometer*. It determines the positions and relative sizes of all the absorptions, or peaks, in the IR region and plots them on a piece of paper. This plot of absorption intensity versus wave-number (or sometimes wavelength) is referred to as the infrared spectrum of the compound.

In Infrared spectra,

Position of band depends on - Mass of atoms

Light atoms give high frequency.

Bond strength

Strong bonds give high frequency.

Strength of band depends on - Change in dipole moment

A large change in dipole moment gives strong absorption.

Width of band depends on - Hydrogen bonding

Strong H-bond gives wide peak.

3.5 Nuclear magnetic resonance

Nuclear magnetic resonance (NMR) is a spectroscopic technique that involves change in nuclear spin energy in the presence of external magnetic field. It is based on the absorption of electromagnetic radiation in the radio-frequency region.

Protons, electrons and neutrons possess a property called spin. Spin comes in multiples of 1/2 and can be + or -. Some atomic nuclei also have spin. If a particular nucleus is composed of p protons and n neutrons, its total mass is p + n, its total charge is +p and its total spin will be a vector combination of p + n spins each of magnitude 1/2. If the number of both the protons and neutrons in a nucleus are even then there is no overall spin.

All nuclei with an even mass number (total number of protons and neutrons in the nucleus) and an even atomic number (number of protons in the nucleus) have a nuclear spin of zero. Any atomic nucleus that possesses either odd mass number, odd atomic number or both has a spin value.

For each nucleus with spin, the number of allowed spin states – it may adopt – is determined by its nuclear spin quantum number (I). The spin quantum number of a nucleus is determined by the number of unpaired protons and neutrons it contains. For example, ^{12}C has even numbers of protons and neutrons: each proton pairs with a proton of opposite sign, as does each neutron, giving a net spin angular momentum of zero ($I = 0$).

Table 3.3 Atomic number, number of neutrons and spin quantum number of some elements

Element	^1H	^{12}C	^{13}C	^{14}N	^{15}N	^{16}O	^{19}F	^{31}P
Atomic number	1	6	6	7	7	8	9	15
Number of neutrons	0	6	7	7	8	8	10	16
Spin quantum number	$\frac{1}{2}$	0	$\frac{1}{2}$	1	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$

Certain atomic nuclei can be considered as having 'spin'; the term spin implies that each nucleus can be considered as a rotating electrical charge and consequently, along with its electrical properties, it also possesses an angular magnetic momentum.

A nucleus of spin $\frac{1}{2}$ has $2j + 1$ allowed spin states. The nuclear spin (j) can have various values such as 0, $1/2$, 1, $3/2$. For example, ${}^1\text{H}$ has the spin quantum number $j = 1/2$ and has two allowed spin states [$2(1/2) + 1 = 2$] for its nucleus, $-1/2$ and $+1/2$. For the chlorine nucleus, $j = 3/2$ and there are four allowed spin states [$2(3/2) + 1 = 4$] for its nucleus, $-3/2$, $-1/2$, $+1/2$ and $+3/2$.

Table 3.4 Spin quantum numbers and number of spin states of some common elements

Element	${}^1\text{H}$	${}^2\text{H}$	${}^{12}\text{C}$	${}^{13}\text{C}$	${}^{14}\text{N}$	${}^{16}\text{O}$	${}^{17}\text{O}$	${}^{19}\text{F}$
Spin quantum number	$1/2$	1	0	$1/2$	1	0	$1/2$	$1/2$
Number of spin states	2	3	0	2	3	0	6	2

Due to spin, nuclei of certain isotopes of some elements (${}^1\text{H}$, ${}^{13}\text{C}$, ${}^{19}\text{F}$) behave like minute magnets. These atomic nuclei have different energy levels when placed in a magnetic field and with the help of nuclear spin value, the number of energy levels can be determined by using formula $2j + 1$. The ${}^1\text{H}$ and ${}^{13}\text{C}$ nuclei possess nuclear spin and have two different energy levels. When a magnetic field is applied to these nuclei, they can either align themselves with it, which would be the lowest energy state, or they can align themselves against the field, which is higher in energy. The spin state $+1/2$ is of lower energy since it is aligned with the field, while the spin state $-1/2$ is of higher energy since it is opposed to the applied field. This energy difference is a function of the strength of the applied magnetic field, B . The stronger the applied magnetic field, the greater the energy difference between the possible spin states.

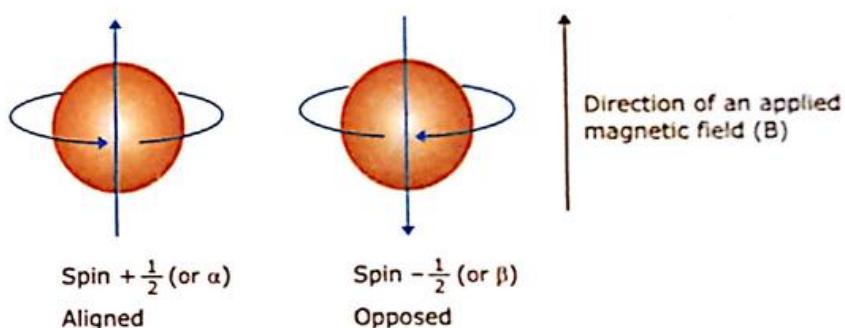


Figure 3.8 Two allowed spin states for a ${}^1\text{H}$ nucleus.

When energy is given to nuclei aligned with an applied field, their spin orientation changes with respect to the applied field. The energy absorption is a quantized process, and the energy absorbed must equal the energy difference between the two energy states involved. But, since the energy difference between the two states even in a very strong external magnetic field is so small, we don't need to do much work. In fact, the amount of energy needed to flip the nucleus can be provided by electromagnetic radiation of radio wave frequency. Radio waves flip the nucleus from the lower energy state to the higher state. The nucleus now wants to return to the lower energy state and, when it does so, the energy comes out again and this is what we detect.

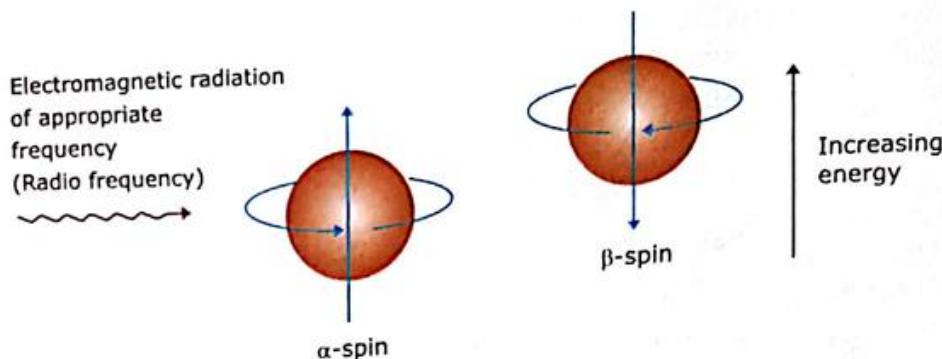


Figure 3.9 A rotating nucleus can take up either of two orientation in an applied electromagnetic field. The energy separation between the α and β spin states is determined by measuring the frequency of electromagnetic radiation needed to induce $\alpha \rightarrow \beta$ spin.

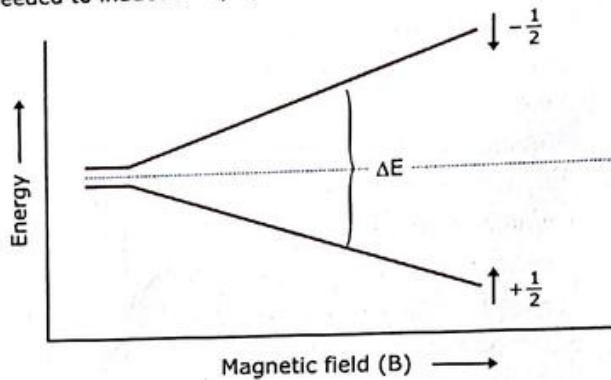


Figure 3.10 The spin-state energy separation as a function of the strength of the applied magnetic field.

Chemical shift

In an applied magnetic field, not all hydrogens and carbons in a molecule resonate at exactly the same frequency. This variability is due to the fact that the hydrogens and carbons in a molecule are surrounded by electrons and exist in slightly different electronic environments from one another. The valence-shell electron densities vary from one hydrogen or carbon to another. In an applied magnetic field, the valence-shell electrons are caused to circulate. This circulation, called a *local diamagnetic current*, generates a counter magnetic field that opposes the applied magnetic field. This is called diamagnetic shielding (or diamagnetic anisotropy). Circulation of electrons around a nucleus can be viewed as being similar to the flow of an electric current in an electric wire. In physics, we know that the flow of a current through a wire induces a magnetic field. In an atom, the local diamagnetic current generates an induced magnetic field that has a direction opposite to that of the applied magnetic field. As a result of diamagnetic anisotropy, each hydrogen or carbon in a molecule is shielded from the applied magnetic field to an extent that depends on the electron density surrounding it. The greater the electron density around a nucleus, the greater the induced counter field that opposes the applied field. The counter field that shields a nucleus diminishes the net applied magnetic field that the nucleus experiences.

Let us take the spectrum of carbon NMR of *propanol*. In this spectrum, 3 peaks form as given below. Each peak represents a different kind of carbon atoms: each one absorbs energy at a different frequency. But why should carbon atoms be different? Two factors that affect the energy difference - the magnetic field strength and what sort of nucleus is being studied. So we might expect all carbon-13 nuclei to resonate at one particular frequency. But they don't.

The chemical shift in absolute terms is defined by the frequency of the resonance expressed with reference to a standard compound which is defined to be at 0 ppm. The scale is made more manageable by expressing it in parts per million (ppm) and is independent of the spectrometer frequency.

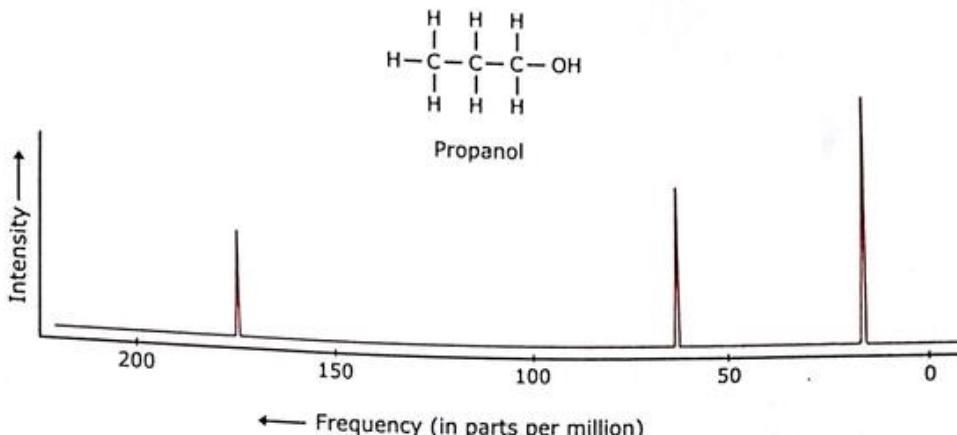
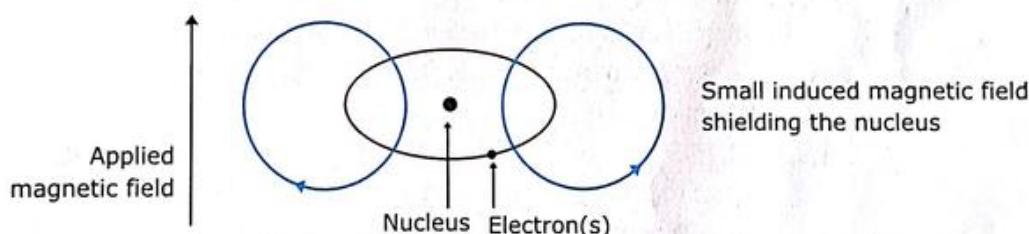


Figure 3.11 Carbon NMR can distinguish between the three different carbon atoms. In this graph, horizontal axis is actually frequency, but it is usually expressed in ppm of the field of the magnet.

The variation in frequency for different carbon atoms must mean that there are different types of carbon atoms and their nuclei experience a magnetic field that is not quite the same as the magnetic field that we apply. Each nucleus is surrounded by electrons, and in a magnetic field these will set up a tiny electric current. This current will set up its own magnetic field, which will oppose the magnetic field that we apply. The electrons are said to *shield* the nucleus from the external magnetic field. If the electron distribution varies from ^{13}C atom to ^{13}C atom, so does the local magnetic field, and so does the resonating frequency of the ^{13}C nuclei.

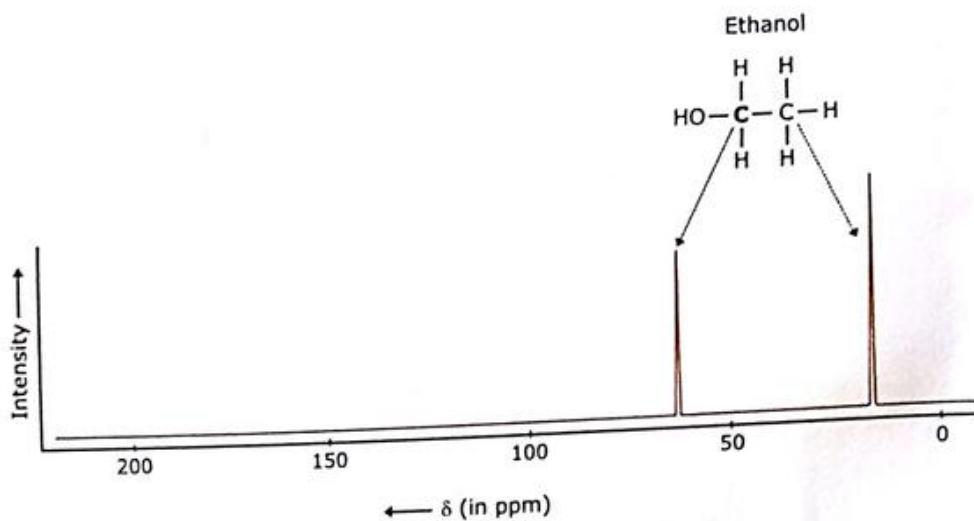


Thus, the changes in the distribution of electrons around a nucleus affect:

- The local magnetic field that the nucleus experiences.
- The frequency at which the nucleus resonates.
- The chemistry of the molecule at that atom.

This variation in frequency is known as the chemical shift and it is denoted by δ . The chemical shift of a nucleus depends on many factors, but the surrounding electron density is often the dominant one. A high electron density causes a large shielding effect.

As an example, consider ethanol. The *bold* carbon attached to the OH group will have relatively fewer electrons around it compared to the *unbold* carbon since the oxygen atom is more electronegative and draws electrons towards it, away from the carbon atom. The external magnetic field that this *bold* carbon nucleus feels will, therefore, be slightly greater than the felt by the *unbold* carbon with more electrons. Since the *bold* carbon is less shielded (*deshielded*) from the applied external magnetic field, it feels a stronger magnetic field and there will be a greater energy difference between the two energy states of the nucleus. The greater the energy difference, the higher the resonance frequency. So, for ethanol, we would expect the *bold* carbon with the OH group attached to resonate at a higher frequency than the *unbold* carbon; indeed this is exactly what the ^{13}C NMR spectrum shows.



The reference sample —Tetramethylsilane

Unlike IR and UV/Vis spectroscopy, where the signals are fixed at particular frequencies or wavelengths, in NMR, the signal is dependent on the field strength. Since no two magnets have the same field, the frequency at which signals are obtained would vary correspondingly. Therefore, there was a need to characterize and specify the location of the signals. In order to avoid this situations, it has been decided to eliminate the problems by keeping a standard reference with respect to which a numerical value can be assigned. The compound we use as a reference sample in ^1H and ^{13}C NMR is usually tetramethylsilane, TMS. This is silane (SiH_4) with each of the hydrogen atoms replaced by methyl groups to give $\text{Si}(\text{CH}_3)_4$. Because of molecular symmetry, all 12 protons of TMS absorb at the same frequency and all 4 carbons absorb at the same frequency. The frequency of absorption for a nucleus of interest relative to the frequency of absorption of a standard is called the *chemical shift* of the nucleus. The chemical shift of the ^1H nuclei in the ^1H NMR spectrum or ^{13}C nuclei in the ^{13}C NMR spectrum of TMS appear at = 0 ppm. Typically, it increases from 0 on the right hand side of the spectrum to 10 ppm on the left hand side of an ^1H NMR spectrum or from 0 on the right hand side to 200 ppm on the left hand side of a ^{13}C NMR spectrum. The reason frequencies of absorption are recorded on the δ -scale relative to those of a standard molecule. It makes the position of absorption independent of the spectrometer used to record the spectrum, in particular, independent of the strength of the magnetic field of the spectrometer.

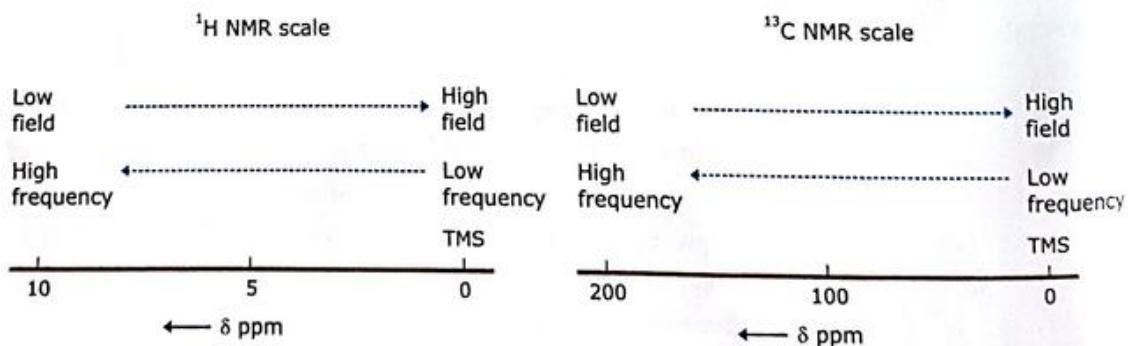


Figure 3.12 The chemical shift scale for ^1H and ^{13}C NMR.

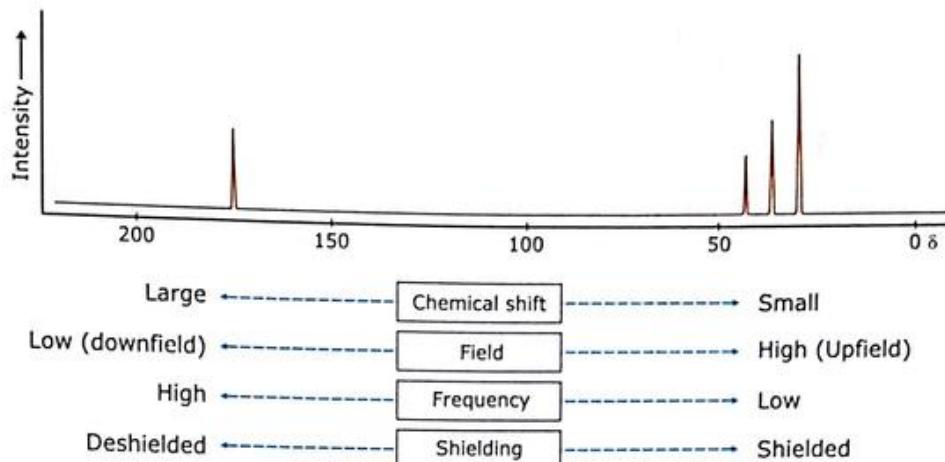


Figure 3.13 Different ways of expressing NMR spectra. The chemical shift scale runs to the left from zero (where TMS resonates).

Magnetic resonance imaging

Light is a transverse electromagnetic wave. The electric and magnetic fields in an electromagnetic wave oscillate along directions perpendicular to the propagation direction of the wave. Light can be unpolarized or polarized. In unpolarized light, the electric field is vibrating in more than one plane; all perpendicular to the direction of propagation. A plane or linearly polarized light vibrates in a single plane perpendicular to the direction of propagation. The orientation of a linearly polarized light is defined by the direction of the electric field vector. When unpolarized light is passed through a polarizer, only light oscillating in one direction is transmitted and gives plane or linearly polarized light.

A very important diagnostic tool in medicine that is based on the principles of NMR is a technique known as Magnetic Resonance Imaging (MRI). MRI uses strong magnetic fields and radio waves to form images of the body. It is used for the diagnosis and evaluation of diseases. In most cases, the ^1H nucleus is the one studied since it is found in the water molecules that are present in and around living tissue. The ^1H nuclei of water molecules that are not bound within living cells have a *relaxation time* different from the nuclei of water molecules bound within tissue. Relaxation processes involve some non-radiative transitions by which nuclei in the higher energy state return to the lower state. Water molecules that appear in a highly ordered state have relaxation times shorter than water molecules that appear in a more random state. The degree of ordering of water molecules within tissues is greater than that of water molecules that are part of the fluid flowing within the body. Furthermore, the degree of order of water molecules may be different in different types of tissue, especially in diseased tissue as compared with normal tissue. Specific pulse sequences detect these differences in relaxation time for the protons of water molecules in the tissue being examined. When the results of the scans are processed, the image that is produced shows different densities of signals, depending on the degree to which the water molecules are in an ordered state. As a result, the 'picture' that we see shows the various types of tissue clearly. The radiologist can then examine the image to determine whether any abnormality exists.

3.6 Optical Rotatory Dispersion and Circular Dichroism

Most of biological compounds are optically active. An optically active compound shows a phenomenon called *optical rotation*. It means that the plane of polarization of a linearly polarized light rotates as it passes through an optically active medium.

The technique of Optical Rotatory Dispersion (ORD) measures the ability of optically active compounds to rotate plane polarized light as a function of the wavelength (i.e. wavelength dependence of optical rotation). The instrument that measures ORD curves is called a *spec-tropolarimeter*, which differs from a polarimeter in that the latter employs only one wavelength.

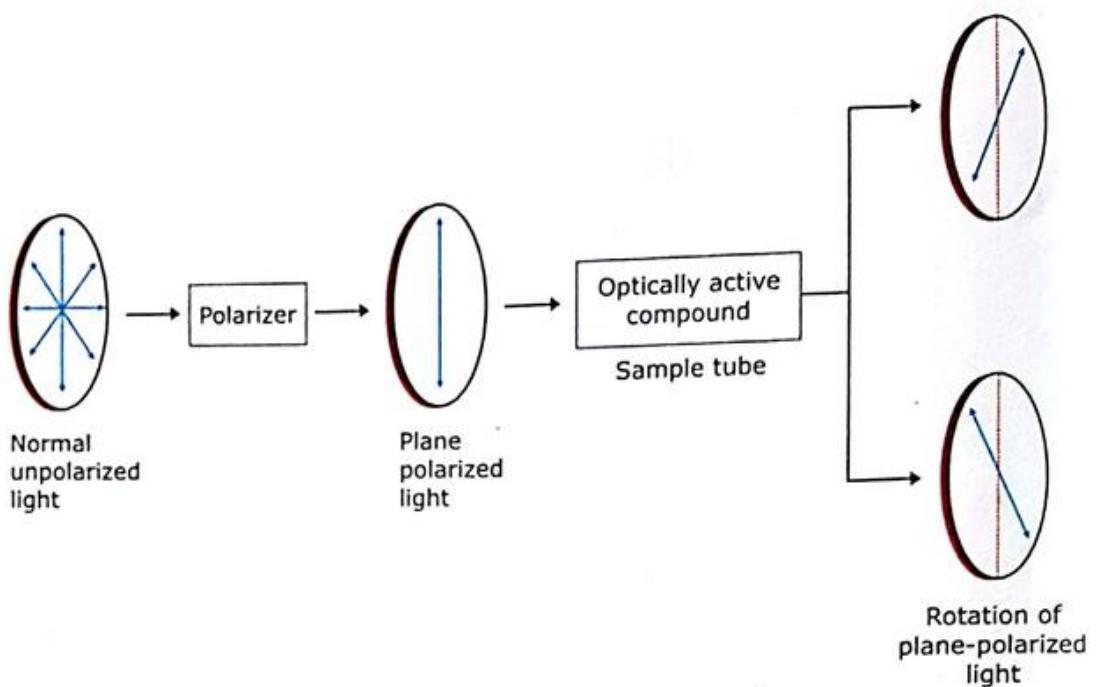


Figure 3.14 When plane polarized light is passed through a solution that contains an optically active compound, there is net rotation of the plane polarized light. The light is rotated either clockwise (dextrorotatory) or counterclockwise (levorotatory) by an angle that depends on the molecular structure and concentration of the compound, the pathlength and the wavelength of the light.

The physical basis of the **Circular Dichroism (CD)** is same as ORD. In case of CD, circularly polarized light is used. *Circularly polarized* light is obtained by superimposing two plane polarized light of same wavelengths and amplitudes which are polarized in two perpendicular planes, but there is a phase difference of 90° between them. A phase difference of 90° means that when one wave is at its peak, then the other one is just crossing the zero line.

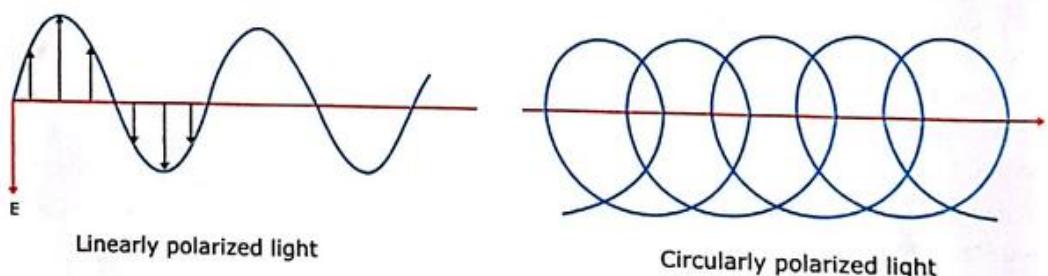


Figure 3.15 Linearly polarized light results when the direction of the E-vector is restricted to a plane perpendicular to the direction of propagation while its magnitude oscillates. In circularly polarized light, the magnitude of the oscillation is constant and the direction oscillates.

A circularly plane polarized light may be right or left handed circularly polarized. For left circularly polarized light with propagation towards the observer, the electric vector rotates counterclockwise. For right circularly polarized light, the electric vector rotates clockwise.

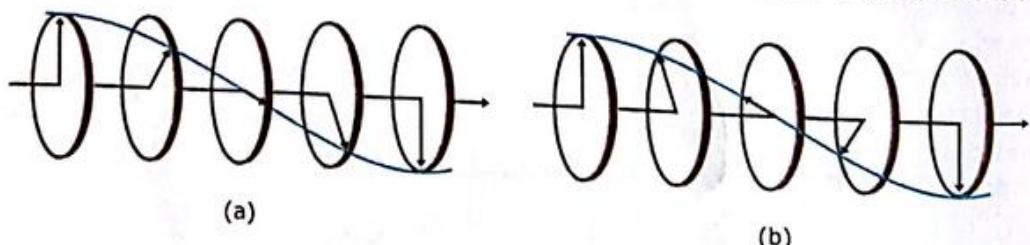


Figure 3.16 Schematic representation of (a) right circularly polarized and (b) left circularly polarized light. In both cases, the length of the vector remain constant. In linearly polarized light, the electric vector stays in the same plane but its length changes.

An elliptically polarized light is obtained by superimposing two plane polarized light vibrating at right angle to each other having same wavelengths and unequal amplitudes but there is a phase difference of 90° between them.

The superposition of left and right circularly polarized light beams of equal amplitudes can result in plane-polarized light (linearly polarized light). Thus, linear polarized light can be viewed as a superposition of opposite circular polarized light of equal amplitude and phase. Dichroism is a word derived from Greek which means *two-colors*, because the sample under analysis has one color if illuminated with the right polarized light and a different color if illuminated with the left one. The color, in fact, depends on light absorption. Circular Dichroism is observed when optically active matter absorbs left and right circularly polarized light slightly different. In fact, circular dichroism is the absorption difference between left and right circularly polarized light at a given wavelength.

Absorption is quantitated by the *molar extinction coefficient* (ϵ). Optically active samples have distinct molar extinction coefficients for left (ϵ_L) and right (ϵ_R) circularly polarized light. The difference in absorbance, A , of the two components, is a measure of Circular Dichroism.

Thus, $CD = A_L - A_R$

The difference between ϵ_L and ϵ_R may be expressed as $\Delta\epsilon$. From Beer-Lambert law the difference in the absorbance of left and right circularly polarized light ΔA , can be given by,

$$\Delta A = \Delta\epsilon cl, \quad c \text{ is the concentration and } l \text{ is the path length.}$$

The differential absorption of the left and right circularly polarized light means that the amplitudes of the right and left circularly polarized components of the transmitted beam will differ. The superposition of the two components is no longer a linearly polarized wave but it rotates along an ellipsoid path. Such a light wave is called an *elliptically polarized light*. How elliptical the plane-polarized wave becomes after traversing the medium is determined by the difference between the absorptions of the two circularly polarized components. In the most extreme case, the material almost completely extincts one left or right component and then the transmitted wave almost becomes a perfect circularly polarized light because the other circular component disappears.

The ellipticity is proportional to the difference in the absorbance of the two components, $A_L - A_R$. Thus, the CD is equivalent to ellipticity. The relationship between CD and ellipticity (θ) is given by:

$$\theta = 2.303 (A_L - A_R) \frac{180}{4\pi} = 33 (A_L - A_R) = 33 \Delta A \text{ degree}$$

Applications

Circular dichroism is an excellent method for the study of the conformations adopted by proteins and nucleic acids in solution. Although not able to provide the detailed residue-specific information as obtained from NMR and X-ray crystallography, CD measurements have two major advantages: they can be made on small amounts of material in physiological buffers and they provide one of the best methods for monitoring any structural alterations that might result from changes in environmental conditions, such as pH, temperature, and ionic strength.

1. Determination of protein's secondary and tertiary structure: Circular dichroism relies on the differential absorption of left and right circularly polarised radiation by chromophores. Proteins possess a number of chromophores which can give rise to CD signals. Secondary structure can be determined by CD spectroscopy in the far-UV spectral region (180–240 nm). At these wavelengths, the chromophore is the peptide bond. The CD spectrum in this region, gives the content of regular secondary structural features such as α -helix and β -sheet.

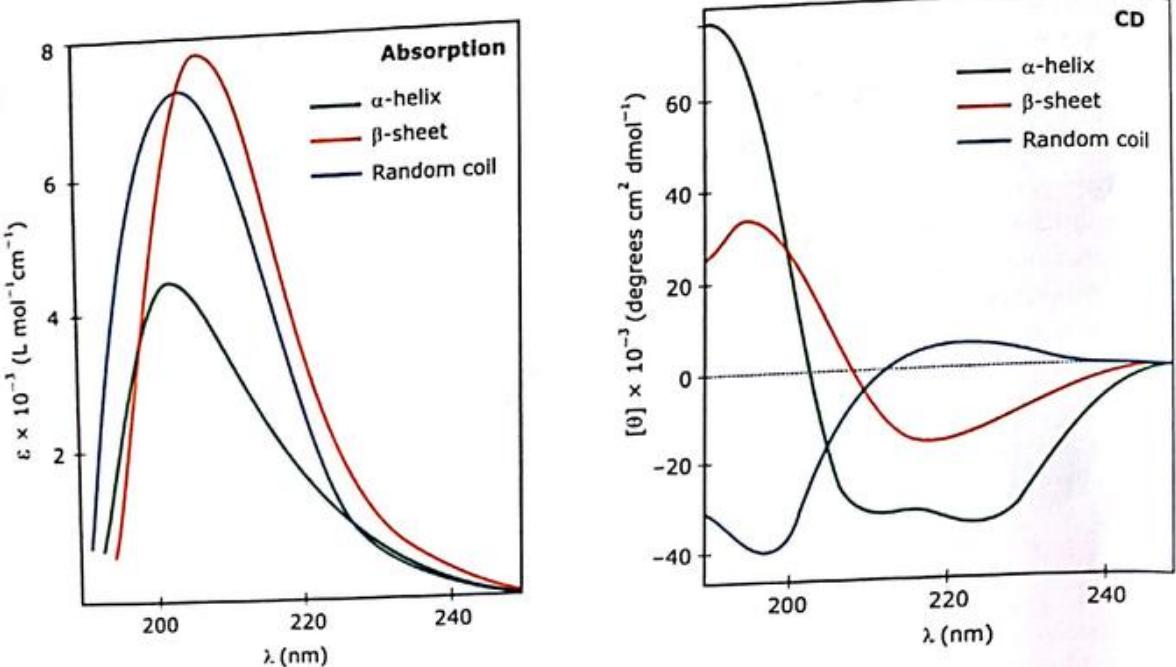


Figure 3.17 UV absorption and circular dichroism spectra in far-UV region and protein secondary structure are shown here. Measurement performed on synthetic homopolyptide, poly-L-lysine, in the far UV region between 180 nm and 250 nm. The α -helix, β -sheet and random coil structures each give rise to a characteristic shape and magnitude in CD spectrum.

The CD spectrum in the near UV region (320–260 nm) gives information about the tertiary structure of the protein. At these wavelengths, the chromophores are the aromatic amino acids and disulfide bonds; and the CD signals they produce are sensitive to the overall tertiary structure of the protein. Signals in the region from 250–270 nm are attributable to phenylalanine residues; signals from 270–290 nm are attributable to tyrosine; and those from 280–300 nm are attributable to tryptophan. Disulfide bonds give rise to broad weak signals throughout the near-UV spectrum. If a protein retains secondary structure but no defined three-dimensional structure (e.g. an incorrectly folded or *molten-globule* structure), the signals in the near-UV region will be nearly zero. On the other hand, the presence of significant near-UV signals is a good indication that the protein is folded into a well-defined structure.

CD spectra can be readily used to estimate the fraction of a molecule that is in the α -helix, the β -sheet, or random coil conformation. Each of the three basic secondary structures of a polypeptide chain shows a characteristic CD spectrum.

2. Comparison of the secondary and tertiary structure of wild type and mutant proteins: It is also an excellent spectroscopic technique to measure the unfolding and folding of proteins as a function of temperature.
3. Nucleic acid structure and changes upon protein binding or melting.
4. Determination of conformational changes due to protein-protein interactions, protein-DNA interactions and protein-ligand interactions.

Analysis by CD has number of advantages. First, molecules of any size can be studied. The experiments are quick to perform and can work with very small concentrations. It allows to study dynamic systems and kinetics. However, it only provides qualitative analysis of data. It does not provide atomic level structural analysis. Also, the observed spectrum is not enough for claiming one and only possible structure.

3.7 Raman spectroscopy

When light interacts with matter, the photons which make up the light may be absorbed or scattered, or may not interact with the material and may pass straight through it. If the energy of an incident photon corresponds to the energy gap between the ground state of a molecule and an excited state, the photon may be absorbed and the molecule promoted to the higher energy excited state. It is this change which is measured in absorption spectroscopy by the detection of the loss of that energy of radiation from the light. However, it is also possible for the photon to interact with the molecule and scatter from it. In this case, there is no need for the photon to have an energy which matches the difference between two energy levels of the molecule.

Raman spectroscopy is based on *inelastic scattering* of monochromatic light. Inelastic scattering means that the frequency of photons in monochromatic light changes upon interaction with a sample. When light encounters molecules in the air, the predominant mode of scattering is elastic scattering, called *Rayleigh scattering*. This scattering is responsible for the blue color of the sky. In this case, the molecules of the substance are excited to a virtual electronic state and immediately fall back to their original state by releasing a photon. The photon energy of this scattered light is equal to that of the incoming light. It is also possible for the incident photons to interact with the molecules in such a way that energy is either gained or lost so that the scattered photons are shifted in frequency. Such inelastic scattering is called *Raman scattering*.

The change in the energy of the scattered photon corresponds exactly to the photon energy. As a result, the wavelength of the scattered photons can be longer (Stokes Raman scattering) or shorter (anti-Stokes Raman scattering). The energy of the scattered radiation is less than the incident radiation for the Stokes line and the energy of the scattered radiation is more than the incident radiation for the anti-Stokes line. The difference in energy between the incoming and scattered photon (Raman shift) corresponds to the energy difference between vibrational energy levels of the molecule. The different vibrational modes of a molecule can therefore be identified by recognizing Raman shifts (or 'bands') in the inelastically scattered light spectrum.

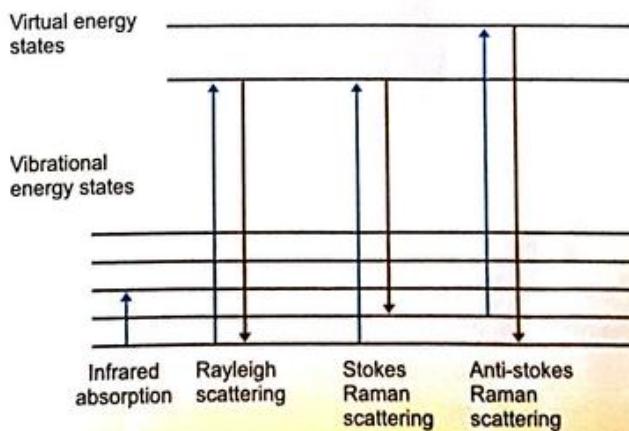


Figure 3.18 Energy level diagram to explain Raman scattering.

IR and Raman are the most common vibrational spectroscopies. However, the way in which radiation is employed in infrared and Raman spectroscopies is different. In infrared spectroscopy, infrared energy covering a range of frequencies is directed onto the sample. Absorption occurs where the frequency of the incident radiation matches that of a vibration; so that

Scattering is a physical process that causes radiation to deviate from a straight trajectory.

the molecule is promoted to a vibrational excited state. The loss of this frequency of radiation from the beam after it passes through the sample is then detected. In contrast, Raman spectroscopy uses a single frequency of radiation to irradiate the sample and it is the radiation scattered from the molecule, one vibrational unit of energy different from the incident beam, which is detected. Thus, unlike infrared absorption, Raman scattering does not require matching of the incident radiation to the energy difference between the ground and excited states. In Raman scattering, the light interacts with the molecule and distorts (polarizes) the cloud of electrons round the nuclei to form a short-lived state called a 'virtual state'. This state is not stable and the photon is quickly reradiated.

References

- Banwell CN, McCash EM (2013), *Fundamentals of Molecular Spectroscopy*, 5th ed. McGraw Hill Education.
- Berg JM, Tymoczko JL and Stryer L (2006), *Biochemistry*, 6th ed. W.H. Freeman and company.
- Clark BJ (1993), *UV Spectroscopy: Techniques, Instrumentation and Data Handling*, Chapman and Hall Inc.
- Holler FJ, Crouch SR (2014), *Fundamentals of analytical chemistry*, 9th ed. Cengage Learning.
- Hollas JM (2002), *Basic Atomic and Molecular Spectroscopy*, RSC Publishing.
- Keeler J (2011), *Understanding NMR Spectroscopy*, 2nd ed. John Wiley and Sons.
- Lakowicz JR (1999), *Principles of Fluorescence Spectroscopy*, 2nd ed. Kluwer, New York.
- Pavia DL, Lampman GM, Kriz GS and Vyvyan JR (2009), *Introduction to spectroscopy*, 4th ed. Brooks/Cole Cengage Learning.

Chapter 4

Mass Spectrometry

A mass spectrometry determines the molecular mass by measuring the mass-to-charge ratio of ions in the gas phase. In this process, the ions are generated in the ionization source by inducing either the loss or the gain of a charge (e.g. electron ejection, protonation or deprotonation). Once the ions are formed in the gas phase, they can be electrostatically directed into a mass analyzer, separated according to mass, and finally detected. Thus, a mass spectrometer has three basic components of: the ionization source, the mass analyzer and the detector.

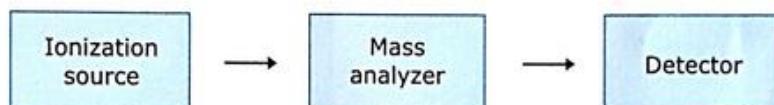


Figure 4.1 Components of a mass spectrometer.

The operation of the mass spectrometer involves the following steps:

1. Production of the sample in an ionized form in the gas phase;
2. Acceleration of the ions in an electric field, each ion emerging with a velocity proportional to its mass-to-charge ratio (m/z);
3. Passage of the ions into a field-free region;
4. Detection of the times of arrival of the ions, the time-of-flight indicating the mass-to-charge ratio of the ions.

Ionization process

A sample molecule studied by mass spectrometer must be converted to gas phase-charged particles by the ionization process before they can be analyzed and detected. The term ionization is misleading because most ionization processes do not perform the ionization of molecules per se. Instead, the term ionization relates to the transfer to gas phase of analytes, while maintaining their charge, and/or acquiring a charge from the sample environment, typically in the form of proton. Several methods are used for converting the sample to ions. These methods include:

Electron ionization

In electron ionization, a beam of high energy electrons strikes the molecules. The electron-molecule collision strips an electron from the molecule, creating a cation.

Chemical ionization

In chemical ionization, the sample molecules are combined with an ionized reagent gas. When the sample molecules collide with the ionized reagent gas, some of the sample molecules are ionized by various mechanisms like proton transfer, electron transfer and adduct formation.

Desorption ionization

In desorption ionization, the sample to be analyzed is dissolved in a matrix and placed in the path of high energy beam of ions or high intensity photons. In case of high intensity photons, it is termed as MALDI (matrix-assisted laser desorption/ionization). In MALDI, analytes are placed in a light-absorbing solid matrix. A matrix is used to protect the analytes from being destroyed by direct laser beam. Solid matrix strongly absorbs the laser radiation and acts as a receptacle for energy deposition. With a short pulse of laser light, the analytes are ionized and then desorbed from the matrix into the vacuum system. The exact mechanism of the origin of ions in MALDI is still not completely understood. Commonly used matrix materials are aromatic compounds that contain carboxylic acid functional groups. The aromatic ring of the matrix acts as a chromophore for the absorption of laser irradiation leading to the desorption of matrix and analytes into the gas phase.

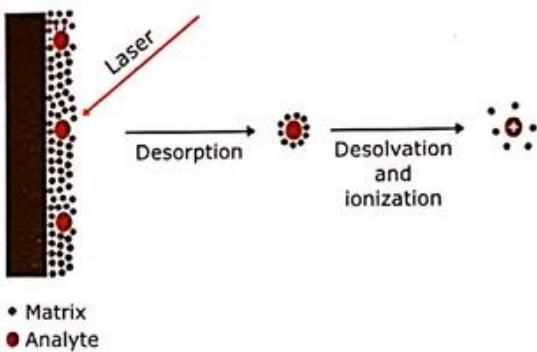
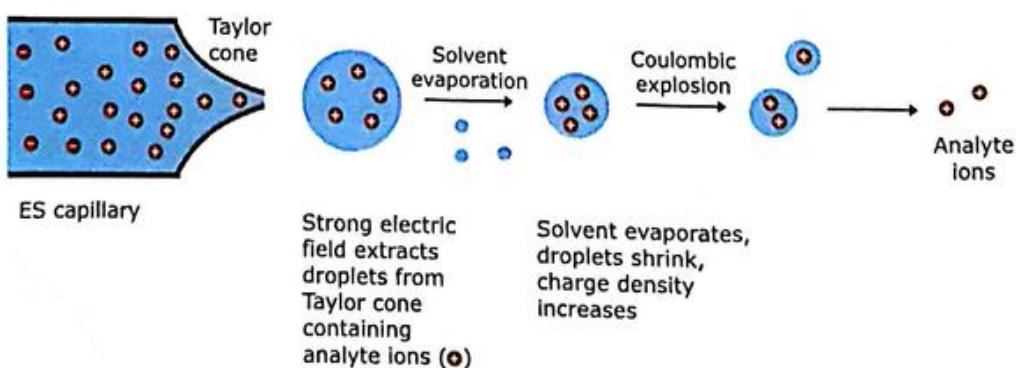


Figure 4.2 Schematic representation of MALDI.

Figure 4.3 Electrospray ionization: The liquid effluent containing the analytes is electrostatically dispersed. This generates highly charged droplets, which are normally positively charged. Once the droplets are airborne, the solvent evaporates, which decreases the size and increases the charge density of the droplets. Desolvated ions are generated by the desorption of analyte ions from the droplet surface due to high electrical fields and/or the formation of very small droplets due to repetitive droplet fission until each droplet contains, on average, only one analyte ion.

Electrospray ionization

During standard electrospray ionization, the sample is dissolved in a polar, volatile solvent and pumped through a narrow, stainless steel capillary. A high voltage of 3 or 4 kV is applied to the tip of the capillary. As a consequence of this strong electric field, the sample emerging from the tip is dispersed into an aerosol of highly charged droplets. The charged droplets diminish in size by solvent evaporation, assisted by a flow of drying gas. Eventually charged sample ions, free from solvent, are released from the droplets.



All the above mentioned ionization processes are fundamentally different techniques; but they achieve essentially the same end result — the generation of gas phase ions via non-destructive vaporization and ionization.

Mass analysis

Once the sample has been ionized, the beam of ions is accelerated by an electric field and then passes into the mass analyzer, the region of the mass spectrometer where the ions are separated according to their mass-to-charge ratio. Just like ionization methods, there are several types of mass analyzer. The most common type of mass analyzer is TOF (time-of-flight) mass analyzer. The TOF mass analyzer measures ion flight time. It is based on the simple idea that velocities of two ions with the same kinetic energy will vary depending on the mass of an ion – the lighter ion will have higher velocity. Because the ions have different velocities, the ions reach the detector at different times. The smaller ions reach the detector first because of their greater velocity as compared to the larger ions. Hence, the analyzer is called TOF because the mass is determined from the ion's time-of-flight. The arrival time at the detector is dependent upon the mass, charge and kinetic energy of the ion.

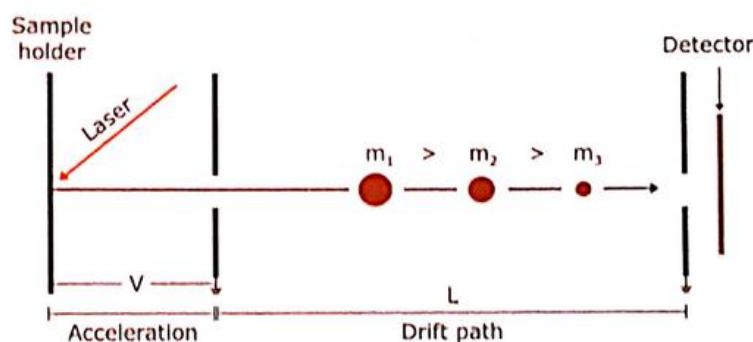


Figure 4.4 Schematic representation of a MALDI/TOF mass spectrometer.

The kinetic energy of an ion accelerated through an electrical potential, V, is:

$$E = zV = \frac{mv^2}{2} \quad \dots(1)$$

where, E = kinetic energy,

V = electrical potential,

v = velocity of an ion,

z = charge of an ion,

m = mass of an ion.

The velocity of the ion is the length of the flight path (L) divided by the time (t) it takes the ion to travel over the distance;

$$v = \frac{L}{t}$$

Replacing the value of v in equation (1) gives:

$$zV = \frac{mL^2}{2t^2}$$

$$\frac{m}{z} = \frac{2Vt^2}{L^2}$$

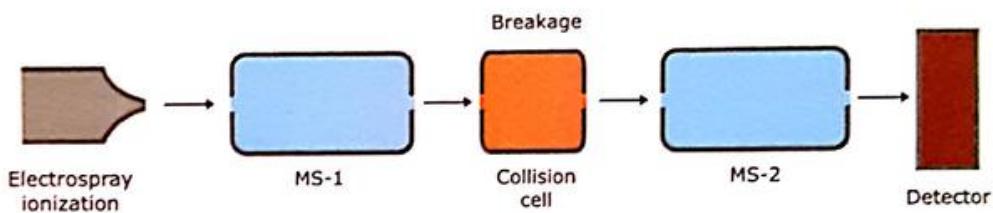
Absolute masses
A mass spectrometer does not measure absolute mass, M . The instrument needs to be calibrated with standard compounds, whose M values are known very accurately.

Application

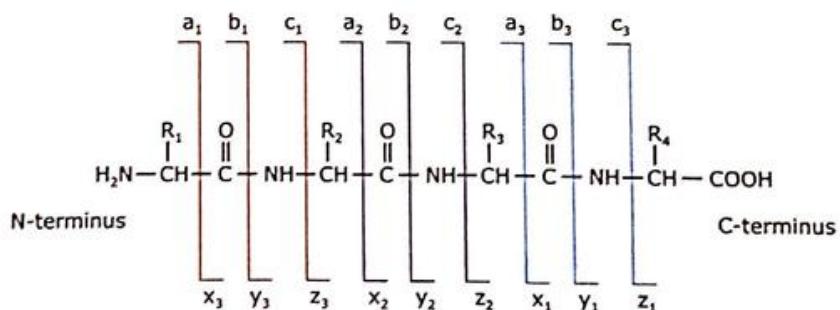
Mass spectrometry is used to measure the molecular mass of biomolecules, polypeptides and nucleic acids sequencing and to determine protein structure. The increased sensitivity and resolution of the instruments, coupled with improvements in the analysis of data, have opened new dimensions in analyses of complex biological systems. It includes drug metabolism, lipid analysis, metabolomics, quantitative proteomics, direct analysis of intact proteins and imaging of proteins in tissues.

Sequence information of a polypeptide is extracted using a technique called tandem MS or MS/MS. MS/MS can sequence not only 20 common amino acids but also known or unknown modified amino acids according to their mass. MS/MS is fast, sensitive and can analyze peptide mixtures directly.

In this process, a solution containing the protein under investigation is first treated with a protease or chemical reagent to hydrolyze it to a mixture of shorter peptides. The mixture is then injected into a device that is essentially two mass spectrometers in tandem.

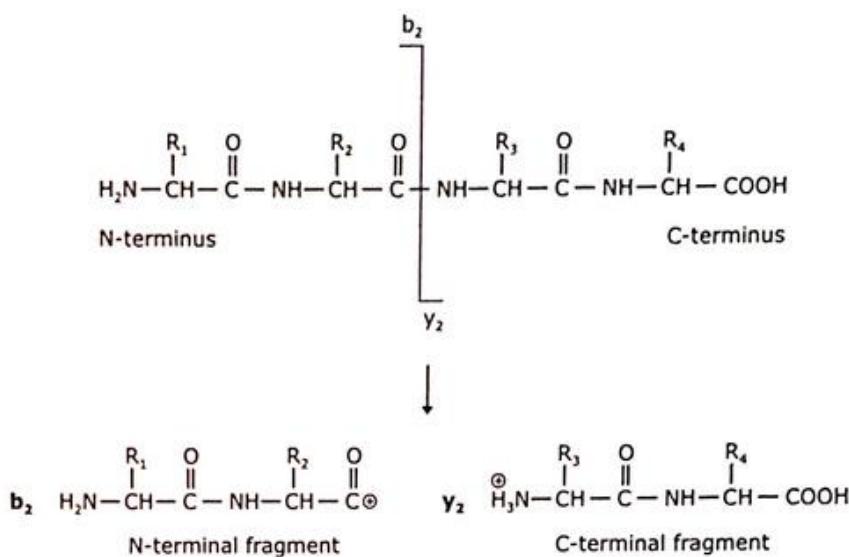


After proteolytic hydrolysis, a protein solution is injected into a mass spectrometer (MS-1). The different peptides are sorted so that only one type is selected for further analysis. The selected peptide is further fragmented in a chamber between the two mass spectrometer. In this collision cell, the peptide is further fragmented by high-energy impact with a collision gas. The m/z for each fragment is measured in the second mass spectrometer (MS-2). There are three different types of bonds that can break along the amino acid backbone: the NH-CH, CH-CO, and CO-NH bonds. Each bond breakage gives rise to two species – one neutral and the other one charged, and only the charged species is monitored by the mass spectrometer. The fragmentation of a peptide is shown in the figure below.



The most common cleavage sites are at the CO-NH bonds (peptide bonds). However, bonds other than the peptide bond can be broken in the fragmentation process. The breakage of the peptide bonds gives two prominent sets of charged fragments. Those fragments that appear to extend from the amino terminus are termed 'b ions'. The fragment containing only the amino terminal amino acid is termed b_1 . The fragment containing the first two amino terminal amino acids is termed b_2 ion, and so forth. Similarly, groups of peptide fragment ions appear to extend from the C-terminus are termed 'y ions'. Because the bond-breaking in the collision cells does not yield full carboxyl and amino groups at the sites of the breaks, the only intact

α -amino and α -carboxyl groups on the peptide fragments are those at the very ends. The two sets of fragments can, thereby, be identified by the resulting slight differences in mass.



From a single, complete ion series we can deduce the amino acid sequence of a peptide, although leucine and isoleucine, and lysine and glutamine, cannot be differentiated in this manner. By comparing the differences between successive masses in each of the series and comparing them to the known masses of the amino acid residue, one can determine the amino acid sequence of the peptide.

An example of peptide sequencing by tandem mass spectrometry is illustrated below. The peptide is fragmented along the peptide backbone to form b and y-ions and resulting fragment ions are measured to produce the MS/MS spectrum.

Let us consider a peptide Ala-Val-Ala-Gly-Cys-Ala-Gly-Ala-Arg to understand the concept. If one can identify either the y-ion or b-ion series in spectrum, the peptide sequence can be determined. The main idea during sequencing is to use the mass difference between two fragment ions to calculate the mass of an amino acid residue on the peptide backbone.

The predicted b- and y-ion fragmentations for above mentioned peptide Ala-Val-Ala-Gly-Cys-Ala-Gly-Ala-Arg (AVAGCAGAR) is illustrated below. Starting from the N-terminus, cleavages are generating an ascending series of fragment ion corresponding to the b-series and a descending order of y-series fragments. For example, the b_4 - and y_5 - fragments are formed by the cleavage of peptide bond between Gly and Cys.

- b_1 A VAGCAGAR y_8
- b_2 AV AGCAGAR y_7
- b_3 AVA GCAGAR y_6
- b_4 AVAG CAGAR y_5
- b_5 AVAGC AGAR y_4
- b_6 AVAGCA GAR y_3
- b_7 AVAGCAG AR y_2
- b_8 AVAGCAGA R y_1

The mass difference between adjacent members of a series can be calculated. For example, the mass differences between $y_7 - y_6 = 605.3 - 534.3 = 71$ Da which is equivalent to amino acid residue alanine; and similarly $y_6 - y_5 = 534.3 - 477.0 = 57$ Da which is equivalent to a

A mass spectrum is a plot of the intensity as a function of mass-to-charge ratio. The peak in the spectrum with highest intensity is called the base peak.

glycine residue. The complete γ -ions from y_6 to y_1 establishes the VAGCAGAR motif. The complete b-ions from b_6 to b_1 corresponds to the AVAGCAGA motif. Combining the results of the γ - and b-ion series, the MS-MS spectrum provides definitive confirmation of the sequence AVAGCAGAR.

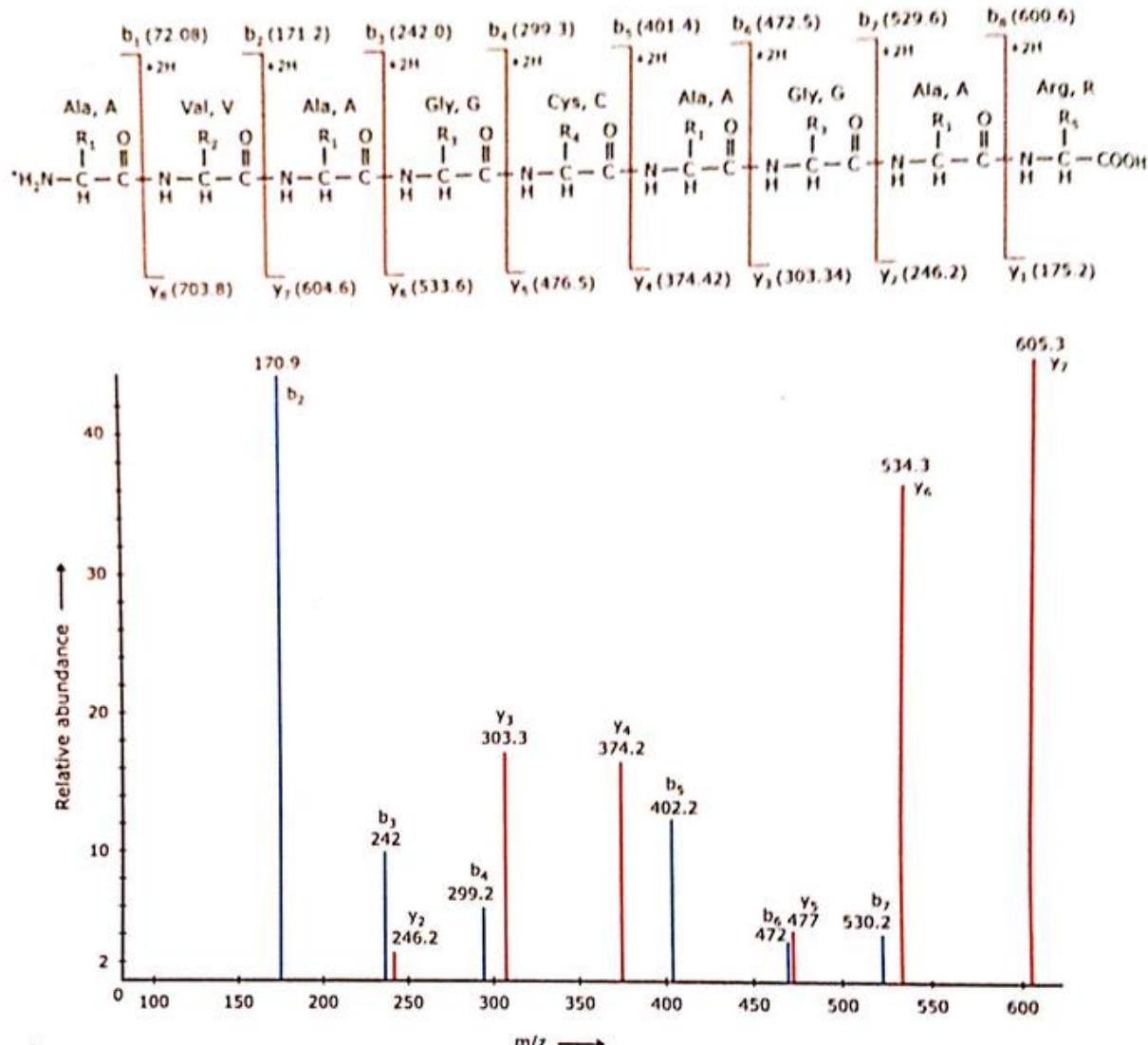


Figure 4.5 In an MS/MS, many copies of the same peptide are fragmented at the peptide backbone to form b- and y-ions. The spectrum consists of peaks at the m/z (mass-to-charge) values of the corresponding fragment ions. The b-series ions have been labeled with blue vertical lines and the y-series ions have been labeled with red vertical lines.

References

- Banwell CN, McCASH EM (2013), *Fundamentals of Molecular Spectroscopy*, 5th ed. McGraw Hill Education.
- Berg JM, Tymoczko JL and Stryer L (2007), *Biochemistry*, 6th ed. W.H. Freeman and Company.
- Branden C and Tooze J (1999), *Introduction to Protein Structure*, 2nd ed. Garland Publishing Inc.
- Domon B and Aebersold R (2006), *Mass Spectrometry and Protein Analysis*. *Science* 312: 212-217.
- Hoffmann E and Stroobant V (2007), *Mass Spectrometry: Principles and Applications*, 3rd ed. John Wiley and sons.
- Liebler DC (2002), *Introduction to Proteomics: tools for the new biology*. Humana Press Inc.
- White FA and Wood GM (1986), *Mass Spectroscopy, Applications and Engineering*, John Wiley.

Chapter 5

Centrifugation

Centrifugation is a process used to separate or concentrate materials suspended in a liquid medium. It is a method to separate molecules based on their sedimentation rate under centrifugal field. It involves the use of the centrifugal force for the sedimentation of molecules. It is also used to measure physical properties (such as molecular weight, density and shape) of molecules. If centrifugation is used for separation of one type of material from others; it is termed as preparative centrifugation; whereas if it is used for measurement of physical properties of macromolecules then termed as analytical centrifugation.

Principle of centrifugation

Particles suspended in a solution are pulled downward by Earth's gravitational force. In a solution, particles whose mass or density is higher than that of the solvent sink or sediment, and particles that are lighter than it float to the top. The greater the difference in mass or density, the faster they sink. This sedimentation movement is partially offset by the buoyancy of the particle. Because the Earth's gravitational field is weak, a solution containing particles of very small masses usually remain suspended due to the random thermal motion. However, sedimentation of these particles can be enhanced by applying centrifugal forces. A centrifuge does the same thing. It increases the sedimentation by generating centrifugal forces as great as 1,000,000 times the force of gravity.

Let us consider a solution being spun in a centrifuge tube. The centrifugal force acting on a solute particle of mass m ,

$$\text{Centrifugal force} = m\omega^2 r$$

$$m\omega^2 r \quad v = \frac{\text{radians}}{\text{second}}$$

where, ω is the angular velocity in radians per second,

r is the distance from the center of rotation to the particle, and

$\omega^2 r$ is the centrifugal acceleration.

A particle will move through a liquid medium when subjected to a centrifugal force. Hence, we must also consider the particle's buoyancy due to the displacement of the solvent molecules by the particle. This buoyancy reduces the force on the particle by $\omega^2 r$ times the mass of the displaced solvent.

$$\text{Buoyant force} = m_0 \omega^2 r$$

where, m_0 is the mass of fluid (solvent) displaced by the particle

$$m_0 = \bar{v} \rho r$$

Where, ρ is the density of the solution (g/mL). \bar{v} is the partial specific volume of the particle. As measuring the volume of a very small particle is difficult, for convenience we use a term called partial specific volume. \bar{v} is the volume in mL that each gram of the particle occupies in solution.

Thus, the net force acting on the particle downward is given by,

$$\text{Net force} = \text{Centrifugal force} - \text{Buoyant force}$$

$$= m \omega^2 r - m_0 \omega^2 r$$

$$= m \omega^2 r - \bar{v} \rho \omega^2 r$$

When particles move downward through the solution, the motion is also opposed by the frictional force. The frictional force is equal to the product of the frictional coefficient, f (Nm⁻¹s) and the sedimentation velocity, v . It acts in the opposite direction to the net force:

$$F_{\text{friction}} = v \cdot f$$

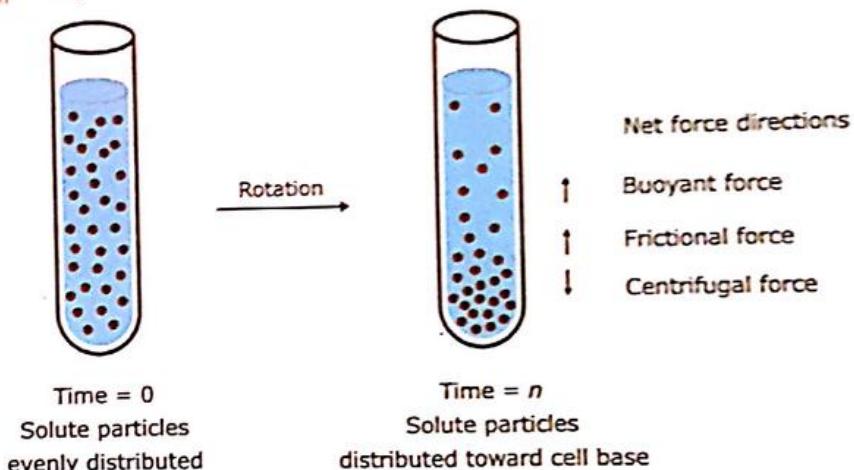


Figure 5.1 A schematic diagram of a centrifugation experiment.

At steady state, then, the *frictional force* is equal to the net force and the molecule moves with velocity v downward:

$$vf = m \omega^2 r - \bar{v} \rho \omega^2 r$$

$$vf = m \omega^2 r (1 - \bar{v} \rho)$$

Since, m (mass in grams of a single particle) = M/N

M is the molar weight of the solute in g/mol

N is Avogadro's number.

$$fv = \frac{M \omega^2 r (1 - \bar{v} \rho)}{N}$$

Now, the *sedimentation coefficient*,

$$s = \frac{v}{\omega^2 r} = \frac{M}{N} \frac{(1 - \bar{v} \rho)}{f}$$

The *sedimentation coefficient* is the ratio of a velocity to the centrifugal acceleration. The sedimentation coefficient has units of second. A sedimentation coefficient of 1×10^{-13} second is defined as one Svedberg, S , ($1S = 10^{-13}$ second). This unit is named for The Svedberg, a pioneer in the field of centrifugation.

Important conclusions drawn from the equation:

- The sedimentation velocity of a particle is proportional to its mass.
- A dense particle moves more rapidly than a less dense one because the opposing buoyant force is smaller for a dense particle.
- Shape, too, is important because it affects the viscous drag. The frictional coefficient of a compact particle is smaller than that of an extended particle of the same mass.
- The sedimentation velocity depends also on the density of the solution (ρ). Particle sinks when $\bar{v}\rho < 1$, floats when $\bar{v}\rho > 1$.

Sedimentation coefficients of biological macromolecules are normally obtained in buffered solutions whose viscosity and density may differ from those of water. It is also measured at different temperatures. Hence, we standardize the sedimentation coefficient value in standard conditions i.e. pure water at 20°C (denoted by $s_{20,w}$).

$$s_{20,w} = s_{exp} \frac{1 - \bar{v}\rho_{20,w}}{1 - \bar{v}\rho_{exp}} \frac{\eta_{exp}}{\eta_{20,w}}$$

where, $s_{20,w}$ is the sedimentation coefficient of the molecule in pure water at 20°C,

s_{exp} is the experimentally measured sedimentation coefficient of the molecule,

η_{exp} is the viscosity of the solvent at the experimental temperature T (°C),

$\eta_{20,w}$ is the viscosity of water at 20°C,

$\rho_{20,w}$ is the density of water at 20°C,

ρ_{exp} is the density of the solvent at given temperature T (°C) and

\bar{v} is the partial specific volume of the molecule.

Relative centrifugal fields are used because different centrifuge rotors have different geometries and the appropriate rpm with one rotor may not be correct if a different rotor is used. It is, therefore, important to know the value of r for the rotor being used. But r is not constant throughout a centrifuge tube, so different parts of the sample will be subjected to different RCFs. Usually the important figure is the r_{max} , which is the maximum radial distance for the rotor. If the rotor is angular or a swinging-bucket type, then r_{max} corresponds to the bottom of the centrifuge tube; if the rotor is an upright one, then r_{max} gives the distance to the outer wall of the tube.

Relative centrifugal field

Particles suspended in a fluid move, under the influence of gravity, towards the bottom of a vessel at a rate that depends, in general, on their size and density. Centrifugation utilizes centrifugal forces which are greater than the Earth's gravitational force to increase the sedimentation rate of particles. This is achieved by spinning the vessel containing the fluid and particles about an axis of rotation so that the particles experience a centrifugal force acting away from the axis. The force is measured in multiples of the Earth's gravitational force and is known as the relative centrifugal field (RCF) or more commonly, the g force. For example, an RCF of 500 xg indicates that the centrifugal force applied is 500 times greater than Earth's gravitational force.

The RCF generated by a rotor depends on the speed of the rotor in revolutions per minute (rpm) and the radius of rotation (i.e. the distance from the axis of rotation). The equations that permit calculation of the RCF from a known rpm and radius of rotation is:

$$RCF = \frac{r\omega^2}{g}$$

where, r is the radius in centimeters,

g is the acceleration due to gravity (cm/sec²) and

ω is the angular velocity.

Angular velocity can be defined as $\omega = \frac{2\pi \times rpm}{60}$ radians / sec

$$RCF = \frac{r}{g} \times \left(\frac{2\pi \times rpm}{60} \right)^2 = r \times (rpm)^2 \times \left(\frac{4 \times 3.14 \times 3.14}{60 \times 60 \times 980} \right) = 1.12 \times r \times (rpm)^2 \times 10^{-5}$$

The force applied to the samples varies according to the size of the centrifuge as a larger centrifuge will have a higher radius and a smaller centrifuge will have a shorter radius. For example, when revolving at 2000 rpm, a larger centrifuge with a longer radius length will spin samples at a higher g force than a smaller centrifuge with a shorter radius length.

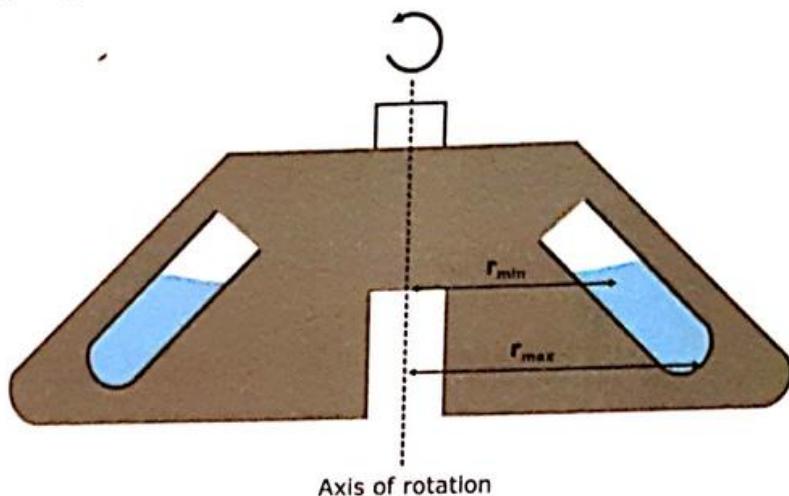


Figure 5.2 A diagram illustrating the variation of RCF with r , the distance of the sedimenting particles from the axis of rotation.

Differential centrifugation

Differential centrifugation separates particles based on differences in sedimentation rate, which reflect differences in sizes and densities. This centrifugation process is used mainly for the separation of sub-cellular components.

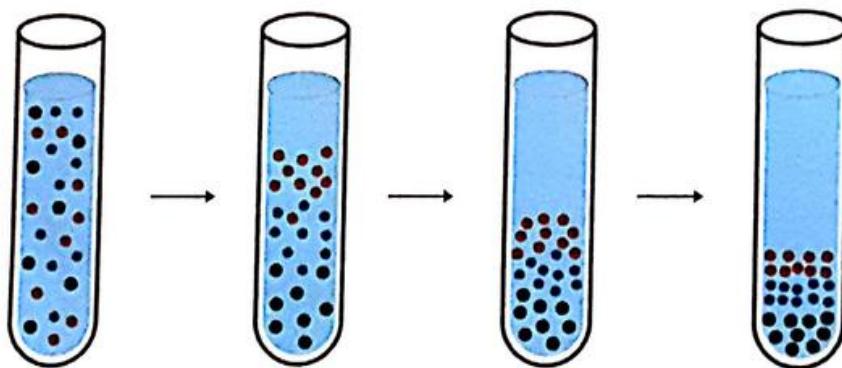


Figure 5.3 Differential centrifugation separates particles based on differences in sedimentation rate, which reflect differences in size and/or density. Particles which are large or dense sediment rapidly, those that are intermediate in size or density sediment less rapidly, and the smallest or least dense particles sediment very slowly. Eventually, all of the particles reach the bottom of the tube.

In the process of separation of sub-cellular components, the preparation of broken cells is poured into a centrifuge tube and is initially centrifuged at low centrifugal force long enough to completely sediment the largest and heaviest sub-cellular component. The supernatant obtained is carefully decanted and is again centrifuged at a higher centrifugal force for sedimenting the next heavier entity in the extract. This process is continued and at each ensuing step the centrifugal force as well as the centrifugal time is increased to successively sediment the lighter components and particles. In this method, the separation occurs due to the differential sedimentation rates of the sub-cellular organelles because of the differences in their sizes and densities.

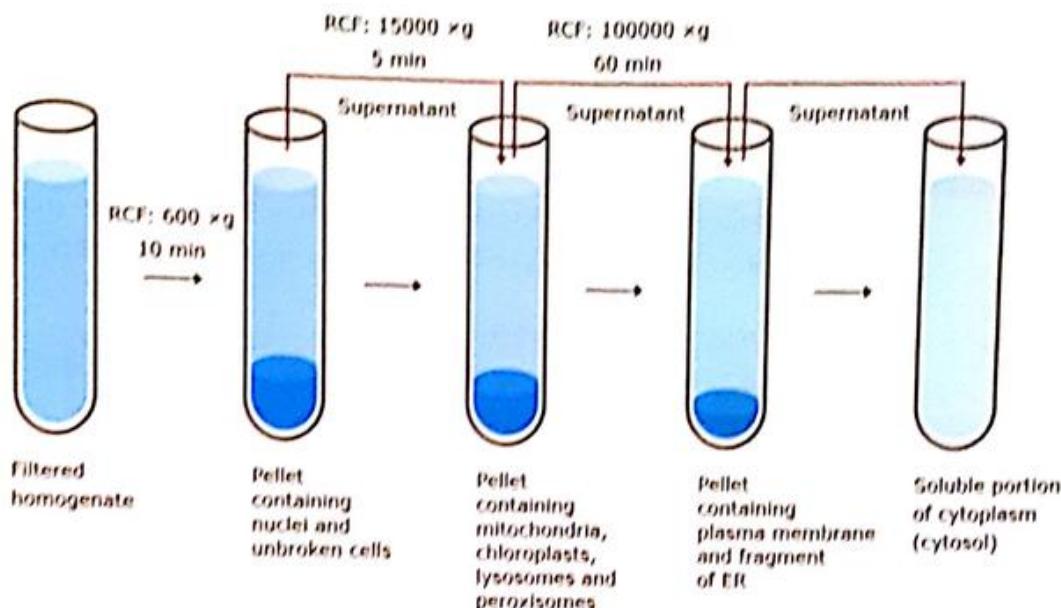


Figure 5.4 Differential centrifugation and separation of cell organelles. The different sedimentation rate of various cellular components make their separation possible. The tissue of interest is first homogenized. Subcellular fractions are then isolated by subjecting the homogenate and subsequent supernatant fractions to successively higher centrifugal forces and longer centrifugation times. During each step of the centrifugation process, particles of a given size and density are removed as a pellet from suspension. Each time, the supernatant from one step is decanted into a new centrifuge tube and subjected to greater centrifugal force to obtain the next pellet.

✓ Density gradient centrifugation

In differential centrifugation, the particles about to be separated were uniformly distributed throughout the solution prior to centrifugation. Density gradient centrifugation is a variation of differential centrifugation in which the sample is centrifuged in a medium that gradually increases in density from top to bottom. The gradient consists of an increasing concentration of solute (and therefore density) from the top of the tube to the bottom. It is of two types: *Rate zonal centrifugation*, in which the sample is centrifuged in a preformed gradient and *Isopycnic centrifugation*, in which a self-generating gradient forms during centrifugation.

✓ Rate zonal centrifugation

Rate zonal centrifugation (also termed as *velocity centrifugation*) is used to separate particles on the basis of differences in their sedimentation rate. Separation of particles occurs according to their size and/or density. Although the rate of sedimentation is strongly influenced by the size (mass) and density of particles, even slight variation in shape also affects the rate of sedimentation. In zonal centrifugation method, materials used for the preparation of density gradients are sucrose, glycerol, ficoll etc. A 5–20% sucrose solution is commonly used to form density gradient. The density range is chosen so that the density of the particles is greater than the density of the medium at all points during the separation.

In this centrifugation process, samples are centrifuged just long enough to separate the molecules of interest into discrete zones. If a sample is centrifuged much longer than necessary, all the components of sample will end up in a pellet at the bottom of the tube.

In addition to the preparation and purification of macromolecules and cellular components, rate zonal centrifugation can be used to determine the sedimentation coefficients and molecular weights of biological macromolecules. If a purified molecule such as a protein is spun

Ultracentrifuges are capable of speeds in excess of 30,000 rpm and RCFs of over 600,000 xg.

in a centrifugal field, the molecule will eventually sediment towards the bottom at a constant velocity. At this point, the molecular weight (M) can be calculated as:

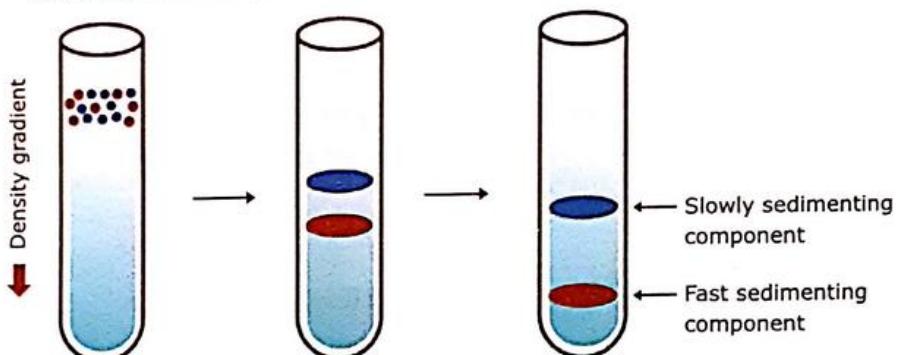
$$M = \frac{f \times v}{\omega^2 r}$$

Where, f is the *frictional coefficient* of the solvent system (which has been calculated from other measurements) and $v/\omega^2 r$ is the *sedimentation coefficient*.

Isopycnic centrifugation

In contrast to rate zonal centrifugation, isopycnic centrifugation (or *equilibrium density gradient centrifugation*) separates particles solely on the basis of buoyant density and is independent of shape and size of particles. It is also independent of time of centrifugation. This technique is used to separate particles of similar size, but different density. Sedimentation of particles in density gradient occurs until the buoyant density of particle is equal to the density of the gradient. A steep density gradient that contains a very high concentration of sucrose (20–70%) or CsCl is generally used.

a. Rate zonal centrifugation



b. Isopycnic centrifugation

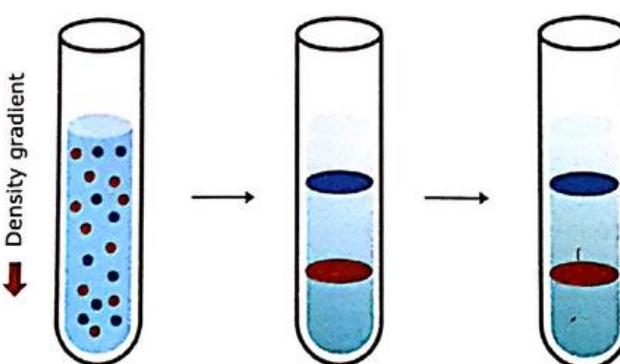


Figure 5.5 (a) Rate-zonal centrifugation is a variation of differential centrifugation in which the sample for fractionation is placed as a thin layer on top of a gradient of solute. The gradient consists of an increasing concentration of solute—and therefore density—from the top of the tube to the bottom. When subjected to a centrifugal force, particles differing in size and/or density move downward as discrete bands, that migrate at different rates. Because of the gradient of solute in the tube, the particles at the leading edge of each zone continually encounter a slightly denser solution and are, therefore, slightly impeded. As a result, each zone remains very compact, maximizing the resolution of different particles.

(b) Isopycnic centrifugation also includes a gradient of solute that increases in concentration and density, but in this case the solute is concentrated so that the density gradient spans the range of densities of the particles about to be separated. During centrifugation, the particles move into the gradient until each reaches its equilibrium (or buoyant) density—the point in the gradient at which the density of the particle is exactly equal to the density of the gradient.

Table 5.1 Density gradient media

	CsCl	Sucrose
Molecular weight	168.36	342.30
Solubility (g ml ⁻¹ in water, 20°C)	1.2	2.0
Maximum density (g ml ⁻¹ in water, 20°C)	1.91	1.32

The zonal and equilibrium centrifugation methods are often confused because both utilize gradients through which the sedimenting molecules move. However, the density gradient in rate zonal centrifugation plays no direct role in the separation of the sedimenting molecules while the density gradient performs the separation in equilibrium centrifugation. Another distinction is that when the centrifugation time of a zonal centrifugation is too long, all of the sedimenting molecules will be found on the bottom of the centrifuge tube. In contrast, lengthening the centrifugation time in equilibrium centrifugation will have no effect, because the distribution of the molecules is at equilibrium.

DNA separation by equilibrium density gradient centrifugation

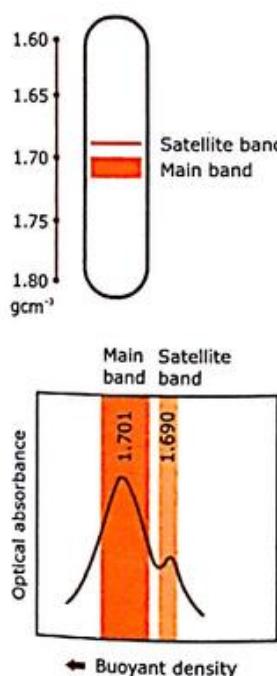
Equilibrium density gradient centrifugation with caesium chloride (CsCl) is a useful tool for fractionating, quantitatively separating and characterizing DNA based on differences in their buoyant densities. CsCl is used because, at a concentration of 1.6 to 1.8 g/mL, it is similar to the density of DNA. During centrifugation, a gradient of the caesium ions is formed.

The DNA molecules having differences in the relative proportions of AT (adenine and thymine base pairs) to GC (guanine and cytosine base pairs) can be separated by CsCl density gradient centrifugation. An AT base pair has a lower molecular weight than a GC base pair and therefore, for two DNA molecules of equal length, the one with the greater proportion of AT base pairs will have a lower density, if all other factors being equal. The density (ρ) of DNA is related to the GC content. A linear relationship exists between the densities of DNA and their GC content.

$$\rho = 1.660 + 0.098 (G+C) \text{ g/cm}^3$$

where, G+C is the mole fraction of G+C in the dsDNA.

ρ is the Buoyant density of DNA.



Mouse DNA is separated into a main band and a satellite band by centrifugation through a density gradient of CsCl.

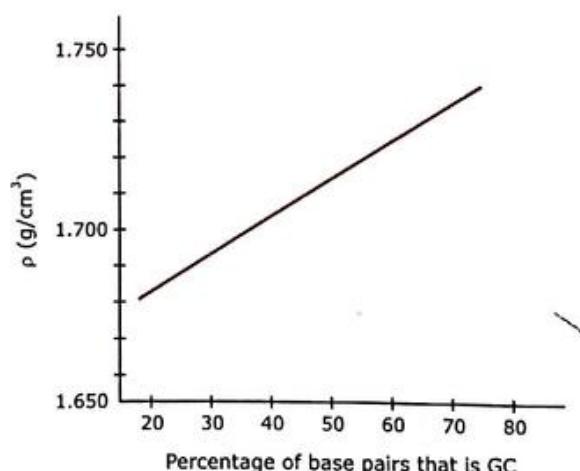


Figure 5.6 Density of DNA as a function of GC content.

The buoyant density of most double stranded linear DNA in CsCl solution is ~1.70 g/cm³. In CsCl solution, the density of ssDNA is ~0.015-0.020 g/cm³ greater than that of dsDNA of the same base composition. The buoyant density of ssRNA in CsCl solution is >1.8 g/cm³.

EtBr is a planar compound which intercalates between base pairs in the DNA double helix. Each molecule of EtBr which intercalates causes the double helix to unwind, decreasing T_w and increasing W_t . When one molecule of EtBr intercalates in the DNA double helix, the helix unwinds by 26° . Addition of 14 EtBr molecules to a DNA molecule results in the unwinding of one full turn. Thus, T_w decreases by one.

Difference in density between highly repetitive satellite DNA from rest of DNA is used for separation of these two by CsCl density gradient centrifugation. DNA that consists of very large numbers of short tandem repeats (termed satellite DNA) may have a base composition (and thus GC content) different from that of the genome as a whole. If so, the satellite DNA will have a different buoyant density from the rest of the DNA, as this property depends on the base composition. DNA may be fractionated according to density by CsCl density gradient centrifugation. Each fraction of DNA forms a band at the position corresponding to its own density. If the GC content varies by 5% or more, separate bands are obtained. For example, when fragments of mouse DNA is centrifuged on a CsCl density gradient, two DNA bands are seen. One contains 92% of the DNA with a density of 1.701 gm/cm^3 (GC content $\sim 42\%$) and the thin satellite band contains 8% of the DNA with a density of 1.690 gm/cm^3 (GC content $\sim 30\%$). Satellite DNA was originally defined by this density separation. However, in cases where the average satellite DNA base composition is close to that of the genome as a whole, the satellite DNA cannot be physically separated using a density gradient.

In laboratory, CsCl equilibrium density gradient centrifugation is also used for the separation of linear DNA (non-supercoiled) molecules from circular supercoiled DNA. It is carried out in the presence of DNA intercalating dye EtBr. DNA-EtBr intercalation causes the unwinding of the DNA helix, which reduces the buoyant density of DNA, by as much as 0.125 g/cm^3 for linear DNA. However, covalently closed circular supercoiled DNA, with no free ends, has very little freedom to unwind, and can only bind a limited amount of EtBr. The decrease in buoyant density of a circular supercoiled DNA is much less, only about 0.085 g/cm^3 . In contrast, linear DNA molecules with free ends are not as topologically constrained and can, therefore, bind more EtBr molecules, resulting in a more decrease in buoyant density. As a consequence, covalently closed circular supercoiled DNA form a band in an EtBr-CsCl gradient at a different position to linear DNA.

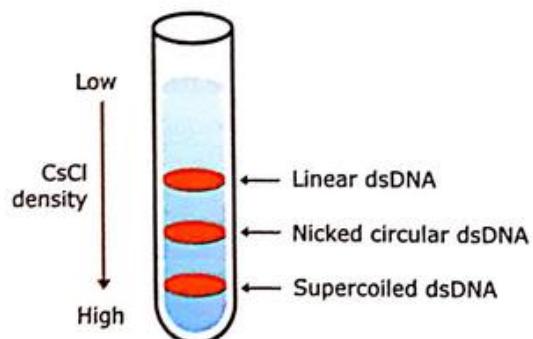


Figure 5.7 Separation of supercoiled DNA from non-supercoiled DNA molecules by density gradient centrifugation in the presence of ethidium bromide (EtBr) is shown here. EtBr binds to DNA molecules by intercalating between adjacent base pairs, causing partial unwinding of the double helix. This unwinding results in a decrease in the buoyant density. However, supercoiled circular DNA, with no free ends, has very little freedom to unwind, and can only bind a less amount of EtBr as compared to linear and nicked circular dsDNA. The decrease in buoyant density of a supercoiled molecule is, therefore, much less. As a consequence, supercoiled molecules form a band in an EtBr-CsCl gradient at a different position to linear and nicked-circular DNA.

Chapter 6

Microscopy

Microscopy is a technique for making very small things visible to the unaided eye. An instrument used to make the small things visible to the naked (unaided) eye is called a microscope. There are two fundamentally different types of microscopes: the light microscope and the electron microscope.

6.1 Light microscope

Light or optical microscope uses visible light as a source of illumination. Because the light travels through the specimen, this instrument can also be called a transmission light microscope. The light microscope creates a magnified image of specimen which is based on the principles of transmission, absorption, diffraction and refraction of light waves.

The simplest form of light microscope consists of a single lens, a magnifying glass. Microscope made up of more than one glass lens in combination is termed compound microscope. Compound microscope includes condenser lens, the objective lens and the eyepiece lens. Condenser lens focuses the light from the light source at the specimen. The one facing the object is called the objective and the one close to the eye is called the eyepiece. The objective has a smaller aperture and smaller focal length than those of the eyepiece (also referred to as the ocular).

The objective lens is responsible for producing the magnified image. It is available in different varieties (4x, 10x, 20x, 40x, 60x, 100x). The power of a lens is described with a number followed by the letter 'x'. For example, if through a microscope one can see something 25 times larger than actual size, its magnification power is 25x. The eyepiece works in combination with the objective lens to further magnify the image. A compound microscope with a single eyepiece is said to be monocular and one with two eyepieces is said to be binocular. Eyepieces usually magnify by 10x, since an eyepiece of higher magnification merely enlarges the image, with no improvement in resolution.

Both living and dead specimens are viewed with a light microscope. The visibility of the magnified specimen depends on contrast and resolution. In general, variation in the light intensity (contrast) within an image occurs because different parts of the specimen absorb light to differing degrees. Brightness contrast arises from different degrees of absorption at different points in the specimen. Color contrast can also arise from absorption when the degree of absorption depends on the wavelength and varies from point-to-point in the specimen.

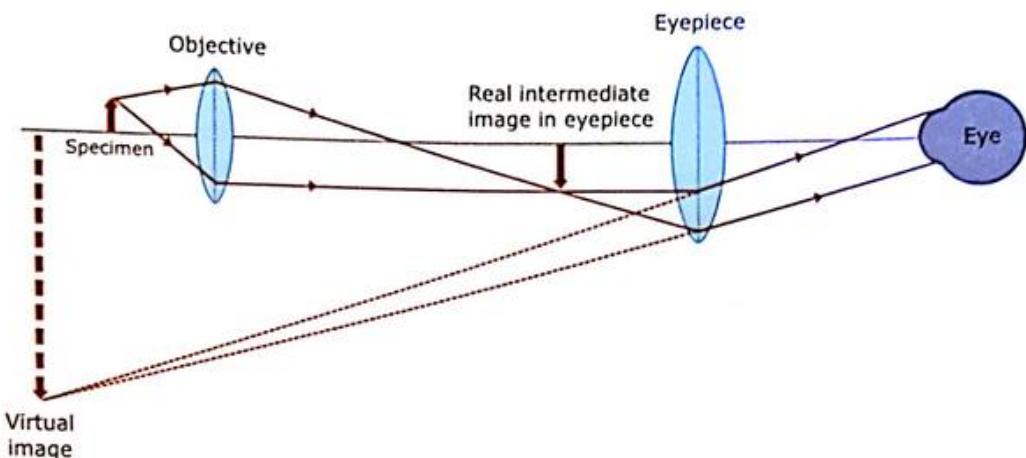


Figure 6.1 Compound microscope. The objective lens forms a real and inverted magnified image of the specimen or object (called the real intermediate image) in the focal plane of eyepiece. This image works as an object for the eyepiece. The final image is, then, formed at infinity. It is erect with respect to the first image and hence, inverted with respect to the object.

For colorless specimens, as is the case for most biological material, contrast is achieved in various ways. The specimen may be stained, thus reducing the amplitude of certain light waves passing through the stained areas. However, this usually requires the fixation and staining of specimens. But fixation and staining kills the specimens. For colorless living specimens, contrast can be achieved in different ways. A living biological specimen is almost uniformly transparent, and therefore the intensity variation in the image will be poor. However, the light does not go through the specimen unaffected. The light going through the specimen will be shifted in phase due to scattering and diffraction. Our eyes cannot detect this phase shift. However, using phase contrast microscopy, we can develop a contrast based on the phase shift.

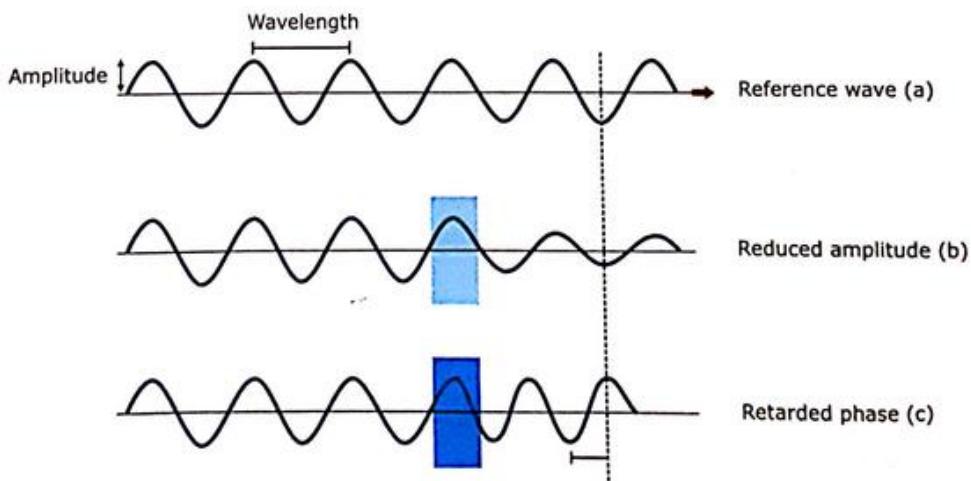


Figure 6.2 Changes in amplitude and phase of light. (a) Reference wave with characteristic amplitude, wavelength and phase. (b) An object absorbs energy and reduces the amplitude, but does not alter the phase of an emergent ray. (c) An object alters velocity and shifts the phase, but not the amplitude of an emergent ray.

Magnification

The magnification or linear magnification of a microscope is defined as the ratio of the image size to the object (specimen) size. If the image and object are in the same medium, then it is just the image distance divided by the object distance. There is a difference in the meaning of the two terms, *magnification* and *magnifying power*. Magnifying power or angular magnification is the ratio of the angle subtended by object and image.

$\lambda > \frac{d}{m}$

The magnification of a compound microscope is the product of the magnification of the objective and the eyepiece. The magnification of the objective is called the linear magnification, because it is measured in linear dimensions. The magnification of the eyepiece is called the angular magnification. The overall magnification is the product of the linear magnification of the objective lens and the angular magnification of the eyepiece with the first image at the focal length.

Resolving power

Resolving power is the ability of magnifying instrument to distinguish two objects that are close together. The resolving power is inversely related to the *limit of resolution*. The limit of resolution is defined as the minimum distance between two points that allows for their discrimination as two separate points. Thus, the higher the resolving power, the smaller the limit of resolution. The limit of resolution of the light microscope depends upon the three factors: the *wavelength* (λ) of the light used to illuminate the specimen, the *angular aperture* (α) and the *refractive index* (n) of the medium surrounding the specimen. The effect of these three variables on the limit of resolution is described quantitatively by the following equation known as the *Abbe equation*:

$$\text{Limit of resolution} = \frac{0.61\lambda}{n \times \sin \alpha}$$

The quantity $n \times \sin \alpha$ is called the *numerical aperture* of the objective lens, abbreviated NA. The NA is a measure of the ability of a lens to collect light from the specimen. Lenses with a low NA collect less light than those with a high NA.

$$\text{Limit of resolution} = \frac{0.61\lambda}{\text{NA}}$$

For small value of limit of resolution, the numerator of the equation should be as small as possible and the denominator should be as large as possible. Equation shows that resolution can be improved by shortening the wavelength of the illuminating light, increasing the index of refraction on the objective lens, and increasing $\sin \alpha$. The angle α can be increased either by shortening the distance between the lens and the object or by increasing the diameter of the lens.

Magnification is how large the image is as compared to real specimen size, whereas resolution is the amount of information that can be seen in the image – defined as the smallest distance below which two discrete objects will be seen as one.

Magnification is a function of the number of lenses. Resolution is a function of the ability of a lens to gather light.

The maximum magnification of compound light microscopes is usually 1500x and has a limit of resolution of about 0.2 μm .

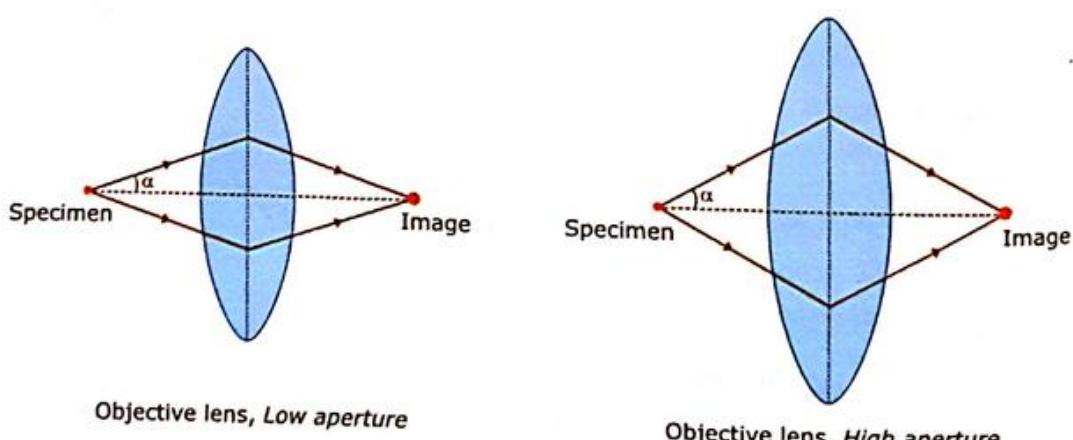


Figure 6.3 The angular aperture of a lens. The angular aperture is the half-angle α of the cone of light entering the objective lens of the microscope from the specimen. It is, therefore, a measure of how much aperture, the more information the lens can transmit.

For minimum value of numerator, the wavelength should be small. Thus, for the best resolution, specimen is illuminated with blue light of 450 nm. The angular aperture for the best objective lenses is about 70° . Hence, the maximum value for $\sin\alpha$ is about 0.94. The refractive index of air is about 1.0, so for a lens designed for use in air, the maximum numerical aperture is about 0.94. In this situation, the limit of resolution for a glass lens in air is roughly 300 nm. To increase the numerical aperture some microscope lenses are designed to be used with a layer of *immersion oil* between the lens and the specimen. Immersion oil has a higher refractive index than air and, therefore, allows the lens to receive more of the light transmitted through the specimen. Since the refractive index of immersion oil is about 1.5, the maximum numerical aperture for an oil immersion lens is about $1.5 \times 0.94 = 1.4$.

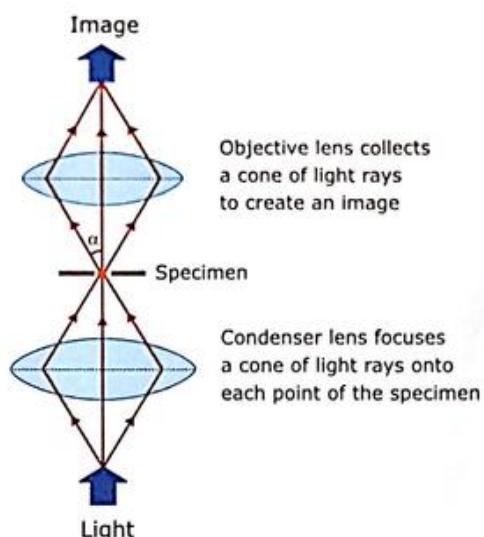


Figure 6.4 The angular aperture is the half-angle α of the cone of light entering the objective lens of the microscope from the specimen.

Thus, the limit of resolution for a microscope that uses visible light is roughly 300 nm in air and 200 nm with an oil immersion lens. The limit of resolution of the unaided human eye is 100 μm .

Oil immersion lens

In most microscopes, air is present as a medium through which light rays pass between the coverslip protecting the sample and front lens of the objective. Objectives of this type are referred to as *dry objectives*. Air has a refractive index of 1.0003, very close to that of a vacuum and considerably lower than most liquids, including water ($n = 1.33$). In light microscopy, *oil immersion lens* is used to increase the resolving power. An objective lens specially designed to be used in this way is termed as an *oil immersion objective*.

In this system, air is replaced by transparent oil (termed *immersion oil*) of high refractive index (very similar to refractive index of glass). Immersion oil such as paraffin oil, cedarwood oil has been placed at the interfaces between the objective lens and the cover slip protecting the specimen (also between the condenser lens and the underside of the specimen slide).

If the air is present between the cover slip and the objective lens, light is refracted, scattered and effectively lost. This happens because the refractive index of air is very different from that of glass and light passing through a glass-air interface is refracted (bent) to a large degree. By reducing the amount of refraction at this point, more of the light can be directed

Refractive index

The refractive index of a medium measures the extent of interaction between electromagnetic radiation and the medium through which it passes. The refractive index of water at room temperature is 1.33, which means that light travels 1.33 times slower in water than it does in vacuum.

to the narrow diameter lens of the high-power objective. The more the light, the clearer will be the image. Placing a material with a refractive index equal to that of glass in the airspace between cover slip and objective, more light can be directed through the objective which improves resolution. Immersion oils improve resolution by performing same function.

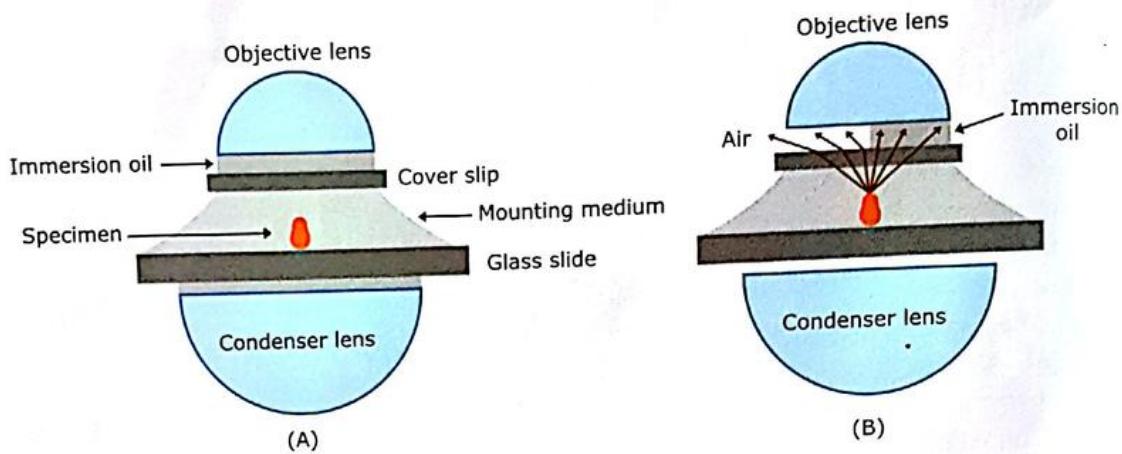


Figure 6.5 (A) An oil immersion lens where immersion oil has been placed at the interfaces between the objective front lens and the specimen covered with a cover slip and also between the condenser lens and the underside of the specimen slide. (B) In case of a dry objective, light rays pass through a specimen that is covered with a cover slip. These rays are refracted at the coverslip-air interface and only the light rays closest to the optical axis of the microscope have the appropriate angle to enter the objective lens. When air is replaced by oil of the same refractive index as glass, the light rays now pass straight through the glass-oil interface without any deviation due to refraction. In this case, the numerical aperture is, thus, increased by the factor of n , the refractive index of oil.

6.2 Types of light microscope

Brightfield microscopy

It is the original and most commonly used form of microscopy in which the specimen is viewed by transmitted light from a condenser lens. Light is aimed towards a condenser, through the specimen, through an objective lens, and to the eye through a second magnifying lens, the ocular or eyepiece. Specimens are visible in the light path because the natural pigmentation or stains absorb light differentially, or because they are thick enough to absorb a significant amount of light despite being colorless.

Darkfield microscopy

A dark-field microscope is a type of microscope in which objects are illuminated at a very low angle from the side so that the background appears dark and the objects show up against this dark background. It is a technique for improving the contrast of unstained, transparent specimens. Darkfield illumination uses a carefully aligned light source to minimize the quantity of directly-transmitted light and collecting only the light scattered by the sample. To view a specimen in a dark field, an opaque disc is placed underneath the condenser lens, that blocks light from entering the objective lens directly; light reflected by specimen enters the objective, and the specimen appears light against a black background.



Phase contrast micrograph of a human cheek cell.

Phase-contrast microscopy

When light passes through a living cell, the phase of the light wave is changed according to the cell's refractive index: light passing through a relatively thick or dense part of the cell, such as the nucleus, is retarded; its phase, consequently, is shifted relative to light that has passed through an adjacent thinner region of the cytoplasm. The phase-contrast microscope (invented by Frits Zernike) exploits the interference effects produced when these two sets of waves recombine, thereby creating an image of the cell's structure. The specimen appears as different degrees of brightness and contrast. It is used for the study of live and unstained cells, which are, in general, transparent to light.

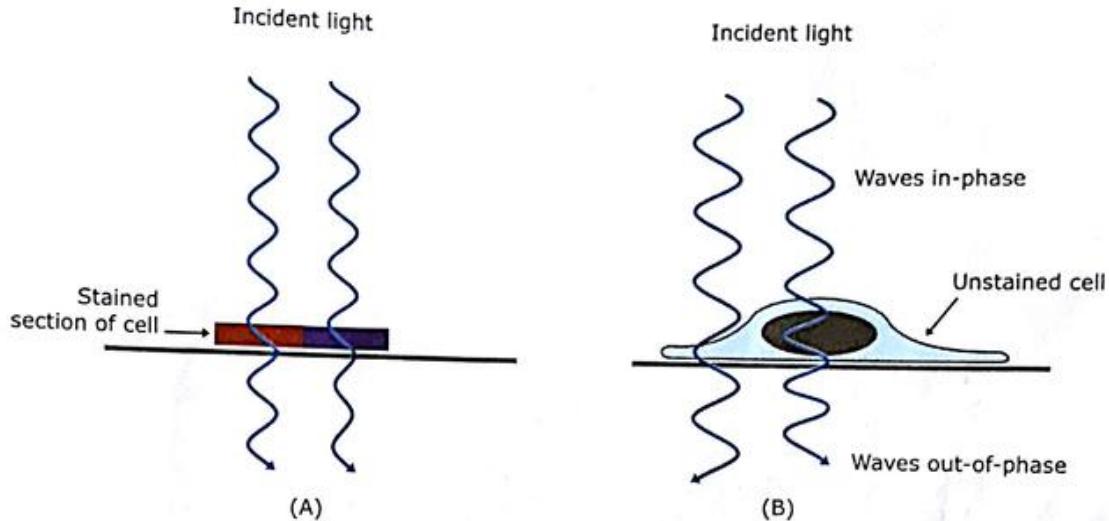


Figure 6.6 Two ways to obtain contrast in light microscopy. The stained portions of the cell in (A) reduce the amplitude of light waves of particular wavelengths passing through them. A colored image of the cell is, thereby, obtained that is visible in the ordinary way. Light passing through the unstained, living cell (B) undergoes very little change in amplitude, and the structural details cannot be seen even if the image is highly magnified. The phase of the light, however, is altered by its passage through the cell, and small phase differences can be made visible by exploiting interference effects using a phase-contrast or a differential-interference-contrast microscope.

Fluorescence microscopy

In fluorescence microscopy, the specimen itself acts as a light source. The specimens used to study are either fluorescent materials or stained with fluorescent dyes. A chemical is said to be fluorescent if it absorbs light at one wavelength and emits light (fluoresces) at a specific and longer wavelength. Most fluorescent dyes (or fluorochromes) emit visible light, but some emit infrared light. Fluorochromes exhibit distinct excitation and emission spectra that depend on their atomic structure and electron resonance properties. Two fluorescent dyes that are commonly used are fluorescein, which emits an intense green fluorescence when excited with blue light, and rhodamine, which emits a deep red fluorescence when excited with green-yellow light.

The fluorescence microscope is similar to an ordinary light microscope except that the illuminating light is passed through two sets of filters – one to filter the light before it reaches the specimen (*excitation filter*) and other to filter the light emitted from the specimen (*barrier or emission filter*). The excitation filter passes only the wavelength that excites the particular fluorescent dye, while the barrier filter blocks out this light and passes only those wavelengths emitted when the dye fluoresces. Only fluorescent light emitted by the fluorescently

stained specimen is used to form an image. The wavelength that excites the specimen and induces the fluorescence is not allowed to pass the filters placed between the objective lens and the eye.

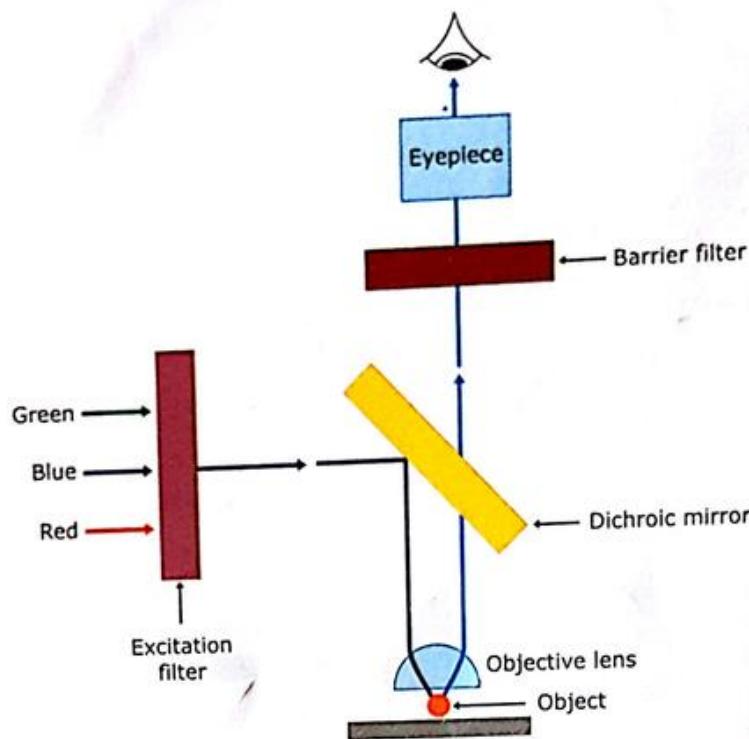


Figure 6.7 The optical system of a fluorescence microscope. Fluorescence microscopes contain special filters and employ a unique method of illumination to produce images of fluorescent light emitted from excited molecules in a specimen. It contains two essential filters – excitation filter and barrier or emission filter. The diagram shows the orientation of filters. The excitation beam (blue line) passes through the excitation filter and is reflected by the dichroic mirror and directed towards the specimen. The return beam of emitted fluorescence wavelengths (cyan line) passes through the dichroic mirror and the emission filter to the eye or camera. Excitation wavelengths that manage to pass through the dichroic mirror are blocked by the barrier (emission) filter.

Confocal microscopy

A confocal microscope creates sharp images of a specimen that would otherwise appear blurred when viewed with a conventional microscope. This is achieved by excluding most of the light from the specimen that is not from the microscope's focal plane. The image has less haze and better contrast than that of a conventional microscope and represents a thin cross-section of the specimen. Thus, apart from allowing better observation of fine details it is possible to build three-dimensional (3D) reconstructions of a volume of the specimen by assembling a series of thin slices taken along the vertical axis.

Confocal is defined as *having the same focus*. What this means in the microscope is that the final image has the same focus as or the focus corresponds to the point of focus in the object. The object and its image are *confocal*. The microscope is able to filter out the out-of-focus light from above and below the point of focus in the object. Normally when an object is imaged in the fluorescence microscope, the signal produced is from the full thickness of the specimen which does not allow most of it to be in focus to the observer. The confocal microscope eliminates this out-of-focus information by means of a confocal *pinhole* situated in front of the image plane which acts as a spatial filter and allows only the in-focus portion of the light

to be imaged. Light from above and below the plane of focus of the object is eliminated from the final image. The confocal microscope uses a laser beam to illuminate a specimen, usually one that has been fluorescently stained. A diagram of the confocal principle is shown below.

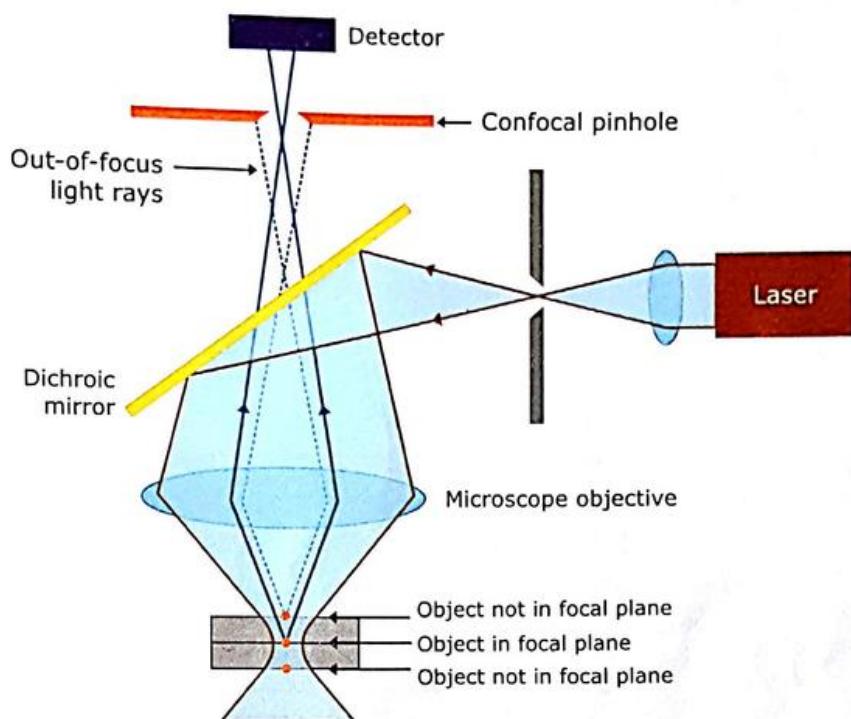


Figure 6.8 Ray path in a confocal microscope. A microscope objective is used to focus a laser beam onto the specimen, where it excites fluorescence. The fluorescent radiation is collected by the objective and efficiently directed onto the detector via a dichroic beam splitter. The wavelength range of the fluorescence spectrum is selected by an emission filter, which also acts as a barrier blocking the excitation laser line. The pinhole is arranged in front of the detector, on a plane conjugate to the focal plane of the objective. Light coming from planes above or below the focal plane is out-of-focus when it hits the pinhole, so most of it cannot pass the pinhole and, therefore, does not contribute to forming the image.

6.3 Electron microscope

The fundamental principles of electron microscopy are similar to those of light microscopy except one major difference of using electromagnetic lenses, rather than optical lenses to focus a high-velocity electron beam instead of visible light. The relationship between the limit of resolution and the wavelength of the illuminating radiation hold true for both – a beam of light or a beam of electrons. Due to the short wavelength of electrons, the resolving power of the electron microscope is very high. Electron microscopy is not used to study live cells. Electron microscopes are of two basic types: transmission electron microscope and scanning electron microscope. The most commonly used type of electron microscope is called the transmission electron microscope (TEM), because it forms an image from electrons that are transmitted through the specimen being examined. Scanning electron microscope (SEM) is fundamentally different from TEM, because it produces image from electrons deflected from a specimen's outer surface (rather than electrons transmitted through the specimen).

Electron microscopes consist largely of a tall, hollow cylindrical column through which the electron beam passes. The top of the column contains the cathode, a tungsten wire filament that is heated to provide a source of electrons. Electrons are accelerated as a fine beam by the high voltage applied between the cathode and anode. Just as a beam of light rays can be focused by a glass lens in light microscopes, a beam of negatively charged electrons can be focused by electromagnetic lenses. The strength of the magnets is controlled by the current provided them. Air is pumped out of the column, producing a vacuum through which the electrons travel. If the air were not removed, electrons would be prematurely scattered by collision with air molecules.

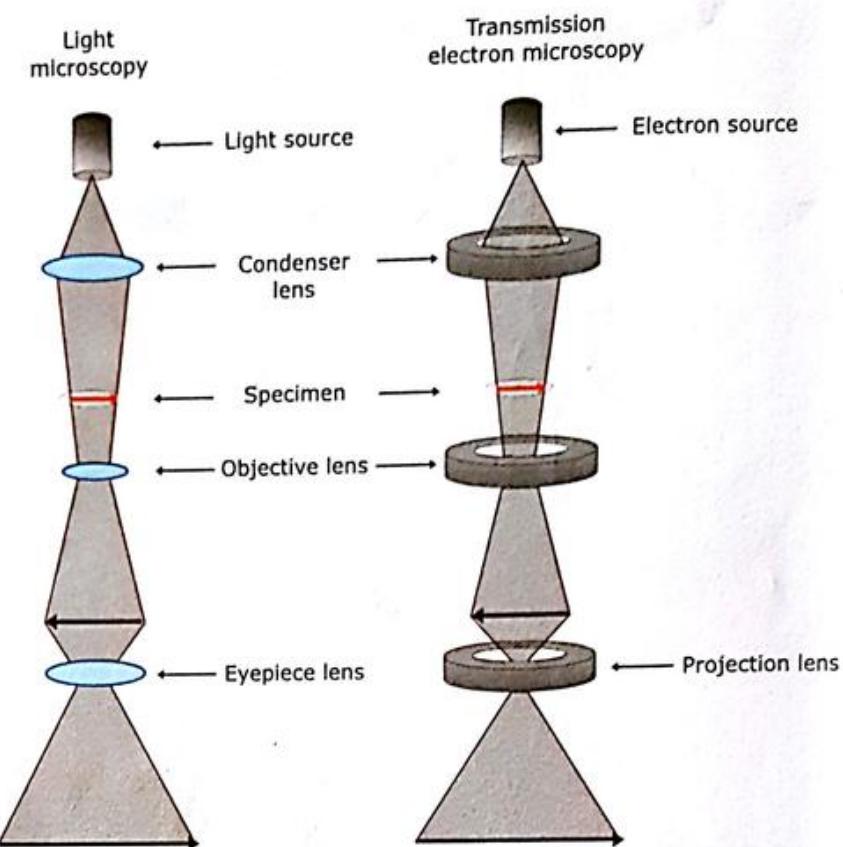


Figure 6.9 Light and electron microscopy.

Table 6.1 Differences between light and electron microscopes

Feature	Light microscope	Electron microscope (TEM)
Highest practical magnification	About 1,000-1,500	Over 100,000
Best resolution	0.2 μm	0.5 nm
Radiation source	Visible light	Electron beam
Medium of travel	Air	High vacuum
Type of lens	Glass	Electromagnet
Source of contrast	Differential light absorption	Scattering of electrons
Specimen mount	Glass slide	Metal grid (usually copper)

Transmission electron microscope

In light microscopy, *differential absorption* of light, which depends mainly on staining the specimen, results in the visible differences in various parts of the image. In the TEM, a condenser lens focuses the electron beam onto the specimen and electrons are transmitted through the specimen. The portion of the beam absorbed by specimen is minimal. To be absorbed, an electron must lose *all* its energy to the specimen. Although not absorbed, these electrons are scattered by the atoms of the specimen. Image formation in the electron microscope depends on differential scattering of electrons by parts of the specimen. Consider a beam of electrons focused on the screen. If no specimen were present in the column, the screen would be evenly illuminated by the beam of electrons, producing an image that is uniformly bright. By contrast, if a specimen is placed in the path of the beam, some of the electrons strike the specimen and are scattered away.

When the electrons come in contact with the sample, it can either be scattered *elastically*, that is, without any loss of energy, or inelastically, i.e. transferring some of that energy to the atom. Interaction between the incoming fast electron and an atomic nucleus gives rise to elastic scattering. Interaction between the fast electron and atomic electrons results in inelastic scattering.

Some of the electrons passing through the specimen are scattered and the remainder is focused to form an image, in a manner analogous to the way an image is formed in a light microscope. The image can be observed on a phosphorescent screen or recorded, either on a photographic plate or with a high-resolution digital camera. Because the scattered electrons are lost from the beam, the dense regions of the specimen show up in the image as areas of reduced electron flux, which look dark. Scattering of electrons contribute to the contrast. Because the electrons are absorbed by atoms in the air, the entire tube between the electron source and the detector is maintained under an ultrahigh vacuum.

The amount of scattering which occurs at any particular specimen point is dependent on its density and overall thickness, and is relatively independent of other specimen properties. Specimen prepared for TEM has fairly uniform thickness, therefore, almost no contrast arises from the thickness. Their atomic number also remains approximately constant, so the overall contrast is very low and the specimen appears featureless in the TEM. To produce scattering contrast, the sample is generally stained with *electron dense materials*. So, before or after slicing, the specimen is immersed in a solution that contains salts of heavy metal such as uranium and lead. The degree of impregnation or staining, with these salts reveals different cellular constituents with various degrees of contrast. Darker areas in the image are where few electrons have been transmitted through the sample due to thickness or staining.

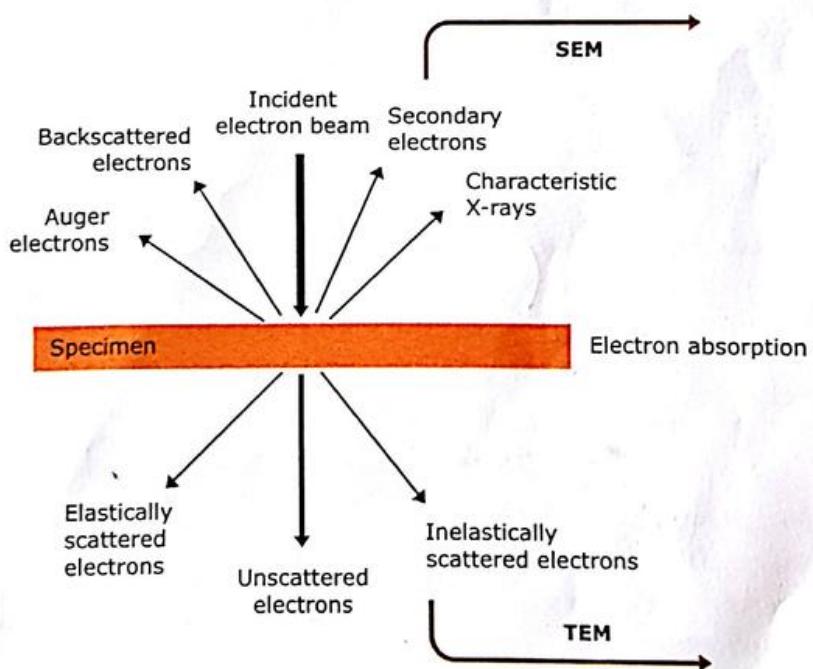


Figure 6.10 Interactions of the electron beam with the specimen.

Scanning electron microscope

The Scanning Electron Microscope (SEM) views the surfaces of specimens. The sample is fixed, dried, and coated with a thin layer of a heavy metal, such as gold or a mixture of gold and palladium. The specimen is, then, scanned with a very narrow beam of electrons. Molecules in the specimen are excited and they release secondary electrons that are captured

by a detector, and generating an image of the specimen's surface. Contrast arises when different parts of specimen generate differing amounts of secondary electrons as the electrons beam strikes them. Areas which generate large numbers of secondary electrons will appear brighter than areas that generate fewer secondary electrons. The resolving power of scanning electron microscopes, which is limited by the thickness of the metal coating, is only about 10 nm, much less than that of transmission electron microscope. The image produced appears three-dimensional in SEM whereas transmission electron microscope produces two-dimensional images.

6.4 Scanning tunneling electron microscopy

The Scanning Tunneling electron Microscopy (STM) utilizes a tiny probe that does not emit an electron beam, but instead possesses a tip made of a conducting material such as platinum-iridium. The tip of the probe is extremely sharp. It can move in three dimensions over a surface. The x and y dimensions scan the surface, while the z dimension governs the distance of the tip above the surface. When the tip and sample are connected to a voltage source, a small tunneling current flows between the tip and specimen surface (the specimen must be an electrical conductor). If the tip is close enough to the surface and the surface is electrically conductive, electrons will begin to leak or *tunnel* across the gap between the probe and the sample. This current can be measured, and the magnitude depends on the distance between the tip and the surface. The tunneling is highly dependent on the distance, so that even small irregularities in the size will affect the rate of electron tunneling. The instrument is designed to move the tip in the z direction to maintain a constant current flow. The movement is, therefore, a function of the tunneling current and is presented on a video screen.

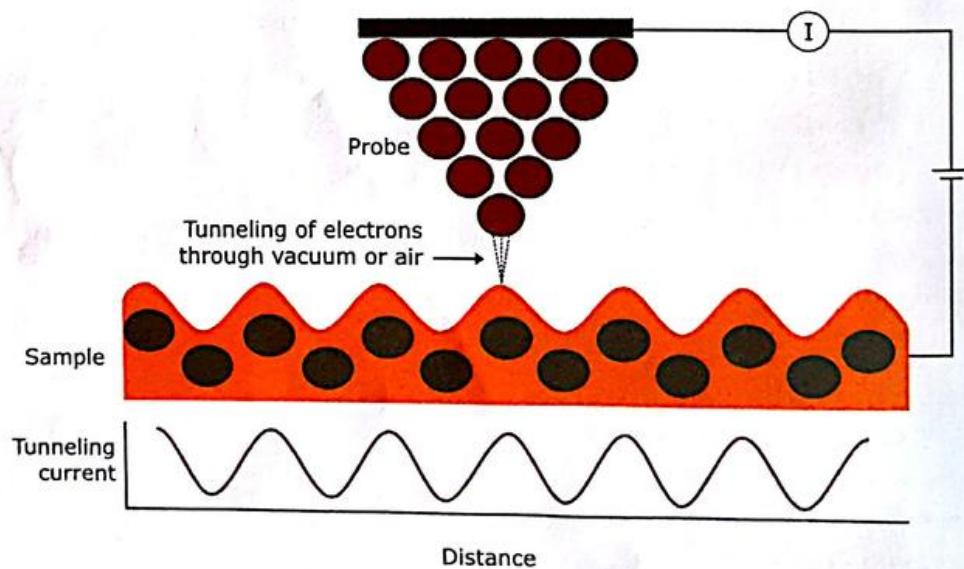


Figure 6.11 When the distance between conductive tip and conductive sample is lowered to a few Å, electrons can traverse the gap with some transmission probability. The STM measures not only distance like the AFM, but also the local density of electronic states.

6.5 Atomic-force microscopy

Atomic-Force Microscopy (AFM) is similar to scanning-tunneling microscopy in that it can image surfaces at atomic-scale resolution. The difference between AFM and STM is that AFM does not require the sample to be an electrically conducting material. Like STM, it uses an atomically-sharp tip that is brought very close to the surface. The tip will feel a chemical attraction or repulsion and will move up or down on its supporting cantilever.

Electrostatic and van der Waals interactions between the tip and the sample produce a force that moves the probe up and down (in the z dimension) as it encounters ups and downs in the sample. The key to the sensitivity of AFM is in monitoring the movement of the tip. A common means of monitoring the tip movement is to use a laser beam that is reflected or diffracted by the tip or cantilever. Up or down movement of the tip is, then, detected by changes in the laser beam position. As in STM, moving the tip across the surface produces a topographic map of the surface with atomic resolution.

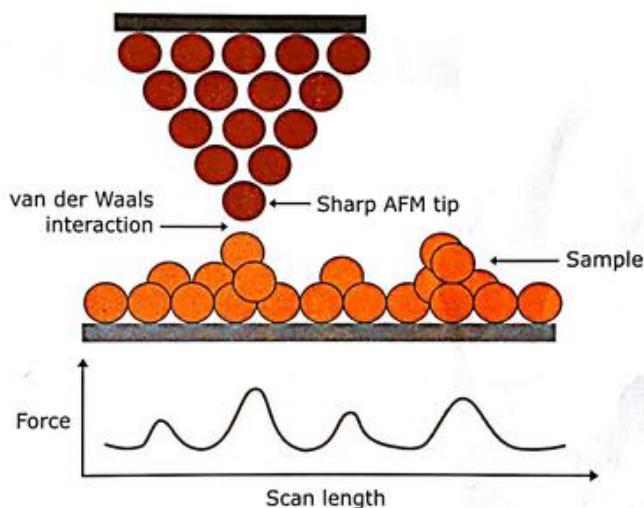


Figure 6.12 When a small tip approaches a surface, it experiences the van der Waals force which is attractive at a distance of a few Å, but repulsive at very short distances. Additional Coulomb forces may play a role when the AFM tip is charged.

6.6 Freeze-fracture electron microscopy

Freeze-fracture electron microscopy technique is used to visualize the features of cell membranes. Cells are frozen and immobilize at the temperature of liquid nitrogen (-196°C) in the presence of a *cryoprotectant* (antifreeze) to prevent distortion from ice crystal formation. The specimen is, then, fractured with a cold microtome knife, which often splits the bilayer into monolayer and exposes the interior of the lipid bilayer and its embedded proteins. The fracture plane is not observed directly rather a replica or cast is made of the fractured surface by shadowing it obliquely with heavy metal (such as platinum) deposition. A carbon layer is then deposited on the top of the metal layer to impart sufficient strength to the replica. The organic material of the cell is dissolved away after shadowing to leave only the thin metal *replica* of the surface of the specimen. Finally, the metal-shadowed replicas are viewed in the electron microscope.

Atomic force microscopy (also scanning force microscopy), does not use lenses to form an image, but instead uses a sharply pointed sensor, or tip, at the end of a flexible cantilever to scan and sense the topography of a sample.

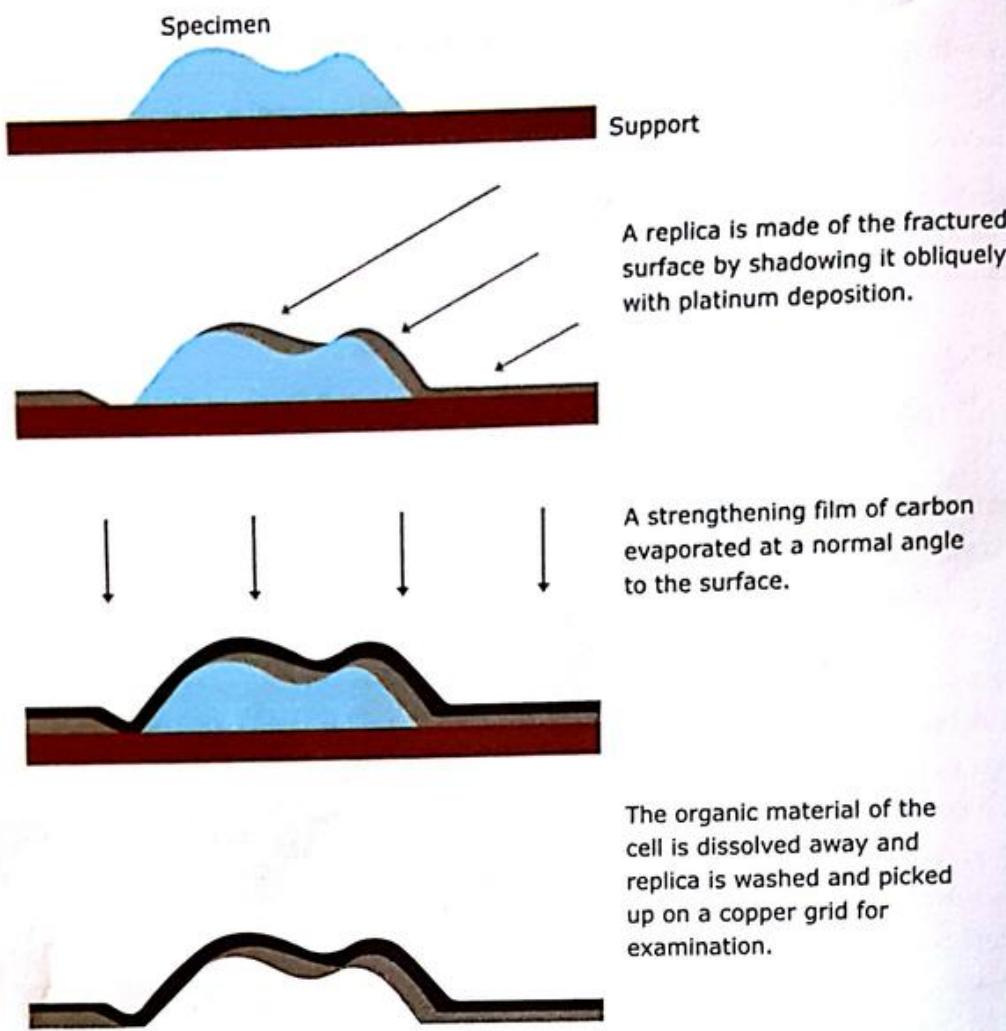


Figure 6.13 Replication of the fractured surface is accomplished by shadow-casting the surface with platinum at an oblique angle (45°). This produces a film 2–5 nm thick on the surface normal to the angle of the shadow. Following shadowing, a strengthening film of carbon evaporates at a normal (90°) angle to the surface. This imparts sufficient strength to the replica so that when it is removed from the surface it will not break. The thickness of the film is directly related to the slope of the shadowed surface. The topology of the surface is converted into variations in the thickness of the shadow and finally into variations in electron density recorded on photographic film.

The freeze fracture method involves the following steps:

- Cryofixation of the specimen by rapid freezing,
- Fracturing of the specimen by cleaving it with a microtome knife,
- Coating of the freeze-fracture plane by heavy metals under high vacuum,
- Deposition of a carbon layer to strengthen the replica,
- Removal of the specimen from the replica and
- Observation of the replicas with a transmission electron microscope.

Freeze etching is a slight modification of freeze-fracture technique. The specimen is rapidly frozen, and the block of ice is fractured with a knife. The frozen fractured specimen is exposed to a vacuum at an elevated temperature for one to a few minutes, during which a layer of ice can evaporate (sublime) from the exposed surface. Once some of the ice has been removed, the surface of the structure can be coated with heavy metal and carbon to create metallic replica. Etching exposes the true surfaces of freeze-fractured membranes, thereby revealing membrane surface features of interest that were obscured previously by the overlying ice.

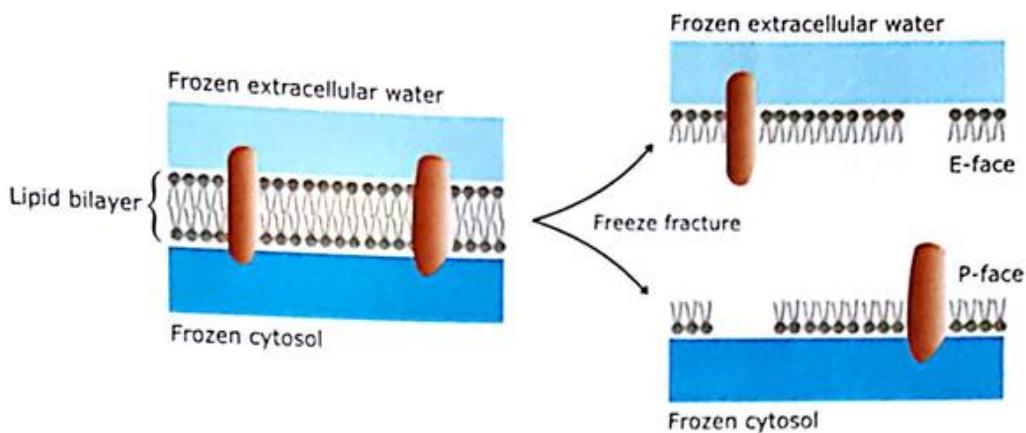


Figure 6.14 When the lipid bilayer splits, both the hydrophobic interior of the cytosolic half of the bilayer (the P-face) and the hydrophobic interior of the external half of the bilayer (the E-face) become exposed. After the fracturing process, the exposed fracture faces are shadowed with platinum and carbon, the organic material is digested away and the resulting platinum replica is examined in an electron microscope.

6.7 Fixation and staining

Fixation

It is the first step in sample preparation and has the aim of preserving structure of living tissue in its original state. Fixation makes cells permeable to staining reagents and cross-links their macromolecules so that they are stabilized and locked in position. Specimens for light and electron microscopy are commonly fixed with a solution containing chemicals that crosslink most proteins and nucleic acids. Fixatives kill the cells while preserve their structural appearance.

The most widely employed fixatives are aldehydes such as formaldehyde and glutaraldehyde. Formaldehyde's mechanism of action is based on the reaction of the aldehyde group with primary amines in proteins.

Glutaraldehyde is the most commonly used primary fixative. It contains two aldehyde groups separated by three methylene bridges. These two aldehyde groups and the flexible methylene bridge greatly increases the cross-linking potential of glutaraldehyde over formaldehyde. It penetrates rapidly and stabilizes proteins by forming cross links, but does not fix lipids. Osmium tetroxide is used as a secondary fixative, reacting with lipids and acting as a stain. It is especially useful for fixing cell membranes since it reacts with the C=C double bonds present in many fatty acids. Following each fixation step, excess fixative must be washed out of the tissue.

Staining

Most biological materials show little contrast with their surroundings unless they are stained. In the case of light microscopy, contrast can be enhanced by using colored stains which selectively absorb certain wavelengths. Light that illuminates the specimen usually consists of a spectrum of colors. These colors become readily visible when the light strikes a specimen that has been stained with various dyes (stains). These dyes bind to specific molecules present in the specimens. For example, hematoxylin binds to basic amino acids (lysine and arginine) of different proteins, whereas eosin binds to acidic molecules such as DNA and side chains of aspartate and glutamate. These dyes absorb certain colors in the spectrum, but transmit or

Light microscopy yields photomicrographs in color as well as in black and white, whereas electron microscopy produces electron micrographs in black and white only.

reflect others to the eye. For example, in a specimen that has been stained with two commonly used biological stains, safranin and fast green, certain parts of the cell will appear red (safranin will stain acidic components of the cell such as DNA in the nucleus), while others will stain green (fast green stains basic components of the cytoplasm).

Contrast in the electron microscope depends on the atomic number of the atoms in the specimen: the higher the atomic number, the more electrons are scattered and the greater the contrast. Biological tissues are composed of atoms of very low atomic number (mainly carbon, oxygen, nitrogen, and hydrogen). Hence, heavy metal salts are used as stain in electron microscope. Colors do not exist in any electron microscope.

An electron stain is said to exhibit *positive contrast* when it increases the density of a particular biological structure as opposed to the adjoining areas. In positive staining, the heavy metal salts attached to various organelles or macromolecules within the specimen increase their density and, thereby, increase contrast differentially. The two most commonly used positive stains are uranyl acetate and lead citrate. The exact mode of action of these stains is not completely understood. It is known that uranyl ions react strongly with phosphate and amino groups so that nucleic acids and certain proteins are highly stained. With lead stains, it is thought that lead ions bind to negatively charged components. Positive staining differs from the *negative stain* situation, where the background area surrounding the specimen is made dense by a heavy metal salt so that the specimen appears lighter in contrast to the darkly stained background. If specimens are coated with stain (positive staining), fine detail may be obscured. Negative staining overcomes this problem by staining the background and leaving the specimens relatively untouched. The negative stain is moulded round the specimen. In the electron micrograph, the sample appears as a light region surrounded by a dark background originating from the stains. Negative stains are not used in sectioned materials, but are used to contrast whole, intact biological structures (viruses, bacteria, cellular organelles, etc.). Negative stains are salts of heavy metals such as uranium, tungsten and molybdenum. Uranyl acetate and phosphotungstic acid are the two most commonly used negative stains.

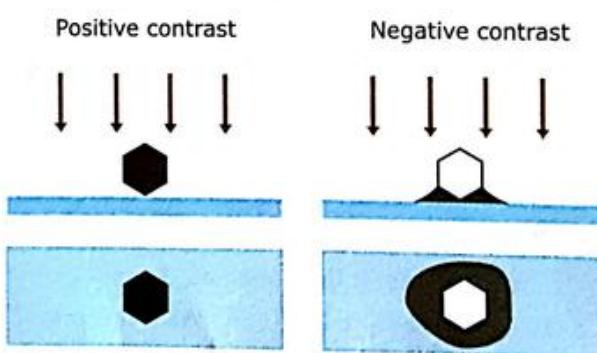


Figure 6.15 Comparison of virus appearance after positive and negative staining. To heighten the contrast between viruses and the background, electron-dense stains are used. These are usually compounds of heavy metals of high atomic number, that serve to scatter the electrons from regions covered with the stain. If virus particles are coated with stain (positive staining), fine detail may be obscured. Negative staining overcomes this problem by staining the background and leaving the virus relatively untouched.

Chapter 7

Flow cytometry

Flow cytometry is a technique for counting, examining, and sorting microscopic objects suspended in a fluid based on their optical properties (scattering and fluorescence). It simultaneously measures and then analyzes multiple physical characteristics of single object (usually cell) as they flow in a fluid stream through a beam of light.

A flow cytometer is made up of three main systems: fluidics, optics, and electronics. The *fluidics system* transports cells in a stream to the laser beam for interrogation. The *optics system* consists of lasers to illuminate the cells in the sample stream and optical filters to direct the resulting light signals to the appropriate detectors. The *electronics system* converts the detected light signals into electronic signals that can be processed by the computer.

Fluidics system

Flow cytometer measures the physical and chemical characteristics of a cell. This can be done by introducing cells present in suspension to go through an observation point one at a time. At this observation point, using a specific light beam, we interrogate each cell and observe the light responses electronically. Flow cytometers use the principle of hydrodynamic focusing for presenting cells one at a time to a light source.

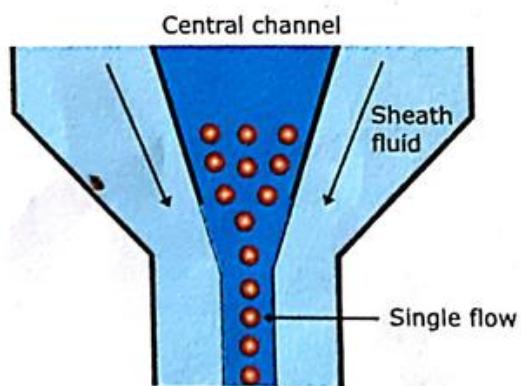


Figure 7.1 Hydrodynamic focusing produces a single stream of particles.

The fluidics system consists of a central channel through which the sample is injected, enclosed by an outer sheath that contains faster flowing fluid. As the sheath fluid moves, it creates a massive drag effect on the narrowing central chamber. This alters the velocity of the central fluid whose flow front becomes parabolic with greatest velocity at its center and

zero velocity at the wall. The effect creates a single stream of particles or cells and is called *hydrodynamic focusing*. Under laminar flow conditions, the fluid in the central chamber will not mix with the sheath fluid.

Optics

After hydrodynamic focusing, each cell passes through a beam of light. Light scattering or fluorescence emission (if the particle is labeled with a fluorochrome) provides information about the cell's properties. Scattered and emitted light from cells are converted to electrical pulses by optical detectors.

Flow cytometers routinely measure light scattered at two different angles (forward and side). Forward scatter is measured in the same direction as the laser beam is moving. Side scatter is measured at the 90° relative to the primary laser beam.

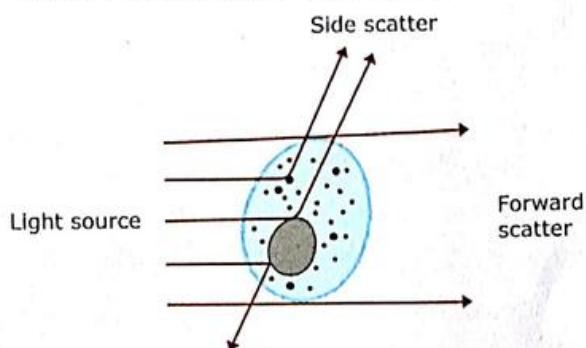


Figure 7.2 Diagram of forward and side scatter. The amount of light scattered in a forward direction is an indicator of the relative size of a cell. This is referred to as *forward scatter*. Likewise, light scattered at higher angles gives rise to information about internal complexity and granularity of a cell, such as cytoplasmic granules, multi-lobed nucleus, cytoplasmic vacuoles, etc. This is referred to as *side scatter*.

Light that is scattered by an object is detected by different detectors. One detector is placed in line with the light beam to measure the *forward scatter* (FSC) from the objects. Detectors perpendicular to the beam measure *side scatter* (SSC) and fluorescence. Both FSC and SSC are unique for every object, and a combination of the two may be used to differentiate different objects in a heterogeneous sample. Forward scatter is based on two properties - size and refractive index.

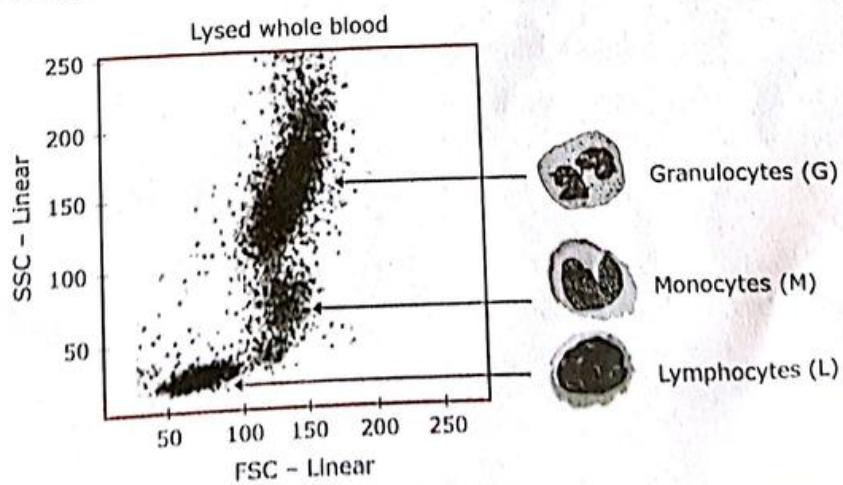


Figure 7.3 The different physical properties of granulocytes, monocytes and lymphocytes allow them to be distinguished from each other and from cellular contaminants. Forward and side scatter can be used to distinguish different cell types in human blood. In this plot, three blood cell populations can be identified based on their light scattering. Each dot represents a cell, and its location in the graph corresponds to the amount of light scattered by the cell.

A fluorescent compound absorbs light energy over a range of wavelengths that is characteristic for that compound. This absorption of light causes an electron in the fluorescent compound to be raised to a higher energy level. The excited electron quickly decays to its ground state, emitting the excess energy as a photon of light. This transition of energy is called fluorescence.

The FSC intensity roughly equates to the particle's size and can also be used to distinguish between cellular debris and living cells. Dead cells have lower FSC and higher SSC than living cells. Side scatter is based on granularity or internal complexity. The more granular the object (cell), the more side scatter light is generated.

Fluorescence measurements taken at different wavelengths can provide quantitative and qualitative data about fluorochrome-labeled cell surface receptors or intracellular molecules such as DNA and cytokines. When a fluorescent dye is conjugated to a monoclonal antibody, it can be used to identify a particular cell type based on the individual antigenic surface markers of the cell. In a mixed population of cells, different fluorochromes can be used to distinguish separate subpopulations. The staining pattern of each subpopulation, combined with FSC and SSC data, can be used to identify which cells are present in a sample and to count their relative percentages. The cells can also be sorted if desired.

Optical detector

Once a cell or particle passes through the laser light, the scattered and fluorescence signals are diverted to the *detectors*. The number of detectors will vary according to the machine. Detectors are either silicon photodiodes or photomultiplier tubes (PMTs). The photodiode is less sensitive to light signals than the PMTs and thus is used to detect the stronger FSC signal. PMTs are used to detect the weaker signals generated by SSC and fluorescence.

Detectors: Scattered light detector - FSC and SSC

Fluorescence light detector - FL-1, FL-2, FL-3, FL-4

All of the signals are routed to their detectors via a system of filters and dichroic mirrors. Each PMT fluorescence detector is placed behind a series of dichroic mirrors and filters, so that it only receives and detects light within a particular range of wavelength. A particular color of light is split off from the incoming mixture and directed to the detectors by dichroic mirrors. Optical filters block certain wavelengths while transmitting others. There are three major filter types:

- Longpass filter: Transmits wavelengths of light equal to or greater than the spectral band of the filter.
- Shortpass filter: Transmits wavelengths of light equal to or shorter than the spectral band of the filter.
- Bandpass filter: Transmits wavelengths of light within a specific range of wavelengths.

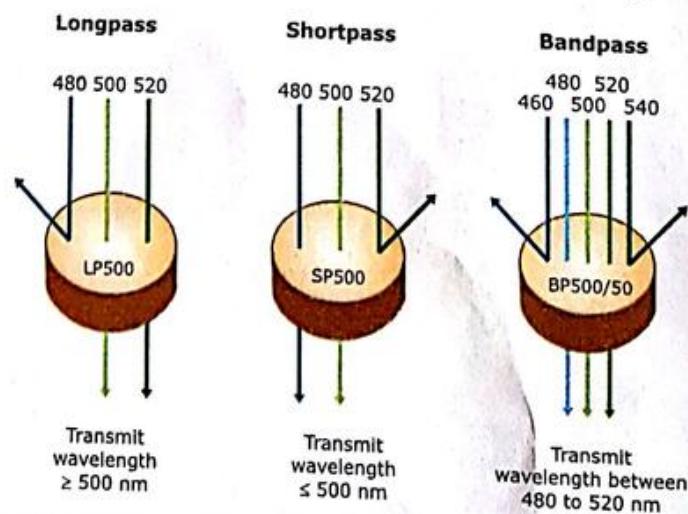


Figure 7.4 Different types of optical filters. All these filters block light by absorption.

Flow cytometers count and evaluate every cell as it passes through the laser beam and record the level of emitted fluorescence at a number of different wavelengths, as well as the amounts of forward- and side scattered light for each cell; an attached computer stores all the data for each cell.

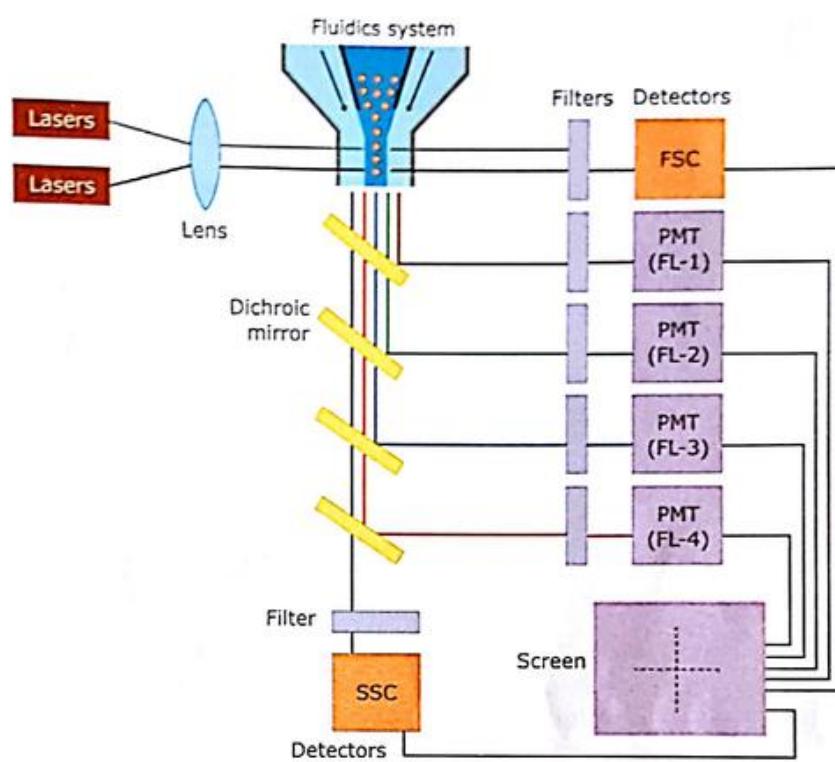


Figure 7.5 Optical system of a typical flow cytometer: The optical system consists of excitation optics and collection optics. The excitation optics consist of the laser and lenses that are used to shape and focus the laser beam. The collection optics consist of a collection lens to collect light emitted from the particle-laser beam interaction and a system of optical mirrors and filters to route specified wavelengths of the collected light to designated optical detectors.

Cell sorting

Applications of flow cytometry
Cell sorting,
Cell's DNA quantification,
Analysis of cell size,
Cell cycle analysis,
Apoptosis,
Cell surface markers analysis,
Immunophenotyping
It is a term applied to the technique of identifying the specific lineage of cells through the use of antibodies that detect antigens or markers on the cell.

A major application of flow cytometry is the physical separation of sub-populations of cells of interest from a heterogeneous population. This process is called **cell sorting** or **Fluorescence Activated Cell Sorting (FACS)**. Most commonly used cell sorting method is the electrostatic deflection of droplets. In this case the stream is focused in a vibrating nozzle and exits in a jet which is broken into regularly spaced droplets. The droplets containing a cell of interest are charged electrically (positively or negatively). Electrostatic charging actually occurs at a precise moment called the 'break-off point', which describes the instant the droplet containing the particle of interest separates from the stream. To prevent the break-off point happening at random distances from the nozzle and to maintain consistent droplet sizes, the nozzle is vibrated at high frequency.

When a charged droplet passes through a high voltage electrostatic field, between the deflection plates, it is deflected and collected into the corresponding collection tube. The deflection of the droplet is towards the oppositely charged plate, so that this droplet is separated from uncharged and oppositely charged droplets.

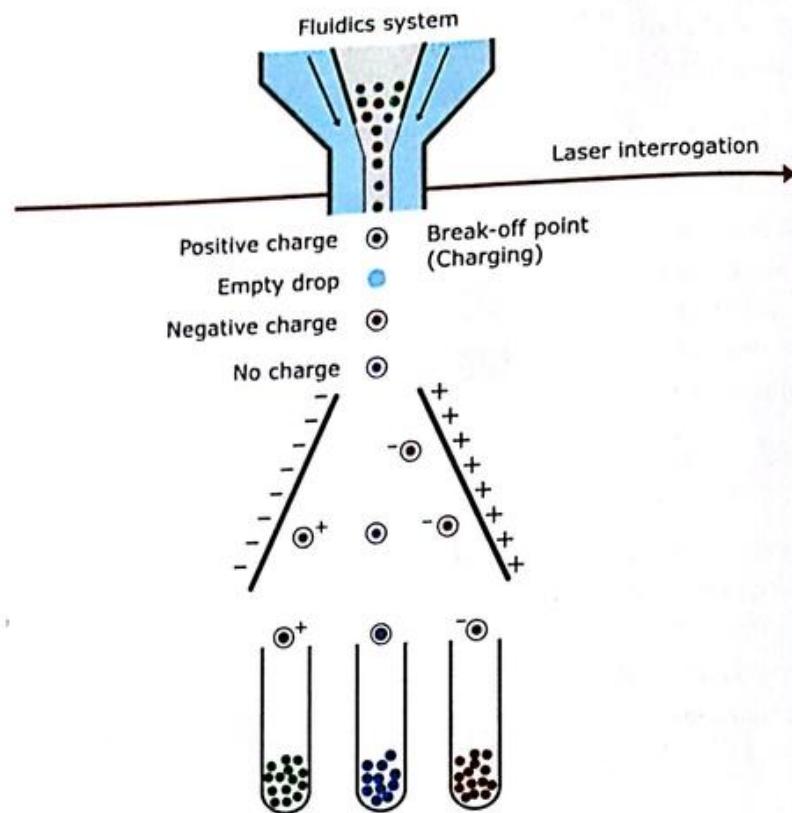
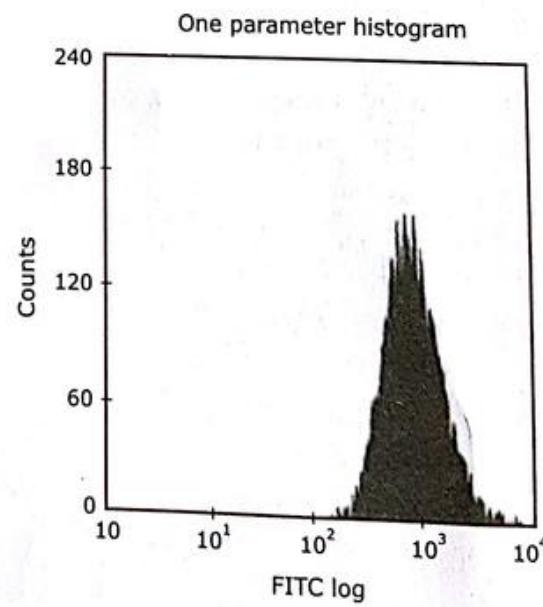


Figure 7.6 Fluorescence-activated cell sorter (FACS) separates cells that are labeled differentially with a fluorescent reagent. A mixture of fluorescently stained cells is passed through a small aperture so that each drop that emerges contains just one cell. The fluorescence detector identifies the signal from drops containing the correct cell and applies an electric charge to these drops. When the drops reach the electric plates, the charged ones are deflected into a separate beaker.

Figure 7.7 The histogram shows the total number of cells in a sample that possess certain physical properties selected for or which express the marker of interest. Cells with the desired characteristics are known as the positive dataset. Ideally, flow cytometry will produce a single distinct peak that can be interpreted as the positive dataset. However, in many situations, flow analysis is performed on a mixed population of cells resulting in several peaks on the histogram. In order to identify the positive dataset, flow cytometry should be repeated in the presence of an appropriate negative isotype control.

Signal processing

Flow cytometry data may be represented as *histograms* or as *dot plots*. A histogram quantifies the intensity of a single parameter, be it fluorescence or scattering (SSC or FSC). Subpopulations are identified as peaks in the histogram. In this plot, the 'X' axis represents scatter or fluorescence intensity and the 'Y' axis represents the number of events. The measurement from each detector is referred to as a 'parameter' e.g. forward scatter, side scatter or fluorescence.



A dot plot is a two parameter histograms representation of a sample's properties. It plots one parameter against another and represents each cell by a dot with a position on the two axes. Three common modes of use for the dot plot are:

1. Forward scatter (FSC) versus side scatter (SSC)
2. Single color versus side scatter
3. Two-color fluorescence plot

In the dot plot of FSC versus SSC, one is able to look at the distribution of cells based upon size and granularity.

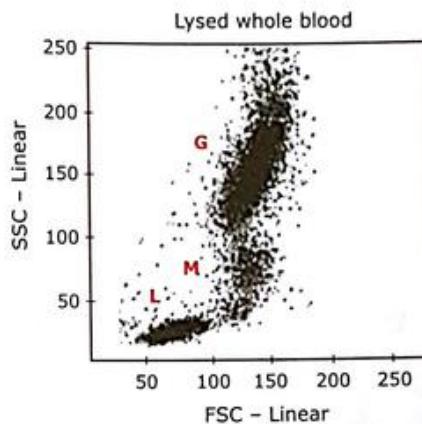


Figure 7.8 A dot plot of flow cytometry data. Lysed whole blood analysis using scattering (SSC versus FSC). Each dot or point represents an individual cell that has passed through the instrument. The clusters labeled G, M and L arise from granulocytes, monocytes and lymphocytes, respectively.

The most common use of dot plots is to analyze multiple colors of fluorescence with respect to each other. This makes it possible to differentiate between those cells that express only one of the particular fluorescent markers, those that express neither, and those that express both. This method is frequently used to discriminate dead cells from the live ones that are expressing the desired fluorescence.

For example, if a researcher was analyzing a sample of peripheral blood that had been stained with fluorochromes to identify the CD4 and CD14 surface markers but were only interested in knowing the percentage of monocytes that contained those markers, they might place a gate around the monocyte population of the FSC versus SSC scatter plot thereby limiting the data they visualize to monocytes as compared to the entire population of cells. Next, using a dot plot to graph the gated region they would be able to gain a clearer understanding of what percentage of monocytes express only CD4, only CD14, neither or both. In this case two different markers are paired and assessed together. Mutually exclusive marking is where a population only expresses one marker and not the other. Co-expression is where a population expresses both markers. Non-expression is when the population of interest does not express either population.

Gating
An important principle of flow cytometry data analysis is called gating. It is a procedure to selectively visualize the cells of interest while eliminating results from unwanted cells and debris. For example, if one has a heterogeneous population of cells which contain lymphocytes, monocytes and granulocytes and is only interested in evaluating the fluorescence of the lymphocyte subpopulation. In this situation one could define an analysis gate around the lymphocyte population. The resulting display would reflect the fluorescence properties of only lymphocytes.

Regions are shapes that are drawn around a population of interest on a one- or two-parameter plot. When a region is used to limit the cells that are drawn on a plot, it is termed a gate.

Whether to use a linear or logarithmic scale on an axis depends mainly on the range of intensity of the fluorescence signal. If the range is narrow then use of a linear scale would be best. Otherwise subtle difference could become obscured if a log scale is used. Conversely, in cases where fluorescence intensities can vary 100-fold, using linear scaling might make two distinct populations hard to visualize as the plot would look too compressed against the axis.

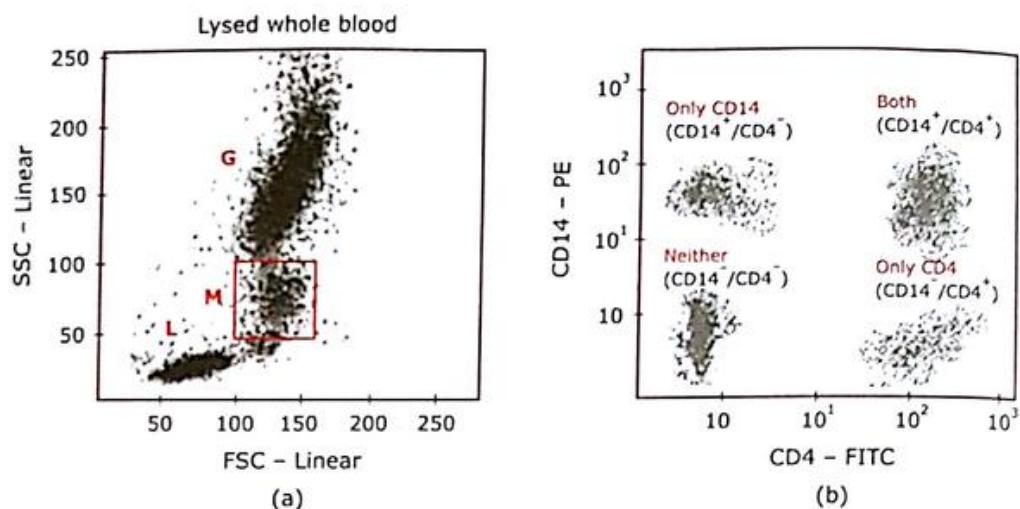


Figure 7.9 (a) Light scatter plot – the clusters labeled G, M and L arise from granulocytes, monocytes and lymphocytes respectively. A gate is applied to the monocyte population of a peripheral blood sample. (b) Fluorescence data from the gated region of monocytes clarifies which cells contain surface markers (CD14 and CD4). The cells are stained with fluorescently labeled antibodies directed against specific cell-surface markers, a FITC-conjugated antibody against CD4 and PE conjugated antibody against CD14. FITC=Fluorescein isothiocyanate dye and PE=Phycoerythrin dye.

Analysis of apoptosis and cell cycle

Apoptosis is a finely tuned mechanism for the control of cell number in eukaryotes. The process is operative during embryogenesis, in tumor regression and in the control of immune response. In most cases, it consists of an ordered sequence of cellular events that start with the transcription of specific genomic sequences, synthesis of specific proteins, and activation of classes of nucleases that cut DNA. These events are paralleled by specific morphological changes in both the cell nucleus and cytoplasm. Cells undergoing apoptosis display typical changes in their morphological, chemical and physical properties, which are well measurable by flow cytometry. The following features of the apoptotic cascade can be observed using flow cytometry:

- Altered phospholipid composition in the plasma membrane
- Activation of caspases
- Cell shrinkage
- Chromatin condensation
- DNA fragmentation
- Expression of proteins involved in apoptosis
- Changes in the mitochondrial membrane potential
- Decrease cytosolic PH
- Altered membrane permeability

Detection of apoptotic cells based on changes in forward scattering

Interaction of a cell with the laser beam produces FSC that correlates with cell size and SSC that correlates with granularity and/or cell density. While necrotic death is characterized by a reduction in both FSC and SSC (probably due to a rupture of plasma membrane and leakage of the cell's content), during apoptosis there is an initial increase in SSC (probably due to the

chromatin condensation) with a reduction in FSC (due to the cell shrinkage). In many cases, however, the forward light scattering histograms of apoptotic and live cells overlap and make it difficult to discriminate apoptotic cells based solely on this parameter. Thus, scatter changes alone are not specific to apoptosis and should be accompanied by an additional characteristic associated with cell death.

Detection of apoptotic cells based on Annexin V binding

Live cells have phospholipids asymmetrically distributed in the inner and outer leaflet of plasma membrane, with most of phosphatidylserine on the inner leaflet. Early during apoptosis cells lose asymmetry, exposing phosphatidylserine on the outer leaflet of plasma membrane. Annexin V is a calcium-dependent phospholipid-binding protein that binds preferentially to negatively-charged phosphatidylserine. Fluorophore (or fluorochrome) conjugated Annexin V is a useful reagent for the detection of apoptotic cells. The assay involves incubating cells briefly in a solution containing FITC-conjugated Annexin V in a buffer that facilitates its binding. The concentration of CaCl_2 in the buffer can be varied to obtain optimal binding for individual cell lines. Apoptotic cells can be detected on the basis of increased binding of FITC-conjugated Annexin V.

The Annexin V binding assay has many of the advantages of other flow cytometric assays in that it is simple, rapid, sensitive, and objective. Unfortunately, it is not specific only for apoptosis in that whenever cell membrane integrity is disrupted (even nonionic detergents), cells may stain with Annexin V, probably because sites on the inner membrane become accessible. Thus, even though loss of membrane asymmetry is an early event characteristic of apoptosis, Annexin V alone cannot distinguish cells dying by apoptosis from those dying by necrosis.

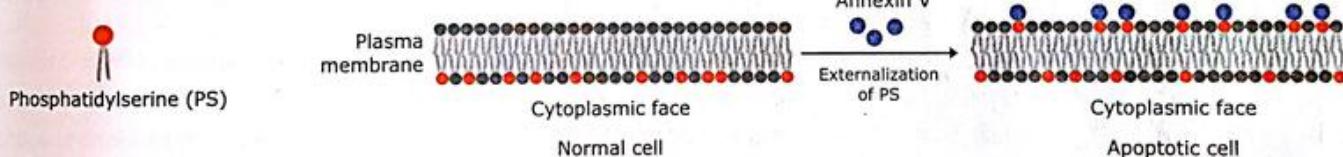


Figure 7.10 Changes in the plasma membrane are one of the first characteristics of the apoptotic process detected in living cells. Apoptosis can be detected by the presence of phosphatidylserine (PS), which is normally located on the cytoplasmic face of the plasma membrane. During apoptosis PS translocates to the outer leaflet of the plasma membrane and can be detected by flow cytometry through binding to fluorochrome-labeled Annexin V when calcium is present.

Detection of apoptotic cells based on PI binding

Another characteristic of plasma membranes associated with live cells is that they exclude charged cationic dyes such as propidium iodide (PI) and 7-amino-actinomycin D (7-AMD). PI staining is a good method to distinguish apoptotic, necrotic and normal live cells. The intact membrane of living cells excludes cationic dyes such as PI. In contrast, due to their extensive membrane damage, necrotic cells are quickly stained by short incubation with PI. Apoptotic cells show an uptake of PI that is much lower than that of necrotic cells. It is therefore possible to distinguish live (PI-negative), apoptotic (PI-dim) and necrotic (PI-bright) cells from each other. However, apparent PI staining only appears in the late phase of apoptosis due to more damaged cell membranes. Thus, the combined use of cationic dyes (e.g. PI) with annexin V allows the discrimination between live cells (annexin V negative/PI negative), early apoptotic cells (annexin V positive/PI negative), late apoptotic cells (annexin V positive/PI positive) and necrotic cells (annexin V negative/PI positive).

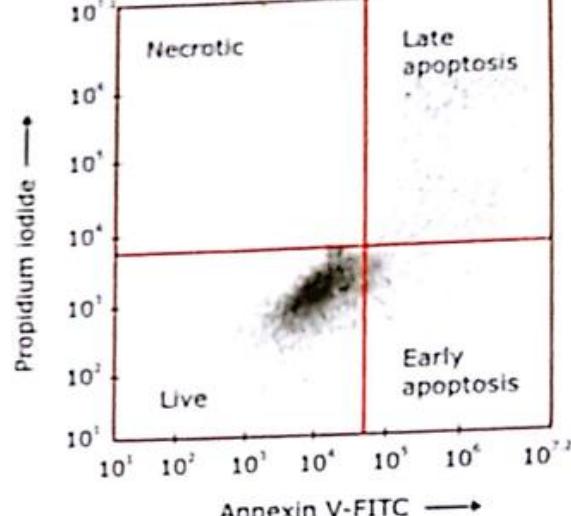


Figure 7.11 Measurement of apoptosis by Annexin V/PI dual stain. Early apoptosis is defined as cells positive for Annexin V only. Late apoptosis is defined as cells positive for Annexin V and PI. Necrotic is defined as cells positive for PI only.

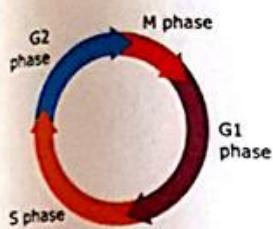
Detection of apoptotic cells based on DNA fragmentation

The late stages of apoptosis are characterized by changes in nuclear morphology, including DNA fragmentation, chromatin condensation, degradation of nuclear envelope, nuclear blebbing and DNA strand breaks. Hence, using flow cytometry a variety of nuclear changes during apoptosis can be analyzed. Cells undergoing apoptosis display an increase in nuclear chromatin condensation. As the chromatin condenses, cell-permeable nucleic acid stains become hyperfluorescent, thus enabling the identification of apoptotic cells when combined with a traditional dead-cell stain.

The late stage of apoptosis is marked with the activation of endonucleases, which translocate into the nucleus and cause DNA fragmentation and internucleosomal cleavage. These nucleases degrade the higher order chromatin structures into fragments of 50–300 kbp and subsequently into smaller DNA pieces of about 200 bp in length, forming a characteristic 'ladder' of DNA fragments. In order to quantify the percentage of cells in the culture that displays such fragmentation pattern, DNA breaks can be labeled at the 3'-hydroxyl termini with labeled nucleotides in a reaction catalyzed by terminal deoxynucleotidyl transferase. This is called Terminal dUTP Nick End Labeling (TUNEL) assay.

The TUNEL assay is based on the following principles: the oligonucleosomal DNA fragments in apoptotic cells contain a 3'-hydroxyl group (and 5'-phosphate) that arise from the cleavage of the phosphodiester bond of the DNA helix. These breaks can be detected on the basis of the ability of the terminal deoxynucleotidyl transferase enzyme to add nucleotides to free 3'-hydroxyl groups in fragmented genomic DNA. Thus, in the TUNEL assay, cells are incubated with this enzyme along with biotinylated nucleoside triphosphate substrates. Incorporation of labeled avidin (note that alternative staining protocols may be used). The number of apoptotic cells is estimated by quantifying the number of cells that exhibit increased fluorescence compared to a live control cell population.

Assessment of mitochondrial transmembrane potential and caspases level within the cell through flow cytometry is also used to analyze apoptotic cells. Cells undergoing apoptosis often lose the electrical potential that normally exists across the inner membrane of their



Cell cycle is the period between successive divisions of a cell. Standard cell cycle of eukaryotic cells is divided into four phases:

G1 phase : Cells prepare for DNA replication and the DNA content is 2C.

S-phase : Replication of DNA occurs at this phase. During this phase, there is a DNA content between G1 and G2 phase. The DNA content is doubled by the end of this phase (2C to 4C).

G2 phase : Replication of the DNA completes in this phase. The cells contain twice the amount of DNA (i.e. 4C) that are found in G1 (i.e. 2C).

M-phase : It is the period of actual division, corresponding to the visible mitosis. At this phase, the cells contain 4C amount of DNA.

mitochondria. Similarly, a distinctive feature of the early stages of apoptosis is the activation of caspase enzymes, which participate in the cleavage of protein substrates and in the subsequent disassembly of the cell.

Cell cycle study

In addition to surface immunophenotyping and cytoplasmic characterization, flow cytometry is also used in cell cycle analysis. One of the important applications of flow cytometry is the measurement of DNA content in cells. Cells spend most of time in the G0/G1 phase. In diploid cell, the amount of DNA in this phase is 2C. The duplication of the DNA at the time of cell division occurs during the S-phase. Once the cell has finished S-phase, it has twice the amount of DNA in each cell (4C). Next the cell goes through G2/M phase.

To measure the DNA content, the cells have to be stained with a fluorescent dye that binds to DNA in a stoichiometric manner (the amount of stain is directly proportional to the amount of DNA within the cell). The cells are treated with a fluorescent dye that stains DNA quantitatively. The fluorescence intensity of the stained cells at certain wavelengths will therefore correlate with the amount of DNA they contain. Some dyes possess an intercalative binding mode, such as propidium iodide, whereas others present an affinity for DNA A.T rich regions such as DAPI or G.C rich regions such as chromomycin A₃.

An accurate method for detection of cell cycle progression also uses the incorporation of thymidine analog bromodeoxyuridine (BrdU) during new DNA synthesis. This thymidine analog is used to directly determine the percentage of S-phase cells and also, when used in kinetic studies it permits a determination of the individual times for the components of the cell cycle. In this technique, BrdU is incorporated into newly synthesized DNA in cells entering and progressing through the S-phase of the cell cycle. The incorporated BrdU is then stained with specific fluorescently labeled anti-BrdU antibodies, and the levels of cell-associated BrdU measured using flow cytometry. A dye that binds to total DNA, such as 7-aminoactinomycin D (7-AAD), is often used in conjunction with immunofluorescent BrdU staining.

Typically, in a population of cells that are all proliferating rapidly but asynchronously, about 30–40% will be in S-phase at any instant and become labeled by a brief pulse of BrdU. From the proportion of cells in such a population that are labeled, we can measure the **labeling index**. Once the labeling index is determined, the duration of S-phase can be calculated by multiplying the labeling index by the mean generation time. Similarly, we can estimate the duration of M-phase, by measuring the **mitotic index** (the proportion of cells in mitosis). The duration of M-phase can be calculated by multiplying the mitotic index by the mean generation time.

If the mean generation time for a given population is known and cells are reasonably homogeneous with respect to their cell cycle kinetics, the length of G1, S and G2 can be estimated by flow cytometry. The mean generation time of cells in a population can be calculated from the change in cell number with incubation time, using the formula:

$$N = N_0 \times 2^{T/mgt}$$

where, N is the final cell number, N_0 is the initial cell number, T is the elapsed time and mgt is the mean generation time.

Fluorescence-based detection
Live cell measurement
 Hoechst 33342
 Binds to the minor groove of double-stranded DNA.
Fixed cell measurement
 Bromodeoxyuridine (BrdU)
 A thymidine analog, is incorporated into the genome during the S-phase of the cell cycle.
 Detected using anti-BrdU antibodies.
 DAPI
 Binds to the A.T rich regions of the DNA.
 Propidium Iodide (PI)
 An intercalating agent that stains the cellular genome upon cell fixation.

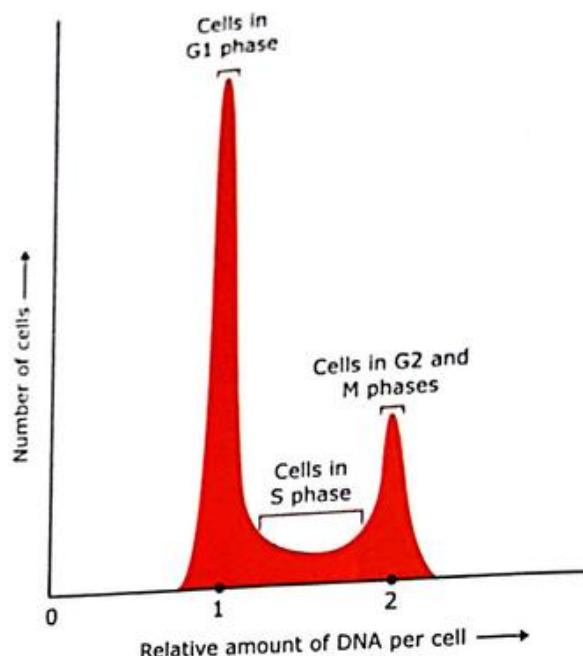


Figure 7.12 Relationship between the cell cycle and the DNA histogram. The cells analyzed here were stained with a fluorescent dye (PI) that binds to DNA, so that the amount of fluorescence is directly proportional to the amount of DNA in each cell. The cells fall into three categories: those that have an unreplicated complement of DNA and are therefore in G1, those that have a fully replicated DNA (twice the G1 DNA content) are in G2 or M-phase, and those that have an intermediate amount of DNA and are in S-phase. The distribution of cells in the case illustrated also indicates that there are greater numbers of cells in G1 than in G2 + M-phase, showing that G1 is longer than G2 and M-phase of this population. Labeling DNA with propidium iodide (PI) allows for fluorescence based analysis of cell cycle. PI stains the nucleus light to dark red in color. Amount of PI fluorescent intensity is correlated to the amount of DNA within each cell. Since the amount of DNA doubles ($2C \rightarrow 4C$) between G1 and G2 phases, the amount of fluorescent intensity of the cell population also doubles.

Chapter 8

X-ray crystallography

X-ray crystallography is a method of determining the arrangement of atoms within a crystal. This technique is based on X-ray diffraction, a nondestructive technique. When a beam of X-ray strikes a crystal, the beam may be diffracted. From the angles and intensities of these diffracted beams, a three-dimensional picture of the density of electrons within the crystal can be derived. From this electron density, the mean positions of the atoms in the crystal can be determined. This method acts as an atomic microscope, using X-rays instead of visible light to determine the three-dimensional structure of crystals. X-ray crystallography has had an enormous impact on chemistry and biology.

A crystal structure is a unique arrangement of atoms or molecules in space. The repeating unit forming the crystal is called the unit cell. Each unit cell may contain one or more molecules. A crystal is built up of many billions of unit cells. These unit cells are packed against each other in three dimensions.

When X-rays interact with a single particle, it scatters the incident beam uniformly in all directions. However, when X-rays interact with a solid material, the scattered beams can add together in a few directions and reinforce each other to yield diffraction. The regularity of the material is responsible for the diffraction of the beams. Hence, for X-ray crystallography, molecule must be crystallized. A crystal is built up of many billions of small identical units called unit cells. The unit cell is the smallest and simplest volume element that is completely representative of the whole crystal. If the internal order of the crystal is poor, then the X-rays will not be diffracted to high angles or high resolution and the data will not yield a detailed structure. If the crystal is well ordered, then diffraction will be measurable at high angles or high resolution; and a detailed structure should result.

A good way to understand X-ray diffraction is to draw an analogy with visible light. Light has certain properties that is best described by considering wave nature. Whenever wave phenomena occur in nature, interaction between waves can occur. If waves from two sources are in same phase with one another, their total amplitude is additive (*constructive interference*); and if they are out of phase, their amplitude is reduced (*destructive interference*). This effect can be seen in figure 8.1 when light passes through two pinholes in a piece of opaque material and then falls onto a white surface. Interference patterns result, with dark regions where light waves are out of phase and bright regions where they are in phase.

If the wavelength of the light (λ) is known, one can measure the angle α between the original beam and the first diffraction peak and then calculate the distance d between the two holes with the formula

$$d = \frac{\lambda}{\sin \alpha}.$$

The terms diffraction and scattering are often used interchangeably and are considered to be almost synonymous. Diffraction describes a specialized case of light scattering in which an object with regularly repeating features produces an orderly pattern.

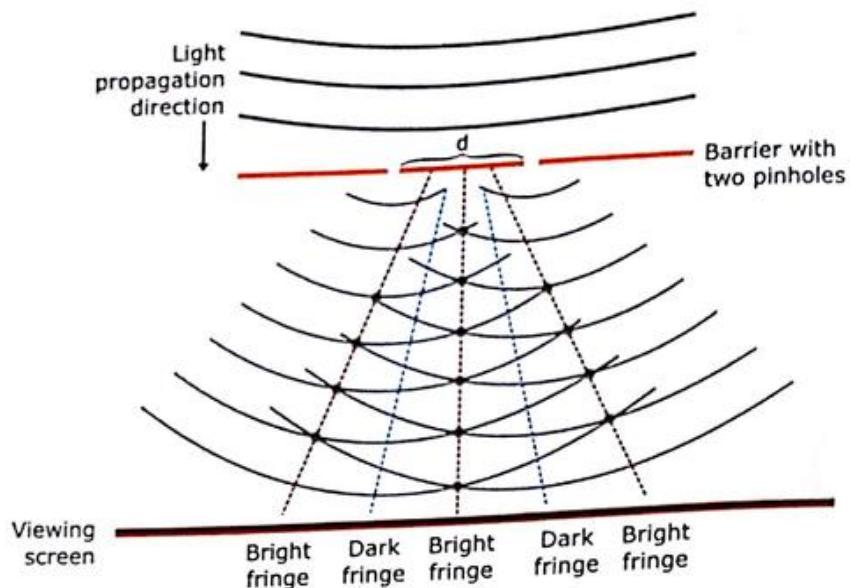


Figure 8.1 Diffraction patterns. Any energy in the form of waves will produce interference patterns if the waves from two or more sources are superimposed in space. One of the simplest patterns can be seen when monochromatic light passes through two neighbouring pinholes and is allowed to fall on a screen. When the light passes through the two pinholes, the holes act as light sources, with waves radiating from each and falling on a white surface. Where the waves are in the same phase, a bright fringe appears (*constructive interference*), but where the waves are out-of-phase they cancel each other out, producing dark fringe (*destructive interference*).

The same approach can be used to calculate the distance between atoms in crystals. Instead of visible light, which has longer wavelength to interact with atoms, we can use a beam of X-rays. X-rays, like light, are a form of electromagnetic radiation, but they have a much smaller wavelength. The wavelengths of X-rays (typically around 0.1 nm) are of the same order of magnitude as the distances between atoms or ions in a molecule or crystal. If a narrow beam of X-rays is directed at a crystalline solid, most of the X-rays will pass straight through it. A small fraction, however, scatters by the atoms in the crystal particularly with the electrons. The scattered waves reinforce one another at the film or detector if they are in phase there; and they cancel one another if they are out of phase. The way in which the scattered waves recombine depends only on the atomic arrangement.

Refraction refers to a change in the direction of a source of light, when the light moves from one medium to another while diffraction is a phenomenon that is often described as 'waves bending around corners'. Diffraction occurs when waves encounter an obstacle.

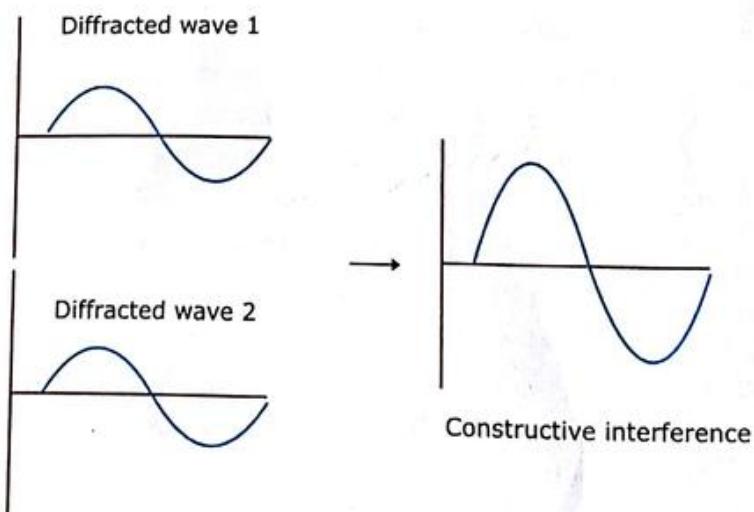
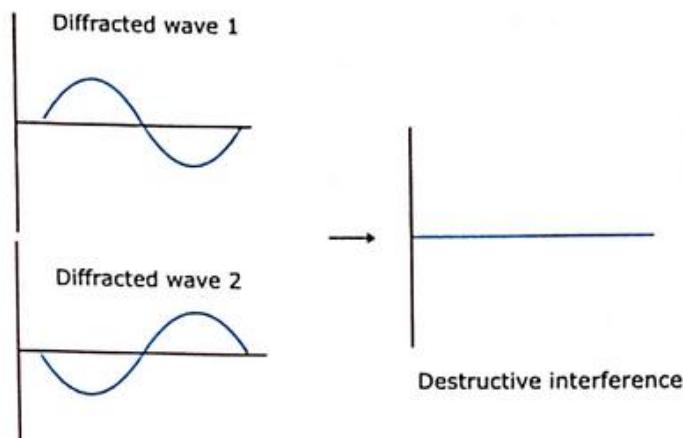


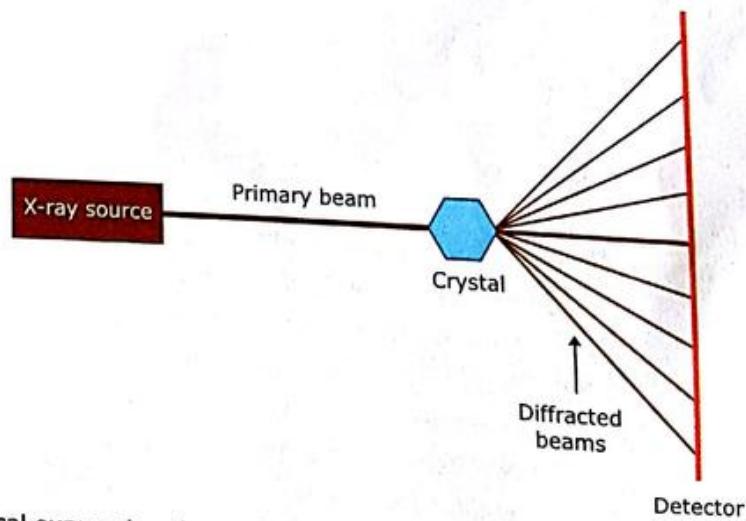
Figure 8.2 Interference occurs among the waves scattered by the atoms when crystalline solids are exposed to X-rays. There are two types of interferences depending on how the waves overlap one another. Constructive interference occurs when the waves are moving in phase with each other; and destructive interference occurs when the waves are out-of-phase.



X-rays are electromagnetic radiation of wavelengths 0.1–100 Å. X-rays can be produced by bombarding a metal target (most commonly copper, molybdenum or chromium) with electrons produced by a heated filament and accelerated by an electric field. A high-energy electron collides with and displaces an electron from a lower orbital in a target metal atom. Then an electron from a higher orbital drops into the resulting vacancy, emitting its excess energy as an X-ray photon.

A crystal is placed in an X-ray beam between the X-ray source and a detector, and a regular array of spots called reflections is generated. The spots are created by the diffracted X-ray beam; and each atom in a molecule makes a contribution to each spot. The electrons that surround the atoms are the entities which physically interact with the incoming X-ray to diffract them, not the atomic nuclei. The diffraction pattern is recorded on a photographic plate and then analyzed to reveal the nature of that lattice. The position and intensity of each spot in the X-ray diffraction pattern contain information about the positions of the atoms in the crystal that gave rise to it. An optical scanner precisely measures the position and the intensity of each reflection and transmits this information in digital form to a computer for analysis. The position of a reflection can be used to obtain the direction in which that particular beam was diffracted by the crystal. The intensity of a reflection is obtained by measuring the optical absorbance of the spot on the film, giving a measure of the strength of the diffracted beam that produced the spot.

Figure 8.3 In diffraction experiments, a narrow beam of X-ray is taken out from the X-ray source and directed onto the crystal to produce diffracted patterns. When the primary beam hits the crystal, most of it passes straight through, but some is diffracted by the crystal. These diffracted beams are recorded on a detector.



The mathematical expression that describes the interactions of X-rays with crystalline solids is Bragg's Law. This law describes the relationship between the angle at which a beam of X-rays

of a particular wavelength diffracts from a crystalline surface. When the X-ray hits an atom it can be diffracted. However, the diffraction could result in either constructive interference or destructive interference. Diffracted beams which have constructive interference will be observed and from that observation, the spacing between planes of atoms, d , can be determined. Knowing the plane spacings can lead us to different pieces of information about the material. To obtain constructive interference, the path difference between the two incidents and the scattered waves, which is $2d \sin\theta$, has to be a multiple of the wavelength λ . For this case, the Bragg's equation gives the relation between interplanar distance, d and reflection angle, θ .

$$n\lambda = 2d \sin\theta$$

where, n is an integer,

λ is the wavelength of X-ray,

d is the spacing between the planes and

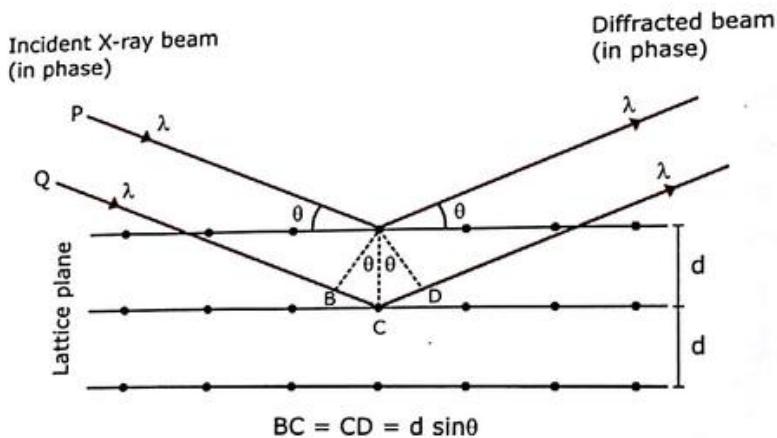
θ is the reflection angle between the incident ray and the scattering planes.

Consider, a plane lattice crystal with interplanar distance, d . Suppose a beam of X-rays of wavelength λ is incident on the crystal at an angle θ , the beam will be diffracted in all possible atomic planes. The path difference between any two diffracted waves is equal to the integral multiple of wavelength. The ray P gets diffracted from the surface, while the ray Q has to undergo some path difference. The extra distance traveled by the ray Q from the figure is $(BC + CD)$. From the diagram, either BC or CD is equal to $d \sin\theta$. So the path difference is:

$$d \sin\theta + d \sin\theta = n\lambda$$

$$2d \sin\theta = n\lambda$$

Here, n is the order = 1, 2, 3,.... This is Bragg's law.



When n is an integer (1, 2, 3 etc.), the diffracted waves from different layers are perfectly in phase with each other and produce constructive interference. Otherwise, the waves are not in phase. Diffracted beams which have constructive interference will be observed and from that observation, the spacing between planes of atoms, d , can be determined.

The relationship between the reflection angle, θ , the distance between the planes, d , and the wavelength, λ , given by Bragg's law can be used to determine the size of the unit cell.

To determine the size of the unit cell, the crystal is oriented in the beam so that diffraction is obtained from the specific set of planes in which any two adjacent planes are separated by the length of one of the unit cell axes. This distance, d , is then equal to $\lambda/(2 \sin\theta)$. The reflection angle, θ , can be calculated from the position of the diffracted spot on the film.

Fiber diffraction
 Many important biological substances do not form crystals. Among these are most membrane proteins and fibrous materials like collagen, DNA and muscle fibers. Like crystals, fibers are composed of molecules in an ordered form. When irradiated by an X-ray beam perpendicular to the fiber axis, fibers produce distinctive diffraction patterns that reveal their dimensions at the molecular level. The order in a fiber is one-dimensional (along the fiber) rather than three dimensional, as in a crystal.

The reflection angle for a diffracted beam can be calculated from the distance (r) between the diffracted spot on a film and the position where the primary beam hits the film. From the geometry shown in the figure 8.4, the tangent of the angle $2\theta = r/A$. A is the distance between crystal and film while r can be measured on the film. r is the distance from the original axis to that point where diffracted spots were observed. Hence reflection angle, θ , can be calculated.

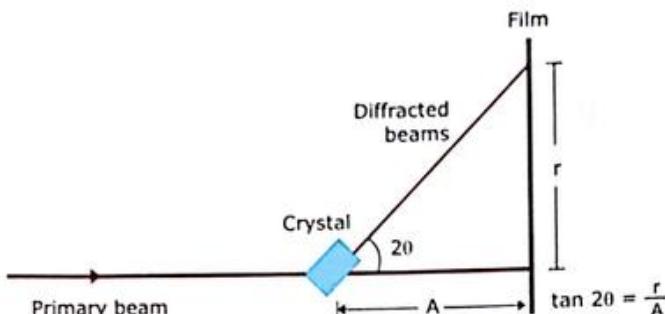


Figure 8.4 Measurement of reflection angle, θ , for a diffracted beam.

Protein crystallization

Crystallographers grow crystals of proteins by slow, controlled precipitation from aqueous solution under conditions that do not denature the protein. A number of substances cause proteins to precipitate. Ionic compounds (salts) precipitate proteins by a process called *salt-ing out*. Organic solvents also cause precipitation, but they often interact with hydrophobic portions of proteins and thereby denature them. The water-soluble polymer polyethylene glycol (PEG) is widely used because it is a powerful precipitant and a weak denaturant.

One simple means of causing slow precipitation is to add denaturant to an aqueous solution of protein until the denaturant concentration is just below that required to precipitate the protein. Then water is allowed to evaporate slowly, which gently raises the concentration of both protein and denaturant until precipitation occurs. Whether the protein forms crystals or instead forms a useless amorphous solid depends on many properties of the solution, including protein concentration, temperature, pH and ionic strength. Finding the exact conditions to produce good crystals of a specific protein often requires many careful trials and is perhaps more art than science.

References

- Berg JM, Tymoczko JL and Stryer L (2006), *Biochemistry*, 6th ed. W.H. Freeman and Company.
- Branden C and Tooze J (1999), *Introduction to protein structure*, 2nd ed. Garland Publishing, Inc.
- Drenth J (2007), *Principles of protein X-ray crystallography*, 3rd ed. Springer.
- Glusker JP, Lewis M and Rossi M (1994), *Crystal structure analysis for chemists and biologists*. New York, VCH Publishers.
- Ladd FC and Palmer RA (2003), *Structure Determination by X-ray Crystallography*, Kluwer 4th ed. Academic/plenum publishers.
- Lodish H, Berk A, Kaiser C et al (2007), *Molecular Biology of the Cell*, 4th ed. W.H. Freeman and Company.
- Verma HC (1999), *Concept of Physics*, Bharati Bhawan, India.

Chapter 9

Patch clamp techniques

The introduction of the patch clamp technique has revolutionized the study of cellular physiology by providing a method of observing the function of individual ion channels in a variety of cell types. It permits high resolution recording of the ionic currents flowing through a cell's plasma membrane.

The patch clamp technique has been invented by Sakmann and Neher in the 1976, for which they received the Nobel Prize in Physiology and Medicine in 1991. This technique is based on a very simple idea. A glass pipette with a very small opening is used to make tight contact with a small area, or patch, of cell membrane. After the application of a small amount of suction to the back of the pipette, the contact between pipette and membrane becomes so tight that no ions can flow between the pipette and the membrane. Thus, all the ions that flow when a single ion channel opens must flow into the pipette. The resulting electrical current, though small, can be measured with an ultra-sensitive electronic amplifier connected to the pipette. Based on the geometry involved, this arrangement usually is called the cell-attached patch clamp recording.

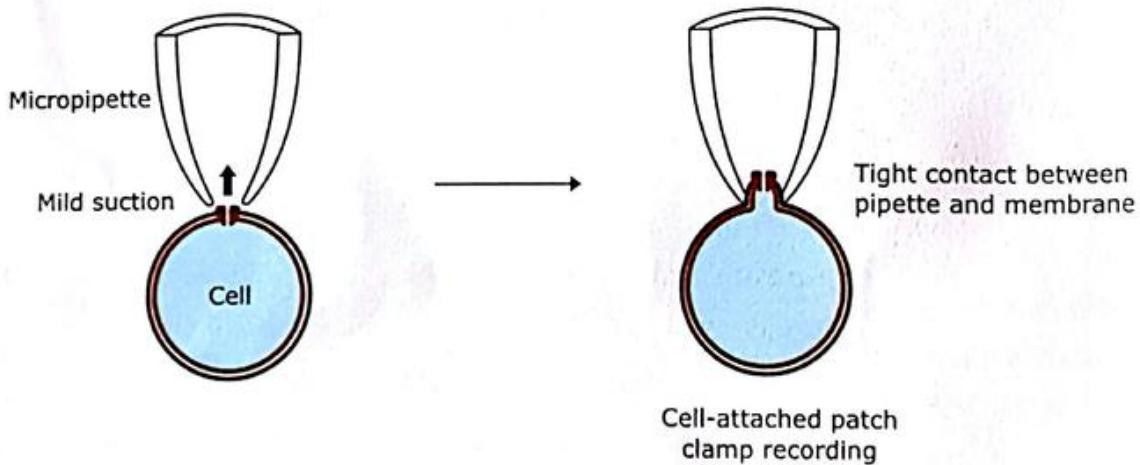


Figure 9.1 Cell-attached patch clamp. When the pipette is in closest proximity to the cell membrane, mild suction is applied to gain a tight seal between the pipette and the membrane. In the figure, two lines represent lipid bilayer of cell membrane.

In order to form the cell-attached mode, a pipette tip is placed on the surface of the cell, forming a low resistance contact (seal) with its membrane. Slight suction applied to the upper end of the pipette results in formation of a tight seal. Such a seal with a resistance in the

range of gigaohms is called 'giga-seal'. In the cell-attached mode, recordings are made from the membrane area under the pipette, while the interior of the cell remains intact. A metal electrode inside the glass pipette, containing a salt solution resembling the fluid normally found within the cell, accomplishes transduction of the ionic current into the electrical current while another one in the bath solution serves as ground. Pt and Ag/AgCl electrodes are especially used for their low junction potentials and weak polarization. The tight seal between pipette and cell membrane isolates the membrane patch electrically, which means that all ions fluxing the membrane patch flow into the pipette and are recorded by an electrode connected to a highly sensitive electronic amplifier. A bath electrode is used to set the zero level. Patch clamp technique can be operated in *voltage clamp* and *current clamp* modes. Most patch-clamp measurements are performed in a voltage-clamp mode. Voltage clamp allows an experimenter to 'clamp' the cell membrane potential (voltage) at a chosen value. This makes it possible to measure voltage specific activity of ion channels. The voltage-gated channels are primarily studied in this mode. Since the introduction of the patch clamp technique by Neher and Sakmann in 1976, patch clamp most often means voltage clamp of a membrane patch. They applied this technique to record for the first time the tiny (pico Ampere) ion currents through single channel in cell membranes.

The voltage-clamp mode is used to control the voltage of the membrane. It takes advantage of a patch-clamp amplifier which allows maintaining (clamping), through a feedback circuit, a specified membrane voltage and measuring, at the same time, the current across the membrane. In practice, during a voltage-clamp experiment, the electronic feedback system of the amplifier measures the membrane voltage and compares it to a pre-set voltage defined by the experimenter. When a current is activated, the voltage of the membrane changes. To compensate for this change and bring the voltage to the pre-set value, a current of equivalent magnitude (but opposite direction) is injected through the pipette.

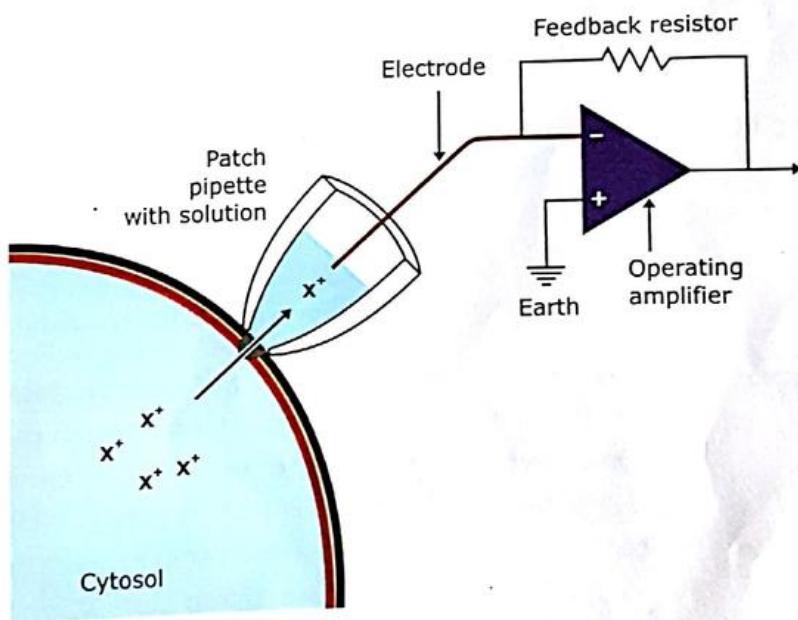


Figure 9.2 General principle of patch-clamp recordings. A glass pipette containing electrolyte solution is tightly sealed onto the cell membrane and, thus, isolates a membrane patch electrically. Currents fluxing through the channels in this patch, hence, flow into the pipette and can be recorded by an electrode that is connected to a highly sensitive differential amplifier. In the voltage-clamp configuration, a current is injected into the cell via a negative feedback loop to compensate changes in membrane potential. Recording of this current allows conclusions about the membrane conductance.

Current clamp records the membrane potential by injecting current pulse into the cell and measuring potential changes in response to it. This technique is used to study how a cell responds when electrical current enters in it. The current clamp mode allows to monitor different forms of cell activity, for example, action potentials, excitatory and inhibitory postsynaptic potentials as well as changes in membrane potentials due to activation of electrogenic membrane transporters.

Variants of the patch clamp method

Whole-cell patch clamp recording

If the membrane patch within the pipette is disrupted by briefly applying strong suction, the interior of the pipette becomes continuous with the cytoplasm of the cell. This arrangement allows measurements of electrical potentials and currents from the entire cell and is, therefore, called the *whole-cell patch clamp recording*. The whole-cell configuration also allows diffusional exchange between the pipette and the cytoplasm, producing a convenient way to inject substances into the interior of a 'patched' cell. In this mode, the pipette solution diffuses into the cell. Since the internal volume of the recording pipette is much larger than that of the cell, the pipette solution will completely substitute the intracellular solution. Depending on size and geometry of the cell and the patch pipette, the diffusion can last from several seconds to 1-2 minutes.

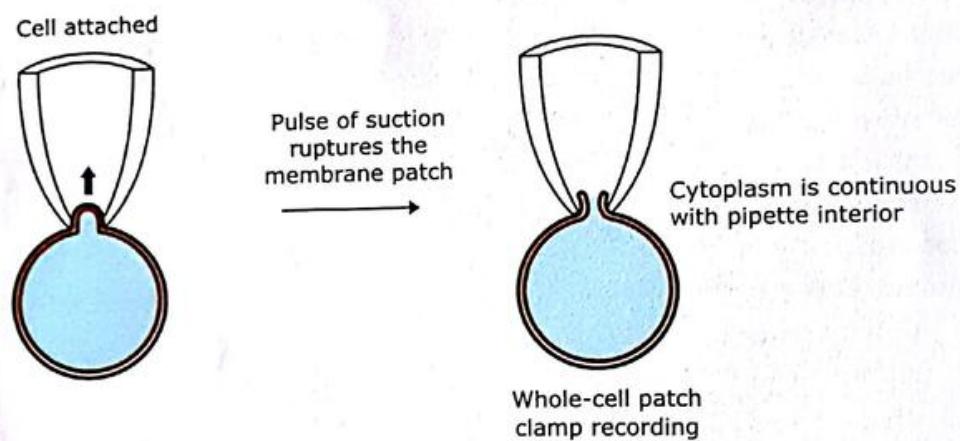


Figure 9.3 Whole-cell patch clamp recording. By applying brief but strong suction, the cell membrane is ruptured and the pipette gains access to the cytoplasm.

This method is used to record the electrical potentials and currents from the entire cell. In whole-cell measurements, the researcher can choose between two configurations: the *voltage-clamp mode* in which the voltage is kept constant and current is recorded, or the *current-clamp mode* in which the current is kept constant and changes in the membrane potential can be observed. In the current clamp mode of the whole cell patch clamp configuration, one electrode is used to inject a known current, while the other electrode is used to sense the potential inside the cell. In voltage clamp mode, the potential inside the cell is compared to a known voltage supplied by the experimenter and an amplifier is used to supply whatever current is required to keep the cell potential at the specified voltage.

There are various ways to *break* the patch membrane to obtain access to the cell interior. In the 'conventional' whole-cell technique, the membrane is disrupted with extra suction or with a brief, high-voltage pulse. In either case, a diffusional pathway is created that allows quite effective dialysis of the cell.

An alternative variant of whole-cell patch-clamp recording is the perforated-patch-clamp recording which was developed to overcome the dialysis of cytoplasmic constituents that occurs with conventional whole-cell recording. In perforated-patch clamp recordings, *perforants*, such as the antibiotics nystatin, amphotericin and gramicidin, are included in the pipette solution. These perforants form channels in the membrane attached to the patch pipette. These pores allow certain monovalent ions to permeate, enabling electrical access to the cell interior, but prevent the dialysis of larger molecules and other ions.

Inside-out and outside-out patch clamp recording

After formation of a tight seal between the membrane and the glass pipette, if small pieces of membrane can be pulled away from the cell without disrupting the seal; this yields a condition where a small patch of membrane with its intracellular surface is exposed. This arrangement is called the *inside-out patch clamp recording*. This is often used to investigate single channel activity with the advantage that the medium that is exposed to the intracellular surface can be modified.

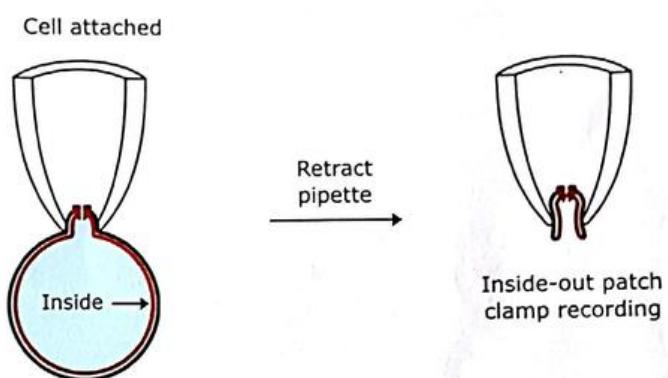


Figure 9.4 Inside-out patch clamp recording. In the cell-attached mode, the pipette is retracted and the patch is separated from the rest of the membrane and exposed to air. The cytosolic surface of the membrane is exposed.

Alternatively, if the pipette is retracted while it is in the whole-cell configuration, a membrane patch is produced that has its extracellular surface exposed. This arrangement is called the *outside-out patch clamp recording*. It is optimal for studying how channel activity is influenced by extracellular chemical signals, such as neurotransmitters.

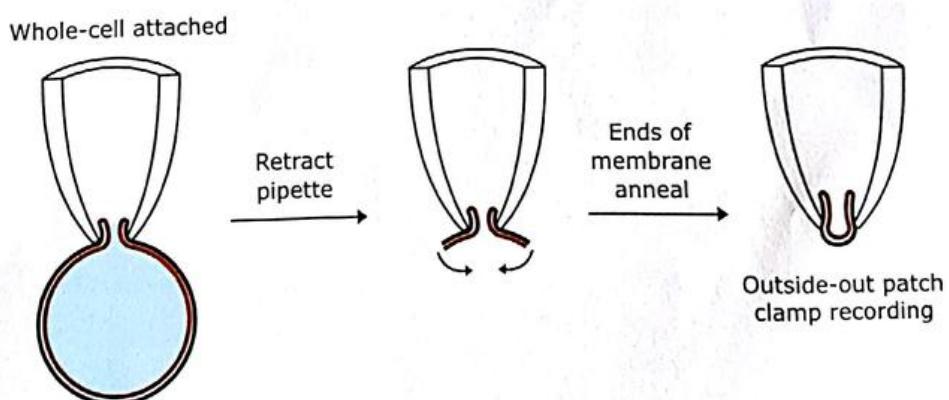


Figure 9.5 Outside-out patch clamp recording. In the whole-cell mode, the pipette is retracted resulting in two small pieces of membrane that reconnect and form a small vesicular structure with the cytosolic side facing the pipette solution.

Chapter 10

Immunotechniques

10.1 Immunoprecipitation

Antigen-antibody interaction is highly specific and occurs in a similar way as a bimolecular association of an enzyme and a substrate. The binding between antigens (Ag) and antibodies (Ab) involve weak and reversible non-covalent interactions consist mainly of van der Waals forces, electrostatic forces, H-bonding and hydrophobic forces. The smallest unit of antigen that is capable of binding with antibodies is called an antigenic determinant (or epitope). The corresponding area on the antibody molecule combining with the epitope is called paratope. The number of epitopes on the surface of an antigen is its valence. The valence determines the number of antibody molecules that can combine with the antigen at one time. If one epitope is present, the antigen is monovalent. Most antigens, however, have more than one copy of the same epitope and are termed polyvalent.

Immunoprecipitation reaction results from the interaction of a soluble antibody with a soluble antigen to form an insoluble complex. Antibodies that aggregate soluble antigens are called precipitins. Formations of an Ag-Ab lattice depend on the valency of both antibody and antigen. The antibody must be bivalent for precipitation reaction to occur. Monovalent Fab fragments cannot form precipitate with antigen. Similarly, the antigen must be either bivalent or polyvalent. If the antigen is bi- or polyvalent, it can bind with multiple antibodies. Eventually, the resulting cross-linked complex becomes so large that it falls out of solution as a precipitate. Immunoprecipitation reaction can be performed in solution or in gel.

Immunoprecipitation reaction in solution

Precipitation occurs maximally only when there are optimal proportions of the two reacting substances – antigen and antibody. Hence, an insoluble antibody-antigen complex formation occurs within a narrow concentration range known as the zone of equivalence. This represents the conditions under which antigen-antibody complexes are formed that are sufficiently large to be precipitated. Outside the equivalence concentration, conditions known as antigen or antibody excess occur, which result in the formation of small, soluble complexes. When increasing concentrations of antigen are added to a series of tubes that contain a constant concentration of antibodies, variable amounts of precipitate form. If the amount of the precipitate is plotted against the amount of antigen added, a precipitin curve, as shown in the

figure 10.1, is obtained. This immunoprecipitation reaction can be used to remove particular antigens from a solution.

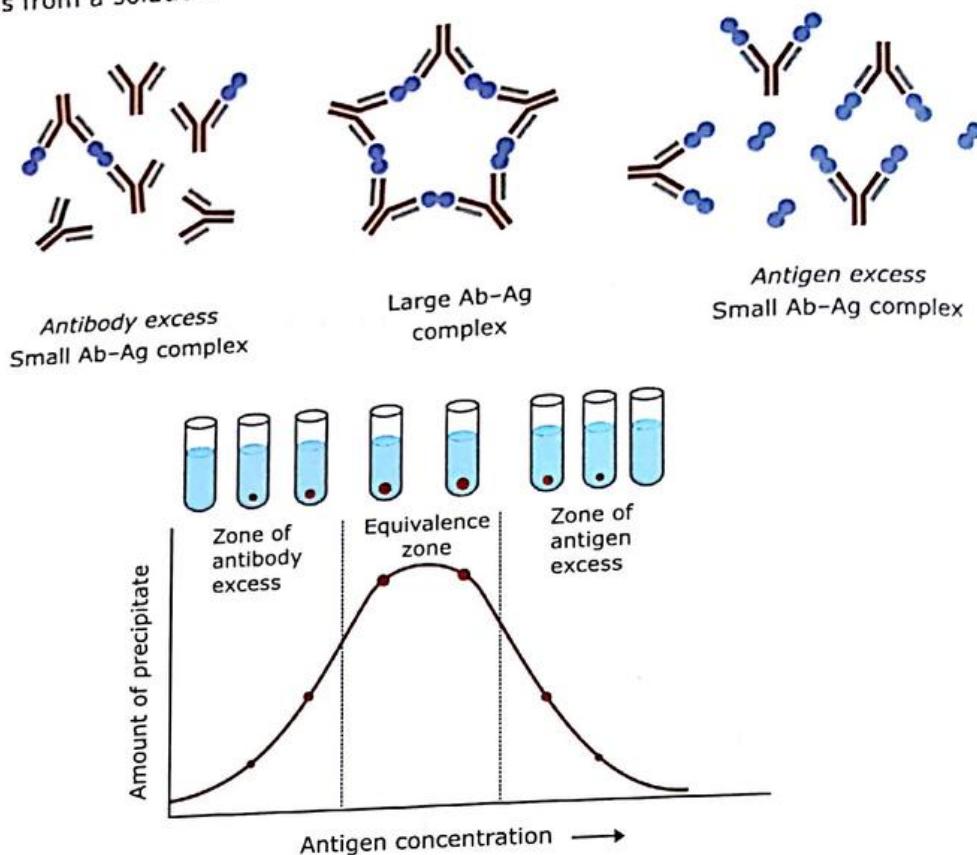


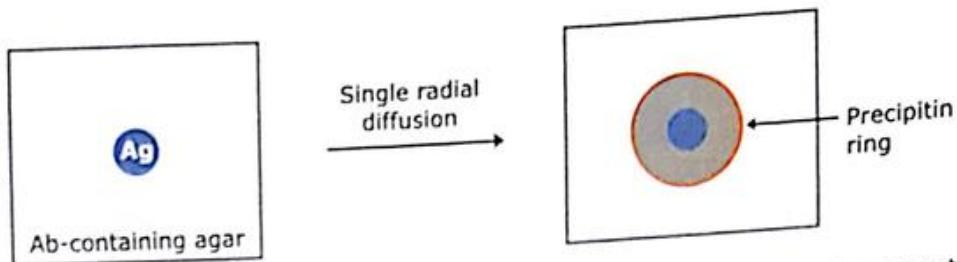
Figure 10.1 A precipitation reaction can be performed by placing a constant amount of antibody in a series of tubes and adding increasing amounts of antigen to the tubes. After the precipitate forms, each tube is centrifuged to pellet the precipitate, the supernatant is poured off, and the amount of precipitate is measured. Plotting the amount of precipitate against increasing antigen concentrations yields a precipitin curve. Under conditions of antibody excess or antigen excess, extensive lattices do not form and precipitation is inhibited. In the equivalence zone, the ratio of antibody to antigen is optimal. As a result, large multimeric lattice is formed at equivalence, the complex increases in size and precipitates out of solution.

Immunoprecipitation reactions in gels

Immunoprecipitation reactions carried out in agar gels are referred to as *immunodiffusion reactions*. When antigen and antibody diffuse toward one another in gel, or when antibody is incorporated into the gel and antigen diffuses into the antibody-containing matrix, a visible line of precipitation (*precipitin line*) will form. Visible precipitation occurs in the region of equivalence. No visible precipitate forms in regions of antibody excess or antigen excess. Two types of immunodiffusion reactions can be used to determine the relative concentrations of antibodies or antigens as well as the identity of antigens. These immunodiffusion techniques are *radial immunodiffusion* and *double immunodiffusion*.

Radial immunodiffusion (Mancini method)

The relative concentration of an antigen can be determined by a simple quantitative assay in which an antigen sample is placed in a well and allowed to diffuse into agar gel containing antibody. In agar gel, antibody is uniformly distributed. At the region of equivalence, a precipitation ring forms around the well. The diameter of the ring is proportional to the log of the concentration of antigen since the amount of antibody is constant.



It is used for the quantitative estimation of antigen. By running different concentrations of a standard antigen on the gel and by measuring the diameters of their precipitin rings, a standard calibration graph is plotted. Antigen concentrations of unknown samples, run on the same gel, can be found by measuring the diameter of precipitin rings and extrapolating this value on the calibration graph.

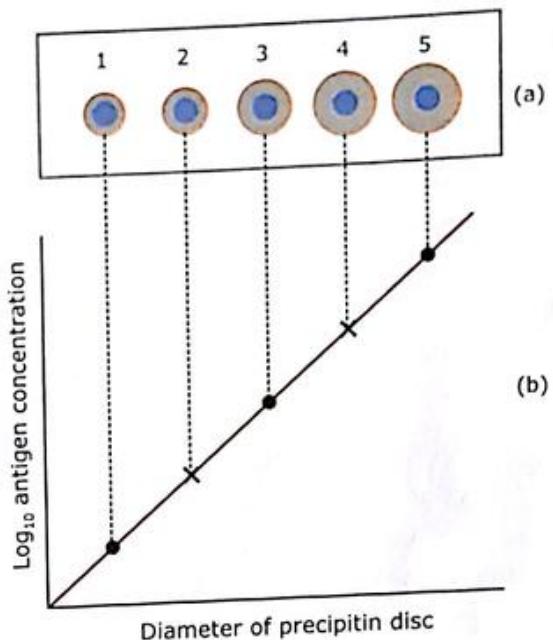


Figure 10.2 Single radial immunodiffusion. (a) Precipitin ring formed in gel containing monospecific antiserum. Wells 1, 3 and 5 contained standard antigen solutions of increasing concentration. Wells 2 and 4 contained samples of the antigen at unknown concentrations. (b) Semilog plot of diameter of the precipitin discs of standard antigen solutions (•) against concentration. Measurement of the diameters of the precipitin discs of the unknown solutions (x) allows an estimation of the antigen concentration to be made by simple interpolation.

Ouchterlony double immunodiffusion

In the Ouchterlony double immunodiffusion (developed by Ouchterlony), both antigen and antibody diffuse radially from wells toward each other, thereby establishing a concentration gradient. Antigens and antibodies diffuse toward each other at rates that increase in proportion to their concentration in the well, but decrease in proportion to their sizes. They form a precipitin line where they meet at equivalence. It is an immunological technique used for the detection, identification and quantification of antibodies and antigens.

Using this technique, the antigenic relationship between two antigens can be analyzed. Distinct precipitation line patterns are formed against the same antibody depending on whether two antigens share all antigenic epitopes or partially share their antigenic epitopes or do not share their antigenic epitopes at all.

- *Identity* occurs when two antigens share identical epitopes.
- *Non-identity* occurs when two antigens are unrelated i.e. share no common epitopes. The antisera form an independent precipitin line with each antigen, and the two lines cross.

- *Partial identity* occurs when two antigens share some epitope but one of the other has a unique epitope. The antiserum forms a line of identity with the common epitope and a curved spur with the unique epitope.

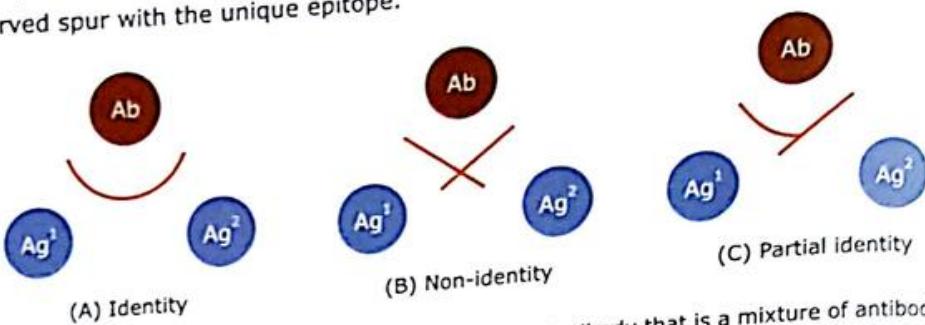


Figure 10.3 Ouchterlony double immunodiffusion patterns. Antibody that is a mixture of antibody-1 and antibody-2 is placed in the central well. Unknown antigens are placed in the outside wells. (A) Identity : The arc indicates that the two antigens (Ag^1 and Ag^2) are identical. (B) Non-identity : Two crossed lines represent two different precipitation reactions. The antigens share no common epitopes. (C) Partial identity : Ag^1 and Ag^2 share some epitopes.

The Ouchterlony method also can be used to estimate the relative concentration of antigens. When an antigen has a relatively higher concentration, the equivalent zone will be formed a little bit away from the antigen well. When an antigen has a relatively lower concentration, the equivalent zone will be formed a little bit closer to the antigen well.

Agglutination reactions

Agglutination reactions
Agglutination is the visible clumping of a particulate antigen when mixed with antibodies specific for the particulate antigens. Antibodies that produce such reactions are called **agglutinins**. The general term agglutinin is used to describe antibodies that agglutinate particulate antigens. When the antigen is an erythrocyte, the term **hemagglutination** is used. All antibodies can theoretically agglutinate particulate antigens; but IgM, due to its high valence, is a particularly good agglutinin.

particularly good agglutinin.

Agglutination reactions are similar, in principle, to precipitation reactions. Just as an excess of antibody inhibits precipitation reactions, such excess can also inhibit agglutination reactions; this inhibition is called the *prozone effect*. In the case of antigen excess, the *postzone effect* occurs. The agglutination reactions can be *direct* (active) or *indirect* (passive). When the antigen is an integral part of the surface of a cell or other insoluble particle, the agglutination reaction is referred to as direct agglutination. The agglutination test only works with particulate antigens. However, it is possible to coat cell or other insoluble particle with a soluble antigen (e.g. viral antigen, a polysaccharide or a hapten) and use the coated cells in an agglutination test for antibody to the soluble antigen. This is called passive agglutination.

Coombs test

Coombs test
When antibodies bind to erythrocytes, they do not always result in agglutination. This can result from the antigen-antibody ratio (either antigen excess or antibody excess); or in some cases, *zeta potential* on the erythrocytes preventing the effective cross-linking of the cells. In order to detect the presence of non-agglutinating antibodies on erythrocytes, one simply adds a second antibody directed against the antibodies attached to their respective epitopes on erythrocytes. This anti-immunoglobulin can now cross-link the erythrocytes and result in agglutination. This test is known as the *Coombs test* (Anti-immunoglobulin test).

Zeta potential
The surfaces of certain particulate antigens may possess an electrical charge, as, for example, the net negative charge on the surface of red blood cells caused by the presence of sialic acid. When such charged particles are suspended in saline solution, an electrical potential, termed the zeta potential, is created between particles, preventing them from getting very close to each other. This introduces a difficulty in agglutination of charged particles by antibodies.

The Coombs test is based on two important facts: 1. that antibodies of one species (e.g. human) are immunogenic when injected into another species (e.g. rabbit) and lead to the production of antibodies against the antibodies, and 2. that many of the anti-immunoglobulins (e.g. rabbit anti-human Ab) bind with antigenic determinants present on the Fc portion of the antibody and leave the Fab portions free to react with antigen. For example, if human IgG antibodies are attached to their respective epitopes on erythrocytes, then the addition of rabbit antibodies to human IgG will result in their binding with the Fc portions of the human antibodies bound to the erythrocytes by their Fab portions. These rabbit antibodies not only bind with the human antibodies that are bound to the erythrocyte but also, by doing so, form cross-links.

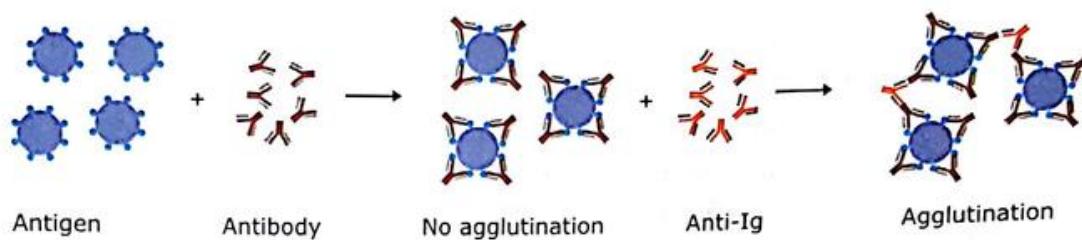


Figure 10.4 Indirect Coombs test.

There are two versions of the Coombs test: *direct* and *indirect Coombs test*. The two versions differ somewhat in the mechanics of the test, but both are based on the same principle. In the direct test, anti-immunoglobulins are added to the erythrocytes that are suspected of having antibodies bound to antigens on their surfaces. The indirect test is used to detect the presence of antibodies specific to antigens on the erythrocytes in the serum.

10.2 Immunoassays

Immunoassays are based on the specific antibody-antigen reactions. All immunoassays depend on the measurement of fractional binding site occupancy of the antibody by analytes (antigens). Immunoassays can be noncompetitive and competitive. *Noncompetitive immunoassays* directly measure antigen-occupied sites, while *competitive assays* are based on the detection of total and unoccupied sites by antigen. In competitive immunoassays, labeled and unlabeled antigens are simultaneously exposed to the antibody. In noncompetitive immunoassays, the labeled antibody detects the bound antigens. Maximal sensitivity will be reached by decreasing the amount of antibody in competitive assays and increasing the antibody concentration in noncompetitive assays.

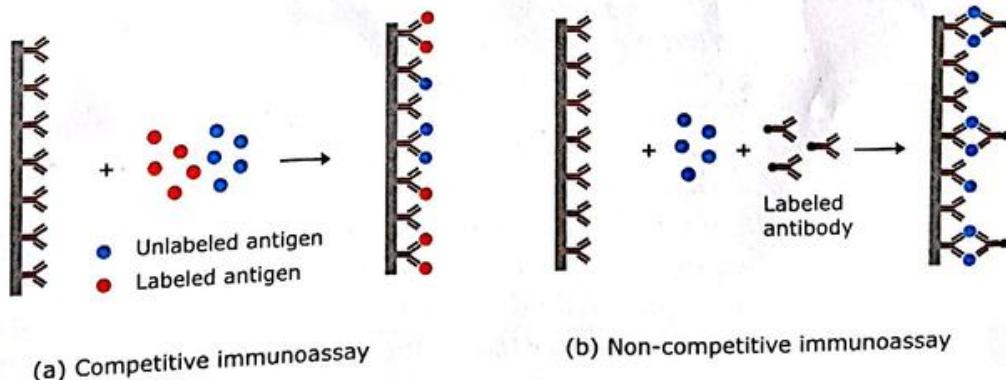


Figure 10.5 Competitive and non-competitive immunoassays.

Radioimmunoassay

Radioimmunoassay (RIA) is a technique for determining the concentration of a particular antigen in a sample, based on competitive binding between unlabeled and radioisotope labeled antigen for its specific antibody. Rosalyn S. Yalow and Solomon A. Berson developed the RIA method, and first applied it to the assay of human and animal insulins in plasma. It is one of the most sensitive techniques for detecting antigen or antibody. The principle of RIA involves competitive binding of radiolabeled antigen and unlabeled antigen to a high-affinity antibody. Typically, a fixed amount of the antibody is attached to the bottom of a tube. Then a fixed amount of radiolabeled antigen is added to the tube. The amount of radiolabeled antigen is typically high enough to saturate all of the antibody molecules attached to the bottom of the tube. Because the interaction between the antibody and the antigen involves very strong non-covalent interactions, the contents of the tube can be discarded without disturbing the radiolabeled antigen that are bound to the antibody molecules in the bottom of the tube. The tube can now be counted in a gamma counter to obtain a Counts Per Minute (CPM). This value represents the total count, and represents a situation when all of the antibody binding sites are occupied by the radiolabeled antigen. This value alone, however, tells us nothing.

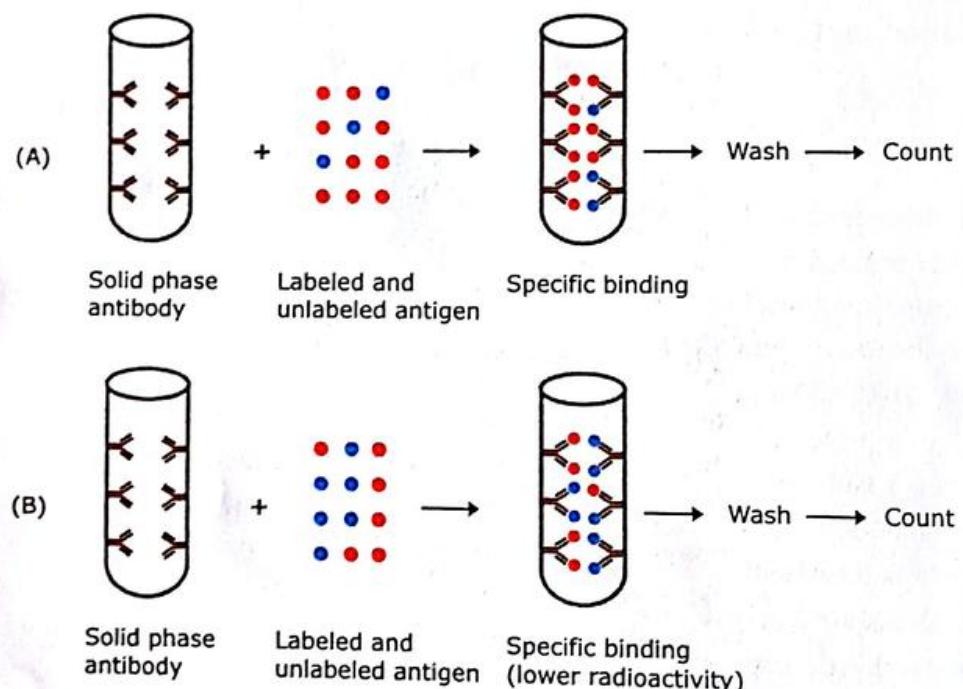


Figure 10.6 Principle of RIA. Labeled antigen competes with unlabeled antigen for a limited number of binding sites on solid-phase antibody. (A) Very little unlabeled antigen is present, making radioactivity of the solid phase high. (B) More unlabeled antigen is present, and the radioactivity of the solid phase is reduced in proportion to the amount of unlabeled antigen bound.

In order to calibrate the measurements, known amounts of unlabeled antigen are added to the tubes that contain the same fixed amount of antibody attached to the bottom, and in addition, the same fixed amount of radiolabeled antigen in the tube. Now in the tube, there is competition between the unlabeled and radiolabeled antigen to bind to the antibody binding sites. The relative amount of unlabeled or radiolabeled antigen that binds to the antibody is strictly a function of their relative amounts in the tube. Because the radiolabeled and the unlabeled antigen bind to the antibody with the same affinity, the higher the concentration of the unlabeled antigen, the better it can compete for the binding sites. Therefore, less radiolabeled antigen will bind, which will result in a smaller CPM. In this fashion, a standard curve

can be obtained by measuring counts from tubes, which have increasing known concentrations of the unlabeled antigen. The concentration of antigen in an unknown sample can be determined by reference to a standard curve.

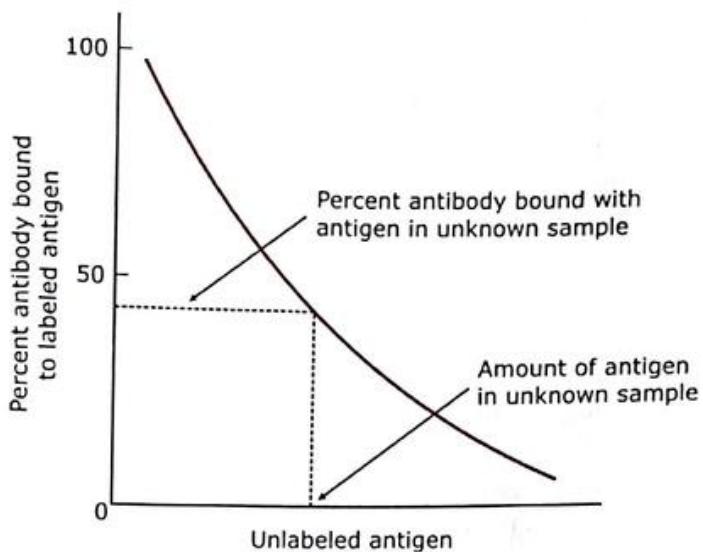


Figure 10.7 A standard curve used in radioimmunoassay for quantitative estimation of antigen in an unknown sample. The concentration of antigen in an unknown sample can be determined by reference to a standard curve constructed from data obtained by allowing varying amounts of unlabeled antigen to compete.

Enzyme-linked immunosorbent assay

Enzyme immunoassays exploit an enzymatic reaction for detecting the immune reaction. In 1971, Engvall and Perlmann and van Weemen and Schuurs described independently the use of enzyme-labeled agents. The most common type of enzyme immunoassays in use is *enzyme-linked immunosorbent assays*.

Enzyme-Linked Immuno Sorbent Assay, commonly known as ELISA (or EIA), is similar in principle to RIA; but depends on an enzyme rather than a radioactive label. An enzyme conjugated with an antibody reacts with a colorless substrate to generate a colored reaction product. Such a substrate is called chromogenic substrate. A number of enzymes have been employed for ELISA, including alkaline phosphatase, horseradish peroxidase, urease and beta galactosidase.

Table 10.1 Enzymes used for conjugation of antibodies

Enzyme	Source	Reaction catalyzed
Peroxidase	Horseradish	$H_2O_2 + \text{Oxidisable substrate} \rightarrow \text{Oxidized product} + 2H_2O$
Alkaline phosphatase	Calf intestine	$R-O-P_i + H_2O \rightarrow R-OH + P_i$
β -Galactosidase	<i>E. coli</i>	β -D-Galactoside + $H_2O \rightarrow$ Galactose + Alcohol
Urease	Jack bean	$(NH_2)_2CO + 3H_2O \rightarrow CO_2 + 2NH_4OH$

Different variants of ELISA have been developed, allowing qualitative detection or quantitative measurement of either antigen or antibody. Each type of ELISA can be used qualitatively to detect the presence of antibody or antigen. Alternatively, a standard curve based on known concentrations of antibody or antigen is prepared from which the unknown concentrations of a sample can be determined.

Indirect ELISA

Ab can be detected or quantitatively determined with an indirect ELISA. Serum or some other sample containing primary antibody (Ab_1) is added to an antigen-coated microtiter well and allowed to react with the antigen attached to the well. After any free antibody is washed away, the presence of antibody bound to the antigen is detected by adding an enzyme-conjugated secondary antibody (Ab_2), which binds to the primary antibody. Any free Ab_2 then is washed away, and a substrate for the enzyme is added. The amount of colored reaction product that forms is measured by specialized spectrophotometric plate readers, which can measure the absorbance of all of the wells.

The indirect ELISA is used to detect the presence of antibody against HIV. In this test, viral core proteins (the antigen) are adsorbed to the bottom of a well. Then antibodies from a patient are added to the coated well and allowed to bind to the antigen. Finally, the enzyme-linked antibodies to the human antibodies are allowed to react in the well and unbound antibodies are removed by washing. Substrate is then applied. An enzyme reaction suggests that the enzyme-linked antibodies were bound to human antibodies, which in turn, implies that the patient had antibodies to the viral antigen.

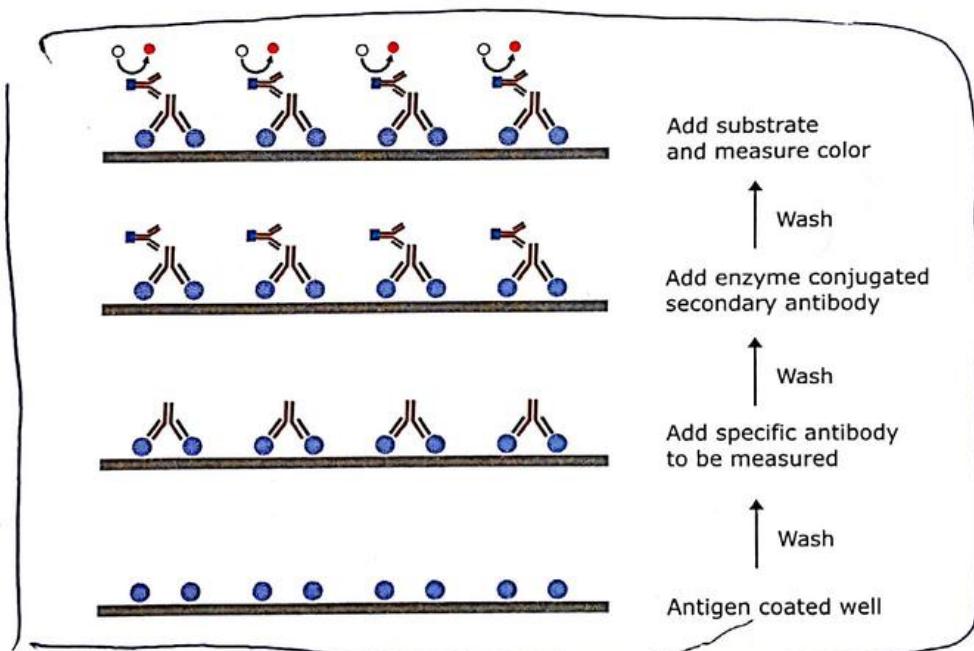


Figure 10.8 Indirect ELISA. The antigen of interest is first immobilized on the bottom of a well. Test antiserum is added and allowed to incubate. If any antibodies in the test antiserum have bound to the immobilized antigen, their presence is detected by adding an enzyme-conjugated secondary antibody. Enzyme substrate is then added and the amount of colored reaction product that forms is measured.



Sandwich ELISA

Antigen can be detected or measured by a sandwich ELISA. In this technique, the antibody (rather than the antigen) is immobilized on a microtiter well. A sample containing antigen is added and allowed to react with the immobilized antibody. After the well is washed, a second enzyme-linked antibody specific to a different epitope on the antigen is added and allowed to react with the bound antigen. Any free second antibody then is washed away and a substrate for the enzyme is added. Finally, the amount of colored reaction product that forms is measured.

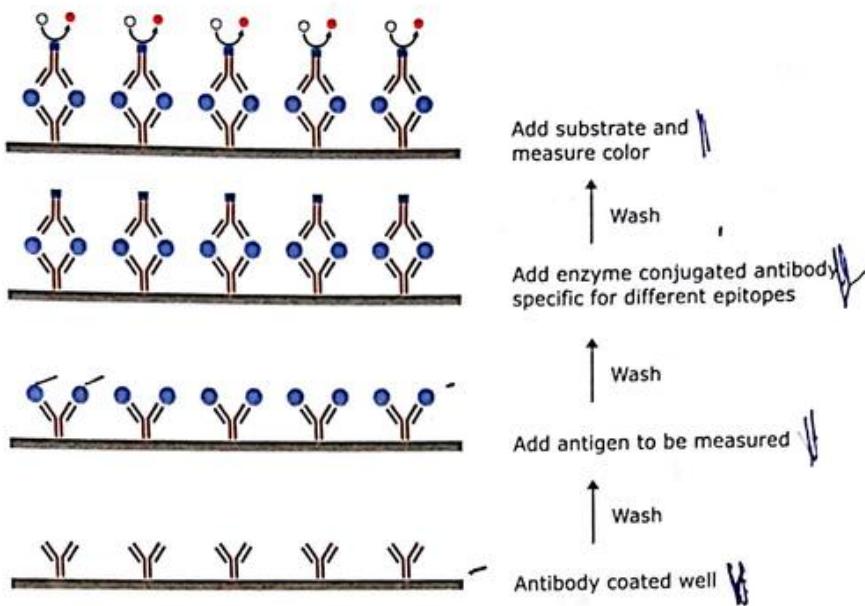


Figure 10.9 Sandwich ELISA. Antibody to a particular antigen is first immobilized on the bottom of a well. Next, the antigen is added to the well which binds to the antibody. Finally, a second, different antibody to the antigen is added. This antibody is enzyme-linked. After any free second antibody is removed by washing, substrate is added, and the colored reaction product is measured. The extent of reaction is directly proportional to the amount of antigen present.

✓ Immunofluorescence

Immunofluorescence is a powerful technique that utilizes fluorescent-labeled antibodies to detect specific target antigens. Fluorescent molecules absorb light of one wavelength (excitation) and emit light of another wavelength (emission). If antibody molecules are tagged with a fluorescent dye or fluorophore, immune complexes containing these fluorescently labeled antibodies can be detected by the colored light emission when excited by the light of the appropriate wavelength. Antibody molecules bound to antigens in cells or tissue sections can similarly be visualized. This technique is known as immunofluorescence. The most commonly used fluorescent dyes are fluorescein and rhodamine. Other dyes such as phycoerythrin and phycobiliproteins are also used now-a-days.

Immunofluorescence is of two types: One is the direct fluorescent antibody method (or direct immunofluorescence) and the other is the indirect fluorescent antibody method (or indirect immunofluorescence). In direct fluorescent antibody method, the specific antibody is directly conjugated with fluorophore whereas in indirect fluorescent antibody method, the primary antibody is unlabeled and is detected with a fluorophore-labeled secondary antibody. It is a two-step process, in which a primary, unlabeled antibody binds to the target, after which a fluorophore-labeled secondary antibody (directed against the Fc portion of the primary antibody) is used to detect the first antibody.

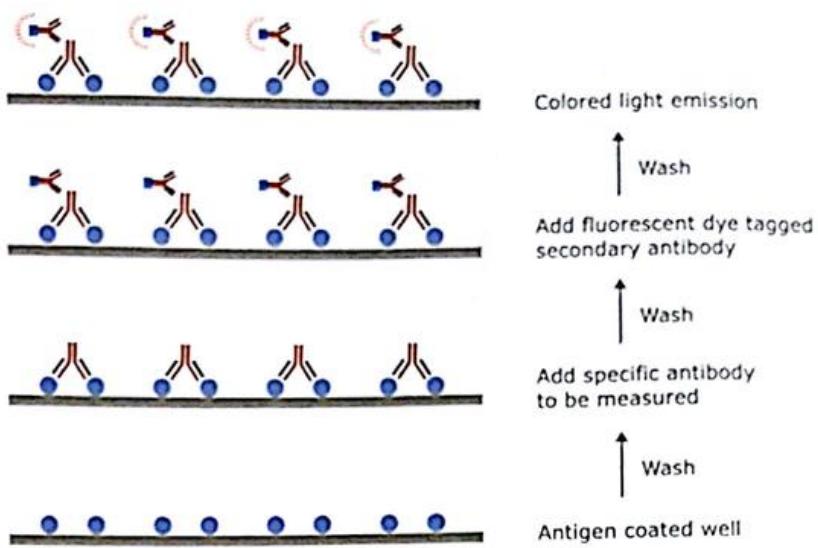


Figure 10.10 Indirect or secondary immunofluorescence. It uses two antibodies; the unlabeled first (primary) antibody specifically binds the target molecule, and the secondary antibody, which carries the fluorophore, recognizes the primary antibody and binds to it. Multiple secondary antibodies can bind a single primary antibody. This provides signal amplification by increasing the number of fluorophore molecules per antigen.

10.3 Immunoelectrophoresis

Immunoelectrophoresis is an immunotechnique that combined electrophoresis and immunoprecipitation for identifying and characterizing antigens within complex mixtures. *Electrophoresis* is conducted in the first stage and *immunoprecipitation* using antibodies against specific antigens in the second stage without removing the antigens from the gel.

There are several variants of immunoelectrophoresis like classical immunoelectrophoresis, crossed immunoelectrophoresis, rocket immunoelectrophoresis and immunofixation electrophoresis.

Classical immunoelectrophoresis

Classical immunoelectrophoresis combines the process of electrophoresis as well as double immunodiffusion. Electrophoresis is used for separation of antigens on the basis of charge and double immunodiffusion for the purpose of their identification. It is also used in clinical laboratories to detect the presence or absence of antigen in the serum.

In this technique, antigen mixture is placed in a well cut in the center of an agar gel. Then the antigen mixture is subjected to electrophoresis, which separates the various components according to their charge in the electrical field. After electrophoresis, a trough is cut along the side of the gel, and antibodies are applied in the trough. The antibodies and the separated antigen mixture diffuse in the agar. At an optimal antigen-to-antibody ratio for each antigen and its corresponding antibodies, precipitin lines form.

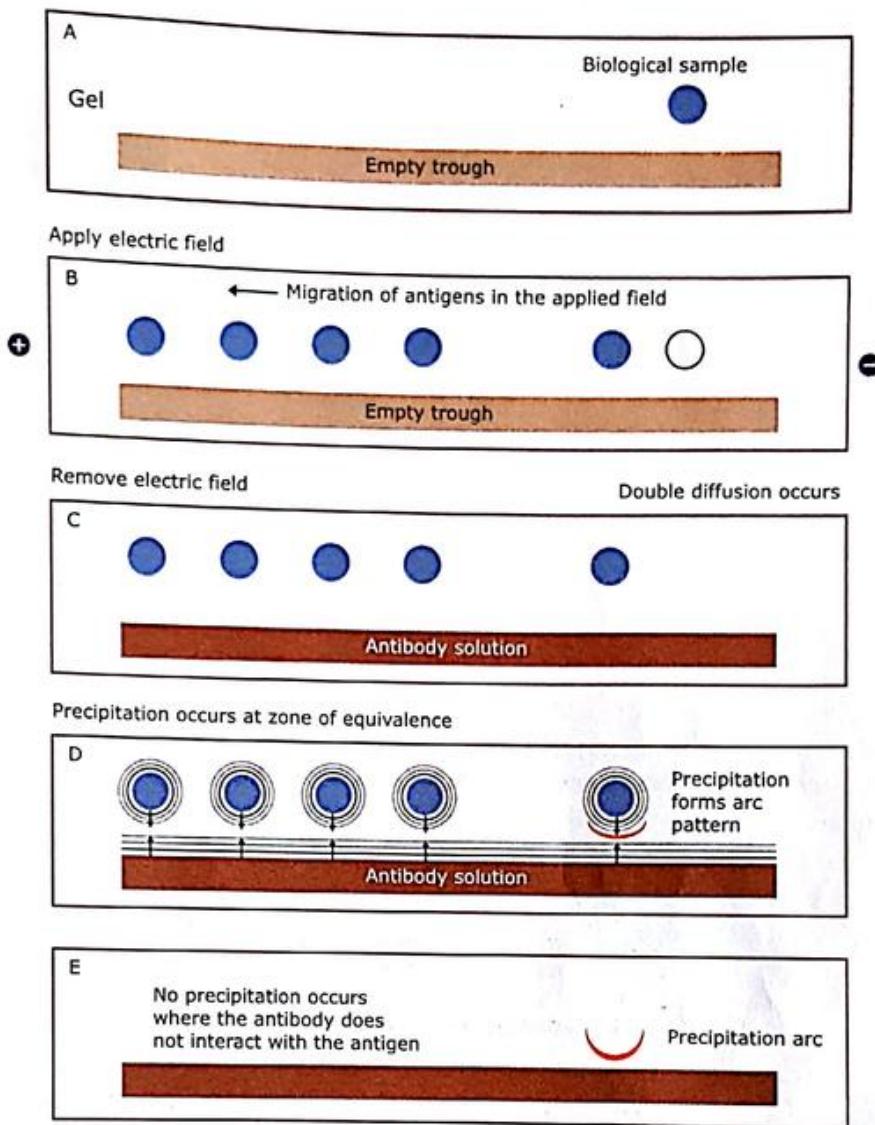


Figure 10.11 Classical immunolectrophoresis. Sample (Ag) is introduced into a well cut in the gel and subjected to a voltage gradient which causes the various antigens to migrate different distances through the gel depending on their charge. After electrophoresis, antibody is introduced in the trough that is located parallel to the separated antigens. Both the antigens that have been electrophoretically separated and the antibodies present in trough diffuse in the gel. When they meet at the appropriate concentrations (equivalence), precipitation occurs.

Crossed immunoelectrophoresis

Crossed immunoelectrophoresis (also known as two-dimensional immunoelectrophoresis) is a two-stage technique in which antigens are electrophoretically separated in one dimension and then subjected to a second electrophoresis perpendicular to the first electrophoresis. The second stage differs from the first in that the antigens move through an agarose gel containing antibodies specific to the antigen. This leads to the formation of a precipitin arc when antigen and antibody meet at equivalence.

This technique is also used for the separation of proteins that have similar electrophoretic mobility. *This technique should not be confused with two-dimensional gel electrophoresis, a technique in which proteins are first separated according to their isoelectric points, and then separated according to their masses.*

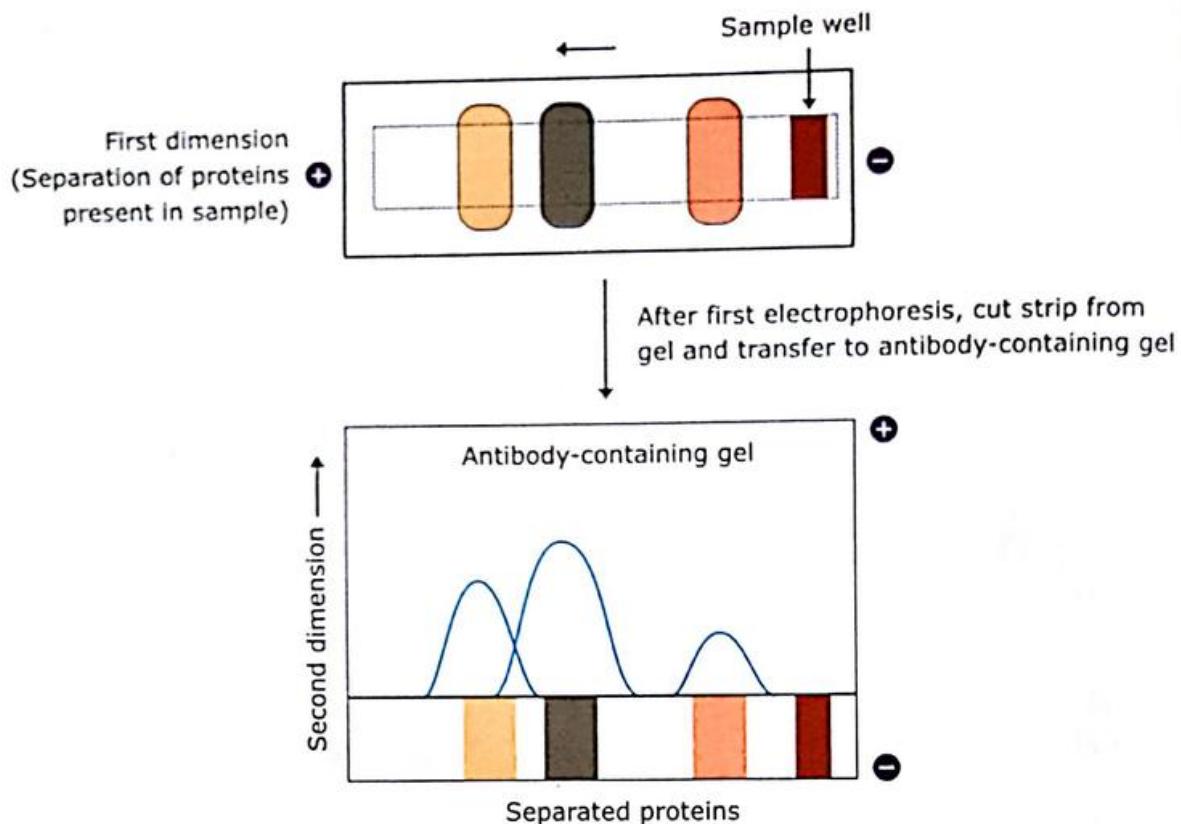


Figure 10.12 Crossed immunoelectrophoresis. The first stage of crossed immunoelectrophoresis consists of protein separation under the influence of an applied field. In preparation for the second stage, liquid agarose, containing the antibody solution of interest, is poured onto a large plate in a manner that allows the liquid agarose to fuse with the piece of gel that has been transferred to the plate. When the gel is hardened, an electric field is applied. The electric field is perpendicular to that applied to the protein sample in the first gel. Precipitation occurs when the proteins (antigens) and antibodies interact in appropriate concentration (equivalence).

Rocket immunoelectrophoresis

Rocket immunoelectrophoresis (also referred to as electroimmunoassay) is a simple, quick, and reproducible method for determining the concentration of a specific protein antigen in a mixture. In this method, a negatively charged antigen is electrophoresed in an antibody-containing gel. During electrophoresis, as the antigen starts to leave the well and move in the gel, antigen molecules will start to interact, and bind with antibody molecules. However, at this early stage, there is considerable antigen excess over antibody and no precipitation occurs. As the antigen samples electrophoreses further through the gel, more antibody molecules are encountered that interact with the antigen, until eventually there is sufficient antibody-antigen cross-linking such that 'equivalence' is reached and the antigen-antibody complex precipitates. The precipitin lines appear like the shape of a rocket. The majority of the antibody-antigen precipitate is, indeed, at the head of this rocket, but the fine precipitation lines up the side of the rockets are formed by a small amount of antigen diffusing sideways as the antigen passes through the gel. This small amount of antigen very quickly meets sufficient antibody to reach equivalence and precipitate.

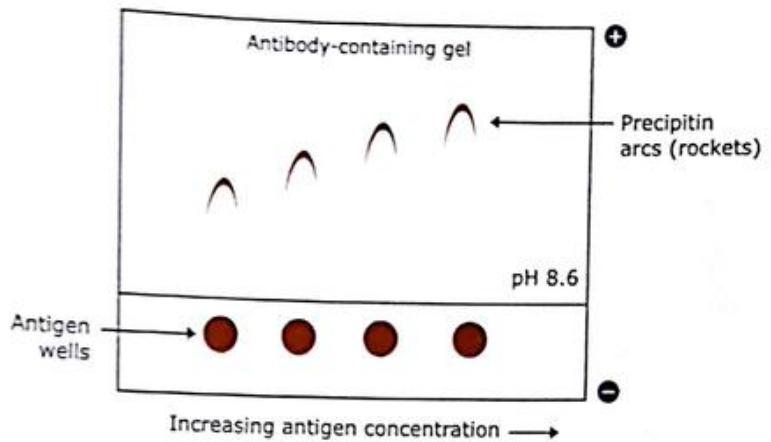


Figure 10.13 Rocket electrophoresis. Antigen is electrophoresed into gel containing antibody. The samples to be compared are loaded side-by-side in small circular wells along the edge of an agarose gel. These samples (antigens) are then electrophoresed into the agarose gel where interaction between antigen and antibody takes place. The distance from the starting well to the front of the rocket shaped arc is related to antigen concentration.

Immunofixation electrophoresis

The principle of *immunofixation electrophoresis* (IFE) is simple. The antibodies are applied to the surface of a gel after electrophoresis, and at antigen-antibody equivalence, they form large precipitates with their counterparts in the gel. Upon washing, unbound antibodies and other proteins are removed from the gel, leaving only the immunoprecipitates which are too large and insoluble to be washed out from the pores. The immunoprecipitates remaining in the gel can be then directly stained or identified by other techniques such as by means of fluorescein, enzyme, or radiolabeled primary or secondary antibodies.

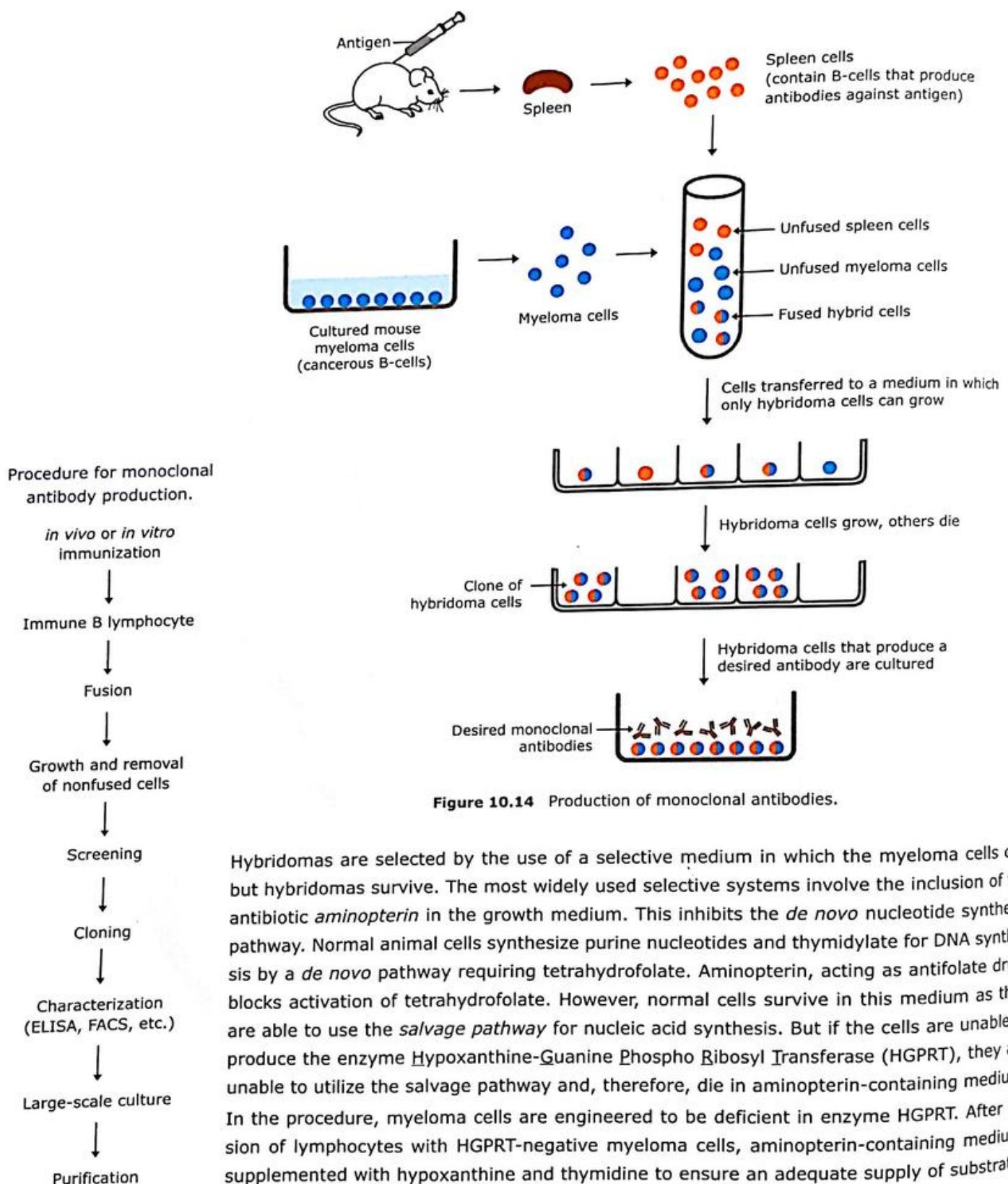
10.4 Monoclonal antibodies and Hybridoma technology

Antibodies produced ordinarily by infection or immunization are *polyclonal* because natural antigens have multiple epitopes, each of which generates clones of lymphocytes. This results in antisera containing antibodies from different clones of lymphocytes with specificities against different epitopes of the antigens. On the other hand, *monoclonal* antibodies (mAb) are monospecific antibodies. These antibodies are produced from clone of single lymphocyte directed against a single antigenic determinant or epitope. Such antibodies produced by a single clone and directed against a single epitope.

Hybridoma technology is a method of forming hybrid cell lines (called hybridomas) by fusing a specific antibody-producing B-cell with a myeloma cell (cancerous B-cell). The antibodies produced by the hybridoma are all of a single specificity and are therefore monoclonal antibodies. The production of monoclonal antibodies was invented by Cesar Milstein and Georges J. F. Köhler in 1975. They shared the Nobel Prize in 1984 for Medicine and Physiology.

Hybridomas are somatic cell hybrids produced by fusing antibodies forming spleen cells with myeloma cells. Antibody-producing B-cells normally die after several weeks in cell culture *in vitro*. Therefore, antibody-producing B-cells are fused with B-cell tumors called *myelomas*. These myelomas are capable of dividing indefinitely and are, therefore, often called immortal cell lines. The immortal cell lines that result from the B cell-myeloma fusion are hybrid cell lines called *hybridomas*. The hybridoma cell lines share the properties of both fusion partners.

They grow indefinitely *in vitro* and produce antibodies. To produce a monoclonal antibody, a mouse is immunized with the antigen of interest. During the next several weeks, antigen-specific B-cells proliferate and begin producing antibodies in the mouse. Spleen tissue, rich in B-cells, is then removed from the mouse, and the B-cells are fused with myeloma cells.



for the salvage pathway (HAT medium) is added, which kills myeloma cells but allows hybridomas to survive as they inherit HGPRT from the lymphocyte parent. Unfused lymphocytes die after a short period of culture, which results in a pure preparation of hybridomas.

References

- Abbas AK, Lichtman AH and Pillai S (2010), *Cellular and Molecular Immunology*, 6th ed. Saunders Elsevier.
- Berg JM, Tymoczko JL and Stryer L (2007), *Biochemistry*, 6th ed. W.H. Freeman and Company.
- Coico R, Sunshine G and Benjamini E (2003), *Immunology, A short course*, 5th ed. John Wiley and Sons, Inc.
- Decker J and Reischl U (2004), *Molecular diagnosis of infectious diseases*, 2nd ed. Humana Press.
- Freifelder D (1982), *Physical Biochemistry*, 2nd ed. W.H. Freeman and Company.
- Law B (2005), *Immunoassay: A practical guide*, Taylor and Francis.
- Owen J, Punt J and Stranford S (2013), *Kuby Immunology*, 7th ed. Macmillan Publishers.
- Paul WE (2008), *Fundamental immunology*, 6th ed. Philadelphia, Wolters Kluwer/Lippincott Williams & Wilkins, c2008.
- Roitt I, Brostoff J and Male D (2012), *Immunology*, 8th ed. Saunders Elsevier.
- Stanley J (2002), *Essentials of immunology and serology*, Cengage Learning.

Chapter 11

FRET and FRAP

11.1 FRET

FRET (Fluorescence Resonance Energy Transfer) is a phenomenon in which an excited donor molecule transfers energy (not an electron) to an acceptor molecule through a non-radiative process. It is a highly distance-dependent radiationless energy transfer process. In this energy transfer process, two fluorophores interact with each other in which one acts as donor and other as acceptor. The donor is a fluorophore that initially absorbs the energy and the acceptor is the fluorophore to which the energy is subsequently transferred. A pair of molecules that interact in such a manner that FRET occurs is often referred to as a donor-acceptor pair. There are some criteria that must be satisfied in order for FRET to occur. The primary condition is the distance between the donor and the acceptor. The donor and acceptor molecules must be in close proximity to one another for FRET to occur.

FRET relies on the distance dependent transfer of energy from the donor fluorophore to an acceptor fluorophore. In FRET, a donor fluorophore is excited by incident light, and if an acceptor is in close proximity, the excited-state energy from the donor can be transferred. The process is non-radiative (not mediated by a photon) and is achieved through dipole-dipole interactions. The donor molecule must have an emission spectrum that overlaps the absorption spectrum of the acceptor molecule.

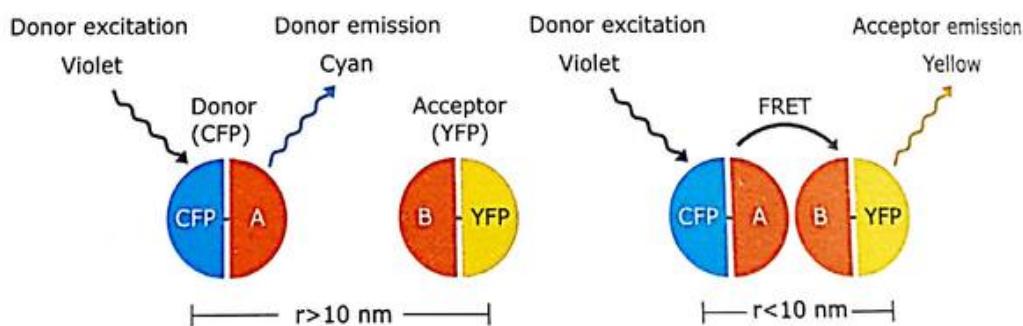


Figure 11.1 The acceptor and donor fluorophores must be very close for the FRET to occur. In this example, protein A is coupled to a CFP (Cyan fluorescent protein) and protein B is coupled to YFP (Yellow fluorescent protein).

The extent of energy transfer for a single donor-acceptor pair at a fixed distance depends on the inverse sixth power of the distance between the donor and acceptor pair (r) and is given by:

$$E = \frac{R_0^6}{R_0^6 + r^6}$$

Where, R_0 is the Förster radius at which half of the excitation energy of donor is transferred to the acceptor chromophore. Therefore, Förster radius is referred to as the distance at which the efficiency of energy transfer is 50%. Förster distances are typically in the range of 15 to 60 Å.

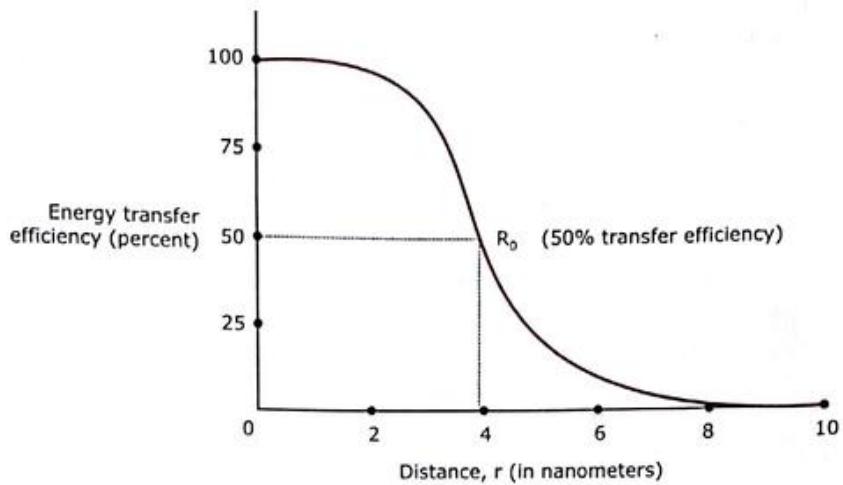


Figure 11.2 Distance and energy transfer efficiency. Energy transfer efficiency is most sensitive to distance between the donor-acceptor molecules. The graph illustrates the exponential relationship between transfer efficiency and the distance separating the donor and the acceptor. The efficiency rapidly increases to 100% as the separation distance decreases below R_0 and conversely, decreases to zero when r is greater than R_0 . Because of the strong (sixth-power) dependence of transfer efficiency on distance, measurements of the donor-acceptor separation distance are only reliable when the donor and acceptor radius lies within the Förster distance by a factor of two.

Secondly, for FRET to happen the absorption or excitation spectrum of the acceptor must overlap the fluorescence emission spectrum of the donor. The degree of overlap is referred to as *spectral overlap integral*.

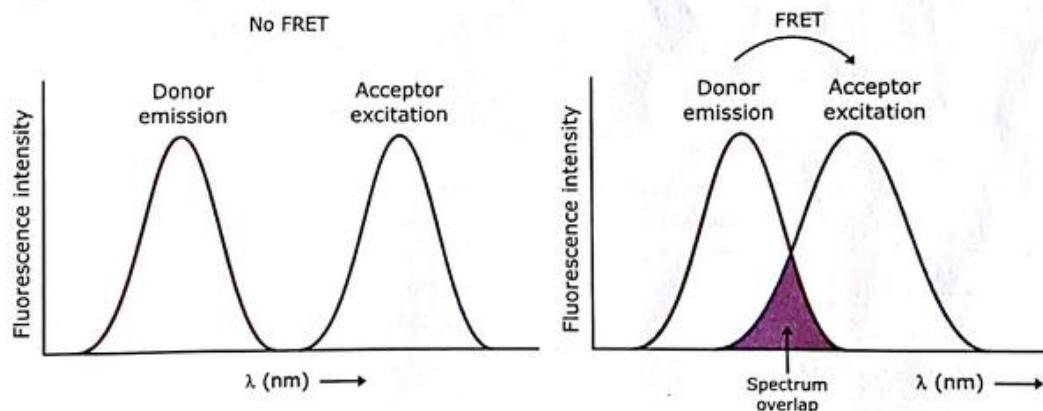


Figure 11.3 The emission spectrum of the donor overlaps with the excitation spectrum of the acceptor. The overlap of emission spectrum of the donor and absorption spectrum of the acceptor means that the energy lost from excited donor to ground state could excite the acceptor group. The energy matching is called the resonance phenomenon. Thus, the more overlap of spectra, the better a donor can transfer energy to the acceptor.

Table 11.1 Examples for common FRET Donor-Acceptor pairs:

Donor (Emission)	Acceptor (Excitation)
FITC (520 nm)	TRITC (550 nm)
Cy3 (566 nm)	Cy5 (649 nm)
EGFP (508 nm)	Cy3 (554 nm)
CFP (477 nm)	YFP (514 nm)
EGFP (508 nm)	YFP (514 nm)

FITC – Fluorescein isothiocyanate; TRITC – Isothiocyanate derivative of rhodamine; Cy3 – Cyanine

Application of FRET

Due to its sensitivity to distance, FRET has been used to investigate molecular interactions. It is used to find out protein-protein interactions inside living cells. To determine whether (and when) two proteins of interest interact inside a cell, they are first produced as fusion proteins attached to different fluorochromes. Two different fluorochromes are selected in such a manner that the emission spectrum of one fluorochrome overlaps the absorption spectrum of the second fluorochrome. If the two proteins and their attached fluorochromes come very close to each other, the energy of the absorbed light is transferred from one fluorochrome to the other. The energy transfer, called FRET, is determined by illuminating the first fluorochrome and measuring emission from the second.

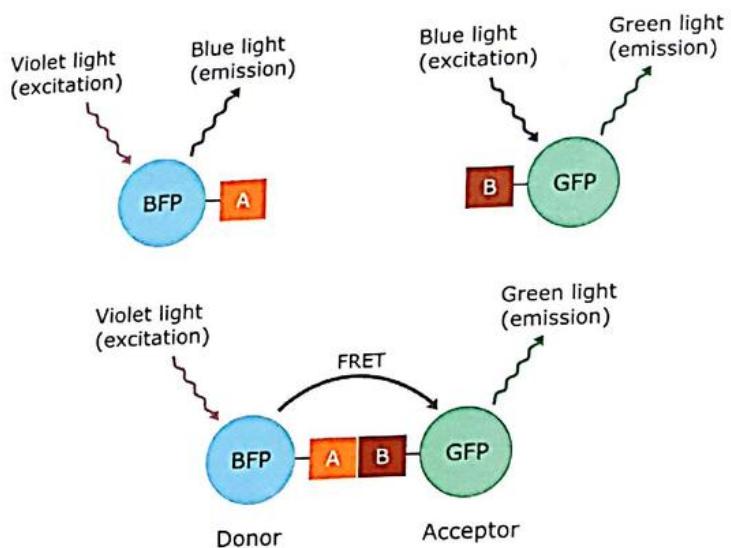


Figure 11.4 Schematic representation of the interaction of two different fusion proteins. In this example, protein A is coupled to a BFP, which is excited by violet light (370–440 nm) and emits blue light (440–480 nm); protein B is coupled to a GFP, which is excited by blue light and emits green light (510 nm). If proteins A and B do not interact, illuminating the sample with violet light yields fluorescence from the BFP only. When proteins A and B interact, FRET can occur. Illuminating the sample with violet light excites the BFP, whose emission, in turn, excites the GFP, resulting in an emission of green light.

FRET is also used to assess nucleic acid annealing. In the figure 11.5, two complementary RNA oligonucleotides are labeled with Cy3 and Cy5 respectively. When these labeled molecules are not annealed, excitation of an RNA oligonucleotide labeled with Cy3 with light at 540 nm results only in the emission of light by Cy3 at 590 nm, while the complementary RNA oligonucleotide labeled with Cy5 does not emit any light at 590 nm or 680 nm. However, when the two oligonucleotides are allowed to anneal, the close proximity of the molecules allows for FRET transfer to occur. This results in the emission of light at 680 nm; when the annealed molecule is excited with 540 nm light.

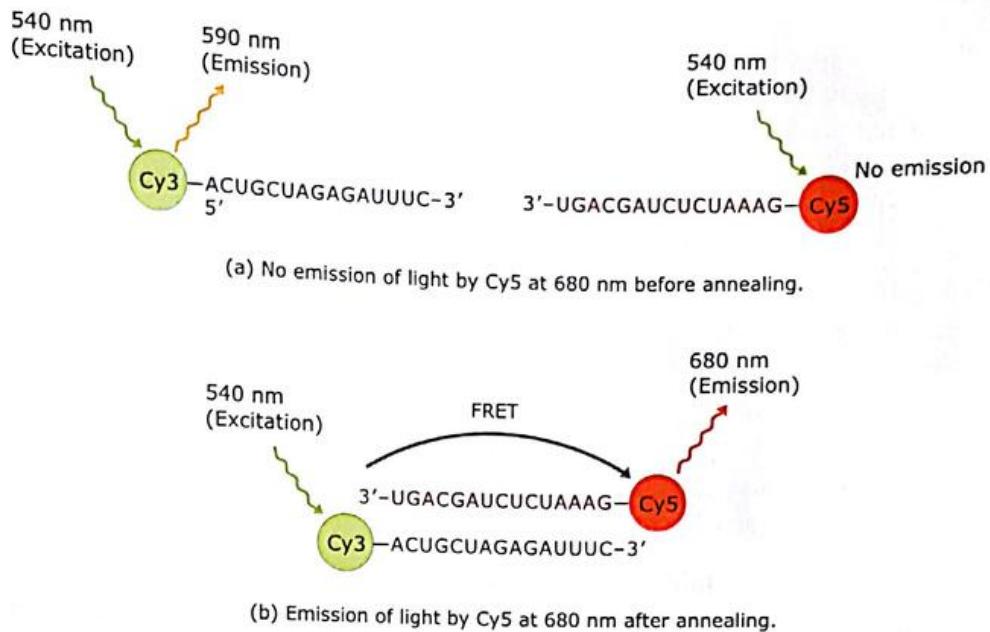


Figure 11.5 Schematic representation of FRET occurring between Cy3 and Cy5 fluorescent moieties when labeled oligonucleotides are annealed.

Green fluorescence protein, GFP

GFP (discovered by Osamu Shimomura) is an autofluorescent protein (238 amino acid residues) of bioluminescent jellyfish *Aequorea victoria*. In GFP, eleven β -strands make up the β -barrel and an α -helix runs through the center. The chromophore is located in the middle of the β -barrel. The chromophore is *p*-hydroxybenzylidene imidazolinone formed from the spontaneous cyclization and oxidation of the amino acid residues, Ser⁶⁵ (or Thr⁶⁵) - Tyr⁶⁶ - Gly⁶⁷. The chromophore of GFP is responsible for its fluorescence.

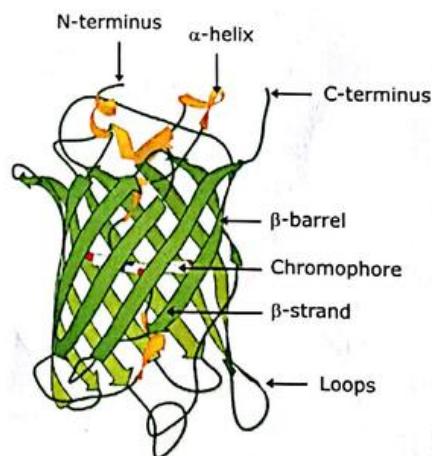


Figure 11.6 Structure of green fluorescence protein.

Wild type GFP from jellyfish has two excitation peaks, a major one at 395 nm and a minor one at 475 nm. Its emission peak is at 509 nm in the lower green portion of the visible spectrum. Several spectral variants of the original wild type green fluorescent protein (wtGFP) have been developed. Examples of these variants include: a blue fluorescent protein known as BFP; a cyan fluorescent variant known as CFP; a yellow fluorescent variant known as YFP;

a violet-excitable green fluorescent variant known as Sapphire and a cyan-excitable green fluorescing variant known as enhanced green fluorescent protein or EGFP.

Application

The most popular applications of GFP involve exploiting them for imaging of the localization and dynamics of specific organelles or recombinant proteins in live cells. It serves as a unique reporter and is used as a fusion tag for monitoring protein localization. GFP is used as a tag in a fusion protein, where it is coupled with the protein whose expression is to be tracked. In such cases, the principal aim is to investigate the subcellular localization of the protein under investigation. Genetic engineering can be used to produce vectors containing a GFP coding sequence into which a coding sequence for an uncharacterized protein, X, can be cloned. The resulting GFP-X fusion construct can be transfected into suitable target cells and expression of the GFP-X fusion protein can be monitored to track the subcellular location of the protein.

Jablonski diagram

The processes that occur between the absorption and emission of light are usually illustrated by the *Jablonski diagram*. A typical Jablonski diagram is shown in figure 11.7. The singlet ground, first and second electronic states are depicted by S_0 , S_1 and S_2 , respectively. A molecular electronic state in which all electron spins are paired is called a *singlet state*.

Each of the electronic states (ground or excited) has a number of vibrational energy levels, depicted by 0, 1, 2, etc. Following light absorption, a molecule is usually excited to higher vibrational level of either S_1 or S_2 . Molecules excited to electronic states S_1 and S_2 , rapidly lose any excess vibrational energy and relax to the ground vibrational level of that electronic state. This non-radiative process is termed *vibrational relaxation*.

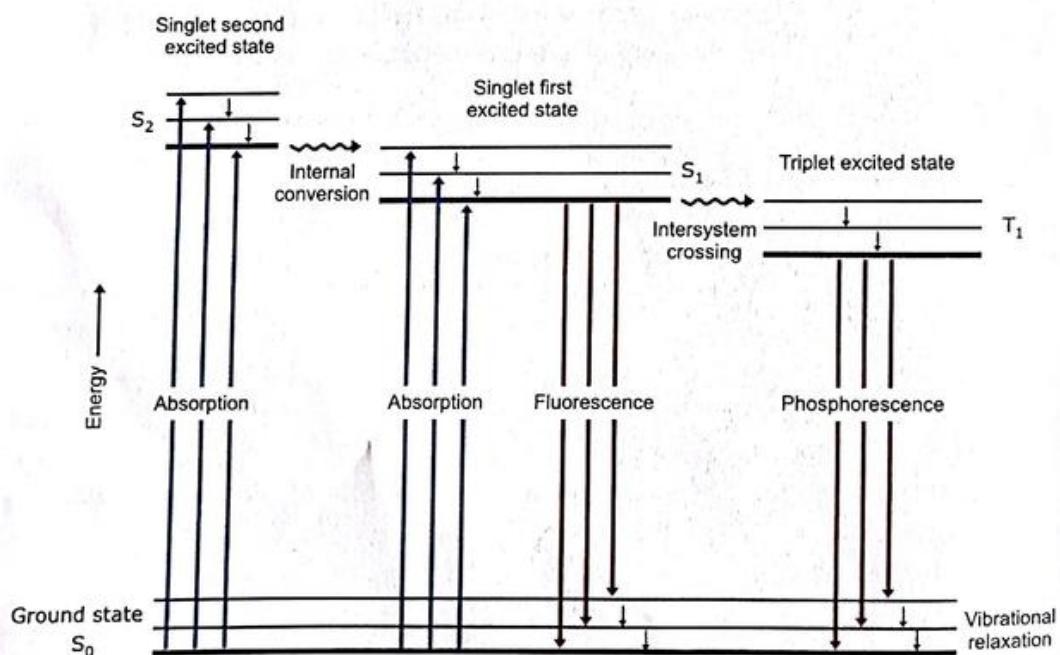


Figure 11.7 Jablonski diagram. The lowest thick horizontal line represents the ground state energy of the molecule, which is normally a singlet state and is labeled S_0 . The upper thick lines are energy levels for the ground vibrational states of three excited electronic states. The two lines on the left represent the first (S_1) and second (S_2) electronic singlet states. The one on the right (T_1) represents the energy of the first electronic triplet state. Numerous vibrational energy levels are associated with each of the four electronic states as represented by the thin horizontal lines.

With a few rare exceptions, molecules rapidly relax to the lowest vibrational level of S_1 . This process is called *internal conversion*. The term internal conversion describes intermolecular processes by which a molecule passes to a lower energy electronic state without emission of radiation. Once a molecule arrives at the lowest vibrational level of an excited singlet state, it can do a number of things, one of which is to return to the ground state by light emission. This process is called fluorescence.

Molecules in the S_1 state can also undergo a spin conversion to the first triplet state, T_1 . When one electron of a pair of electrons of a molecule is excited to a higher energy level, a *singlet* or a *triplet state* is formed. In the *excited singlet state*, the spin of the promoted electron is still paired with the ground state electron. In the triplet state, however, the spins of the two electrons have become unpaired and are thus parallel. Conversion of S_1 to T_1 is called *inter-system crossing*. It is a spin-dependent internal conversion process. Emission of light when electron comes from T_1 to the ground state is termed phosphorescence, and is generally shifted to longer wavelengths relative to the fluorescence.

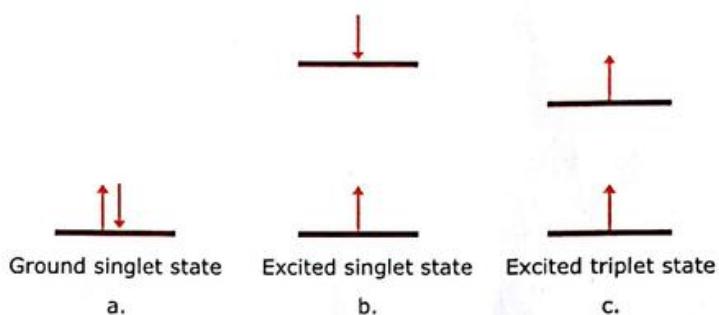


Figure 11.8 Electronic spin states of molecules. In (a), the ground electronic state is shown. In the ground state, the spins are always paired, and the state is said to be a *singlet state*. In (b) and (c), excited electronic states are shown. If the spins remain paired in the excited state, the molecule is in an *excited singlet state* (b). If the spins become unpaired, the molecule is in an *excited triplet state* (c).

11.2 FRAP

FRAP (Fluorescence Recovery After Photobleaching) is used to measure the dynamics of two or three dimensional movement of fluorescently labeled molecules within or between cells. The study of molecular mobility is an important parameter in the understanding of cell physiology. The principle of FRAP is to photobleach the fluorescently labeled molecules in a small region of the sample. Then the mobility of the fluorescently labeled molecules is evaluated from recovery of fluorescence due to exchange of fluorescently labeled molecules from surrounding unbleached area. In FRAP experiments, the photobleached area is restricted and as the techniques' name suggests, the recovery of fluorescence back into it, is monitored. Photobleaching is an irreversible process that involves the irradiation of the fluorophore with light. This results in the destruction of the fluorophore; and with it its ability to emit fluorescence. Recovery of the fluorescence signal is a result of the exchange of bleached fluorophores with those unbleached from the surrounding area. The fractions of molecules that are able and unable to participate in this exchange are termed the *mobile fraction* and *immobile fraction* respectively. In addition to this, observing the rate of fluorescent recovery can provide important understandings of the movement and interaction of intracellular molecules.

FRAP was developed by Axelrod and coworkers as a technique to study protein mobility in the membrane of living cells. However, with the advancement of technique, this technique became popular for studying protein mobility in the cell interior.

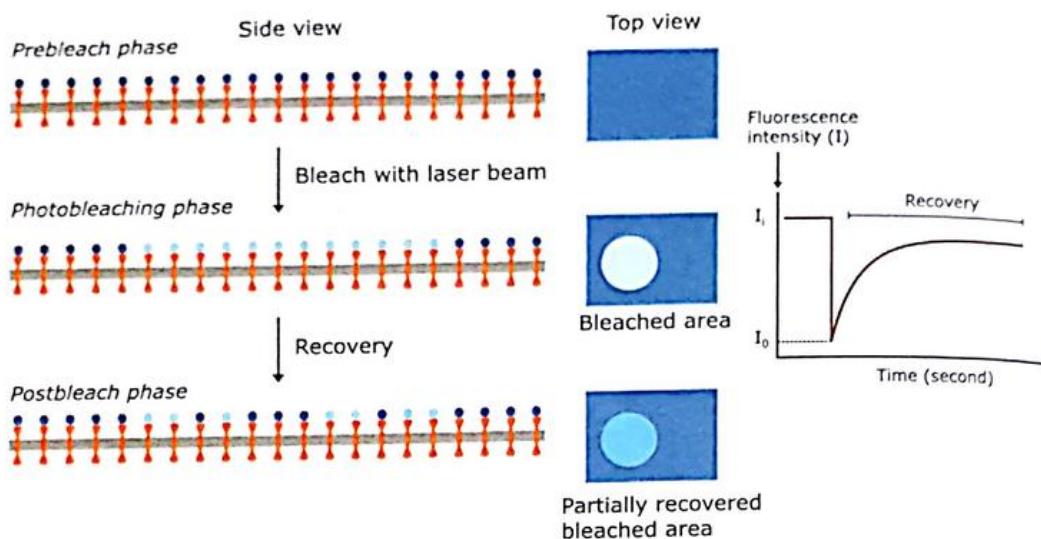
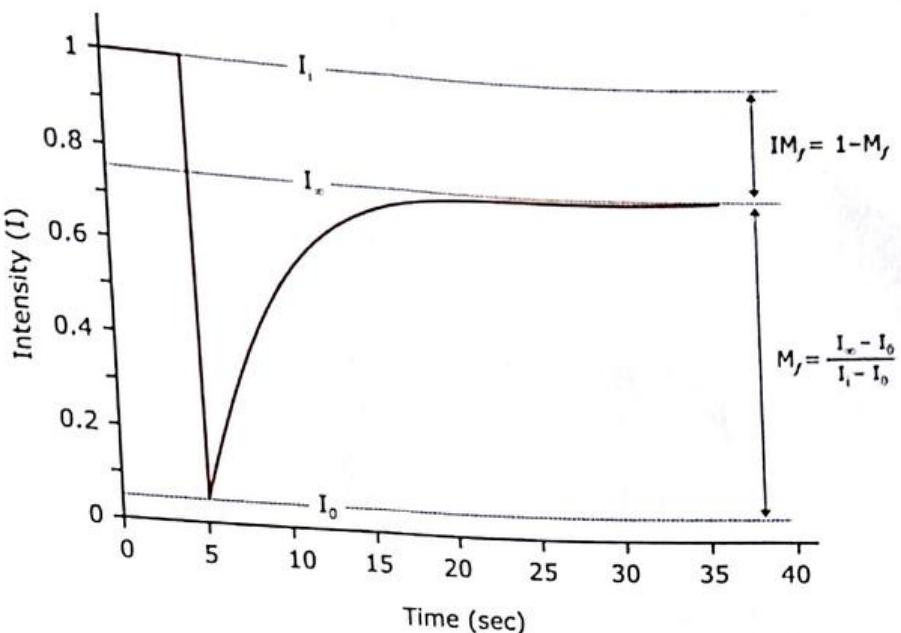


Figure 11.9 Measuring the rate of lateral diffusion of a membrane protein by FRAP technique. This technique is composed of three phases – prebleach phase, photobleaching phase and postbleach phase. A specific protein of interest can be labeled with a fluorescent molecule. Fluorescent molecules are bleached in a small area using a laser beam. The fluorescence intensity recovers as the bleached molecules diffuse away and unbleached molecules diffuse into the irradiated area. The diffusion coefficient is calculated from a graph of the rate of recovery: the greater the diffusion coefficient of the membrane protein, the faster the recovery. From this plot, the mobile and immobile fractions can be determined by calculating the ratios of the final to the initial fluorescence intensity. In the graph, I_0 is the fluorescence intensity immediately after the photobleaching, I_∞ the fluorescence intensity after full recovery and I_i the initial fluorescence intensity before photobleaching.

The lateral diffusion rates of membrane proteins can be measured by using the FRAP. The method usually involves marking the membrane protein of interest with a specific fluorescent group. The fluorescent group is then bleached in a small area by a laser beam. This bleached area gradually retrieves its fluorescent signal, not by a reversal of the bleaching effect, but by migration of fluorescent molecules into the bleached region from the unbleached area. The time taken for adjacent membrane proteins carrying unbleached fluorescent group to diffuse into the bleached area is measured. From such measurement, one can calculate the diffusion coefficient for the particular cell-surface protein that was marked.

Analysis of typical FRAP curve

From the initial (prebleach) fluorescence intensity (I_i), the signal drops to a particular low value (I_0) as the high intensity laser beam bleaches fluorochromes in the region of interest. Over time, the signal recovers from the post-bleach intensity (I_0) to a maximal plateau value (I_∞). From this plot, the mobile fraction (M_f) and immobile fraction (IM_f) can be calculated. The information from the recovery curve (from I_0 to I_∞) can be used to determine the diffusion constant and the binding dynamics of fluorescently labeled proteins.



A complementary technique is fluorescence loss in photobleaching (FLIP). FLIP experiments differ from FRAP by the repetitive bleaching of the same region in the specimen, thereby preventing recovery of fluorescence in that region. Here, a laser beam continuously irradiates a small area to bleach all the fluorescent molecules that diffuse into it, thereby gradually depleting the surrounding membrane of fluorescently labeled molecules.

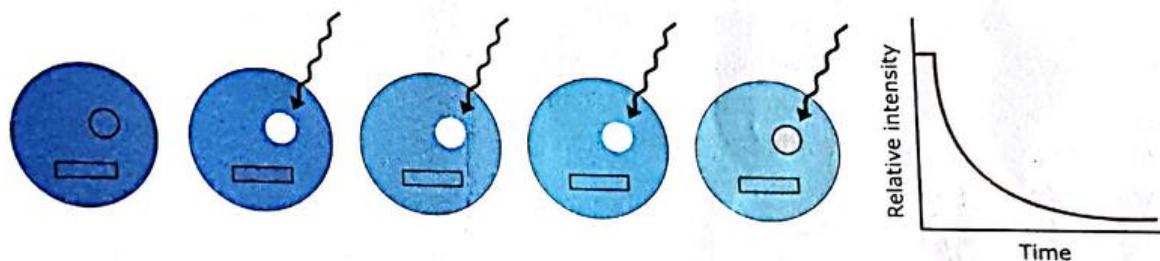


Figure 11.10 FLIP experiments involve repetitive bleaching of a selected region of interest during the entire monitoring period and the fluorescence intensity in regions outside the selected bleached area is measured. The decline in fluorescence intensity in the surrounding regions is due to bleaching of fluorochromes that move through the region of interest during the repetitive bleaching process. The drop in fluorescence intensity outside the bleached region is caused by a steadily increasing population of bleached, non-fluorescent molecules within the cell and thus provides quantitative data on their molecular mobility.

References

- Alberts B, et al (2008), *Molecular Biology of the Cell*, 5th ed. Garland Science Publishing.
- Chang R (2005), *Physical Chemistry for the Biosciences*, 1st ed. University Science Books.
- Lakowicz JR (1999), *Principles of Fluorescence Spectroscopy*, 2nd Ed. Kluwer, New York.
- Tsien RY (1998), The green fluorescent protein. *Annu. Rev. Biochem.* 67, 509–544.
- van Royen ME, Dinant C, Farla P, Trapman J and Houtsmailler AB (2009), FRAP and FRET methods to study nuclear receptors in living cells. *Methods Mol. Biol.* 505, 69–96.
- Zimmer M (2002), Green fluorescent protein (GFP): Applications, structure, and related photophysical behavior. *Chem. Rev.* 102, 759–781.

Chapter 12

Molecular Biology Techniques

12.1 Polymerase chain reaction

Polymerase chain reaction (PCR) is a rapid and versatile *in vitro* method for amplifying defined target DNA sequences present within the source of DNA. This technique was formulated by Kary Mullis in 1985. Usually, the method is designed to permit *selective amplification* of a specific target DNA sequence(s) within a heterogeneous collection of DNA molecules (e.g. total genomic DNA or a complex cDNA population). To permit such selective amplification, some prior DNA sequence information from the target sequences is required. This information is used to design two oligonucleotide primers (amplifiers), which are specific to the target sequence and are often about 15–25 nucleotides long. After the primers are added to denatured template DNA, they bind specifically to complementary DNA sequences at the target site. In the presence of a suitably heat-stable DNA polymerase and DNA precursors (the four deoxynucleoside triphosphates, dATP, dCTP, dGTP and dTTP), primer initiates the synthesis of new DNA strands which are complementary to the individual DNA strands of the target DNA segment, and will overlap each other.

Primer design

In PCR, primer designing is the most important aspect for selective amplification. For primer design, some prior DNA sequence information from the target DNA is required. The information is used to design two primers (amplifiers), which are specific to sequences flanking the target DNA sequence. So, for most PCR reactions, it is very important to reduce the chance of the primers binding to other locations in the DNA than the desired one. Hence, certain rules for primer design are important to consider. These rules follow:

Primer length

Length of the primers should not be very short or long. If the primers are too short, they might hybridize to non-target sites and give undesired amplification products.

To illustrate this point, imagine a DNA molecule of 10 Mbp is used in a PCR experiment with a pair of primers eight nucleotides in length. The attachment sites for these primers are expected to occur, on average, once every $4^8 = 65,536$ bp. This means that it would be very unlikely that a pair of 8 nucleotides long primers would give a single, specific amplification product. If the primer length will be 15 nucleotides, then the expected frequency of attachment site for the primer will be once every $4^{15} = 1073741824$ bp. This figure is higher than

the length of DNA molecule, so 15 nucleotides long primer would be expected to have just one attachment site. But we cannot take very long primer because the long primer influences the rate at which it hybridizes to the template DNA; long primers hybridize at a slower rate. It is generally accepted that the optimal length of primers is 18-20 nucleotides. This length is long enough for adequate specificity and short enough for primers to bind easily to the template at the annealing temperature.

Calculation of length of primers: The longer the primer, the higher its specificity for a particular target. The following equation can be used to calculate the probability that a sequence exactly complementary to the primer sequence will occur by chance in a DNA molecule that consists of a random sequence of nucleotides.

$$K = [g/2]^{G+C} \times [(1-g)/2]^{A+T}$$

where, K is the expected frequency of occurrence of target sequence in a DNA molecule,

g is the relative G + C content of the DNA molecule and

G, C, A and T are the number of specific nucleotides in the primer.

For a double-stranded genome of size N (in nucleotides), the expected number (n) of sites complementary to the primer is $n = 2NK$.

Nature of sequences

Inverted repeats or any self-complementary sequences >3bp in length should not be present. Sequence of this type tends to form hairpin structure and prevents the primer annealing to its target. There should be no complementarity between the two primers. Because primers are present at high concentration in PCR, even small complementarity between them lead to hybrid formation.

T_m values 25 - 95

T_m (melting temperature) is the temperature at which half of the DNA strands are in double stranded state. T_m values for two primers used together should not differ by $>5^\circ\text{C}$ and the T_m of the amplification product should not differ from those of the primers by $>10^\circ\text{C}$.

Base composition

The GC content should be between 40 to 60% with an even distribution of all four nucleotides.

Degenerate primers

In order to make the PCR primers, some sequence information is required. Degenerate primers are used when partial sequence information is available, but the complete sequence is unknown. A degenerate primer is a mixture of primers, all of similar sequence but with variations at one or more positions. Degenerate DNA primers are generally used if only a protein sequence is available. In this case, the protein sequence is translated backwards to give the corresponding DNA sequence. Due to the degeneracy of the genetic code, several possibilities will exist for the sequence of DNA that corresponds to any particular polypeptide sequence. Again, most of the ambiguity is in the third codon position. This ambiguous sequence may be used to make degenerate primers.

Partial sequence of polypeptide:

Met — Tyr — Cys — Asn — Thr — Arg — Pro — Gly

Possible codons in DNA:

ATG	TAC	TGT	AAT	ACT	AGA	GCT	GGT
TAT	TGC	AAC	ACC	AGG	GCC	GGC	
			ACA		GCA	GGA	
			ACG		GCG	GGG	

Corresponding degenerate primer:

ATG	TAC	TGT	AAT	ACT	AGA	GCT	GGT
T	C	C	C	G	C	C	
			A		A	A	
			G		G	G	

Bases in the third codon position are shown in red. The degenerate primer consists of a mixture of primers with these bases varied as shown.

Figure 12.1 Degenerate DNA primers are used if only partial DNA sequence information is available. Often, as here, a short amino acid sequence from a protein is known. Because many amino acids are encoded by several alternative codons, the deduced DNA coding sequence is ambiguous. For example, the amino acid tyrosine is encoded by TAC or TAT. Hence, the third base is ambiguous and when the primer is synthesized, a 50:50 mixture of C and T will be inserted at this position. This ambiguity occurs for all the bases shown in bold letters, resulting in a pool of primers with different, but related sequences. Hopefully, one of these primers will have enough complementary bases to anneal to the target sequence that is to be amplified.

Reaction cycle

The PCR is a chain reaction because newly synthesized DNA strands will act as templates for further DNA synthesis in subsequent cycles. It consists of a series of cycles of three successive reactions:

- Denaturation, typically at about 93–95°C for human genomic DNA.
- Primer annealing at temperatures usually from about 50°C to 70°C depending on the T_m of the expected duplex. Annealing temperature must be low enough to enable hybridization between primer and template, but high enough to prevent mismatched hybrids from forming. The annealing temperature is typically about 5°C below the calculated T_m .

Several equations are used to calculate the T_m value of duplex formed between an oligonucleotide primer and its complementary target sequence. One common equation, known as *Wallace rule*, can be used to calculate the T_m for perfect duplex 15–20 nucleotides in length in solvents of high ionic strength (e.g. 1M NaCl):

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

A + T is the sum of A and T residues and G + C is the sum of G and C residues in the oligonucleotide primer.

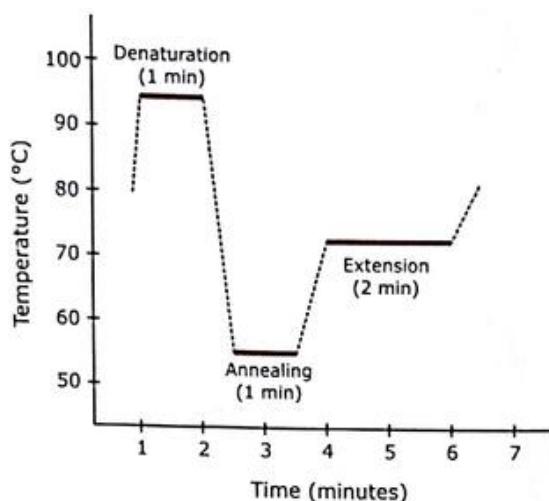
- DNA synthesis, typically at about 70–75°C.

These three steps constitute one cycle of the PCR amplification and can be carried out repetitively just by changing the temperature of the reaction mixture. The thermostability of the polymerase makes it feasible to carry out PCR. Suitable heat-stable DNA polymerases have been obtained from microorganisms whose natural habitat is hot springs.

Essential components of PCRs:

- Thermostable DNA pol
- Primers
- dNTP
- Divalent cations (usually Mg²⁺)
- Buffer (to maintain pH)
- Template DNA

Figure 12.2 A typical temperature profile for a PCR. The denaturation temperature is usually 94°C, which denatures the dsDNA and releases single-stranded DNA to act as templates in the next round of DNA synthesis. The annealing temperature, at which the primers attach to the templates, is dependent on the nucleotide composition of the primers. The extension temperature, at which the bulk of DNA synthesis occurs, is usually set at 74°C, optimum for Taq polymerase.



For example, the widely used *Taq* DNA polymerase is obtained from *Thermus aquaticus* and is thermostable up to 94°C, with an optimum working temperature of 75–80°C. It has an extension rate of 35 to 100 nucleotides per second at 72°C. *Taq* polymerase lacks a 3'-5' exonuclease (proofreading) activity. When the *Taq* polymerase incorporates a wrong dNTP, subsequent extension of the strand either proceeds very slowly or stops completely. This problem can be overcome by using other thermostable DNA polymerases with 3'-5' proof-reading exonuclease activity.

DNA Pol	5' to 3' exonuclease	3' to 5' exonuclease	Source
<i>Tli</i> (or <i>Vent</i>)	No	Yes	<i>Thermococcus litoralis</i>
<i>Pfu</i>	?	Yes	<i>Pyrococcus furiosus</i>
<i>Pwo</i>	No	Yes	<i>Pyrococcus woesei</i>
<i>Tth</i>	?	No	<i>Thermus thermophilus</i>

After each cycle, the number of templates doubles, so that if one starts with a single dsDNA molecule, after 20 cycles, the number of molecules synthesized by the PCR is 1×10^6 , and after 30 cycles the number increases to 1×10^9 . This number can be calculated by applying the following formula:

$$N_f = N_i \times 2^n$$

where, N_f is the final number of DNA molecules produced by the PCR,
 N_i is the initial number of molecules (template), and
 n is the number of cycles performed.

When amplifying a small segment of a long double stranded DNA template, the desired blunt-ended target fragments first appear in the third cycle of PCR. The number of blunt-ended target products synthesized after 25 cycles of PCR with different numbers of starting molecule (assuming that amplification is 100% efficient) is given below:

Number of starting molecules	Number of blunt-ended target products
1	4,194,304
2	8,388,608
5	20,971,520
10	41,943,040

In theory, each amplification cycle should double the number of target molecules, resulting in an exponential increase in PCR product. However, even before substrate or enzyme becomes limiting, the efficiency of exponential amplification is less than 100% due to sub-optimal DNA polymerase activity, poor primer annealing and incomplete denaturation of the templates. So, the actual amount of PCR product can be calculated by considering efficiency term. This PCR efficiency formula can be expressed as:

$$\text{PCR product} = (\text{Initial input amount}) \times (1 + \% \text{ efficiency})^{\text{cycle number}}$$

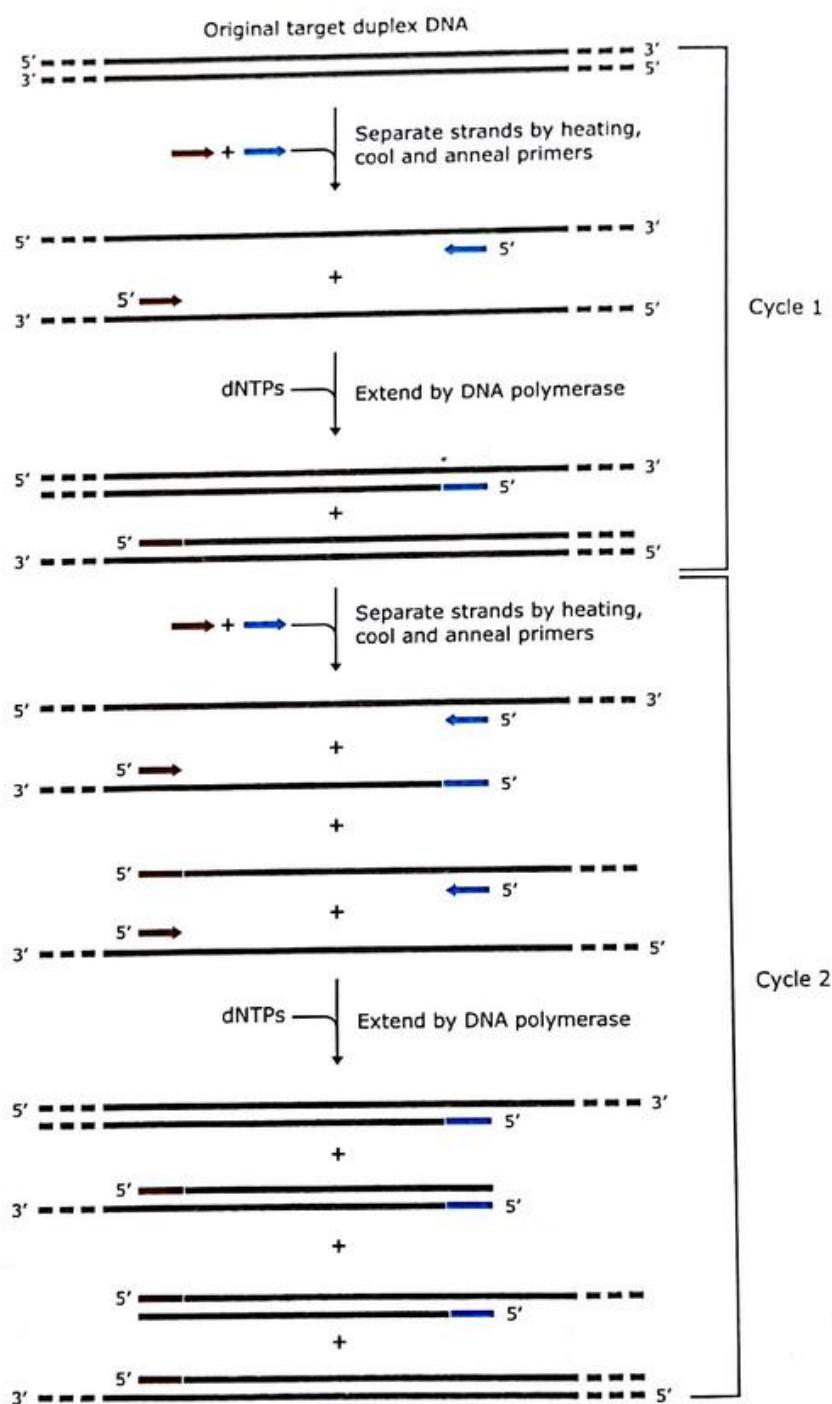


Figure 12.3 The polymerase chain reaction (PCR). In each cycle of the reaction, the strands of the duplex DNA are separated by heat denaturation, the preparation is cooled to such level that synthetic DNA primers anneal to a complementary segment on each strand, and the primers are extended by DNA polymerase. The process is then repeated for numerous cycles. The number of "unit-length" strands doubles with every cycle after the second cycle.

Modifications to the PCR techniques

Nested PCR

Nested PCR increases the specificity of DNA amplification, by reducing non-specific amplification of DNA. Two sets of primers are being used in two successive PCR reactions. In the first reaction, one pair of primers is used to generate DNA products, which besides the intended target, may still consist of non-specifically amplified DNA fragments. The product(s) is, then, used in a second PCR reaction with a set of primers whose binding sites are just downstream of the first primer, or *nested* between the original set of primers. Binding sites are completely or partially different from the primer pair used in the first reaction, but are completely within the DNA target fragment.

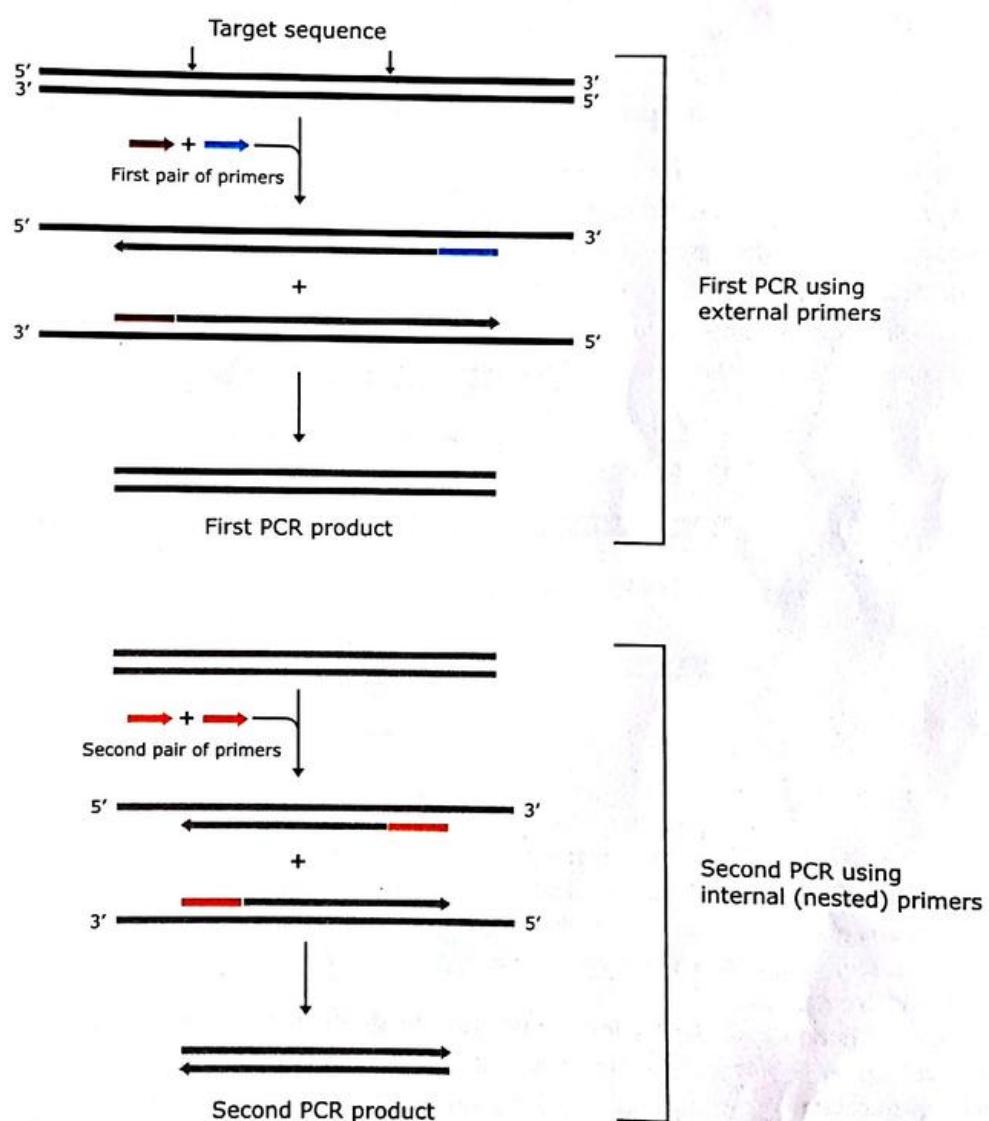


Figure 12.4 Nested PCR. It involves using two sets of primers. The first external set generates a normal PCR product. Primers that lie inside the first set are, then, used for a second PCR reaction. These internal or nested primers generate a shorter product. It is used to increase the specificity and fidelity of the PCR.

Quantitative Real Time PCR

Quantitative Real Time PCR is based on the general principle of PCR, which is used to amplify and simultaneously quantify a target DNA molecule. This is called Real Time PCR because it allows the scientist to actually view the increase in the amount of DNA as it is amplified. Real Time PCR systems rely upon the detection and quantitation of a fluorescent reporter, whose signal increases in direct proportion to the amount of PCR product in a reaction. These fluorescent reporter molecules include dyes that bind to the dsDNA (i.e. SYBR Green) or sequence specific probes. SYBR Green binds to the minor groove of the dsDNA only. In the solution, the unbound dye exhibits very little fluorescence. This fluorescence is substantially enhanced when the dye is bound to dsDNA. SYBR Green has, however, a limitation that includes preferential binding to G.C rich sequences.

SYBR Green monitors the total amount of double-stranded DNA, but cannot distinguish between different sequences. To be sure that the correct target sequence is being amplified, a sequence-specific fluorescent probe is needed. An example is the **TaqMan probe** which is oligonucleotide designed to hybridize to an internal region of a PCR product.

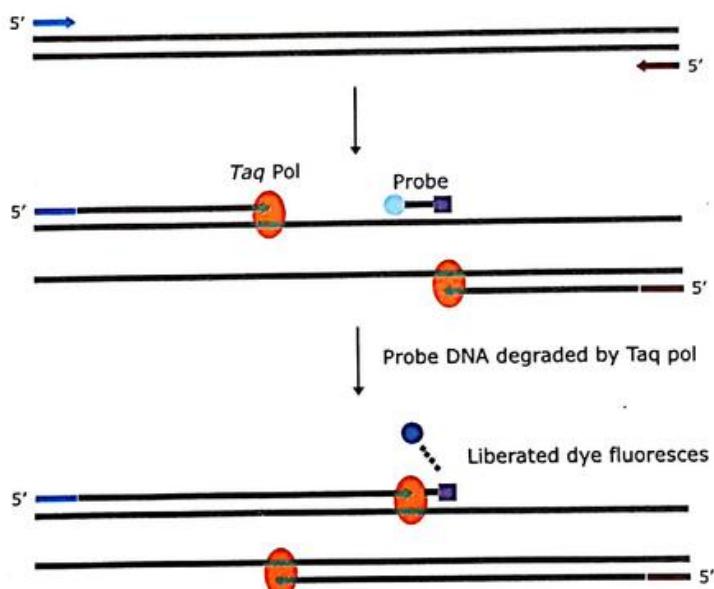


Figure 12.5 Real-time PCR with TaqMan probe. The TaqMan probe binds to the target sequence after the denaturation step that separates the two DNA strands. As the *Taq* polymerase extends the primer during the next PCR cycle, it will eventually bump into the TaqMan probe. The *Taq* polymerase has a 5' to 3' exonuclease activity that degrades the DNA strand of the probe. This breaks the linkage between the fluorophore and quencher. The fluorophore is now free from quenching and its fluorescence increases.

Quenching is the decrease in intensity of fluorescence. It can occur by different mechanisms. This is an example of collisional quenching which occurs when the excited-state fluorophore is deactivated upon contact with some other molecule in solution, which is called the quencher.

The TaqMan probe has three elements: a short-wavelength fluorophore on one end, a sequence that is specific for the target DNA, and a quencher at the other end. As long as the fluorophore and the quencher are close to each other, fluorescence is quenched and no fluorescent light is emitted. This probe is designed to anneal to the center of the target DNA. When *Taq* polymerase elongates the second complementary strand during PCR, its 5' to 3' exonuclease activity cuts the probe into single nucleotides. This removes the close proximity between the fluorophore and the quencher and abolishes quenching. The short-wavelength fluorophore can now fluoresce and a signal will be detected that is proportional to the number of newly synthesized strands.

RT-PCR

RT-PCR (Reverse Transcription PCR) is a method used to amplify, isolate or identify a known sequence from a cellular or tissue RNA. The PCR reaction is preceded by a reaction using reverse transcriptase to convert RNA to cDNA. Some thermostable DNA polymerase can use RNA templates as substrate. An example of this is the recombinant form of *Tth polymerase* from *Thermus thermophilus*, which can catalyze high-temperature reverse transcription of RNA in the presence of MnCl₂. It has both intrinsic reverse transcriptase and thermostable DNA-dependent DNA polymerase activities. This has led to the development of protocols for single-enzyme reverse transcription and PCR amplification.

Inverse PCR

Standard PCR is used to amplify a segment of DNA that lies between two inward-pointing primers. In contrast, inverse PCR (also known as inverted or inside-out PCR) is used to amplify unknown DNA sequences that flank one end of a known DNA sequence and for which no primers are available. The inverse PCR method involves a series of restriction digestion and ligation, resulting in a circular DNA that can be primed for PCR from a single section of known sequence. It involves isolating a restriction fragment that contains the known sequence plus flanking sequences. The restriction fragments circularize to form circular DNA under very low concentrations in presence of DNA ligase. To perform PCR, two primers that bind specifically to the known sequence but they are oriented in opposite directions. Successful PCR with these primers produces a linear product in which central unknown region remain flanked by two short known sequences.

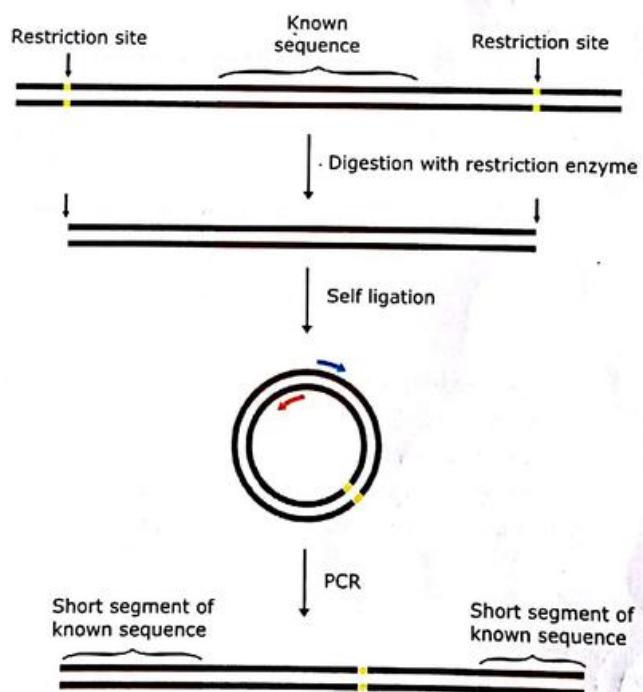


Figure 12.6 Inverse PCR. A region of DNA in which part of the sequence is known. If the areas of interest lie outside the known region, inverse PCR can be used to amplify the flanking region. It involves the digestion of genomic DNA with appropriate restriction endonucleases, intramolecular ligation to circularize the DNA fragments and PCR amplification. First, the DNA is cut with a restriction enzyme that does not cut within the region of known sequence. This generates a fragment of DNA containing the known sequence flanked by two regions of unknown sequence. Under low DNA concentrations, self-ligation is induced to give a circular DNA product. Finally, PCR is performed on the circular fragments of DNA. Two primers are used that face outwards from the known DNA sequence. PCR amplification gives linear products.

Anchored PCR

In the basic PCR technique and the inverse PCR, one has to use two primers representing the sequences lying at both ends of sequence to be amplified. But sometimes, we may have knowledge about the sequence at only one of the two ends of the DNA sequence to be amplified. In such cases anchored PCR may be used, which will utilize only one primer instead of two primers. In this technique, due to the use of one primer, only one strand will be copied first, after which a poly G tail will be attached at the end of the newly synthesized strand. This newly synthesized strand with poly G tail at its 3'-end will, then, become the template for the daughter strand synthesis utilizing an anchor primer with which a poly C sequence is linked to complement with poly G of the template. In the next cycle, both the original primer and anchored primer will be used for gene amplification.

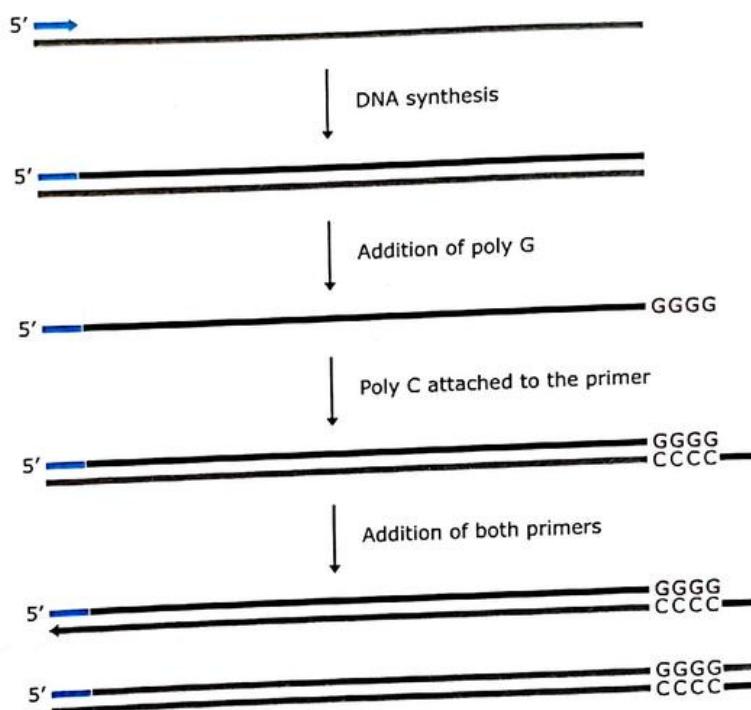


Figure 12.7 Anchored PCR. In this technique, due to the use of one primer, only one strand will be copied first, after which a poly G will be attached at the end of the newly synthesized strand. This newly synthesized strand with poly G tail at its 3'-end will, then, become the template for the daughter strand synthesis utilizing an anchor primer with which a poly C sequence is linked to complement with poly G of the template. In the next cycle, both the original primer and anchored primer will be used for gene amplification.

RACE

RACE (Rapid Amplification of cDNA Ends) is a PCR-based method for locating the precise start and end points of gene transcripts. It is of two types- 5'-RACE and 3'-RACE.

In **5'-RACE**, the first step is the conversion of mRNA into cDNA with enzyme reverse transcriptase. The primer used during reverse transcription is specific for an internal region (gene specific primer) close to 5'-end of the gene under study. Since only a small segment of mRNA is copied, the cDNA so generated will correspond exactly with the start of the mRNA. After synthesis of above cDNA, a short poly(A) tail is added to its 3'-end using enzyme terminal deoxynucleotidyl transferase and it is subjected to normal PCR. The second primer will

anneal to this poly(A) sequence and convert the single stranded cDNA into a double stranded product will reveal the precise position of the start or 5'-end of the transcript. This is known as 5'-RACE, because it results in amplification of the 5'-end of the starting RNA.

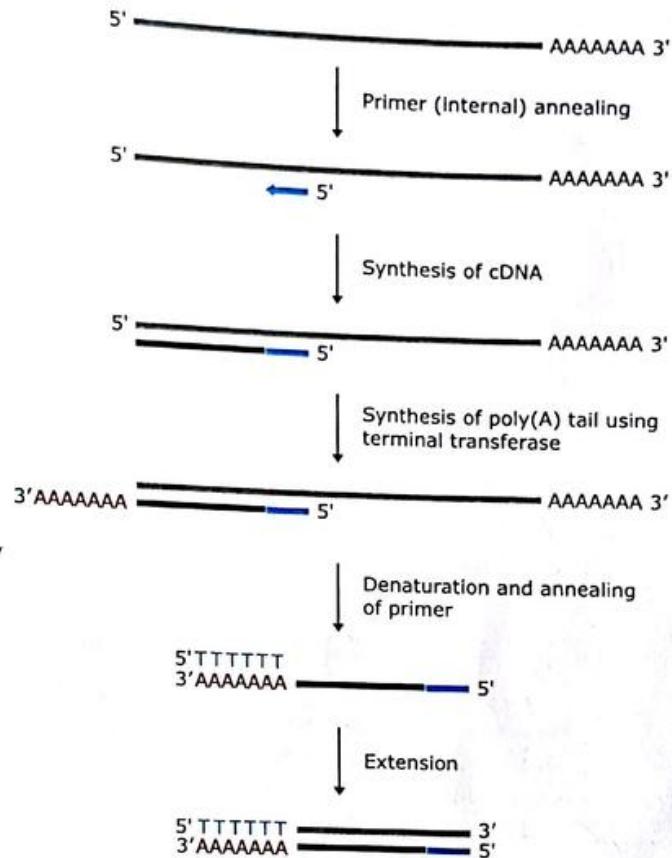


Figure 12.8 5'-RACE PCR begins using mRNA as a template for a first round of cDNA synthesis (or reverse transcription) reaction using an antisense primer that recognizes a known sequence in the gene of interest; the primer is called a gene specific primer, and it copies the mRNA template in the 3' to the 5' direction to generate a specific single-stranded cDNA product. Following cDNA synthesis, the enzyme terminal deoxynucleotidyl transferase is used to add homopolymer poly(A) tail to the 3'-end of the cDNA. A PCR reaction is then carried out, by using primers that hybridize to the 3' poly(A) tail of the cDNA and convert single-stranded cDNA into a double stranded molecule.

3'-RACE like 5'-RACE, requires knowledge of a sequence within the target RNA. A population of mRNAs is transcribed into cDNA with an adapter-primer consisting oligo(dT) primer linked with adapter sequence. It uses the natural polyA tail that exists at the 3'-end of all eukaryotic mRNAs for primer formation during reverse transcription. After first strand cDNA synthesis, the original mRNA template is destroyed with RNase H, which is specific for RNA-DNA hybrid molecules. Since the internal sequence is known, an internal primer (*gene specific primer*) is used for synthesis of second strand. Further, the same internal primer and a primer corresponding to the adapter sequence are used in a standard PCR reaction to amplify just the 3'-end of the cDNA.

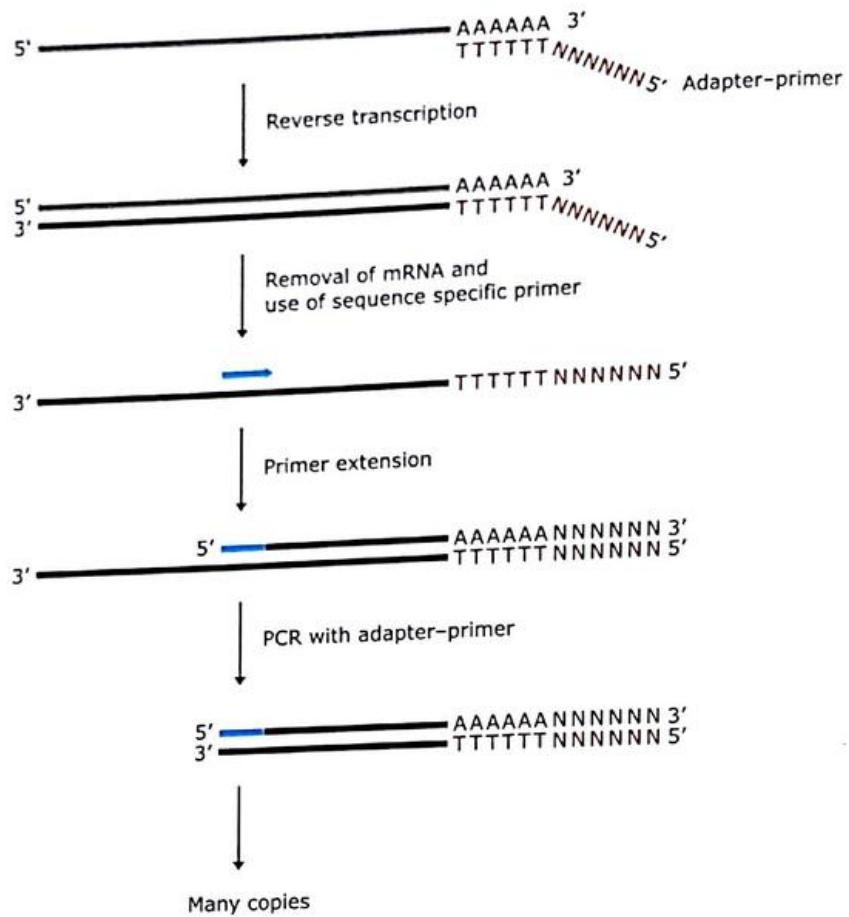


Figure 12.9 3'-RACE. This requires an oligo(dT) primer that has an adapter sequence at the 5'-end i.e. oligo(dT) adapter-primer. This primer is used with reverse transcriptase to make the mRNA. The mRNA is removed from resultant DNA-RNA hybrid molecule. After removal of mRNA, a second strand of DNA is synthesized. The second strand is synthesized by using internal sequence specific primer. The same internal primer and a primer corresponding to the adapter sequence are then used in a standard PCR reaction to amplify just the 3'-end of the cDNA.

Touchdown PCR

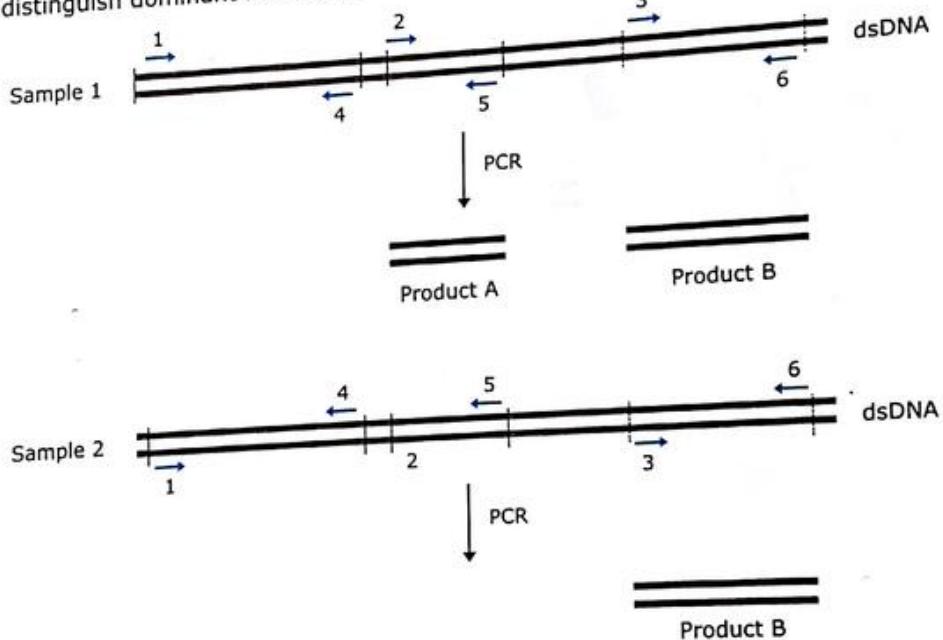
The specificity of conventional PCRs decreases as the sequence complexity of the template DNA increases. The greater the complexity, the greater is the chance that the primers will bind promiscuously to sequences other than the intended target. The number of mispriming events can be reduced by optimizing the concentrations of the components of the PCR, in particular, the concentrations of Mg^{2+} , primers, dNTPs and template. Further minimization of off-target amplification can be obtained by using stringent annealing temperature.

Touchdown PCR is a method used to increase specificity without compromising the yield. The principle is to initiate synthesis at very high annealing temperatures which permit only perfectly matched primer-template hybrids to form. The annealing temperature is dropped in a stepwise fashion with each cycle (1-2°C/every second cycle). Once copies of the target sequence have begun to accumulate over the first few cycles, high temperature annealing becomes much less critical for specificity, as it's the previous products which form the major template. These products are unlikely to have any sites for mispriming. The benefit of decreasing the annealing temperature is to increase the probability of stable primer-target interaction.

RAPD

RAPD stands for Random Amplification of Polymorphic DNA. It is a type of PCR, but the segments of DNA that are amplified are random. Unlike traditional PCR analysis, RAPD does not require any specific knowledge of the DNA sequence of the target organism. In RAPD, several arbitrary, short primers (8–12 nucleotides) and a large template of genomic DNA are used. No knowledge of the DNA sequence for the targeted gene is required, as the primers will bind somewhere in the sequence, but it is not certain exactly where. This makes the method popular for comparing the DNA of biological systems that have not had the attention of the scientific community, or in a system in which relatively few DNA sequences are compared. In recent years, RAPD has been used to characterize and trace, the phylogeny of diverse plant and animal species.

RAPD involves amplification of DNA fragments from any species by use of a single arbitrary oligonucleotide primer without prior sequence information. As the approach requires no prior knowledge of the genome that is being analyzed, it can be employed across species using universal primers. The major limitation of the RAPD method is the reproducibility and dominant inheritance. Several factors influence the reproducibility of RAPD reactions such as quality and quantity of template DNA, PCR buffer, concentration of magnesium chloride, primer to template ratio and annealing temperature. RAPD markers are dominant markers, and hence do not distinguish dominant homozygotes from heterozygotes.



Unlike traditional PCR analysis, RAPD (pronounced 'rapid') does not require any specific knowledge of the DNA sequence of the target organism: the identical 10-mer primers will or will not amplify a segment of DNA, depending on positions that are complementary to the primers' sequence. For example, no fragment is produced if primers annealed too far apart or 3'-ends of the primers are not facing each other. Therefore, if a mutation has occurred in the template DNA at the site that was previously complementary to the primer, a PCR product will not be produced, resulting in a different pattern of amplified DNA segments on the gel.

Figure 12.10 Schematic drawing of reaction conditions for RAPD. The primers must anneal in a particular orientation (such that they point towards each other) and within a reasonable distance of one another. The arrows represent multiple copies of a single primer and the direction of the arrow indicates the direction in which DNA synthesis will occur. The numbers represent primer annealing sites on the DNA template. For sample 1, primers anneal to sites 1, 2, and 3 on the top strand of the DNA template and to sites 4, 5, and 6 on the bottom strand of the DNA template. In this example, only 2 RAPD products are formed. For sample 1: (i) product A is produced by PCR amplification of the DNA sequence which lies in between the primers bound at positions 2 and 5; and (ii) product B is produced by PCR amplification of the DNA sequence which lies in between the primers bound at positions 3 and 6. No PCR product is produced by the primers bound at positions 1 and 4 because these primers are too far apart to allow completion of the PCR reaction. No PCR products are also produced by the primers bound at positions 4 and 2 or positions 5 and 3 because these primer pairs are not oriented towards each other. For sample 2, the primer failed to anneal at position 2 and PCR product was obtained only for primers bound at position 3 and 6.

AFLP

To overcome the limitation of reproducibility associated with RAPD, AFLP (Amplified Fragment Length Polymorphism) technology was developed. Like RAPD, AFLP does not require any DNA sequence information from the organism under study and also dominant marker. However, it is highly reliable and reproducible. It combines the power of RFLP with the flexibility of PCR-based technology by ligating primer recognition sequences (adaptors) to the restricted DNA and selective PCR amplification of restriction fragments using a limited set of primers. The first step in AFLP analysis involves restriction digestion of genomic DNA with a combination of rare cutter (*Eco*RI or *Pst*I) and frequent cutter (*Mse*I or *Taq*I) restriction enzymes. Double-stranded oligonucleotide adaptors are, then, designed in such a way that the initial restriction site is not restored after ligation. Such adaptors are ligated to both ends of the fragments to provide known sequences for PCR amplification. PCR amplification will only occur where the primers are able to anneal to fragments which have the adaptor sequence plus the complementary base pairs to the additional nucleotides called selective nucleotides. An aliquot is then subjected to two subsequent PCR amplifications under highly stringent conditions with primers complementary to the adaptors, and possessing 3' selective nucleotides of 1-3 bases.

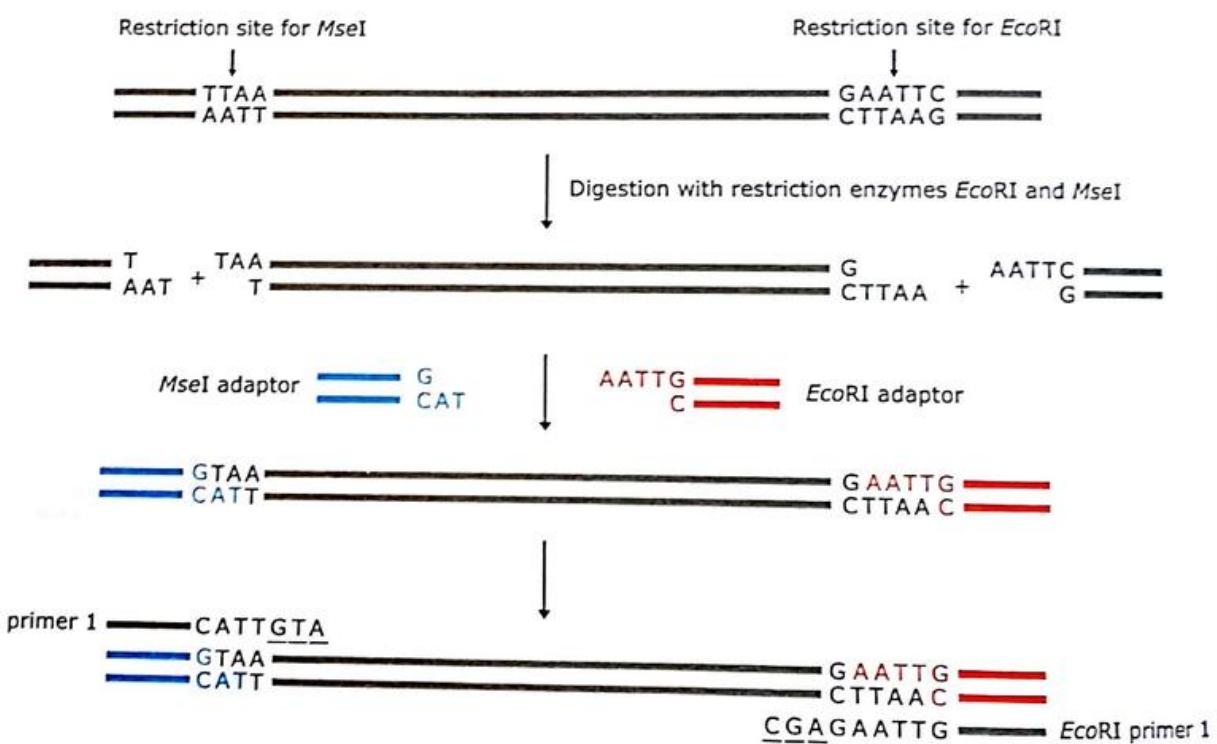


Figure 12.11 AFLP procedure. The procedure of this technique is divided into two steps:

1. Digestion of cellular DNA with two restriction enzymes and ligation of restriction half-site specific adaptors. The adaptor is designed in such a way that ligation of a fragment to an adaptor does not reconstitute the reaction site.
2. Selective amplification of some of these fragments with two PCR primers that have corresponding adaptor and restriction site specific sequences.

To achieve selective amplification of a subset of these fragments, primers are extended into the unknown part of the fragments [underlined base], usually one to three arbitrarily chosen bases beyond the restriction site. The first is performed with a single-bp extension, followed by a more selective base in the AFLP extension amplifies a different subset of fragments.

The first PCR (preamplification) is performed with primer combinations containing a single bp extension, while final (selective) amplification is performed using primer pairs with up to 3-bp extension. Because of the high selectivity, primers differing by only a single base in the AFLP extension amplify a different subset of fragments. A primer extension of one, two or three bases reduces the number of amplified fragments by factors of 4, 16 and 64, respectively. Ideal primer extension lengths will vary with genome size of the species and result in an optimal number of bands: not too many bands to cause smears or high levels of band co-migration during electrophoresis, but sufficient to provide adequate polymorphism. AFLP fragments are visualized either on agarose gel or on denaturing polyacrylamide gels with autoradiography.

12.2 Nucleic acid hybridization

Nucleic acid hybridization is a fundamental tool in molecular biology which takes advantage of the ability of individual single-stranded nucleic acid molecules to form double-stranded molecules by standard base pairing. For this to happen, the interacting single-stranded molecules must have a sufficiently high degree of base complementarity. Nucleic acid hybrids can be formed between two strands of DNA, two strands of RNA, or one strand of DNA and one of RNA.

Nucleic acid hybridization is used to identify related or identical molecules on the basis of base complementarity. The hybridization assays involve a labeled nucleic acid *probe* to identify complementary DNA or RNA molecules present within a complex mixture of nucleic acid molecules, the *target* nucleic acid. It can be used to detect homologous DNA or RNA sequences not only in cell extracts, but also in chromosomes or intact cells – a procedure called *in situ* hybridization.

Stringency of a hybridization reaction

Stringency is a term that describes the number of mismatched bases that are allowed in a hybridization reaction and still have a double-stranded hybrid. The more stringent the hybridization conditions are, the fewer the mismatched bases that are allowed before the two strands come apart.

Stringency is mainly influenced by the *temperature* and the *salt concentration* of the hybridization mixture. All these conditions affect the stability of the double-stranded molecule. High stringency conditions are usually achieved with high temperatures (approaching the melting temperature of the DNA-DNA hybrid) and low salt concentrations (<0.1 M). At high temperature and low salt concentration, only closely matched sequences will hybridize.

The optimum hybridization temperature is experimentally determined, starting with temperatures 5°C below the melting temperature (T_m). The T_m is the temperature at which the probe and target are 50% dissociated. A variety of equations have been derived that can be used to estimate the T_m under a range of conditions.

Factors influencing the hybridization reaction of nucleic acid in solution:

- Temperature
- Ionic strength
- Destabilizing agents
- Mismatched base pairs
- Duplex length
- Viscosity

Salt concentration also influences the stringency of hybridization reaction. The phosphate groups are moderately strong acids, and thus they are always ionized at physiological pH and bear the negative charges. These negative charges tend to repel the strands in double helical structure apart. Binding of positive ions (cations) with phosphates neutralize the negative charge of phosphates and thus screens the repulsive force between the sugar phosphate backbones in dsDNA. Lowering the salt concentration of a DNA solution promotes denaturation by removing the cations that shield the negative charges on the two strands from each other.

other. At low ionic strength, the mutually repulsive forces of these negative charges from the phosphoryl groups are enough to denature the DNA, even at a relatively low temperature. So, at low salt concentrations, the two strands will only stay double-stranded if the two strands are well matched (complementarity). High salt concentrations tend to stabilize the double-stranded molecule and allow more mismatches to take place when two strands hybridize. The pH of the solution in which hybridization takes place also affects the stability of hybridizing molecules. Under slightly acidic conditions (pH 6–7), the double-stranded DNA molecule is more stable because the increased H⁺ concentration also helps to reduce the repulsive forces between the two backbones of the DNA molecule.

Nucleic acid probe

Nucleic acid hybridization with a labeled probe is the only practical way to detect a complementary target sequence in a complex nucleic acid mixture. Nucleic acid probes are oligonucleotides or polynucleotides that can bind with high specificity to complementary sequences. Probes can be complementary to either DNA or RNA and can be from as few as 20 nucleotides to hundred of nucleotides long. A nucleic acid probe may be DNA (DNA probes), RNA (RNA probes) and synthetic oligonucleotide. Probes are usually labeled (isotopically or non-isotopically) to aid their easy detection when bound to the target nucleic acids.

Oligonucleotide probes are synthesized chemically and end-labeled. DNA probes, which are cloned DNAs and may either be end-labeled or internally labeled during *in vitro* replication. RNA probes are internally labeled during *in vitro* transcription from cloned DNA templates. DNA probes can be double-stranded or single-stranded. Prior to hybridization, the double-stranded probe will be denatured. RNA probes and oligonucleotide probes are generally single-stranded. DNA probes, RNA probes and oligonucleotide may be *heterologous* and *homologous*.

A heterologous probe is a probe that is similar to, but not exactly the same as, the nucleic acid sequence of interest. If the gene being sought is known to have a similar nucleotide sequence to a second gene that has already been cloned, then it is possible to use this known sequence as a probe. For example, a mouse probe could be used to search a human genomic library. A homologous probe is a probe that is exactly complementary to the nucleic acid sequence of interest.

12.3 Labeling of nucleic acids

Nucleic acids may be modified with tags that enable detection or purification. The resulting nucleic acid can be used to identify or recover other interacting molecules. Nucleic acids can be labeled by *isotopic* and *non-isotopic labeling* methods:

Isotopic labeling

Isotopic labeling of nucleic acids has been conducted by incorporating nucleotides containing radioisotopes (radiolabeling). Such radiolabeled probes contain nucleotides with a radioisotope (often ³²P, ³³P, ³⁵S or ³H), which can be detected specifically in solution or, much more commonly, within a solid specimen. In molecular biology, two are especially important: the radioactive isotopes of phosphorus, ³²P, and sulfur, ³⁵S. Since sulfur is not a normal component of DNA or RNA, we use *phosphorothioate* derivatives. A normal phosphate group has four oxygen atoms around the central phosphorus. In a phosphorothioate, one of these is replaced by sulfur. To introduce ³⁵S into DNA or RNA, phosphorothioate groups containing radioactive sulfur atoms are used to link together the nucleotides.

Isotopic labeling involves the replacement of specific atoms by their isotopes. The isotopes may be stable or radioisotopes. The labeling with radioisotopes or radioactive isotopes is called *radiolabeling*.

Atoms that have the same atomic number, but have different masses are known as isotopes. Some isotopes of an atom have unstable nuclei which after nuclear reaction emit characteristic radiation. These isotopes are called radioisotopes, or more commonly *radionuclides*.

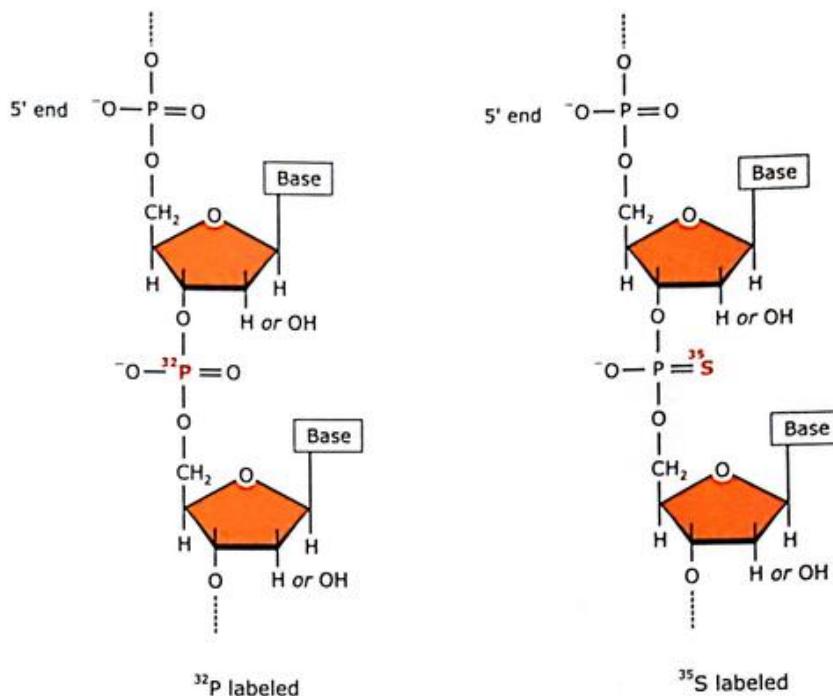
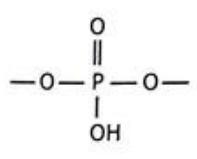


Table 12.1 Radioisotopes which are commonly used in biological research

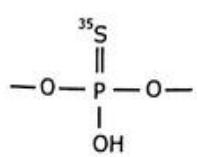
Isotope	Half-life
³² P	14 days
¹³¹ I	8.1 days
³⁵ S	87 days
¹⁴ C	5570 years
⁴⁵ Ca	164 days
³ H	12.3 years

The intensity of signal produced by radioisotopes is dependent on the intensity of the radiation emitted by the radioisotopes, and the time of exposure, which may often be long (one or more days, or even weeks in some applications). ³²P has been used widely in Southern blot hybridization, dot-blot hybridization, colony and plaque hybridization because it emits high energy β -particles which provide a high degree of sensitivity of detection.

The incorporation of radioisotopes in a sample can be detected by two common methods. In *autoradiography*, radiolabeled material is allowed to expose a photographic emulsion. Development of the emulsion reveals the distribution of labeled material. In the second detection method, the amount of radioactivity in radiolabeled samples is directly measured, either by a *Geiger counter* or by a *scintillation counter*. In a Geiger counter, emission from radioisotopes is detected by the ionization they produce in gas. In scintillation counting, the sample is mixed with a material that will fluoresce upon interaction with a radiation emitted by radioactive decay. The scintillation counter quantifies the resulting flashes of light.



Phosphate group



Phosphorothioate group

Autoradiography is a technique which is used for detecting radioisotopes present in a solid sample on gels or membranes. It involves the production of an image in a photographic emulsion. Such emulsions consist of silver halide crystals in gelatin base. When a β -particle or γ -ray from a radionuclide passes through the emulsion, the silver ions are converted to silver atoms. The resulting latent image can then be converted to a visible image once the image is

developed; an amplification process in which entire silver halide crystals are reduced to give metallic silver.

Scintillation counting relies on special chemicals called **scintillants**. The scintillant molecules absorb the radiations emitted by the radioisotopes in the samples, and in turn emit a flash of light. The light pulses from the scintillant are detected by a photocell. To use the scintillation counter (machine that detects and counts pulses of light), radioactive samples to be measured are added to a vial containing scintillant fluid and loaded into the counter. The counter prints out the number of light flashes it detects within a designated time.

Label location

There are two ways to label a DNA molecule – by the ends (*end labeling*) or all along the molecule (*uniform labeling*).

Labeling DNA by nick translation

Nick translation is one method of labeling DNA, which uses the enzymes pancreatic DNase I, *E. coli* DNA polymerase I and DNA ligase. The endonuclease DNase I is used to create nicks at random sites in both strands of double stranded target DNA. Following DNase I treatment, DNA polymerase I is used to add nucleotide residues to the free 3'-hydroxyl ends created during the DNase I nicking process. As the DNA polymerase I extends the 3'-ends, the 5'-to 3' exonuclease activity of the enzyme simultaneously removes bases from the 5'-end of the nick. The sequential addition of base onto the 3'-end with the simultaneous removal of bases from the 5'-end results in translation of the nick along the DNA molecule. When performed in the presence of a radioactive deoxynucleoside triphosphate such as ($[\alpha-^{32}\text{P}] \text{dCTP}$), the newly synthesized strand becomes radioactively labeled. For nonradioactive labeling procedures, a digoxigenin or a biotin moiety attached to a dNTP analog is used.

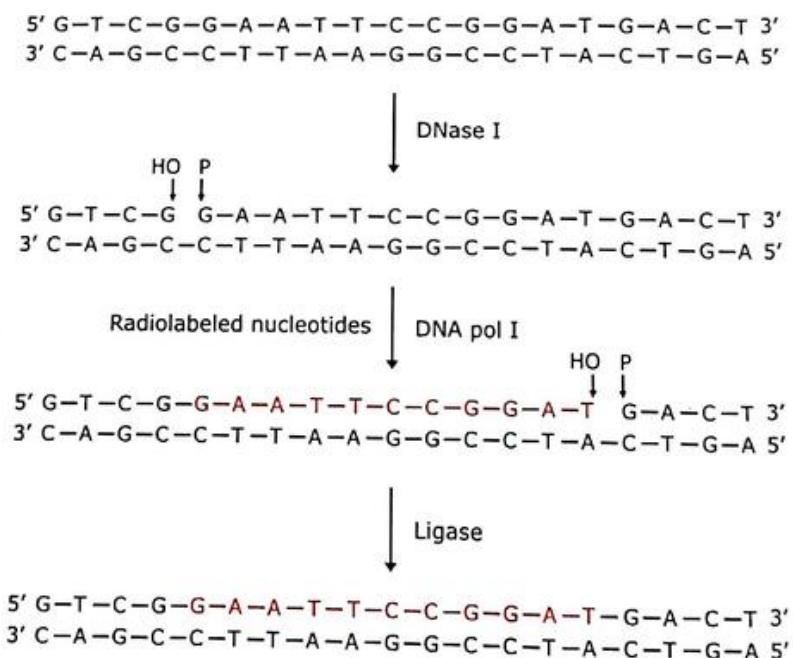


Figure 12.12 Nick translation. DNase I introduces single-stranded nicks by cleaving internal phosphodiester bonds, generating a 5' phosphate group and a 3' hydroxyl terminus. Addition of DNA pol I contributes two enzyme activities — a 5'-3' exonuclease attacks the exposed 5' termini of the nick and sequentially removes nucleotides in the 5'-3' direction and a DNA polymerase adds new nucleotides to the exposed 3' hydroxyl group in the 5'-3' direction.

Random priming

An alternative method for preparing uniformly labeled DNA is by oligonucleotide-primed DNA synthesis with hexanucleotides (or longer oligomers) of random sequence. Oligonucleotides of random sequence will anneal to a variety of homologous locations on a single-stranded DNA. Once annealed, they serve as primers for DNA synthesis by DNA polymerases. The Klenow fragment is used as this enzyme lacks the 5'-3' exonuclease activity of DNA polymerase I; and so only fills in the gaps between adjacent primers. Labeled nucleotides are incorporated into the new DNA that is synthesized.

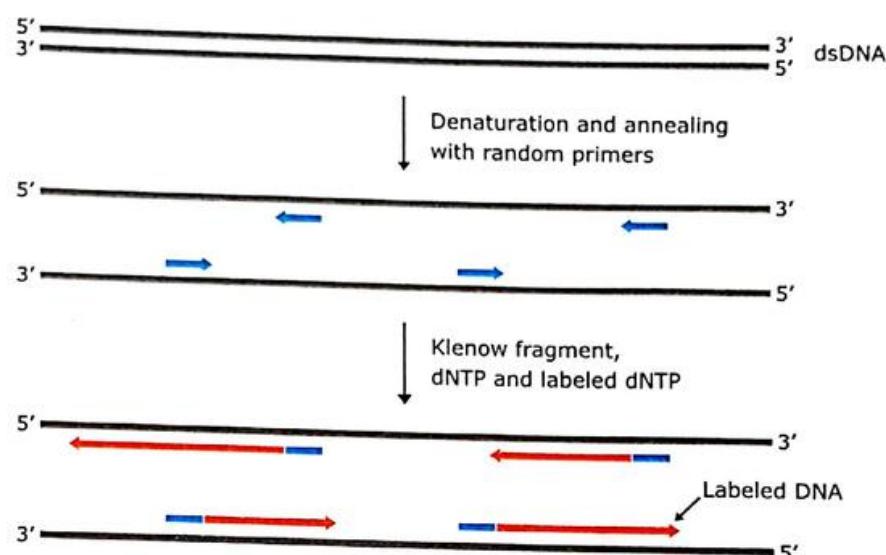


Figure 12.13 Labeling by random priming.

End-labeling of DNA

End labeling can be performed at the 3'- or 5'-end.

3'-end labeling

Template-independent polymerization of $[\alpha^{32}\text{P}]$ NTP to the 3' terminus of DNA is catalyzed by calf thymus terminal deoxynucleotidyl transferase. Terminal deoxynucleotidyl transferase (TdT) is a template independent DNA polymerase that incorporates dNTPs to the 3'-OH end of single or double-stranded DNA and RNA in an irreversible manner. This enzyme is used for the formation of homo or heteropolymeric tail at the 3'-end and also for incorporating a single nucleotide analog such as $[\alpha^{32}\text{P}]$ cordycepin-5'-triphosphate. Terminal dideoxynucleotidyl transferase labels 3'-protruding ends more efficiently than blunt ends.

5'-end labeling

5'-end labeling is performed by enzymatic methods (T4 polynucleotide kinase), by chemical modification of sensitized oligonucleotides with phosphoramidite, or by combined methods. 5'-end labeling is usually performed using polynucleotide kinase (*kinase end-labeling*). The polynucleotide kinase utilizes two types of reactions: forward reaction and exchange reaction. In the *forward reaction*, a hydroxyl group is first created by removing the unlabeled phosphate residue from the 5'-end of the DNA with an alkaline phosphatase. T4 polynucleotide kinase is then used to transfer the labeled gamma phosphate from ATP to the 5'-end of DNA.

In the *exchange reaction*, polynucleotide kinase first transfers the phosphate from the 5'-end of DNA to ADP, forming ATP and leaving a dephosphorylated target. Then enzyme performs a forward reaction and transfers a labeled gamma phosphate from ATP onto the target DNA.

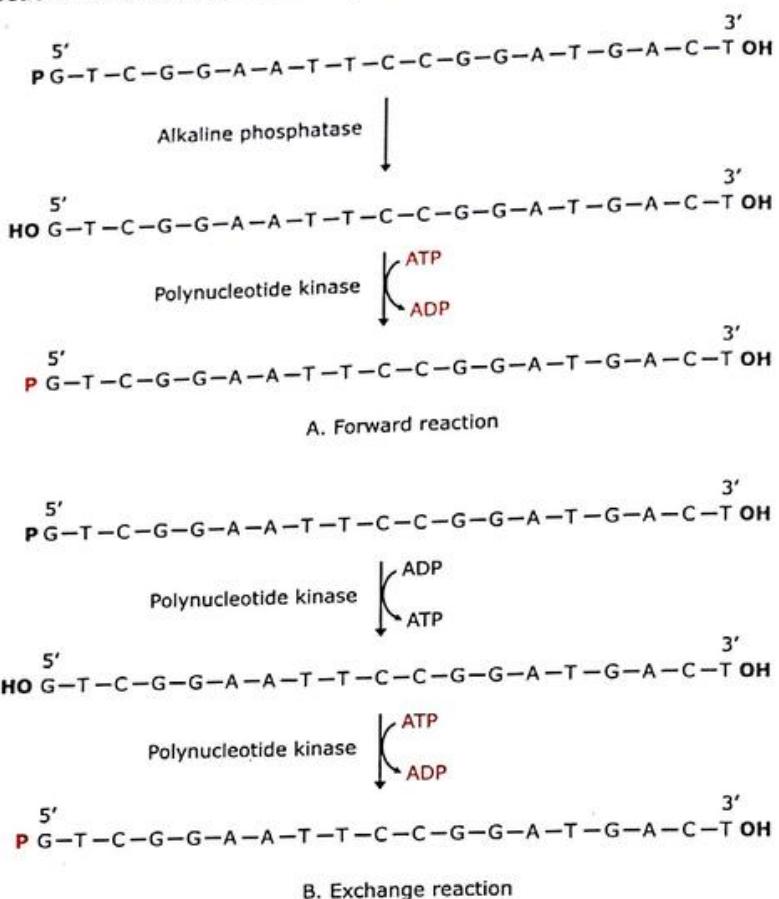


Figure 12.14 5'-end labeling. (A) In the *forward reaction*, polynucleotide kinase transfers the gamma phosphate from ATP to the 5'-end of a polynucleotide (DNA or RNA). (B) In the *exchange reaction*, target DNA or RNA that has a 5' phosphate is incubated with an excess of ADP. Polynucleotide kinase first transfers the phosphate from the nucleic acid onto an ADP, forming ATP and leaving a dephosphorylated target. Polynucleotide kinase then performs a forward reaction and transfers a phosphate from ATP onto the target nucleic acid.

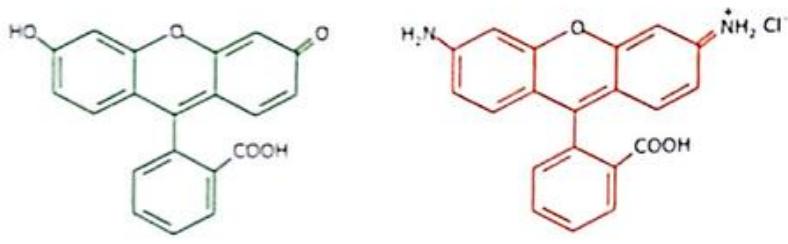
Nonisotopic labeling

Compared to radioactive labels, the use of nonradioactive labels have several advantages:

- Safety.
- Higher stability of probe.
- Efficiency of the labeling reaction.
- *In situ* detection.
- Less time taken to detect the signal.

Nonisotopic labeling systems involve the use of non-radioactive probes. Two types of non-radioactive labeling are conducted – direct and indirect. Direct labeling strategies utilize probes that are directly conjugated to a dye or an enzyme, which generates the detection signal. Indirect labeling systems utilize probes that contain a hapten that will bind to a secondary agent generating the detection signal; the probe itself does not generate signal.

Direct nonisotopic labeling, where a nucleotide which contains the label that will be detected is incorporated. Often such systems involve incorporation of modified nucleotides containing a fluorophore, a chemical group which can fluoresce when exposed to light of a certain wavelength. Most commonly used fluorophores for direct labeling are *fluorescein*, a pale green fluorescent dye and *rhodamine*, a red fluorescent dye.



Fluorescein

Rhodamine, core structure

Indirect nonisotopic labeling, usually featuring the chemical coupling of a modified reporter molecule to a nucleotide precursor. After incorporation into DNA, the reporter groups can be specifically bound by an affinity molecule, a protein or other ligand which has a very high affinity for the reporter group. Conjugated to the latter is a marker molecule or group which can be detected in a suitable assay. The reporter molecules on modified nucleotides need to protrude sufficiently far from the nucleic acid backbone to facilitate their detection by the affinity molecule and so a spacer of 4-16 carbon atoms long is required to separate the nucleotide from the reporter group.

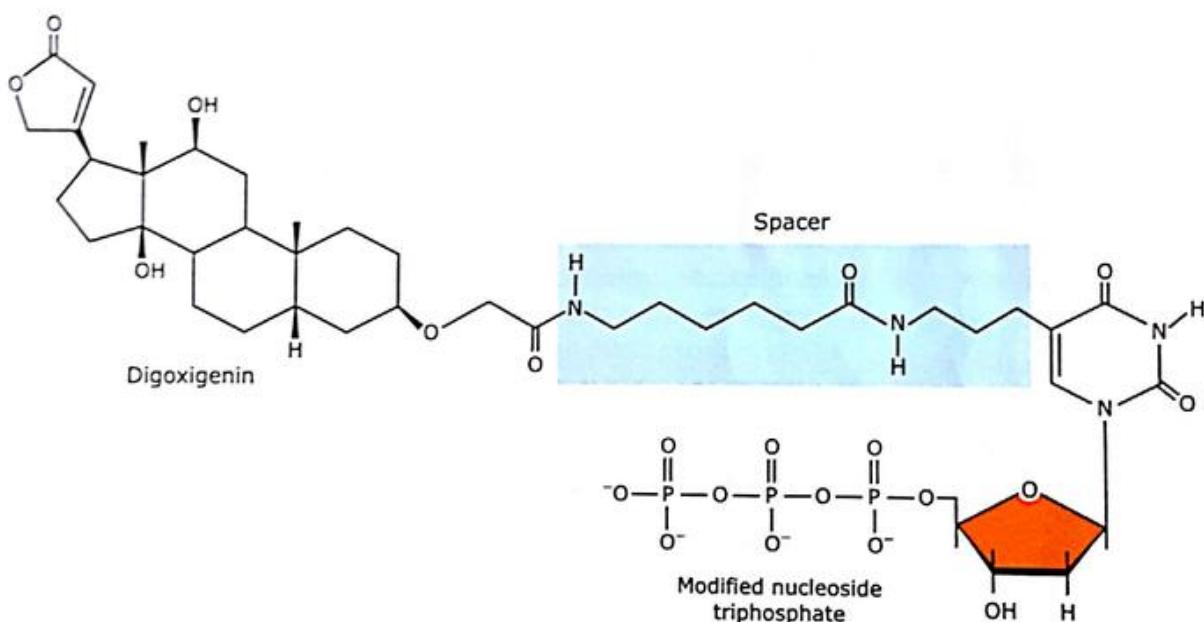


Figure 12.15 The base of the nucleoside triphosphate depicted is an analog of thymine in which the methyl group has been replaced by a spacer arm linked to the plant steroid digoxigenin. The digoxigenin is detected by a specific antibody coupled to a visible marker such as a fluorescent dye.

Two indirect nonisotopic labeling systems are widely used:

The biotin-streptavidin system utilizes the extremely high affinity of two ligands – biotin (vitamin H) which acts as the reporter; and the bacterial protein streptavidin which is the affinity molecule. Biotin and streptavidin bind together extremely tightly with a dissociation constant in the order of 10^{-14} mol/litre, one of the strongest known in biology.

Digoxigenin is a plant steroid (obtained from *Digitalis* plants) to which a specific antibody has been raised. The digoxigenin-specific antibody permits detection of nucleic acid molecules which have incorporated nucleotides containing the digoxigenin reporter molecule.

Both biotin and digoxigenin are linked to uracil, which is normally a component of RNA not DNA. Therefore, to label DNA, uracil must be incorporated into the DNA instead of thymine.

If deoxyUTP labeled with biotin or digoxigenin is added to the polymerization reaction, DNA polymerase will incorporate the labeled uridine where thymidine would normally be inserted. The biotin or digoxigenin tags stick out from the DNA without disrupting its structure.

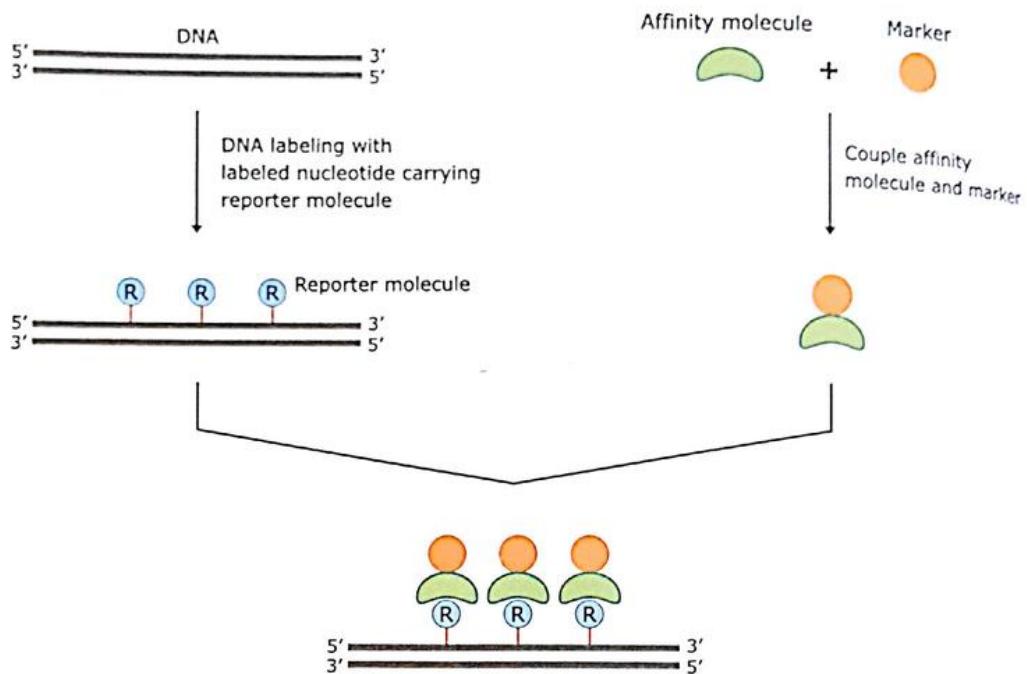


Figure 12.16 Indirect non-isotopic labeling involves chemical linkage between reporter molecule and a nucleotide. When this modified nucleotide is incorporated into DNA, then it binds specifically to an affinity molecule which has high affinity against the reporter molecule. Long spacer is introduced between nucleotide and reporter molecule so as to reduce steric hindrance for binding of affinity molecule.

Molecular beacons

The *molecular beacon* is a structured fluorescent probe. A structured probe contains stem-loop structure regions that confer enhanced target specificity when compared with a traditional linear probe. Molecular beacons form a stem-loop structure, where the central-loop-sequence is complementary to the target of interest and the stem arms are complementary to each other.

It is an oligonucleotide (about 25 nucleotides long) that contains both a fluorophore and a quenching group at opposite ends. Its detection mechanism relies on the principle of FRET, in which a fluorophore in the excited state, can transfer energy to acceptor fluorophore, with subsequent emission of a fluorescent signal from the acceptor, or to a quencher dye, which dissipates the energy without emission of a detectable fluorescence signal. In order for this energy transfer to take place, the donor and acceptor molecules must be situated in close physical proximity.

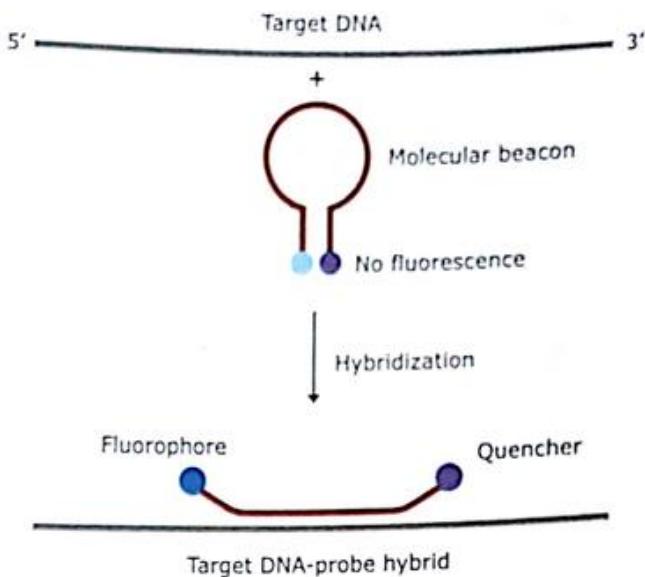


Figure 12.17 Structure and operation of molecular beacons. A typical molecular beacon probe is about 25 nucleotides long. About 15 nucleotides in the middle of probe are complementary to the target DNA or RNA and 5–7 nucleotides at each terminus are complementary to each other; rather than to the target DNA. At the 5'-end of the probe, a fluorescent dye is covalently attached. The quencher dye is covalently attached to the 3'-end. When the beacon is in closed loop shape, the quencher resides in proximity to the fluorophore, which results in quenching the fluorescent emission of the latter. Hence, these molecules are non-fluorescent. But when the probe sequence in the loop hybridizes to its target, forming a double helix, a conformational reorganization occurs that separates the quencher from the fluorophore, restoring fluorescence.

Molecular beacon is a sequence-specific probe designed to fluoresce only when it binds to a specific DNA target sequence. The central region of the probe is complementary to the target sequence. The terminal 5 to 7 nucleotides at each end of the probe are complementary and form a short double-stranded region. In the stem and loop conformation, the quenching group is next to the fluorophore and so prevents fluorescence. When the molecular beacon binds to the target sequence, it is linearized. This separates the quenching group from the fluorophore, which is now free to fluoresce.

Scorpions

Scorpions are bifunctional molecules containing a PCR primer covalently linked to a probe. The fluorophore in the probe interacts with a quencher, which reduces fluorescence. During a PCR reaction, the fluorophore and quencher separate, which leads to an increase in fluorescence. Scorpion probes are of two basic designs. Both types attach to the 5'-end of a PCR primer through a compound that inhibits PCR extension into the probe. In the first design, the probe consists of a stem-loop structure with a quencher and fluorophore similar to a molecular beacon (unimolecular stem-loop format). In the second design, the fluorophore-containing probe base-pairs with a complementary oligonucleotide that contains the quencher (bimolecular linear scorpion format).

In unimolecular stem-loop format, the probes contain the fluorophore at 5'-end, stem-loop, quencher, a PCR blocker (which prevents read-through by DNA polymerase) and primer.

Scorpions are used as fluorescent reporter molecules that enable the quantification of PCR products in real time. During the extension step of PCR, the primer portion of the probe anneals to the template and *Taq* polymerase makes new DNA. During the next denaturation step, the whole probe plus new DNA strand become single-stranded. During annealing step,

the loop section of the scorpion probe is able to base-pair with its complementary sequence within the target DNA, releasing the fluorophore from the quencher. The resultant fluorescence emission gives the direct measure of the amount of PCR product produced.

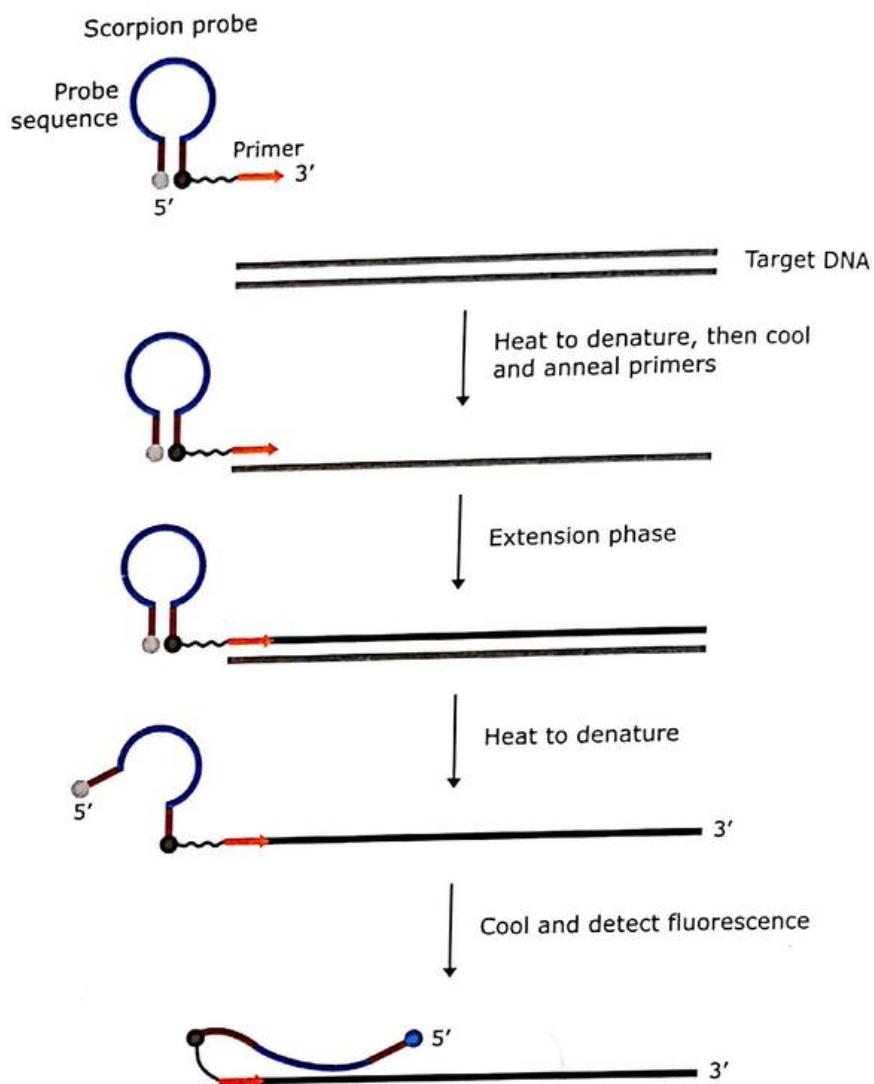


Figure 12.18 Scorpions incorporate two distinct structures - a target-specific DNA probing sequence and a target-specific PCR primer. It contains a stem-loop structure with a fluorophore molecule and quencher. The loop of the Scorpions probe includes a sequence that is complementary to an internal portion of the target sequence. During the first amplification cycle, the Scorpions PCR primer is extended, and the sequence complementary to the loop sequence is generated on the same strand. After subsequent denaturation and annealing, the loop of the Scorpions probe hybridizes to the internal target sequence, and the reporter is separated from the quencher. The resulting fluorescence signal is proportional to the amount of amplified product in the sample. The Scorpion probe contains a PCR blocker, just downstream of the quencher, to prevent read-through during the extension of the opposite strand.

✓ Fluorescence *in situ* hybridization (FISH)

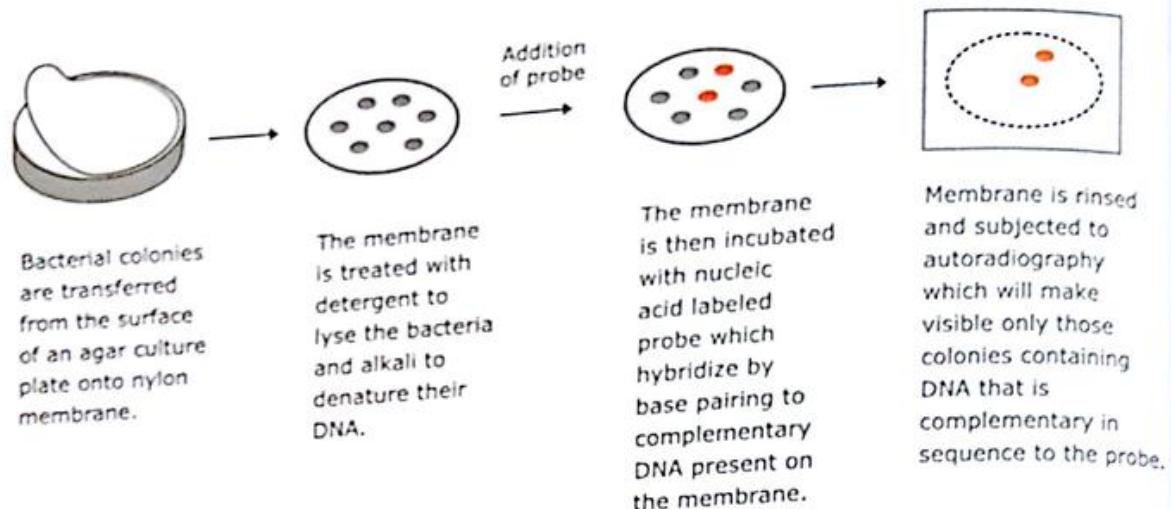
FISH is essentially based upon the same principle as a Southern blot analysis that exploits the ability of single-stranded DNA to anneal to complementary DNA. In the case of FISH, the target is the nuclear DNA of either interphase stage or metaphase stage. It was originally used with metaphase chromosomes. It involves hybridization of a fluorescent-labeled DNA probe to denatured DNA of metaphase chromosomes. In the past, probes were labeled with radioisotopes. But now-a-days, fluorescent dyes are used for increased sensitivity and resolution.

In FISH, the DNA probe is either labeled directly by incorporation of a fluorescent-labeled nucleotides precursor, or indirectly by incorporation of a nucleotides precursor containing a reporter molecule which after incorporation into the DNA is then bound by a fluorescently labeled affinity molecule. The position at which the probe hybridizes to the chromosomal DNA is visualized by detecting the fluorescent signal emitted by the labeled DNA. If FISH is carried out simultaneously with two DNA probes, each labeled with a different fluorochrome, the relevant positions of the two markers to which probes hybridize on the chromosome can be visualized. Hence, FISH enables the position of a marker on a chromosome or extended DNA molecule to be directly visualized.

FISH involves two main components: the DNA probe and the target DNA to which the probe will be hybridized. One of the most important considerations in FISH analysis is the choice of probe. A wide range of probes can be used, from whole genomes to small cloned probes. There are broadly three types of probes, each with a different range of applications: whole-chromosome painting probes; repetitive sequence probes and locus-specific probes. For the method to work, the DNA in the chromosome must be made single-stranded by denaturation of double helix. Only then will the chromosomal DNA be able to hybridize with the probe. The standard method for denaturing chromosomal DNA without destroying the morphology of the chromosome is to dry the preparation onto a glass microscope slide and then treat with *formamide*. The most common methods for FISH visualization are: flow cytometry systems and slide-based systems. In flow systems, FISH-stained cells are prepared in suspension, and the suspension flows in a narrow stream across a laser beam wherein the detector records their fluorescent intensities. In slide-based systems, FISH-stained cells are fixed to a conventional wide-field or a confocal fluorescence microscope slide and observed as a static image. Prior to fixation, the cells can be trapped in the metaphase by colchicine, which interferes with mitosis. The fixed preparations are, then, incubated with various solvents and at elevated temperatures to allow the probe DNA to hybridize with the chromosomes.

Colony hybridization

Colonies of bacteria which contain specific DNA can be selected or identified by colony hybridization. In this hybridization process, bacterial colonies are transferred from the surface of an agar culture plate onto a nylon membrane. The colonies transferred to the membrane are subjected to alkali hydrolysis and detergent treatment to release the DNA content from the bacterial cells, which would then bind to the membrane. The DNA on the membrane is denatured with an alkali to produce single strands which covalently binds with the membrane by UV irradiation. The membrane is, then, immersed in a solution containing a labeled nucleic acid probe and incubated to allow the probe to hybridize to its complementary sequence. After hybridization, the membrane is washed extensively to remove unhybridized probe, and regions where the probe has hybridized are then visualized. By comparing the membrane with the original dish and lining up the regions of hybridization, the original group of colonies can be identified.



12.4 Blotting

Blotting describes the immobilization of sample nucleic acids/proteins onto a solid support. It involves the transfer of nucleic acids and/or proteins from a gel strip to a specialized, chemically reactive matrix called *blotting membrane* (typically nitrocellulose or activated nylon) on which the nucleic acids/proteins may become immobilized in a pattern similar to that present in the original gel.

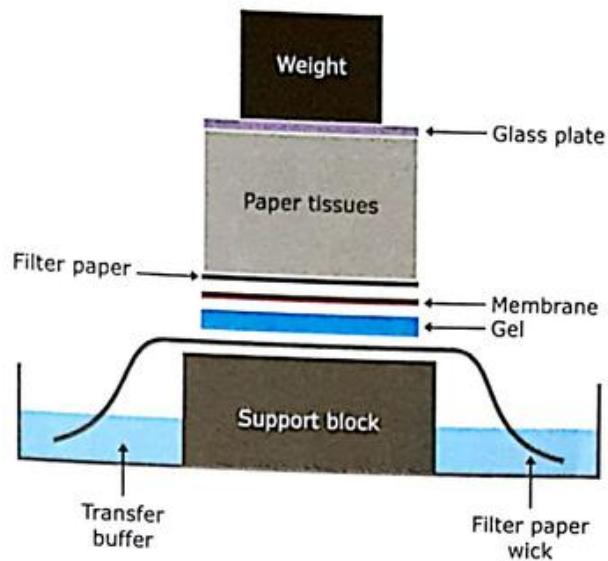
Southern blotting

Southern blotting (or Southern blot hybridizations) immobilizes and detects target DNA fragments onto a membrane that have been size-fractionated by gel electrophoresis. This technique was invented in 1975 by E. M. Southern. In the procedure, the DNA fragments separated on an agarose gel are denatured, transferred and immobilized onto a membrane.

Following electrophoresis, the test DNA fragments are denatured in strong alkali. As electrophoretic gels are fragile, and the DNA in them can diffuse within the gel, it is usual to transfer the denatured DNA fragments by blotting onto a durable nitrocellulose paper or nylon membrane, to which single-stranded DNA binds readily. Now-a-days, nylon membranes are commonly used. After transfer, the DNA fragments need to be fixed to the membrane so that they cannot detach. In case of nitrocellulose paper, nucleic acid immobilization occurs non-covalently after baking for 2 hrs at 80°C. In case of nylon membrane either it is baking for 1 hour at 70°C or UV irradiation at 254 nm. Nucleic acid binds covalently with nylon membrane after UV irradiation for 5 minutes. UV irradiated covalent-linking is based on the formation of cross-links between the T-residues in the DNA and the positively charged amino groups on the surface of the nylon membrane. The individual DNA fragments become immobilized on the membrane at positions which are a faithful record of the size separation achieved by gel electrophoresis.

Following the fixation step, the membrane is placed in a solution of labeled (radioactive or non-radioactive) RNA, single-stranded DNA or oligodeoxynucleotide which is complementary in sequence to the blot transferred DNA band or bands to be detected. Since this labeled nucleic acid is used to detect and locate the complementary sequence, it is called the *probe*. The probe is allowed to hybridize to its complementary single-stranded target DNA sequences on the membrane. Conditions are chosen which maximize the rate of hybridization, compatible with a low background of non-specific binding on the membrane. After the hybridization, reaction has been carried out, the membrane is then washed extensively to remove non-specifically bound probe.

Figure 12.19 Blotting apparatus. The gel is placed on a filter paper wick and a nitrocellulose paper or nylon membrane placed on top. Further, sheets of filter paper and paper tissues complete the setup. Transfer buffer is drawn through the gel by capillary action, and the nucleic acid fragments are transferred out of the gel and onto the membrane.



If the probe is radioisotope labeled, then the membrane is exposed to photographic film. If the probe is non-isotopically labeled with biotin or digoxigenin, the membrane may be treated with chemiluminescent substrate to detect the labeled probe, and then exposed to photographic film. The probe will form a band on the film at a position corresponding to the complementary sequence on the membrane.

Northern blotting

Northern blotting or Northern blot hybridization is a variant of Southern blotting in which the target nucleic acid is RNA instead of DNA. This method is used to measure the amount and size of RNAs transcribed from genes and to estimate their abundance. In this technique, an RNA extract is electrophoresed in an agarose gel, using a denaturing buffer to ensure that the RNAs do not form inter or intra molecular base pairs. After electrophoresis, the gel is blotted onto a reactive DBM (diazobenzyloxymethyl) paper, and hybridized with a labeled probe. RNA bands can also be blotted onto nitrocellulose paper under appropriate conditions and suitable nylon membranes.

Nylon is a generic name for any long-chain synthetic polymer having recurring polyamide ($-\text{CONH}-$) groups. Two types of nylon membranes are available commercially: unmodified (or neutral) nylon and charge-modified nylon which carries amine groups and is, therefore, also known as *positively charged* nylon. Both types of nylon bind single- and double-stranded nucleic acids. Charge-modified nylon has a greater capacity to bind nucleic acids.

Table 12.2 Properties of materials used for blotting of nucleic acids:

Materials	Binding capability
Nitrocellulose	ssDNA, RNA
Nylon (Neutral)	ssDNA, dsDNA, RNA
Positively charged nylon	ssDNA, dsDNA, RNA
Activated papers (DBM and DPT)	ssDNA, RNA
DBM-diazobenzyloxymethyl, DPT-Diazophenylthioether	

Dot blot assay

In a dot blot assay, a specified volume of nucleic acid mixture is spotted onto a small area of a nylon membrane. In this technique, the nucleic acid molecules are not first separated by electrophoresis. Instead, a specimen containing nucleic acid mixture to be detected is applied directly on a membrane as a small dot. The dot is treated with an alkaline solution to denature DNA molecules. Finally, the dot-blot membrane is allowed to hybridize with a labeled probe. The membranes are processed and exposed to film. If the dot of DNA contains a sequence similar to the probe, the film will turn dark in that area. If no dark spot appears on the film, it can be inferred that no DNA hybridization has occurred.

Slot blot is fundamentally similar to dot blot. The difference between dot and slot blot procedures is in the way that the nucleic acid mixture is blotted onto the membrane. In dot blot, the nucleic acids are blotted as circular blots, whereas in slot blot they are blotted in rectangular slots.

12.5 Phage display

Phage display is a molecular technique that allows expression of foreign polypeptides or peptides on the surface of phage particles. This method was first described by George Smith in 1985. In this technique, the DNA encoding the protein of interest is fused with a gene encoding one of the proteins that forms the viral coat. Genetic engineering techniques are used to insert foreign DNA fragments into a suitable phage coat protein gene. Phage display involved the use of filamentous phages such as fd, f1, M13, where the foreign gene was incorporated into a gene specifying a minor coat protein. Filamentous phage M13 is the most popular choice for phage display. M13 is a filamentous phage contains 6.4 kb single-stranded circular DNA. M13 enters *E. coli* through the bacterial sex pilus, a protein appendage that permits the transfer of DNA between bacteria. The single-stranded DNA in the virus particle [called the (+) strand] is replicated through an intermediate circular double-stranded replicative form containing (+) and (-) strands. Only the (+) strand is packaged into new virus particles.

Protein-protein interaction
Study of interaction between different proteins is very useful as it yields information about the tentative function of a protein through its interaction with well characterized protein with known role or function. Phage display and yeast two hybrid system are two well known and most useful methods for studying protein-protein interactions.

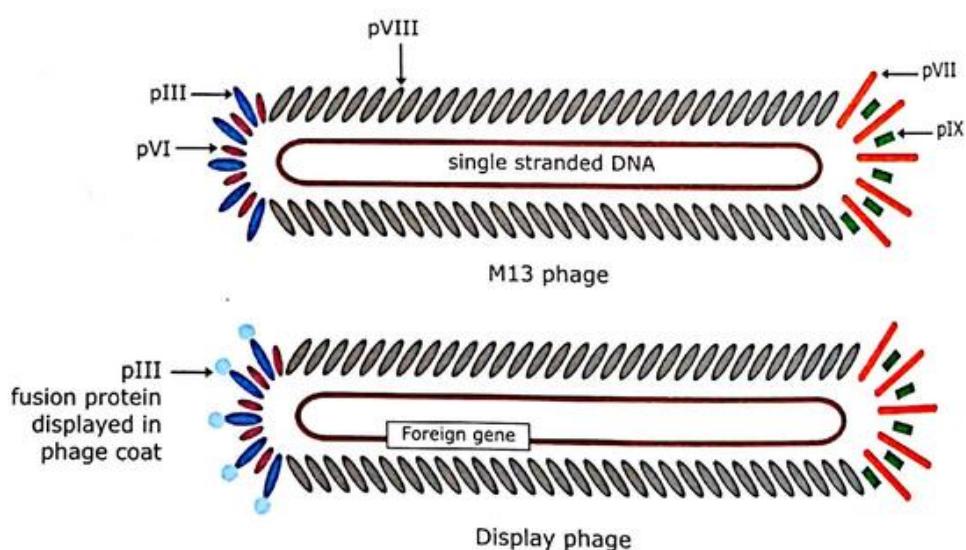


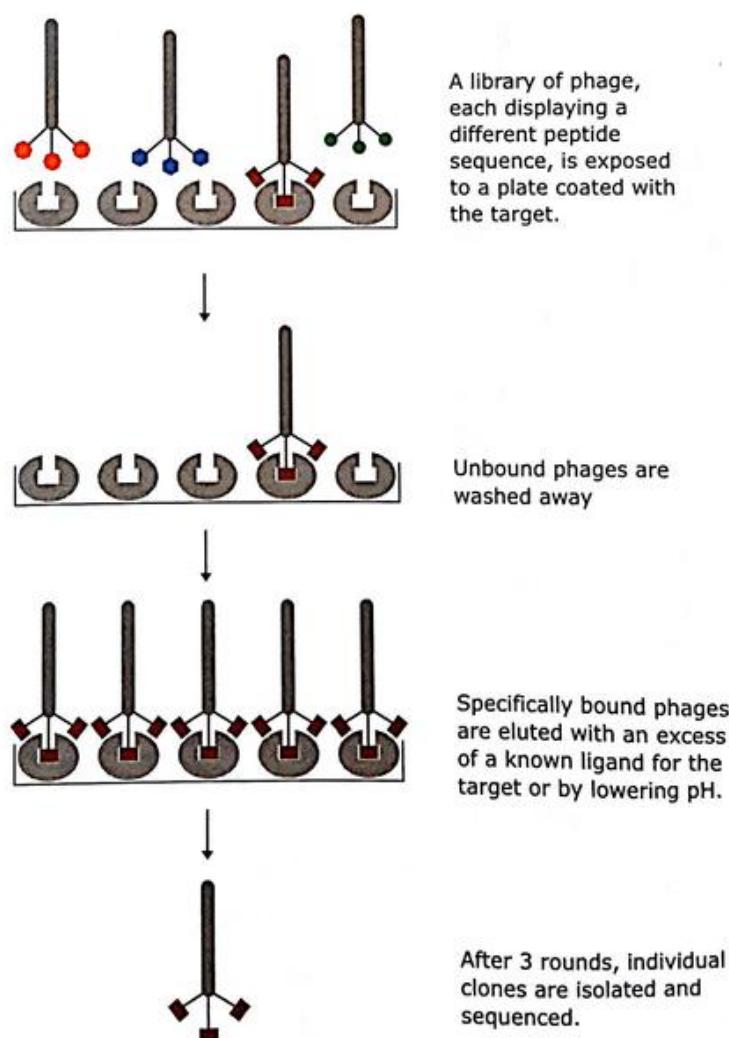
Figure 12.20 Phage display. In order to display a peptide on the surface of a bacteriophage, the DNA sequence encoding the peptide must be fused to the gene for a bacteriophage coat protein (either to the pVIII gene or the pIII gene). The coat protein is expressed as a fusion protein on the virion surface without disturbing the infectivity of the phage.

M13 is preferred since it is non-lytic and does not destroy the host bacteria during phage production. Instead, phage particles are secreted through the bacterial cell envelope. The absence of cell debris simplifies purification of the phage. The M13 phage particle consists of a single-stranded DNA molecule surrounded by a protein coat. The coat consists of major coat protein (pVIII) and minor coat proteins (pIII, pVI, pVII and pIX). At one end of the particle are five copies each of the two minor coat proteins, pIX and pVII and at the other end five copies each of pIII and pVI. The minor coat protein pIII is located on the tip of the phage, which is responsible for attaching the phage to its bacterial host during infection.

Gene encoding the minor coat protein pIII is most commonly used for recombinant formation with the DNA encoding the foreign target peptide or polypeptide. *E. coli* transfected with the recombinant DNA molecules produce phage particles that display peptide or polypeptide as a fusion protein with endogenous pIII coat protein on the surface of the phage particles.

There are two main methods in which phage display can be used to study protein interaction. In one method, test protein is displayed and its interaction sought with a series of purified proteins or protein fragments of known function. In the second method, a phage display library is prepared. The library is made up of many recombinant phages, each displaying a different protein. These libraries can be prepared by cloning a mixture of cDNAs from a particular tissue by cloning genomic DNA fragments. The library consists of the phage displaying a range of different proteins and is used to identify those that interact with a test protein.

Figure 12.21 Phage display libraries and biopanning. Phage display libraries consist of a large number of modified phages displaying a library of different peptide sequences. They are screened to find peptides that bind to specific target molecules, such as a particular antibody, enzyme or cell surface receptor. The peptide of interest is found by a selection procedure referred to as *biopanning*. The phage display library is incubated with target molecules that are attached to a solid support. Unbound phage is washed away. After several washings, only the phages that contain a protein fragment with high affinity for the ligands will remain bound to the container; the rest will be washed off.



Vector used for phage display

In the bacteriophage display system, a segment of foreign DNA is inserted into gene III or gene VIII, a few nucleotides downstream from the cleavage site. *E. coli* transfected with the recombinant viral DNA synthesize and secrete *fusion phage* particles that display on their surface the amino acids encoded by the foreign DNA. Every copy of pIII or pVIII on the surface of an infectious bacteriophage particle carries the sequences encoded by the foreign DNA, which are, therefore, displayed in a densely packed, 'multivalent' fashion.

The phagemid display system consists of a plasmid that carries a single copy of gene III or gene VIII and the viral origin of DNA replication. In phagemid display, as in conventional phage display, a segment of foreign DNA is inserted into gene III or gene VIII just downstream from the cleavage site that separates the hydrophobic signal sequence from the mature protein. The recombinant plasmid is then used to transform an appropriate strain of *E. coli*. Bacteriophage particles displaying the amino acid sequences encoded by the segment of foreign DNA are obtained by superinfecting the transformed cells with helper phages. Because replication and packaging of the helper phages are less efficient than that of the phagemid, the population of bacteriophages secreted from the superinfected cells consists mainly of particles that display the cloned target sequence.

12.6 Yeast two-hybrid assay

Yeast two-hybrid assay is used to detect protein interactions inside a yeast cell nucleus. It is based on the fact that a eukaryotic gene expression is promoted by transcription factors called *activators* that contain two distinct domains. One is the *DNA binding domain* (DBD) that binds to regulatory sequences and another is the *activation domain* (AD) that interacts with RNA polymerase and stimulates expression of the associated gene. These two domains can be physically separated from each other in which case the protein loses its activity, but is re-activated when the two domains are brought together.

Yeast two hybrid system makes use of a *Saccharomyces cerevisiae* strain which lacks activator for a particular reporter gene. Gene for one of the interacting proteins called *bait* is fused to the gene coding for DNA binding domain (DBD) of the activator and specifies synthesis of fusion protein made up of DNA-binding domain and test protein. The recombinant yeast strain is co-transformed with a second construct carrying gene, for the second interacting protein called *prey* and gene for activation domain (AD). The reporter gene (such as *lacZ*) is expressed only in transformed cells in which *bait* and *prey* genes code for proteins that interact with each other. No gene expression will be seen if the *bait* and *prey* do not interact with each other.

In the first yeast two-hybrid system, protein-protein interactions were tested by fusing one test protein to the DNA-binding domain of the yeast GAL4 transcription factor, and the second protein to the GAL4 activation domain. The fusion proteins were expressed in a suitable yeast strain and the interaction detected by assaying for expression of a GAL4 responsive reporter gene. The most commonly used systems are the GAL4 system (in which the DNA-binding and activation domains of the yeast GAL4 protein are used) and the LexA system (DNA-binding domain of the bacterial repressor protein LexA used in combination with the *Escherichia coli* B42 activation domain).

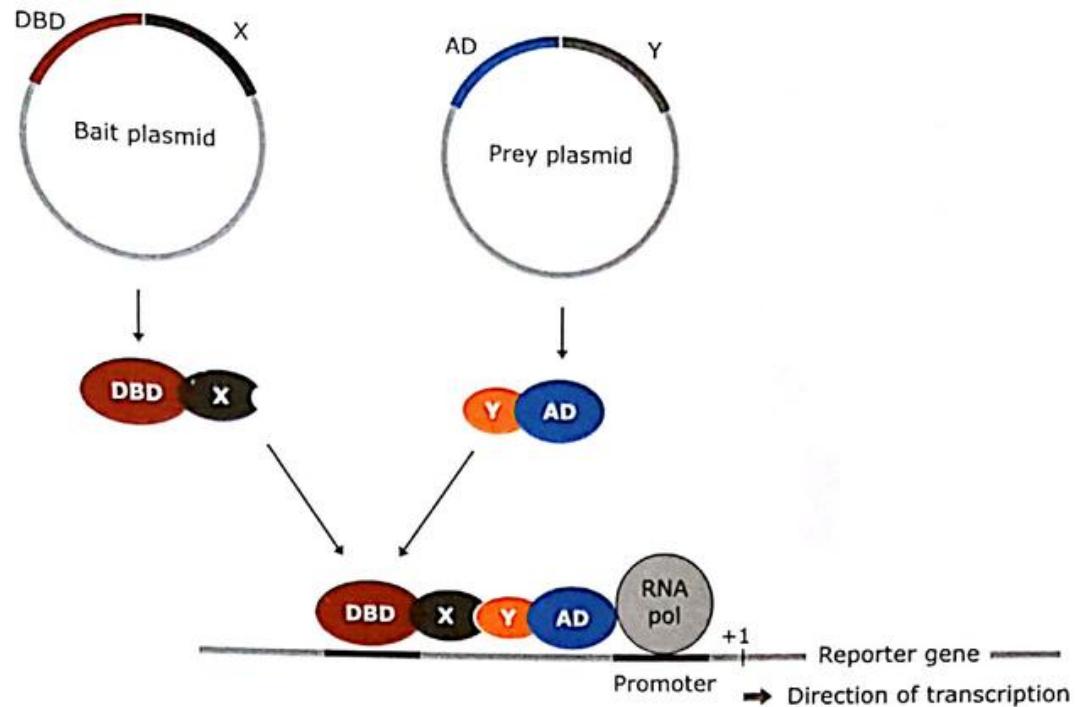


Figure 12.22 Yeast two-hybrid assay. Using the yeast two-hybrid system, the protein of interest (X) is expressed as a fusion protein to the DBD (DBD-X) and the AD is fused to the second protein of interest, Y as AD-Y. The AD-Y fusion vector is introduced into a yeast strain containing the DBD-X fusion partner by transformation. Only if proteins X and Y physically interact with each other, the DBD and AD will be brought together to activate expression of the downstream reporter gene.

A range of modifications to this basic principle have been made to expand the range of molecular interactions that can be identified. The yeast one-hybrid system screens for proteins that bind to DNA; for instance, to identify factors that bind specific gene promoter sequences. In addition, a variety of yeast three-hybrid systems exist. One such system assesses tertiary protein complexes and can be used to identify proteins that act as a 'bridge' between two other proteins, or that stabilize the interaction of proteins that would otherwise only interact very weakly. Another three-hybrid system assesses RNA–protein interactions, which identifies proteins that bind to a specific RNA species.

The yeast two-hybrid system characterizes the novel protein-protein interaction. The two-hybrid system uses the bi-functional nature of transcription factors to allow protein-protein interactions to be monitored through changes in transcription of reporter genes. Once a positive interaction has been identified, either of the interacting proteins can mutate (either by site-specific or randomly introduced changes) to produce proteins with a decreased ability to interact. Mutants generated using this strategy are very powerful in defining the residues involved in the interaction. Such techniques are termed reverse two-hybrid system.

12.7 Transcript analysis

A number of protocols are available for the analysis of mRNA transcripts: Northern blotting, S1 nuclease mapping and prime extension analysis. These protocols vary in the degree of sensitivity of detection and the information generated. In Northern blotting, an RNA extract is electrophoresed in an agarose gel, using denaturing electrophoresis buffer to ensure that the RNAs do not form inter- or intramolecular base pairs (because base pairing affects the rate at which the molecules migrate through the gel). After electrophoresis, the gel is blotted onto a nylon or nitrocellulose membrane and hybridized with a labeled probe. Once a

transcript has been identified, cDNA synthesis can be used to convert it into a dsDNA copy, which can be cloned and sequenced. Comparison between the sequence of the cDNA and the sequence of its gene will reveal the position of introns and possibly the start and end point of the transcripts.

S1 nuclease mapping

S1 nuclease mapping is used to detect 5'-end of mRNA on DNA template. This technique is very sensitive and can be used for quantification of the amount of mRNA. In S1 nuclease mapping, mRNA is hybridized to ssDNA that overlaps the start of the target transcript. The resulting RNA/DNA hybrid molecule has an overlap of DNA at the 3'-end that is digested by the single-strand specific S1 nuclease. The size of the processed ssDNA molecule, which is labeled at its 5'-end, is determined by denaturing polyacrylamide gel electrophoresis.

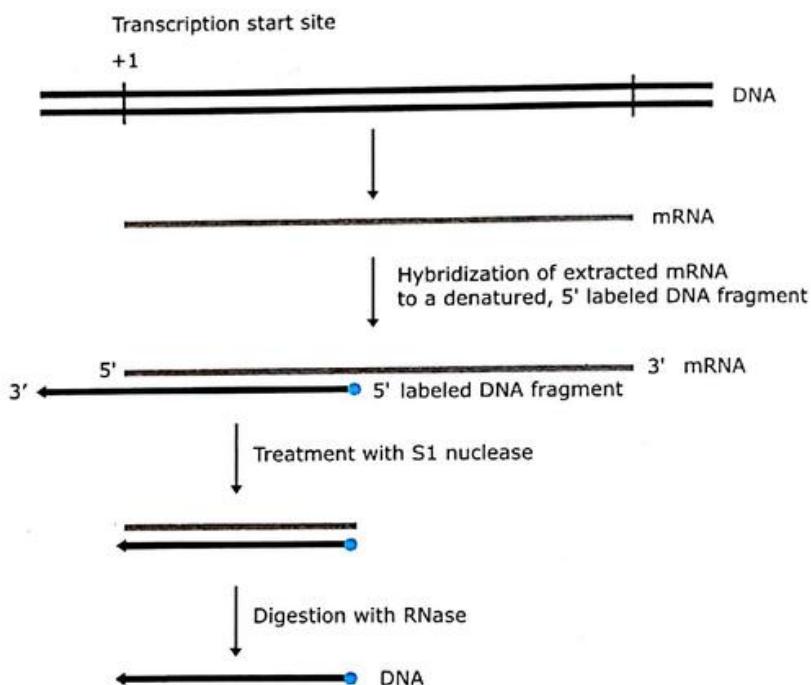


Figure 12.23 In S1 nuclease mapping, RNA is hybridized to the denatured DNA fragment that has been labeled at its 5'-end. The DNA fragment is chosen so that its 5'-end is internal to the target mRNA, while the 3'-end extends beyond the putative mRNA start point. The RNA-DNA hybrid molecule has single-stranded extensions that are degraded by the single-stranded specific S1 nuclease. The 3'-end of the DNA fragment is finally determined.

Primer extension

In the primer extension assay, the transcription start site for a gene is determined experimentally by identifying the 5'-end of the encoded messenger RNA (mRNA).

This approach involves binding of a labeled primer, usually a synthetic oligonucleotide of about 20 residues, that is complementary to an mRNA sequence downstream of the anticipated 5'-end. The primer is then extended by reverse transcriptase, thus synthesizing DNA that is complementary to the mRNA. The 3'-end of this newly synthesized strand of DNA, therefore, corresponds with the 5' terminus of the transcript. The resulting radiolabeled cDNA products are analyzed by denaturing polyacrylamide gel electrophoresis, followed by autoradiography. The sizes of the bands detected on the gel, as compared to an adjacent sequencing ladder or molecular weight standards, provide a measure of the distance from

the 5'-end of the synthetic oligonucleotide to the beginning of the mRNA transcripts. In theory, the 3'-end of the cDNA will coincide with the 5'-end of the mRNA. Thus, the size of the radiolabeled cDNAs should represent the distance from the labeled 5'-end of the primer to the 5'-end of the mRNA (i.e. the 3'-end of the cDNA).

Primer extension analysis has two main applications. First, it is used for mapping the 5'-end of transcripts. This allows one to determine the transcription initiation site. Second, it can be used to quantify the amount of transcript in an *in vitro* transcription system.

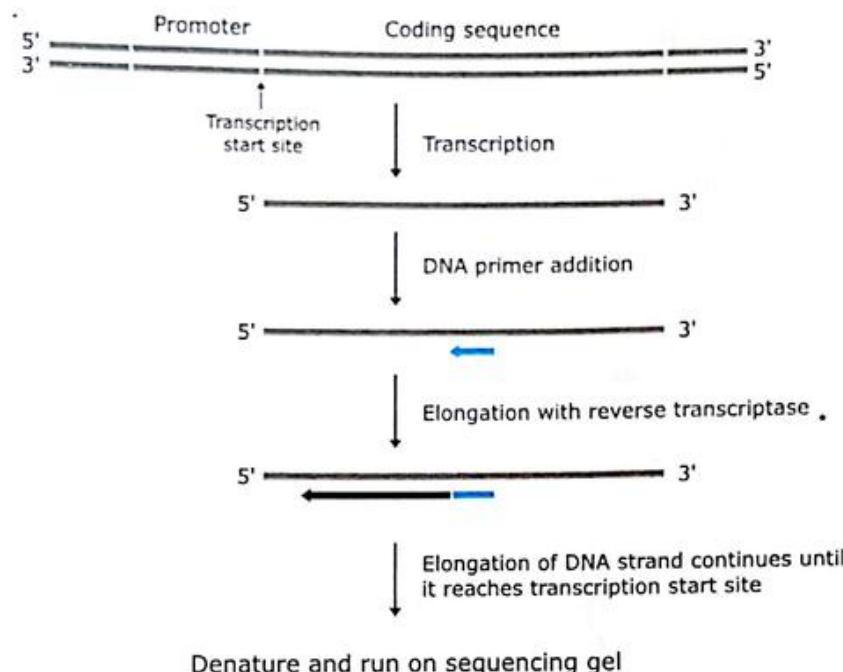


Figure 12.24 A primer extension assay involves four steps: First, selection and preparation of a labeled primer complementary to the RNA transcript of interest. Second, hybridization of the primer complementary to a region of the RNA under study. Third, extension from the primer that is catalyzed by reverse transcriptase using RNA as the template to synthesize a cDNA strand. Fourth, analysis of the extended cDNA products on denaturing polyacrylamide gels and autoradiography.

12.8 DNA microarray

Although all of the cells in the human body contain identical genetic material, but the same genes are not active in every cell. Studying which genes are active and which are inactive in different cell types helps scientists to understand both how these cells function normally and how they are affected when various genes do not perform properly. In the past, scientists were able to conduct these genetic analyses on a few genes at once. With the development of DNA microarray technology, however, scientists can now examine how thousands of genes express and their product interact at any given time.

A DNA microarray (also commonly known as *gene chip*, *DNA chip*, *bioarray* or *gene array*) is an orderly arrangement of thousands of identified genes fixed on a solid support, usually glass, silicon chips or nylon membrane for the purpose of expression profiling, monitoring expression levels for thousands of genes simultaneously. In expression profiling, it detects the presence and abundance of labeled nucleic acids in a biological sample, which will hybridize to the DNA on the array via Watson-Crick base pairing, and which can be detected via the label. The most well known use of DNA microarrays is for profiling mRNA levels; however, DNA microarrays have also been used to detect DNA-protein interactions (e.g. transcription factor-

binding site and transcription factor), alternatively spliced variants, the epigenetic status of the genome (such as methylation patterns), DNA copy number changes and sequence polymorphisms.

In DNA microarray, a large number of DNA probes, each one with a different sequence, are immobilized at defined positions on a solid surface (such as glass, silicon chips or nylon membrane). The probe can be *synthetic oligonucleotides*, short DNA molecules, such as cDNA or PCR products. These can be spotted onto a glass microscope slide or a piece of nylon membrane (*low density arrays*) or on the surface of a wafer of silicon (*high density array*). High-density microarrays may have up to 10^6 spots in a $1\text{--}2\text{ cm}^2$ area.

An array is an orderly arrangement of samples. In general, arrays are described as *macroarrays* or *microarrays*, the difference being the size of the sample spots. Macroarrays contain sample spot sizes of about 300 microns or larger. The sample spot sizes in microarray are typically less than 200 microns in diameter and the arrays usually contain thousands of spots. Usually, a single DNA microarray slide/chip may contain thousands of spots each representing a single gene and collectively the entire genome of an organism.

Types of DNA microarrays

DNA microarrays can be classified into two broad categories depending on whether they are manufactured using presynthesized DNA sequences or *in situ* synthesis methods.

1. Glass DNA microarrays : It involves the micro spotting of pre-fabricated cDNA fragments on a glass slide. This DNA microarray was the first type of DNA microarray technology developed (pioneered by Patrick Brown and his colleagues at Stanford University). It is fabricated by either pen tip deposition (*spotted arrays*) or inkjet deposition (*sprayed arrays*).
2. High-density oligonucleotide microarrays often referred to as a 'chip' which involves *in situ* oligonucleotide synthesis. The most prominent microarrays with *in situ* synthesized probes are the 'GeneChips' manufactured by Affymetrix. They are produced by chemical synthesis of the oligonucleotides directly on the coated quartz surface of the array. This technology allows very high spot densities. Therefore, they are called *high-density oligonucleotide arrays*. GeneChips are produced in a unique photolithographic process in combination with chemical reactions developed for combinatorial chemistry. A quartz wafer is coated with a narrow layer of a light sensitive compound. This light-directed synthesis has enabled the large-scale manufacture of arrays containing hundreds of thousands of oligonucleotide probe sequences on glass slides or 'chips,' less than 2 cm^2 in size.

How does DNA microarray work?

Standard Watson-Crick base-pairing (i.e. A-T and G-C for DNA; A-U and G-C for RNA) between *target* and *probe* is the underlining principle of DNA microarray. Microarray technology evolved from Southern blotting, whereby fragmented DNA is attached to a substrate and then probed with a known gene or fragment.

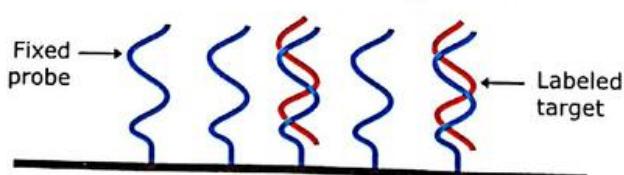


Figure 12.25 Hybridization of the target to the probe.

In a gene expression profiling experiment, the expression levels of thousands of genes are simultaneously monitored. There are four laboratory steps in using a microarray to measure gene expression in a sample:

1. Sample preparation and labeling
2. Hybridization
3. Washing
4. Image acquisition

To determine which genes are turned on and which are turned off in a given cell, a researcher must first collect the mRNA molecules present in that cell. The researcher then labels each mRNA molecule by attaching a fluorescent dye. Next, the researcher places the labeled mRNA onto a DNA microarray slide. The mRNA that was present in the cell will then hybridize - or bind - to its complementary DNA on the microarray, leaving its fluorescent tag. A researcher must then use a special scanner to measure the fluorescent areas on the microarray.

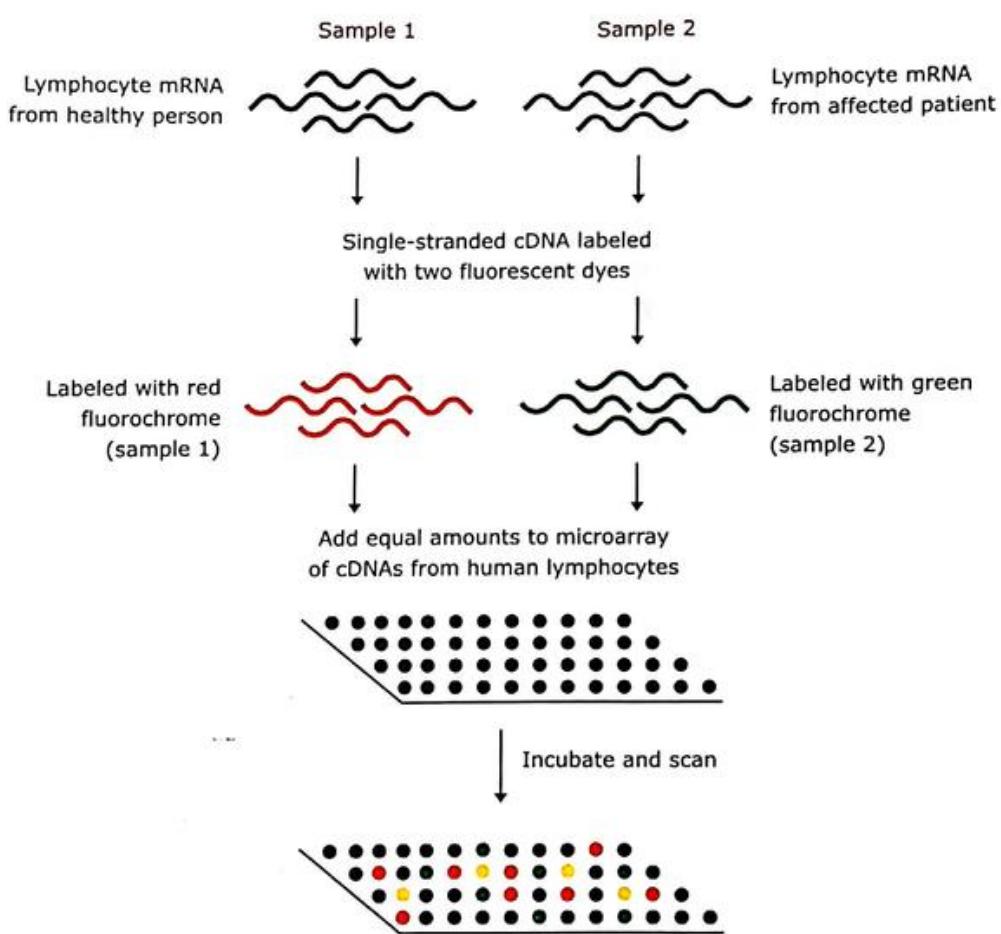


Figure 12.26 DNA microarray. To prepare the microarray, DNA fragments—each corresponding to a gene—are spotted onto a slide. In this example, mRNA is collected from two different lymphocytes for a direct comparison of their relative levels of gene expression. These samples are converted to cDNA and labeled, one with a red fluorochrome, the other with a green fluorochrome. The labeled samples are mixed and then allowed to hybridize on the microarray. After incubation, the array is washed and the fluorescence scanned. Red spots indicate that the gene in sample 1 is expressed at a higher level than the corresponding gene in sample 2. Green spots indicate that expression of the gene is higher in sample 2 than in sample 1. Yellow spots reveal genes that are expressed at equal levels in both samples. Black spots indicate little or no expression in either sample.

If a particular gene is very active, it produces many molecules of mRNA, which hybridize to the DNA on the microarray and generate a very bright fluorescent area. Genes that are comparatively less active produce fewer mRNAs, which result in dimmer fluorescent spots. If there is no fluorescence, none of the messenger molecules have hybridized to the DNA, indicating that the gene is inactive. Researchers frequently use this technique to examine the activity of various genes at different times.

DNA hybridization analysis on microarrays usually involves detecting the signal generated by the binding of a reporter probe (fluorescent, chemiluminescent, colorimetric, radioisotope, etc.) to the target DNA sequence. The microarray is scanned or imaged to obtain the complete hybridization pattern.

Principal applications of microarrays

- *Investigating cellular states and process.* Patterns of expression that change with the cellular state or growth conditions can give clues to the mechanisms of processes such as sporulation, or the change from aerobic to anaerobic metabolism.
- *Diagnosis of disease.* Testing for the presence of mutation can confirm the diagnosis of a suspected genetic disease, including detection of a late-onset condition such as Huntington disease.
- *Drug selection.* Allows detection of genetic factors that govern responses to drugs, that in some patients render treatment ineffective and in other cause unusual serious adverse reactions.
- *Specialized diagnosis of disease.* Different types of leukemia can be identified from different patterns of gene expression.
- *Pathogen resistance.* Comparisons of genotypes of expression patterns, between bacterial strains susceptible and resistant to an antibiotic, point to the proteins involved in the mechanism of resistance.
- *Investigating cellular states in responses to pathogen infection and environmental change.*
- *Investigating cellular states during the cell cycle.*

12.9 Electrophoretic mobility shift assay

An electrophoretic mobility shift assay (EMSA) or mobility shift electrophoresis, also referred to as a gel shift assay, gel mobility shift assay, band shift assay, or gel retardation assay, is a common affinity electrophoresis technique used to study protein-DNA or protein-RNA interactions. It is a rapid and sensitive method to detect protein-nucleic acid interactions.

How far a DNA molecule migrates during gel electrophoresis varies with size: the smaller the molecule the more easily it moves through the gel. In addition, if a given DNA molecule has a protein bound to it, migration of that DNA-protein complex through the gel is retarded compared to migration of the unbound DNA molecule. This forms the basis of an assay to detect specific DNA binding activities. EMSA is based on the observation that the electrophoretic mobility of a protein-nucleic acid complex is typically less than that of the free nucleic acid in the gel matrix. This process tests the ability of a protein to bind a DNA fragment as it migrates through a non-denaturing gel under the influence of an electric current. Binding of the protein will reduce the mobility of the DNA. If the fragment travels further in the absence of a protein than it does in its presence, one can conclude that the protein has an affinity for that piece of DNA. Different proteins may retard a fragment to different extents,

depending on their size and shape. If two proteins bind the same piece of DNA, they will reduce its mobility even further.

The mobility shift assay has a number of strengths. The basic technique is simple to perform, yet it is robust enough to accommodate a wide range of binding conditions. Using radioisotope-labeled nucleic acids, the assay is highly sensitive, allowing assays to be performed with small protein and nucleic acid concentrations and small sample volumes. It is very popular because it is easy, quick, versatile and very sensitive.

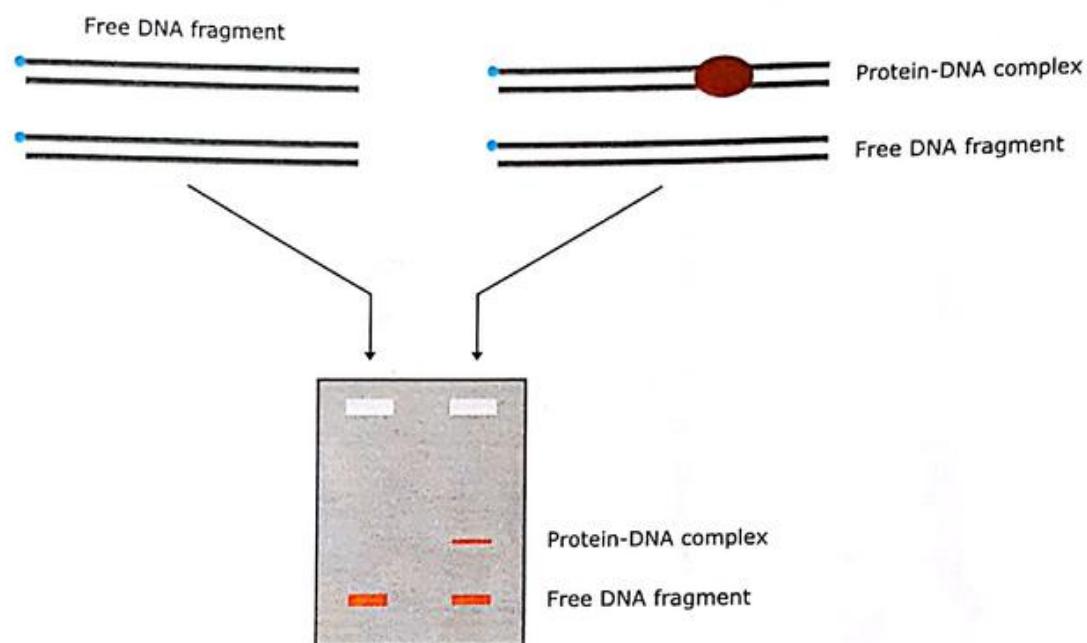


Figure 12.27 Gel mobility shift assay. The effect of the proteins binding on the mobility of the DNA fragment is analyzed by polyacrylamide-gel electrophoresis followed by autoradiography. The free DNA fragments migrate rapidly to the bottom of the gel, while those fragments bound to proteins are retarded. In the figure, a protein is mixed with radiolabeled DNA containing a binding site for that protein. The mixture is resolved by polyacrylamide gel electrophoresis and visualized using autoradiography. DNA not mixed with protein runs as a single band corresponding to the size of the DNA fragment (left lane). In the mixture with the protein, a proportion of the DNA molecules (but not all of them at the concentrations used) binds the DNA molecule. Thus, in the right-hand lane, there is a band corresponding to free DNA and another corresponding to the DNA fragment in complex with the protein.

12.10 Footprinting assay

Gel retardation assay gives a general indication about the protein binding with a DNA fragment. However, it does not reveal where a protein is binding within a particular DNA fragment. Footprinting assay allows us to identify the sites, where protein binds. It is of two main types: protection and interference. The *protection footprinting* consists of identifying the sites of protection from – cleavage of phosphodiester bonds or modification of nucleotide – of the DNA by its interaction with the protein, whereas the *interference footprinting* identifies the DNA sites essential to the interaction by the fact that their modification inhibits protein binding.

Nuclease protection footprinting

DNase I footprinting assay is a method of studying *DNA-protein interaction* and identifying the DNA sequence to which a protein binds. It measures the ability of a protein to protect a

radiolabeled DNA fragment against digestion by DNase I, which is an endonuclease with little sequence specificity that will cut almost anywhere within a DNA molecule. DNase I treatment is carried out in the presence of a manganese salt, which induces the enzyme to make random, double strand cuts in the target molecules, leaving blunt-ended fragments.

The nuclease treatment is carried out under limiting conditions, such as a low temperature and/or very little enzyme, so that on average, each copy of the DNA fragment suffers a single 'hit'—meaning that it is cleaved at just one position along its length. This generates a population of fragments of all possible sizes that appear as a continuous ladder when resolved according to size on a denaturing gel. If a protein (suppose X) binds the DNA at a particular site, then this protects that site against digestion by DNase I; as a consequence, a gap or *footprint* appears in the ladder of DNA molecules. The position of the footprint corresponds to the positions where on the DNA molecule the protein X is bound. Chemical cleavage agents are sometimes used for footprinting instead of DNase I.

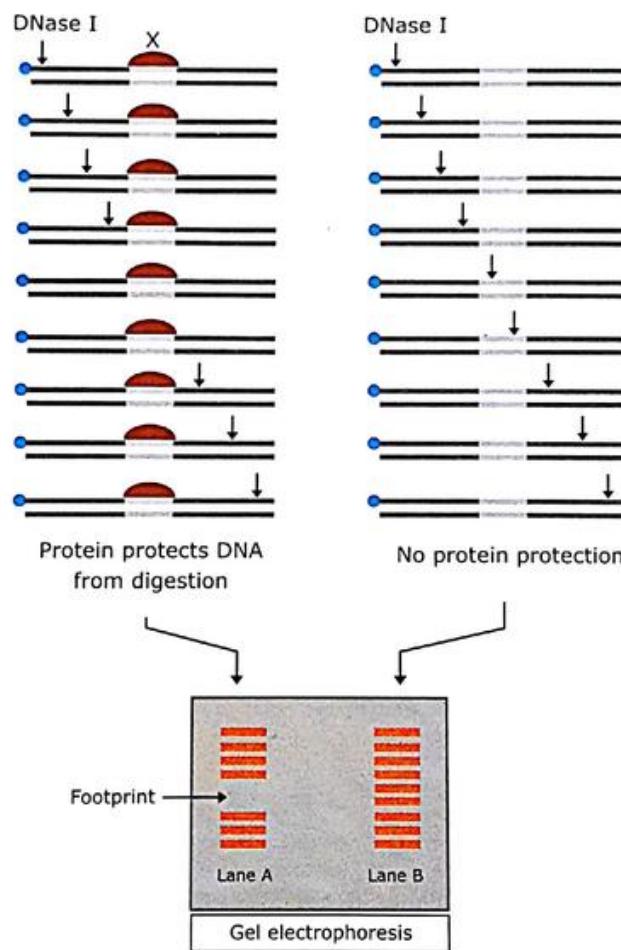


Figure 12.28 DNase I footprinting assay. X represents the protein having specific binding site on DNA. During the footprinting reaction, a sample of the DNA fragment containing the binding site is mixed with purified protein (X) and DNase I. The DNase I cleaves the DNA randomly. The amount of DNase I is controlled so that each DNA fragment is only digested once. Since DNase does not cut the DNA where the protein is bound, sequence of DNA fragments bound with protein remain protected from DNase I action. The samples are run on a sequencing gel to separate the fragments. When lane B (no protein) and lane A (plus protein) are compared, it can be seen that lane A shows no bands in the boxed region. Therefore, the protein binds to DNA in this region and protects it from cutting by the DNase. Alignment with a sequencing ladder (lane B) allows the precise region of binding to be deduced.

Modification protection footprinting

Modification protection footprinting is similar to DNase I footprinting. The basis of this technique is that if a DNA molecule carries a bound protein, then part of its nucleotide sequence will be protected from modification. In this case instead of DNase I digestion, the fragments are treated with limited amounts of base modifying agent, such as dimethyl sulfate (DMS) which adds methyl groups to G nucleotides. If a protein binds the DNA at a particular site, then this protects guanines at that site from the action from dimethyl sulfate. In DMS modification protection assay, the fragments are treated with limited amounts of DMS so that a single G nucleotide is methylated in each fragment.

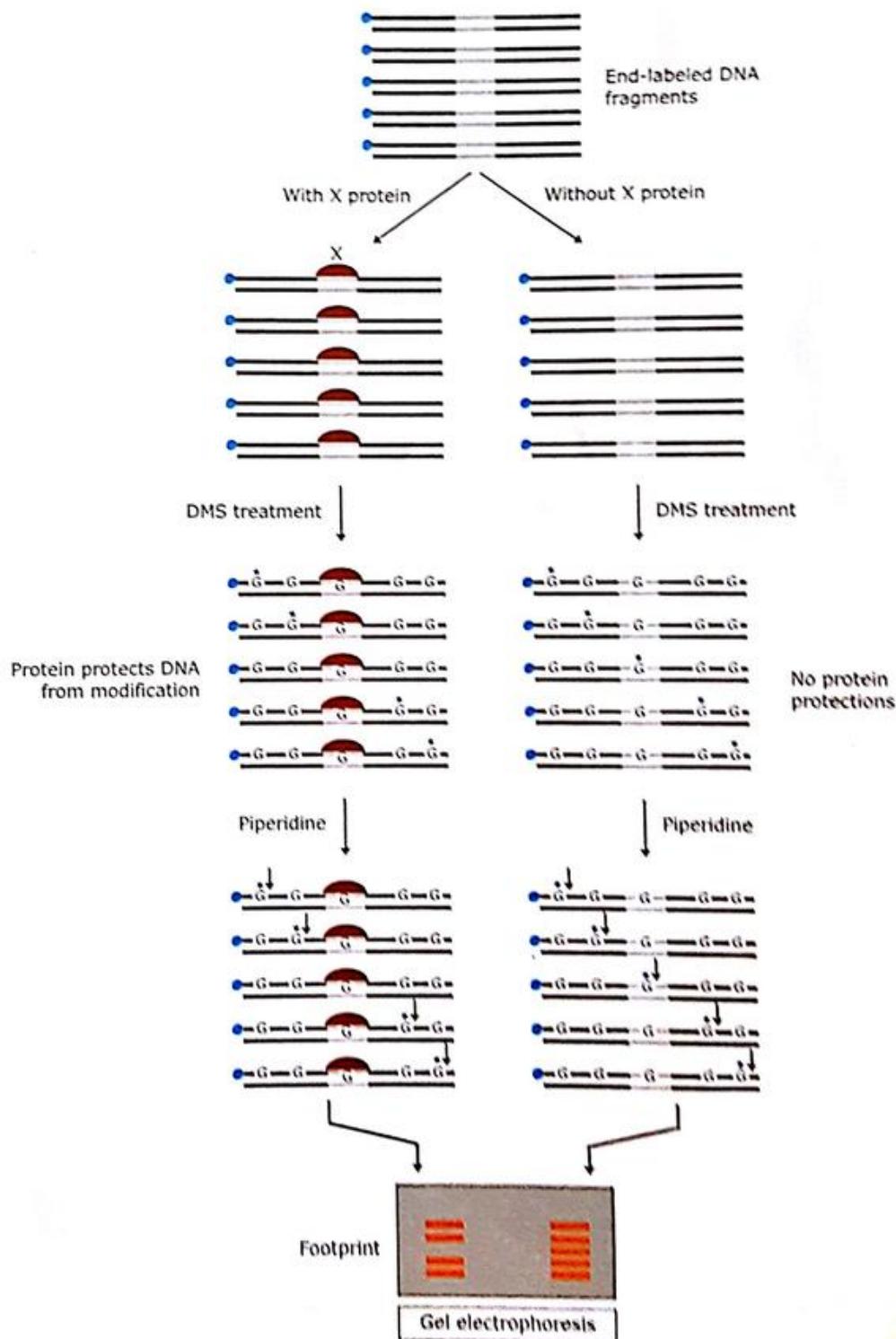


Figure 12.29 The modification protection footprinting experiment with dimethylsulfate (DMS) as the probe. The black circle denotes a radioactive end-label which is required for visualization of the final products on the gel. A portion of the DNA is combined with the protein and other portion is kept free from protein. Both are then treated with the DMS, which modifies G nucleotide on the DNA. The region of the DNA bound with protein remain protected from modification with the DMS. The DNA is, then, cleaved at the sites of modification by treatment with piperidine. The end-labeled fragments of the free and protein-bound DNA are then compared by separation on a denaturing sequencing gel, and visualized with an autoradiogram.

After removal of the protein, the DNA is treated with piperidine, which cleaves the phosphodiester bond at the modified nucleotide positions. DNA fragments are then resolved according to size by gel electrophoresis. Piperidine only cuts the strand that is modified, rather than making a double-stranded cut. The samples are, therefore, examined by *denaturing gel* electrophoresis so that the two strands are separated. The resulting autoradiograph shows the sizes of the strands that have one labeled end and one end created by piperidine nicking. The banding pattern for the control DNA strands - those not incubated with the proteins - indicates the positions of G nucleotide in the DNA fragment, and the footprint.

Modification interference assay footprinting

Modification interference is different from modification protection. Modification interference works on the basis that if a nucleotide critical for protein binding is altered, for example by addition of a methyl group, then binding may be prevented. In this case, the DNA is exposed to the modifying agent prior to complexing with the protein. The subsequent binding reaction yields populations of free DNA and DNA bound to protein, defined by whether or not the modification interferes with the binding.

12.11 Site-directed mutagenesis

Site-directed mutagenesis is an *in vitro* method for creating a specific mutation in a known sequence. It is a fundamental tool that is widely used in molecular biology and protein engineering. Because it allows the researcher to design selective changes to the DNA, site-directed mutagenesis provides a powerful method for probing gene regulation as well as the relationship between protein structure and function. The functional and structural roles of amino acid residues in a protein of interest can be studied by comparing the mutant protein carrying changes in amino acid residues to the wild-type protein.

Conventional mutagenesis is a random process that introduces changes at unspecified positions in a DNA molecule. Hence, screening of large numbers of mutated organisms is necessary to find a mutation of interest. Unlike natural mutations or random mutagenesis, mutations constructed by genetic engineering do not require extensive screening.

The methods utilized for *in vitro* site-directed mutagenesis are generally of two types: polymerase chain reaction (PCR)-based and non-PCR-based. The PCR-based site-directed mutagenesis methods are used more frequently than the non-PCR-based methods.

Non-PCR-based site-directed mutagenesis

Cassette mutagenesis

In cassette mutagenesis, a restriction fragment from the cloned DNA of interest is replaced by another restriction fragment containing the mutated sequence. It relies on the presence of two restriction enzyme recognition sites flanking the DNA that is to be mutated. With the help of restriction enzymes, the target site in the wild-type gene is removed. A double-stranded oligonucleotide cassette containing the desired mutant sequence is used to replace the corresponding sequence in the wild-type gene. These are relatively large-scale changes. It is a simple method and the efficiency of mutagenesis is very high. The disadvantages are the requirement for unique restriction sites flanking the region of interest.

Design of mutagenic oligonucleotides

A crucial step in site-directed mutagenesis is the design of the mutagenic oligonucleotide. By definition, mutagenic oligonucleotides must contain at least one base change, but they may incorporate far more complicated mutations including insertions, deletions, and multiple substitutions. The minimum length of the mutagenic oligonucleotide is defined by the complexity of the mutation. Simple single-base substitutions can be accomplished with oligonucleotides ~25 bases in length. More complicated mutations may require oligonucleotides 80 bases or more in length.

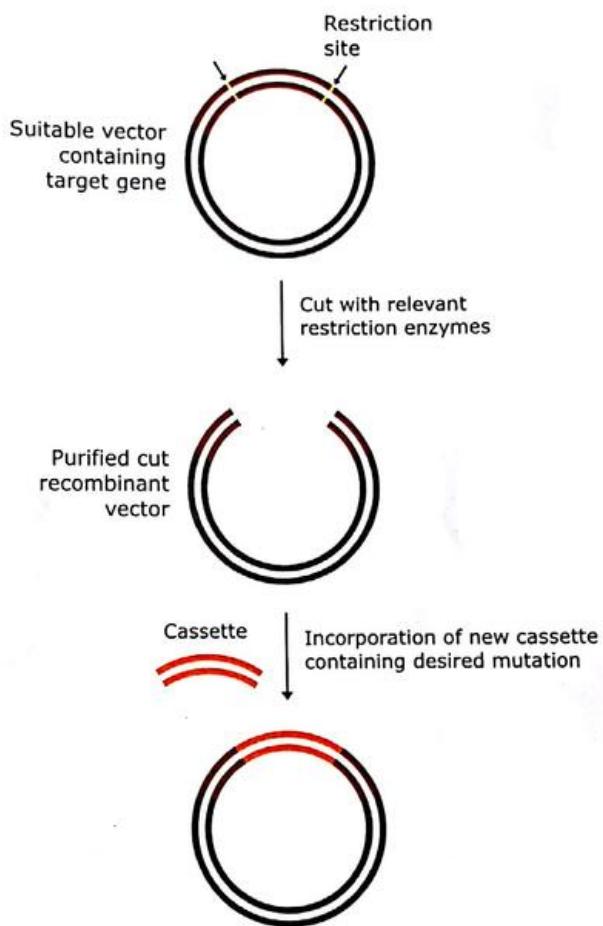


Figure 12.30 Cassette mutagenesis. The gene contained in a suitable vector is cleaved with two restriction enzymes. This releases a small section of DNA from the gene. These restriction enzymes cut at sites that flank the area of DNA to be changed. In this way, only the desired cutting occurs. A synthetic duplex is then ligated in place of the released cassette. This technique can be used to make single or multiple changes. The changes are only limited by the available size of the synthetic DNA cassette.

Primer extension mutagenesis

The simplest method of site-directed mutagenesis is the primer extension mutagenesis (*oligonucleotide directed mutagenesis*). The method involves DNA synthesis with a chemically synthesized oligonucleotide (7-20 nucleotides long) that carries a base mismatch with the complementary sequence. As shown in figure 12.31, the method requires that the DNA to be mutated is available in single-stranded form. A single-stranded version of the gene is obtained by cloning in an M13 vector. The synthetic oligonucleotide which acts as primer, primes DNA synthesis and is itself incorporated into the resulting heteroduplex molecule.

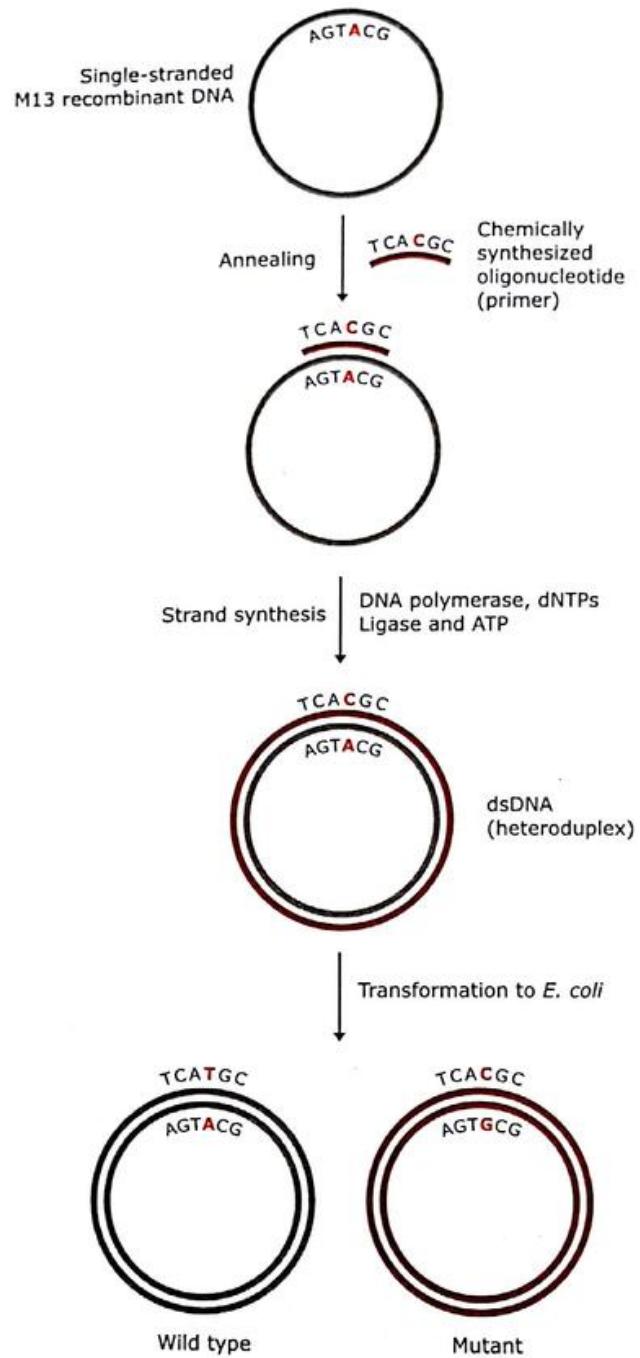


Figure 12.31 Primer extension mutagenesis. Step 1, an oligonucleotide primer annealing to template DNA. The gene or the DNA fragment is inserted in an M13 phage vector. Step 2, Primer extension by DNA polymerase and synthesis of second strand. Step 3, Sealing of nick by ligase. DNA ligase seals the nick and a completely closed, circular double-stranded DNA molecule is produced. One strand of this molecule contains the original sequence while the other strand contains the mutated sequence, that is, a heteroduplex. Step 4, Transformation to *E. coli*. The dsDNA is transformed into *E. coli* cells. In *E. coli*, subsequent replication will produce some double-stranded circular DNA molecules with wild-type sequences and some molecules with mutated sequences. Step 5, Screening of phages with mutant sequence. Bacteriophages containing either the wild-type or the mutant sequence can be distinguished from each other through hybridization screening.

The newly formed heteroduplex is used to transform cells. After introduction into *Escherichia coli*, DNA replication produces numerous copies of this recombinant DNA molecule. Half of the resulting double-stranded molecules are copies of the original strand of DNA, and the remaining are copies of the strand that contains the mutated sequence.

A variation of the procedure outlined above involves oligonucleotides containing inserted or deleted sequences. As long as stable hybrids are formed with single-stranded wild-type DNA, priming of *in vitro* DNA synthesis can occur, ultimately giving rise to clones corresponding to the inserted or deleted sequence.

PCR based site-directed mutagenesis

PCR approaches have become the method of choice to generate arrays of predefined mutations within the gene of interest. There are numerous PCR-based approaches to DNA mutagenesis. PCR-mediated nucleotide changes, deletions, or insertions can be accomplished by performing PCR synthesis reactions with carefully chosen or modified PCR reaction components.

One strategy for this kind of PCR mutagenesis is based on the principle of 'mispriming'. Because mismatches between templates and primers are tolerated under certain PCR conditions, primers can be designed to include predefined changes (so-called *mutagenic primers*). Other strategies make use of the built-in high error rate of *Taq* DNA polymerase in the PCR reaction. In general, mutagenic primer PCR approaches are used for introducing site-directed mutagenesis into the genes of interest, whereas *Taq* DNA polymerase or base analog PCR approaches are useful in creating random and extensive mutagenesis in the target gene.

Overlap extension method

A specialized technique developed for the introduction of mutations into the center of a PCR fragment is known as the *overlap extension method*. In this technique, primers are designed for the synthesis of two different PCR fragments with a region of common sequence. Each reaction uses one flanking primer and one internal primer containing the desired mutation. The two overlapping pieces of DNA are then annealed to generate a heteroduplex that can be extended by *Taq* polymerase and amplified by subsequent rounds of PCR to produce a full-length mutant segment. By varying the templates and primers, both mutant and chimeric DNA can be generated. The technique is applicable to the introduction of site-specific insertion and deletion mutations. It has also been modified to allow for the introduction of multiple mutations on the same template without loss of efficiency.

The overlap extension method can be modified to simplify the removal of excess primers and wild-type template after each PCR step. Biotinylated primers and streptavidin-coated magnetic beads are employed to purify PCR products from reactants. This approach improves the speed and efficiency of overlap extension mutagenesis.

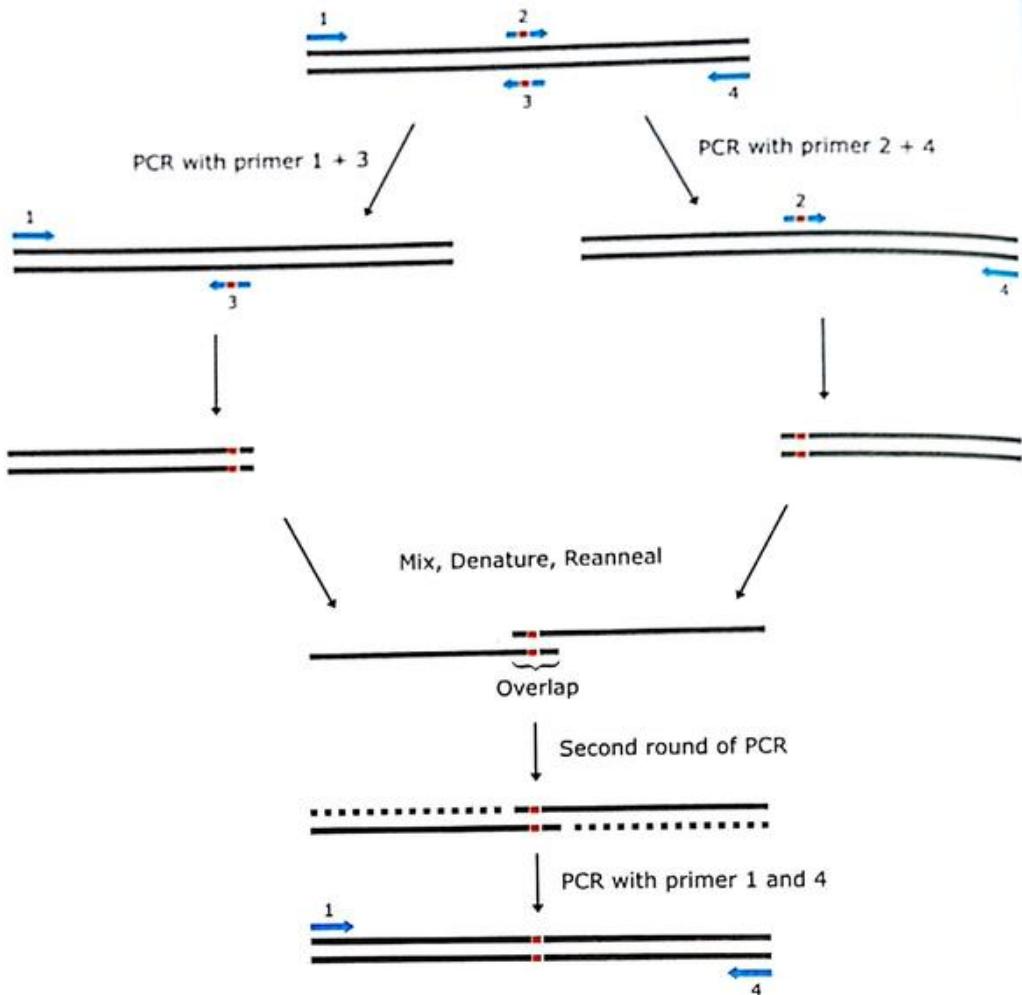


Figure 12.32 Overlap extension method of mutagenesis. Two outermost primers (1 and 4) are designed, which are complementary to the antisense strand and the sense strand of the target DNA, respectively. In addition, two internal primers 2 and 3 containing the desired mutations are taken, so that the required mutation can be introduced to each strand. Two PCR reactions are carried out in two separate tubes, but with the same reaction conditions (as shown in the figure). The products of the two reactions are then mixed, denatured, and re-annealed; some of the strands from the first PCR product will anneal with that obtained from the second reaction especially at the regions of overlap, that is, at the regions corresponding to the sequences of primers 2 and 3. The hybrid molecule containing a short 3'-end can act as a primer and can be extended in the next phase by DNA polymerase to form a complete double-stranded DNA molecule with mutation on both strands. The subsequent and final PCR carried out with flanking primers (primers 1 and 4) will then amplify the full-length product to yield the mutated DNA.

Megaprimer PCR

Although primers are typically short, single-stranded oligonucleotides that a long, double-stranded DNAs could also serve as primers. These larger primers, known as *megaprimers* can be used to introduce any combination of point mutations, deletions or insertions. The megaprimer method of mutagenesis uses three oligonucleotide primers (two flanking and one megaprimer) and two rounds of PCR performed on a DNA template. The first PCR uses one outside primer and the middle mutagenic primer to form a double-stranded product (the megaprimer) containing the desired mutations.

The amplified product is then used in a second round of PCR with wild-type template and the other flanking primer to create a fragment of the same length as the original target DNA.

containing the desired mutation. The key to this method is that the amplified product from the first round of PCR is used as a primer in the second round of PCR.

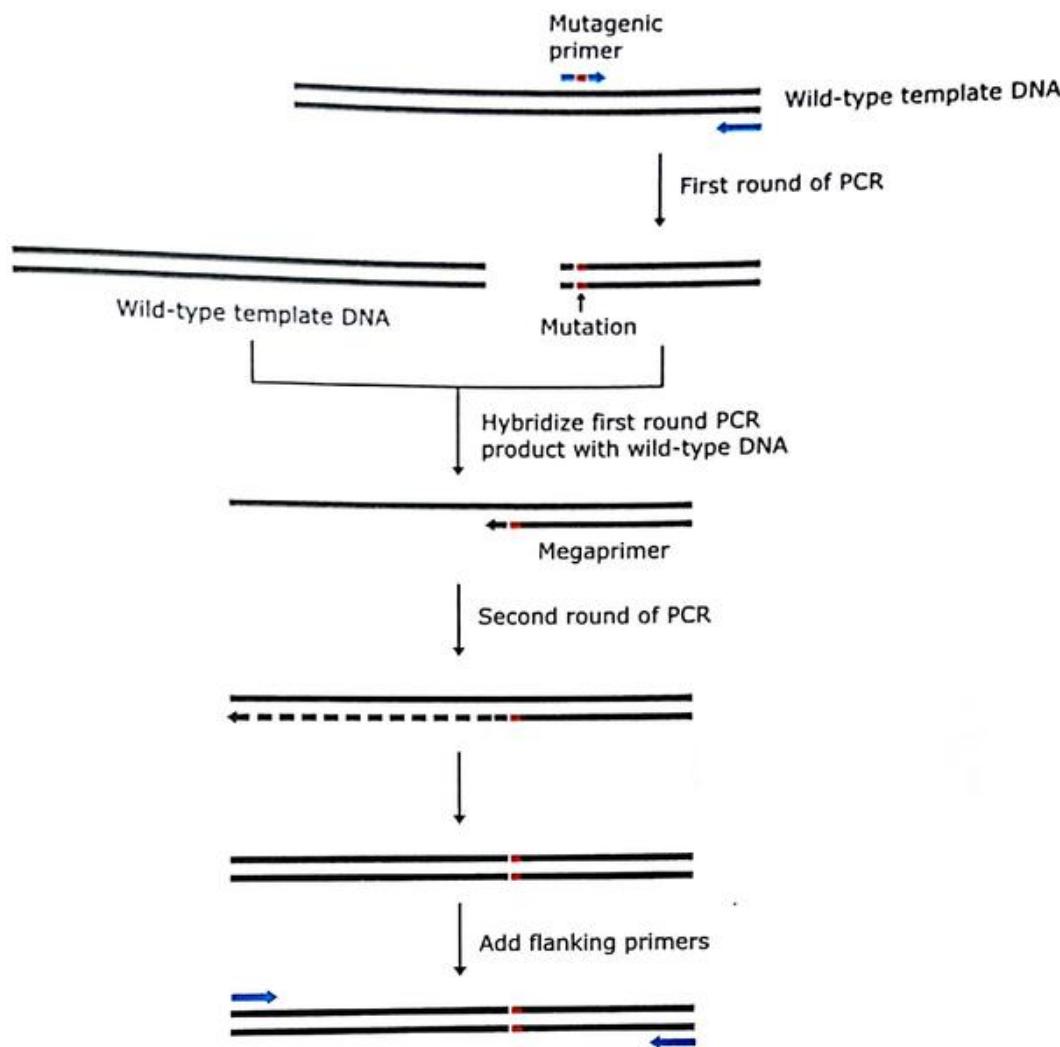


Figure 12.33 Schematic illustration of the megaprimer method for site-directed mutagenesis. The first round of PCR is used to make a fragment of the template DNA containing the desired mutation. This 'megaprimer' is then hybridized to the wild-type template DNA, and a second round of PCR is carried out, to generate the entire molecule with the mutation.

Inverse PCR method

The inverse PCR method uses only two primers to create the desired mutation. The key feature of this method is that in making the mutation, the entire vector is amplified. The two primers, one containing the desired mutation, extend on the circular template DNA in opposite directions. Amplification ultimately yields a linear, double-stranded DNA molecule containing the mutation at one end. Following amplification, the ends are ligated and the resulting circular DNA molecule is transformed into *E. coli*. There are a number of variations of this method that improve the efficiency of mutagenesis.

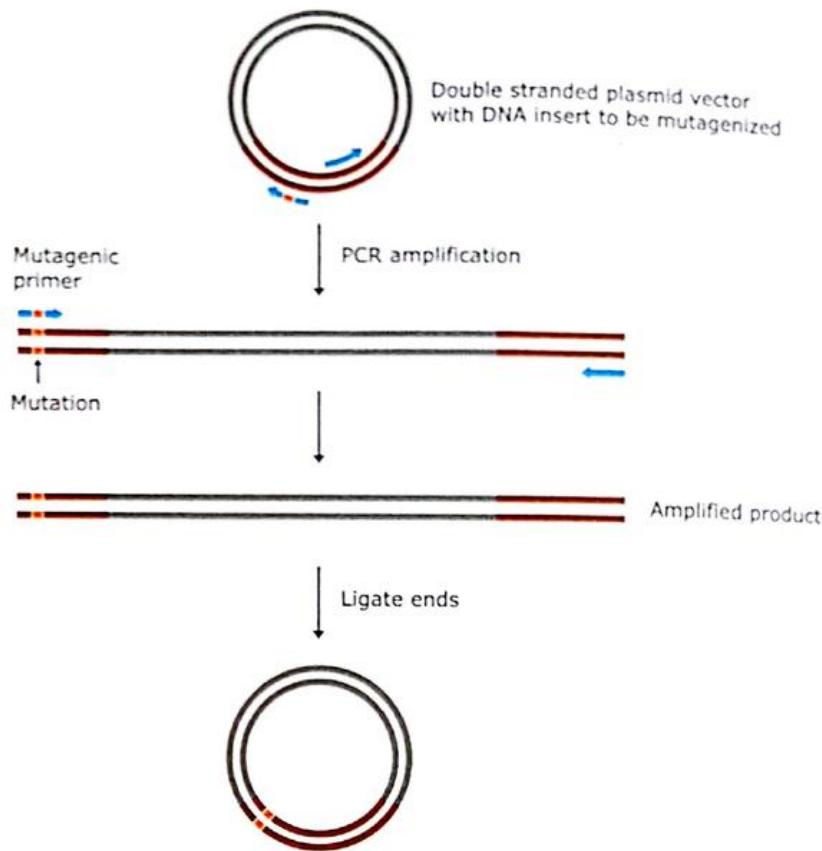


Figure 12.34 Inverse PCR method. A pair of primers is used, the 5' termini of which hybridize end-to-end on opposite strands of the plasmid DNA to be mutagenized. One of the primers contains the desired mutation. PCR amplification of the entire double-stranded plasmid results in a linear PCR amplicon that can be circularized by ligation.

Random and extensive mutagenesis

All of the approaches discussed so far are designed to create either single or a few site-directed mutations. Sometimes, it is desirable to obtain random and extensive mutations in the gene of interest or to generate a library of such molecules. There are two basic approaches to achieve these goals, and both are based on selected uses of PCR reaction components. The first approach makes use of error-prone DNA polymerases. Certain thermostable DNA polymerases, like *Taq* DNA polymerase, have an intrinsic error rate due to the lack of a 3'-5' exonuclease activity. Mismatch passes by the polymerase during PCR allows the possibility of mutations. The cumulative error rate can be substantial. This tendency is further enhanced by other factors such as buffer composition (e.g. high-magnesium concentration, or high pH) and other experimental conditions (e.g. a large amount of polymerase, a great number of cycles, a low-annealing temperature).

The second approach to generate random and extensive mutations in the gene of interest is based on the base pairing property of degenerate base analog which can form base pairing with nucleotides A, C, G and T under normal reaction conditions. In the presence of degenerate base analog and a biased ratio of dNTPs, DNA polymerase tends to randomly incorporate a substantial amount of base analog in the newly synthesized DNA strand. This base analog-containing DNA can serve as a template in subsequent PCR amplifications and allows random base insertion sites. As a result, the final PCR products will have base substitution random sites.

12.12 DNA sequencing

The term DNA sequencing encompasses biochemical methods for determining the order of the nucleotide bases, adenine, guanine, cytosine and thymine, in a DNA molecule. The methodologies for DNA sequencing are as follows:

✓ Chain termination method

The enzymatic chain termination method (also dideoxy method) was developed by Frederick Sanger and coworkers in year 1977. The method is based on the DNA polymerase-dependent synthesis of a complementary DNA strand in the presence of natural deoxynucleotides (dNTPs) and dideoxynucleotides (ddNTPs) that serve as terminators. The DNA synthesis reaction is randomly terminated whenever a ddNTP is added to the growing oligonucleotide chain, resulting in truncated products of varying lengths with an appropriate ddNTP at their 3'-terminus. The products are separated by size using polyacrylamide gel electrophoresis and the terminal ddNTPs are used to reveal the DNA sequence of the template strand.

Chemical degradation method

The chemical degradation method was introduced by Allan Maxam and Walter Gilbert in year 1977. In this method, the sequence of a dsDNA molecule is determined by treatment with chemicals that cut the molecule at specific nucleotide positions.

Pyrosequencing method

The pyrosequencing method was developed by Mostafa Ronaghi and Pal Nyrén at the Royal Institute of Technology in Stockholm in 1996. In this method, the addition of a deoxynucleotide to the end of the growing strand is detectable because it is accompanied by the release of a flash of light.

Chain termination method

Chain termination method relies on the use of dideoxyribonucleoside triphosphates (ddNTP), derivatives of the normal deoxyribonucleoside triphosphates that lack the 3' hydroxyl group. Purified DNA is synthesized *in vitro* in a mixture that contains single-stranded molecules of the DNA to be sequenced, the enzyme DNA polymerase, a short primer DNA to enable the polymerase to start DNA synthesis, and the four deoxyribonucleoside triphosphates (dATP, dCTP, dGTP, dTTP: A, C, G and T). If a dideoxyribonucleotide analog of one of these nucleotides is also present in the nucleotide mixture, it can become incorporated into a growing DNA chain. Because this chain now lacks a 3' OH group, the addition of the next nucleotide is blocked, and the DNA chain terminates at that point.

To determine the complete sequence of a DNA fragment, the double-stranded DNA is first separated into its single strands and one of the strands is used as the template for sequencing. Four different chain-terminating dideoxyribonucleoside triphosphates (ddATP, ddCTP, ddGTP and ddTTP) are used in four separate DNA synthesis reactions on copies of the same single-stranded DNA template.

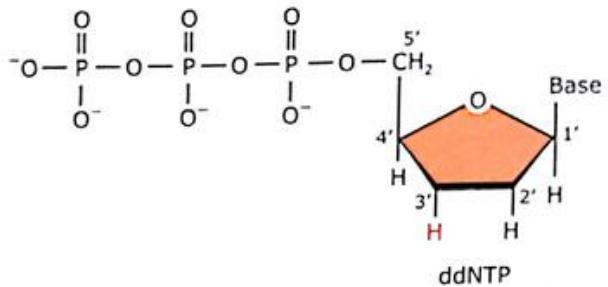
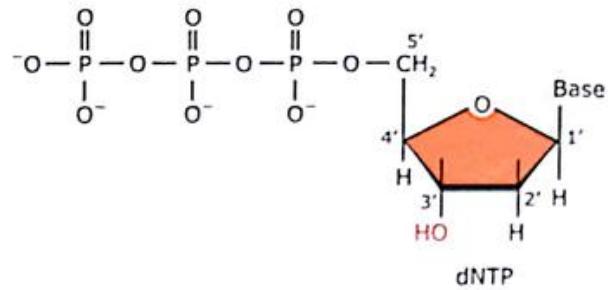


Figure 12.35 The structures of dNTP and ddNTP differ in the number and location of hydroxyl groups on the 2' and 3' carbons. DNA polymerase cannot add another nucleotide to a chain ending in dideoxyribose because its 3' carbon does not have a hydroxyl group.

Each reaction produces a set of DNA copies that terminate at different points in the sequence. The products of these four reactions are separated by electrophoresis in four parallel lanes of a polyacrylamide sequencing gel. Sequencing gels are run in the presence of denaturing agents, urea and formamide. They routinely contain 6–20% polyacrylamide and 7 mol/l urea.

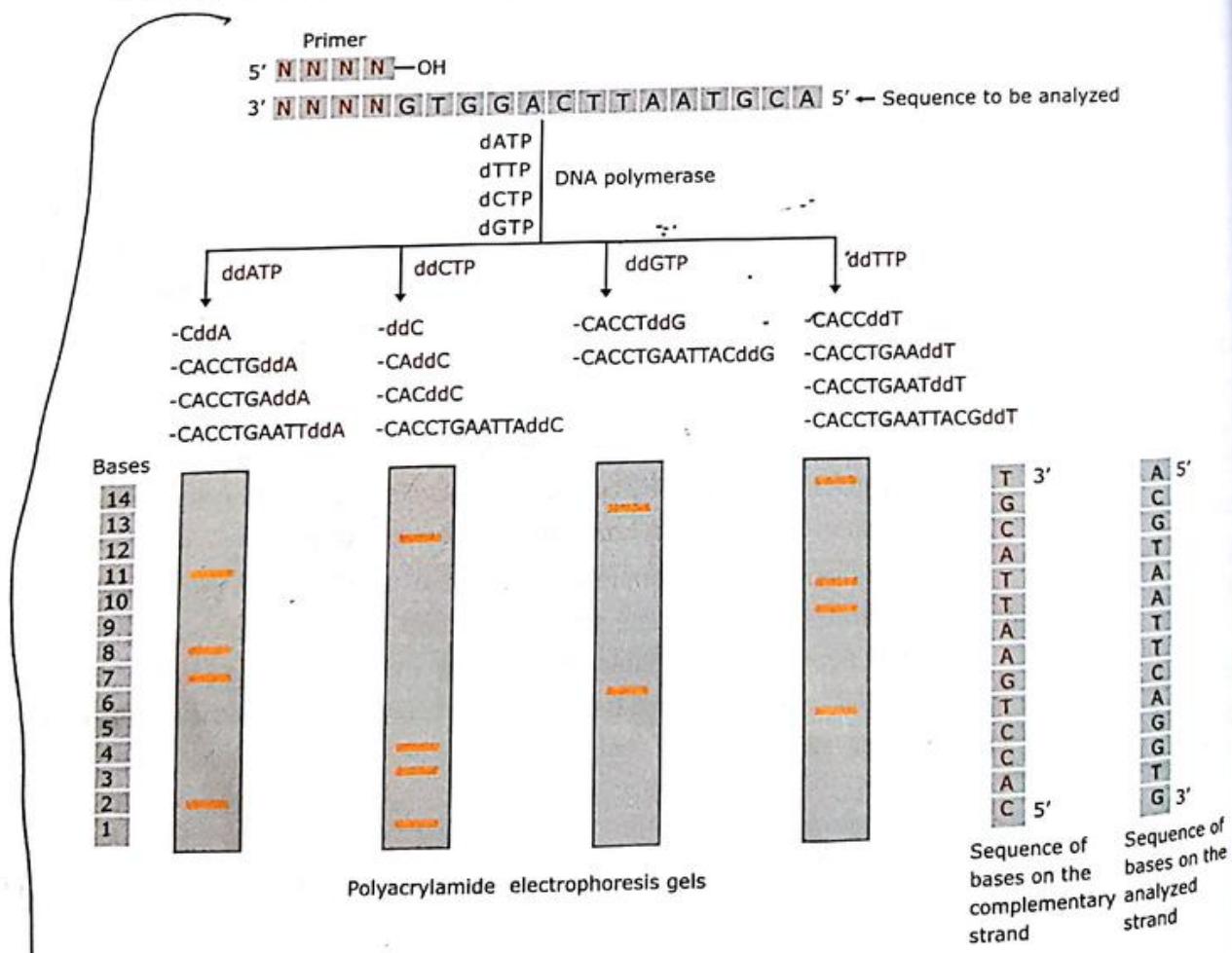


Figure 12.36 DNA sequencing by chain termination method (Sanger method).

The function of the urea is to prevent DNA secondary structure due to intrachain base pairing, which affects electrophoretic mobility. This is important because the change in conformation resulting from base pairing alters the electrophoretic mobility of a single-stranded molecule. DNA sequencing gels tend to be very long (100 cm) to maximize the separation achieved. The newly synthesized fragments are detected by a label either radioactive or fluorescent that has been incorporated either into the primer or into one of the deoxyribonucleoside triphosphates used to extend the DNA chain. In each lane, the bands represent fragments that have terminated at a given nucleotide, but at different positions in the DNA. By reading off the bands in order, starting at the bottom of the gel and working across all lanes, the DNA sequence of the newly synthesized strand can be determined.

Automated sequencing

The standard chain termination sequencing methodology employs radioactive labels, and the banding pattern in the polyacrylamide gel is visualized by autoradiography. This approach is not well suited to automation. To automate the process, it is desirable to acquire sequence data in real time by detecting the DNA bands within the gel during the electrophoretic separation.

Fluorescent dideoxynucleotides are the basis of automated sequencing. Each of the four dideoxynucleotides carries a spectrally different fluorophore. Chains terminated with A are, therefore, labeled with one fluorophore, chains terminated with C are labeled with a second fluorophore, and so on. Now it is possible to carry out the four sequencing reactions – A, C, G and T – in a single tube and to load all four families of molecules into just one lane of the polyacrylamide gel, because the fluorescent detector can discriminate between the different labels and determine if each band represents an A, C, G or T. The sequence can be read directly as the bands pass in front of the detector; and either printed out in a form readable by eye or sent straight to a computer for storage.

5' A C G T A A T T C A G G T G 3' Sequence to be analyzed

T Red G Yellow C Blue A Green

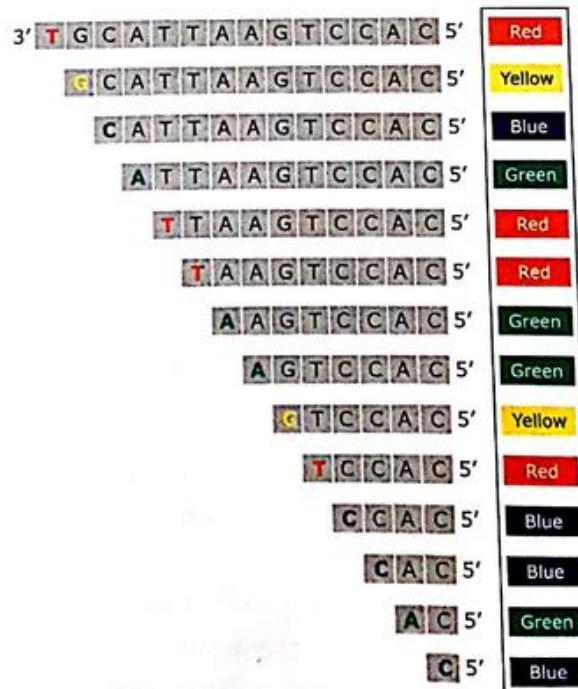


Figure 12.37 Replication of the DNA template strand proceeds with a reaction mixture including the four standard dNTPs and all four ddNTP, each labeled with a different fluorescent dye (ddATP, ddCTP, ddGTP, and ddATP). Random incorporation of the labeled ddNTPs produces a series of DNA fragments in which chain growth has been terminated at each successive position, each one nucleotide longer than the previous. Separation of the fragments by size produces a sequencing ladder as a series of colored bands.

Sequencing enzymes

Naturally occurring polymerases have features that are often not optimal for DNA sequencing. The relevant and desirable properties of sequencing polymerases include:

High processivity

Processivity is the degree to which chain extension continues before the enzyme dissociates from a primer-template annealing complex. T7 Sequenase is the most processive of the current catalog of sequencing enzymes, whereas the Klenow fragment is the least.

Thermostability

Resistance to inactivation or dissociation at high temperatures is the most important factor in the cycle-sequencing reactions that are the basis for modern high-throughput sequencing. *Taq* polymerase or variants are the only feasible options.

Incorporation of nucleotide analogs such as dye terminators

The ability to incorporate analogs is a critical factor for the dideoxy chain-termination method. The efficiency of chain termination with each of the dye-labeled terminators must be similar to avoid low-quality data with uneven peaks.

Exonuclease activities

Polymerases often have 3'-exonuclease 'proofreading' and/or 5'-exonuclease activities that remove RNA primers after DNA replication. Because neither activity is desirable for sequencing, variants of the polymerases should be used that lack these activities (e.g. thermal sequenase).

Chemical degradation method

Chemical degradation method is based on preferential base specific modification followed by chemical cleavage to generate a nested set of end-labeled derivatives. In this method, the starting material is double-stranded DNA, which is first labeled by attaching a radioactive phosphorus group to the 5' end of each strand. The strands are then separated by electrophoresis under denaturing conditions, and analyzed separately. DNA labeled at one end is divided into four aliquots and each is treated with chemicals that act on specific bases. On average, each molecule is modified at only one position along its length; every base in the DNA strand has an equal chance of being modified. After the modification reactions, the separate samples are cleaved by specific chemical, which breaks phosphodiester bonds of nucleotide whose base has been modified. The cleavage reactions in this method have not changed significantly since its initial development, although additional chemical cleavage reactions have been devised. The chemical treatment used for base modification are:

- G only - Methylation of G with dimethyl sulfate (DMS) at pH 8.0.
- A+G - Treatment with piperidine formate at pH 2.0.
- C+T - Hydrazine treatment.
- C only - Hydrazine treatment in the presence of 1.5 M NaCl.

After base modification, the volatile secondary amine piperidine is used to cleave the sugar-phosphate chain of DNA at the sites of base modifications. Thus, these cleavage reactions are carried out in two stages:

First, chemical modification of specific bases, and

Second, removal of the modified base from its sugar, and cleavage of the phosphodiester bonds 5' and 3' to the modified base.

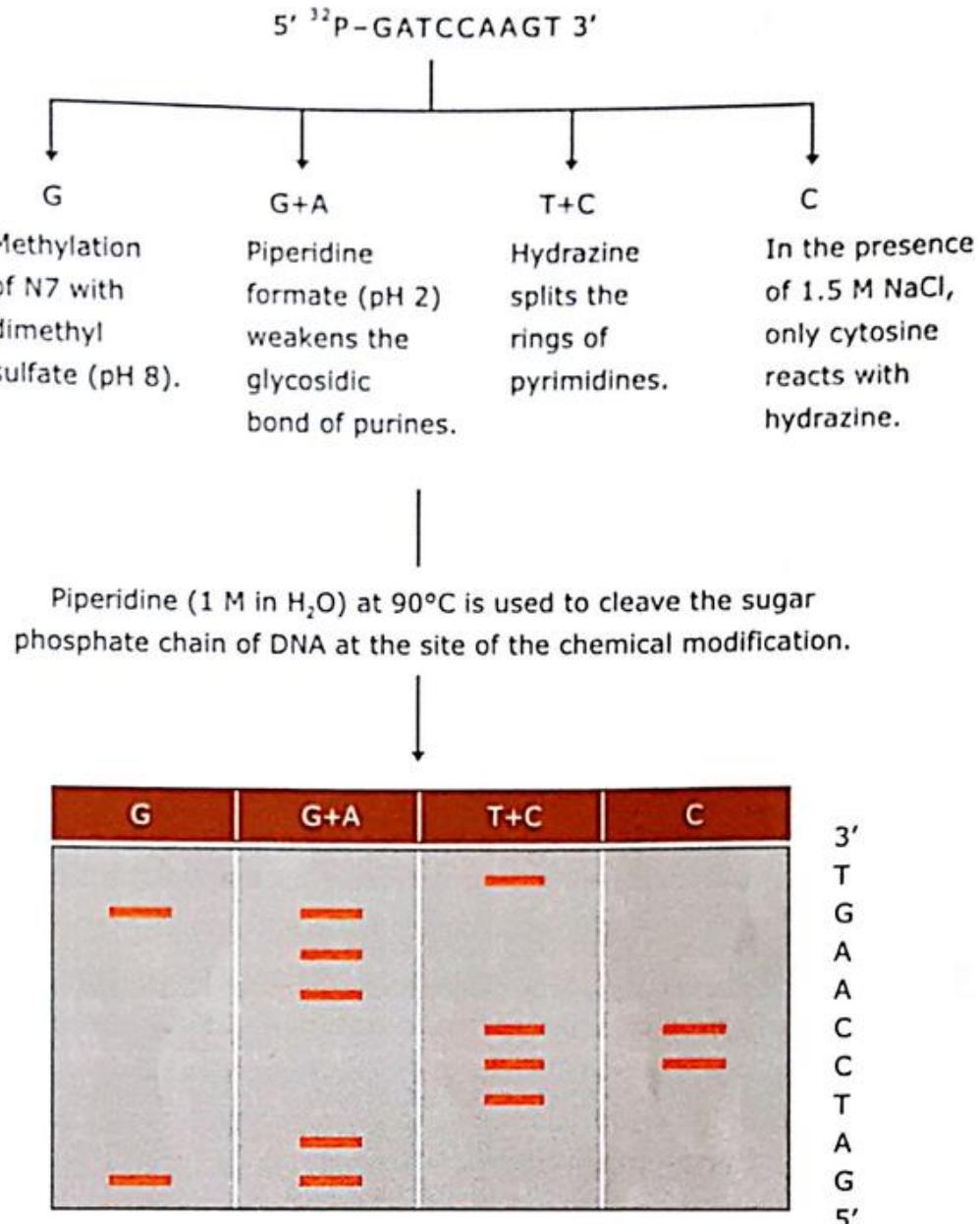


Figure 12.38 Sequencing by the Maxam-Gilbert method. The oligonucleotide in this example (5'-GATCCAAGT-3') is subjected to four chemical cleavage reactions (G, G+A, T+C and C). The resulting products of each of the reactions are separated by electrophoresis through a denaturing acrylamide gel and visualized by autoradiography.

To illustrate the procedure, we will follow the 'G' cleavage reaction. First, the molecules are treated with dimethyl sulfate, which attaches a methyl group to the purine ring of G nucleotides. Only a limited amount of dimethyl sulfate is added, the objective being to modify, on average, just one G per polynucleotide. At this stage the DNA strands are still intact, cleavage not occurring until a second chemical – piperidine – is added. Piperidine removes the modified purine after cleaving the phosphodiester bonds. The result is a set of cleaved DNA molecules, some of which are labeled one and some are not. The labeled molecules all have one end in common and one end determined by the cut sites, the latter indicating the positions of the G nucleotides in the DNA molecules that were cleaved. The cleaved molecules are electrophoresed in a polyacrylamide gel and the sequence read in a similar way to that described for chain termination sequencing.

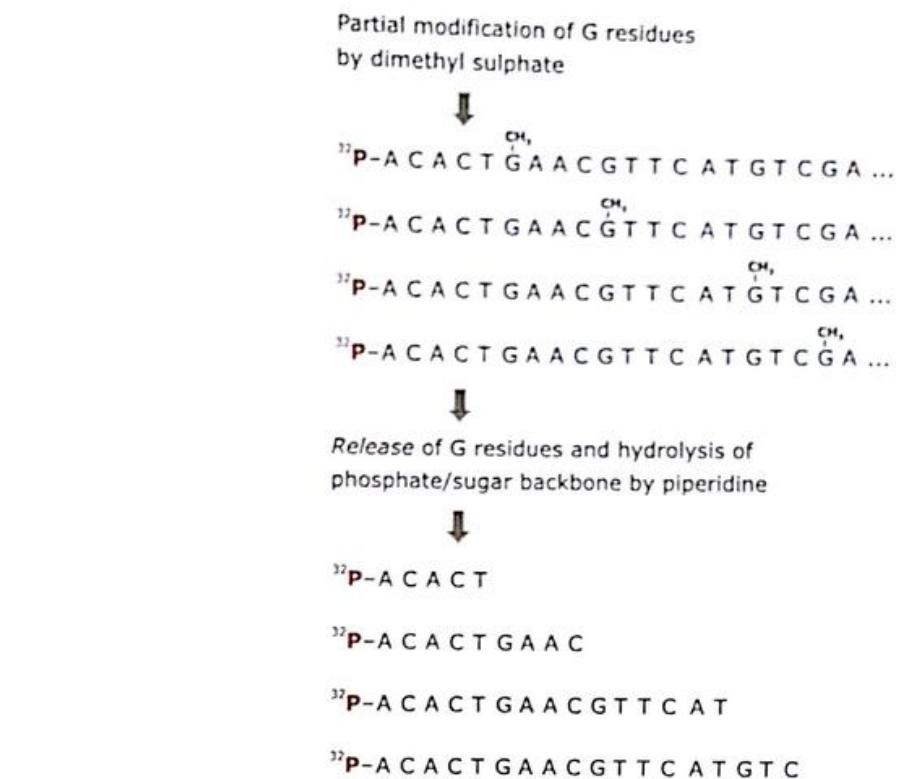
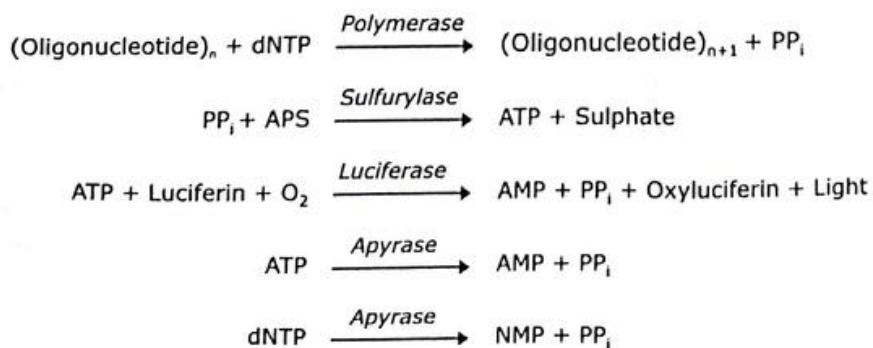


Figure 12.39 DNA sequencing by chemical degradation reaction-specific for guanine residues. The guanine base is first modified with dimethyl sulfate, which makes the chain susceptible to cleavage by piperidine.

Pyrosequencing

Pyrosequencing is a DNA sequencing technology based on the sequencing-by-synthesis principle. The four enzymes included in the pyrosequencing system are the Klenow fragment of DNA polymerase I, ATP sulfurylase, luciferase and apyrase. Apyrase is included in the pyrosequencing technology for degradation of unincorporated nucleotides and excess ATP between base additions. The reaction mixture also contains the enzyme substrates adenosine phosphosulfate (APS), luciferin and the sequencing template with an annealed primer to be used as starting material for the DNA polymerase. The first reaction, DNA polymerization, occurs if the added nucleotide forms a base pair with the sequencing template and, thereby, is incorporated into the growing DNA strand. The inorganic pyrophosphate, released by the DNA polymerase serves as substrate for ATP sulfurylase, which produces ATP. Finally, the ATP is used by luciferase to generate light and the light signal is detected. Hence, only if the correct nucleotide is added to the reaction mixture, light is produced. Apyrase removes unincorporated nucleotides and ATP between the additions of different bases.



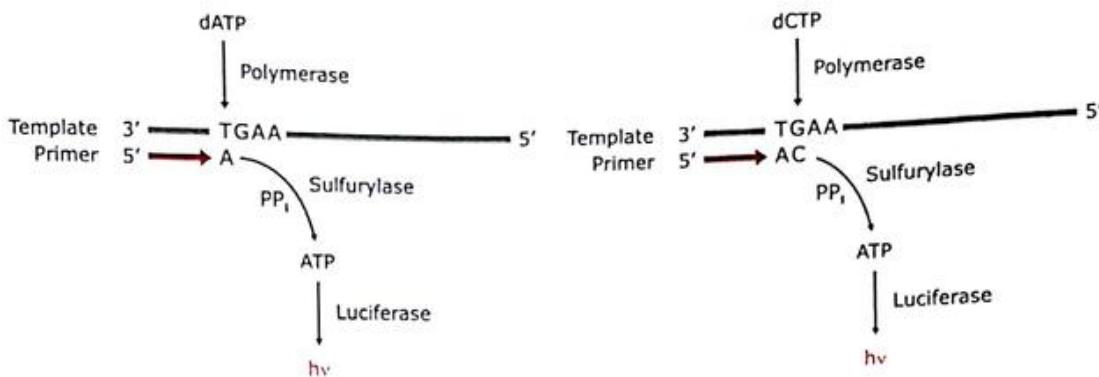


Figure 12.40 Schematic representation of the pyrosequencing enzyme system. If the added dNTP forms a base pair with the template, DNA polymerase incorporates it into the growing DNA strand and pyrophosphate is released. ATP sulfurylase converts the PP_i into ATP which serves as substrate for the light producing enzyme luciferase. The produced light is detected as evidence of that nucleotide incorporation has taken place.

Next-generation sequencing technologies

Nucleic acid sequencing is a method for determining the exact order of nucleotides present in a given DNA or RNA molecule. The chain-termination method (also commonly referred to as Sanger or dideoxy sequencing method), published in 1977, has remained the most commonly used DNA sequencing technique to date. The Human Genome Project, led by the International Human Genome Sequencing Consortium and Celera Genomics, was accomplished with *first-generation* Sanger sequencing. Since completion of the first human genome sequence, demand for cheaper and faster sequencing methods has increased greatly. This demand has driven the development of *second-generation* sequencing methods (or next-generation sequencing).

Next-generation sequencing platforms perform massively parallel sequencing, during which millions of fragments of DNA from a single sample are sequenced in unison. Massively parallel sequencing technology facilitates high-throughput sequencing, which allows an entire genome to be sequenced in less than one day. In the past decade, several next-generation sequencing platforms have been developed.

To illustrate how next-generation sequencing process works, consider a single genomic DNA (gDNA) sample. The gDNA is first fragmented into a library of small segments that can be uniformly and accurately sequenced in millions of parallel reactions. The newly identified strings of bases, called *reads*, are then reassembled using a known reference genome as a scaffold (resequencing), or in the absence of a reference genome (*de novo* sequencing). The full set of aligned reads reveals the entire sequence of each chromosome in the gDNA sample.

Illumina (Solexa) sequencing

This system was initially developed in 2007 by Solexa and was subsequently acquired by Illumina, Inc. Illumina next-generation sequencing utilizes a fundamentally different approach from the classic Sanger chain-termination method. It is based on sequencing by synthesis (SBS) technology – tracking the addition of labeled nucleotides as the DNA chain is copied – in a massively parallel fashion. Illumina sequencing systems can deliver data output ranging from 300 kilobases up to 1 terabase in a single run, depending on instrument type and configuration.

First generation sequencing technologies
 include sequencing by synthesis developed by Sanger and sequencing by cleavage pioneered by Maxam and Gilbert.
Next-generation sequencing refers to non-Sanger-based high-throughput DNA sequencing technologies. Millions or billions of DNA strands can be sequenced in parallel, yielding substantially more throughput and minimizing the need for the fragment-cloning methods that are often used in Sanger sequencing of genomes. The four main advantages of next-generation sequencing over classical Sanger sequencing are:

- speed
- cost
- sample size
- accuracy

In this sequencing technique, random fragments of the genome to be sequenced are immobilized in a flow cell, and then amplified *in situ*, resulting in localized clusters of around 1000 identical copies of each fragment. This system utilizes a sequencing by-synthesis approach in which all four fluorescently labeled and 3'-OH blocked nucleotides (reversible terminators) are added simultaneously to the flow cell channels, along with DNA polymerase, for incorporation into the oligo-primed cluster fragments. The cluster strands are extended by one nucleotide. Following the incorporation step, the unused nucleotides and DNA polymerase molecules are washed away, a scan buffer is added to the flow cell, and the optics system scans each lane of the flow cell by imaging units called tiles. Once imaging is completed, chemicals that affect cleavage of the fluorescent labels and the 3'-OH blocking groups are added to the flow cell, which prepares the cluster strands for another round of fluorescent nucleotide incorporation.

Ion Torrent sequencing

Ion Torrent sequencing does not make use of optical signals. Instead, they exploit the fact that the incorporation of a dNTP into a growing DNA strand involves the formation of a covalent bond and the release of pyrophosphate and a positively charged hydrogen ion. The release of hydrogen ions results in changes in the pH of the solution, which is detected by a detector. This technology differs from other sequencing technologies in that the system does not require nucleotide labeling and no optical detection is involved.

Ion Torrent sequencing may also be referred to as Ion semiconductor sequencing, pH-mediated sequencing, silicon sequencing, or semiconductor sequencing. This technology, perhaps the current most versatile and low cost technology, has been delivered in the form of a personal genomic machine (PGM) as a bench top instrument to research and clinical laboratories. The first step in this approach includes library construction which involves DNA fragmentation and adapter ligation. These fragments are clonally amplified on the small beads by emulsion PCR. Enriched beads are primed for sequencing by annealing a sequencing primer and are deposited into the wells of an Ion Chip, a specialized silicon chip designed to detect pH changes within individual wells of the sequencer as the reaction progresses stepwise.

Each microwell contains many copies of single-stranded template DNA molecule to be sequenced and DNA polymerase are sequentially flooded with a single species of unmodified dNTP. If the introduced dNTP is complementary to the leading template nucleotide, it is incorporated into the growing complementary strand. The hydrogen ion that is released during the reaction changes the pH of the solution. Beneath each microwell the Ion-Sensitive Field Effect Transistor (ISFET) detects the pH change and a potential change is recorded as direct measurement of nucleotide incorporation events. The unattached dNTP molecules are washed out before the next cycle when a different dNTP species is introduced. If the introduced dNTP is not complementary there is no incorporation and no biochemical reaction. If homopolymer repeats are present in the template sequence, multiple dNTP molecules will be incorporated in a single cycle. This leads to a corresponding number of released hydrogen ions and a proportionally higher electronic signal.

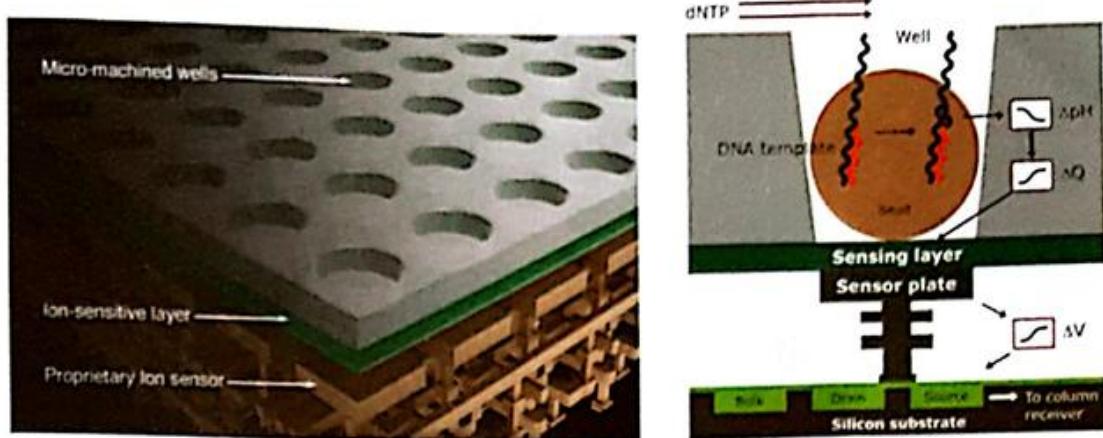


Figure 12.41 Ion Torrent technology- Sensor, well and chip architecture. An image of a well, a bead containing DNA template, and the underlying sensor and electronics. Cross section view of a single well that houses ion sphere particles with a clonal amplified DNA template. A hydrogen ion (proton) is released when a nucleotide is incorporated by DNA polymerase. The proton is then detected by the sensing layer due to the change of pH, therefore translating the chemical signal to a digital input.

12.13 Chromatin immunoprecipitation

Chromatin immunoprecipitation:
Chemical crosslinking of DNA and proteins and immunoprecipitation using a specific antibody to determine DNA-protein associations *in vivo*.

ChIP entails a series of steps:

- DNA and associated proteins are chemically cross-linked;
- Nuclei are isolated, lysed, and the DNA is fragmented;
- An antibody specific for the DNA binding protein (transcription factor, histone, etc.) of interest is used to selectively immunoprecipitate the associated protein-DNA complexes and
- The chemical crosslinks between DNA and protein are reversed and the DNA is claimed for downstream analysis.

The chromatin immunoprecipitation (ChIP) assay is a powerful method for analyzing epigenetic modifications and genomic DNA sequences bound to specific regulatory proteins under a particular set of conditions. In this method, protein-DNA complexes are crosslinked, immunoprecipitated, purified, and amplified for gene- and promoter-specific analysis of known targets. When performing the ChIP assay, cells are first fixed with formaldehyde. It acts as a reversible protein-DNA and protein-protein cross-linking agent that serves to fix the protein-DNA interactions occurring in the cell. Cells are then lysed and chromatin is released and fragmented using either sonication or enzymatic digestion. Antibodies directed against a given gene regulatory protein are then used to purify the DNA that are covalently cross-linked to that protein in the cell. After immunoprecipitation, the protein-DNA cross-links are reversed and the DNA is purified. The purified DNA is analyzed by dot blot or Southern blot using a radiolabeled probe derived from the cloned DNA fragment of interest.

As formaldehyde inactivates cellular enzymes immediately upon addition to cells, ChIP provides snapshots of protein-protein and protein-DNA interactions at a particular time point, and hence is useful for kinetic analysis of events occurring on chromosomal sequences *in vivo*. In addition, if the immunoprecipitated DNA is hybridized to microarrays that contain the entire genome displayed as a series of discrete DNA fragments, the precise genomic location of each precipitated DNA fragment can be determined. In this way, all the sites occupied by the gene regulatory protein in the original cells can be mapped on a genome-wide basis. ChIP has been widely used to map the localization of post-translationally modified histones, histone variants, transcription factors, or chromatin modifying enzymes on the genome or on a given locus.

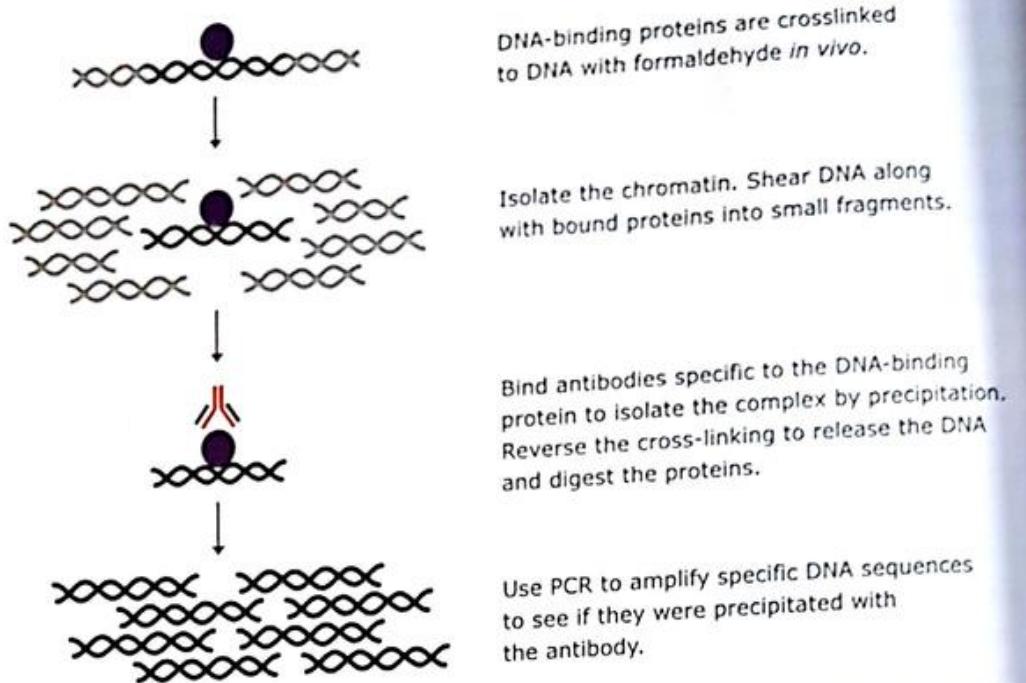


Figure 12.42 Cells grown under the desired experimental condition are fixed with formaldehyde which forms heat-reversible DNA-protein crosslinks. After crosslinking, cells were lysed and sonicated to fragment the chromatin. These fragments are purified and then used to perform ChIP. It is performed by incubation of fractionated chromatin with an antibody directed to a protein of interest. The antibody recognizes its targeted protein and precipitates protein-DNA from solution. In this way, only DNA fragments crosslinked to the protein of interest are enriched, while DNA-protein complexes that are not recognized by the antibody are washed away.

12.14 Biosensors

A biosensor is a measuring device which is used to detect chemical compounds by converting a biological response into an electrical signal. The electrical signal it produces carries the necessary information about the process under investigation. The most widely accepted definition of a biosensors is: 'a self-contained analytical device that incorporates a biologically active material in contact with an appropriate transducer for the purpose of detecting the concentration or activity of chemical species in any type of sample.'

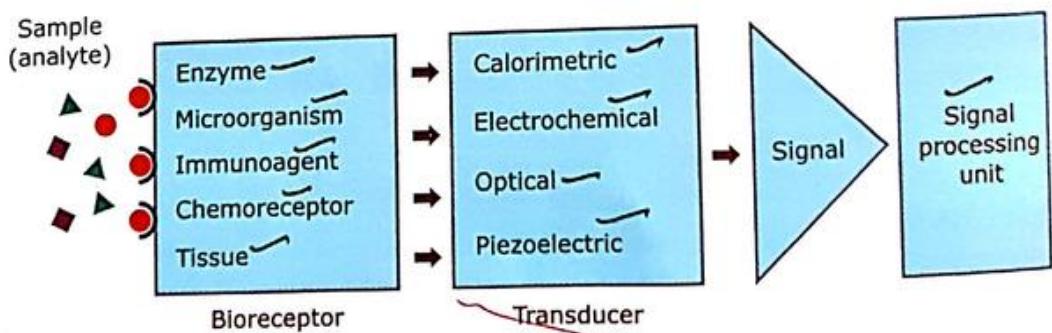


Figure 12.43 Schematic diagram showing the main components of a biosensor. The bioreceptor – interacts with the specific analyte of interest to produce signal; the transducer – transforms the signal resulting from the interaction of the analyte with the bioreceptor into another signal (i.e. transduces) that can be more easily measured and quantified; and signal processing unit – converts the signal into a workable form (amplified, processed and displayed).

There are three main parts of a biosensor:

1. Biological detection system, a bioreceptor. Bioreceptor may be an enzyme, an antibody or similar binding molecule, a living cell or organelles.
2. A transducer, which transforms the signal resulting from the interaction of the analyte with the bioreceptor into a measurable signal.
3. A signal processing system, which converts the signal into a workable form.

Types of biosensors

Biosensors are usually classified into various groups either by type of transducer employed or by the kind of bioreceptor utilized. The bioreceptors include enzymes, antibodies, nucleic acids, microorganisms, biological tissues and organelles. Its role is to interact specifically with the target analyte and the result of this interaction is consequently transformed through transducer to measurable signal. Enzymes are the most commonly used bioreceptors in biosensors.

There are two classes of bio-recognition processes - Bioaffinity recognition and Biocatalytic recognition. Both processes involve the selective binding of an analyte with receptor. In bioaffinity recognition, the binding is very strong and the transducer detects the presence of the bound receptor-analyte pair. In biocatalytic recognition, the analytes are chemically altered to form the product molecules.

Another key part of a biosensor is the transducer. It transforms the physical and chemical change accompanying the biorecognition event into measurable electrical signals. The transducers can be electrochemical, optical, piezoelectric and calorimetric.

Electrochemical transducers: The basic principle for this class of biosensors is that chemical reactions between immobilized bioreceptors and target analyte produce or consume ions or electrons, which affects measurable electrical properties of the solution. Amperometric and potentiometric transducers are the most commonly used electrochemical transducers. In amperometric transducers, the potential between the two electrodes is set and the current produced by the oxidation or reduction of electroactive species is measured and correlated to the concentration of the analyte of interest. Potentiometric transducers measure electrical potential due to changes in the distribution of charges.

Calorimetric (thermometric) transducers measure the heat of a biochemical reaction. Once the analyte comes in contact with the bioreceptor, the heat of reaction which is proportional to the analyte concentration is measured.

Optical transducers can employ a number of techniques to detect the presence of a target analyte and are based on well-founded methods including chemiluminescence, fluorescence, light absorbance, phosphorescence, photothermal techniques, surface plasmon resonance (SPR), light polarization and rotation, and total internal reflectance.

Piezoelectric (mass-sensitive) transducers are based on the coupling of the bioreceptors with a piezoelectric component, usually a quartz-crystal coated with gold electrodes. Crystals, such as those of quartz, have no center of symmetry and produce an electrical signal when stressed mechanically (i.e. by applying some pressure on them). A crystal oscillates at a certain frequency, which can be modulated by its environment. When the crystal is coated with some material, the actual frequency depends on the mass of the crystal and the coating. The resonant frequency can be measured with great accuracy hence making it possible to calculate the mass of analyte adsorbed onto the crystal surface.

~~X~~ Bioreceptor immobilization

Very important part of a biosensor fabrication is the immobilization of the bioreceptor in the vicinity of the transducer. Two important considerations have to be taken into account during immobilization - operational stability and long term use. The immobilization is done either by physical adsorption and entrapment or chemical methods. Chemical attachment often involves covalent bonding to transducer surface by suitable reagents. The physical adsorption utilizes a combination of van der Waals and hydrophobic forces, hydrogen bonds and ionic interactions to attach the bioreceptors to the surface of the transducer.

~~X~~ Physical methods

Physical methods of immobilization include entrapment and adsorption. Biorecognition molecules such as enzymes can be entrapped in polyacrylamide, calcium alginate, agarose or chitosan polymer network or gel. Chief disadvantages of this technique are irregular pore size of the gel, lack of mechanical strength and diffusional limitations encountered by substrates and products. Direct physical adsorption of enzymes on a surface is an alternative method. However, immobilization using adsorption (utilizes a combination of van der Waals forces, hydrophobic forces, H-bonds and ionic interactions to attach the biorecognition molecules on transducer or support matrices) alone generally leads to poor long-term stability.

~~X~~ Chemical methods

Chemical methods of enzyme immobilization include covalent binding and crosslinking using multifunctional reagents, such as glutaraldehyde and cyanuric chloride.

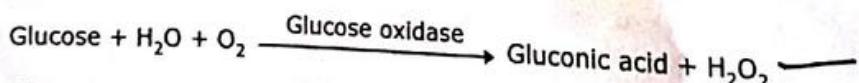
~~X~~ Biosensor characteristics

Biosensors are characterized by eight parameters. These are:

1. *Sensitivity* is the response of the sensor to per unit change in analyte concentration.
2. *Selectivity* is the ability of the sensor to respond only to the target analyte. That is, lack of response to other interfering chemicals is the desired feature.
3. *Range* is the concentration range over which the sensitivity of the sensor is good.
4. *Response time* is the time required by the sensor to produce responses.
5. *Reproducibility* is the accuracy with which the sensor's output can be obtained.
6. *Detection limit* is the lowest concentration of the analyte to which there is a measurable response.
7. *Life time* is the time period over which the sensor can be used without significant deterioration in performance characteristics.
8. *Stability* characterizes the change in its baseline or sensitivity over a fixed period of time.

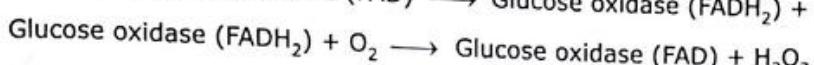
~~X~~ Enzyme-based glucose biosensor

In 1962, Leland C. Clark (father of biosensor) first developed enzyme-based electrochemical glucose biosensor. The basic operation of glucose biosensor is based on the fact that the enzyme *glucose oxidase* catalyses the oxidation of glucose to δ -gluconolactone, which then hydrolyzes to gluconic acid and hydrogen peroxide.



Glucose oxidase is a dimeric protein composed of two identical subunits. Each subunit folds into two domains: one domain binds to substrate β -D-glucose, while the other domain binds

to FAD. The catalytic reaction involves reduction of the FAD present in the enzyme into FADH_2 . FADH_2 finally reoxidizes by molecular oxygen to regenerate the FAD.



The enzyme glucose oxidase acts as a bioreceptor molecule and once it binds with the glucose molecule, it oxidizes glucose to gluconic acid and hydrogen peroxide. To measure the glucose in aqueous solutions, three different transducers can be used:

1. An oxygen sensor that measures oxygen concentration.
2. A pH sensor that measures the acid (gluconic acid), a reaction product.
3. A peroxidase sensor that measures H_2O_2 concentration, a reaction product.

Note that an oxygen sensor is a transducer that converts oxygen concentration into electrical current. A pH sensor is a transducer that converts pH change into voltage change. Similarly, a peroxidase sensor is a transducer that converts peroxidase concentration into an electrical current. In the first biosensor (invented by L.C. Clark), biosensor was made from a thin layer of glucose oxidase entrapped on a Clark oxygen electrode using a dialysis membrane. Using this amperometric electrochemical glucose biosensor, the amount of glucose was estimated by the reduction in the dissolved oxygen concentration. The decrease in measured oxygen concentration was proportional to glucose concentration.

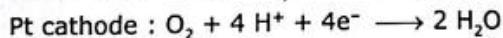
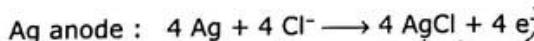
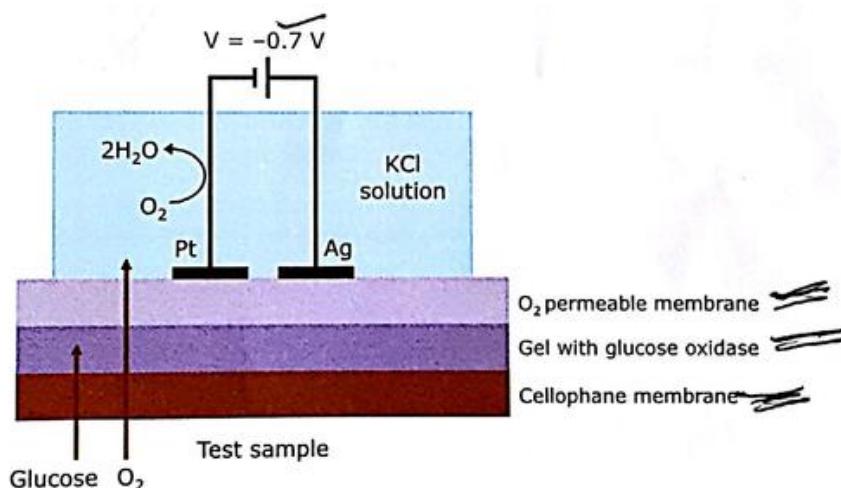


Figure 12.44 Clark glucose biosensor: O_2 in test sample diffuses through membrane into KCl electrolyte. O_2 is reduced at the Pt cathode. Four electrons used by each reduced O_2 results in current flow proportional to $[\text{O}_2]$. A voltage of -0.7 V is applied between the platinum cathode and the silver anode and this voltage is sufficient to reduce the oxygen. The cell current is proportional to the oxygen concentration and the current is measured (amperometric method of detection has been employed). The concentration of glucose is then proportional to the decrease in current (oxygen concentration).

It is also possible to use glucose dehydrogenase (GDH) instead of glucose oxidase for amperometric biosensing of glucose. However, the construction of glucose biosensors based on GDH requires a source of NAD^+ and a redox mediator to lower the overvoltage for oxidation of the NADH product.

Evolution from first to third generation biosensors

There are three so-called 'generations' of biosensors: *first generation* biosensors where the normal product of the reaction diffuses to the transducer and causes the electrical response, *second generation* biosensors which involve specific 'mediators' between the reaction and the transducer in order to generate improved response, and *third generation* biosensors where the reaction itself causes the response and no product or mediator diffusion is directly involved.

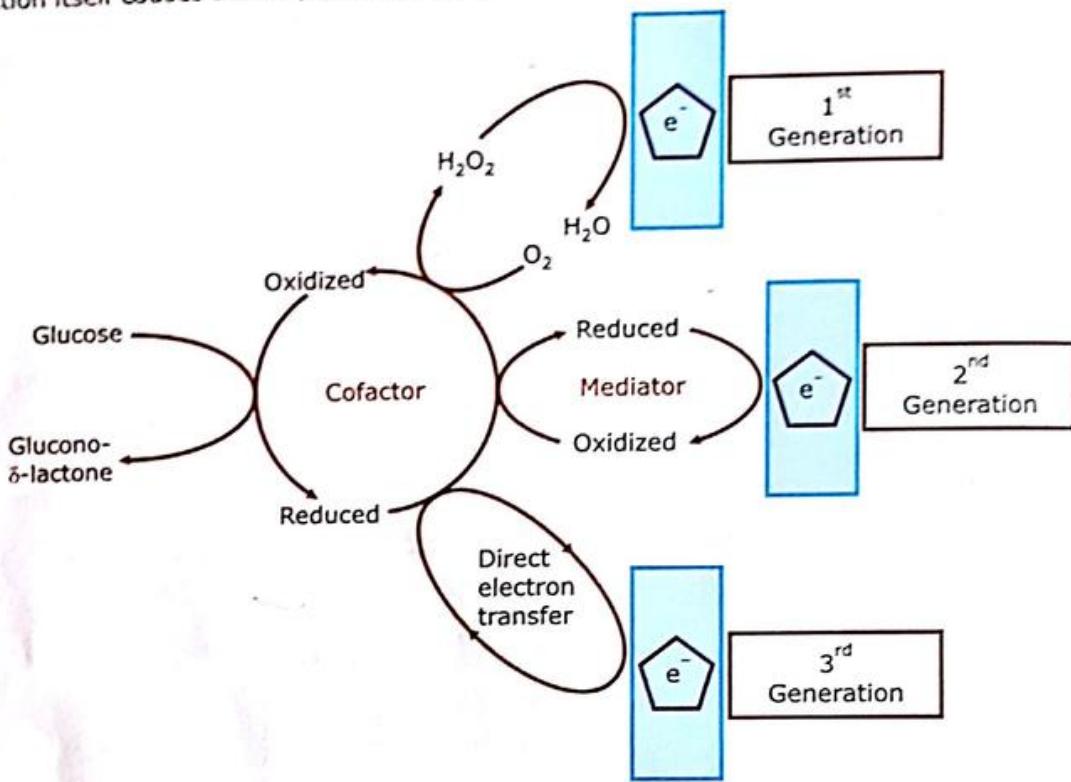


Figure 12.45 The evolution from 1st to 3rd generation electrochemical biosensors. The figure highlights modifications in the biosensor layout with each generation using glucose sensors as an example.

References

- Alberts B, et al (2008), *Molecular Biology of the Cell*, 5th ed. Garland Science.
- Bakker E (2004), *Electrochemical sensors*, Analytical Chemistry 2004, 76 (12), 3285–3298.
- Berk AJ (1989), Characterization of RNA molecules by S1 nuclease analysis. *Methods in Enzymology*, 180, 334–347
- Brown TA (2010), *Gene cloning and DNA analysis*, 6th ed. Wiley-Blackwell.
- Carrigan PE, Ballar P, Tuzmen S (2011), Site-directed mutagenesis, *Methods Mol Biol.* 700:107-24.
- Divan A and Royds J (2013), *Tools and techniques in biomolecular science*, Oxford University Press.
- Egging B (2002) *Chemical sensors and biosensors*, Analytical Techniques in the Sciences. John Wiley & Sons, West Sussex.
- Ekins R and Chu FW (1999). Microarrays: their origins and applications. *Trends Biotechnol.* 17(6):217–18
- Erlich HA, Gelfand D and Sninsky JJ (1991), Recent advances in the polymerase chain reaction. *Sciences*, 252: 1643-1651.

Self Test

Question 1

A biochemist is attempting to separate a DNA binding protein (X) in a solution. Only three other proteins (A, B and C) are present. The proteins have the following properties:

	Isoelectric point	Size	Bind to DNA
Protein A	7.4	82,000	Yes
Protein B	3.8	21,500	Yes
Protein C	7.9	23,000	No
Protein X	7.8	22,000	yes

What type of protein separation techniques might she use to separate?

- P. Protein X can be separated from protein A by size-exclusion chromatography.
- Q. Protein X can be separated from protein B by ion-exchange chromatography.
- R. Protein X can be separated from protein A by isoelectric focusing.
- S. Protein X can be separated from protein C by affinity chromatography.
- a. P and Q
- b. Q and S
- c. P, Q and S
- d. R and S

Question 2

A mixture of amino acids can be analyzed by first separating the mixture into its components through ion exchange chromatography. Amino acids placed on a cation exchange resin containing sulfonate groups flow down the column at different rates because of two factors that influence their movement:

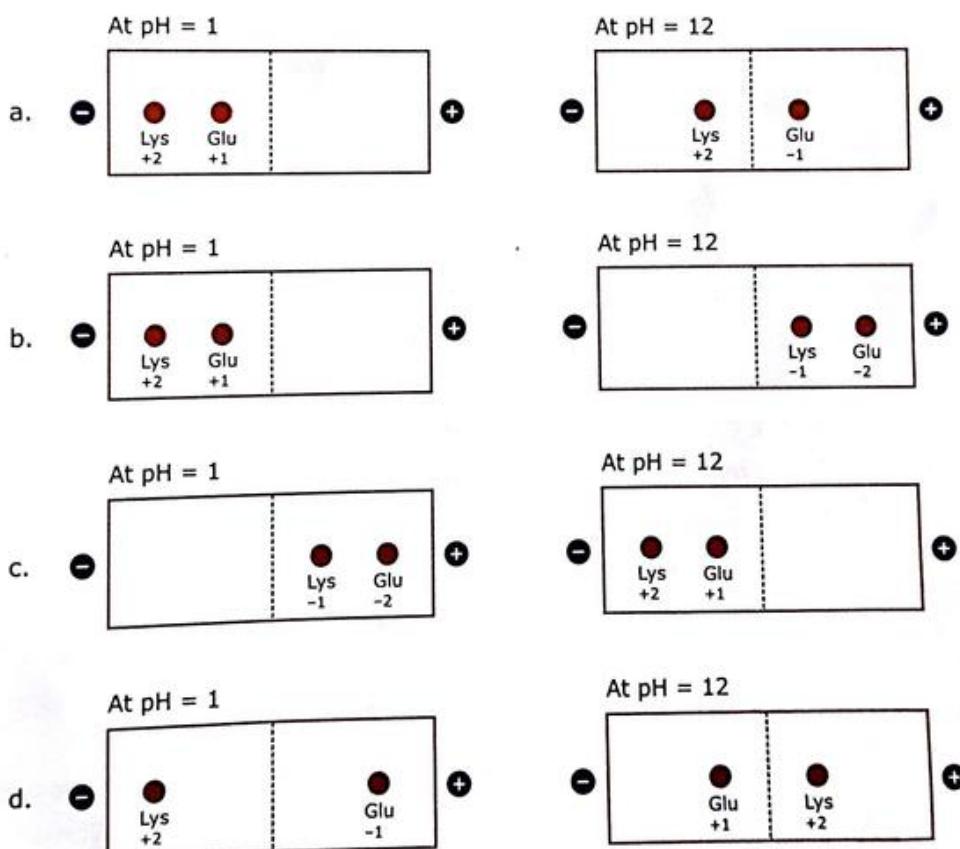
1. Ionic attraction between the sulfonate residues on the column and positively charged functional groups on the amino acids.
 2. Hydrophobic interactions between amino acid side chains and the strongly hydrophobic backbone of the polystyrene resin. For each pair of amino acids listed, determine which will be eluted first from the cation-exchange column by a pH 7 buffer.
- P. Asp and Lys
 - Q. Glu and Val
 - R. Gly and Leu
 - S. Ser and Ala
 - a. P-Asp, Q-Glu, R-Gly, S-Ser
 - b. P-Asp, Q-Val, R-Leu, S-Ala
 - c. P-Lys, Q-Glu, R-Leu, S-Ser
 - d. P-Lys, Q-Val, R-Gly, S-Ala

Question 3

A protein has a molecular mass of 400 kDa when measured by gel filtration chromatography. When subjected to gel electrophoresis in the presence of SDS, the protein gives three bands with molecular masses of 180, 160 and 60 kDa. When electrophoresis is carried out in the presence of SDS and β -mercaptoethanol, three bands are again formed, this time with molecular masses of 160, 90 and 60 kDa. What is the subunit composition of the protein?

Question 4

A mixture of glutamic acid and lysine is separated by paper electrophoresis in two experiments, one at pH 1 and other at pH 12. After the separation, the papers are treated with ninhydrin to reveal the location of the amino acids (colored spots). The results are shown below. Which of the following represent the *correct* result?



Question 5

From the data given below, identify the protein pair that would (P) give the least mobility band on a SDS polyacrylamide gel electrophoresis experiment and (Q) elute last on an anion exchange e.g. DEAE column respectively:

Protein	pI	Subunit MW	Native MW
A	9	10,000	20,000
B	8	35,000	35,000
C	6	15,000	90,000
D	5	20,000	80,000
E	3	30,000	30,000

a. P-protein C ; Q-protein A b. P-protein A ; Q-protein E
 c. P-protein B ; Q-protein E d. P-protein B ; Q-protein A

Question 6

You have a mixture of proteins with the following properties:

Protein 1 : molecular mass = 12 kDa and pI = 10

Protein 2 : molecular mass = 62 kDa and pI = 4

Protein 2 : molecular mass = 28 kDa and pI = 8

σ -value = 4.0; molecular mass = 98 kDa and PI = 5.

Predict the order of emergence of these proteins when a mixture of the four is chromatographed in DEAE-cellulose at pH 7.0

Question 7

A mixture of four proteins of pIs 11, 7, 5 and 3 are loaded on DEAE anion-exchange column equilibrated with low ionic strength buffer of pH 8. Which of the four proteins would be expected to be retained on the column?

- a. Protein with pI 11 but not the others
 - b. Proteins with pIs 11 and 7 but not 5 and 3
 - c. Proteins with pIs 7, 5 and 3
 - d. Protein with pI 7 but not the others

Question 8

The void volume of a gel filtration column, $V_0 = 30$ mL. A monomeric protein with a known molecular mass of 25 kDa elutes at a volume of 45 mL. The protein that you are trying to purify elutes at a volume of 35 mL. Which of the following is a valid conclusion?

- a. Your protein is repelled by the gel filtration material
 - b. Your protein molecular mass = 25 kDa
 - c. Your protein molecular mass > 25 kDa
 - d. Your protein molecular mass < 25 kDa

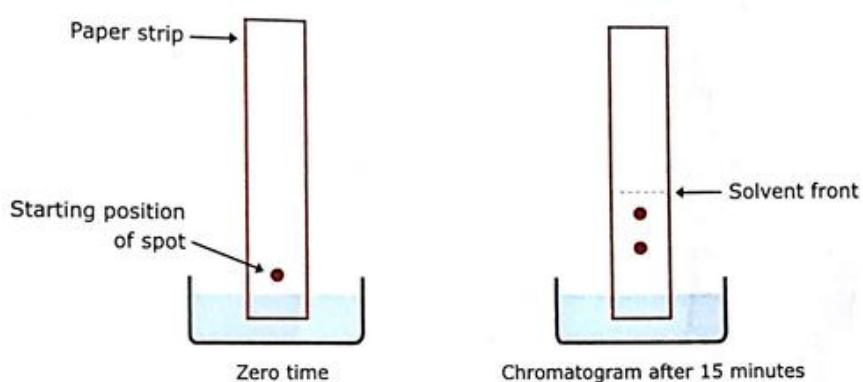
Question 9

A mixture containing proteins-1, -2, -3, -4 and -5 with molecular weights 5,000, 10,000, 25,000, 65,000 and 100,000, respectively, were separated as a Sephadex G-50 column. The order of elution of these proteins from the column will be

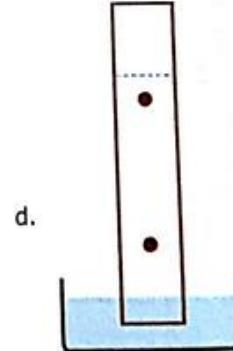
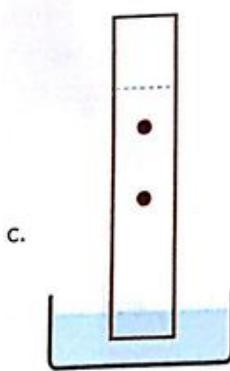
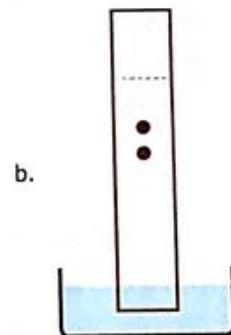
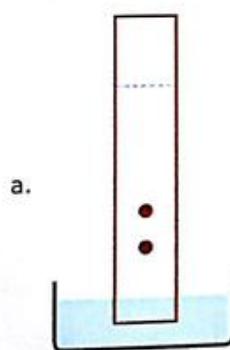
- protein-1, protein-2, protein-3, protein-4 and protein-5
- protein-5, protein-4, protein-3, protein-2 and protein-1
- protein-1, -2, and -3 elute first, followed by protein-5 and -4
- protein-4 and -5 elute first, followed by protein-3, -2 and -1

Question 10

Some students used paper chromatography to separate the pigments in purple ink. They set up a chromatogram and after 15 minutes the colors had separated as shown in the diagram.



Which one of the following diagrams is most likely to indicate the appearance of the chromatogram after a further 30 minutes?



Question 11

- You are given a mixture that contains glutamic acid ($pI = 3.2$), arginine ($pI = 10.8$) and valine ($pI = 6.0$) and you subject the mixture to electrophoresis.
- P. Arginine migrates toward the cathode when the electrophoresis is carried out at a pH of 7.1.
- Q. Valine and glutamic acid migrate toward the anode when the electrophoresis is carried out at a pH of 7.1.
- R. Glutamic acid migrates farthest towards the anode at a pH of 7.1.
- S. Since amino acids themselves are colorless, ninhydrin is used to locate the amino acids which forms a purple derivative in the presence of an amino acid.
- a. Q and S
b. Q, R and S
c. P, Q and R
d. P, Q, R and S

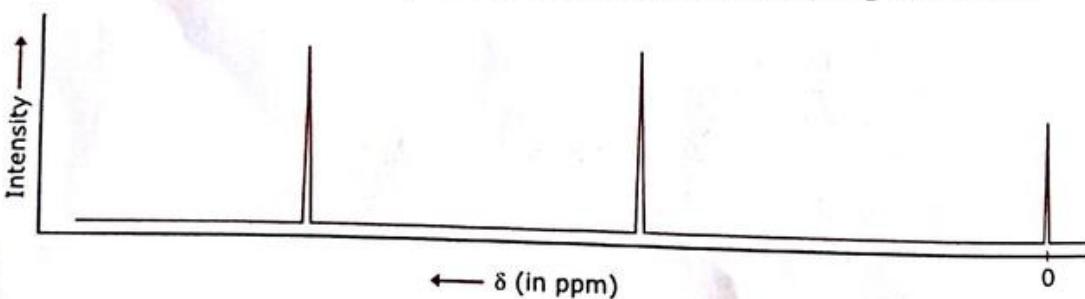
Question 12

The separation and identification of proteins that can be used as disease markers is an exciting area of research. Researchers must separate and identify proteins that could be used as disease markers from the many thousands of proteins that exist in our bodies. Which of the following sequence of techniques could be used to

1. separate these molecules
 2. accurately determine their molecular mass, and
 3. determine their molecular structure.
- a. NMR spectroscopy, followed by mass spectrometry, followed by high-performance liquid chromatography.
- b. high-performance liquid chromatography, followed by mass spectrometry, followed by NMR spectroscopy.
- c. high-performance liquid chromatography, followed by infrared spectroscopy, followed by mass spectrometry.
- d. mass spectrometry, followed by high-performance liquid chromatography, followed by infrared spectroscopy.

Question 13

The $^1\text{H-NMR}$ spectrum of methyl acetate (not drawn to scale) is given below.



Which of the following statements are *correct* about the spectrum?

- P. The zero δ -value corresponds to TMS.
- Q. Higher the value of δ , the higher the deshielding and lower the electron density around the proton.
- R. Number of peaks indicating the number of chemically different protons present in the molecule.

- S. The two signals are of equal intensity because the number of protons responsible for them is equal.
- P and Q
 - Q and S
 - Q, R and S
 - P, Q, R and S

Question 14

The absorption spectrum of a chromophore is primarily determined by the chemical structure of the molecule. However, the polarity of the solvent also affects the electronic spectra. Which of the following statements are correct about the polarity effect?

- The $\pi-\pi^*$ transition shows red shift in more polar solvents.
 - The $n-\pi^*$ transition shows a red shift in more polar solvents.
 - The $n-\pi^*$ transition occurs at a shorter wavelength in polar solvents than in nonpolar solvents.
 - Solvent effect for $n-\pi^*$ transitions is opposite to that found for $\pi-\pi^*$ transitions. Polar solvents cause a shift to lower wavelengths relative to non-polar solvents.
- Q and S
 - P and R
 - P, R and S
 - P, Q, R and S

Question 15

Solutions containing tryptophan and tyrosine have characteristic absorption spectra. The molar absorption coefficients at 240 nm and 280 nm are the following:

Wavelength (nm)	ϵ_{Tyr} ($M^{-1} \text{ cm}^{-1}$)	ϵ_{Trp} ($M^{-1} \text{ cm}^{-1}$)
240	11,300	1960
280	1,500	5380

A 10 mg sample of a protein is hydrolyzed to its constituent amino acids and diluted to 100 ml. The absorbance of this solution, using a 1 cm light path, is 0.717 at 240 nm and 0.239 at 280 nm. Estimate the content of tryptophan and tyrosine using units of $\mu\text{mol/g}$ protein.

- 322 and 586 $\mu\text{mol/g}$
- 281 and 586 $\mu\text{mol/g}$
- 586 and 686 $\mu\text{mol/g}$
- 181 and 506 $\mu\text{mol/g}$

Question 16

A solution containing two substances, A and B, has an absorbance in a 1 cm cuvette of 0.36 at 350 nm and 0.225 at 400 nm. The molar absorption coefficients ($M^{-1} \times \text{cm}^{-1}$) of A and B at the two wavelengths are given below. The concentration of B is 3×10^{-5} M. Calculate the concentrations of A in the solution.

Compound	ϵ at 350 nm	ϵ at 400 nm
A	15,000	300
B	7,000	6500
a. 3×10^{-5} M	b. 3×10^{-3} M	c. 1×10^{-5} M
c. 1×10^{-5} M	d. 1×10^{-3} M	

Question 17

A solution containing NAD⁺ and NADH had an optical density (i.e. absorbance) of 0.311 at 340 nm and 1.2 at 260 nm in a 1 cm cuvette. Calculate the concentrations of the NAD⁺ in the solution. Both NAD⁺ and NADH absorb at 260 nm, but only NADH absorbs at 340 nm. The extinction coefficients (ϵ) are given below.

ϵ (M⁻¹ × cm⁻¹)

Compound	260 nm	340 nm
NAD ⁺	18,000	~0
NADH	15,000	6220
a. 5×10^{-5} M		b. 2.5×10^{-5} M
c. 0.75×10^{-5} M		d. 0.5×10^{-5} M

Question 18

The product of an enzyme reaction is estimated in a colorimeter. For a 1 cm path length cuvette, 20% of the incident light intensity is absorbed. If the path length is increased to 2 cm, the percentage of light transmitted is

- | | |
|-------|-------|
| a. 36 | b. 40 |
| c. 64 | d. 90 |

Question 19

IR spectroscopy is one type of vibrational spectroscopy. It measures the vibrational frequency of bonds. Which of the following statements concerning vibrational frequencies are correct?

1. It is unique for each type of bond.
 2. Stretching frequencies are lower than corresponding bending frequencies.
 3. Bonds to lighter atoms usually vibrate at higher frequencies than heavier atoms.
 4. The frequency of the stretching vibration of a bond depends on the masses of the atoms and the stiffness of the bond.
 5. The C—O bond has a lower frequency than the C—N bond because an O-atom has more mass than N-atom.
- | | |
|------------------|------------------|
| a. 1, 4 and 5 | b. 2, 3 and 4 |
| c. 1, 3, 4 and 5 | d. 2, 3, 4 and 5 |

Question 20

Separation of small peptides, electrophoresis on a paper support is effected on the basis of the charge on a peptide at different pH values. Predict the direction of migration for the peptide (Lys-Gly-Ala-Glu) at the given pH values. 'C' is mentioned for migration toward the cathode, A for migration toward the anode and 'O' if the peptide remains stationary.

	pH			
	2.0	4.0	6.0	11.0
a. Lys-Gly-Ala-Glu	C	C	O	A
b. Lys-Gly-Ala-Glu	A	C	O	A
c. Lys-Gly-Ala-Glu	C	A	A	A
d. Lys-Gly-Ala-Glu	C	C	A	A

Question 21

The most important property of compound light microscope is its power of resolution, which is numerically equivalent to D, the minimum distance between two distinguishable objects. D depends on three parameters namely, the angular aperture, α , the refractive index, N, and wavelength, λ , of the incident light. Below are given few possible options to increase the resolution of the microscope.

- P. Decrease the value of λ or increase either N or α to improve resolution.
- Q. Moving the objective lens closer to the specimen will decrease $\sin\alpha$ and improve the resolution.
- R. Using a medium with high refractive index between the specimen and the objective lens to improve the resolution.
- S. Increase the wavelength of the incident light to improve the resolution.

Which of the following combination of above statements is *correct*?

- a. P and R
- b. Q and R
- c. P and S
- d. R and S

Question 22

Using FRAP (fluorescence recovery after photo-bleaching) techniques, diffusion coefficient of three integral membrane proteins M_1 , M_2 and M_3 in a kidney cell is calculated as 1 $\mu\text{m}/\text{s}$, 0.05 $\mu\text{m}/\text{s}$ and 0.005 $\mu\text{m}/\text{s}$, respectively. Considering fluid-mosaic nature of biological membrane and relationship of structural organization of integral membrane protein with diffusion coefficient, which protein(s) will have highest number of integral membrane domain?

- a. M_2 and M_3
- b. M_2 only
- c. M_3 only
- d. M_1 and M_3

Question 23

The DNA mixture containing ^{14}N and ^{15}N labelled DNA from phage T4 was denatured and allowed to reanneal. How many bands one would observe upon CsCl_2 density gradient centrifugation of the above mixture?

- a. 0
- b. 1
- c. 2
- d. 3

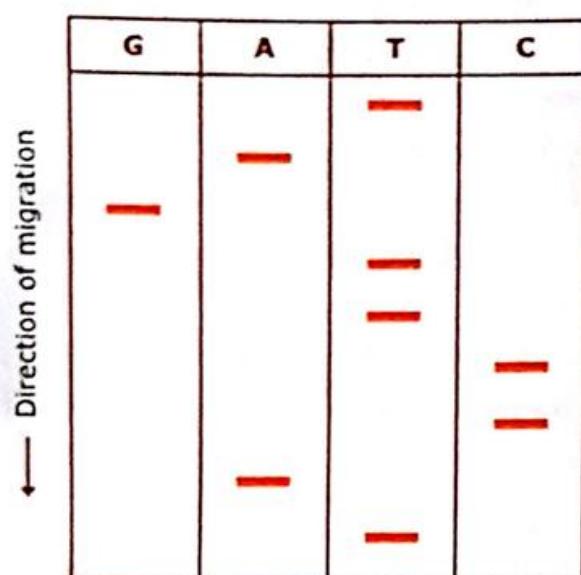
Question 24

A DNA sequencing reaction was performed with the fragment 5'-XXXXGCGATCGYYYY-3' as the template, dideoxy GTP, all the four dNTPs, and the required primers and enzyme. XXXX and YYYY in the given DNA fragment represent primer binding sites. The set of fragments obtained during the reaction will be (the primers are not shown in the amplified fragments)

- a. 5'-CGATCGC-3' only
- b. 5'-CG-3', 5'-CGCTAG-3', 5'-CGCTAGC-3'
- c. 5'-CG-3', 5'-CGATCG-3', 5'-CGATCGC-3'
- d. 5'-G-3', 5'-GCG-3', 5'-GCGATCG-3'

Question 25

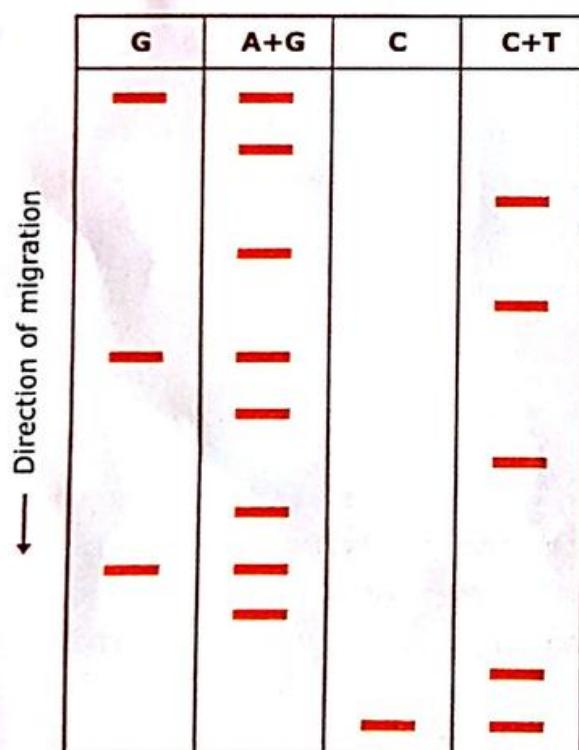
The following gel pattern was obtained in an attempt to sequence a DNA molecule using the Sanger method. What is the sequence of the original DNA molecule?



- a. 5' ATCAAGGTA 3'
- b. 3' ATCAAGGTA 5'
- c. 5' ATGGAACATA 3'
- d. 5' TACCTTGAT 3'

Question 26

A restriction fragment, obtained with a type II endonuclease, was subjected to Maxam-Gilbert sequencing with results as shown in the autoradiogram below. Find out the sequence of DNA fragment.



- a. 5' GATATGATAGATC 3'
- b. 5' CTAGATAGTATAG 3'
- c. 5' GGTGTGGTGGGTC 3'
- d. 5' CTATACTATCTAG 3'

Question 27

Given below are the experimental protocols to find out the exact location of repetitive DNA sequence in mitotic chromosome by FISH (fluorescence in situ hybridization). Which one of the protocols will give the correct result?

- a. Mitotic chromosomes were fixed on glass slide → incubated with biotinylated telomeric DNA → denatured → incubated with fluorescently labeled.
- b. Mitotic chromosomes were fixed on glass slide → denatured → incubated with FITC labeled unrelated non-repetitive DNA sequence → counterstained with propidium iodide → localization observed under fluorescence microscope.
- c. Mitotic chromosomes were fixed on glass slide → denatured → incubated with biotinylated satellite DNA → localization observed under fluorescence microscope.
- d. Mitotic chromosomes were fixed on glass slide → denatured → incubated with repetitive DNA sequence binding protein → FITC labeled antibody against the protein → localization observed under fluorescence microscope.

Question 28

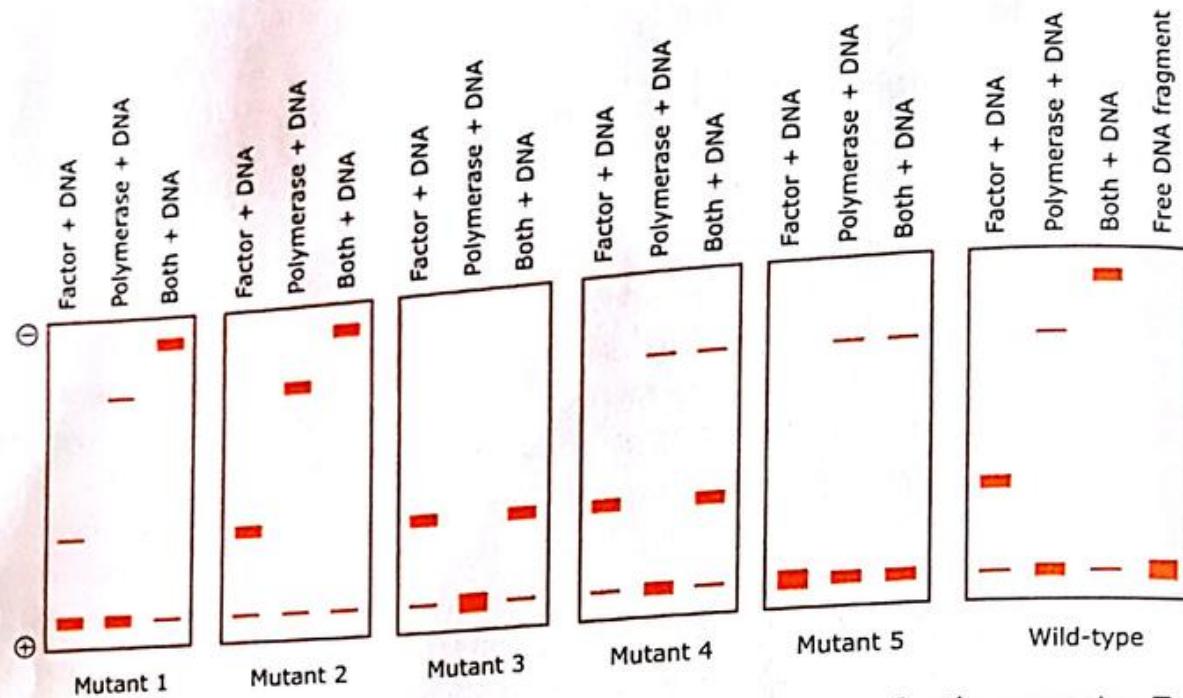
An investigator discovers a new receptor for a known ligand and wanted to identify the binding partner of the receptor i.e. its co-receptor. The antireceptor antibody is not available but anti GFP-antibody is available. Which one of the following strategies is most likely to identify the co-receptor?

- a. The GFP-receptor fusion protein is expressed in a cell line and analyzed by LC-MS/MS.
- b. The GFP-receptor fusion protein is expressed in a cell line and the cells positive for GFP were sorted out, lysed and run on a polyacrylamide gel.
- c. The GFP-receptor protein is coated on ELISA plate, followed by ELISA with anti-GFP antibody.
- d. The receptor is cloned as a fusion protein of GFP and expressed in stimulated cells. The immunoprecipitated complex obtained by anti-GFP antibody was analyzed by LC-MS/MS.

Question 29

Five *E. coli* strains have been identified, each of which has a different mutation that disrupts the normal regulation of a particular operon. For each mutant strain, the mutation has been mapped to the promoter or the operator region; however, the exact sequence changes are not known for these mutations. It is known that the normal promoter/operator consists of a single binding site for a positively acting transcription factor located just upstream of the promoter itself. Short DNA fragments containing the promoter and the operator were subcloned from each of the five mutant strains and from the wild type, purified, and radiolabeled. These fragments were then incubated under conditions of DNA excess with either purified regulatory factor or RNA polymerase or with both polymerase and regulatory factor.

The resulting protein-DNA complexes were separated by electrophoresis, and the radioactive DNA fragments were detected by exposure to X-ray film, giving the results shown below. Electrophoresis is from top to bottom; the largest complexes run slowest.



- A. One of the mutations increases the affinity of the polymerase for the promoter. Transcription of the operon is not stimulated by the regulatory factor in this mutant. Which mutant is most likely to show this effect?
- Mutant 1
 - Mutant 2
 - Mutant 3
 - Mutant 4
- B. One of the mutations maps to the operator. Transcription of the operon is not stimulated by the regulatory factor in this mutant. Which mutant is most likely to show this effect?
- Mutant 1
 - Mutant 2
 - Mutant 3
 - Mutant 5

Question 30

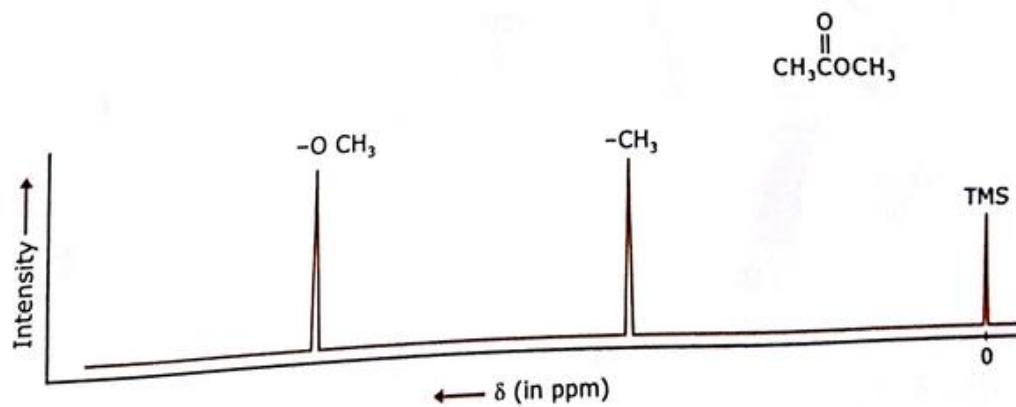
In purifying an enzyme that converts B into A in a single step, you discover that your enzyme activity disappears when you subject it to ion-exchange chromatography; none of the fractions from the ion-exchange column contains detectable enzyme activity. However, if you combine the material that passed straight through the column with a fraction that came off the column later, enzyme activity is again detectable. This is true even if you boil the first fraction (the 'pass-through') before combining it with the second. Give a reasonable explanation for these observations.

Answers

1. c
2. a
3. a
4. b

At pH 1, the average charges on both amino acids is positive, all migrate towards the cathode, although Lys migrates more rapidly than Glu. At pH 12, the average charge on two amino acids is negative, and thus all migrate toward the anode. In this case, however, Glu migrates more rapidly than Lys.

5. c
6. a
7. c
8. c
9. d
10. c
11. d
12. b
13. d



Besides showing the number of different kinds of protons, the ^1H -NMR spectrum also tells how many protons of each kind are present in a molecule. This is shown by the intensity of the signal. The intensity of a signal is measured in terms of the area

under that peak. The area under an NMR signal is directly proportional to the number of protons giving rise to that signal.

14. c
The wavelengths of absorption of chromophores in electronic spectra are often influenced by the solvent. When a polar solvent is used, the dipole-dipole interaction with the solvent molecules lowers the energy of the excited state more than that of the ground state. This is due to the fact that excited states are more polar than ground states. The energy difference between the excited and ground states is reduced. This leads to a small red shift of the absorption maximum in polar solvents. Thus, the $\pi-\pi^*$ transition shows a red shift of the order of 10–20 nm when the solvent is changed from hexane (nonpolar) to ethanol (polar).

Solvent effect for $n-\pi^*$ transitions is opposite to that found for $\pi-\pi^*$ transitions. Polar solvents cause a shift to lower wavelengths (blue shift) relative to non-polar solvents. The lone-pair electrons in the non-bonding orbital hydrogen bond or otherwise interact strongly with the polar solvent, leading to a lowering in energy of the non-bonding orbital whereas π^* orbital is affected much less. The result is an increase in the transition energy on going from a less polar to a more polar solvent. For example, in hexane solution, acetone shows absorption maximum at 279 nm whereas in aqueous solution, the absorption maximum is at 264.5 nm.

15. b

According to Beer-Lambert law, $A = \epsilon \cdot c \cdot l$

One solves the pair of equations,

$$11,300 [\text{Tyr}] + 1960 [\text{Trp}] = 0.717 \quad \dots(1)$$

$$1500 [\text{Tyr}] + 5380 [\text{Trp}] = 0.239 \quad \dots(2)$$

Equation (1) can be written as,

$$[\text{Tyr}] = \frac{0.717 - 1960 [\text{Trp}]}{11300}$$

Putting this value in equation (2)

$$1500 \times \frac{(0.717 - 1960 [\text{Trp}])}{11300} + 5380 [\text{Trp}] = 0.239$$

$$10.755 - 29400 [\text{Trp}] + 607940 [\text{Trp}] = 27.007$$

$$578540 [\text{Trp}] = 16.252$$

$$[\text{Trp}] = 0.00002809 = 28.1 \mu\text{mol}$$

$$[\text{Tyr}] = \frac{0.717 - 1960 \times 0.00002809}{11300} = \frac{0.66212}{11300} = 0.0000586 = 58.6 \mu\text{mol.}$$

The solution is at a concentration of 0.1 g/l. Thus, the content of tryptophan and tyrosine is 281 and 586 $\mu\text{mol/g}$ protein, respectively.

16. c

Because both compounds show absorption at both wavelengths, we can set up two simultaneous equations.

$$A_{350} = \epsilon_{A350} \times [A] + \epsilon_{B350} \times [B]$$

$$0.36 = 15000 \times [A] + 7000 \times [B]$$

$$0.36 - 7000 \times [B] = 15000 \times [A]$$

$$15000 \times [A] = 0.36 - 0.21$$

$$[A] = \frac{0.15}{15000} = \frac{15}{1500000} = 1 \times 10^{-5} M$$

17. b

The concentration of NADH from its absorbance at 340 nm,
 $A = \epsilon \cdot c \cdot l$

$$0.311 = 6220 \times c \times l; \quad \text{where, } l = 1 \text{ cm}$$
$$c_{\text{NADH}} = 5 \times 10^{-5} M$$

The absorbance at 260 nm resulting from the NADH,
 $A = \epsilon \cdot c \cdot l$

$$A = (15000) (5 \times 10^{-5}) (1) = 75 \times 10^{-2} = 0.75$$

$$\begin{aligned} \text{Now, the absorbance at 260 nm from the NAD}^+ &= \text{total absorbance at 260 nm} - \text{absorbance of NADH at 260 nm} \\ &= 1.20 - 0.75 = 0.45 \end{aligned}$$

Finally, from the absorbance of the NAD⁺ at 260 nm, we can calculate the concentration of NAD⁺.

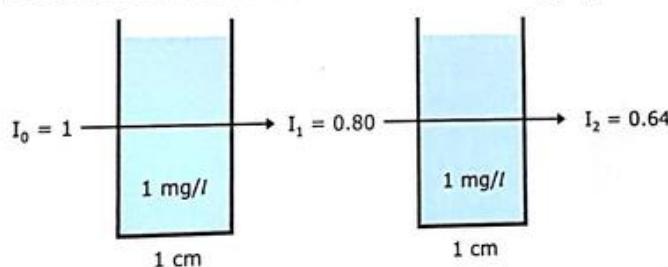
$$A = \epsilon \cdot c \cdot l$$

$$0.45 = (18 \times 10^3) (c) (1)$$

$$c = 2.5 \times 10^{-5} M$$

18. c

The Lambert-Beer law is not a linear relationship. Consider a beam of light passing through a 1 cm cuvette containing 1 mg/liter of a light-absorbing compound. Suppose 80% of the incident light is transmitted (20% of the incident light is absorbed). Now let us place a second identical cuvette in the light path directly behind the first cuvette.



What is the intensity of the light transmitted through both cuvettes? The Lambert-Beer law is not a linear relationship. Thus, I_2 is not 0.64. Each centimeter of path length does not absorb a constant amount of light. Instead, each centimeter absorbs 20% of the incident light. However, it absorbs 20% of 0.8, which is 0.16. It transmits 80% of the incident light. However, it transmits 80% of 0.8, which is 64% of the original incident light. We could obtain exactly the same result if we use a single cuvette with 2 cm thickness.

19. c

20. a

The peptide carries a net charge of +1.5 at pH 2.0 (Lys side chain, +1; α -amino group, +1; Glu side chain, 0; and terminal carboxyl, -0.5, since the pH coincides with its pK value). At pH 4.0, the net charge is +0.5; the Glu side chain is half ionized (-0.5), but the terminal carboxyl is almost completely ionized (-1). At pH 6.0, the net charge is 0 due to a +2 charge contributed by the Lys residue and a -2 charge contributed by the

Glu residue. At pH 11.0, the α -amino group is deprotonated (charge of 0) and the Lys side chain is half-protonated (charge of +0.5); thus, the net charge is -1.5.

21. a
22. c
23. d
24. c
25. a
Moving up the gel corresponds to reading the sequence of the extended primer from 5' to 3'. The sequence is 5' TACCTTGAT 3'.
The template sequence is the reverse complement of this sequence.
5' ATCAAGGTA 3'.
26. b
5' CTAGATAGTATAG 3'
Bands appearing in the G and C lanes can be read directly. Bands in the A+G lane that are not duplicated in the G lane are read as A. Bands in the C+T lane that are not duplicated in the C lane are read as T. The sequence is read from the bottom of the gel to top.
27. c
28. d
29. A. b B. d
30. The 'pass-through' contains a heat-stable coenzyme or cofactor, and the later-eluting fraction contains the enzyme itself. Neither is active without the other.

Fundamentals and Techniques of
**Biophysics and
Molecular Biology**



Pranav Kumar

Fundamentals and Techniques of

Biophysics and Molecular Biology

This text covers essential fundamentals and their applications of current techniques and methodologies. It is presented in a sharply focused manner without overwhelming or excessive detail. Each chapter contains information in a systematic and logically organized manner. It is a perfect introductory text for any reader needing to learn the basics of the subject, assuming they have little prior knowledge of the subject.

Pathfinder Publication

www.pathfinderpublishation.in