17MBU603B MOLECULAR BIOLOGY (3H – 3C)

Instruction Hours / week: L: 3 T: 0 P: 0 Marks: Internal: **40** External: **60** Total: **100**

End Semester Exam: 3 Hours

SCOPE

This course explores technologies using molecular biology, embryo manipulation, cell and tissue culture to manipulate the genomes of animals for ways to improve the live stock for food production and biomedical purpose.

OBJECTIVE

To provide an experience for the students in an interdisciplinary research program connecting animal genomics with animal reproduction and biotechnology.

Unit I

DNA Structure: Miescher to Watson and Crick - historic perspective, DNA Structure, Salient features of double helix, Types of DNA, Types of genetic material, denaturation and renaturation, cot curves. DNA topology - linking number, topoisomerases; Organization of DNA Prokaryotes, Viruses, Eukaryotes. RNA Structure, Organelle DNA -- mitochondria and chloroplast DNA.

Unit II

Bidirectional and unidirectional replication, semi- conservative, semi- discontinuous replication Mechanism of DNA replication: Enzymes and proteins involved in DNA replication –DNA polymerases, DNA ligase, primase, telomerase – for replication of linear ends. Various models of DNA replication including rolling circle, D- loop (mitochondrial), (theta) mode of replication and other accessory protein, Mismatch and excision repair.

Unit III

Transcription: Definition, difference from replication, promoter - concept and strength of promoter RNA Polymerase and the transcription unit, Transcription in Eukaryotes: RNA polymerases, general Transcription factors. Split genes, concept of introns and exons, RNA splicing, spliceosome machinery, concept of alternative splicing, Polyadenylation and capping, Processing of rRNA, RNA interference: si RNA, miRNA and its significance

Unit IV

Translational machinery, Charging of tRNA, aminoacyl tRNA synthetases, Mechanisms of initiation, elongation and termination of polypeptides in both prokaryotes and eukaryotes, Fidelity of translation, Inhibitors of protein synthesis in prokaryotes and eukaryote

Unit V

Principles of transcriptional regulation, regulation at initiation with examples from *lac* and *trp* operons, Sporulation in *Bacillus,* Yeast mating type switching, Changes in Chromatin Structure - DNA methylation and Histone Acetylation mechanisms.

SUGGESTED READINGS

1. Watson JD, Baker TA, Bell SP, Gann A, Levine M and Losick R (2008) Molecular Biology of the Gene, 6th edition, Cold Spring Harbour Lab. Press, Pearson Publication.
2. Becker WM, Kleinsmith LJ, Hardin J and Bertoni GP (2009) The World of the Cell, 7th edition, Pearson Benjamin Cummings Publishing, San Francisco.
3. De Robertis EDP and De Robertis EMF (2006) Cell and Molecular Biology, 8th edition. Lippincott Williams and Wilkins, Philadelphia.
4. Karp G (2010) Cell and Molecular Biology: Concepts and Experiments, 6th edition, John Wiley & Sons. Inc.
5. Sambrook J and Russell DW. (2001). Molecular Cloning: A Laboratory Manual. 4th Edition, Cold Spring Harbour Laboratory press.
6. Krebs J, Goldstein E, Kilpatrick S (2013). Lewin's Essential Genes, 3rd Ed., Jones and Bartlett Learning.
7. Gardner EJ, Simmons MJ, Snustad DP (2008).). Principles of Genetics. 8th Ed. Wiley-India.

UNIT I

Duration	Topic to be covered	Reference
1	DNA Structure: Miescher to Watson and Crick- historic perspective	T1: 88-90
1	DNA Structure, Salient features of double helix	R1 :287
1	Types of DNA, Types of genetic material	R1: 281
1	Denaturation and renaturation, cot curves	T1: 85-88
1	DNA topology - linking number, topoisomerases	T1:1209-1213
1	Organization of DNA Prokaryotes, Viruses, Eukaryotes	R1: 962- 965
1	Organization of DNA in Eukaryotes	R1: 962- 965
1	RNA Structure, Organelle DNA	T1:1264
1	Mitochondria and chloroplast DNA.	T2: 99- 103
1	Revision	

TOTAL HOURS – 10 h

References

T1: Donald Voet and Judith Voet; 2012, Biochemistry, 4th Edition, John Wiley and Sons. Inc.

T2: R.C. Dubey and D.K. Maheshwari; 2014, TextBook of Microbiology, S. Chand Publication, New Delhi.

R1: Daniel L. Nelson and Michael M. Cox; 2005 Biochemistry – Lehininger 5th Edition, W. H. Freeman and Company, Newyork.

UNIT II

Duration	Topic to be covered	Reference
1	Bidirectional and unidirectional replication,	T1: 1190-1195;
1	Semi- conservative, semi- discontinuous replication	T1: 1188; T1: 1208
1	Mechanism of DNA replication	T2: 149-156
1	Enzymes and proteins involved in DNA replication	T2: 149-156
1	DNA polymerases, DNA ligase, primase, telomerase – for replication of linear ends.	T1: 1176-1181; T1: 1126- 1127; T1: 1181-1187
1	Various models of DNA replication including rolling circle, D-loop,	T2:156-157
1	Theta mode of replication and other accessory protein	T2:156-157
1	Mismatch and excision repair.	T1: 1220 T1: 1216-1217
1	Revision	
1	Unit Test	

TOTAL HOURS – 10 h

References

T1: Donald Voet and Judith Voet; 2012, Biochemistry, 4th Edition, John Wiley and Sons. Inc.

T2: R.C. Dubey and D.K. Maheshwari; 2014, TextBook of Microbiology, S. Chand Publication, New Delhi.

UNIT III

Duration	Topic to be covered	Reference
1	Transcription: Definition, difference from replication	T3: 297-314
1	Promoter - concept and strength of promoter RNA Polymerase and the transcription unit	T1: 1265; T1:1266; T1: 1267-69; T1: 1275-1277
1	Transcription in Eukaryotes	T1: 1277-1279
1	RNA polymerases, general Transcription factors.	T1: 1277-1279
	Split genes, concept of introns and exons,	T1: 1302-1304
1	RNA splicing, spliceosome machinery,	T1: 1307
1	Concept of alternative splicing, Polyadenylation and capping	T1: 1288- 1302
1	Processing of rRNA, RNA interference: si RNA, miRNA and its significance	T1: 1307
1	Revision	
1	Unit Test	

TOTAL HOURS – 10 h

References

T1: Donald Voet and Judith Voet; 2012, Biochemistry, 4th Edition, John Wiley and Sons. Inc.

T3: Joanne M. Willey; Linda M. Sherwood and Christopher J. Woolverton; Prescott's Microbiology, Eighth Edition; Mc Graw Hill International Edition.

UNIT IV

Duration	Topic to be covered	Reference
1	Translational machinery,	T1: 1363-1365
1	Charging of tRNA, aminoacyl tRNA synthetases,	T1: 1365
1	Mechanisms of initiation, elongation of polypeptides in prokaryotes	T1: 1373-1394
1	Termination of polypeptides in prokaryotes	T1: 1373-1394
1	Mechanisms of initiation, elongation of polypeptides in Eukaryotes	T1: 1373-1394
1	Termination of polypeptides in Eukaryotes	T1:1395-1396
1	Fidelity of translation	T1:1395-1396
1	Inhibitors of protein synthesis in prokaryotes	T1:1408-1421
1	Inhibitors of protein synthesis in eukaryotes	T1:1408-1421
1	Revision	

TOTAL HOURS – 10 h

References

T1: Donald Voet and Judith Voet; 2012, Biochemistry, 4th Edition, John Wiley and Sons. Inc.

UNIT V

Duration	Topic to be covered	Reference
1	Principles of transcriptional regulation	T2: 308-314
1	Regulation at initiation with examples from *lac* operons	T1:1262-1266
1	Regulation at initiation with examples from *trp* operons	T1:1262-1266
1	Sporulation in *Bacillus,*	T2: 560-561
1	Yeast mating type switching	T2: 560-561
1	Changes in Chromatin Structure	T2: 373
1	DNA methylation	T2: 373
1	Histone Acetylation mechanisms.	T2: 71, 97-98, 765
1	Revision	
1	Discussion of last 3 year ES-QP	

TOTAL HOURS – 10 h

References

T1: Donald Voet and Judith Voet; 2012, Biochemistry, 4th Edition, John Wiley and Sons. Inc.

T2: R.C. Dubey and D.K. Maheshwari; 2014, Text Book of Microbiology, S. Chand Publication, New Delhi.

Unit I

DNA Structure: Mieschertо Watson and Crick-historic perspective, DNA Structure, Salient features of double helix, Types of DNA, Types of genetic material, denaturation and renaturation, cot curves. DNA topology - linking number, topoisomerases; Organization of DNA Prokaryotes, Viruses, Eukaryotes. RNA Structure, Organelle DNA -- mitochondria and chloroplast DNA.

DNA: Types, Structure and Function

Historical:

Nucleic acids were first isolated by Friedrich Miescher (1869) from pus cells.

They were named nuclein. Hertwig (1884) proposed nuclein to be the carrier of hereditary traits. Because of their acidic nature they were named nucleinic acids and then nucleic acids (Altmann, 1899).

Fisher (1880s) discovered the presence of purine and pyrimidine bases in nucleic acids. Levene (1910) found deoxyribose nucleic acid to contain phosphoric acid as well as deoxyribose sugar.

He characterised four types of nucleotides present in DNA. In 1950, Chargaff found that purine and pyrimidine content of DNA was equal. By this time W.T. Astbury had found through X-ray diffraction that DNA is a polynucleotide with nucleotides arranged perpendicular to the long axis of the molecule and separated from one another by a distance of 0.34 nm.

In 1953, Wilkins and Franklin got very fine X-ray photographs of DNA. The photographs showed that DNA was a helix with a width of 2.0 nm. One turn of the helix was 3.4 nm with 10 layers of bases stacked in it. Watson and Crick (1953) worked out the first correct double helix model from the X-ray photographs of Wilkins and Franklin. Wilkins, Watson and Crick were awarded Nobel Prize for the same in 1962.

Watson and Crick (1953) built a 3D, molecular model of DNA that satisfied all the details obtained from X-ray photographs. They proposed that DNA consisted of a double helix with two chains having sugar phosphate on the outside and nitrogen bases on the inner side.

The nitrogen bases of the two chains formed complementary pairs with purine of one and pyrimidine of the other held together by hydrogen bonds (A-T, C-G). Complementary base

pairing between the two polynucleotide chains is considered to be hall mark of their proposition. It is of course based on early finding of Chargaff that A = T and C = G Their second big proposal was that the two chains are antiparallel with 5'→ 3′ orientation of one and У → 5'orientation of the other.

The two chains are twisted helically just as a rope ladder with rigid steps twisted into a spiral. Each turn of the spiral contains 10 nucleotides. This double helix or duplex model of DNA with antiparallel polynucleotide chains having complementary bases has an implicit mechanism of its replication and copying.

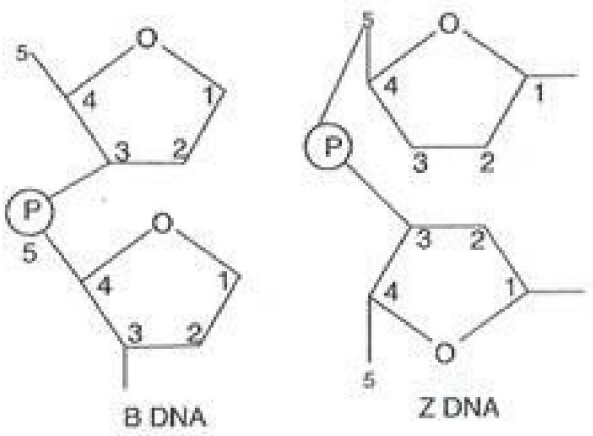
Here both the polynucleotide chains function as templates forming two double helices, each with one parent chain and one new but complementary strand. The phenomenon is called semi conservative replication. In vitro synthesis of DNA has been carried out by Kornberg in 1959.

Types of DNA:

DNA duplex model proposed by Watson and Crick is right handed spiral and is called B-DNA (Balanced DNA). In the model the base pairs lie at nearly right angles to the axis of helix (Fig. 6.5 D). Another right handed duplex model is A-DNA (Alternate DNA). Here, a single turn of helix has 11 base pairs.

The base pairs lie 20° away from perpendicular to the axis. C-DNA has 9 base pairs per turn of spiral while in D-DNA the number is only 8 base pairs. Both are right handed. Z-DNA (Zigzag DNA) is left-handed double helix with zigzag back-bone, alternate purine and pyrimidine bases, single turn of 45 A length with 12 base pairs and a single groove.

B-DNA is more hydrated and most frequently found DNA in living cells. It is physiologically and biologically active form. However, it can get changed into other forms. Right handed DNA is known to change temporarily into the left handed form at least for a short distance. Such changes may cause changes in gene expression.



Orientation of adjacent sugar molecules in B and Z DNA.

Circular and Linear DNA:

In many prokaryotes the two ends of a DNA duplex are covalently linked to form circular DNA. Circular DNA is naked, that is, without association with histone proteins, though polyamines do occur. In linear DNA the two ends are free. It is found in eukaryotic nuclei where it is associated with histone proteins.

Linear DNA, without association with histone proteins, also occurs in some prokaryotes, e.g., Mycoplasma. In semi-autonomous cell organelles (mitochondria, plastids) DNA is circular, less commonly linear. It is always naked.

Chargaff's Rules:

Chargaff (1950) made observations on the bases and other components of DNA. These observations or generalizations are called Chargaff's base equivalence rule.

(i) Purine and pyrimidine base pairs are in equal amount, that is, adenine + guanine = thymine + cytosine. [A + G] = [T + C], i.e., [A+G] / [T+C] = 1

(ii) Molar amount of adenine is always equal to the molar amount of thymine. Similarly, molar concentration of guanine is equalled by molar concentration of cytosine.

[A] = [T], i.e., [A] / [T] = 1; [G] = [C], i.e., [G] / [C] = 1

(iii) Sugar deoxyribose and phosphate occur in equimolar proportions.

(iv) A-T base pairs are rarely equal to C—G base pairs.

(v) The ratio of [A+T] / [G+C] is variable but constant for a species (Table 6.2). It can be used to identify the source of DNA. The ratio is low in primitive organisms and higher in advanced ones.

Table 6.2. Base Composition of DNA from Various Sources:

Species	A	G	C	T	A+T/C+G
1. Man	30.4	19.0	19.9	30.1	1.55
2. Calf	29.0	21.2	21.2	28.5	1.35
3. Wheat germ	28.1	21.8	22.7	27.4	1.25
4. Pea	30.8	19.2	18.5	30.5	1.62
5. Euglena	22.6	27.7	25.8	24.4	0.88
6 Escherichia coli	24.7	26.0	25.7	23.6	0.93

Structure of DNA:

DNA or deoxyribonucleic acid is a helically twisted double chain polydeoxyribonucleotide macromolecule which constitutes the genetic material of all organisms with the exception of rhinoviruses. In prokaryotes it occurs in nucleoid and plasmids. This DNA is usually circular. In eukaryotes, most of the DNA is found in chromatin of nucleus.

It is linear. Smaller quantities of circular, double stranded DNA are found in mitochondria and plastids (organelle DNA). Small sized DNAs occur in viruses, ϕ x 174 bacteriophage has 5386 nucleotides. Bacteriophage lambda (Phage X) possesses 48502 base pairs (bp) while number of base pairs in Escherichia coli is 4.6×10^6. A single genome (haploid set of 23 chromosomes) has about 3.3×10^9 bp in human beings. Single-stranded DNA occurs as a genetic material in some viruses (e.g., phage ϕ x 174, coliphage fd, M_{13}). DNA is the largest

macromolecule with a diameter of 2 nm (20Å or 2 x 10^{-9}m) and often having 3 length in millimetres

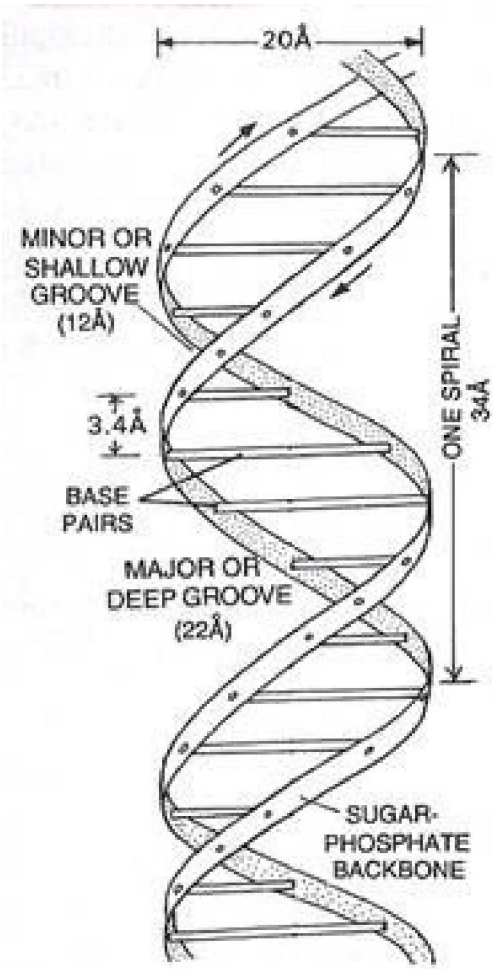
It 15 negatively charged due to phosphate groups. It is a long chain polymer of generally several hundred thousands of deoxyribonucleotides. A DNA molecule has two un-branched complementary strands. They are spirally coiled. The two spiral strands of DNA are collectively called DNA duplex (Fig. 6.5).

The two strands are not coiled upon each other but the whole double strand (DNA duplex) is coiled upon itself around a common axis like a rope stair case with solid steps twisted into a spiral. Due to spiral twisting, the DNA duplex comes to have two types of alternate grooves, major (22 Å) and minor (12 Å).

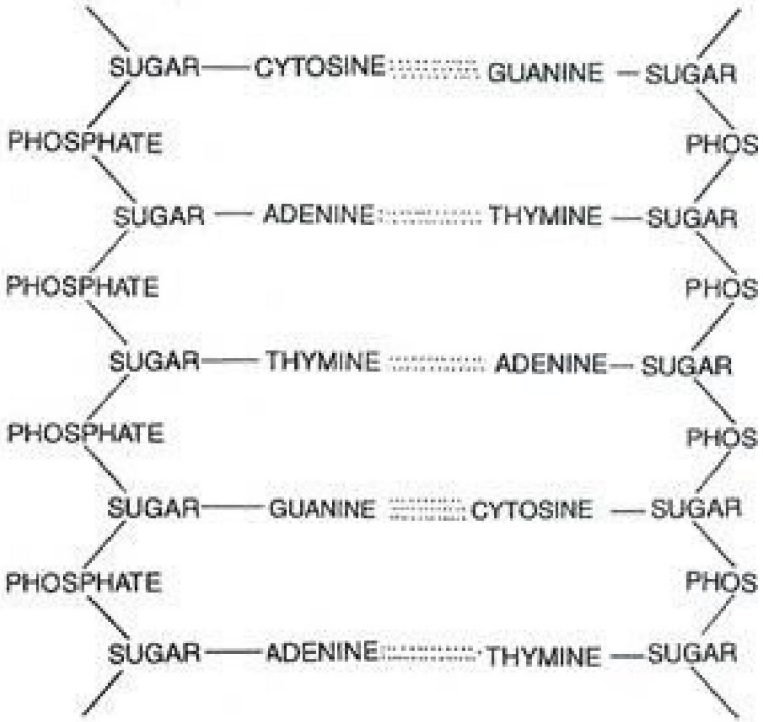
In B-DNA, one turn of the spiral has about 10 nucleotides on each strand of DNA. It occupies a distance of about 3.4 nm (34 Å or 3.4×10^{-9} m) so that adjacent nucleotides or their bases are separated by a space of about 0.34 nm (0.34×10^{-9} m or 3.4 Å).

A deoxyribonucleotide of DNA is formed by cross-linking of three chemicals ortho-phosphoric acid (H_3PO_4), deoxyribose sugar ($C_5H_{10}O_4$) and a nitrogen base. Four types of nitrogen bases occur in DNA. They belong to two groups, purines (9-membered double rings with nitrogen at 1,3,7 and 9 positions) and pyrimidines (six membered rings with nitrogen at 1 and 3 positions). DNA has two types of purines (adenine or A and guanine or G) and two types of pyrimidines (cytosine or C and thymine or T).

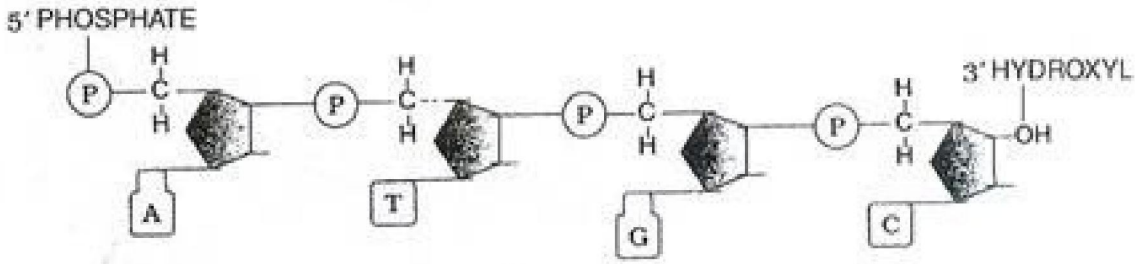
.


20Å
MINOR OR SHALLOW GROOVE (12Å)
3.4Å
BASE PAIRS
MAJOR OR DEEP GROOVE (22Å)
ONE SPIRAL 34Å
SUGAR-PHOSPHATE BACKBONE

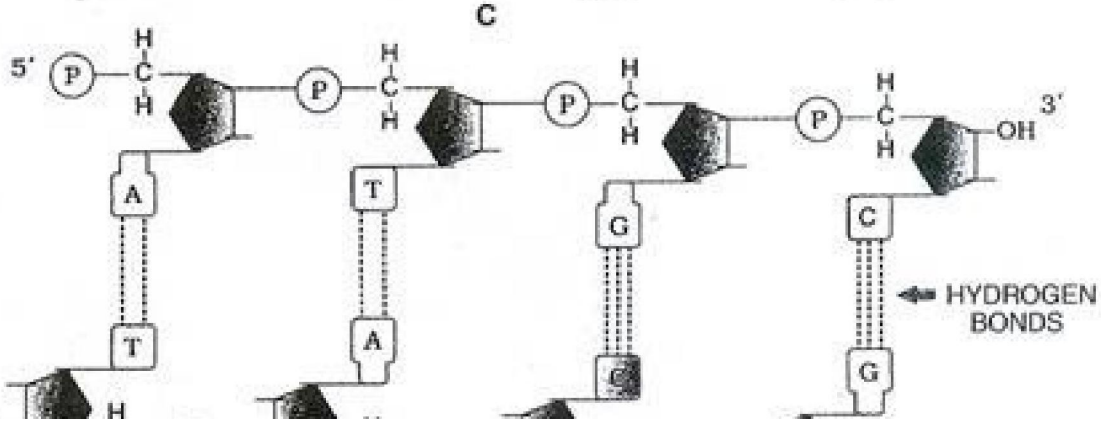
A

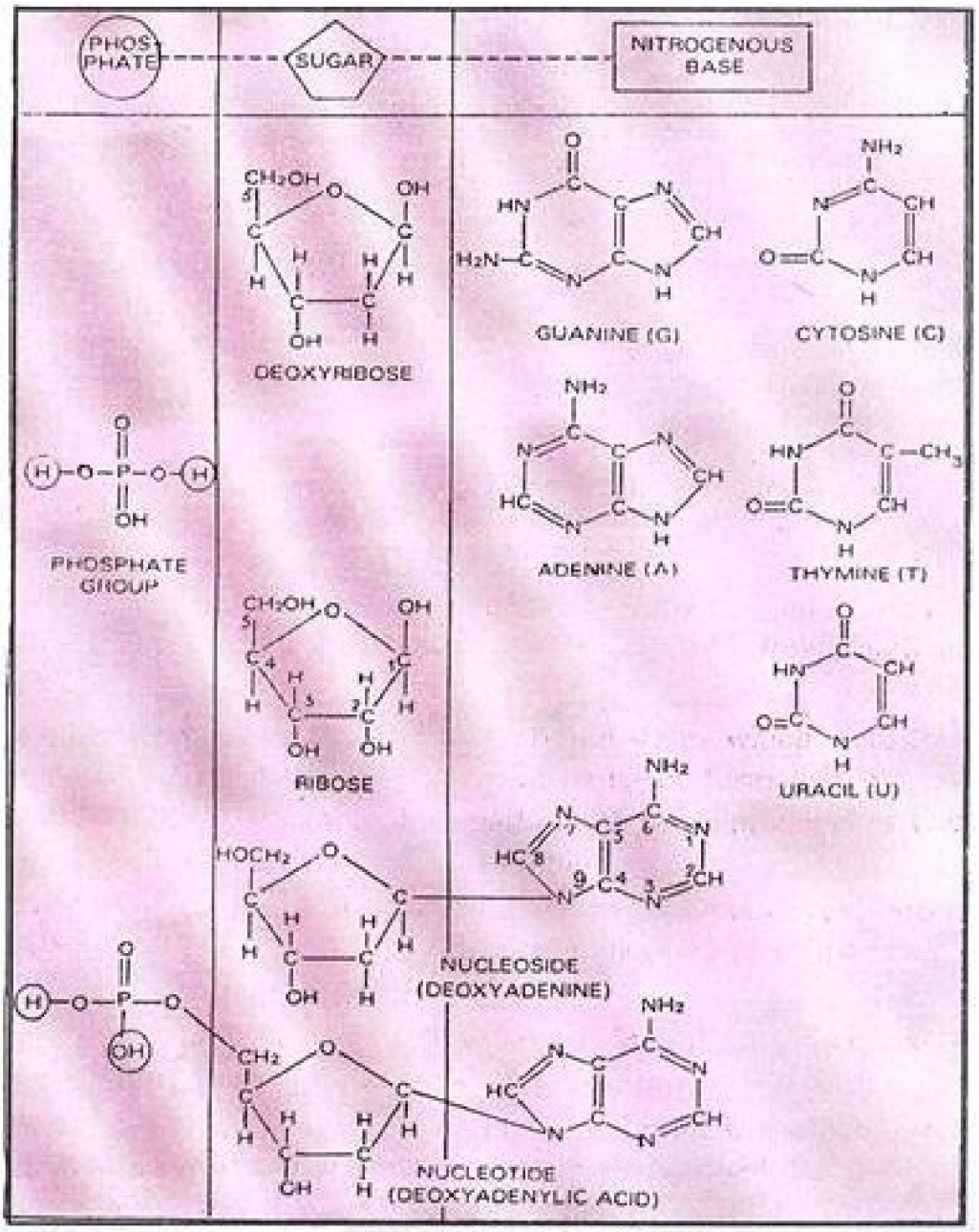

SUGAR — CYTOSINE ····· GUANINE — SUGAR
PHOSPHATE PHOS
SUGAR — ADENINE ····· THYMINE — SUGAR
PHOSPHATE PHOS
SUGAR — THYMINE ····· ADENINE — SUGAR
PHOSPHATE PHOS
SUGAR — GUANINE ····· CYTOSINE — SUGAR
PHOSPHATE PHOS
SUGAR — ADENINE ····· THYMINE — SUGAR

B


5' PHOSPHATE
3' HYDROXYL
OH
A
T
G
C

C


5'
3'
OH
A
T
G
C
T
A
G
C
HYDROGEN BONDS



Building blocks of nucleotides

Depending upon the type of nitrogen base, DNA has four kinds of deoxyribonucleotides — deoxy adenosine 5- monophosphate (d AMP), deoxy guaninosine 5-monophosphate (d GMP), deoxy thymidine 5-monophosphate (d TMP) and deoxy cytidine 5- monophosphate (d CMP).

The back bone of a DNA chain or strand is built up of alternate deoxyribose sugar and phosphoric acid groups. The phosphate group is connected to carbon 5′ of the sugar residue of its own nucleotide and carbon Ɣ of the sugar residue of the next nucleotide by (3 '—5') phosphodiester bonds. -H of phosphate and -OH of sugar are eliminated as H_20 during each ester formation.

Phosphate group provides acidity to the nucleic acids because at least one of its side group is free to dissociate. Nitrogen bases lie at right angles to the longitudinal axis of DNA chains. They are attached to carbon atom 1 of the sugars by N-glycosidic bonds. Pyrimidine (C or T) is attached to deoxyribose by its N-atom at 1 position while a purine (A or G) does so by N-atom at 9 position.

The two DNA chains are antiparallel that is, they run parallel but in opposite directions. In one chain the direction is 5'→ Ɣ while in the opposite one it is 3′ →5′ (Fig. 6.5). The two chains are held together by hydrogen bonds between their bases. Adenine (A), a purine of one chain lies exactly opposite thymine (T), a pyramidine of the other chain. Similarly, cytosine (C, a pyrimidine) lies opposite guanine (G a purine). This allows a sort of lock and key arrangement between large sized purine and small sized pyrimidine.

It is strengthened by the appearance of hydrogen bonds between the two. Three hydrogen bonds occur between cytosine and guanine (C = G) at positions 1'-1', 2′- 6′ and 6′-2′. There are two such hydrogen bonds between adenine and thymine (A=T) which are formed at positions 1'-3′ and 6′-4′. Hydrogen bonds occur between hydrogen of one base and oxygen or nitrogen of the other base. Since specific and different nitrogen bases occur on the two DNA chains, the latter are complementary.

Thus the sequence of say AAGCTCAG of one chain would have a complementary sequence of TTCGAGTC on the other chain. In other words, the two DNA chains are not identical but complementary to each other. It is because of specific base pairing with a purine lying opposite a pyrimidine. This makes the two chains 2 nm thick.

A purine- purine base pair will make it thicker while a pyrimidine- pyrimidine base pair will make it narrower than 2 nm. Further, A and C or G and T do not pair because they fail to form hydrogen bonds between them. 5′ end of each chain bears phosphate radical while the 3′ end possesses a sugar residue (3'-OH).

Salient Features of B model of DNA of Watson and Crick:

1. DNA is the largest biomolecule in the cell.

2. DNA is negatively charged and dextrorotatory.

3. Molecular configuration of DNA is 3D.

4. DNA has two polynucleotide chains.

5. The two chains of DNA have antiparallel polarity, 5′ —> 3′ in one and 3′ —> 5′ in other.

6. Backbone of each polynucleotide chain is made of alternate sugar-phosphate groups. The nitrogen bases project inwardly.

7. Nitrogen bases of two polynucleotide chains form complementary pairs, A opposite T and C opposite G.

8. A large sized purine always comes opposite a small sized pyrimidine. This generates uniform distance between two strands of helix.

9. Adenine (A) of one polynucleotide chain is held to thymine (T) of opposite chain by two hydrogen bonds. Cytosine (C) of one chain is similarly held to guanine of the other chain by three hydrogen bonds.

10. The double chain is coiled in a helical fashion. The coiling is right handed. This coiling produces minor and major grooves alternately.

11. The pitch of helix is 3.4 nm (34 A) with roughly 10 base pairs in each turn. The average distance between two adjacent base pairs comes to about 0.34 nm (0.34×10^{-9} m or 3.4 A).
12. Planes of adjacent base pairs are stacked over one another. Alongwith hydrogen bonding, the stacking confers stability to the helical structure.

13. DNA is acidic. For its compaction, it requires basic (histone) proteins. The histone proteins are +vely charged and occupy the major grooves of DNA at an angle of 30° to helix axis.

Sense and Antisense Strands:

Both the strands of DNA do not take part in controlling heredity and metabolism. Only one of them does so. The DNA strand which functions as template for RNA synthesis is known as template strand, minus (-) strand or antisense strand.

Its complementary strand is named nontemplate strand, plus (+) strand, sense and coding strand. The latter name is given because by convention DNA genetic code is written according to its sequence.

(5') GCATTCGGCTAGTAAC (3′)

DNA Nontemplate, Sense (+) or Coding Strand

(3′) CGTAAGCCGATCATTG (5′)

DNA Template, Antisense, or Noncoding or (-) Strand

(5′) GCAUUCGGCUAGUAAC (3′)

RNA Transcript

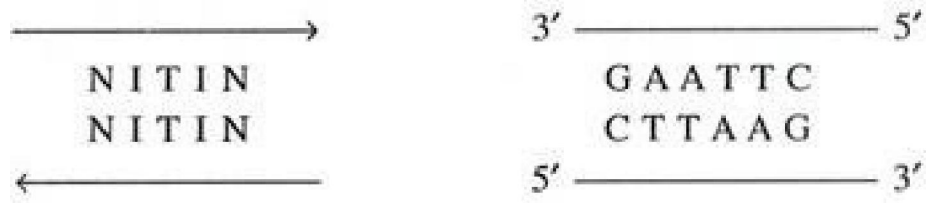
RNA is transcribed on 3'→5′ (-) strand (template/antistrand) of DNA in 5 → 3 direction.

The term antisense is also used in wider prospective for any sequence or strand of DNA (or RNA) which is complementary to mRNA.

Types of DNA

Palindromic DNA:

DNA duplex possesses areas where sequence of nucleotides is the same but opposite in the two strands. These sequences are recognised by restriction endonucleases and are used in genetic engineering. Given hereunder sequence of bases in one strand (3′ → 5′) is GAATTC. It is same in other strand when read in 5′ → 3′ direction. It is similar to palindrome words having same words in both forward and backward direction, e.g., NITIN, MALAYALAM.



Repetitive DNA:

It is the DNA having multiple copies of identical sequences of nitrogen bases. The number of copies of the same base sequence varies from a few to millions. DNA having single copy of base sequences is called unique DNA. It is made of functional genes. rRNA genes are, however, repeated several times. Repetitive DNA may occur in tandem or interspersed with unique sequences.

It is of two types, highly repetitive and moderately repetitive. Highly repetitive DNA consists of short sequences of less than 10 base pairs which are repeated millions of times. They occur in preeentromeric regions, heterochromatic regions of Y-chromosomes and satellite regions. Moderately repetitive dna consists of a few hundred base pairs repeated at least 1000 times. It occurs in telomeres, centromeres and transposons.

Tandemly repeated sequences are especially liable to undergo misalignments during chromosome pairing, and thus the size of tandem clusters tends to be highly polymorphic, with wide variations between individuals. Smaller clusters of such sequences can be used to characterize individual genomes in the technique of "DNA-finger-printing".

Satellite DNA:

It is that part of repetitive DNA which has long repetitive nucleotide sequences in tandem that forms a separate fraction on density ultracentrifugation. Depending upon the number of base pairs involved in repeat regions, satellite DNA is of two types, microsatellite sequences (1-6 bp repeat units flanked by conserved sequences) and minisatellite sequences (11-60 bp flanked by conserved restriction sites). The latter are hyper variable and are specific for each individual. They are being used for DNA matching or finger printing as first found out by Jeffreys et al (1985).

Genetic Information:

The arrangement of nitrogen bases of DNA (and its product mRNA) determines the sequence of amino acid groups in polypeptides or proteins formed over ribosomes. One amino acid is

specified by the sequence of three adjacent nitrogen bases. The latter is called codon. The segment of DNA which determines the synthesis of complete polypeptide is known as cistron.

In procaryotes, a cistron has a continuous coding sequence from beginning to end. In eucaryotes a cistron contains noncoding regions which do not produce part of gene product. They are called introns. Introns are often variable. The coding parts are known as exons. Cistrons having introns are called split genes.

Coding and Noncoding DNA:

Depending on the ability to form functional or nonfunctional products, DNA has two types of segments, coding and noncoding. In eukaryotes a greater part of DNA is noncoding since it does not form any functional product. They often possess repeated sequences or repetitive DNA. Most of them have fixed positions.

Some can move from one place to another. The mobile sequences are called jumping genes or transposons. In prokaryotes the amount of noncoding or nonfunctional DNA is small. Coding DNA consists of coding DNA sequences. These are of 2 types — protein coding sequences coding for all proteins except histone and nonprotein coding sequences for tRNA, rRNA and histones.

Denaturation and Renaturation:

The H-bonds between nitrogen bases of two strands of DNA can break due to high temperature (82-90°C) or low or high pH, so that the two strands separate from each other. It is called denaturation or melting. Since A-T base pair has only 2H bonds, the area rich in A-T base pairs can undergo easy denaturation (melting). These areas are called low melting areas because they denature at comparatively low temperature. The area rich in G- C base pairs (called high melting area) is comparatively more stable and dense because three hydrogen bonds connect the G-C bases.

These areas have high temperature of melting (Tm). On melting the viscosity of DNA decreases. The denatured DNA has the tendency to reassociate, i.e., the DNA strands separated by melting at 82-90°C can reassociate and form duplex on cooling to temperature at 65°C. It is called renaturation or annealing.

Denatured or separated DNA strands absorb more light energy than the intact DNA double strand. The increased absorption of light energy by separated or denatured DNA strands is called hyperchromatic effect. The effect is used in knowing whether DNA is single or double stranded.

Functions of DNA:

1. Genetic Information (Genetic Blue Print):

DNA is the genetic material which carries all the hereditary information. The genetic information is coded in the arrangement of its nitrogen bases.

2. Replication:

DNA has unique property of replication or production of carbon copies (Autocatalytic function). This is essential for transfer of genetic information from one cell to its daughters and from one generation to the next.

3. Chromosomes:

DNA occurs inside chromosomes. This is essential for equitable distribution of DNA during cell division.

4. Recombination's:

During meiosis, crossing over gives rise to new combination of genes called recombinations.

5. Mutations:

Changes in sequence of nitrogen bases due to addition, deletion or wrong replication give rise to mutations. Mutations are the fountain head of all variations and evolution.

6. Transcription:

DNA gives rise to RNAs through the process of transcription. It is heterocatalytic activity of DNA.

7. Cellular Metabolism:

It controls the metabolic reactions of the cells through the help of specific RNAs, synthesis of specific proteins, enzymes and hormones.

8. Differentiation:

Due to differential functioning of some specific regions of DNA or genes, different parts of the organisms get differentiated in shape, size and functions.

9. Development:

DNA controls development of an organism through working of an internal genetic clock with or without the help of extrinsic information.

10. DNA Finger Printing:

Hypervariable microsatellite DNA sequences of each individual are distinct. They are used in identification of individuals and deciphering their relationships. The mechanism is called DNA finger printing.

11. Gene Therapy:

Defective heredity can be rectified by incorporating correct genes in place of defective ones.

12. Antisense Therapy:

Excess availability of anti-mRNA or antisense RNAs will not allow the pathogenic genes to express themselves. By this technique failure of angioplasty has been checked. A modification of this technique is RNA interference (RNAi).

Cot Curve:

Principle behind Cot curve: The rate at which a particular sequence of a DNA reassociates(renatures) is proportional to the number of times it is found in the DNA molecule.

That means,

Rate of Renaturation of a sequence of a DNA molecule ® number of times the sequence is present in the DNA molecule.

As greater the number of repetitive sequences, the more the chances of two complementary strands colliding with each other.

STEPS IN COT ANALYSIS

1. A genomic DNA is first cut randomly into fragments about 1kb in length.
2. Fragments are completely denatured by heating above the Melting temperature Tm(temperature at which 50% of the DNA is melted.
3. Fragments are allowed to renature(renaturation at a temperature about 10°C below the Tm).

4. Renaturation at different time intervals can be monitored by two ways:

a)by using Spectrophotometer(by taking absorbance at 260nm- as ssDNA absorbs light more efficiently than ds DNA (due to base stacking it has more constrained bases).

b) by using Hydroxylapitite chromatography(HC)- an adsorption chromatography, Hydroxylapitite (modified form of crystalline calcium phosphate) adsorb double stranded DNA as the positively charged calcium in hydroxylapitite will have more affinity towards double stranded DNA as dsDNA consists of more negatively charged phosphate groups than ssDNA.

Cot curve: When the extent of renaturation of DNA is plotted against log ***Cot*** , this is known as Cot curve.

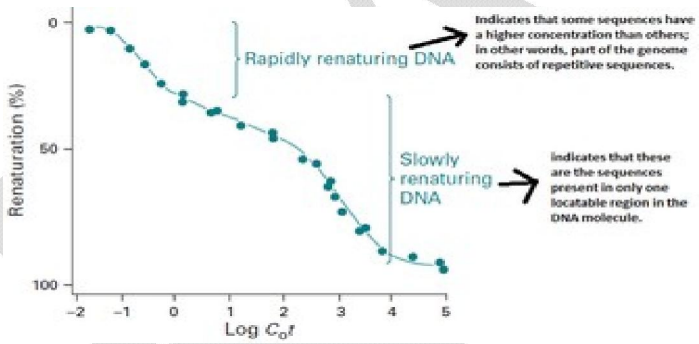
Degree of Renaturation after a given time depends upon ***C***o and ***t.***

Cot value = ***C***o × ***t*** × ***buffer factor***(if applicable)

Co = nucleotide concentration in moles per liter prior to denaturation.

t = reassociation time in seconds.

buffer factor= a factor based upon the cation concentration of the buffer.



SIGNIFICANCE OF COT CURVE:

1. Estimation of genome size.
2. Estimation of the proportion of single-copy(sequences present in only one locatable region in the dna molecule) and repetitive sequences.
3. Estimation of kinetic complexity of genome.
4. Interspecific comparison of Cot data has provided considerable insight into the structure and evolution of eukaryotic genomes.

DNA Topology:

DNA is essentially an extremely long double-stranded rope in which the two strands are wound about one another. As a result, **topological** properties of the genetic material, including **DNA** underwinding and overwinding, knotting, and tangling, profoundly influence virtually every major nucleic acid process.

Linking Number:

The helix-helix crossovers (i.e., nodes) are assigned a positive or negative value based on the orientation (i.e., handedness) of the **DNA** axis. The numerical term that describes the sum of the twist and the **writhe** is called the **linking number**, which represents the total **linking** within a **DNA** molecule.

Discovery

In the 1970s, James C. Wang was the first to discover a topoisomerase when he identified *E. coli* topoisomerase I.

Function

The double-helical configuration of DNA strands makes them difficult to separate, which is required by helicase enzymes if other enzymes are to transcribe the sequences that encode proteins, or if chromosomes are to be replicated. In circular DNA, in which double-helical DNA is bent around and joined in a circle, the two strands are topologically linked, or knotted. Otherwise identical loops of DNA, having different numbers of twists, are topoisomers, and cannot be interconverted without the breaking of DNA strands. Topoisomerases catalyze and guide the unknotting or unlinking of DNA by creating transient breaks in the DNA using a conserved tyrosine as the catalytic residue.

The insertion of (viral) DNA into chromosomes and other forms of recombination can also require the action of topoisomerases.

Topologically linked circular molecules, aka catenanes, adopt a positive supercoiled form during the process of replication of circular plasmids. The unlinking of catenanes is performed by type IIA topoisomerases, which were recently found to be more efficient unlinking positive supercoiled DNA.The conformational properties of negative vs. positive supercoiled catenanes affects their features in respect to their corresponding enzymatic reaction catalyzed by topoisomerases. Experiments have demonstrated that positive supercoiled DNA provides a sharp DNA bend in the first bound DNA segment, which allows the topoisomerase to bind successfully and therefore carry on its enzymatic reaction to the following segment in a specific inside-to-outside matter. On the other hand, negative supercoiled DNA does not provide such bend and the access of the enzyme to the first segment is nearly impossible, therefore inhibiting unlinking

Topoisomerase is also found in the mitochondria of cells. The mitochondria generates ATP as well as plays a role in programmed cell death and aging.The mitochondrial DNA is a circular, double-stranded DNA that requires the activity of topoisomerase to be replicated. The classes of topoisomerase found in the mitochondria are I, IIβ, IIIα.

Clinical significance

Many drugs operate through interference with the topoisomerases. The broad-spectrum fluoroquinolone antibiotics act by disrupting the function of bacterial type II topoisomerases. These small molecule inhibitors act as efficient anti-bacterial agents by hijacking the natural ability of topoisomerase to create breaks in chromosomal DNA.

Some chemotherapy drugs called topoisomerase inhibitors work by interfering with mammalian-type eukaryotic topoisomerases in cancer cells. This induces breaks in the DNA that ultimately lead to programmed cell death (apoptosis). This DNA-damaging effect, outside of its potential curative properties, may lead to secondary neoplasms in the patient.

Topoisomerase I is the antigen recognized by Anti Scl-70 antibodies in scleroderma.

Topological problems

There are three main types of topology:

- Supercoiling
- Knotting
- Catenation

Outside of the essential processes of replication or transcription, DNA must be kept as compact as possible, and these three states help this cause. However, when transcription or replication occurs, DNA must be free, and these states seriously hinder the processes. In addition, during replication, the newly replicated duplex of DNA and the original duplex of DNA become intertwined and must be completely separated in order to ensure genomic integrity as a cell divides. As a transcription bubble proceeds, DNA ahead of the transcription fork becomes overwound, or positively supercoiled, while DNA behind the transcription bubble becomes underwound, or negatively supercoiled. As replication occurs, DNA ahead of the replication bubble becomes positively supercoiled, while DNA behind the replication fork becomes entangled forming precatenanes. One of the most essential topological problems occurs at the very end of replication, when daughter chromosomes must be fully disentangled before mitosis occurs. Topoisomerase IIA plays an essential role in resolving these topological problems.

Classes

Topoisomerases can fix these topological problems and are separated into two types depending on the number of strands cut in one round of action:[7] Both these classes of enzyme utilize a conserved tyrosine. However these enzymes are structurally and mechanistically different.

A type I topoisomerase cuts one strand of a DNA double helix, relaxation occurs, and then the cut strand is re-ligated. Cutting one strand allows the part of the molecule on one side of the cut to rotate around the uncut strand, thereby reducing stress from too much or too little twist in the helix. Such stress is introduced when the DNA strand is "supercoiled" or uncoiled to or from higher orders of coiling. Type I topoisomerases do not require ATP for hydrolysis are subdivided into three subclasses:

- Type IA topoisomerases, which share many structural and mechanistic features with the type II topoisomerases. Examples of type IA topoisomerases include prokaryotic Topoisomerase I and III, eukaryotic Topoisomerase IIIα and Topoisomerase IIIβ and Reverse Gyrase. Like type II topoisomerases, type IA topoisomerases form a covalent intermediate with the 5' end of DNA.
- Type IB topoisomerases, which utilize a controlled rotary mechanism. Examples of Type IB topoisomerases include Eukaryotic and eukaryal viral Topoisomerase I. In the past, type IB topoisomerases were referred to as eukaryotic topoisomerase I, but IB topoisomerases are present in all three domains of life. Type IB topoisomerases form a covalent intermediate with the 3' end of DNA.
- Type IC topoisomerase (also called Topoisomerase V) has been identified. i t is structurally unique from type IA and IB topoisomerases, It shares a similar mechanism with type IB topoisomerase.
- A type II topoisomerase cuts both strands of one DNA double helix, passes another unbroken DNA helix through it, and then re-ligates the cut strands. Type II topoisomerases utilize ATP hydrolysis and are subdivided into two subclasses which possess similar structure and mechanisms:
- Type IIA topoisomerases which include eukaryotic and eukaryal viral Topoisomerase IIα and Topoisomerase IIβ, bacterial gyrase, and topoisomerase IV.
- Type IIB topoisomerases, which include Topoisomerase VI found in archaea.

Both type I and type II topoisomerases change the linking number (L) of DNA. Type IA topoisomerases change the linking number by one, type IB and type IC topoisomerases change the linking number by any integer, whereas type IIA and type IIB topoisomerases change the linking number by two.Organisation of DNA in Prokaryotes and Eukaryotes:

The DNA is twisted by what is known as supercoiling. Supercoiling means that DNA is either under-wound (less than one turn of the helix per 10 base pairs) or over-wound (more t han 1 turn per 10 base pairs) from its normal relaxed state. Some proteins are known to be

involved in the supercoiling; other proteins and enzymes such as DNA gyrase help in maintaining the supercoiled structure.

Eukaryotes, whose chromosomes each consist of a linear DNA molecule, employ a different type of packing strategy to fit their DNA inside the nucleus (Figure 2). At the most basic level, DNA is wrapped around proteins known as histones to form structures called nucleosomes. The histones are evolutionarily conserved proteins that are rich in basic amino acids and form an octamer. The DNA (which is negatively charged because of the phosphate groups) is wrapped tightly around the histone core. This nucleosome is linked to the next one with the help of a linker DNA. This is also known as the "beads on a string" structure. This is further compacted into a 30 nm fiber, which is the diameter of the structure. At the metaphase stage, the chromosomes are at their most compact, are approximately 700 nm in width, and are found in association with scaffold proteins.

In interphase, eukaryotic chromosomes have two distinct regions that can be distinguished by staining. The tightly packaged region is known as heterochromatin, and the less dense region is known as euchromatin. Heterochromatin usually contains genes that are not expressed, and is found in the regions of the centromere and telomeres. The euchromatin usually contains genes that are transcribed, with DNA packaged around nucleosomes but not further compacted.

DNA in Virus (With Diagram)

A virus is a pathogenic agent, a infections particle, consisting of a core of nucleic acid (DNA or RNA) enclosed in a protein shell. Viruses are, in general, much smaller than bacteria and have a much smaller genome. But their genetic material is of many varieties, unknown in cellular organisms. Some of these are discussed here.

Bacteriophage φX174 which attacks E. coli has a single-stranded circular DNA genome consisting of 5,386 nucleotides coding altogether 11 genes. Some of the genetic messages are overlapping which means that parts of DNA segments are common to two different genes.

Another bacteriophage M13 also capable of infecting E. coli has a similar genome as that of φX174. But the phage has a filamentous structure and does not cause lysis of infected cells, as φX174 does.

Bacteriophages MS2 and Qβ are icosahedral lytic phages of E. coli and their genomes are small single-stranded RNA molecules. The genomes are 36,000 nucleotide long and encode only 3 viral genes.

In contrast to the small phages, the T-series bacteriophages which also infect E. coli causing lysis have much larger genomes consisting of linear ds-DNA molecules. The T-even phages, like T_2, T_4, T_6 etc. have genomes of about 166,000 base pairs measuring about 55 µm in length. This DNA molecule has to be packaged into the phage-head measuring only 0.06 x 0.09 µm which necessitates an elaborate folding of the DNA molecule into a compact structure.

A special feature of the DNA of these phages is that it contains hydroxymethyl cytosine in place of cytosine. The phage T_7 is a representative of T-odd series. It has a linear ds-DNA genome containing 39,936 base pairs which code for more than 55 genes. The genes are arranged on the T_7-chromosome in an order according to their function and sequence of expression in the life cycle of the lytic phage.

The λ-phage of E. coli is a DNA-virus which has a linear ds-DNA consisting of 48,514 base pairs. After entry into the host cell, the linear molecule circularizes with the help of a segment of 12 unpaired bases at the 5′-ends of each strand. These ends of the DNA molecule of λ-phage are complementary to each other and during circularization they form base pairs. The free ends are ligated to form a covalently closed circular DNA.

These single-stranded extensions of the linear 1-DNA are called cohesive ends or cohesive site (cos)

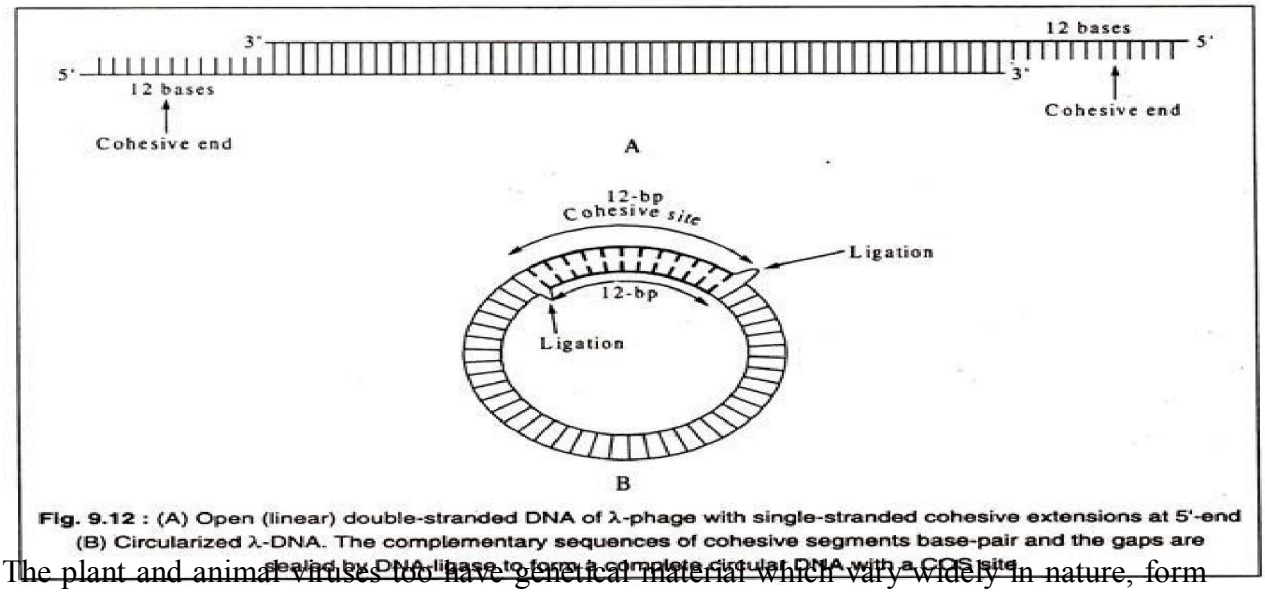


Fig. 9.12 : (A) Open (linear) double-stranded DNA of λ-phage with single-stranded cohesive extensions at 5'-end (B) Circularized λ-DNA. The complementary sequences of cohesive segments base-pair and the gaps are sealed by DNA ligase to form a complete circular DNA with a COS site

The plant and animal viruses too have genetical material which vary widely in nature, form and size. The genome may be single-stranded RNA, typical examples being influenza and tobacco mosaic viruses. Again, the ss-RNA of the virus can either be used directly as m-

RNA when it is designated (+)RNA, or the viral ss-RNA can act as a template for m-RNA synthesis.
In this case, the viral RNA is called (-)RNA. The enzyme which copies (-)RNA into another RNA (m-RNA) is RNA-dependent RNA polymerase. Some RNA viruses, like the retrovirus (HIV or Rous- Sarcoma virus) have two identical RNA molecules in virions.

During replication in host a DNA intermediate is synthesized with the help of a virion-borne enzyme, reverse transcriptase which copies the base-sequences of RNA into DNA. Thus, in this case the flow of genetic information is in the reverse direction. Normally, information coded in DNA base sequence is copied into RNA in the process of transcription.

The ds-RNA viruses have always a segmented genome. Thus, the Reoviruses have 10 segments, all of which are packaged into individual virions (isocapsidic virions). Some of the ss-RNA viruses, like influenza virus have also segmented genome; each segment is different and codes for a different polypeptide.

The alfalfa mosaic virus (a ss-RNA virus) has also four non-identical segments, but each segment is packaged into a separate virion (heterocapsidic virion). Infection of the host requites the presence of all the four-types of virions.

The ss-DNA viruses, like simian virus 40 (SV40) have a circular DNA containing 5,200 bases. The DNA of SV40 codes for 6 proteins. The infected cells contain the viral DNA associated with cellular histones to form beaded structures similar to the eukaryotic nucleosomes.

SV40 DNA replicates in the nucleus of the infected cells with the help of a host DNA polymerase to form long concatemers of viral DNA. These concatemers are opened into mono-melic viral DNA and these are then packaged into individual virions.

Large DNA viruses, like Adenovirus have double-stranded linear DNA with 35,000-36,000 base pairs. Both 3′ and 5′ ends of each strand contain 103 to 162 inverted repeated sequences which are complementary to those of the other strand. Another interesting feature of adenovirus DNA is the presence of a covalently linked 55 kilo-Dalton protein at the 5′-end of each strand. This protein can cause circularization of the ds-DNA in vitro.

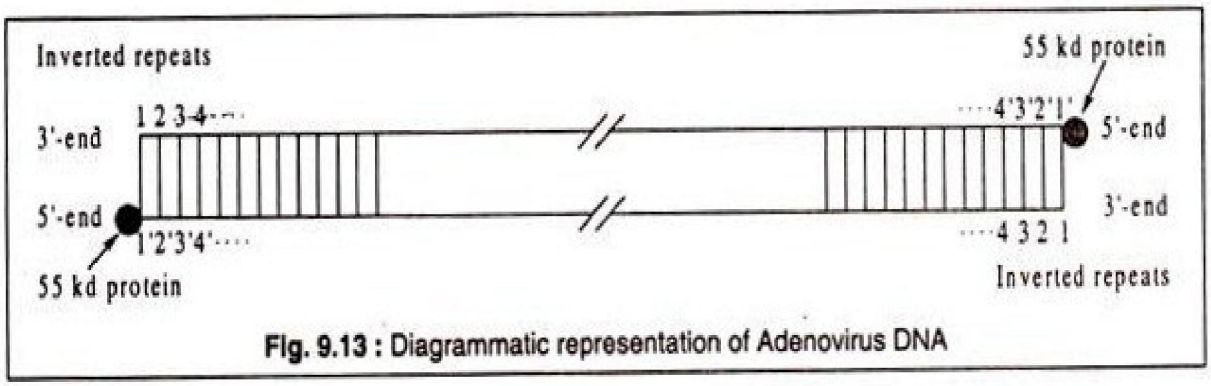


Fig. 9.13 : Diagrammatic representation of Adenovirus DNA

RNA Structure:

Structurally speaking, **ribonucleic acid (RNA)**, is quite similar to DNA. However, whereas DNA molecules are typically long and double stranded, RNA molecules are much shorter and are typically single stranded. RNA molecules perform a variety of roles in the cell but are mainly involved in the process of **protein synthesis (translation)** and its regulation.

RNA is typically single stranded and is made of **ribonucleotides** that are linked by phosphodiester bonds. A ribonucleotide in the RNA chain contains ribose (the pentose sugar), one of the four nitrogenous bases (A, U, G, and C), and a phosphate group. The subtle structural difference between the sugars gives DNA added stability, making DNA more suitable for storage of genetic information, whereas the relative instability of RNA makes it more suitable for its more short-term functions. The RNA-specific pyrimidine **uracil** forms a complementary base pair with adenine and is used instead of the thymine used in DNA. Even though RNA is single stranded, most types of RNA molecules show extensive intramolecular base pairing between complementary sequences within the RNA strand, creating a predictable three-dimensional structure essential for their function.

Organelle DNA:

Mitochondrial DNA (**mtDNA** or **mDNA** is the DNA located in mitochondria, cellular organelles within eukaryotic cells that convert chemical energy from food into a form that cells can use, adenosine triphosphate (ATP). Mitochondrial DNA is only a small portion of the DNA in a eukaryotic cell; most of the DNA can be found in the cell nucleus and, in plants and algae, also in plastids such as chloroplasts.

In humans, the 16,569 base pairs of mitochondrial DNA encode for only 37 genes. Human mitochondrial DNA was the first significant part of the human genome to be sequenced. In most species, including humans, mtDNA is inherited solely from the mother.

Since animal mtDNA evolves faster than nuclear genetic markers, it represents a mainstay of phylogenetics and evolutionary biology. It also permits an examination of the relatedness of populations, and so has become important in anthropology and biogeography.

Mitochondrial DNA is the small circular chromosome found inside mitochondria. These organelles found in cells have often been called the powerhouse of the cell.[1] The mitochondria, and thus mitochondrial DNA, are passed only from mother to offspring through the egg cell.

Human mitochondrial DNA with the 37 genes on their respective H- and L-strands.

Electron microscopy reveals mitochondrial DNA in discrete foci. Bars: 200 nm. (A) Cytoplasmic section after immunogold labelling with anti-DNA; gold particles marking mtDNA are found near the mitochondrial membrane (black dots in upper right). (B) Whole mount view of cytoplasm after extraction with CSK buffer and immunogold labelling with anti-DNA; mtDNA (marked by gold particles) resists extraction. From Iborra et al., 2004.

Structure

Circular versus linear

In most multicellular organisms, the mtDNA – or mitogenome – is organized as a circular, covalently closed, double-stranded DNA. But in many unicellular (e.g. the ciliate *Tetrahymena* or the green alga *Chlamydomonas reinhardtii*) and in rare cases also in multicellular organisms (e.g. in some species of Cnidaria) the mtDNA is found as linearly organized DNA. Most of these linear mtDNAs possess telomerase-independent telomeres (i.e. the ends of the linear DNA) with different modes of replication, which have made them interesting objects of research, as many of these unicellular organisms with linear mtDNA are known pathogens.

In mammals

For human mitochondrial DNA (and probably for that of metazoans in general), 100–10,000 separate copies of mtDNA are usually present per somatic cell (egg and sperm cells are exceptions). In mammals, each double-stranded circular mtDNA molecule consists of 15,000–17,000[base pairs. The two strands of mtDNA are differentiated by their nucleotide content, with a guanine-rich strand referred to as the heavy strand (or H-strand) and a cytosine-rich strand referred to as the light strand (or L-strand). However, confusion of labeling of these strands is widespread, and appears to originate with a identification of the majority coding strand as the heavy in one influential article in 1999. The light strand encodes 28 genes, and the heavy strand encodes 9 genes for a total of 37 genes. Of the 37 genes, 13 are for proteins (polypeptides), 22 are for transfer RNA (tRNA) and two are for the small and large subunits of ribosomal RNA (rRNA).The human mitogenome contains overlapping genes (ATP8 and ATP6 as well as ND4L and ND4: see the human mitochondrial genome map), a feature that is rare in animal genomes. The 37-gene pattern is also seen among most metazoans, although in some cases one or more of these genes is absent and the mtDNA size range is greater.

The 37 genes of the Cambridge Reference Sequence for human mitochondrial DNA and their locations

Gene	Type	Product	Positions in the mitogenome	Strand
MT-ATP8	protein coding	ATP synthase, Fo subunit 8 (complex V)	08,366–08,572 (overlap with MT-ATP6)	L
MT-ATP6	protein coding	ATP synthase, Fo subunit 6 (complex V)	08,527–09,207 (overlap with MT-ATP8)	L
MT-CO1	protein coding	Cytochrome c oxidase, subunit 1 (complex IV)	05,904–07,445	L
MT-CO2	protein coding	Cytochrome c oxidase, subunit 2 (complex IV)	07,586–08,269	L
MT-CO3	protein coding	Cytochrome c oxidase, subunit 3 (complex IV)	09,207–09,990	L
MT-CYB	protein coding	Cytochrome b (complex III)	14,747–15,887	L
MT-ND1	protein coding	NADH dehydrogenase, subunit 1 (complex I)	03,307–04,262	L
MT-	protein	NADH dehydrogenase,	04,470–05,511	L

The 37 genes of the Cambridge Reference Sequence for human mitochondrial DNA and their locations

Gene	Type	Product	Positions in the mitogenome	Strand
ND2	coding	subunit 2 (complex I)		
MT-ND3	protein coding	NADH dehydrogenase, subunit 3 (complex I)	10,059–10,404	L
MT-ND4L	protein coding	NADH dehydrogenase, subunit 4L (complex I)	10,470–10,766 (overlap with MT-ND4)	L
MT-ND4	protein coding	NADH dehydrogenase, subunit 4 (complex I)	10,760–12,137 (overlap with MT-ND4L)	L
MT-ND5	protein coding	NADH dehydrogenase, subunit 5 (complex I)	12,337–14,148	L
MT-ND6	protein coding	NADH dehydrogenase, subunit 6 (complex I)	14,149–14,673	H
MT-RNR2	protein coding	Humanin	—	—
MT-TA	transfer RNA	tRNA-Alanine (Ala or A)	05,587–05,655	H

The 37 genes of the Cambridge Reference Sequence for human mitochondrial DNA and their locations

Gene	Type	Product	Positions in the mitogenome	Strand
MT-TR	transfer RNA	tRNA-Arginine (Arg or R)	10,405–10,469	L
MT-TN	transfer RNA	tRNA-Asparagine (Asn or N)	05,657–05,729	H
MT-TD	transfer RNA	tRNA-Aspartic acid (Asp or D)	07,518–07,585	L
MT-TC	transfer RNA	tRNA-Cysteine (Cys or C)	05,761–05,826	H
MT-TE	transfer RNA	tRNA-Glutamic acid (Glu or E)	14,674–14,742	H
MT-TQ	transfer RNA	tRNA-Glutamine (Gln or Q)	04,329–04,400	H
MT-TG	transfer RNA	tRNA-Glycine (Gly or G)	09,991–10,058	L
MT-TH	transfer RNA	tRNA-Histidine (His or H)	12,138–12,206	L

The 37 genes of the Cambridge Reference Sequence for human mitochondrial DNA and their locations

Gene	Type	Product	Positions in the mitogenome	Strand
MT-TI	transfer RNA	tRNA-Isoleucine (Ile or I)	04,263–04,331	L
MT-TL1	transfer RNA	tRNA-Leucine (Leu-UUR or L)	03,230–03,304	L
MT-TL2	transfer RNA	tRNA-Leucine (Leu-CUN or L)	12,266–12,336	L
MT-TK	transfer RNA	tRNA-Lysine (Lys or K)	08,295–08,364	L
MT-TM	transfer RNA	tRNA-Methionine (Met or M)	04,402–04,469	L
MT-TF	transfer RNA	tRNA-Phenylalanine (Phe or F)	00,577–00,647	L
MT-TP	transfer RNA	tRNA-Proline (Pro or P)	15,956–16,023	H
MT-TS1	transfer RNA	tRNA-Serine (Ser-UCN or S)	07,446–07,514	H

The 37 genes of the Cambridge Reference Sequence for human mitochondrial DNA and their locations

Gene	Type	Product	Positions in the mitogenome	Strand
MT-TS2	transfer RNA	tRNA-Serine (Ser-AGY or S)	12,207–12,265	L
MT-TT	transfer RNA	tRNA-Threonine (Thr or T)	15,888–15,953	L
MT-TW	transfer RNA	tRNA-Tryptophan (Trp or W)	05,512–05,579	L
MT-TY	transfer RNA	tRNA-Tyrosine (Tyr or Y)	05,826–05,891	H
MT-TV	transfer RNA	tRNA-Valine (Val or V)	01,602–01,670	L
MT-RNR1	ribosomal RNA	Small subunit : SSU (12S)	00,648–01,601	L
MT-RNR2	ribosomal RNA	Large subunit : LSU (16S)	01,671–03,229	L

In plants

Great variation in mtDNA gene content and size exists among fungi and plants, although there appears to be a core subset of genes that are present in all eukaryotes (except for the few that have no mitochondria at all).Some plant species have enormous mitochondrial genomes,

with *Silene conica* mtDNA containing as many as 11,300,000 base pairs. Surprisingly, even those huge mtDNAs contain the same number and kinds of genes as related plants with much smaller mtDNAs. The genome of the mitochondrion of the cucumber (*Cucumis sativus*) consists of three circular chromosomes (lengths 1556, 84 and 45 kilobases), which are entirely or largely autonomous with regard to their replication.

In protists

The smallest mitochondrial genome sequenced to date is the 5,967 bp mtDNA of the parasite *Plasmodium falciparum*.

nicotiana tabacum

Chloroplast DNA Interactive gene map of chloroplast DNA from *Nicotiana tabacum*. Segments with labels on the inside reside on the B strand of DNA, segments with labels on the outside are on the A strand. Notches indicate introns.

Chloroplasts have their own DNA, often abbreviated as**cpDNA**. It is also known as the **plastome** when referring to genomes of other plastids. Its existence was first proved in 1962, and first sequenced in 1986—when two Japanese research teams sequenced the chloroplast DNA of liverwort and tobacco Since then, hundreds of chloroplast DNAs from various species have been sequenced, but they are mostly those of land plantsand green algae—glaucophytes, red algae, and other algae groups are extremely underrepresented, potentially introducing some bias in views of "typical" chloroplast DNA structure and content.

Molecular structure

The 154 kb chloroplast DNA map of a model flowering plant (*Arabidopsis thaliana*: Brassicaceae) showing genes and inverted repeats.

Chloroplast DNAs are circular, and are typically 120,000–170,000 base pairs long.They can have a contour length of around 30–60 micrometers, and have a mass of about 80–130 million daltons.

Most chloroplasts have their entire chloroplast genome combined into a single large ring, though those of dinophyte algae are a notable exception—their genome is broken up into about forty small plasmids, each 2,000–10,000 base pairs long.[6] Each minicircle contains one to three genes,but blank plasmids, with no coding DNA, have also been found.

Inverted repeats

Many chloroplast DNAs contain two *inverted repeats*, which separate a long single copy section (LSC) from a short single copy section (SSC).

The inverted repeats vary wildly in length, ranging from 4,000 to 25,000 base pairs long each. Inverted repeats in plants tend to be at the upper end of this range, each being 20,000–25,000 base pairs long.The inverted repeat regions usually contain three ribosomal RNA and two tRNA genes, but they can be expanded or reduced to contain as few as four or as many as over 150 genes. While a given pair of inverted repeats are rarely completely identical, they are always very similar to each other, apparently resulting from concerted evolution.

The inverted repeat regions are highly conserved among land plants, and accumulate few

mutations.Similar inverted repeats exist in the genomes of cyanobacteria and the other two chloroplast lineages (glaucophyta and rhodophyceæ), suggesting that they predate the

chloroplast,though some chloroplast DNAs like those of peas and a few red algae have since lost the inverted repeats.Others, like the red alga *Porphyra* flipped one of its inverted repeats (making them direct repeats).It is possible that the inverted repeats help stabilize the rest of the chloroplast genome, as chloroplast DNAs which have lost some of the inverted repeat segments tend to get rearranged more.

Linear structure

Chloroplast DNA has long been thought to have a circular structure, but some evidence suggests that chloroplast DNA more commonly takes a linear shape. Over 95% of the chloroplast DNA in corn chloroplasts has been observed to be in branched linear form rather than individual circles.

Nucleoids

Each chloroplast contains around 100 copies of its DNA in young leaves, declining to 15–20 copies in older leaves. They are usually packed into nucleoids which can contain several identical chloroplast DNA rings. Many nucleoids can be found in each chloroplast.

Though chloroplast DNA is not associated with true histones,[15] in red algae, a histone-like chloroplast protein (HC) coded by the chloroplast DNA that tightly packs each chloroplast DNA ring into a nucleoid has been found.

In primitive red algae, the chloroplast DNA nucleoids are clustered in the center of a chloroplast, while in green plants and green algae, the nucleoids are dispersed throughout the stroma.

DNA replication

Leading model of cpDNA replication

Chloroplast DNA replication via multiple D loop mechanisms. Adapted from Krishnan NM, Rao BJ's paper "A comparative approach to elucidate chloroplast genome replication."

The mechanism for chloroplast DNA (cpDNA) replication has not been conclusively determined, but two main models have been proposed. Scientists have attempted to observe chloroplast replication via electron microscopy since the 1970s.[17][18] The results of the microscopy experiments led to the idea that chloroplast DNA replicates using a double displacement loop (D-loop). As the D-loop moves through the circular DNA, it adopts a theta intermediary form, also known as a Cairns replication intermediate, and completes replication with a rolling circle mechanism.Transcription starts at specific points of origin. Multiple replication forks open up, allowing replication machinery to transcribe the DNA. As replication continues, the forks grow and eventually converge. The new cpDNA structures separate, creating daughter cpDNA chromosomes.

In addition to the early microscopy experiments, this model is also supported by the amounts of deamination seen in cpDNA. Deamination occurs when an amino group is lost and is a mutation that often results in base changes. When adenine is deaminated, it becomes hypoxanthine. Hypoxanthine can bind to cytosine, and when the XC base pair is replicated, it becomes a GC (thus, an A → G base change).

Over time, base changes in the DNA sequence can arise from deamination mutations. When adenine is deaminated, it becomes hypoxanthine, which can pair with cytosine. During replication, the cytosine will pair with guanine, causing an A → G base change.

In cpDNA, there are several A → G deamination gradients. DNA becomes susceptible to deamination events when it is single stranded. When replication forks form, the strand not being copied is single stranded, and thus at risk for A → G deamination. Therefore, gradients in deamination indicate that replication forks were most likely present and the direction that they initially opened (the highest gradient is most likely nearest the start site because it was single stranded for the longest amount of time).This mechanism is still the leading theory today; however, a second theory suggests that most cpDNA is actually linear and replicates through homologous recombination. It further contends that only a minority of the genetic material is kept in circular chromosomes while the rest is in branched, linear, or other complex structures.

Alternative model of replication

One of the main competing models for cpDNA asserts that most cpDNA is linear and participates in homologous recombinationand replication structures similar to bacteriophage T4.It has been established that some plants have linear cpDNA, such as maize, and that more still contain complex structures that scientists do not yet understand however, the predominant view today is that most cpDNA is circular. When the original experiments on cpDNA were performed, scientists did notice linear structures; however, they attributed these linear forms to broken circles.If the branched and complex structures seen in cpDNA experiments are real and not artifacts of concatenated circular DNA or broken circles, then a D-loop mechanism of replication is insufficient to explain how those structures would replicate. At the same time, homologous recombination does not explain the multiple A → G gradients seen in plastomes.This shortcoming is one of the biggest for the linear structure theory.

POSSIBLE QUESTIONS PART B

1)What are the types of DNA?
2)What are the functions of DNA?
3)Write notes on mitochondrial DNA?
4)Write notes on Chloroplast DNA?
5)Write about the organisation of DNA in Eucaryotes?

PART C

1)Write in detail about the Structure of DNA?
2)Write in detail about the Types of DNA?
3)Explain the salient features of DNA?
4)Explain the functions of DNA?
5)Give short notes on mitochondrial and chloroplast DNA?
6)Give short note on Organisation of DNA in Prokaryotes, Eucayotes and Viruses?

KARPAGAM ACADEMY OF HIGHER EDUCATION

CLASS: III B.Sc MB COURSE NAME: MOLECULAR BIOLOGY

COURSE CODE: 17MBU603B UNIT: I BATCH-2017-2020

Sl. No	Question	Option A	Option B	Option C	Option D	Correct Ans
1	Enzyme that adds methyl group to the newly formed DNA	Gyrase	Topoisomerase	Helicase	Methylase	Methylase
2	Heat Killed S cells + Live R cells produced	Death in mice + S cells	Live mice + S cells	Death in mice + R cells	Live mice + R cells	Death in mice + S cells
3	Initiation of replication occurs	Bidirectionally	Cross sectionally	Unidirectionally	Parallely	Bidirectionally
4	The replication origin of *E.coli* is approximately	245 bp	425 bp	254 bp	524 bp	245 bp
5	Transformation in *Pneumococci* was discovered by	Friedrick Griffith	Erwin Chargaff	Hershey & Chase	Watson & Crick	Friedrick Griffith
6	Bacteria contains	Single circular DNA	Single linear DNA	Double Linear DNA	Double circular DNA	Single circular DNA
7	Who proposed the molecular struccture of DNA	Hershey & Chase	Erwin Chargaff	Jim Latham	Watson & Crick	Watson & Crick
8	The most common form of DNA is	B-DNA	Z-DNA	Y-DNA	SS-DNA	B-DNA
9	Sequencing and molecular characterization of genome	Genetics	Molecular biology	Proteomics	Genomics	Genomics
10	Adenine always pair with	Guanine	Cytosine	Thymine	Uracil	Thymine
11	Initiation of replication is carried out by	DnaA	DnaC	DnaB	DnaE	DnaA
12	Which of the following is not associated with DNA replication?	Polymerase	Promoter	Primer	RepA protein	Promoter

13	Father of genetics	Mendel	Morgan	Watson	McLeod	Mendel
14	If a free phosphate is found at the 5' end of a DNA strand, what is found at the other end of the same strand?	A hydroxyl group on the 5' carbon of a deoxyribose sugar	A phosphate group on the 3' carbon of a deoxyribose sugar.	A base attached to the 3' carbon of a deoxyribose sugar	A hydroxyl group on the 3' carbon of a deoxyribose sugar	A hydroxyl group on the 3' carbon of a deoxyribose sugar
15	The replication origins of higher eukaryotes are made up of	Different AT-rich regions	Similar AT-rich regions	Different GC-rich regions	Similar GC-rich regions	Similar AT-rich regions
16	Transfer of genes from one generation to the next is	Inheritance	Carrying over	Subheritance	Gene transport	Inheritance
17	Basically, flow of genetic material is accompolished by	Replication	Transformation	Transduction	Conjugation	Replication
18	Purines are	A, T	G, C	C, T	A, G	A, G
19	The DNA of E.coli is _____ times longer than the cell	1	10	1000	100	1000
20	SSB protein helps in	Degradation of protein	Keep the two strands separated after unwinding	Elongation of DNA	Uncoiling of RNA	Keep the two strands separated after unwinding
21	The experiments of Avery, McLeod and McCarty was based on	Protein coupling	Enzymatic reactions	Synthetic reaction	DNA binding	Enzymatic reactions
22	The contribution of Rosalind Franklin towards structure of DNA was	X-ray crystallography	Electron microscopy	NMRspectroscopy	Gas chromatography	X-ray crystallography

23	Proof reading activity of DNA polymerase is in the direction	5' to 3'	3' to 5'	Parallel	Centre	3' to 5'
24	Length of primer during replication is	2-10 nucleotides	10-20 nucleotides	5-15 nucleotides	10-25 nucleotides	2-10 nucleotides
25	Distance between the two base pairs is	3.4Å	34Å	10Å	20Å	3.4Å
26	Formation of pre-replicating complex is seen in replication mechanism of	Prokaryotes	Plants	Virus	Eukaryotes	Eukaryotes
27	Triple bonding is seen in	G-T	G-C	A-T	A-C	G-C
28	Repair and insertion of DNA is carried out by	Endonucleases	Ribozyme	Primase	Exonucleases	Endonucleases
29	Nucleoside is	Base + Sugar	Sugar + Phosphate	Base + Phosphate	A+T & G+C	Base + Sugar
30	Left handed DNA	B-DNA	C-DNA	Y-DNA	Z-DNA	Z-DNA
31	Bonding between two bases	Hydrogen bond	Hydrophobic bond	Nitrogen bond	Van Der waals	Hydrogen bond
32	DNA absorbs UV light at _____ wavelength	220 nm	240 nm	260 nm	280 nm	260 nm
33	The Pioneer work on nucleic acid discovery was carried out by	Friedrick Miescher	Watson & Crick	Griffith	Milstein	Friedrick Miescher
34	The enzyme that copies RNA from DNA template	Dnase	Rnase	DNA polymerase	RNA polymerase	RNA polymerase
35	The size of a major groove is	34Å	3.4Å	43Å	20Å	34Å
36	Chromosomal theory of inheritance was formulated by	Mendel	Miescher	Metchinikoff	Morgan	Morgan

37	Eukaryotic DNA replication is a conserved mechanism that restricts DNA replication to	Never	Only once per cell cycle	Only twice per cell cycle	Only thrice per cell cycle	Only once per cell cycle
38	Semiconservative mechanism of DNA replication was demonstrated by	Meselson & Stahl	Beedle & Tatum	Hershy & Chase	Avery & McLeod	Meselson & Stahl
39	The most widely studied origin recognition complex of eukaryotes is that of	*Bacillus*	*Staphylococcus*	*Escherichia coli*	*Saccharomyces cerevisiae*	*Saccharomyces cerevisiae*
40	Number of base pairs per helical turn of B form DNA	13	12	11	10	10
41	RNA primer is removed by	DNA pol	RNA pol	Terminase	Caspase	DNA pol
42	Longest DNA is seen in	Human	Lung fish	Yeast	Bacteria	Lung fish
43	Joining of DNA fragments	DNA ligase	Gyrase	RNA polymerase	DNA polymerase	DNA ligase
44	Synthesis of DNA always moves from	3' to 5'	5' to 3'	Ffrom the centre	Anywhere	5' to 3'
45	Dihybrid ratio	3 : 3 : 9 : 1	9 : 3 : 3 : 1	9 : 3 : 1 : 3	1 : 3 : 3 : 9	9 : 3 : 3 : 1
46	Experiments of Hershey and Chase was based on	Virus	Bacteriophage	Bacteria	Fungi	Bacteriophage
47	In eukaryotes, the vast majority of DNA synthesis occurs during _____ of the cell cycle	G phase	H phase	R phase	S phase	S phase
48	Key enzyme in rolling circle replication	DNA Polymerase-IV	DNA Polymerase-III	DNA Polymerase-II	DNA Polymerase-I	DNA Polymerase-III

49	Mendels pioneer work was with	Monkey	Human	Garden pea	Mice	Garden pea
50	Okazaki fragments are	RNA strands	Enzymes	Leading strands	Lagging strands	Lagging strands
51	The ability to remove incorrectly matched nucleotides or Proof-reading	RNA polymerase	DNA ligase	DNA polymerase	DNA helicase	DNA polymerase
52	The enzyme that unwinds DNA	Polymerase	Ligase	Gyrase	Helicase	Helicase
53	_____ million base pairs of nucleotides are seen in *E.coli*	64 million base pairs	46 million base pairs	4.6 million base pairs	6.4 million base pairs	4.6 million base pairs
54	Chargaff's rule	No complementarity	Partial complementarity	No such rule	Complementarity of one strand with the other	Complementarity of one strand with the other
55	Eukaryotic DNA damage or replication errors are corrected during	G_1 phase	S phase	G_2 phase	R phase	G_2 phase
56	Semiconservative DNA replication model	Daughter molecule contains both from parent	Daughter molecule entirely new	Daughter molecule contains one from parent and one newly synthesized	Some sections from parent and some newly synthesized	Daughter molecule contains one from parent and one newly synthesized
57	The negative charge of DNA is due to	Deoxyribose Sugar	Phosphate bond	Hydrogen bond	Nitrogenous base	Phosphate bond
58	Thymine in DNA is replaced by _____ in RNA	Adenine	Cytosine	Guanine	Uracil	Uracil
59	Polymerase	Primase	RNase	DNase	DNase	Polymerase

60	Longest DNA is seen in	Human	Lung fish	Yeast	Bacteria	Lung fish

Unit II

Bidirectional and unidirectional replication, semi- conservative, semi- discontinuous replication Mechanism of DNA replication: Enzymes and proteins involved in DNA replication –DNA polymerases, DNA ligase, primase, telomerase – for replication of linear ends. Various models of DNA replication including rolling circle, D-loop (mitochondrial), (theta) mode of replication and other accessory protein, Mismatch and excision repair.

Unidirectional and Bidirectional Replication:

Replication starts at a point on the chromosome where the two parental strands begin to separate; this point is called the origin. Addition of complementary nucleotides to form two new strands takes place along both parent strand templates starting from that point (Fig. 21-4).

In unidirectional replication, growth proceeds along both strands in the same direction leading from the origin. Along one of the parental template strands, synthesis of the new complementary strand takes place by the continuous addition of nucleotides to the available 3′ end of the forming strand. The growing strand is called the leading strand or continuous strand. The 5′ end of this strand is located at the origin and its 3′ end at the moving replication fork (i.e., the progressing point of separation of the parental strands).

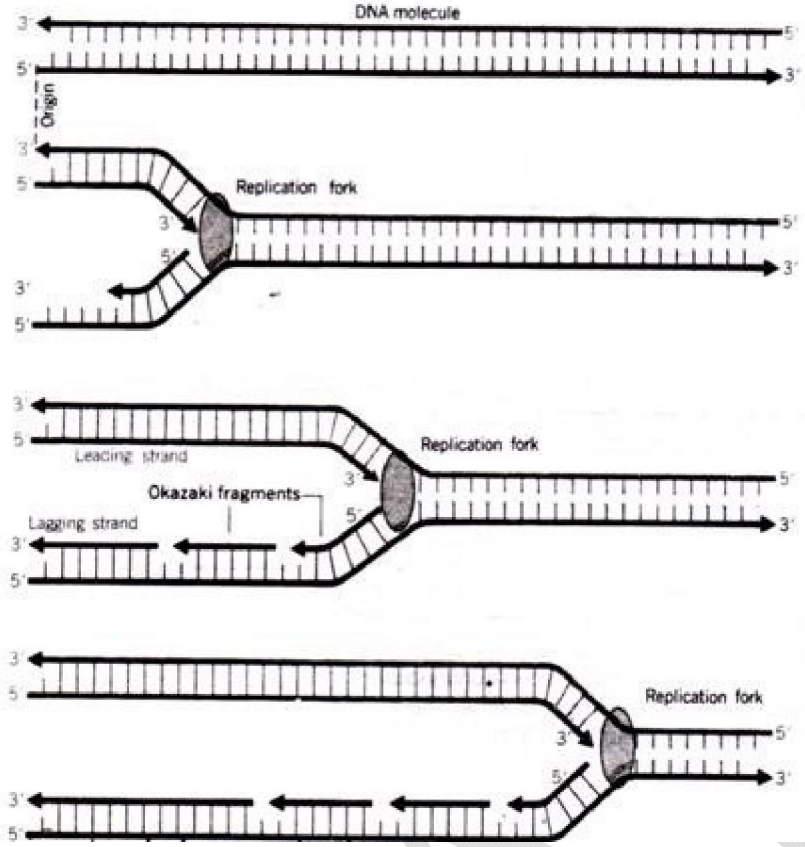


FIGURE 21–4 *Unidirectional* replication of DNA. Behind the *replication fork*, one new DNA strand grows continuously and the other grows as a series of fragments that are eventually linked together. The overall growth of both new strands is in the direction of the moving fork.

The other polynucleotide strand being formed is called the lagging strand or discontinuous strand. The elongation of this strand takes place by a somewhat modified mechanism. In contrast to the leading strand, the lagging strand has its 3′ position at the origin and its 5′ position at the replication fork. If nucleotides were sequentially added to the end of the lagging strand at the replication fork, then this strand's growth would proceed in a 3'→5′ direction.

This does not occur. Instead, growth takes place by the synthesis of a number of short polynucleotide chains be- tween the replication fork and the origin. Each short chain is laid down in the direction 5′ to 3′ and these are later linked together and to the 5′ end of the lagging strand.

As a result, the overall direction of growth of the lagging strand is the same as that of the leading strand. The unusual growth pattern that characterizes the synthesis of the lagging

strand explains why it is also referred to as the "discontinuous" strand.

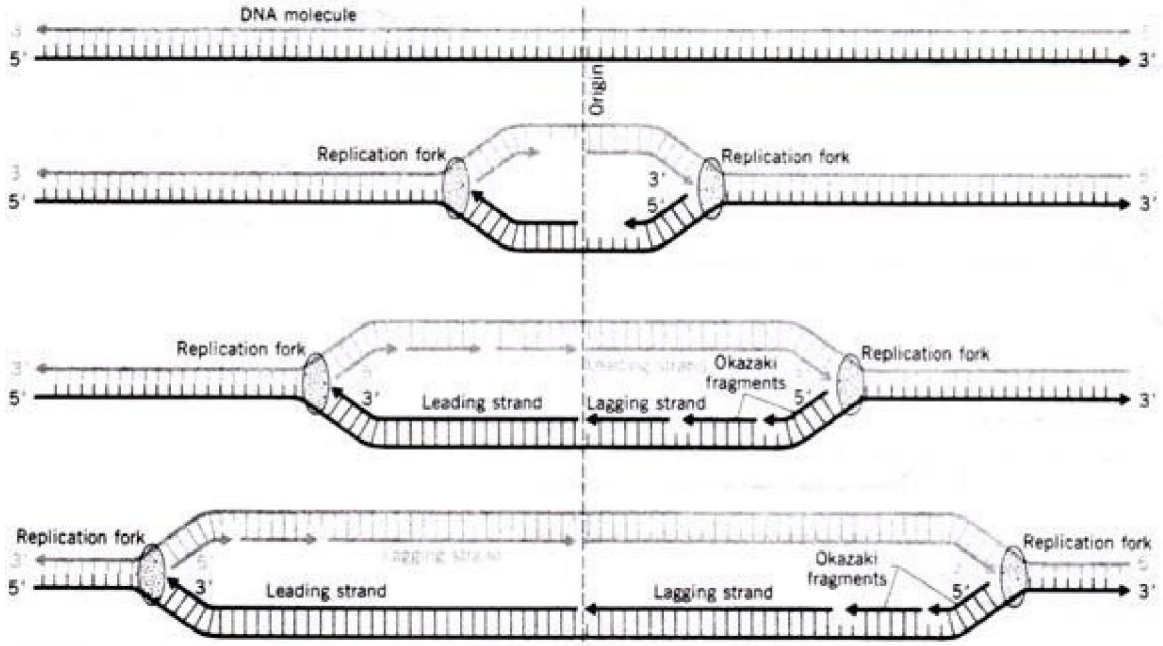


FIGURE 21–5 *Bidirectional* replication of DNA. Two replication forks move away from the origin in opposite directions. Behind each fork, one new DNA strand grows continuously and the other grows as a series of fragments that are eventually linked together. The overall growth of both sets of new strands is in the direction of the moving forks.

In bidirectional replication (Fig. 21-5), two replication forks are formed at the origin and these move away from the origin in both directions as the parental double helix is separated. The synthesis of the complementary strands also occurs in both directions. Behind each fork there is a set of leading and lagging strands. As in the case of unidirectional replication, elongation of the two leading strands is continuous, whereas elongation of the two lagging strands is discontinuous.

It is to be noted that regardless of whether replication is unidirectional or bidirectional, the addition of nucleotides always occurs in the direction from 5′ to 3′, as new nucleotides are added to available 3′ ends of either the continuous strand or the discontinuous strand. Discontinuous synthesis of lagging strands was first demonstrated by R. Okazaki. Okazaki incubated E. coli cells in a medium containing ^{3}H-thymidine for very short periods of time (a pulse of only 15 seconds) and then examined the distribution of the radioisotope in newly synthesized DNA.

The radioisotope was found in a number of polynucleotides (1000-2000 nucleotides long), now referred to as Okazaki fragments (Figs. 21-4 and 21-5). When pulsed cells

were transferred to unlabeled medium for varying lengths of time prior to analysis, the radioactive label was recovered in much longer stretches of DNA. This is because the Okazaki fragments produced during the short tritium pulse had been linked together and connected to the 5′ end of the lagging strand.

In eukaryotic cells, Okazaki fragments are usually smaller (about 100-200 nucleotides long). Bidirectional replication of DNA is the mechanism employed in all eukaryotic and most prokaryotic cells. Unidirectional replication is rare and appears to occur in only a limited number of prokaryotes.

Replication as a "Semiconservative" Process:

Although semiconservative replication of DNA was predicted by the original Watson-Crick model, it was not verified until the classic studies of M. S. Meselson and F. W. Stahl. At the time of their experiments, two other modes of replication were deemed equally feasible (Fig. 21-1):

(1) Conservative replication, in which both strands of the parent double helix would be conserved and the new DNA molecule would consist of two newly synthesized strands; and

(2) Dispersive replication, in which replication would involve fragmentation of the parent

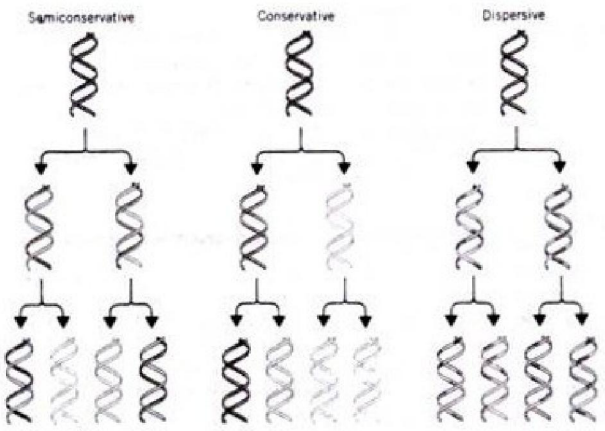


FIGURE 21–1 Three possible modes of DNA replication: semiconservative (left), conservative (center), and dispersive (right) (see text for explanation). (From E. J. Gardner and D. P. Snustad, *Principles of Genetics* (7th ed.). John Wiley & Sons, 1984, p. 97.)

double helix and the interspersing of pieces of the parent strands with newly synthesized pieces, thereby forming the two new double helices.

Meselson and Stahl verified the semiconservative nature of DNA replication in a series of elegant experiments using isotopically labeled DNA and a form of isopycnic density gradient centrifugation (see Chapter 12). They cultured Escherichia coli cells in a medium in which the nitrogen was ^{15}N (a "heavy" isotope of nitrogen, but not a

radioisotope) instead of the commonly occurring and lighter ^{14}N.

In time, the purines and pyrimidines of DNA in new cells contained ^{15}N (where ^{14}N normally occurs) and thus the DNA molecules were denser. DNA in which the nitrogen atoms are ^{15}N can be distinguished from DNA containing ^{14}N because during isopycnic centrifugation, the two different DNAs band at different density positions in the centrifuge tube (Fig. 21-2).

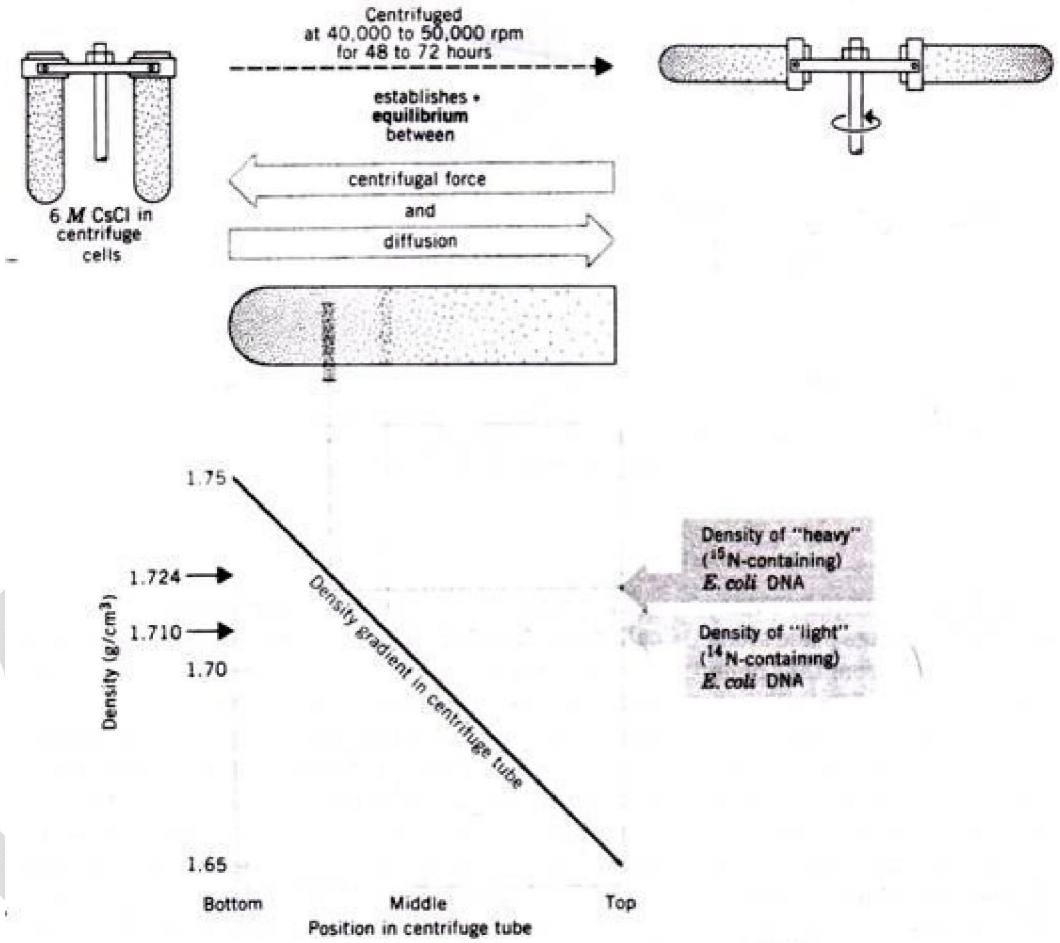


FIGURE 21–2 Separation of ^{15}N-containing DNA from ^{14}N-containing DNA by equilibrium isopycnic centrifugation (see text for details). (From E. J. Gardner and D. P. Snustad, *Principles of Genetics* (7th ed.), John Wiley & Sons, 1984, p. 98.)

Meselson and Stahl centrifuged DNA isolated from the cells for 2-3 days at very high rotational speeds in centrifuge tubes initially containing a uniform solution of CsCl. During centrifugation, density gradients were automatically formed in the tubes as a result of the equilibrium that was established between the sedimentation of CsCl toward the bottom of the tube and diffusion of the salt toward the top of the tube. This form of

centrifugation, called equilibrium isopycnic centrifugation,

Depending on its content of ^{15}N and ^{14}N, the DNA bands at a specific position in the density gradient. Because the DNA synthesized by cells grown in ^{15}N would be denser than ^{14}N-containing DNA it would band further down the tube (Fig. 21-2).Cells grown for some time in the presence of ^{15}N- medium were washed free of the medium and transferredto ^{14}N-containing medium and allowed to continue to grow for specific lengths of time (i.e., for various numbers of generation times). DNA isolated from cells grown for one generation of time in the ^{14}N medium had a density intermediate to that of the DNA from cells grown only in ^{15}N-containing medium (identified as generation 0 in Fig. 21-3) and that of DNA from cells grown only in ^{14}N-containing medium (the controls of Fig. 21-3).

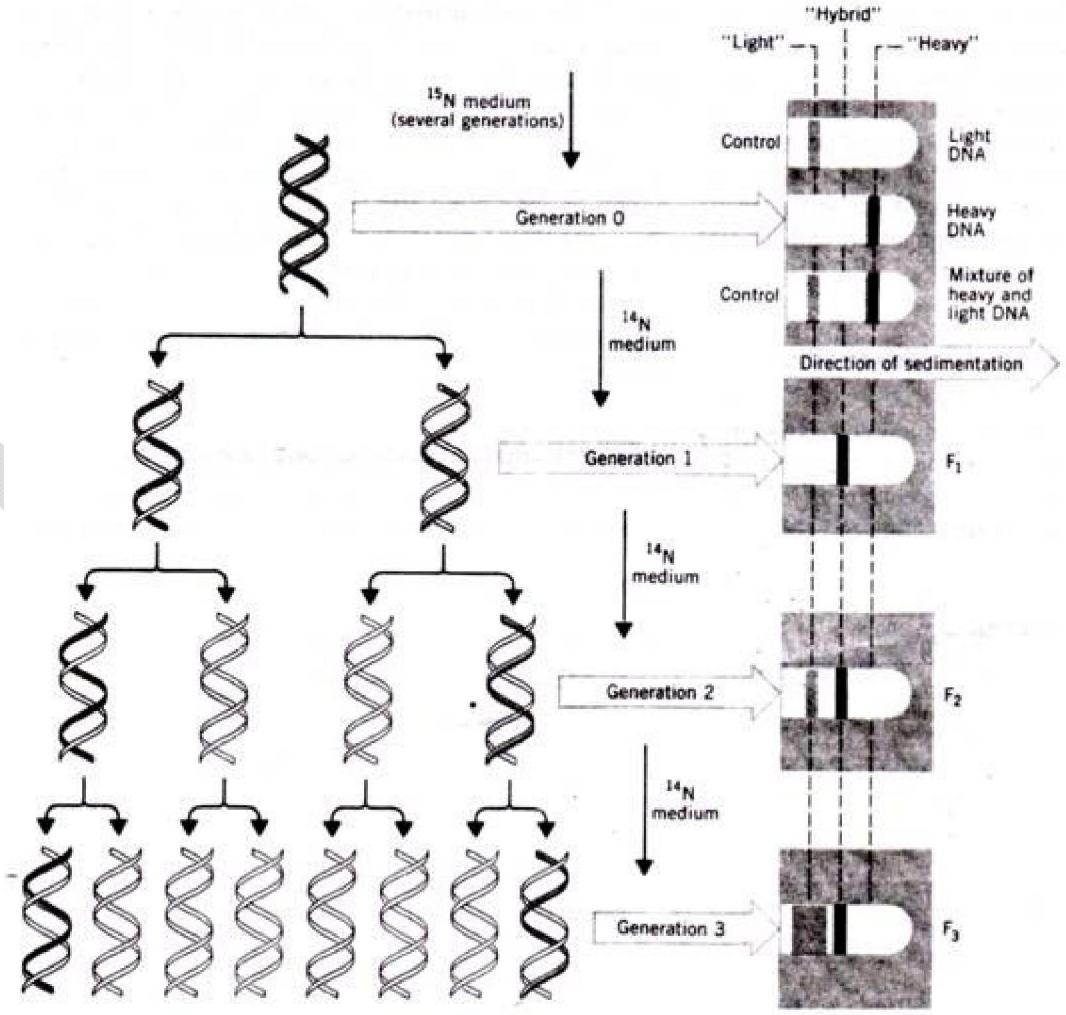


FIGURE 21–3 Results of the Meselson–Stahl experiments (right) and their interpretation (left) (see text for explanation). (From E. J. Gardner and D. P. Snustad, *Principles of Genetics* (7th ed.), John Wiley & Sons, 1984, p. 99.)

Such a result immediately ruled out the possibility that DNA replication was conservative, as conservative replication would have yielded two DNA bands in the density gradient for generation 1 (i.e., F,) cells. The single band of intermediate density (identified as "hybrid" DNA in Fig. 21-3) consisted of DNA molecules in which one strand contained ^{15}N and the other contained ^{14}N.

When the incubation in ^{14}N medium was carried out for two generations of time (i.e., generation 2), two DNA bands were formed—one at the same density position as the DNA from cells grown exclusively in ^{14}N medium (i.e., "light controls") and one of intermediate density. Subsequent generations produced greater numbers of DNA molecules that banded at the "light" (^{14}N- containing DNA) position in the density gradient. These results are consistent only with the model of semiconservative replication.

Dispersive replication would have produced a single band for each generation and the band would have been found at successively lighter density positions in the gradient. Studies using other prokaryotes as well as eukaryotes indicate that semiconservative replication of DNA is probably the universal mechanism.

Semi discontinuous replication

Semi-discontinuous nature of replication is attributed to the fact that one strand is replicated continuously whereas the other strand in discontinuous fashion. The continuous synthesis here means that the replication only needs one primer, and the replication occurs till the termination site. This strand is called Leading strand. On the counterpart, the discontinuous strand, also called Lagging strand, will have the formation of Okazaki fragments.

The Trombone model can well explain the semi-discontinuous replication. It gives an explanation how the Polymerase III holoenzyme manage to replicate in the opposite direction without falling off from the strand.

The DNA pol III holoenzyme has 3 Polymerase III core enzymes. Two DNA pol III core enzyme is dedicated to Lagging strand (as the replication is slow) and one is devoted to Leading strand.The β-clamp or Sliding DNA clamp binds to the Polymerase III core enzyme to increase the processivity of the DNA polymerase enzyme. It is also believed that the β-clamp also helps in the formation of a loop on the lagging strand. When a loop is formed, the direction of the DNA polymerase core enzyme on the lagging strands will

be same to that of the leading strand polymerase. Thus the strain on the Tau-proteins will be decreased.To simplify, Tau-proteins are the proteins which attach the polymerase core enzyme to the γ- complex or Sliding clamp loaders. There are three Tau-proteins for three DNA polymerase core enzymes to bind the γ-complex.Now, imagine that the leading DNA polymerase is adding nucleotides and moving forward and on the counterpart, the lagging DNA polymerases are adding nucleotides and heading inthe opposite direction. It would increase the strain on the Tau-proteins, and the DNA polymerase core enzyme may fall off the replicating strand. Thus the β-clamp or Sliding DNA clamp facilitates the formation of a loop in such a way that the direction of both polymerases i.e. leading and lagging is almost in a similar direction.

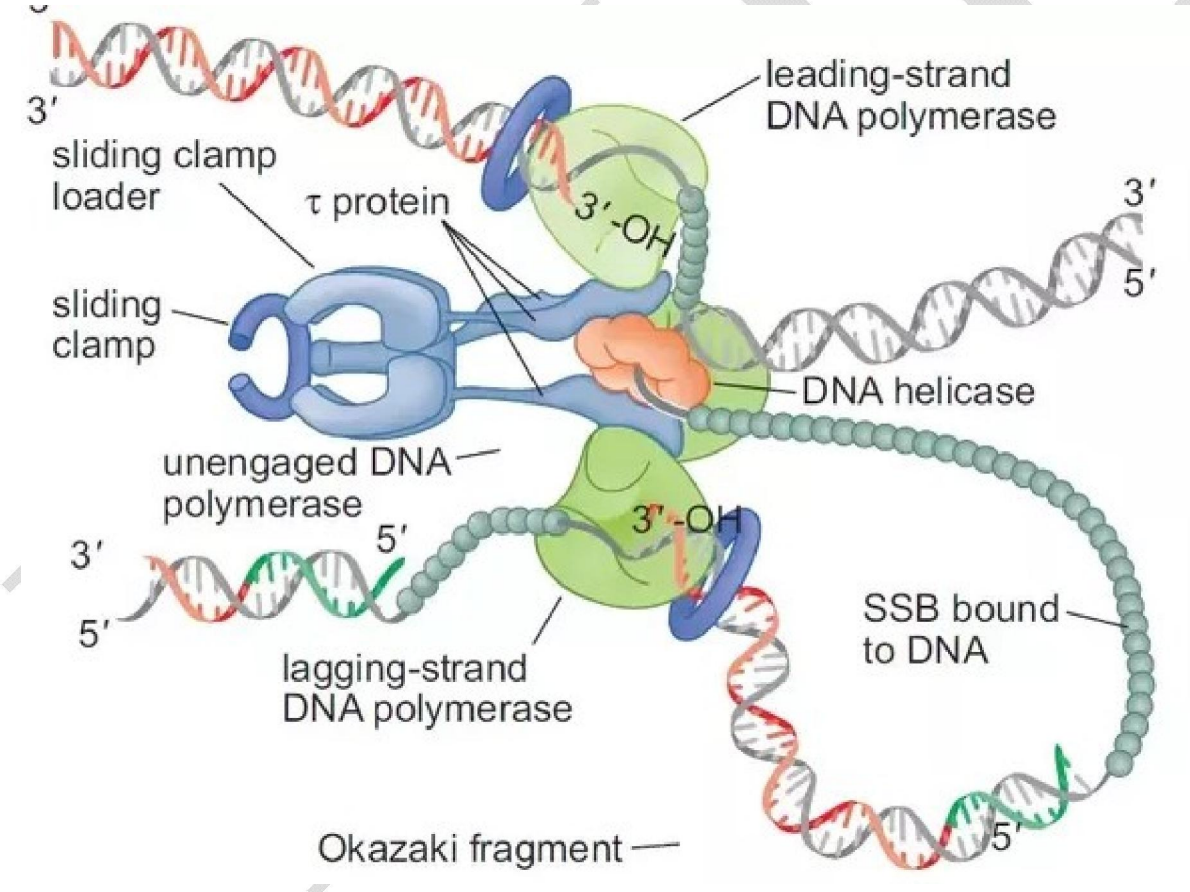


Image Source: Molecular Biology by Watson et al, 7th edition; Chapter 9; page 285.

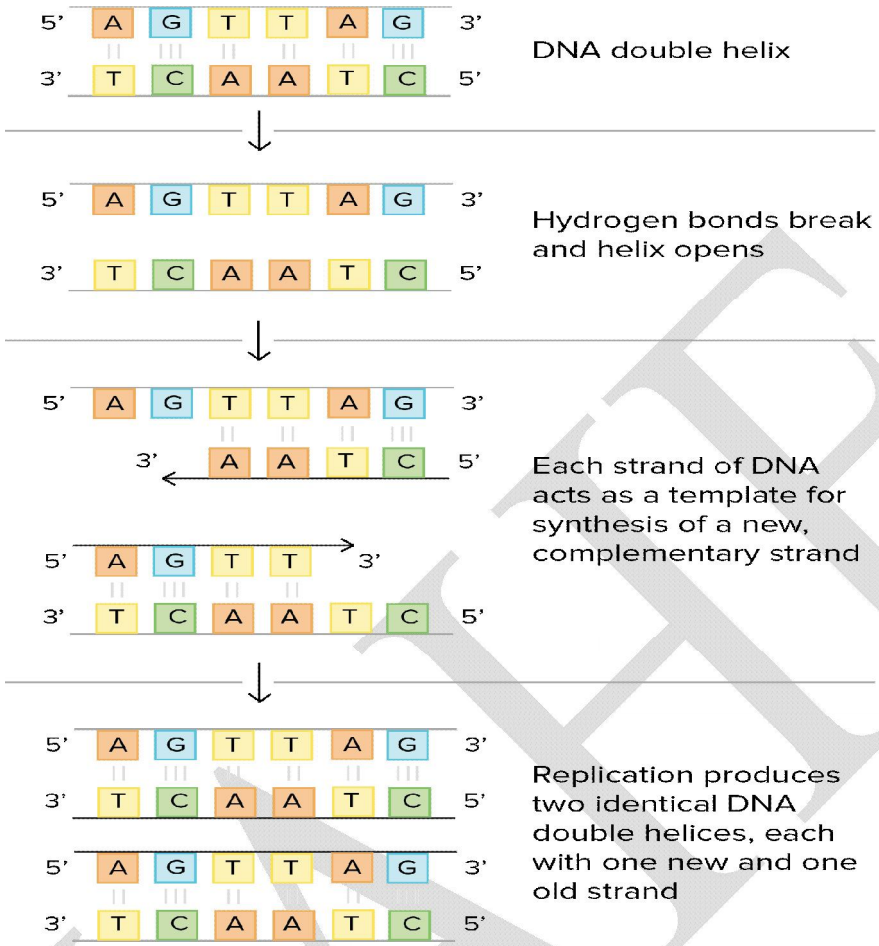
In the above image, you can observe the three DNA polymerase core enzymes shown in Green. The lagging strand will have two dedicated DNA polymerase core enzyme.

The DNA polymerase core enzymes work simultaneously. In the lagging strand, the replication is discontinuous, and hence the Polymerase core enzyme would fall at the end of Okazaki fragment and then loaded to a new primer-template junction. This process takes a lot of time and hence the speed of the whole DNA polymerase holoenzyme (i.e. including all three core enzymes, sliding clamp and its loader) is compromised.Thus, there are two polymerase enzymes dedicated to lagging strand. While on of the DNA polymerase core enzyme is synthesising the Okazaki fragment, the other DNA polymerase core enzyme prepares itself and may start the replication from a new primer template junction. Thus when the first DNA Polymerase core enzymes reach the end of Okazaki fragments, the other core enzyme is already replicating the strand. And thus, the detached polymerase core enzyme will get the time to identify and load itself to the new primer- template junction As the multiple primers are formed, the DNA synthesised on the lagging strand will be discontinuous, and the fragments are called Okazaki fragments. Later, RNase H enzyme would degrade the RNA primers of the lagging strand and DNA polymerase I enzyme would fill the gaps by adding dNTPs. The nicks will be filled by the formation of phosphodiester bond by DNA ligase enzyme.

Introduction DNA replication, or the copying of a cell's DNA, is no simple task! There are about 33 \text{billion}billion base pairs of DNA in your genome, all of which must be accurately copied when any one of your trillions of cells divides. The basic mechanisms of DNA replication are similar across organisms. In this article, we'll focus on DNA replication as it takes place in the bacterium *E. coli*, but the mechanisms of replication are similar in humans and other eukaryotes Let's take a look at the proteins and enzymes that carry out replication, seeing how they work together to ensure accurate and complete replication of DNA.

The basic idea

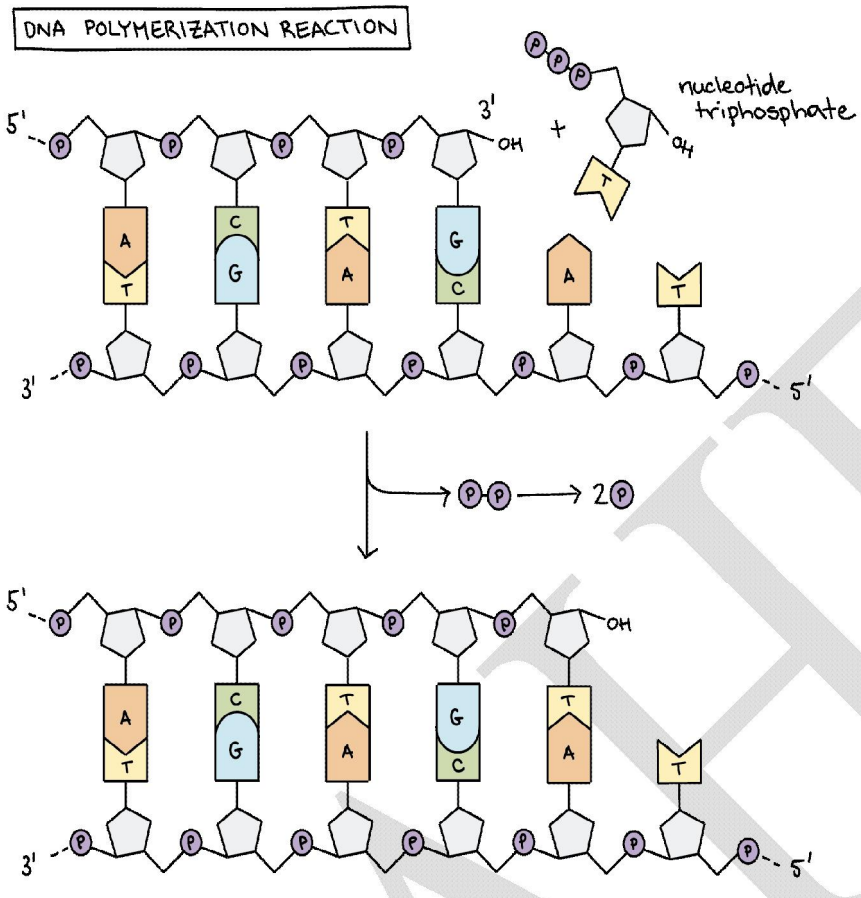
DNA replication is **semiconservative**, meaning that each strand in the DNA double helix acts as a template for the synthesis of a new, complementary strand. This process takes us from one starting molecule to two "daughter" molecules, with each newly formed double helix containing one new and one old strand



Here are some key features of DNA polymerases:

- They always need a template
- They can only add nucleotides to the 3' end of a DNA strand
- They can't start making a DNA chain from scratch, but require a pre-existing chain or short stretch of nucleotides called a **primer**
- They **proofread**, or check their work, removing the vast majority of "wrong" nucleotides that are accidentally added to the chain

The addition of nucleotides requires energy. This energy comes from the nucleotides themselves, which have three phosphates attached to them (much like the energy-carrying molecule ATP). When the bond between phosphates is broken, the energy released is used to form a bond between the incoming nucleotide and the growing chain.



[

DNA polymerization reaction

The diagram shows a template DNA strand paired up with a new DNA strand that is currently being synthesized. The bases of the new strand and the template form complementary pairs held together by hydrogen bonds. The two strands are antiparallel. Their sequences are:

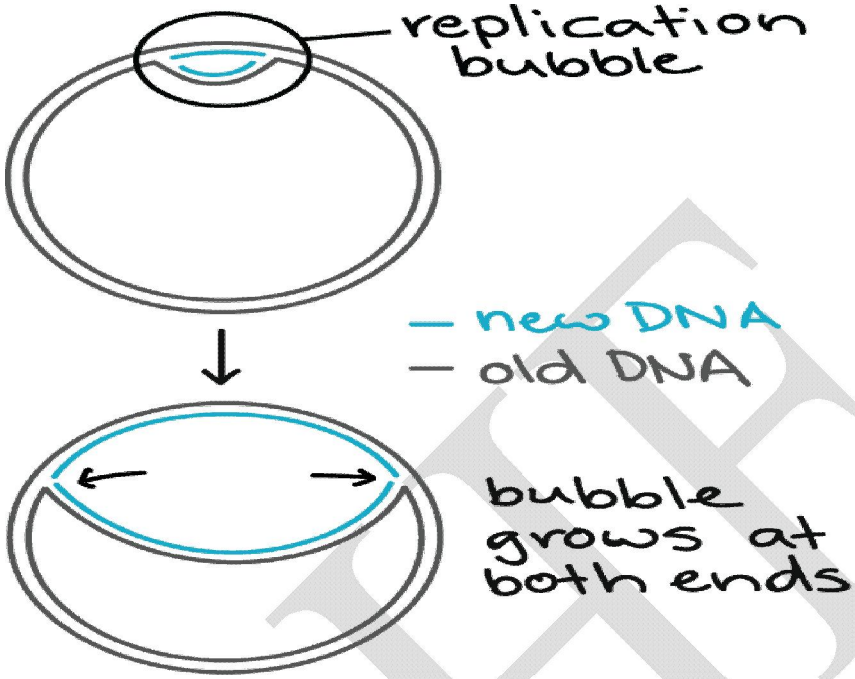
New strand: 5' ACTG... 3' Template strand: 3' TGACAT 5'

At the end of the new strand, a 3' hydroxyl group is exposed on the final nucleotide of the strand. This hydroxyl group will undergo a chemical reaction with a phosphate group in an incoming nucleotide, resulting in the formation of a new bond (and the addition of the nucleotide to the end of the chain). Specifically, the next nucleotide exposed on the template is an A. The nucleotide complementary to an A is a T, so when a T pairs with the A on the

template strand, it will undergo a reaction with the 3' hydroxyl at the end of the chain and be added on. The nucleotide consists of a sugar with a chain of three phosphates, a base, and a hydroxyl attached. The 3' hydroxyl exposed at the end of the growing strand will form a bond to the innermost phosphate in the chain of the new nucleotide, a reaction that release a two-phosphate unit called pyrophosphate. The pyrophosphate will be further cleaved into two individual phosphate ions. The result of the reaction is the addition of the T nucleotide to the growing strand of DNA. The 3' hydroxyl of the T nucleotide is now exposed at the end of the chain. This hydroxyl can participate in a new reaction with the phosphate of the next nucleotide to be added to the chain. In prokaryotes such as *E. coli*, there are two main DNA polymerases involved in DNA replication: DNA pol III (the major DNA-maker), and DNA pol I, which plays a crucial supporting role we'll examine later.

Starting DNA replication

How do DNA polymerases and other replication factors know where to begin? Replication always starts at specific locations on the DNA, which are called **origins of replication** and are recognized by their sequence. *E. coli*, like most bacteria, has a single origin of replication on its chromosome. The origin is about 245245 base pairs long and has mostly A/T base pairs (which are held together by fewer hydrogen bonds than G/C base pairs), making the DNA strands easier to separate. Specialized proteins recognize the origin, bind to this site, and open up the DNA. As the DNA opens, two Y-shaped structures called **replication forks** are formed, together making up what's called a **replication bubble**. The replication forks will move in opposite directions as replication proceeds



Bacterial chromosome. The double-stranded DNA of the circular bacteria chromosome is opened at the origin of replication, forming a replication bubble. Each end of the bubble is a replication fork, a Y-shaped junction where double-stranded DNA is separated into two single strands. New DNA complementary to each single strand is synthesized at each replication fork. The two forks move in opposite directions around the circumference of the bacterial chromosome, creating a larger and larger replication bubble that grows at both ends.

Diagram based on similar illustration in Reece et al. ^22.

How does replication actually get going at the forks? **Helicase** is the first replication enzyme to load on at the origin of replication^33. Helicase's job is to move the replication forks forward by "unwinding" the DNA (breaking the hydrogen bonds between the nitrogenous base pairs).Proteins called **single-strand binding proteins** coat the separated strands of DNA near the replication fork, keeping them from coming back together into a double helix.

Primers and primase

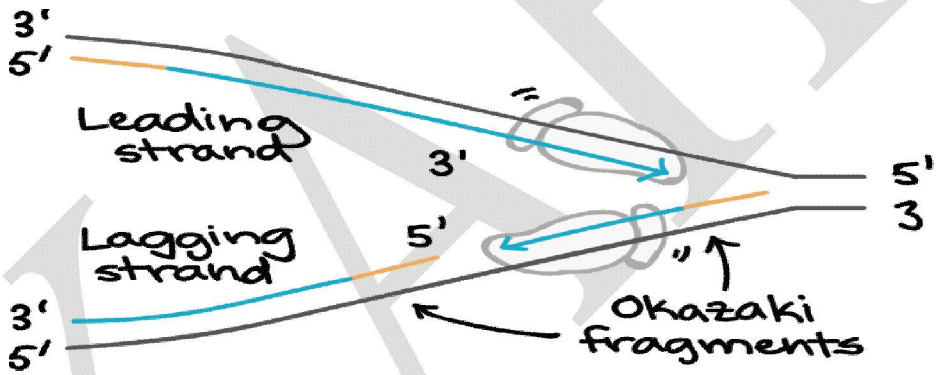
DNA polymerases can only add nucleotides to the 3' end of an existing DNA strand. (They use the free -OH group found at the 3' end as a "hook," adding a nucleotide to this group in the polymerization reaction.) How, then, does DNA polymerase add the first nucleotide at a new replication fork? Alone, it can't! The problem is solved with the help of an enzyme

called **primase**. Primase makes an RNA **primer**, or short stretch of nucleic acid complementary to the template, that provides a 3' end for DNA polymerase to work on. A typical primer is about five to ten nucleotides long. The primer *primes* DNA synthesis, i.e., gets it started. Once the RNA primer is in place, DNA polymerase "extends" it, adding nucleotides one by one to make a new DNA strand that's complementary to the template strand.

Leading and lagging strands

In *E. coli*, the DNA polymerase that handles most of the synthesis is DNA polymerase III. There are two molecules of DNA polymerase III at a replication fork, each of them hard at work on one of the two new DNA strands.

DNA polymerases can only make DNA in the 5' to 3' direction, and this poses a problem during replication. A DNA double helix is always anti-parallel; in other words, one strand runs in the 5' to 3' direction, while the other runs in the 3' to 5' direction. This makes it



necessary for the two new strands, which are also antiparallel to their templates, to be made in slightly different ways One new strand, which runs 5' to 3' towards the replication fork, is the easy one. This strand is made continuously, because the DNA polymerase is moving in the same direction as the replication fork. This continuously synthesized strand is called the **leading strand**.

The other new strand, which runs 5' to 3' away from the fork, is trickier. This strand is made in fragments because, as the fork moves forward, the DNA polymerase (which is moving away from the fork) must come off and reattach on the newly exposed DNA. This tricky strand, which is made in fragments, is called the **lagging strand**.

The small fragments are called **Okazaki fragments**, named for the Japanese scientist who discovered them. The leading strand can be extended from one primer alone, whereas the

lagging strand needs a new primer for each of the short Okazaki fragments.

The maintenance and cleanup crew

Some other proteins and enzymes, in addition the main ones above, are needed to keep DNA replication running smoothly. One is a protein called the **sliding clamp**, which holds DNA polymerase III molecules in place as they synthesize DNA. The sliding clamp is a ring-shaped protein and keeps the DNA polymerase of the lagging strand from floating off when it re-starts at a new Okazaki fragment.

Topoisomerase also plays an important maintenance role during DNA replication. This enzyme prevents the DNA double helix ahead of the replication fork from getting too tightly wound as the DNA is opened up. It acts by making temporary nicks in the helix to release the tension, then sealing the nicks to avoid permanent damage.Finally, there is a little cleanup work to do if we want DNA that doesn't contain any RNA or gaps. The RNA primers are removed and replaced by DNA through the activity of **DNA polymerase I**, the other polymerase involved in replication. The nicks that remain after the primers are replaced get sealed by the enzyme **DNA ligase**.

Summary of DNA replication in *E. coli*

Let's zoom out and see how the enzymes and proteins involved in replication work together to synthesize new DNA.

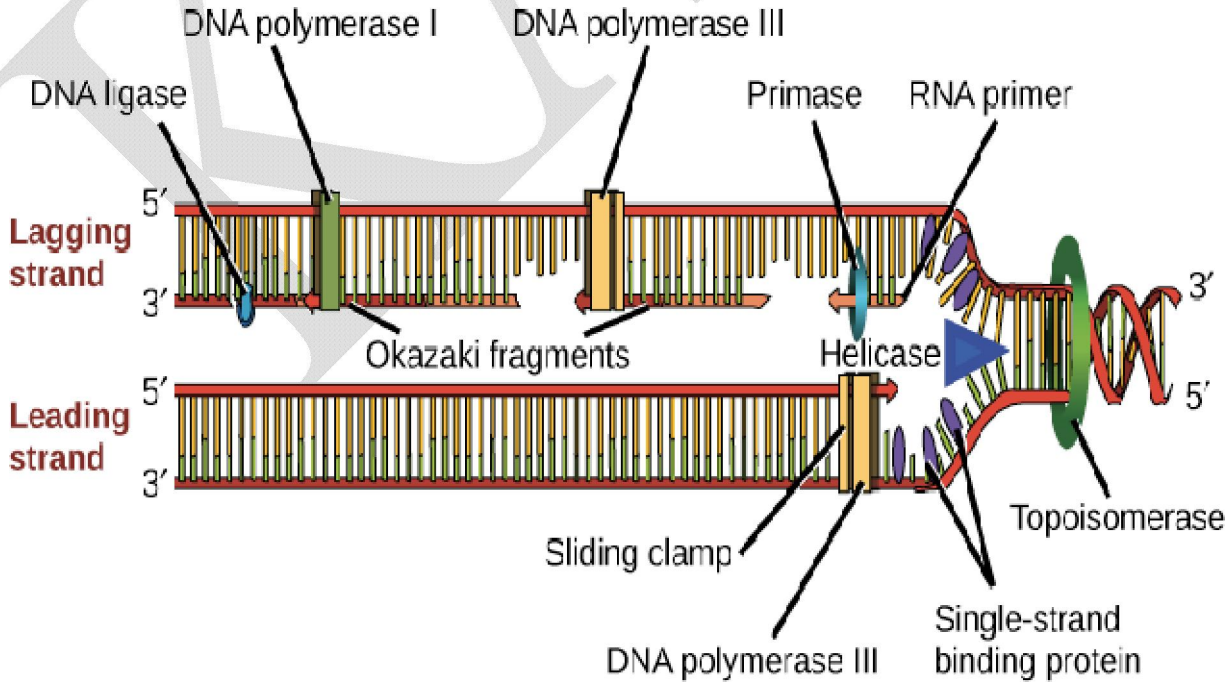


Illustration shows the replication fork. Helicase unwinds the helix, and single-strand binding proteins prevent the helix from re-forming. Topoisomerase prevents the DNA from getting

too tightly coiled ahead of the replication fork. DNA primase forms an RNA primer, and DNA polymerase extends the DNA strand from the RNA primer. DNA synthesis occurs only in the 5' to 3' direction. On the leading strand, DNA synthesis occurs continuously. On the lagging strand, DNA synthesis restarts many times as the helix unwinds, resulting in many short fragments called "Okazaki fragments." DNA ligase joins the Okazaki fragments together into a single DNA molecule.

- **Helicase** opens up the DNA at the replication fork.
- **Single-strand binding proteins** coat the DNA around the replication fork to prevent rewinding of the DNA.
- **Topoisomerase** works at the region ahead of the replication fork to prevent supercoiling.
- **Primase** synthesizes RNA primers complementary to the DNA strand.
- **DNA polymerase III** extends the primers, adding on to the 3' end, to make the bulk of the new DNA.
- RNA primers are removed and replaced with DNA by **DNA polymerase I**.
- The gaps between DNA fragments are sealed by **DNA ligase**.

DNA replication in eukaryotes

The basics of DNA replication are similar between bacteria and eukaryotes such as humans, but there are also some differences:

- Eukaryotes usually have multiple linear chromosomes, each with multiple origins of replication. Humans can have up to 100,100,000000 origins of replication!
- Most of the *E. coli* enzymes have counterparts in eukaryotic DNA replication, but a single enzyme in *E. coli* may be represented by multiple enzymes in eukaryotes. For instance, there are five human DNA polymerases with important roles in replication^55.
- Most eukaryotic chromosomes are linear. Because of the way the lagging strand is made, some DNA is lost from the ends of linear chromosomes (the telomeres) in each round of replication.

Enzymes and Proteins involved in DNA replication

The following points highlight the seven important enzymes involved in the process of DNA replication

of prokaryotes. The enzymes are: 1. DNA Polymerase 2. Primase 3. Polynucleotide Ligase 4. Endonucleases 5. Pilot Proteins 6. Helicase 7. Single-Strand Binding (SSB)

Protein.

Enzyme # 1. DNA Polymerase:

DNA polymerase is the chief enzyme of DNA replication. DNA polymerase activity was discovered by Kornberg in 1956; this activity was due to DNA polymerase I. E. coli has four more enzymes, DNA polymerase II, III (Table. 28.1), IV and V; DNA polymerase III (Pol III) is concerned with DNA replication, while the remaining four enzymes are involved in DNA repair.

All DNA polymerases require the following:

(1) A template DNA strand,

(2) A short primer (either RNA or DNA), and

(3) A free 3′ -OH in the primer.

They add one nucleotide at a time to the free 3′ -OH of the primer, and extend the primer chain in 5′ → 3′ direction.

A. DNA Polymerase I:

DNA polymerase I enzyme provides the major part of activity in E. coli. It is chiefly a DNA repair enzyme, and is used for in vitro DNA replication.

This enzyme has the following three activities:

(i) The 5′ → 3′ polymerase activity is responsible for primer extension or DNA synthesis.

(ii) The 5′ → 3′ exonuclease activity is involved in excision of DNA strands during DNA repair; it removes ~ 10 bases at a time. An exonuclease digests nucleic acids (here DNA) from one end, and it does not cut DNA internally.

(iii) The 3′ → 5′ exonuclease activity is responsible for proof-reading.

In this case, only one nucleotide is removed at a time. The polymerase action does commit errors in DNA synthesis. DNA polymerase is known to scrutinize the new bases roof-

TABLE 28.1. Properties of DNA polymerases I, II and III of *E. coli*

	DNA polymerase		
	I	***II***	***III***
Molecular weight (Daltons)	1,09,000	1,20,000	>2,50,000
Constitution	Monomer	Not known	Heteromultimer
Molecules/cell	400	Not known	10-20
Nucleotides polymerised at 37° C/min/molecule	Upto 1,000	Upto 50	Upto 15,000
Affinity for 5′-triphosphates of deoxyribonucleosides	Low	Low	High
Activities:			
5′ → 3′ polymerase	present	present	present
3′ → 5′ exonuclease	present	present	present
5′ → 3′ exonuclease	present	absent	absent
Functions in	DNA repair; excision of RNA primers	DNA repair	***DNA replication; it is the real replicase***
Mutant loci	*PolA*	*PolB*	*dnaE (polC), dnaN, dnaX, dnaQ, dnaT*

tiate

lytic

now

l.

Prep

B. DNA Polymerase II:

DNA polymerase II enzyme functions in DNA-repair. It has 5′ → 3′ polymerase and 3′ → 5′ exonuclease activities, and uses as template only such DNA duplexes that have short gaps.

C. DNA Polymerase III:

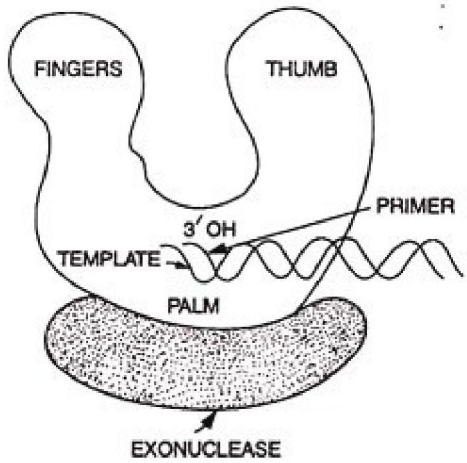


FIG. 28.15. Schematic representation of crystal structure of DNA polymerase belonging to *E. coli* pol I family. These subdomains create a DNA binding cleft, which is the polymerase catalytic domain; the active site lies in the palm. Proof-reading is provided by the exonuclease domain located below the palm.

DNA polymerase III enzyme is responsible for DNA replication in vivo. It has 5'→ 3′ polymerase and 3'→ 5′ exonuclease activities. It catalyzes DNA synthesis at very high rates, e.g., 15,000 bases/min at 37°C. It is composed of several subunits. A DNA polymerase molecule has the following 4 functional sites involved in polymerase activity (Fig. 28.15).

(i) Template site binds the strand serving as template during replication.

(ii) Primer site binds to the primer used for DNA replication.

(iii) Primer terminus site binds only to such primers that have free 3′ -OH.

(iv) The nucleotide triphosphate site binds to the deoxynucleotide 5′-triphosphate that is complementary to the corresponding nucleotide of the template. It also catalyzes the formation of phosphodiester bond between the 5′ phosphate of this nucleotide and the 3′ -OH of the terminal primer nucleotide.

[In addition, the polymerase mole-cule has (5) a 3′ → 5′ exonuclease site and (6) a 5'→ 3′ exonuclease site (in case of DNA polymerase I only)].

In case of eukaryotes, at least nine different DNA polymerases are found; Table 28.2 lists the properties of five of these enzymes. DNA polymerase δ replicates the leading strand, while DNA polymerase ϵ synthesizes the lagging strand.

TABLE 28.2. Properties of the α, β, ε, γ and δ polymerases of eukaryotes

Characteristic	*Polymerase*				
	*α**	*β*	*ε*	*γ*	*δ*
Location	Nuclear	Nuclear	Nuclear	Mitochondrial	Nuclear
Function	Replication initiation of both strands and priming	Repair	Replication of lagging strand	Replicaton of mt DNA	Replication elongation of leading strands
Mass (Daltons)	3,00,000	40,000	250,000	180,000-300,000	170,000-230,000
Number of subunits	4 (180 kDa catalytic core, 2 primases, 1 unknown)	1 (catalytic)	2 (catalytic core, 1 unknown)	3 (catalytic core, 2 unknown)	2 (catalytic core, 1 unknown)
Relative activity	about 80%	10-15%	—	2-15%	—

*Sometimes called Polα/primase

DNA polymerase α catalyzes priming of both the strands. DNA polymerases ξ, η, τ, and k are all nuclear DNA repair enzymes. DNA polymerase y is found in mitochondria and catalyzes replication of mtDNA.

Enzyme # 2. Primase:

This enzyme activity catalyzes the synthesis of RNA primers to initiate DNA replication. In E. coli, DnaG functions as primase. But in eukaryotes, DNA polymerase α provides this function. There are, however, several other ways in which primers are produced, e.g., the 3′-OH generated by a nick in the template DNA molecule.

Enzyme # 3. Polynucleotide Ligase:

DNA ligase or polynucleotide ligase catalyzes the formation of phosphodiester linkage between two immediate neighbour nucleotides of a DNA strand. Thus it seals the nicks remaining in a DNA strand either following DNA replication or DNA repair. However, this enzyme cannot fill the gaps in DNA strands.

Enzyme # 4. Endonucleases:

An endonuclease produces an internal cut (single- or double-stranded) in a DNA molecule. But a restriction endonuclease produces cuts only at those sites that have a specific base sequence. During DNA replication, the origin to produce one single strand copy of the phage genome (Fig. 28.8). endonuclease may induce a nick to initiate DNA

replication, or it may induce nicks to generate a swivel for DNA unwinding. Restriction endonucleases are required for DNA repair.

Enzyme # 5. Pilot Proteins:

Pilot proteins are produced by most viruses. The type of pilot proteins associated with viral genome determines whether the viral DNA will undergo replication or it would support transcription.

Enzyme # 6. Helicase:

Helicase effects strand separation at the forks and uses one ATP molecule for each base that is separated. In E. coli, DNA functions as helicase; this protein is a hexamer and it moves with the replication fork.

Enzyme # 7. Single-Strand Binding (SSB) Protein:

SSB protein binds to single-stranded DNA, and prevents it from forming duplex DNA or secondary structures. SSB binds as a monomer, but it binds cooperatively in that binding of one SSB molecule facilitates binding of more SSB monomers to the same DNA strand. E. coli SSB is a tetramer.

Models of DNA replication

Rolling circle replication

Rolling circle replication occurs in many small bacterial and eukaryotic viruses, e.g., in bactriophage φX174 φ, Greek letter 'phi' here stands for phage). φX174 has a single stranded genome, i.e., the plus-strand, which is converted through discontinuous replication into the double-stranded replicative form (RF).

Phage protein A induces a nick, i.e., break in a single strand, in one strand (the + strand) of the double-stranded RF chromosome. The nick occurs at the origin and protein A binds to the 5′-end so produced. The 3′-end produced at the nick serves as primer and DNA synthesis proceeds from this end.

As a result, the + strand is progressively displaced from the (-) strand beginning at the 5′-end. Replication produces a single strand copy of the + strand, which functions as the viral genome. When a complete positive strand has been synthesized, and a new origin is produced in the (+) strand, protein A nicks in

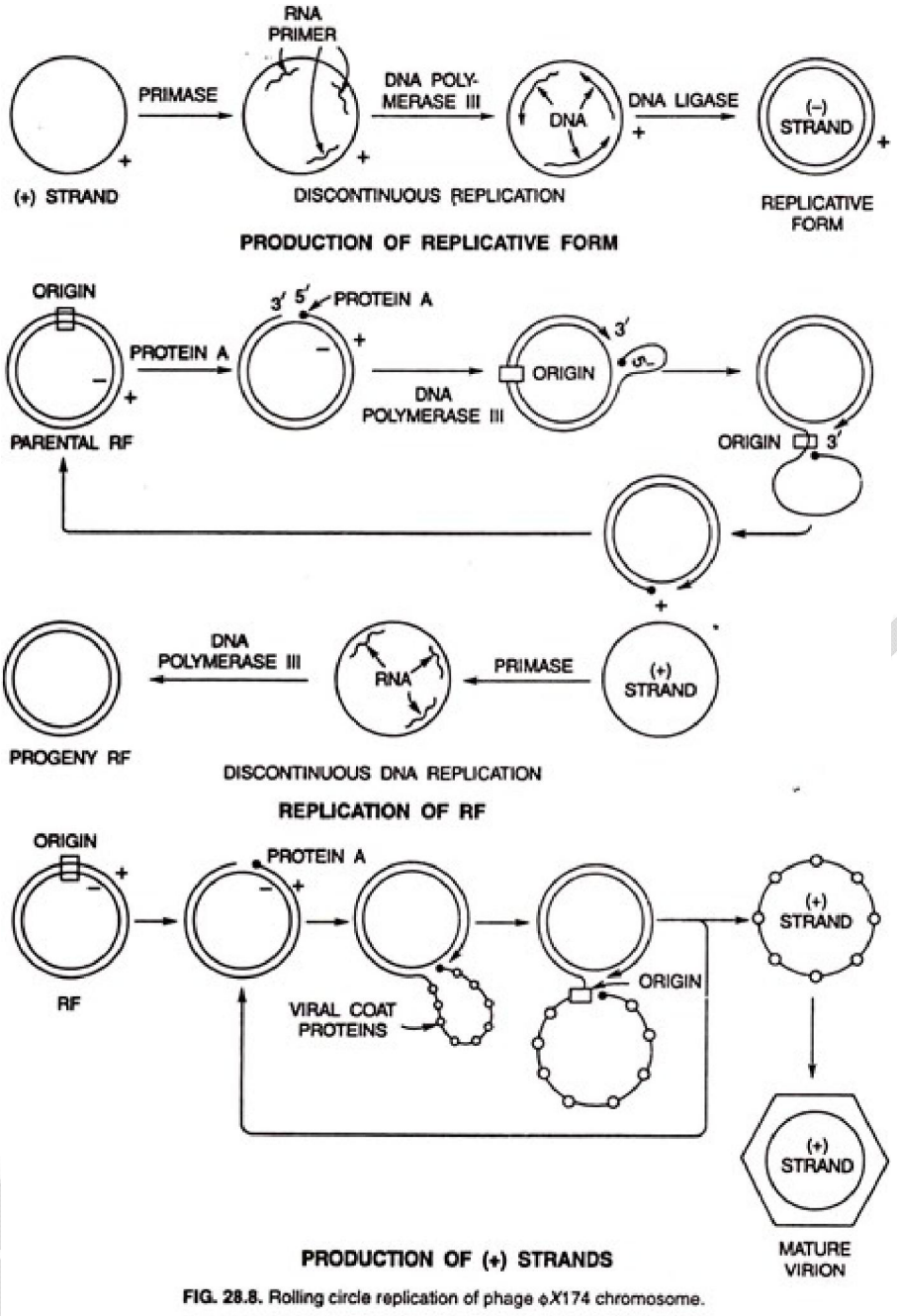


FIG. 28.8. Rolling circle replication of phage $\phi X174$ chromosome.

(3) Displacement Loop DNA Replication:

This mode of replication occurs in mammalian mitochondria; the mitochondrial chromosome in circular. One strand of the chromosome is denoted as H strand, while its complementary strand is named L strand. Replication begins at a specific origin, but only one strand, the H strand, is replicated; the other strand (the L strand) of the chromosome is displaced forming a loop, called displacement loop or D loop.When replication of (the H strand) has progressed up to about 2/3 of the chromosome, the replication of the

displaced single strand (L strand) begins.

This replication begins at a different origin and replication proceeds in the opposite direction. Thus both the strands of DNA are replicated in a continuous manner, their replications begin at different origins, and replication of one strand (H strand) begins much earlier than that of the other (L strand).In this case, replication is not only unidirectional, but the replication of only one of the two strands takes place at each of the replication forks (Fig. 28.9). The same is the case with rolling circle replication.

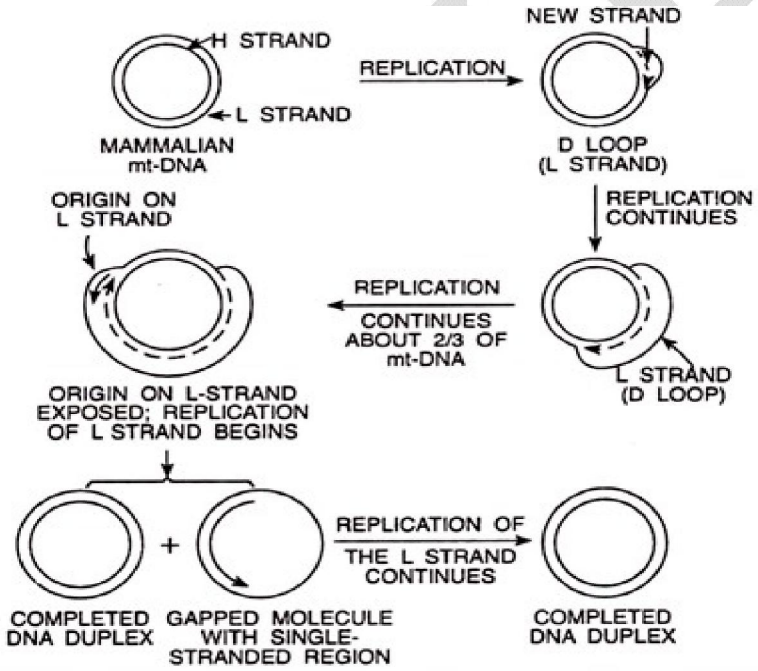


FIG. 28.9. Replication of mammalian mtDNA through the formation of D loops. The replication of only H strand takes place initially in one direction only; the L strand replication begins much later in the opposite direction.

Excision Repair

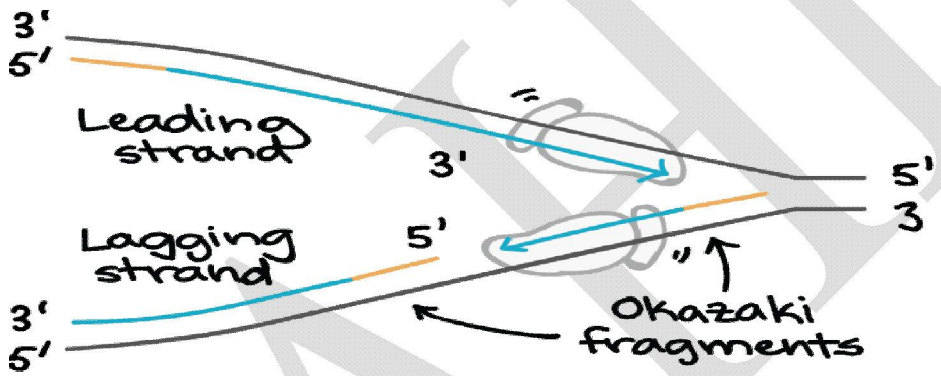
The most common means of repairing damage or a mismatch is to cut it out of the duplex DNA and recopy the remaining complementary strand of DNA, as outlined in Fig. 7.12. Three different types of excision repair have been characterized: nucleotide excision repair, base excision repair, and mismatch repair. All utilize a **cut, copy, and paste** mechanism. In the **cutting**stage, an enzyme or complex removes a damaged base or a string of nucleotides from the DNA. For the **copying**, a DNA polymerase (DNA polymerase I in *E. coli*) will copy the template to replace the excised, damaged strand. The DNA polymerase can initiate synthesis from 3' OH at the single-strand break (nick) or gap in the DNA remaining at the site of damage after excision. Finally, in the

pastingstage, DNA ligase seals the remaining nick to give an intact, repaired DNA.

Figure 7.12. A general scheme for excision repair, illustrating the cut (steps 1 and 2), copy (step 3) and paste (step 4) mechanism.

Nucleotide Excision Repair (NER)

In *nucleotide excision repair*, damaged bases are cut out within a string of nucleotides, and replaced with DNA as directed by the undamaged template strand. This repair system is used to remove pyrimidine dimers formed by UV radiation as well as



nucleotides modified by bulky chemical adducts. The common feature of damage that is repaired by nucleotide excision is that the modified nucleotides cause a significant distortion in the DNA helix. NER occurs in almost all organisms examined.Some of the best-characterized enzymes catalyzing this process are the UvrABC excinuclease and the UvrD helicase in *E. coli.* The genes encoding this repair function were discovered as mutants that are highly sensitive to UV damage, indicating that the mutants are defective in UV repair. As illustrated in Fig. 7.13, wild type *E. coli* cells are killed only at higher doses of UV radiation. Mutant strains can be identified that are substantially more sensitive to UV radiation; these are defective in the functions needed for UV-resistance, abbreviated *uvr*. By collecting large numbers of such mutants and testing them for their ability to restore resistance to UV radiation in combination, complementation groups were identified. Four of the complementation groups, or genes, encode proteins that play major rules in NER; they are *uvrA, uvrB, uvrC*and *uvrD*. One new strand, which runs 5' to 3' towards the replication fork, is the easy one. This strand is made continuously, because the DNA polymerase is moving in the same

direction as the replication fork. This continuously synthesized strand is called the **leading strand**.

One new strand, which runs 5' to 3' towards the replication fork, is the easy one. This strand is made continuously, because the DNA polymerase is moving in the same direction as the replication fork. This continuously synthesized strand is called the **leading strand**.

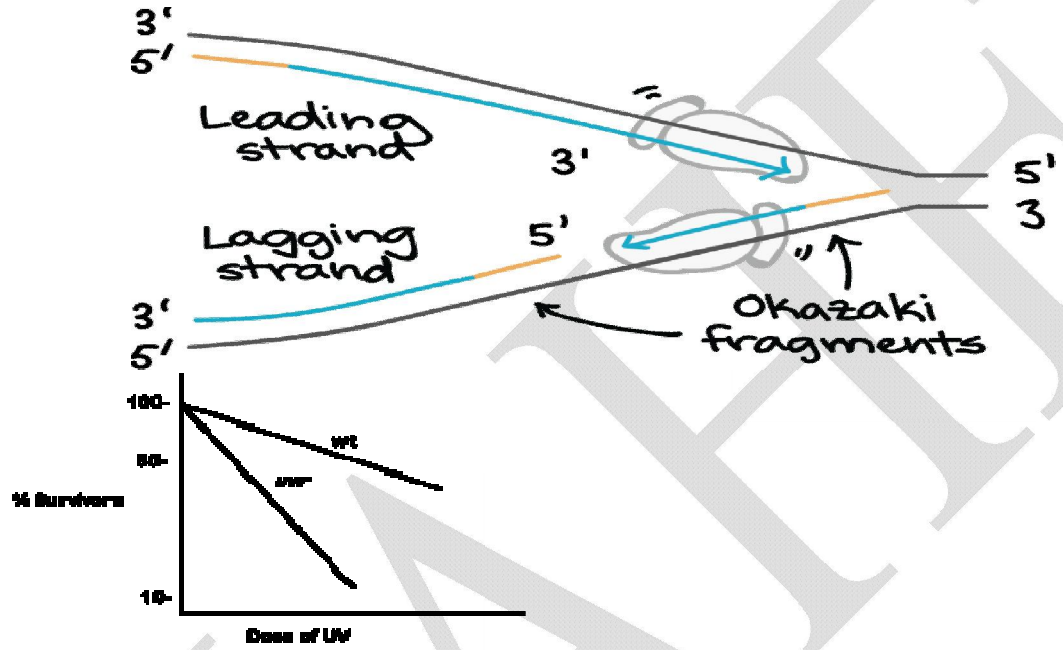


Figure 7.13. Survival curve of bacteria exposed to UV radiation. Cultures of bacteria are exposed to increasing doses of UV radiation, plotted along the horizontal axis. Samples of each irradiated culture are then plated and the number of surviving colonies are counted (plotted as a logarithmic function on the vertical axis). Mutant strains that are more sensitive to UV damage are defective in the genes that confer UV-resistance, i.e. they are defective in *uvr* functions.

The enzymes encoded by the *uvr*genes have been studied in detail. The polypeptide products of the *uvrA*, *uvrB*, and *uvrC*genes are subunits of a multisubunit enzyme called the **UvrABC excinuclease**. UvrA is the protein encoded by *uvrA*, UvrB is encoded by *uvrB*, and so on. The UvrABC complex recognizes damage-induced structural distortions in the DNA, such as pyrimidine dimers. It then cleaves on both sides of the damage. Then UvrD (also called helicase II), the product of the *uvrD*gene, unwinds the DNA, releasing the damaged segment. Thus for this system, the UvrABC and UvrD proteins

carry out a series of steps in the cutting phase of excision repair. This leaves a gapped substrate for copying by DNA polymerase and pasting by DNA ligase.

The UvrABC proteins form a dynamic complex that recognizes damage and makes endonucleolytic cuts on both sides. The two cuts around the damage allow the single-stranded segment containing the damage to be excised by the helicase activity of UvrD. Thus the UvrABC dynamic complex and the UvrBC complex can be called **excinucleases**. After the damaged segment has been excised, a gap of 12 to 13 nucleotides remains in the DNA. This can be filled in by DNA polymerase and the remaining nick sealed by DNA ligase. Since the undamaged template directs the synthesis by DNA polymerase, the resulting duplex DNA is no longer damaged.

In more detail, the process goes as follows (Fig. 7.14). $UvrA_2$ (a dimer) and Uvr B recognize the damaged site as a (UvrA)2UvrB complex. $UvrA_2$ then dissociates, in a step that requires ATP hydrolysis. This is an autocatalytic reaction, since it is catalyzed by UvrA, which is itself an ATPase. After UvrA has dissociated, UvrB (at the damaged site) forms a complex with UvrC. The UvrBC complex is the active nuclease. It makes the incisions on each side of the damage, in another step that requires ATP. The phosphodiester backbone is cleaved 8 nucleotides to the 5' side of the damage and 4-5 nucleotides on the 3' side. Finally, the UvrD helicase then unwinds DNA so the damaged segment is removed. The damaged DNA segment dissociates attached to the UvrBC complex. Like all helicase reactions, the unwinding requires ATP hydrolysis to disrupt the base pairs. Thus ATP hydrolysis is required at three steps of this series of reactions.

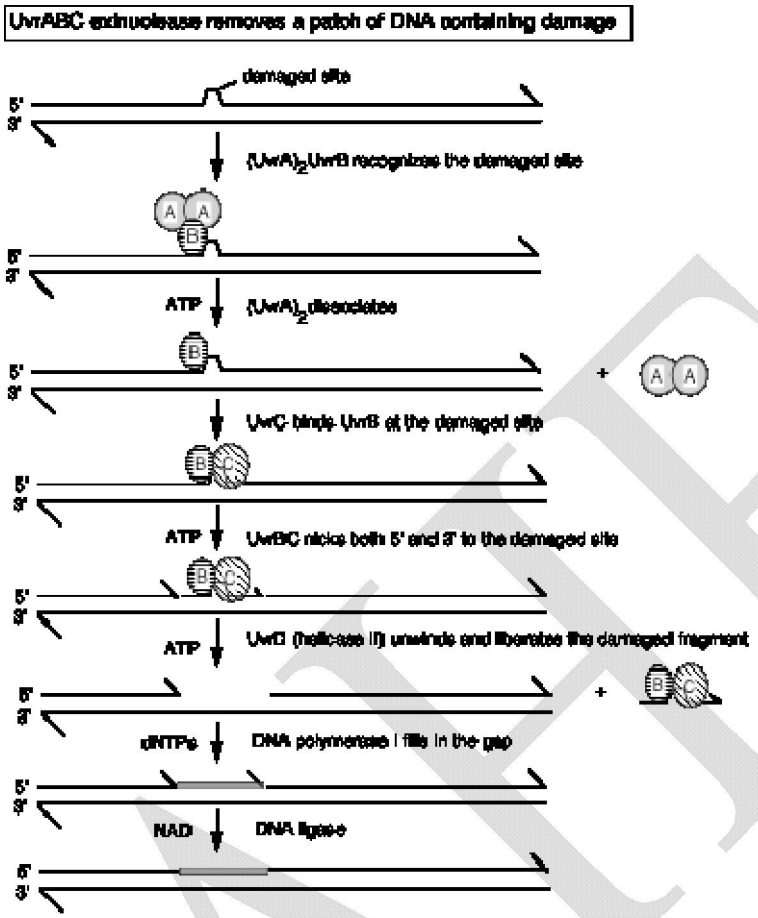


Figure 7.14.*The Uvr(A)BC excinuclease of E. coli recognizes AP sites, thymine dimers, and other structural distortions and makes nicks on both sides of the damaged region. The 12-13 nucleotide-long fragment is released together with the excinuclease by helicase II action.*

Base Excision Repair

Base excision repair differs from nucleotide excision repair in the types substrates recognized and in the initial cleavage event. Unlike NER, the base excision machinery recognizes damaged bases that do not cause a significant distortion to the DNA helix, such as the products of oxidizing agents. For example, base excision can remove uridines from DNA, even though a G:U base pair does not distort the DNA. Base excision repair is versatile, and this process also can remove some damaged bases that do distort the DNA, such as methylated purines. In general, the initial recognition is a specific damaged base, not a helical distortion in the DNA. A second major difference is that the initial cleavage is directed at the glycosidic bond connecting the purine or pyrimidine base to a deoxyribose in DNA. This contrasts with the initial cleavage of a phosphodiester bond in NER.

Cells contain a large number of specific **glycosylases** that recognize damaged or inappropriate bases, such as uracil, from the DNA. The glycosylase removes the damaged or inappropriate base by catalyzing cleavage of the N-glycosidic bond that attaches the base to the sugar-phosphate backbone. For instance, uracil-N- glycosylase, the product of the *ung*gene, recognizes uracil in DNA and cuts the N-glycosidic bond between the base and deoxyribose (Fig. 7.15). Other glycosylases recognize and cleave damaged bases. For instance methylpurine glycosylase removes methylated G and A from DNA. The result of the activity of these glycosylases is an apurinic/apyrimidinic site, or AP site (Fig. 7.15). At an AP site, the DNA is still an intact duplex, i.e. there are no breaks in the phosphodiester backbone, but one base is gone.

Next, an **AP endonuclease** nicks the DNA just 5' to the AP site, thereby providing a primer for DNA polymerase. In *E. coli*, the 5' to 3' exonuclease function of DNA polymerase I removes the damaged region, and fills in with correct DNA (using the 5' to 3' polymerase, directed by the sequence of the undamaged complementary strand).

Additional mechanisms have evolved for keeping U's out of DNA. *E. coli*also has a dUTPase, encoded by the *dut*gene, which catalyzes the hydrolysis of dUTP to dUMP. The product dUMP is the substrate for thymidylate synthetase, which catalyzes conversion of dUMP to dTMP. This keeps the concentration of dUTP in the cell low, reducing the chance that it will be used in DNA synthesis. Thus the combined action of the products of the *dut*+ *ung*genes helps prevent the accumulation of U's in DNA.

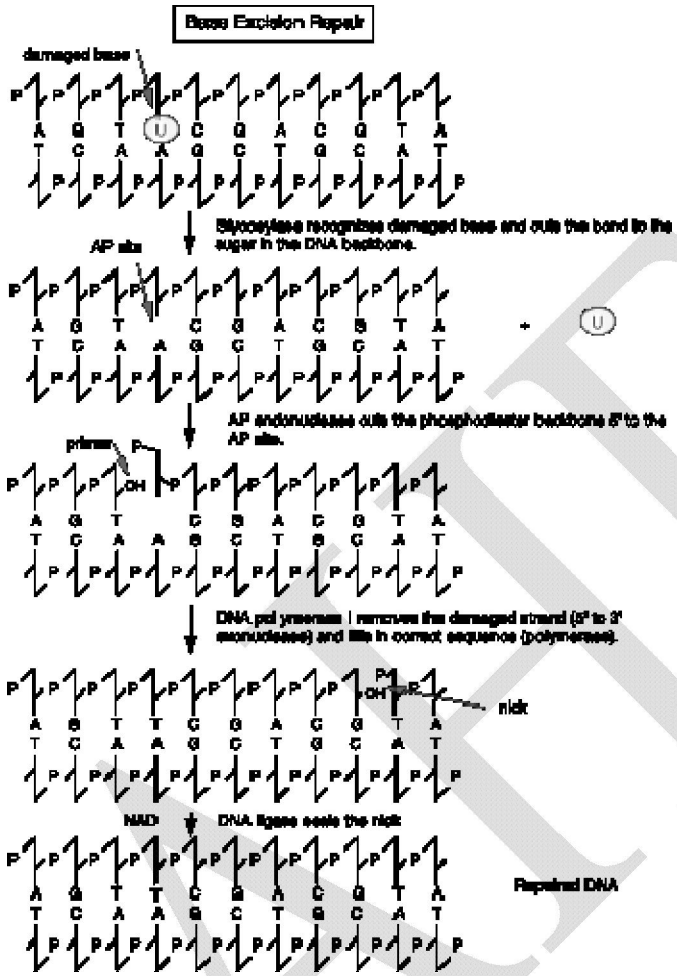


Figure 7.15. *Base excision repair is initiated by a glycosylase that recognizes and removes chemically damaged or inappropriate bases in DNA. The glycosylase cleaves the glycosidic bond between the base and the sugar, leaving an apurinic/apyrimidinic site. The AP endonuclease can then nick the phosphodiester backbone 5' to the AP site. When DNA polymerase I binds the free primer end at the nick, its 5'-3' exonuclease activity cuts a few nucleotides ahead of the missing base, and its polymerization activity fills the entire gap of several nucleotides.*

Mismatch Repair

The third type of excision repair we will consider is **mismatch repair**, which is used to repair errors that occur during DNA synthesis. Proofreading during replication is good but not perfect. Even with a functional e subunit, DNA polymerase III allows the wrong

nucleotide to be incorporated about once in every 108 bp synthesized in *E. coli*. However, the measured mutation rate in bacteria is as low as one mistake per 1010 or 1011 bp. The enzymes that catalyze **mismatch repair**are responsible for this final degree of accuracy. They recognize misincorporated nucleotides, excise them and replace them with the correct nucleotides. In contrast to nucleotide excision repair, mismatch repair does not operate on bulky adducts or major distortions to the DNA helix. Most of the mismatches are substitutes within a chemical class, e.g. a C incorporated instead of a T. This causes only a subtle helical distortions in the DNA, and the misincorporated nucleotide is a normal component of DNA. The ability of a cell to recognize a mismatch reflects the exquisite specificity of **MutS**, which can distinguish normal base pairs from those resulting from misincorporation. Of course, the repair machinery needs to know which of the nucleotides at a mismatch pair is the correct one and which was misincorporated. It does this by determining which strand was more recently synthesized, and repairing the mismatch on the nascent strand.

In *E. coli*, the methylation of A in a GATC motif provides a covalent marker for the parental strand, thus methylation of DNA is used to discriminate parental from progeny strands. Recall that the ***dam* methylase** catalyzes the transfer of a methyl group to the A of the pseudopalindromic sequence GATC in duplex DNA. Methylation is delayed for several minutes after replication. IN this interval before methylation of the new DNA strand, the mismatch repair system can find mismatches and direct its repair activity to nucleotides on the unmethylated, newly replicated strand. Thus replication errors are removed preferentially.

The enzyme complex MutH-MutL-MutS , or MutHLS, catalyzes mismatch repair in *E. coli*. The genes that encode these enzymes, *mutH*, *mutL*and *mutS*, were discovered because strains carrying mutations in them have a high frequency of new mutations. This is called a **mutator phenotype**, and hence the name *mut*was given to these genes. Not all mutator genes are involved in mismatch repair; e.g., mutations in the gene encoding the proofreading enzyme of DNA polymerase III also have a mutator phenotype. This gene was independently discovered in screens for defects in DNA replication (*dnaQ*) and mutator genes (*mutD*). Three complementation groups within the set of mutator alleles have been implicated primarily in mismatch repair; these are *mutH*, *mutL*and *mutS*.

MutS will recognize seven of the eight possible mismatched base pairs (except for C:C)

and bind at that site in the duplex DNA (Fig. 7.16). **MutH**and **MutL** (with ATP bound) then join the complex, which then moves along the DNA in either direction until it finds a hemimethylated GATC motif, which can be as far a few thousand base pairs away. Until this point, the nuclease function of MutH has been dormant, but it is activated in the presence of ATP at a hemimethylated GATC. It cleaves the unmethylated DNA strand, leaving a nick 5' to the G on the strand containing the unmethylated GATC (i.e. the new DNA strand). The same strand is nicked on the other side of the mismatch. Enzymes involved in other processes of repair and replication catalyze the remaining steps. The segment of single-stranded DNA containing the incorrect nucleotide is to be excised by UvrD, also known as helicase II and MutU. SSB and exonuclease I are also involved in the excision. As the excision process forms the gap, it is filled in by the concerted action of DNA polymerase III (Fig. 7.16.).

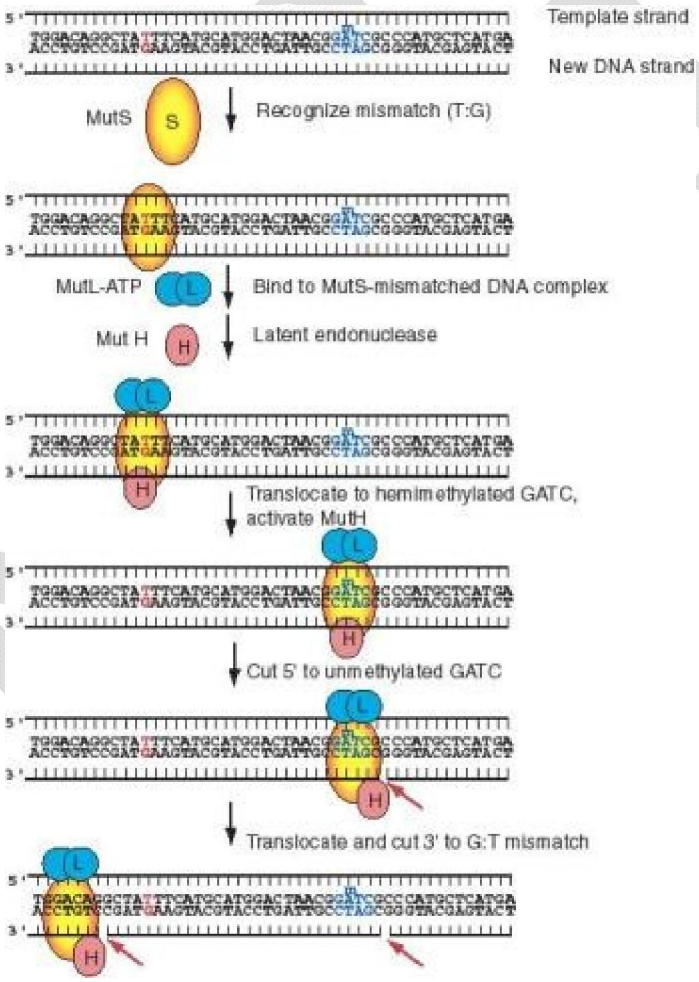


Figure 7.16 (part 1). *Mismatch Repair by MutHLS: recognition of mismatch (shown in*

red), identifying the new DNA strand (using the hemimethylated GATC shown in blue) and cutting to encompass the unmethylated GATC and the misincorporated nucleotide (red G).

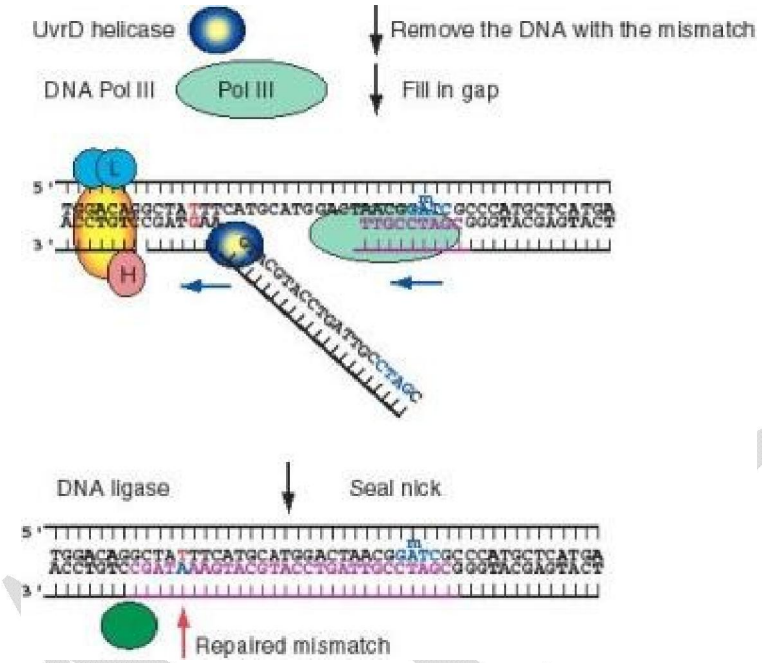


Figure 7.16 (part 2). *Mismatch Repair: excision of the DNA with the misincorporated nucleotide bu Uvr D (aided by exonuclease I and SSB), gap filling by DNA polymerase III and ligation.*

Mismatch repair is highly conserved, and investigation of this process in mice and humans is providing new clues about mutations that cause cancer.Homologs to the *E. coli* genes *mutL*and *mutS*have been identified in many other species, including mammals. The key breakthrough came from analysis of mutations that cause one of the most common hereditary cancers, *hereditary nonpolyposis colon cancer*(HNPCC). Some of the genes that, when mutated, cause this disease encode proteins whose amino acid sequences are significantly similar to those of two of the *E. coli*mismatch repair enzymes. The human genes are called *hMLH1*(for human *mutL*homolog 1), *hMSH1*, and *hMSH2*(for human *mutS* homolog 1 and 2, respectively). Subsequent work has shown that these enzymes in humans are involved in mismatch repair. Presumably the increased frequency of mutation in cells deficient in mismatch repair leads to the

accumulation of mutations in proto- oncogenes, resulting in dysregulation of the cell cycle and loss of normal control over the rate of cell division.

POSSIBLE

QUESTIONS PART A

1) Define Unidirectional and Bidirectional replication?
2) What are the enzymes involved in replication?
3) What are the functions of Polymerase enzyme?
4) What are the various models of replication?
5) What is a Mismatch repair mechanism?
6) What is a Excision repair mechanism?

PART C

1) Explain in detail about Unidirectional and Bidirectional replication?
2) Explain in detail about enzymes involved in replication?
3) Write in detail about the models of replication?
4) Write in detail about Mismatch repair and Excision repair mechanism?
5) Explain in detail about the semi conservative mode of replication with diagram?

KARPAGAM ACADEMY OF HIGHER EDUCATION

CLASS: III B.Sc MB COURSE NAME: MOLECULAR BIOLOGY

COURSE CODE: 17MBU603B UNIT: II BATCH-2017-2020

Sl. No	Question	Option A	Option B	Option C	Option D	Correct Ans
1	Which of the following chemicals induce depurination	Methyl ethane sulphonate	Guanidine	Ethyl sulphonate	Dichlor	Methyl ethane sulphonate
2	Xeroderma pigmentosum is a genetic disorder of	Skin	Hair	Nail	Tongue	Skin
3	UV radiation causes	Adenine dimers	Cytosine dimers	Guanine dimers	Thiamine dimers	Thiamine dimers
4	The result of addition or deletion of one or more base pair in a gene is	Frameshift	Base pair substitution	Misense mutation	Nonsense mutation	Frameshift
5	Repairing mechanism of _____ depends on absorption of visible light by the enzyme.	DNA helicase	DNA ligase	DNA gyrase	DNA photolyase	DNA photolyase
6	Herman J.Mueller reported results of induced mutations on	Yeast	*Drosophila*	Fish	Pea plant	*Drosophila*
7	Chromosomal mutation is	Abberation	Change over	Variation	Genetic change	Abberation
8	_____ is a non-ionizing radiation	Alpha	UV	Gamma	Beta	UV
9	Mutation involving single-base changes are	Induced mutations	Point mutations	Silent mutations	Inverse mutations	Point mutations
10	Transposons was first reported by	Louise pasteur	Koch	Barbara McClintock	Lister	Barbara McClintock
11	Nonsense mutation leads to	Termination of DNA synthesis	Termination of protein synthesis	Termination of cell wall synthesis	Termination of RNA synthesis	Termination of protein synthesis

12	Alkylation is	Addition of methyl or ethyl group	Deletion of ethyl and addition of methyl group	Deletion of methyl and addition of ethyl group	Deletion of methyl or ethyl group	Addition of methyl or ethyl group
13	Chemical mutagens leading to addition of nucleotides to the DNA are	Thimers	Base analogs	Alkylating agents	Interchelating agents	Interchelating agents
14	Human bladder cancer is brought about by a change in single point mutation of	Valine to glycine	Isoleucine to leucine	Leucine to isoleucine	Glycine to valine	Glycine to valine
15	Which of the following dimer formation is most common	Cytidine dimer	Uracil dimer	Thymidine dimer	Adenosine dimer	Thymidine dimer
16	Daughter strand repair is also called as	Recombination repair	SOS repair	Photo repair	Excision repair	Recombination repair
17	An intercalating dye	Sunset yellow	Safranin	India ink	Acridine orange	Acridine orange
18	Most common repair system is	SOS	Photoreactive	Mismatch	Excision	Excision
19	Virus capable of causing mutations is	Bacteriophage Ru	Bacteriophage Mu	Bacteriophage Nu	Bacteriophage Ly	Bacteriophage Mu
20	Potent oxidizing agent that can alter structure of purine and pyrimidine	Free radicals	Water	Dyes	Acids	Free radicals
21	Reverse mutation is	Wild type to mutant	Mutant to wild type	A new gene introduced	A gene deleted	Mutant to wild type
22	Alkylation of guanine causes its removal from DNA in a reaction called	Deamination	Depyrimidation	Degradation	Depurination	Depurination

23	In *E.coli*, parental DNA is methylated at an adenine residue found in the sequence	5' TAGC 3'	5' ATGC 3'	5' CATG 3'	5' GATC 3'	5' GATC 3'
24	Mutation in which there is an amino acid substitution is called	Missense	Nonsense	Silent	Point	Missense
25	Nutritional mutans of neurospora are known as	Phototrophs	Auxotrophs	Heterotrophs	Isotrophs	Auxotrophs
26	UV radiation at 260 nm cross-links adjacent thiamine that produces	Butane ring	Cyclane ring	Butocyclane ring	Phenyl alanine	Cyclobutane ring
27	When a part of chrosome is moved to another chromosome, it is called as	Point mutation	Induced mutation	Spontaneous mutation	Translocation mutation	Translocation mutation
28	Mismatch repair cannot take place if there is a mutation in	Helicase	Polymerase	Ligase	Methylase enzyme	Methylase enzyme
29	Common chemical events that produce spontaneous mutation	Deamination	Depurination	Dimerization	Isomerization	Depurination
30	Bacterium used in Ames test	*Salmonella*	*Shigella*	*Streptococcus*	*Staphylococcus*	*Salmonella*
31	Naturally occurring mutations are	Induced	Spontaneous	Nonsense	Frameshift	Spontaneous
32	Recombinational repair is often due to	many cytidine dimer and associated large gaps in a strand	incorporation of many incorrect nucleotides by DNA polymerase	many thymidine dimer formation and associated large gaps in a strand	DNA breaking	incorporation of many incorrect nucleotides by DNA polymerase

33	The most common error prone repair mechanism is	Mismatch	Excision	SOS	Recombination	SOS
34	Site that mutates at a rate significantly greater thag statistical probability is referred to as	Hotspots	Blackspots	Dotspots	DNA spots	Hotspots
35	Transposition is	Movement of a phage	Movement of a virus	Movement of a transposon	Movement of a plasmid	Movement of transposon
36	Radiation that causes cross chromosomal mutations in humans	UV	Visible	Ionozing	X-rays	Ionozing
37	Frameshift mutation is caused by	Proflavin	Nitrous acid	UV	X-rays	Proflavin
38	Change of purine to pyrimidine base in an mutation	Transition mutation	Transverse mutation	General mutation	Transformation	Transverse mutation
39	Detection of silent mutations require	Aminoacid analysis	Peptide analysis	RNA analysis	Nucleotide analysis	Nucleotide analysis
40	Most frequently employed technique in the study of mutations	Analysis of phenotypes	Analysis of genotypes	Analysis of proteins	Analysis of both phenotypes and genotypes	Analysis of phenotypes
41	Natural phenomena of changes in chemical structure of nitrogenous bases is called	Complementary	Conservative	Tautomeric	Telomeric	Tautomeric
42	Mutation that has no detectable effect on the phenotype of a cell	Point	Induced	Silent	Leaky	Silent
43	Chemical agent that resembles thiamine	5-bromothiamine	5-bromoadenine	5-bromoguanine	5-bromouracil	5-bromouracil

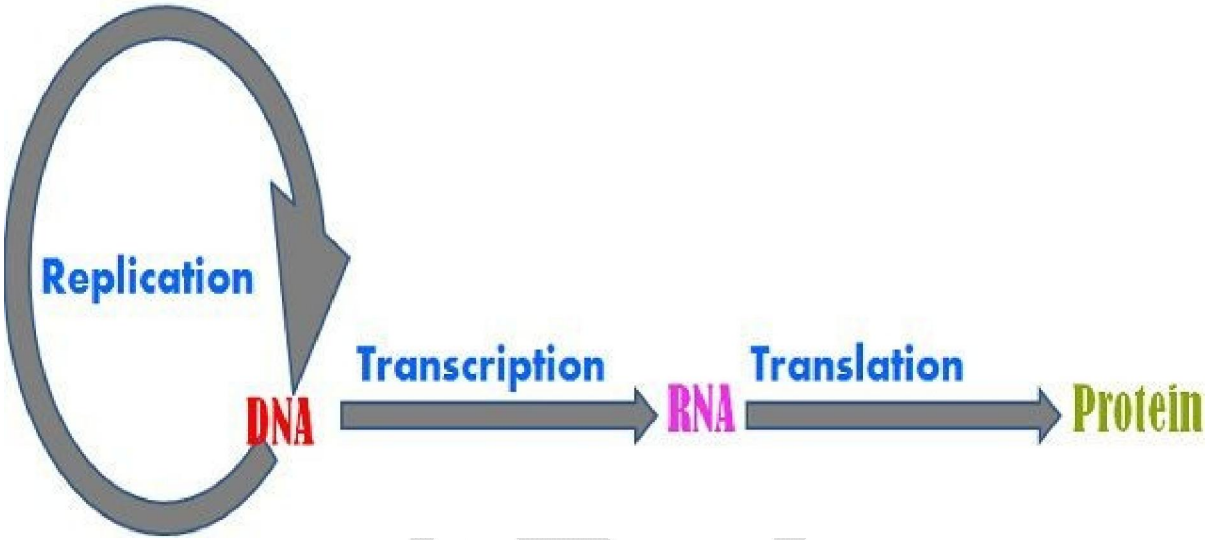
44	Duplication mutation is	Segments of nucleotides sequences are repeated	Segments of nucleotides sequences are deleted	Segments of nucleotides sequences are inserted	Segments of nucleotide sequences are inserted & deleted evenly	Segments of nucleotides sequences are repeated
45	Mutations that result from treatment with mutagens are called	Induced mutation	Uninduced	Spontaneous	Frameshift	Induced mutation
46	Oncogenes are found in certain	Bacteria	Fungi	Viruses	Algae	Viruses
47	Mutation generating new stop codon are called	Nonsense mutation	Misense mutation	Point mutation	Silent mutation	Nonsense mutation
48	UV induced dimers are separated using light energy by	Primase	Photolyase	Dnase	Rnase	Photolyase
49	The function of DNA glycosylase in base excision repair is	Addition of correct base	Addition of correct nucleotide	Removal of incorrect base	Removal of phosphodiester bond	Removal of phosphodiester bond
50	Which of the following biomolecule has self¬repair mechanisms?	DNA,RNA and protein	DNA and RNA	DNA and proteins	DNA only	DNA only
51	Rapid screening technique for mutagens and carcinogens	Aims test	Sima test	Ames test	Sema test	Ames test
52	Mutation resulting from deamination of 5-methylcytosine produces Thymine which pairs with	Uracil	Adenine	Cytosine	Guanine	Adenine
53	Converts amino groups to keto groups by oxidative deamination	Hydrochloric acid	Nitrous acid	Sulphuric acid	Oxalic acid	Nitrous acid

54	In bacteria, a small circle of DNA found outside the main chromosome is called	Cosmid	Bacmid	Transposon	Plasmid	Plasmid
55	Biological agents of mutagenesis are	Transposable elements	Lipids	Bacteria	Carbohydrates	Transposable elements
56	Most common proto-oncogene implicated in human cancers	*s-rac*	*a-src*	*r-cas*	*c-ras*	*c-ras*
57	Deficiency in phenylalanine hydroxidase results in	Cancer	Phenylketonuroa	Melanoma	Asthma	Phenylketonuroa
58	Cancer that results from deletion of a portion of chrosome 13 is	Eye cancer	Bone cancer	Skin cancer	Lung cancer	Eye cancer
59	Mutation in which a purine base is substituted with another purine base is	Transverse mutation	General mutation	Transition mutation	Transduction	Transition mutation
60	Chemical agent that resembles thiamine	5-bromothiamine	5-bromoadenine	5-bromoguanine	5-bromouracil	5-bromouracil

Unit III

Transcription: Definition, difference from replication, promoter - concept and strength of promoter RNA Polymerase and the transcription unit, Transcription in Eukaryotes: RNA polymerases, general Transcription factors. Split genes, concept of introns and exons, RNA splicing, spliceosome machinery, concept of alternative splicing, Polyadenylation and capping, Processing of rRNA, RNA interference: si RNA, miRNA and its significance

Difference Between Replication and Transcription



Replication is processed inside the nucleus and involves the copying of the genetic material so that the new daughter cell thus form contains the identical copies as their parent cells. While **transcription** is processed in the cytoplasm where a segment of DNA is transcribed into RNA. Both the process occurs inside the cell.The flow of biological information from DNA to RNA and then synthesis of proteins is considered as the **'central dogma of life'**. These involve the three main processes which are replication, transcription, and translation. Replication is the process of **duplicating** the ones own genetic materials into two more identical copies, so that the similar information may get further transfer to the new daughter cells.

Transcription involves the **conversion of DNA into RNA**, it is helpful in gene expression of the selected segment of DNA. **Translation** is said as the final step where protein formation takes place. Below we will be discussing the important difference between replication and transcription, and the process involved in it.

Replication Vs Transcription

1. Comparison Chart
2. Definition
3. Key Differences
4. Conclusion

Comparison Chart

BASIS FOR COMPARISON	REPLICATION	TRANSCRIPTION
Definition	Replication is the duplication of strands of Deoxyribonucleic acids (DNA), which gives two daughter strands and each strand contains half of the original DNA.	Transcription is the formation of only single identical Ribonucleic acid (RNA) from the double stranded DNA, which means transcription is the process after replication.
Principle	The main function of replication is to maintain the entire set of the genome for the next generation.	The main function of transcription is to make RNA copies of one's genes and here the genes are expressed of the replicated DNA.
In which phase it occurs	It occurs in the S phase of cell cycle.	It occurs in G1 and G2 phases of cell cycle.
Enzymes involved	DNA helicase, DNA polymerase enzymes, gyrase (eukaryotes).	RNA polymerase, Transcriptase.
It comprises	The unwinding and splitting of the entire DNA molecule (chromosome).	The unwinding and splitting of those genes only, which are to be transcribed.

BASIS FOR COMPARISON	REPLICATION	TRANSCRIPTION
	Also copying of the entire genome.	Copying of few selected genes only.
	There is a hydrogen bonding between the replicated DNA strand and template strand.	Transcribed RNA strands gets separate from its DNA template strand.
	Products do not degrade after their function.	Products gets degraded after their function get complete.
The site of process	Product remains in the nucleus.	Product move from the nucleus to the cytoplasm.
Primer requirement	Requires RNA primer.	No primer required.
Material required	Deoxyribonucleoside triphosphate like dATP, dTTP, dCTP, dGTP serves as raw material.	Ribonucleoside triphosphate like ATP, CTP, GTP, UTP serve as raw materials.
Final Result	It results in the formation of two double-stranded DNA molecule from one DNA molecule and thus giving rise to the two new identical daughter cells.	It results in the formation of RNA molecule from a section of one strand which includes tRNA, rRNA, mRNA and non-coding RNA (like microRNA).

Definition of Replication

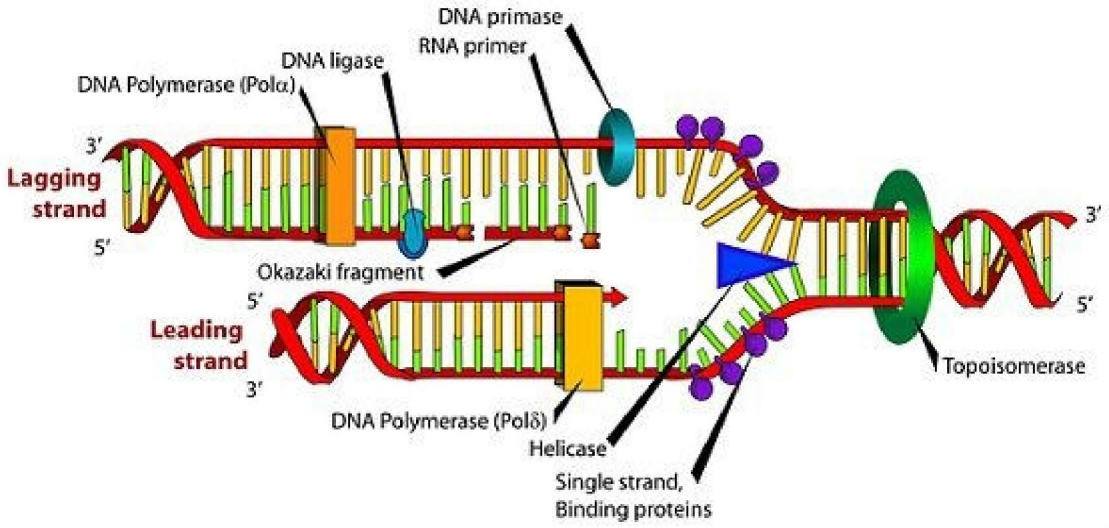
DNA is a macromolecule, which carries genetic information from one generation to next generations. DNA can be regarded as **reserve bank of geneticin formation**. It is responsible for **preserving the identity** of the species over several years.

In the process of cell division, when the cell divides into two identical daughter cell, it also transfers the genetic information from the parent cell. So we can say that replication is a process where DNA copies itself and produces identical daughter molecules of DNA.

The process of **replication is different** in prokaryotes and in eukaryotes. Though it involves the few common steps like the **origin of replication,** it is the site from where the replication begins, at this site the enzyme gets attached and unwind the double helical structure into single and accessible form assisted by the enzyme **DNA helicase**.

The one strand is called **leading** (continuous or forward strand) strand while the another one is called **lagging** (discontinuous or retrograde) strand. This unwinding exposes the unpaired bases to serve as a template for the formation of new strands. The strand ends have their name as **5′ and 3′**, and the process of replication starts from 5′ to 3′ directions, simultaneously on both the strands.

It is said that in prokaryotes synthesis of DNA is **semi-discontinuous**. The primer (a small segment of RNA) is added, eventually proceeding to the addition of nucleotides, which are the complementary base pair with the unpaired base.



The enzyme called **DNA polymerase** helps in the formation of this assembly. Also, the pattern of replication in prokaryotes and in eukaryotes is same, that is the semi-conservative type, where half of the original DNA is conserved and the other is newly formed DNA. This evidence for semi-conservative DNA replication was given by **Meselson and Stahl (1958).**

Now the **difference** between the process of the two is due to the complexity of the cells where eukaryotes are more complex and hence they have multiple origins of replication,

while prokaryotes have a single origin of replication. Also, replication is **unidirectional** in eukaryotes, which is **bi-directional** in prokaryotes.

Enzymes like **DNA polymerase** are only two in number in prokaryotes, while in eukaryotes it is four to five like (α, β,γ, δ, ε). The rate of replication is much faster in prokaryotes than eukaryotes. The DNA in prokaryotes is circular and don't have ends to synthesize. The process of short replication in prokaryotes goes on continuously, whereas the DNA replication of the eukaryotes gets completes in **S-phase** of the cell cycle.

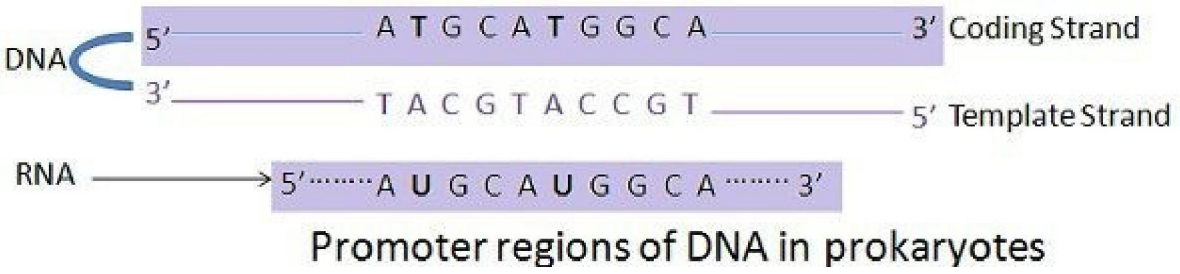
The process is carried out with **high fidelity**, so that the genetic information may get transfer correctly from one generation to generation. **Proofreading** activity is also done by **DNA polymerase III**, which checks the attachment of the nucleotides to the correct base pair. DNA polymerase corrects the mistakes of any mismatch found between the base pairing of the complementary bases.

Definition of Transcription

The **intermediate product** of DNA is RNA, where the genes are expressed after the replication. So it is called the **site of expression** of the genetic information. In this process, one of the two strands formed after replication works as a**template** (noncoding strand or sense strand) and another as **antisense** (coding strand or antisense strand). The almost whole process is same in both prokaryotes as well as in eukaryotes, but there exist some basic differences between them.

The whole molecule of DNA is not expressed in transcription, rather some selected part of DNA is only synthesized as RNA. The reason for this is unknown, but it is said that it might be due to the internal signaling.

The product formed in transcription is referred to as **primary transcript,** as these are **inactive**. So to make them functionally active they undergo certain kind of **alterations** like splicing, base modifications, terminal additions, etc. These are known as **post – transcriptional modifications**.

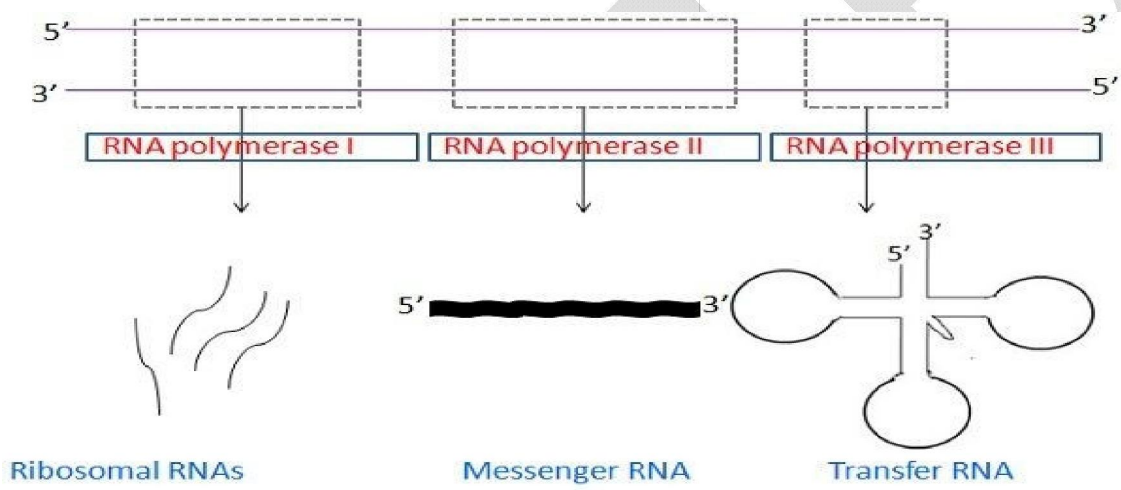


Promoter regions of DNA in prokaryotes

Some of the **similarities** between the prokaryotes and eukaryotes transcription process are like in both the kind DNA act as the template for the process, chemical composition (base pairs) is same, RNA polymerase plays a major role in both the groups.

While the **difference** lies in the process, which is simple in prokaryotes and it is much complexed in eukaryotes. In prokaryotes, only one type of RNA polymerase produces all three types of RNA (mRNA, tRNA, rRNA), while in eukaryotes different types of RNA produces different types of RNA-like type I produces **rRNA**, type II is **mRNA** and type III for **tRNA** and **5S rRNA**.

Apart from this, there are other differences like in initiation site, Rho factor, promoter region, termination point, the presence of introns, post-transcriptional modifications, etc.



An overview of transcription in eukaryotes

Though in many viruses, the genetic material is contained by RNA also and has the ability to perform other cellular function like DNA. But it is chemically found that**DNA is more stable than RNA**, hence DNA is only preferred as more suitable macromolecule for the storing the genetic information long life.

Key Differences Between Replication and Transcription

1. Replication is the duplication of strands of Deoxyribonucleic acids (DNA), which gives two daughter strands and each strand contains half of the original DNA double helix; Transcription is the formationof only single identical Ribonucleic acid (RNA) from the double stranded DNA, which means transcription is the process of replication.

2. The principle function of replication is to maintain and send the copy of the entire set of the genome for the next generation; While transcription work is to make RNA copies and where the genes are expressed of the replicated DNA.

3. Replication occurs in the S phase of cell cycle while Transcription occurs in G1 and G2 phases of cell cycle.

4. Enzymes involved in replication are DNA helicase, DNA polymerase, gyrase (in eukaryotes) and in transcription RNA polymerase, Transcriptase plays a major role.

5. The process of replication and transcription comprises:

- The unwinding and splitting of the entire DNA molecule (chromosome) whereas transcription involves the unwinding and splitting of those genes only, which are to be transcribed.

- The process engages in the copying of the entire genome whereas transcription is copying of few selected genes only.

- There is a hydrogen bonding between the replicated DNA strand and template strand, while transcribed RNA strands get separate from its DNA template strand.

- Products do not degrade after their function, but in transcription process products gets degraded after their function gets complete.

6. The site of the process of replication remains in the nucleus, but during the process, product move from the nucleus to the cytoplasm.

7. Requires RNA primer in the replication process, there is no requirement of primer

8. Deoxyribonucleoside triphosphate like dATP, dTTP, dCTP, dGTP serve as raw material in replication, Ribonucleoside triphosphate like ATP, CTP, GTP, UTP serve as raw materials in transcription.

9. Replication results in the formation of two double-stranded DNA molecule from one DNA molecule and thus giving rise to the two new identical daughter cells while transcription results in the formation of RNA molecule from a section of one strand which includes tRNA, rRNA, mRNA and non-coding RNA (like microRNA).

Conclusion

From the above article, we can say that cell division is a vital and essential process for all the living beings to grow. Prior to cell division involves the most important process called as replication of the DNA. In this process, the genetic material gets divides and is ready to transfer it further to the new daughter cells.

While transcription involves the formation of RNA. This two process involves the enzymes like helicase, DNA polymerase, RNA polymerase, primase, transcriptase.So precisely we can say that DNA makes RNA and RNA make protein, which is the central dogma of all kinds of life.

A promoter conceives an idea for setting-up a particular business at a given place and performs various formalities required for starting a company. A promoter may be a individual, firm, association of persons or a company. The persons who assist the promoter in completing various legal formalities are professional people like Counsels, Solicitors, Accountants etc. and not promoters.

Definitions:

Following definitions of a promoter clarify his status and role:
"A promoter is the one, who undertakes to form a company with reference to a given object and sets it going and takes the necessary steps to accomplish that purpose." —Justice C.J. Cokburn

"A promoter is the person conscious of the possibility of transforming an idea into a business capable of yielding a profit; who brings together various persons concerned and who finally, superintendents the various steps necessary to bring the new business into existence." — Arthur Dewing

Characteristics of a Promoter:

The above given definitions bring out the following characteristics or features of a promoter:
1. A promoter conceives an idea for the setting-up a business.

2. He makes preliminary investigations and ensures about the future prospects of the business.

3. He brings together various persons who agree to associate with him and share the business responsibilities.

4. He prepares various documents and gets the company incorporated.

5. He raises the required finances and gets the company going.

Kinds of Promoters:

The promoters may be of the following types:
1. Professional Promoters:

These are the persons who specialise in promotion of companies. They hand over the companies to shareholders when the business starts. In India, there is lack of professional promoters. In many other countries, professional promoters have played an important role and helped the business community to a great extent. In England, Issue Houses; In U.S.A., Investment Banks and in Germany, Joint Stock Banks have played the role of promoters very appreciably.

2. Occasional Promoters:
These promoters take interest in floating some companies. They are not in promotion work on a regular basis but take up the promotion of some company and then go to their earlier profession. For instance, engineers, lawyers, etc. may float some companies.

3. Financial Promoters:
Some financial institutions of financiers may take up the promotion of a company. They generally take up this work when financial environment is favourable at the time.

4. Managing Agents as Promoters:
In India, Managing Agents played an important role in promoting new companies. These persons used to float new companies and then got their Managing Agency rights. Managing Agency system has since long been abolished in India.

Legal Position of Promoter:

The company law has not given any legal status to promoters. A promoter is neither an agent nor a trustee of the company because it is a non entity before incorporation. Some legal cases have tried to specify the status of a promoter. He stands in a fiduciary position.

The promoter moulds and creates the company and under his supervision it comes into existence. It is the duty of the promoter to get maximum benefits for the company. He should not get secret profits from the company. If he sells his property to company, then he should explain his interest in such property.

Liabilities of a Promoter:

Following are the liabilities of a promoter:
(i) A promoter should not make secret profits out of the dealings of the company.

(ii) He must deposit with the company all money received on its behalf.

(Hi) He must exercise due diligence and care while performing the work of a promoter.

(iv) He will be personally responsible for all the preliminary contracts till all these are approved by the company.

(v) He will compensate any person who made investments in the company on the basis of untrue statements made by the promoter.

Formation of a Company:

The formation of a company involves the following stages:

1. Promotion
2. Incorporation
3. Capital Subscription Stage
4. Commencement of Business

This stage includes all the processes starting from the inception of an idea to the completion of the project. It is the process of planning and arranging various inputs required for running an enterprise. Promotion involves identification of an opportunity, studying of feasibility, assembling the requirements, financing the proposition, etc. The promoter undertakes various stages of promoting an enterprise.

The promoter performs various stages of promotion. Following are the steps in promotion:

1. Identification of Business Opportunity:

The first stage in promotion of a business is the identification of a business opportunity. The promoter visualises that there are opportunities for a particular type of business and it can be run profitability. The idea may be to exploit a new area of natural resources or a venture in the existing line of business. He develops the ideas with the help of technical experts of that field. When the promoter feels that there are opportunities in taking up a particular venture then the idea is taken further.

2. Detailed Investigation:

At the second stage, various factors relating to the business are studied from a practical point of view. The demand for the product is estimated and the likely business share is determined. After determining the prospective demand, the promoter thinks of arranging finances, labour, raw materials, power, etc. The cost structure of the product is analysed to find out profitability from the venture. An expert opinion is sought upon the viability of the project.

3. Approval of Name:

It is necessary to get the name of the company approved from the Registrar of Companies. This is done in order to avoid duplication of the name. Generally, a company submits a list of names in order of preference. The Registrar matches the names with the names of existing companies and then one name is approved.

4. Signatories to Memorandum:

The promoters decide the names of persons to be the signatories to the memorandum of association. Usually, the first signatories to the memorandum become the first directors of the company. The written consent of the persons to act as directors is taken and they are asked to take qualifying shares of the company.

5. Appointment of Professionals:
The next stage is of raising funds and deciding about various contracts. So, promoters appoint the brokers and underwriters to ensure the availability of capital by sale of company's securities. They also appoint solicitors to deal with legal matters of the company.

6. Preparing necessary Documents:
The promoters take steps to prepare various legal documents of the company which have to be submitted to the Registrar of Companies at the time of incorporation. The documents which are required to be prepared include Memorandum of Association, Articles of Association, Prospectus, etc.

Transcription

Transcription is the first step in gene expression. It involves copying a gene's DNA sequence to make an RNA molecule.

- Transcription is performed by enzymes called **RNA polymerases**, which link nucleotides to form an RNA strand (using a DNA strand as a template).
- Transcription has three stages: initiation, elongation, and termination.
- In eukaryotes, RNA molecules must be processed after transcription: they are **spliced** and have a **5' cap** and **poly-A tail** put on their ends.
- Transcription is controlled separately for each gene in your genome.

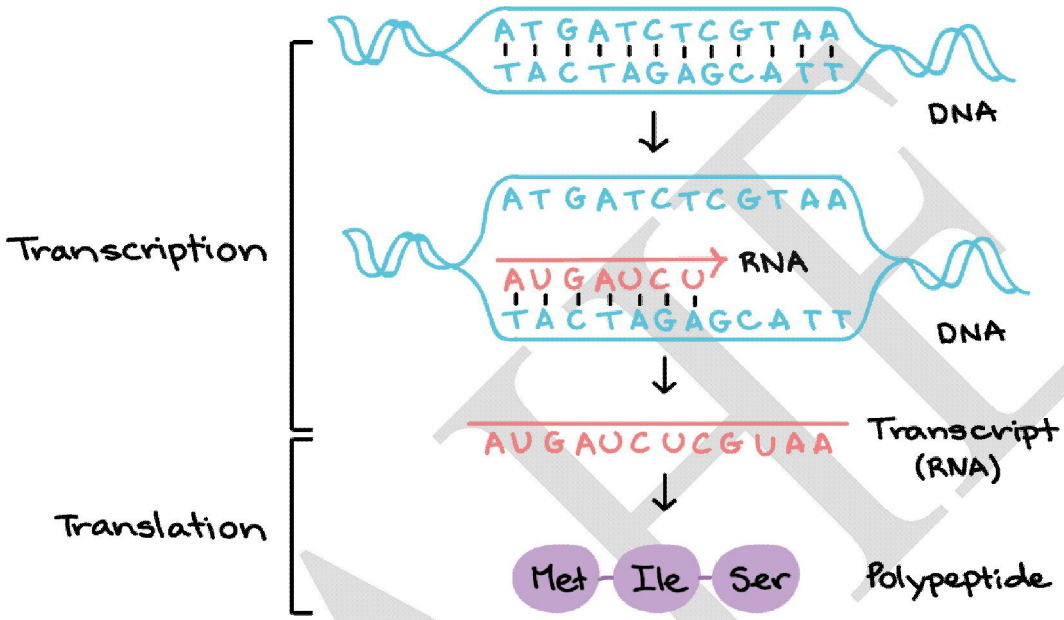
Introduction

Have you ever had to transcribe something? Maybe someone left a message on your voicemail, and you had to write it down on paper. Or maybe you took notes in class, then rewrote them neatly to help you review.

As these examples show, *transcription* is a process in which information is rewritten. Transcription is something we do in our everyday lives, and it's also something our cells must do, in a more specialized and narrowly defined way. In biology, **transcription** is the process of copying out the DNA sequence of a gene in the similar alphabet of RNA.

Overview of transcription

Transcription is the first step in gene expression, in which information from a gene is used to construct a functional product such as a protein. The goal of transcription is to make a RNA copy of a gene's DNA sequence. For a protein-coding gene, the RNA copy, or **transcript**, carries the information needed to build a polypeptide (protein or protein subunit). Eukaryotic transcripts need to go through some processing steps before translation into proteins.



In transcription, a region of DNA opens up. One strand, the template strand, serves as a template for synthesis of a complementary RNA transcript. The other strand, the coding strand, is identical to the RNA transcript in sequence, except that it has uracil (U) bases in place of thymine (T) bases.

Example:

Coding strand: 5'-ATGATCTCGTAA-3' Template strand: 3'-TACTAGAGCATT-5' RNA transcript: 5'-AUGAUCUCGUAA-3'

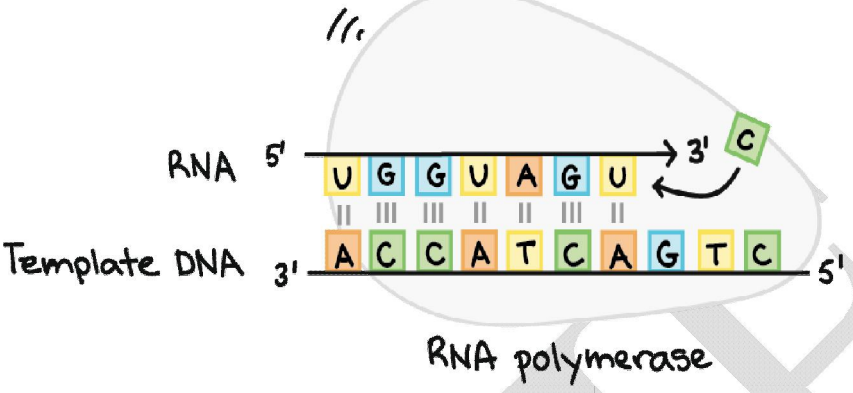
For a protein-coding gene, the RNA transcript contains the information needed to synthesize a polypeptide (protein or protein subunit) with a particular amino acid sequence. In this case:

RNA transcript (acting as messenger RNA): 5'-AUGAUCUCGUAA-3' Polypeptide: Met-Ile-Ser-STOP

RNA polymerase

The main enzyme involved in transcription is **RNA polymerase**, which uses a single-stranded DNA template to synthesize a complementary strand of RNA. Specifically, RNA

polymerase builds an RNA strand in the 5' to 3' direction, adding each new nucleotide to the 3' end of the strand.



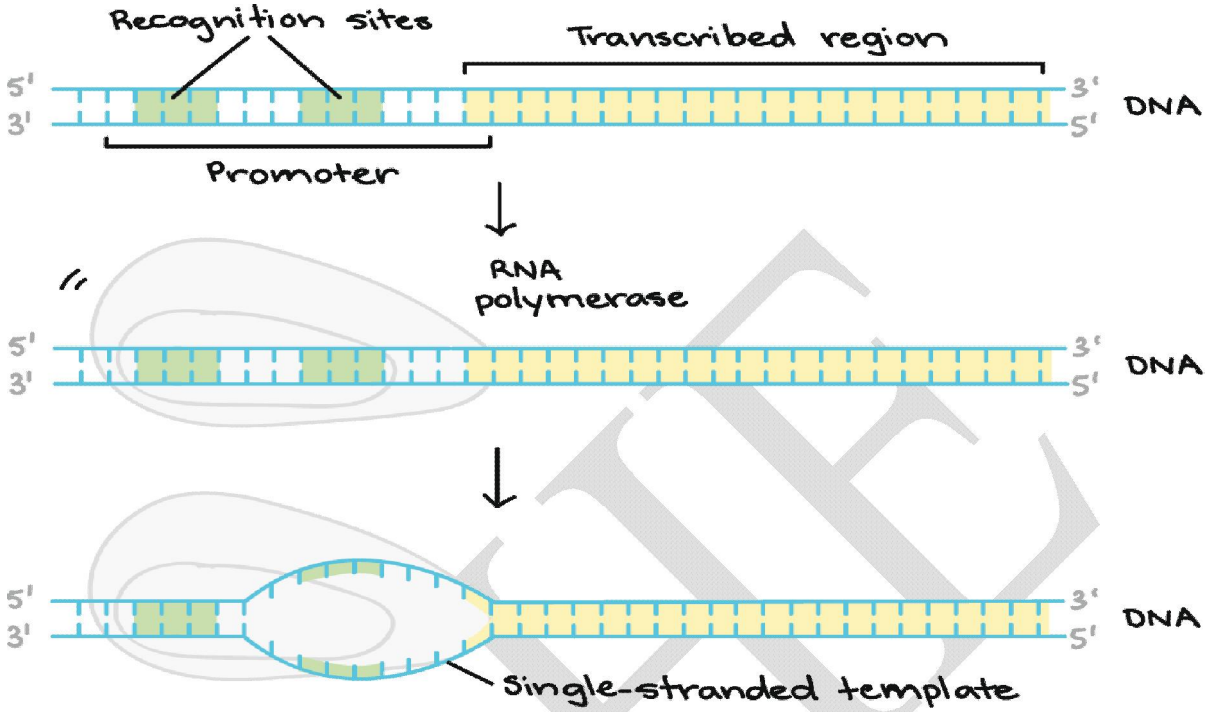
RNA polymerase synthesizes an RNA strand complementary to a template DNA strand. It synthesizes the RNA strand in the 5' to 3' direction, while reading the template DNA strand in the 3' to 5' direction. The template DNA strand and RNA strand are antiparallel.

RNA transcript: 5'-UGGUAGU...-3' (dots indicate where nucleotides are still being added at 3' end) DNA template: 3'-ACCATCAGTC-5'

Stages of transcription

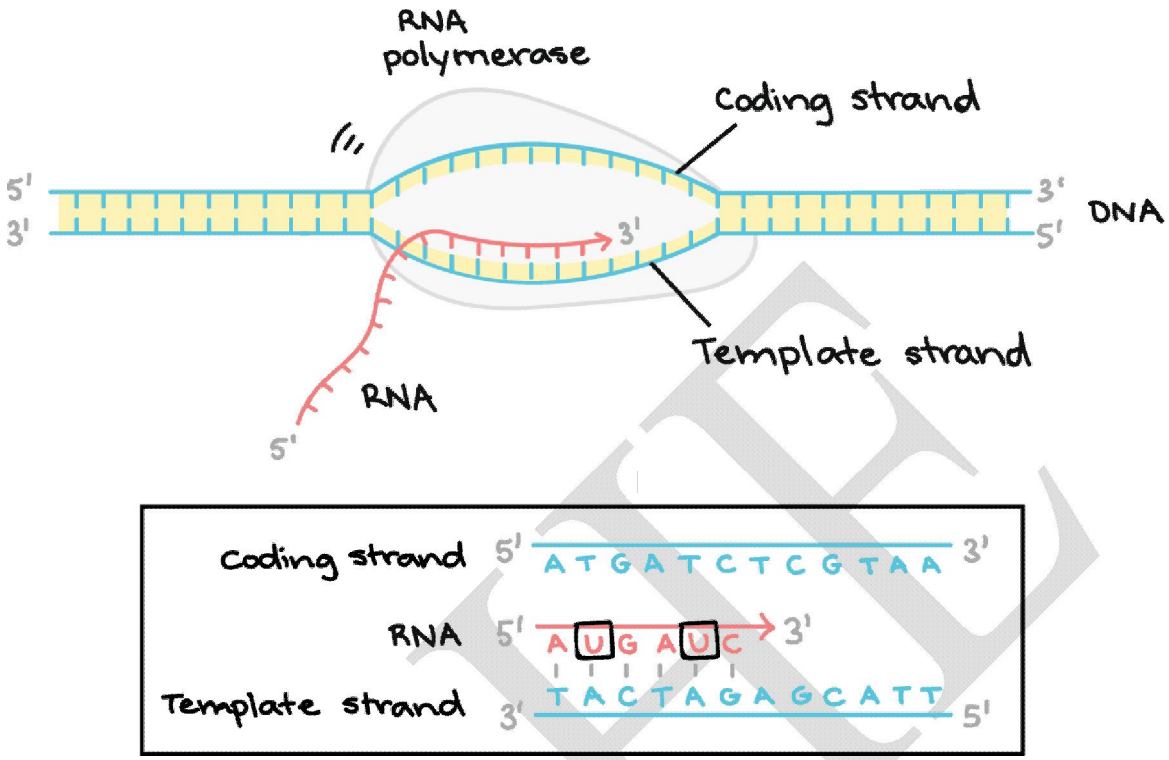
Transcription of a gene takes place in three stages: initiation, elongation, and termination. Here, we will briefly see how these steps happen in bacteria. You can learn more about the details of each stage (and about how eukaryotic transcription is different) in the stages of transcription article.

1. **Initiation.** RNA polymerase binds to a sequence of DNA called the **promoter**, found near the beginning of a gene. Each gene (or group of co-transcribed genes, in bacteria) has its own promoter. Once bound, RNA polymerase separates the DNA strands, providing the single-stranded template needed for transcription.



The promoter region comes before (and slightly overlaps with) the transcribed region whose transcription it specifies. It contains recognition sites for RNA polymerase or its helper proteins to bind to. The DNA opens up in the promoter region so that RNA polymerase can begin transcription.

2. **Elongation.** One strand of DNA, the **template strand**, acts as a template for RNA polymerase. As it "reads" this template one base at a time, the polymerase builds an RNA molecule out of complementary nucleotides, making a chain that grows from 5' to 3'. The RNA transcript carries the same information as the non-template (**coding**) strand of DNA, but it contains the base uracil (U) instead of thymine (T).



RNA polymerase synthesizes an RNA transcript complementary to the DNA template strand in the 5' to 3' direction. It moves forward along the template strand in the 3' to 5' direction, opening the DNA double helix as it goes. The synthesized RNA only remains bound to the template strand for a short while, then exits the polymerase as a dangling string, allowing the DNA to close back up and form a double helix.

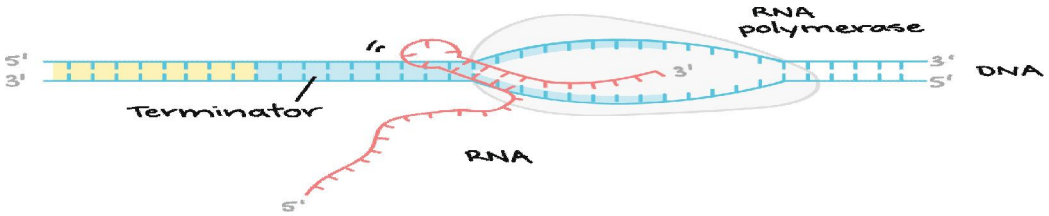
In this example, the sequences of the coding strand, template strand, and RNA transcript are:

Coding strand: 5' - ATGATCTCGTAA-3'

Template strand: 3'-TACTAGAGCATT-5'

RNA: 5'-AUGAUC...-3' (the dots indicate where nucleotides are still being added to the RNA strand at its 3' end)

3. **Termination.** Sequences called **terminators** signal that the RNA transcript is complete. Once they are transcribed, they cause the transcript to be released from the RNA polymerase. An example of a termination mechanism involving formation of a hairpin in the RNA is shown below.

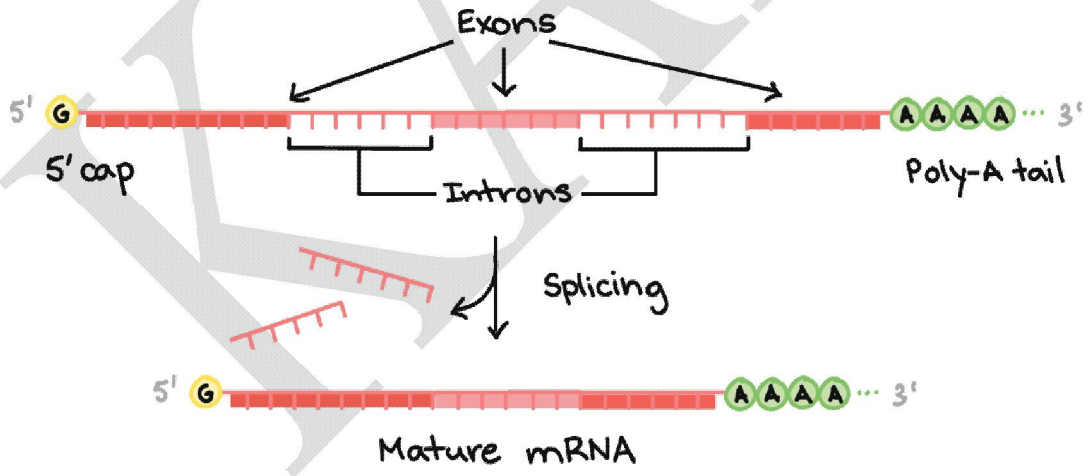


The terminator DNA encodes a region of RNA that forms a hairpin structure followed by a string of U nucleotides. The hairpin structure in the transcript causes the RNA polymerase to stall. The U nucleotides that come after the hairpin form weak bonds with the A nucleotides of the DNA template, allowing the transcript to separate from the template and ending transcription.

Eukaryotic RNA modifications

In bacteria, RNA transcripts can act as **messenger RNAs** (**mRNAs**) right away. In eukaryotes, the transcript of a protein-coding gene is called a **pre-mRNA** and must go through extra processing before it can direct translation.

- Eukaryotic pre-mRNAs must have their ends modified, by addition of a **5' cap** (at the beginning) and **3' poly-A tail** (at the end).
- Many eukaryotic pre-mRNAs undergo **splicing**. In this process, parts of the pre-mRNA (called **introns**) are chopped out, and the remaining pieces (called **exons**) are stuck back together.



Top of image: Diagram of a pre-mRNA with a 5' cap and 3' poly-A tail. The 5' cap is on the 5' end of the pre-mRNA and is a modified G nucleotide. The poly-A tail is on the 3' end of the pre-mRNA and consists of a long string of A nucleotides (only a few of which are shown).

The pre-mRNA still contains both exons and introns. Along the length of the mRNA, there is an alternating pattern of exons and introns: Exon 1 - Intron 1 - Exon 2 - Intron 2 - Exon 3. Each consists of a stretch of RNA nucleotides.

During splicing, the introns are removed from the pre-mRNA, and the exons are stuck together to form a mature mRNA.

Bottom of image: Mature mRNA that does not contain the intron sequences (Exon 1 - Exon 2 - Exon 3 only).

End modifications increase the stability of the mRNA, while splicing gives the mRNA its correct sequence. (If the introns are not removed, they'll be translated along with the exons, producing a "gibberish" polypeptide.)

To learn more about pre-mRNA modifications in eukaryotes, check out the article on pre-mRNA processing.

Transcription happens for individual genes

Not all genes are transcribed all the time. Instead, transcription is controlled individually for each gene (or, in bacteria, for small groups of genes that are transcribed together). Cells carefully regulate transcription, transcribing just the genes whose products are needed at a particular moment.

For example, the diagram below shows a "snapshot" of an imaginary cell's RNAs at a given moment in time. In this cell, genes 1, 2 and 3, are transcribed, while gene 4 is not. Also, genes 1, 2, and 3 are transcribed at different levels, meaning that different numbers of RNA molecules are made for each.

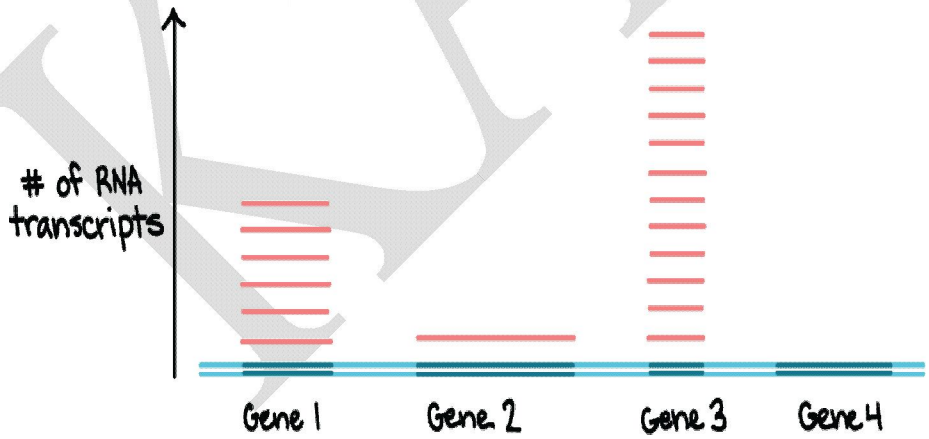


Diagram showing that individual genes are transcribed in different amounts.

A region of DNA containing four genes is shown, with the transcribed region of each gene highlighted in dark blue. The number of transcripts of each gene is indicated above the DNA

(on a Y- axis). There are six transcripts of gene 1, one transcript of gene 2, twelve transcripts of gene 3, and no transcripts of gene 4.

This is not an illustration of any actual set of genes and their transcription levels, but rather, illustrates that transcription is controlled individually for genes and other transcription units.

In the following articles, we'll take a more in-depth look at RNA polymerase, the stages of transcription, and the process of RNA modification in eukaryotes. We'll also consider some important differences between bacterial and eukaryotic transcription.

Split genes

An **interrupted gene** (also called a **split gene**) is a gene that contains sections of DNA called exons, which are expressed as RNA and protein, interrupted by sections of DNA called introns, which are not expressed.

The DNA sequence in the exon provides instructions for coding proteins. The function of the intron was not understood at first, and they were called noncoding or junk DNA. Split genes were independently discovered by Richard J. Roberts and Phillip A. Sharp in 1977, for which they shared the 1993 Nobel Prize in Physiology or Medicine Their discovery implied the existence of then-unknown machinery for splicing out introns and assembling genes; namely, the spliceosome. It was soon accepted that 94% of human genes were interrupted, and perhaps 50% of hereditary diseases involved errors in splicing introns out of interrupted genes. The best-known example of a disease caused by a splicing error is Beta-thalassemia, in which extra intronic material is erroneously spliced into the gene for making hemoglobin.

Lower eukaryotes, including yeast, have many *uninterrupted* regions, as they contain long stretches of exons that create the mRNA necessary for the synthesis of proteins. This does not mean, however, that these sections are fully uninterrupted, as tRNA synthesis requires excision of a nucleotide sequence, followed by ligation. Nevertheless, gene interruption is the rule.

Most bacteria have some interruption of some genes. Interrupted genes are universal in eukaryotes; yeasts may display single interruptions of a minority of genes, while in higher organisms most genes are interrupted, some multiple times and with introns that can be longer than exons. Introns are well-conserved across evolutionary history, suggesting their structure has some importance for the organism, and they are longer in advanced organisms (higher plants and animals), whose longer growth and development requires longer sequences of gene activation and down-regulation. Details of the role of introns in the regulation of gene accessibility and transcription have yet to be worked out.

The architecture of the interrupted gene allows for the process of alternative splicing, where various mRNA products can be produced from a single gene.

Concept of intron and exon

Introns and exons are nucleotide sequences within a gene. Introns are removed by RNA splicing as RNA matures, meaning that they are not expressed in the final messenger RNA

(mRNA) product, while exons go on to be covalently bonded to one another in order to create mature mRNA.

Introns can be considered as ***int***ervening sequences, and exons as ***ex***pressed sequences.

There are an average of 8.8 exons and 7.8 introns per human gene.

Exons

Exons are nucleotide sequences in DNA and RNA that are conserved in the creation of mature RNA. The process by which DNA is used as a template to create mRNA is called transcription.

mRNA then works in conjunction with ribosomes and transfer RNA (tRNA), both present in the cytoplasm, to create proteins in a process known as translation.

Exons usually include both the 5'- and 3'- untranslated regions of mRNA, which contain start and stop codons, in addition to any protein coding sequences.

Introns

Introns are nucleotide sequences in DNA and RNA that do not directly code for proteins, and are removed during the precursor messenger RNA (pre-mRNA) stage of maturation of mRNA by RNA splicing.

Introns can range in size from 10's of base pairs to 1000's of base pairs, and can be found in a wide variety of genes that generate RNA in most living organisms, including viruses.

Four distinct types of introns have been identified:

- Introns in protein coding genes, removed by spliceosomes
- Introns in tRNA genes, which are removed by proteins
- Self-splicing introns, which catalyse their own removal from mRNA, tRNA, and rRNA precursors using guanosine-5'-triphosphate (GTP), or another nucleotide cofactor (Group 1)
- Self-splicing introns, which do not require GTP in order to remove themselves (Group 2)

It is vital for the introns to be removed precisely, as any left-over intron nucleotides, or deletion of exon nucleotides, may result in a faulty protein being produced. This is because the amino acids that make up proteins are joined together based on codons, which consist of

three nucleotides. An imprecise intron removal thus may result in a frameshift, which means that the genetic code would be read incorrectly.

This can be explained by using the following phrase as a metaphor for an exon: "BOB THE BIG TAN CAT". If the intron before this exon was imprecisely removed, so that the "B" was no longer present, then the sequence would become unreadable: "OBT HEB IGT ANC AT…"

RNA splicing

RNA splicing is the method by which pre-mRNA is made into mature mRNA, by removal of introns and joining together of exons. Several methods of splicing exist, depending on the organism, type of RNA or intron structure, and the presence of catalysts.

Introns possess a highly conserved GU sequence at their 5' end, known as the donor site, and a highly conserved AG sequence at the 3' end, called the acceptor site. A large RNA-protein complex, the spliceosome, made up of five small nuclear ribonucleoproteins (snRNPs) recognise the start and end points of the intron thanks to these sites, and catalyse the removal of the intron accordingly. The spliceosome forms the intron into a loop that can be cleaved easily, and the remaining RNA on each side of the intron is connected. Other types of spliceosomes that recognise unusual or mutated intron sequences also exist, known as minor spliceosomes.

tRNA splicing is far rarer, though does occur in all three major domains of life, bacteria, archaea and eukarya. Multiple enzymes fill the role of snRNPs in a step-wise process, which can vary wildly between organisms.

Self-splicing introns are usually found in RNA molecules that are intended to catalyse biochemical reactions, ribozymes. Group 1 introns are attacked at the 5' splice site by a nucleotide cofactor, which may be free in the biological milieu or a part of the intron itself, leading to the 3'OH of the adjacent exon to become nucleophilic and thus bond to the 5' end of another exon, following the formation of the intron into a loop. Group 2 introns are spliced in a similar way, though with the use of a specific adenosine that attacks the 5' splice site.

Alternative Splicing

Alternative splicing refers to the way that different combinations of exons can be joined together, resulting in a single gene coding for multiple proteins. Walter Gilbert first put this idea forward, and he proposed that the different permutations of exons could produce different protein isoforms. These in turn would have different chemical and biological activities.

It is now thought that between 30 and 60% of human genes undergo alternative splicing. Moreover, over 60% of disease-causing mutations in humans are related to splice errors, rather than mistakes in coding sequences.

One example of a human gene that undergoes alternative splicing is fibronectin, a glycoprotein that extends from the cell into the extracellular matrix. Over 20 different isoforms of fibronectin have been discovered. These have all been produced from different combinations of fibronectin gene exons.

I. Capping

Besides splicing, the processing of mRNA contains 2 events: capping and polyadenylation. In capping, a special blocking nucleotide (a cap) is added to the 5'-end of a pre-mRNA. In polyadenylation, a string of AMPs (polyA) is added to the 3'-end of the pre-mRNA.

1. Cap structure:

Moss and others found caps in vaccinia virus mRNAs.

1) The capping substance is 7-methylguanosine (m^7G).

2) m^7G is linked to 2'-O-methyl-nucleotide by 5' to 5'.

3) This linkage is a triphosphate (a-phosphate of m^7G plus b- and a-phosphates of 2'-O-methylated nucleotide).

4) There are 3 types of cap:

a) cap0 (only in certain viral RNAs): has no 2'-O-methylated nucleotide.

b) cap1 (cellular and viral RNAs): has one 2'-O-methylated nucleotide.

c) cap2 (only in eukaryotic RNAs): has two 2'-O-methylated nucleotides.

2. Cap synthesis:

1) Sequence of capping (cap1) formation

a) A nucleotide phosphohydrolase (also called RNA triphosphatase) cleaves the g-phosphate off the triphosphate at the 5'-end of the growing RNA.

b) A guanylyl transferase attaches GMP from GTP to the diphosphate at the end of the RNA, forming the 5'-5' triphosphate linkage.

c) A methyl transferase adds the methyl group from S-adenosylmethionine (AdoMet) to the 7-nitrogen of the capping guanine.

d) A methyl transferase adds the methyl group from another AdoMet to the 2'-hydroxyl of the next nucleotide.

2) Moss and others showed pppY is a real intermediate in the capping scheme.

3) James Darnell and others found capping occurs early in the transcription process, before the chain length reaches 30 nts.

3. Functions of caps:

1) Protection

The cap may protect the mRNA from attack by RNases that begin at the 5'-end of their substrates and that cannot cleave triphosphate linkages. In 1977, Furuichi and others showed capped reovirus RNAs are much more stable than uncapped RNAs. They injected 3 kinds (capped with m^7GpppG, blocked with GpppG, and uncapped) of reovirus RNAs into *Xenopus* oocytes or wheat germ extract, and measured RNAs in 3 size classes (large, medium, and small).

2) Translatability

The cap can enhance the translatablitiy of mRNAs. An eukaryotic mRNA gains access to the ribosome for translation via a cap-binding protein that recognizes the cap.

Gallie demonstrated the stimulatory effect of the cap on translation by an *in vitro* assay of firefly luciferase mRNA into tobacco cells.

3) Transportation

The cap facilitates the transport of mRNAs from the nucleus into the cytoplasm.

4) Splicing efficiency

The cap may enhance the efficiency of splicing of mRNAs.

II. Polyadenylation

hnRNA and mRNA have a long chain of poly A, while rRNA and tRNA do not.

Polyadenylation: Addition of polyA to the 3'-end of an RNA.

1. Poly A:

The average size of fresh polyA is about 250 nucleotides. Polyadenylation is catalyzed by polyA polymerase (PAP) that adds AMP residues one at a time to mRNA precursors. Thus, actinomycin D does not inhibit polyadenylation.

2. Function of polyA:

Most mRNAs contain polyA, except histone mRNA.

1) Protection

PolyA helps protect mRNAs from degradation.

2) Translatability

PolyA binding protein I (PAB I) may boost the efficiency with which an mRNA is translated. Thus, excess polyA added to an *in vitro* reaction inhibits translation of capped and polyadenylated mRNAs, due to competition for PAB I.

3. Basic mechanism of polyadenylation:

A cleavage of an mRNA precursor before transcription has terminated, PAP adds polyA to the 3'-end of the mRNA, and the extra RNA lying beyond the polyadenylation site is degraded.

4. Polyadenylation signal:

An efficient polyadenyaltion signal has an AAUAAA motif about 20-30 nts upstream of a polyadenylation site in a pre-mRNA.

5. Cleavage and polyadenylation of a pre-mRNA:

1) Pre-mRNA cleavage

CPSF (cleavage and polyadenyaltion specificity factor) binds to the AAUAAA motif. CstF (cleavage stimulation factor) binds to the G/U region. CF I (cleavage factor I), CF II, and PAP itself are required for cleavage.

2) Initiation of polyadenylation

The first phase, slow addition of at least 10 As to the pre-mRNA. The second phase, elongation, is independent of the AAUAAA motif, involves the rapid addition of 200 or more As to the RNA.

3) Elongation

PolyA binding protein II (PAB II, 49 kDa) binds specificity to a preinitiated oligoA and helps PAP in elongating the polyA up to 250 nts or more. PAB II acts independently of the AAUAAA motif, and its activity is enhanced by CPSF.

6. PAP:

In 1991, Manley and others cloned bovine PAP, and found 2 different PAPs (PAP I and PAP II) that differ in their C-termini. PAP contains an RNA-binding domain (RBD), a polymerase module (PM), 2 nuclear localization signals (NLS 1 and NLS 2), and several Ser/Thr-rich regions (S/T).

7. Turnover of polyA:

Sheiness and Darnell found a slight difference in size between nuclear (210 ± 20 nts) and cytoplasmic (190 ± 20 nts) polyA. In the cytoplasm, polyA turns over by RNase degradation and PAP polyadenylation.

1) PolyA degradation in the cytoplasm

The shortening of polyA in the cytoplasm is deadenylation, carried out by polyA nuclease (PAN).

2) Cytoplasmic polyadenylation

Maturation of *Xenopus* oocytes occurs *in vitro* on stimulation by progesterone. The immature oocyte cytoplasm contains lots of maternal mRNAs which are almost fully deadenylated. Maturation-specific polyadenylation of Xenopus maternal mRNAs in the cytoplasm depends on the AAUAAA and UUUUUAU (cytoplasmic polyadenylation element, CPE) motifs.

III. The Effects of the Cap and PolyA on Splicing

1. Dependence of splicing on the cap:

The cap selectively enhances splicing of the pre-mRNA, while uncapped precursor inhibits splicing.

2. Effect of polyA on splicing:

Polyadenylation stimulates removal of the last intron in a pre-mRNA, but not upstream introns.

Ribosomal RNA (rRNA)

The four rRNAs in eukaryotes are first transcribed as two long precursor molecules. One contains just the pre-rRNA that will be processed into the 5S rRNA; the other spans the 28S, 5.8S, and 18S rRNAs. Enzymes then cleave the precursors into subunits corresponding to each rRNA. In bacteria, there are only three rRNAs and all are transcribed in one long precursor molecule that is cleaved into the individual rRNAs. Some of the bases of pre-rRNAs are methylated for added stability. Mature rRNAs make up 50-60% of each ribosome. Some of a ribosome's RNA molecules are purely structural, whereas others have catalytic or binding activities.

The eukaryotic ribosome is composed of two subunits: a large subunit (60S) and a small subunit (40S). The 60S subunit is composed of the 28S rRNA, 5.8S rRNA, 5S rRNA, and 50 proteins. The 40S subunit is composed of the 18S rRNA and 33 proteins. The bacterial ribosome is composed of two similar subunits, with slightly different components. The bacterial large subunit is called the 50S subunit and is composed of the 23S rRNA, 5S rRNA, and 31 proteins, while the bacterial small subunit is called the 30S subunit and is composed of the 16S rRNA and 21 proteins.

The two subunits join to constitute a functioning ribosome that is capable of creating proteins.

RNA interference

(RNAi) is a biological process in which RNA molecules inhibit gene expression or translation, by neutralizing targeted mRNAmolecules. Historically, RNAi was known by other names, including *co-suppression*, *post-transcriptional gene silencing* (PTGS), and *quelling*. The detailed study of each of these seemingly different processes, elucidated that the identity of these phenomena were all actually RNAi. Andrew Fire and Craig C. Mello shared the 2006 Nobel Prize in Physiology or Medicine for their work on RNA interference in the nematode worm *Caenorhabditis elegans*, which they published in 1998. Since the discovery of RNAi and its regulatory potentials, it has become evident that RNAi has immense potential in suppression of desired genes. RNAi is now known as precise, efficient, stable and better than antisense technology for gene suppression.[1] However, antisense RNA produced intracellularly by an expression vector may be developed and find utility as novel therapeutic agents.

Two types of small ribonucleic acid (RNA) molecules – microRNA (miRNA) and small interfering RNA (siRNA) – are central to RNA interference. RNAs are the direct products of genes, and these small RNAs can direct enzyme complexes to degrade messenger RNA (mRNA) molecules and thus decrease their activity by preventing translation, via post-transcriptional gene silencing. Moreover, transcription can be inhibited via the pre-transcriptional silencing mechanism of RNA interference, through which an enzyme complex catalyzes DNA methylation at genomic positions complementary to complexed siRNA or miRNA. RNA interference has an important role in defending cells against parasitic nucleotide sequences – viruses and transposons. It also influences development.

The RNAi pathway is found in many eukaryotes, including animals, and is initiated by the enzyme Dicer, which cleaves long double-stranded RNA (dsRNA) molecules into short double-stranded fragments of ~21 nucleotide siRNAs. Each siRNA is unwound into two single-stranded RNAs (ssRNAs), the passenger strand and the guide strand. The passenger strand is degraded and the guide strand is incorporated into the RNA-induced silencing complex (RISC). The most well-studied outcome is post-transcriptional gene silencing, which occurs when the guide strand pairs with a complementary sequence in a messenger RNA molecule and induces cleavage by Argonaute 2 (Ago2), the catalytic component of the RISC. In some organisms, this process spreads systemically, despite the initially limited molar concentrations of siRNA.

RNAi is a valuable research tool, both in cell culture and in living organisms, because synthetic dsRNA introduced into cells can selectively and robustly induce suppression of specific genes of interest. RNAi may be used for large-scale screens that systematically shut down each gene in the cell, which can help to identify the components necessary for a particular cellular process or an event such as cell division. The pathway is also used as a practical tool in biotechnology, medicine and insecticides.

siRNA vs.miRNA

Before you can understand the ways in which siRNA and miRNA are similar and how they're different, it helps to know just what they are. Both siRNA and miRNA are proteomics tools used to study various aspects of gene expression. Proteomics is the study of proteins by which a cell's complete complement of proteins is examined at once. Technological advances have made such study possible.

So are siRNA and miRNA similar or different The jury is still somewhat out on that question, depending on whom you ask. Some sources feel that siRNA and miRNA are the same things, while others indicate that they're separate entities entirely.

The disagreement comes about because the two are both formed in the same manner. They emerge from longer RNA precursors. They're also both processed in the cytoplasm by an enzyme called Dicer before becoming part of the protein complex RISC. Enzymes are proteins that can improve the rate of reaction between biomolecules.

There Are Slight Differences Between the Two

The process of RNA interference (RNAi) can be moderated by either siRNA or miRNA, and there are subtle differences between the two. As mentioned, both are processed inside the cell by the enzyme Dicer and incorporated into the complex RISC.

siRNA is considered exogenous double-stranded RNA that is taken up by cells. In other words, it enters through vectors, such as viruses. Vectors arise when geneticists use bits of DNA to clone a gene to produce a genetically modified organism (GMO). The DNA used in this process is called a vector.

Although siRNA is thought to be exogenous double-stranded RNA, miRNA is single-stranded. It comes from endogenous noncoding RNA, meaning that it's made inside the cell. This RNA is found within the introns of larger RNA molecules.

A Few Other Differences

Another difference between siRNA and miRNA is that siRNA typically binds perfectly to its mRNA target in animals. It's a perfect match to the sequence. In contrast, miRNA can inhibit translation of many different mRNA sequences because its pairing is imperfect. Translation occurs after messenger RNA is altered and binds to a particular site on a ribosome. In plants, miRNA tends to have a more perfectly complementary sequence, which induces mRNA cleavage as opposed to just repression of translation.

siRNA and miRNA can both play a role in epigenetics through a process called RNA-induced transcriptional silencing (RITS). Epigenetics is the study of heritable genetic information in which the nucleotide sequence of DNA is not altered but manifested as chemical marks. These marks are added to DNA or chromatin proteins after replication. Likewise, both are important targets for therapeutic use because of the roles they play in the controlling gene expression

POSSIBLE QUESTIONS

PART B

1) Define Transcription?
2) What is the difference between transcription and replication?
3) What are split genes?
4) What are RNA Splicing?
5) Define Polyadenylation and capping?

PART C

1) Explain in detail about the Promoter and the strength of RNA?
2) Explain in detail about the transcription in eukaryotes?
3) Write short notes on spliceosome machinery?
4) Write short notes on RNA interference: and its significance
5) Explain in detail about the general Transcription factors.

KARPAGAM ACADEMY OF HIGHER EDUCATION

CLASS: III B.Sc MB COURSE NAME: MOLECULAR BIOLOGY

COURSE CODE: 17MBU603B UNIT: III BATCH-2017-2020

Sl. No	Question	Option A	Option B	Option C	Option D	Correct Ans
1	Non codon specifies more than ___ amino acid	1	2	3	4	1
2	Molecular weight of egg lysozyme is	19300 daltons	13900 daltons	31900 daltons	91300 daltons	13900 daltons
3	*lac* operon is an example for	Repressible operon	Inducible operon	Mutated operon	Neutral operon	Inducible operon
4	In RNA, thiamine is replaced by	Uracil	Adenine	Cytosine	Guanine	Uracil
5	Genes are located in specialized structures called	Histone	RNA	Chrosomes	Genomes	Chrosomes
6	Enzyme activity is regulated by changes in the confirmation of enzymes except	Polymerase	Ribozymes	Chimozymes	Nuclease	Ribozymes
7	Codon/Anticodon consists of _____ nucleotides	4	6	3	9	3
8	Amino acid that have largest number of codons	Proline	Cysteine	Serine	Valine	Serine
9	7-methylguanosine cap is an important site	For eukaryotic transcription initiation factor	For prokaryotic translation initiation factor	For eukaryotic translation initiation factor	For prokaryotic translation initiation factor	For eukaryotic translation initiation factor
10	Capping in mRNA is addition of the group	7-ethylguanosine	7-methylguanosine	7-methylcytosine	7-ethylcytosine	7-methylguanosine
11	_____ is the first amino acid during translation of proteins	Threonine	Leucine	Methionine	Valine	Methionine
12	Control of gene expression was proposed by	Beedle & Tatum	Avery & McLeod	Jacob & Monad	Hershey & Chase	Jacob & Monad

13	In prokaryotes, AUG is translated in to	Methionine	N-acetyl-methionine	N-formamyl-aspargine	N-formamyl-methionine	N-formamyl-methionine
14	In trp operon, the genes *trp*E & *trp*D codes for	Arginase	Tryptophan synthase	Anthranilate isomerase	Anthranilate synthase	Anthranilate synthase
15	Enzyme that lactose in to glucose and galactose	Lactosidase	Glucanse	α-galactosidase	β-galactosidase	β-galactosidase
16	The number of nitrogenous bases codes by 9 amino acids would be	27	36	18	9	27
17	The termination of transcription is signaled by rich	AT containing inverted repeat	AC containing inverted repeat	GC containing inverted repeat	CT containing inverted repeat	GC containing inverted repeat
18	tRNA's are matched with their aminoacids by a group of enzymes collectively called as	Aminoacyl DNA synthatases	Aminoacyl synthatases	Amino synthatases	Aminoacyl tRNA synthatases	aminoacyl tRNA synthatases
19	Translation is	rRNA to protein	tRNA to protein	DNA to protein	mRNA to protein	mRNA to protein
20	Stop codons in mammalian mitochodria are	GAA & GAG	AGA & AGG	CGA & AGC	CGG & GCG	AGA & AGG
21	Ribosomal site most frequently occupied by the tRNA carrying the growing peptide chain	A-site	P-site	E-site	G-site	P-site
22	Non-coding regions are called as	Exons	Introns	Cistrons	Positrons	Exons
23	Mammalian mitochondrion not only uses AUG as initiation codon but also	AUA, AUU, AUC	UAA, UAU, UAC	AAU, UAU, CAU	GUA, GUU, GUC	AUA, AUU, AUC
24	Region that comprise the core prokaryotic promoter	Klenow box	Pribnow box	TAGTAG box	Polypeptide box	Pribnow box

25	All aminoacids have more than one codon except	Methionine & Tryptophan	Valine & Leucine	Threonine & Alanine	Lysine & Arginine	Methionine & Tryptophan
26	Site to which substrate molecules are attached	Catalytic site	Effector site	Alleosteric site	Binding site	Catalytic site
27	The main function of nonsense codons is to	Initiate protein synthesis	Elongate protein synthesis	Terminate protein synthesis	Regulate protein synthesis	Terminate protein synthesis
28	Which is astop codon	UAA	AAU	AUA	AAA	UAA
29	The promoter sequence in eukaryotes is	TATAAA	TAATAA	TTGACA	GTTAAA	TATAAA
30	Structure of proteins may be classified into ___ types	2	3	4	5	4
31	rRNA is also called	Rnase	Ribase	Ribulase	Ribozyme	Ribozyme
32	7-methylguanosine cap is an important site	For eukaryotic transcription initiation factor	For prokaryotic translation initiation factor	For eukaryotic translation initiation factor	For prokaryotic translation initiation factor	For eukaryotic translation initiation factor
33	2006 Nobel Prize in Physiology & Medicine for studies on molecular basis of eukaryotic transcription	Arthur Nirenberg	Roger D. Kornberg	David Osborne	Michael Whitney	Roger D. Kornberg
34	Transcription is	DNA to rRNA	DNA to tRNA	DNA to mRNA	DNA to protein	DNA to mRNA
35	tRNA is responsible for the transfering	Protein	Aminoacid	Codon	Anticodon	Anticodon
36	Monad & Cohen-Bazire first reported the evidence for the repression of the enzyme	Tryptophan synthase	Gluconase synthetase	Arabinase trimutase	Tryptophanase	Tryptophan synthase

37	A-site is the ribosomal site most frequently occupied by the	Aminoacyl-rRNA	Aminoacyl-mRNA	Iminoacyl-tRNA	Aminoacyl-tRNA	Aminoacyl-tRNA
38	Action of repressor protein in *Lac* operon is called as	Positive control	Negative control	Neutral control	No control	Negative control
39	Other than methionine _____ is the amino acid that appear rarely in proteins	Arginine	Tryptophan	Glutamic acid	Threonine	Tryptophan
40	In post translational modification of RNAs, trimming is	Removal of excess nucleotides	Removal of excess proteins	Removal of excess lipids	Removal of excess carbohydrates	Removal of excess nucleotides
41	Common method of covalent modification of enzyme in regulation of gene expression is	to methylate the enzyme at a proline residue	to phosphorylate the enzyme at a proline residue	to phosphorylate the enzyme at a serine residue	to methylate the enzyme at a serine residue	to phosphorylate the enzyme at a serine residue
42	Short sequence of aminoacids are called	Peptides	Proteins	Polypeptides	Palindromes	Peptides
43	In the absence of effector molecule, the enzyme is said to be in	Relaxed state	Tense state	Free state	Degrading state	Tense state
44	Allosteric enzymes that are controlled by a molecule other than it's substrate	Cohesive molecules	Systematic molecules	Effector molecules	Affector molecules	Segments of nucleotides sequences are repeated
45	The first and best example of control of gene expression was proposed by	Khorana & Nirenberg	Hershey & Chase	Avery & McLeod	Jacob and Monad	Jacob and Monad
46	Addition of poly A tail to 3' end of mRNA is mediated by the enzyme	RNA polymerase	DNA polymerase	Rnase	poly A polymerase	poly A polymerase
47	Confirmational changes in protein is brought about by	Systematic	Cohesive molecules	Affector molecules	Effector molecules	Effector molecules

		molecules				
48	In Rho-independent transcription termination, the termination sequence is usually	Palindromic sequence	Paliomic sequence	Panoramic sequence	Pandemic sequence	Palindromic sequence
49	Model example for gene regulation by repression	*trp* operon	*lac* operon	*ara* operon	*gal* operon	*trp* operon
50	Operon model that demonstrates both positive and negative control of gene regulation	*lac* operon	*ara* operon	*gal* operon	*trp* operon	*ara* operon
51	Repressor molecule in lac operon is a	Dimer	Trimer	Tetramer	Pentamer	Tetramer
52	Stop codon UAA is also called	Amber	Opal	Acre	Ochre	Ochre
53	The first codon during translation is	AGU	AUG	GUA	UGA	AUG
54	Transcription initiation site starts from	-1	Plus 1	-10	Plus 10	Plus 1
55	What are the possible number of codons that can be generated using possible nucleotide combinations	46	64	20	30	64
56	Which transports lactose across the cell membrane	Galactosidase permease	β-galactosidase	Glucanse	Glucose permease	Galactosidase permease
57	Who deciphered the genetic code	Hershey & Chase	Avery & McLeod	Beedle & Tatum	Nirenberg & Khorana	Nirenberg & Khorana
58	The stop codons are called as	Missense	Nonsense codons	Central codons	Last codons	Nonsense codons
59	The ability of the cell to choose between glucose and	Catabolic	Catabolic expression	Metabolic	Metabolic	Catabolic repression

	other sugars is termed as	repression		repression	expression	
60	Sequence of codons in mRNA between a start and a stop sequence is called as	Close reading frame	Open reading frame	Central reading frame	Last reading frame	Open reading frame

Unit IV

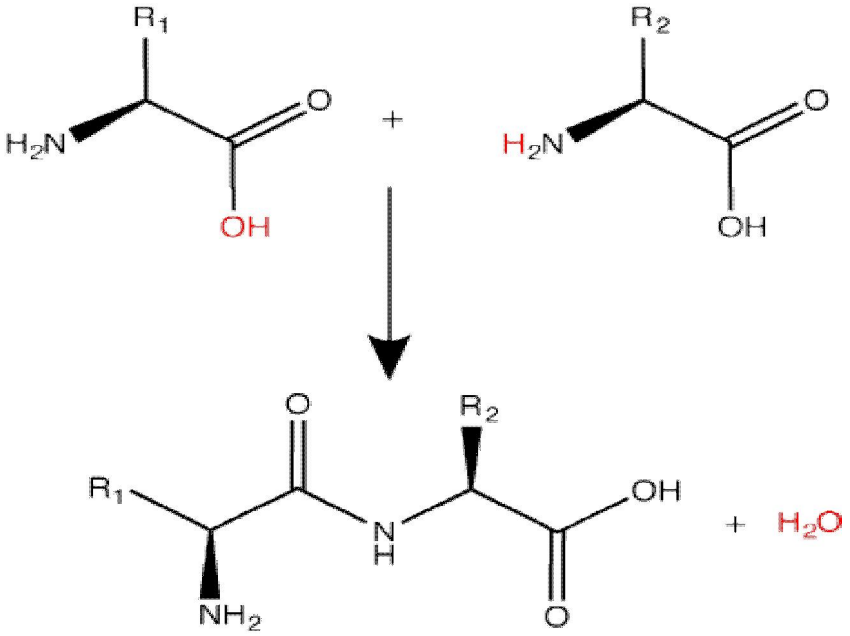
Translational machinery, Charging of tRNA, aminoacyl tRNA synthetases, Mechanisms of initiation, elongation and termination of polypeptides in both prokaryotes and eukaryotes, Fidelity of translation, Inhibitors of protein synthesis in prokaryotes and eukaryote

Protein Synthesis

Introduction

The process of translation in biology is the decoding an mRNA message into a polypeptide product. Put another way, a message written in the chemical language of nucleotides is "translated" into the chemical language of amino acids. Amino acids are linearly strung together via covalent bonds (called peptide bonds) between amino and carboxyl termini of adjacent amino acids. The decoding and "linking" process is catalyzed by a ribonucleoprotein complex called the ribosomes and can result in chains of amino acids of lengths ranging from tens to more than 1,000.

The resulting proteins are so important to the cell that their synthesis consumes more of a cell's energy than any other metabolic process. Like DNA replication and transcription, translation is a complex molecular process that we can approach using both the Energy Story and Design Challenge rubrics. Describing the overall process, or steps in the process, requires the accounting of the matter and energy before the process and after the process and a description of how that matter is transformed and energy transferred during the process. From a Design Challenge standpoint, we can - even before digging any further into what is or is not understood about translation - try to infer some of the basic questions that we will need to answer regarding this process.



A peptide bond links the carboxyl end of one amino acid with the amino end of another, expelling one water molecule. The R1 and R2 designation refer to side chain of amino acid the two amino acids.

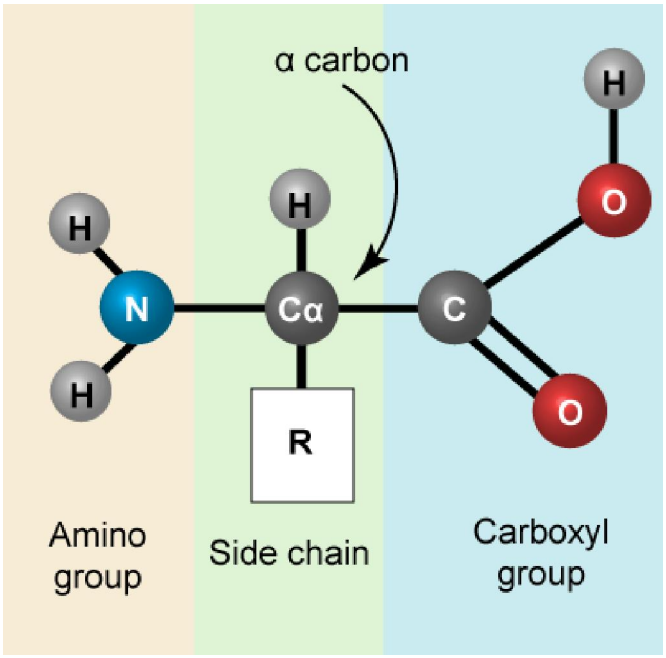
Protein Synthesis Machinery

The components that go into the process

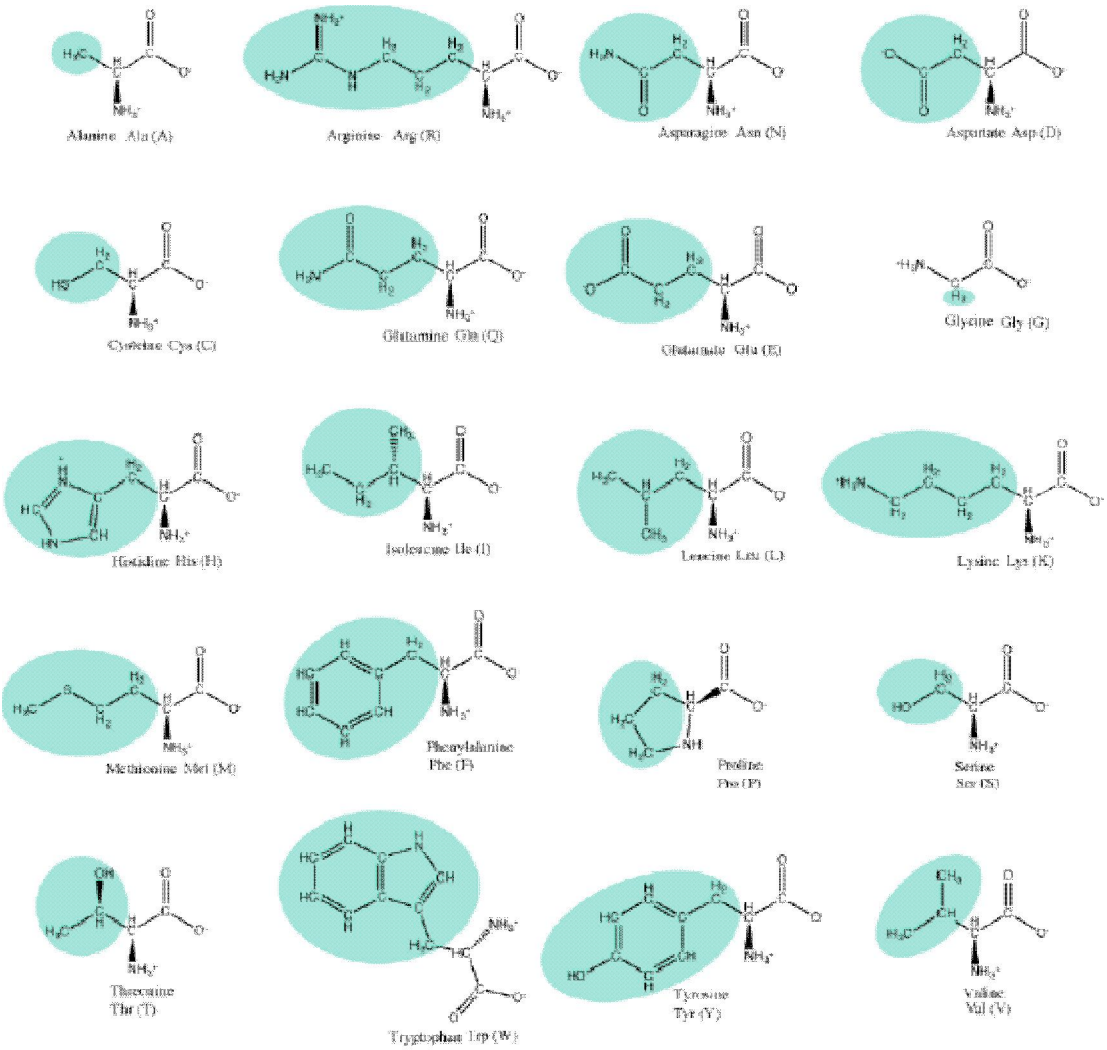
Many different molecules and macromolecules contribute to the process of translation. While the exact composition of "the players" in the process may vary from species to species - for instance, ribosomes may consist of different numbers of *rRNAs* (ribosomal RNAs) and polypeptides depending on the organism - the general functions of the protein synthesis machinery are comparable from bacteria to human cells. We focus on these similarities. At a minimum, translation requires an **mRNA template**, **amino acids**, **ribosomes**, **tRNAs**, an energy source, and various additional accessory enzymes and small molecules.

Reminder: Amino acids

Let us simply recall that the basic structure of amino acids is composed of a backbone composed of an amino group, a central carbon (called the α-carbon), and a carboxyl group. Attached to the α-carbon is a variable group that helps determine some of the chemical properties and reactivity of the amino acid.



A generic amino acid

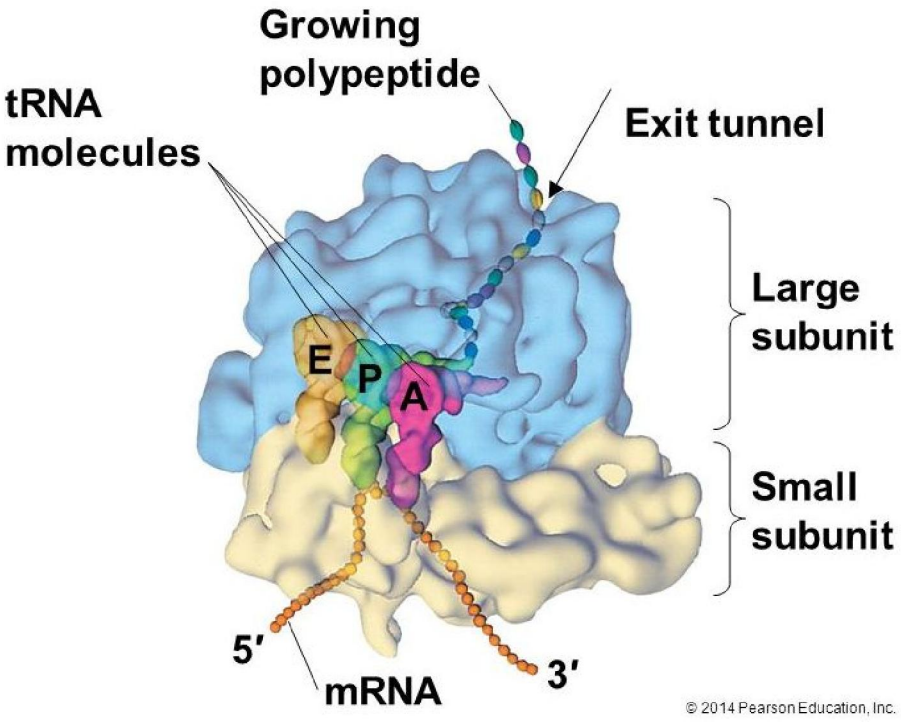


The 20 common amino acids

Ribosomes

A **ribosome** is a complex macromolecule composed of structural and catalytic rRNAs, and many distinct polypeptides. As we start to try thinking about energy accounting in the cell it is worth noting that ribosomes do not come "free". Even before an mRNA is translated, a cell must invest energy to build each of its ribosomes. In *E. coli*, there are between 10,000 and 70,000 ribosomes present in each cell at any given time.

Ribosomes exist in the cytoplasm in bacteria and archaea and in the cytoplasm and on the rough endoplasmic reticulum in eukaryotes. Mitochondria and chloroplasts also have their own ribosomes in the matrix and stroma, which look more similar to bacterial ribosomes (and have similar drug sensitivities), than the ribosomes just outside their outer membranes in the cytoplasm. Ribosomes dissociate into large and small subunits when they are not synthesizing proteins and reassociate during the initiation of translation. In *E. coli*, the small subunit is described as 30S, and the large subunit is 50S. Mammalian ribosomes have a small 40S subunit and a large 60S subunit. The small subunit is responsible for binding the mRNA template, whereas the large subunit sequentially binds tRNAs. Each mRNA molecule is simultaneously translated by many ribosomes, all synthesizing protein in the same direction: reading the mRNA from 5' to 3' and synthesizing the polypeptide from the N terminus to the C terminus. The complete mRNA/poly-ribosome structure is called a **polysome**.

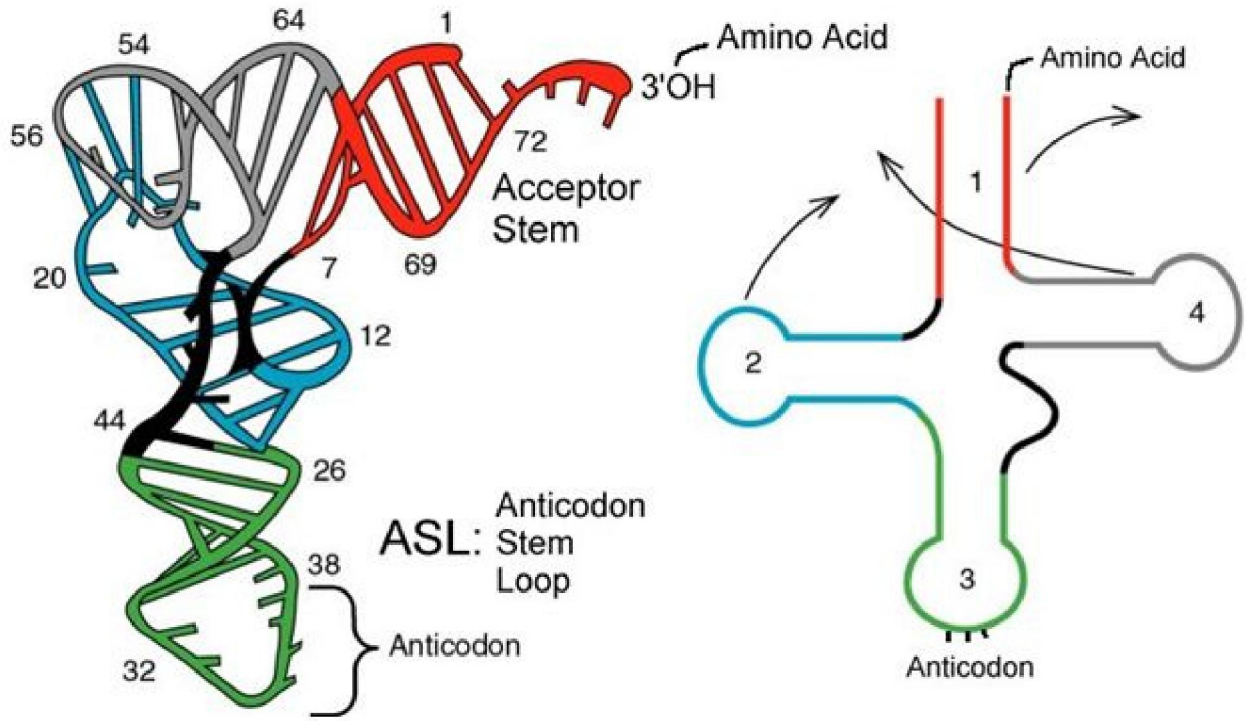


The protein synthesis machinery includes the large and small subunits of the ribosome, mRNA, and tRNA.

tRNAs

tRNAs are structural RNA molecules that were transcribed from genes. Depending on the species, 40 to 60 types of tRNAs exist in the cytoplasm. Serving as adaptors, specific tRNAs bind to sequences on the mRNA template and add the corresponding amino acid to the polypeptide chain. Therefore, tRNAs are the molecules that actually "translate" the language of RNA into the language of proteins.

Of the 64 possible mRNA **codons**—or triplet combinations of A, U, G, and C, three specify the termination of protein synthesis and 61 specify the addition of amino acids to the polypeptide chain. Of these 61, one codon (AUG) also encodes the initiation of translation. Each tRNA **anticodon** can base pair with one of the mRNA codons and add an amino acid or terminate translation, according to the genetic code. For instance, if the sequence CUA occurred on an mRNA template in the proper reading frame, it would bind a tRNA expressing the complementary sequence, GAU, which would be linked to the amino acid leucine.



The folded secondary structure of a tRNA. The anticodon loop and amino acid acceptor stem are indicated. Aminoacyl tRNA Synthetases

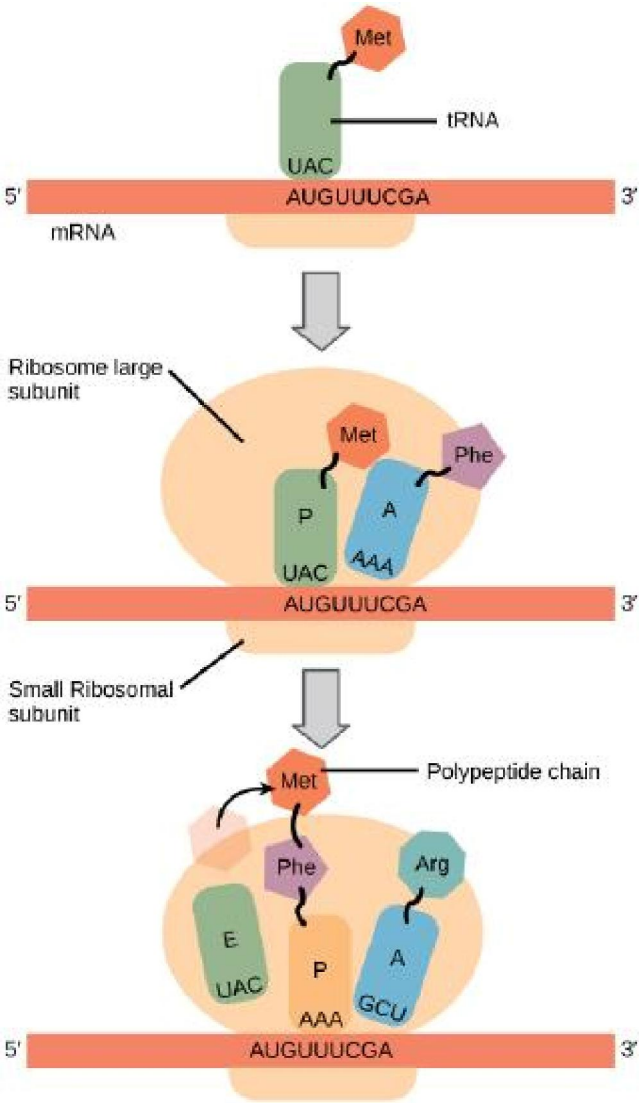
The process of pre-tRNA synthesis by RNA polymerase III only creates the RNA portion of the adaptor molecule. The corresponding amino acid must be added later, once the tRNA is processed and exported to the cytoplasm. Through the process of tRNA "charging," each tRNA molecule is linked to its correct amino acid by a group of enzymes called **aminoacyl tRNA synthetases**. At least one type of aminoacyl tRNA synthetase exists for each of the 20 amino acids; the exact number of aminoacyl tRNA synthetases varies by species. These enzymes first bind and hydrolyze ATP to catalyze a high-energy bond between an amino acid and adenosine monophosphate (AMP); a pyrophosphate molecule is expelled in this reaction. The activated amino acid is then transferred to the tRNA, and AMP is released.

The Mechanism of Protein Synthesis

Just as with mRNA synthesis, protein synthesis can be divided into three phases: initiation, elongation, and termination. The process of translation is similar in bacteria, archaea and eukaryotes.

Translation Initiation

In general, protein synthesis begins with the formation of an initiation complex. The small ribosomal subunit will bind to the mRNA at the **ribosomal binding site**. Soon after, the methionine-tRNA will bind to the AUG start codon (through complementary binding with its anticodon). This complex is then joined by large ribosomal subunit. This initiation complex then recruits the second tRNA and thus translation begins.



Translation begins when a tRNA anticodon recognizes a codon on the mRNA. The large ribosomal subunit joins the small subunit, and a second tRNA is recruited. As the mRNA moves relative to the ribosome, the polypeptide chain is formed. Entry of a release factor into the A site terminates translation and the components dissociate. Bacterial vs Eukaryotic initiation

In *E. coli* mRNA, a sequence upstream of the first AUG codon, called the **Shine-Dalgarno sequence** (AGGAGG), interacts with a rRNA molecule. This interaction anchors the 30S ribosomal subunit at the correct location on the mRNA template. Stop for a moment to appreciate the repetition of a mechanism you've encountered before. In this case, getting a

protein complex to associate - in proper register - with a nucleic acid polymer is accomplished by aligning two antiparallel strands of complementary nucleotides with one another. We also saw this in the function of telomerase.

Instead of binding at the Shine-Dalgarno sequence, the eukaryotic initiation complex recognizes the 7-methylguanosine cap at the 5' end of the mRNA. A cap-binding protein (CBP) assists the movement of the ribosome to the 5' cap. Once at the cap, the initiation complex tracks along the mRNA in the 5' to 3' direction, searching for the AUG start codon. Many eukaryotic mRNAs are translated from the first AUG, but this is not always the case. According to **Kozak's rules**, the nucleotides around the AUG indicate whether it is the correct start codon. Kozak's rules state that the following consensus sequence must appear around the AUG of vertebrate genes: 5'-gccRccAUGG-3'. The R (for purine) indicates a site that can be either A or G, but cannot be C or U. Essentially, the closer the sequence is to this consensus, the higher the efficiency of translation.

Translation Elongation

During translation elongation, the mRNA template provides specificity. As the ribosome moves along the mRNA, each mRNA codon comes into 'view', and specific binding with the corresponding charged tRNA anticodon is ensured. If mRNA were not present in the elongation complex, the ribosome would bind tRNAs nonspecifically. Note again the use of base pairing between two antiparallel strands of complementary nucleotides to bring and keep our molecular machine in register and in this case also to accomplish the job of "translating" between the language of nucleotides and amino acids.

The large ribosomal subunit consists of three compartments: the A site binds incoming charged tRNAs (tRNAs with their attached specific amino acids), the P site binds charged tRNAs carrying amino acids that have formed bonds with the growing polypeptide chain but have not yet dissociated from their corresponding tRNA, and the E site which releases dissociated tRNAs so they can be recharged with another free amino acid.

Elongation proceeds with charged tRNAs entering the A site and then shifting to the P site followed by the E site with each single-codon "step" of the ribosome. Ribosomal steps are induced by conformational changes that advance the ribosome by three bases in the 3'

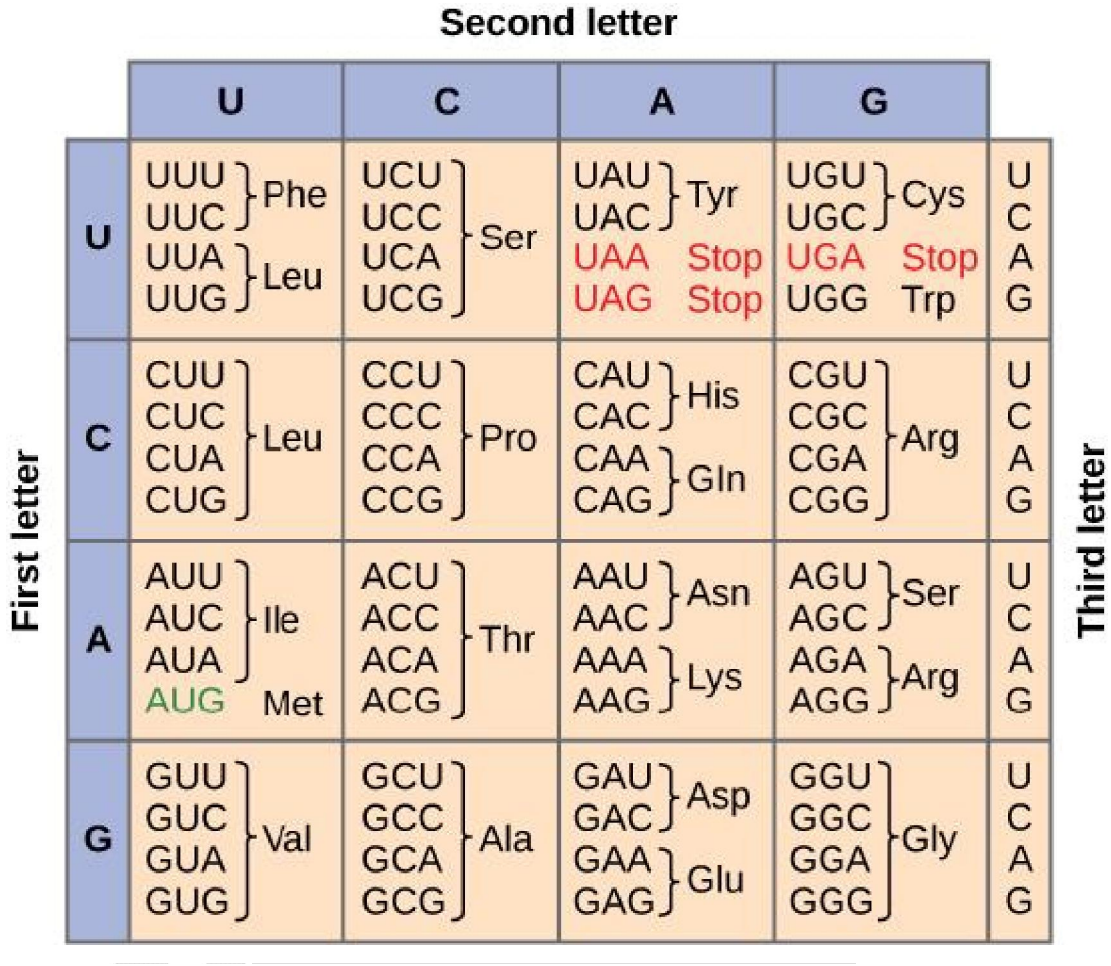
direction. The energy for each step of the ribosome is donated by an elongation factor that hydrolyzes GTP. Peptide bonds form between the amino group of the amino acid attached to the A-site tRNA and the carboxyl group of the amino acid attached to the P-site tRNA. The formation of each peptide bond is catalyzed by **peptidyl transferase**, an RNA-based enzyme that is integrated into the 50S ribosomal subunit. The energy for each peptide bond formation is derived from GTP hydrolysis, which is catalyzed by a separate elongation factor. The amino acid bound to the P-site tRNA is also linked to the growing polypeptide chain. As the ribosome steps across the mRNA, the former P-site tRNA enters the E site, detaches from the amino acid, and is expelled. The ribosome moves along the mRNA, one codon at a time, catalyzing each process that occurs in the three sites. With each step, a charged tRNA enters the complex, the polypeptide becomes one amino acid longer, and an uncharged tRNA departs. Amazingly, this process occurs rapidly in the cell, the *E. coli* translation apparatus takes only 0.05 seconds to add each amino acid, meaning that a 200-amino acid polypeptide could be translated in just 10 seconds.

The Genetic Code

To summarize what we know to this point, the cellular process of transcription generates messenger RNA (mRNA), a mobile molecular copy of one or more genes with an alphabet of A, C, G, and uracil (U). Translation of the mRNA template converts nucleotide-based genetic information into a protein product. Protein sequences consist of 20 commonly occurring amino acids; therefore, it can be said that the protein alphabet consists of 20 letters. Each amino acid is defined by a three-nucleotide sequence called the triplet **codon**. The relationship between a nucleotide codon and its corresponding amino acid is called the **genetic code**. Given the different numbers of "letters" in the mRNA and protein "alphabets," means that there are a total of 64 (4 × 4 × 4) possible codons; therefore, a given amino acid (20 total) must be encoded for by more than one codon.

Three of the 64 codons terminate protein synthesis and release the polypeptide from the translation machinery. These triplets are called **stop codons**. Another codon, AUG, also has a special function. In addition to specifying the amino acid methionine, it also serves as the **start codon** to initiate translation. The reading frame for translation is set by the AUG start codon near the 5' end of the mRNA. The genetic code is universal. With a few

exceptions, virtually all species use the same genetic code for protein synthesis, which is powerful evidence that all life on Earth shares a common origin.



This figure shows the genetic code for translating each nucleotide triplet, or codon, in mRNA into an amino acid or a termination signal in a nascent protein.

Redundant, not Ambiguous

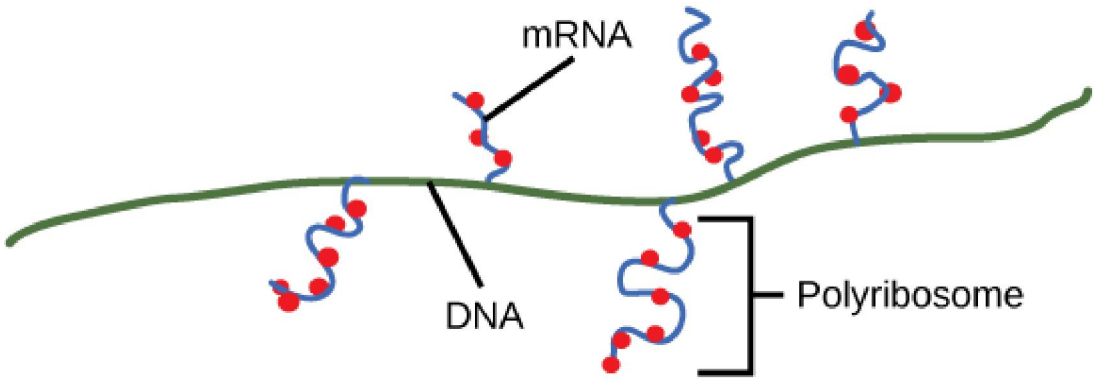
The information in the genetic code is redundant. Multiple codons code for the same amino acid. For example, using the chart above, you can find 4 different codons that code for Valine, likewise, there are two codons that code for Leucine, etc. But the code is not ambiguous, meaning, that if you were given a codon you would know definitively which amino acid it is coding for, a codon will only code for a specific amino acid. For example, GUU will always code for Valine, and AUG will always code for Methionine. This is

important, you will be asked to translate an mRNA into a protein using a codon chart like the one shown above.

Translation Termination

Termination of translation occurs when a stop codon (UAA, UAG, or UGA) is encountered. When the ribosome encounters the stop codon no tRNA enters into the A site. Instead a protein know as a **release factor** binds to the complex. This interaction destabilizes the translation machinery, causing the release of the polypeptide and the dissociation of the ribosome subunits from the mRNA. After many ribosomes have completed translation, the mRNA is degraded so the nucleotides can be reused in another transcription reaction.

As discussed previously, bacteria and archaea do not need to transport their RNA transcripts between a membrane bound nucleous and the cytoplasm. The RNA polymerase is therefore transcribing RNA directly into the cytoplasm. Here ribosomes can bind to the RNA and begin the process of translation, in some instances while transciption is still occurring. The coupling of these two processes, and even mRNA degradation, is facilitated not only because transcription and translation happen in the same compartment but also because both of the processes happen in the same direction - synthesis of the RNA transcript happens in the 5' to 3' direction and translation reads the transcript in the 5' to 3' direction. This "coupling" of transcription with translation occurs in both bacteria and archaea and is, in fact, essential for proper gene expression in some instances.



Multiple polymerases can transcribe a single bacterial gene while numerous ribosomes concurrently translate the mRNA transcripts into polypeptides. In this way, a specific protein can rapidly reach a high concentration in the bacterial cell.

Protein Sorting

In context of a protein synthesis Design Challenge we can also raise the question/problem of how proteins get to where they are supposed to go. We know that some proteins are destined for the plasma membrane, others in eukaryotic cells need to be directed to various organelles, some proteins, like hormones or nutrient scavenging proteins, are intended to be secreted by cells while others may need to be directed to parts of the cytosol to serve structural roles. How does this happen?

Since various mechanisms have been uncovered, the details of this process are not easily summarized in a brief paragraph or two. However, some key common elements of all mechanisms can be mentioned. First, is the need for a specific "tag" that can provide some molecular information about where the protein of interest is destined. This tag usually takes the form of a short string of amino acids - a so called signal peptide - that can encode information about where the protein is intended to end up. The second required component of the protein sorting machinery must be a system to actually read and sort the proteins. In bacterial and archaeal systems this usually consists of proteins that can identify the signal peptide during translation, bind to it, and direct the synthesis of the nascent protein to the plasma membrane. In eukaryotic systems, the sorting is by necessity more complex, and involves a rather elaborate set of mechanisms of signal recognition, protein modification, and trafficking of vesicles between organelles or the membrane. These biochemical steps are initiated in the endoplasmic reticulum and further "refined" in the Golgi apparatus where proteins are modified and packaged into vesicles bound for various parts of the cell.

Some of the various specific mechanisms may be discussed by your instructor in class. The key for all students it so appreciate the problem and to have a general idea of the high-level requirements that cells have adopted to solve them.

Post-translational Protein Modification

After translation individual amino acids may be chemically modified. These modifications add chemical variation and new properties that are rooted in the chemistries of the functional

groups that are being added. Common modifications include phosphate groups, methyl, acetate, and amide groups. Some proteins, typically targeted to membranes will be lipidated - a lipid will be added. Other proteins will be glycosylated - a sugar will be added. Another common post-translational modification is cleavage or linking of parts of the protein itself. Signal-peptides may be cleaved, parts may be excised from the middle of the protein, or new covalent linkages may be made between cysteine or other amino acid side chains. Nearly all modifications will be catalyzed by enzymes and all change the functional behavior of the protein.

Inhibitors of protein synthesis in prokaryotes and eukaryote

Many substances are known to act as inhibitors of various stages of protein synthesis. Included among these are a number of antibiotics produced by one strain of microorganism and lethal to other strains of the same or a different species.

Some of the best understood inhibitors of protein synthesis are listed in Table 22-10. Because the actions of many of these inhibitors are quite specific, they have proved to be extremely useful tools in the step-by-step elucidation of the mechanism of protein synthesis.

TABLE 22–10 INHIBITORS OF PROTEIN SYNTHESIS IN PROKARYOTIC AND EUKARYOTIC CELLS

Inhibitor	Effective in	
	Prokaryotes	Eukaryotes
Anisomycin	–	+
Aurintricarboxylic acid	+	+
Chloramphenicol	+	–
Colicin E3	+	–
Cycloheximide	–	+
Diphtheria toxin	–	+
Edeine	+	+
Erythromycin	+	–
Fusidic acid	+	+
Pactamycin	–	+
Puromycin	+	+
Ricin	–	+
Sodium fluoride	–	+
Sparsomycin	–	+
Streptomycin	+	–
Tetracycline	+	+
Trichodermin	–	+

Inhibitors of both Prokaryotic and Eukaryotic Protein Synthesis:

Aurintricarboocylic acid inhibits formation of the initiation complex by preventing the association of mRNA with the small ribosomal subunit. Inhibitors of initiation are readily

distinguished from inhibitor's blocking other stages of protein synthesis because of the delay effect that follows their administration, that is, protein synthesis continues for a short time after administration of the inhibitor, because peptide chains whose growth had already begun are unaffected and grow to completion.

Edeine, a polypeptide isolated from Bacillus brevis, inhibits the binding of aminoacyl-tRNA and N-formylmet-tRNA$^{Met}_{f}$ (in prokaryotes) to the small subunit. Fusidic acid is a steroidal antibiotic; in prokaryotes, it inhibits the binding of aminoacyl-tRNA to the ribosome, whereas in eukaryotes, it inhibits translocation by reacting with elongation factor. Puromycin was one of the first inhibitors of protein synthesis to have its specific effect determined. This antibiotic mimics aminoacyl-tRNA and binds to the free A site of ribosomes engaged in protein synthesis. Catalytic formation of a bond between the nascent polypeptide and puromycin is followed by the release of the peptidyl-puromycin from the ribosome, as no further elongation is possible.

The specific effects of puromycin have been used to advantage for studies of nascent chain length, the kinetics of chain elongation, and the identification of the effects of other antibiotics. Tetracycline inhibits protein synthesis by blocking aminoacyl-tRNA binding to the small subunit.

Inhibitors Specific for Prokaryotes

Chloramphenicol (Chloromycetin) binds to the large subunit of prokaryotic ribosomes and interferes with the functioning of peptide synthetase, thereby inhibiting chain elongation. Colicin E3 inhibits protein synthesis in prokaryotes by interfering in some manner with the functioning of the small subunit. Erythromycin binds to ribosomes that are not engaged in protein synthesis, preventing their potential participation, but does not bind to ribosomes containing nascent chains (i.e., ribosomes that are part of a functioning polysome). Streptomycin was one of the earliest discovered antibiotics and was employed as an agent against bacterial infection for many years before its specific chemical actions were known. Streptomycin binds to protein S12 of the small ribosome subunit, causing release of N-formylmet-tRNA$^{Met}_{f}$ from initiation complexes (thereby preventing initiation of chain growth) and also causing misreading of the codons of mRNA by ribosomes already involved in chain elongation.

Inhibitors Specific for Eukaryotes:

Anisomycin is an antibiotic produced by Streptomyces that inhibits peptide bond formation when bound to the small ribosomal subunit. Cycloheximide binds to the large subunit, preventing the translocation of tRNA in the A site to the P site. Diphtheria toxin (produced by a strain of Corynebacterium diphtherial) inhibits protein synthesis through its action on EF-2 (translocase). EF-2 exists in cells in two forms— ribosome-bound and free. Diphtheria toxin acts enzymatically to alter free EF-2, rendering the factor inactive. Ribosome-bound EF-2 is not susceptible to inactivation by the toxin.

Pactamydn (produced by a strain of Streptomyces) binds to free small subunits (not to small subunits already part of polysomes), where it prevents initiation by inhibiting binding of met-tRNA j and formation of the initiation complex. The toxic effects of ricin (a protein present in the castor bean) have been known for nearly a century.

Ricin consists of two polypeptide chains (linked by disulfide bridges), one of which acts as the inhibitor once incorporated into the cell. Ricin acts on the large subunit, preventing formation of the 80 S initiation complex. Like ricin, sodium fluoride acts as an inhibitor or initiation; NaF blocks addition of the large subunit to mRNA.

Sparsomycin, another antibiotic produced by Strep- tomyces, inhibits the association of the amino acid moiety of aminoacyl-tRNA from binding to the large sub- unit and, in so doing, blocks peptide synthetase. THchodermin is the only chemical compound so far identified as a specific inhibitor of the termination stage of polypeptide synthesis.

Inhibitors of Organeiiar Protein Synthesis:

Protein synthesis by mitochondrial and chloroplast ribosomes is also subject to inhibition by certain antibiotics and other chemicals. Shortly after the initial demonstration of organellar protein synthesis, it was found that chloramphenicol, a strong inhibitor of prokaryote protein synthesis, blocks synthesis in mitochondria and chloroplasts, whereas cycloheximide, which blocks eukaryote cytoplasmic ribosomal protein synthesis, is without effect on mitochondrial and chloroplast synthesis.

These observations provided added credence for the notion that prokaryotic cells, mitochondria, and chloroplasts have a common evolutionary origin. It is now clear, however,

that the picture is considerably more complex. For example, streptomycin, which inhibits prokaryotic protein synthesis, fails to inhibit mitochondrial protein synthesis in yeast cells. Other antibiotics inhibit mitochondrial protein synthesis but have no effect on prokaryotes. Erythromycin inhibits the synthesis of proteins in prokaryotes, yeast mitochondria, and chloroplasts but fails to block protein synthesis in mammaliam mitochondria.

The latter observation suggests that the nature of mitochondrial protein synthesis varies among different groups of eukaryotes. In general, mitochondria from higher eukaryotes are more resistant to inhibitors of prokaryotic protein synthesis than are mitochondria from lower eukaryotes. In chloroplasts, protein synthesis is inhibited by the same agents that block this process in prokaryotic cells.

The differential sensitivity of eukaryotic cytoplasmic and mitochondrial ribosomes to specific inhibitors provides a means for examining the sources of certain mitochondrial proteins. The synthesis of a mitochondrial protein in the presence of cycloheximide (a cytoplasmic inhibitor) indicates that the mitochondria are the source of the protein, whereas synthesis of the protein in the presence of chloramphenicol indicates that the mitochondrial protein is produced in the cytoplasm and then moves to the mitochondria.

Experimentally, determinations of this sort are carried out by incubating cells in media containing both radioactively labeled amino acids and inhibitor. The synthesis of the mitochondrial protein is manifested by the appearance of radioactivity in the protein later isolated from the mitochondria.

Using this approach, it has been possible to show that perhaps 85 to 90% of all mitochondrial ribosomal proteins are synthesized in the cytoplasm and then are imported into the mitochondria, where, together with mitochondrial rRNA, they are assembled into ribosomes. Of the seven polypeptides that make up the enzyme cytochrome oxidase, four are synthesized in the cytoplasm and three are synthesized in the mitochondria.

POSSIBLE QUESTIONS

Part – B

1. Define Translational,
2. What is meant by Charging of tRNA,
3. Define aminoacyl tRNA synthetases,
4. What are the steps in translation?

5. Differentiate prokaryotic and eukaryotic translation

Part – C

1. Write in detail about translational machinery
2. Write a note on Charging of tRNA, aminoacyl tRNA synthetases
3. Describe about the mechanisms of initiation, elongation and termination of polypeptides in eukaryotes
4. Describe about mechanisms of initiation, elongation and termination of polypeptides in prokaryote
5. Write in detail about the inhibitors of protein synthesis in prokaryotes and eukaryote

KARPAGAM ACADEMY OF HIGHER EDUCATION

CLASS: III B.Sc MB COURSE NAME: MOLECULAR BIOLOGY

COURSE CODE: 17MBU603B UNIT: IV BATCH-2017-2020

Sl. No	Question	Option A	Option B	Option C	Option D	Correct Ans
1	The phenomenon of linkage was first observed in the plant	*Lathyrus odoratus*	*Pisum sativum*	*Datura*	*Mirabilus jalapa*	*Lathyrus odoratus*
2	The Competence of a cell in the process of transformation is aided by	$CaCl_2$	$MgCl_2$	KCl	$AgCl_2$	$CaCl_2$
3	The viral genome integrated to the bacterial genome is called	Plasmid	Capsid	Prophage	Virion	Prophage
4	Pneumococcal 'S' cells produce _____ colonies during growth on agar plates	Smooth	Rough	Elongated	Flat	Smooth
5	In conjugation, the donor always carries on	(F-)	(F+)	F neutral	No F	(F+)
6	Occurs when new DNA does not integrate into the chromosome, not replicated and is eventually lost	Abortive transduction	Specialized Transduction	Generalized Transduction	Transfusion	Abortive transduction
7	Tendency of alleles located close together on a chromosome to be inherited together during the meiosis	Linkage	Crossing over	Gene overlapping	Recessive genes	Linkage
8	Virulence in *Streptococcus pneumoniae* is attributed to	Flagella	Capsules	Pili	Fimbriae	Capsules
9	The phenomenon in which genes are present on the same chromosomes is	Cross over	Segregation	Linkage	Assortment	Linkage
10	Specialised transduction is effected by _____	T_2	Mu	P_1	T_7	T_2

11	Genetic recombination in phages was discovered by _____	Hershey and Rotma	Hershey and Chase	Hershey and Wollmer	Hershey and Singer	Hershey and Rotma
12	Conjugation can only occur between cells of	F positive types	F negative types	Same mating types	Opposite mating types	Opposite mating types
13	An example for specialized transducing particle	No infection	T2 phage infects *Staphylococcus*	Phage P22 infects *Salmonella typhimurium*	Phage lambda infects *E.coli*	Phage lambda infects *E.coli*
14	_____ involves finding a contiguous series of cloned DNA fragments which contain overlapping portions of the genome	Physical mapping	Chemical mapping	Marker mapping	Loci mapping	Physical mapping
15	_____ refers to a genetic changes in different genomes of same cell.	Trans type	Cis type	Same type	Different type	Cis type
16	Conditions that favor the termination of the lysogenic state	Desiccation	Decomposition	Nutrient Media	Macronutrient	Desiccation
17	Direct way of observing the physical arrangement of markers along the chromosomes	Fluorescence in situ hybridization	Fluorescence invitro hybridization	Fluorescence invivo hybridization	Fluorescence in cell hybridization	Fluorescence in situ hybridization
18	Metalloproteins found in all eukaryotes	Zinc fingers	Iron fingers	Lead fingers	Copper fingers	Zinc fingers
19	The gene linkage minimize the chances of	Cross over	Segregation	Recombination	Assortment	Recombination
20	Transfer of DNA from one bacterium to another through the action of viruses	Transduction	Conjugation	Transformation	Gene expression	Transduction

21	The frequency at which two genes are - - - - - - - - by population of phages can be used to estimate their relative distance	Transduced	Co transduced	Co repressor	Co operator	Co transduced
22	Who coined the term linkage	Mendel	Morgan	de Vries	Correns	Morgan
23	The complex of DNA, RNA and protein is	Chromatin	Somatin	Pigmentin	Fromatin	Chromatin
24	The process to identify a genetic element that is responsible for a disease is also referred as	Mapping	Linkage	Sequencing	Genome data mining	Mapping
25	Cells carrying non-integrated transducing fragments are called _____	Specialized	Abortive transductants	Generalized	Conjugation	Abortive transductants
26	Crossing over occurs during	Pachytene	Diplotene	Diakinesis	Haplotene	Pachytene
27	Avery, MacLeod & McCarty used enzymes and solvents to destroy molecules such as	Anything except DNA	RNA	Lipids & proteins	Polysaccharide	Anything except DNA
28	A bacterium harboring a prophage is called	Lytic phage	Helper phage	Transducing phage	Lysogency	Lysogency
29	Genome of T4 phage is	ds DNA	ss DNA	ds RNA	ss RNA	ds DNA
30	Non sex chrosomes are called	Rhizomes	Lysosomes	Mesosomes	Autosomes	Autosomes
31	Uptake of DNA molecules from environmental surrounding	Transduction	Conjugation	Transformation	Gene expression	Transformation
32	Results from inaccurate excision of an integrated prophage with addition of some bacterial genes	Specialized Transduction	Generalized Transduction	Abortive transduction	Transfusion	Specialized Transduction

33	Give the full form for Hfr	High fertility recombination	High fundamental recombination	High frequency recombination	Heavy frequency recombination	High frequency recombination
34	Genes that cause suppression of mutations in other genes are called -------- genes	Reverse genes	Control genes	Suppressor genes	Inducer genes	Suppressor genes
35	Conjugation is predominant in _____	Spirochaetes	G+ bacteria	G-bacteria	Cyanobacteria	G-bacteria
36	_____ is a DNA associated protein	Protone	Histone	Chromotome	Cistron	Histone
37	Conjugation involves the use of _____for mapping	Interrupted mating	Direct mapping	Contact mapping	Linkage	Interrupted mating
38	During insertion of lambda DNA in to host, a viral protein called _____ is required along with integration host factor (IHF)	Integrase	Caspase	Helicase	Polymerase	Integrase
39	Linkage prevents	Segregation of alleles	Homozygous condition	Hybrid formation	Heterozygous condition	Segregation of alleles
40	T4 is capable of undergoing only a	Lytic cycle	Lysogenic cycle	Both Lytic & Lysogenic cycle	Other cycle	Lytic cycle
41	A cell carrying an integrated F factor is called an	F	Hfr	Hfr 1+	trans	Hfr
42	Capsule of *Streptococcus pneumoniae* are made up of	Protein	Lipid	Glycoprotein	Polysaccharide	Polysaccharide
43	Conjugation involves the use of _____for mapping	Interrupted mating	Direct mapping	Contact mapping	Linkage	Interrupted mating

44	Genetic fine structure mapping of T4 was studied by _____	Benzon	Mendel	Colins	Bennazir	Benzon
45	In genetic mapping, the measurement of distance between the genes is expressed as	Centimorgan	Centimeter	Millimorgan	Millimeter	Centimorgan
46	Methods used to identify the locus of a gene and the distances between genes	Gene mapping	Chromosomal linkage	Gene walking	Chromosomal walking	Gene mapping
47	T4 bacteriophages generally parasitizes	*Bacillus*	*E.coli*	*Psuedomonas*	*Agrobacterium*	*E.coli*
48	The non specific transduction is also called as - - - - - - - - - -	Restricted transduction	Generalized transduction	Non specific transduction	Specialised transduction	Generalized transduction
49	Transfer of a portion of chromosome to a recipient with direct contact is termed	Gene expression	Transformation	Transduction	Conjugation	Conjugation
50	The first demonstration of bacterial transformation was done with	*Streptococcus pyogenes*	*Staphylococcus aureus*	*Streptococcus pneumoniae*	*Klebsiella pneumoniae*	*Streptococcus pneumoniae*
51	Size of T4 phage genome	169 kbp	196 kbp	619 kbp	916 kbp	169 kbp
52	Map distance is equal to the percentage of _____	Recombinant meiotic product	Reproducible meiotic product	Recombinant mitotic product	Reproducible mitotic product	Recombinant meiotic product
53	Genome of different bacteria suggest that genes have in the past moved from one species to another. This phenomenon is called lateral	DNA transfer	RNA transfer	Gene transfer	Protein transfer	Gene transfer
54	Experiment on transformation	Monad	Griffith	Morgan	Hersehy	Griffith

55	Conjugational genetic exchange has been - - - - - - - - frequently encountered among gram positive than gram negative bacteria.	More	Less	Very High	Very low	Less
56	Complete linkage has been reported in	Male Drosophila	Human female	Female Drosophila	Maize	Male Drosophila
57	Bacteriophages were jointly discovered by	Frederick Twort and by Felix d'Herelle	Hershey and Chase	Luria and Delbruck	McKay and McCartney	Frederick Twort and by Felix d'Herelle
58	Capsules help bacteria in escaping	Inflammation	RBC's	Phagocytosis	Antibodies	Phagocytosis
59	Genes responsible for antibiotic resistance are mainly transferred across bacterial population by	Conjugation	Transformation	Transduction	Gene expression	Conjugation
60	Conjugation involves the use of _____for mapping	Interrupted mating	Direct mapping	Contact mapping	Linkage	Interrupted mating

Unit V

Principles of transcriptional regulation, regulation at initiation with examples from *lac* and *trp* operons, Sporulation in *Bacillus,* Yeast mating type switching , Changes in Chromatin Structure - DNA methylation and Histone Acetylation mechanisms.

The Protein Synthesis Machinery

In addition to the mRNA template, many molecules and macromolecules contribute to the process of **translation**. The composition of each component varies across taxa; for instance, ribosomes may consist of different numbers of ribosomal RNAs (rRNAs) and polypeptides depending on the organism. However, the general structures and functions of the protein synthesis machinery are comparable from bacteria to human cells. Translation requires the input of an mRNA template, ribosomes, tRNAs, and various enzymatic factors.

Ribosomes: A **ribosome** is a complex macromolecule composed of catalytic rRNAs (called **ribozymes**) and structural **rRNAs**, as well as many distinct polypeptides. Mature rRNAs make up approximately 50% of each ribosome. Prokaryotes have 70S ribosomes, whereas eukaryotes have 80S ribosomes in the cytoplasm and rough endoplasmic reticulum, and 70S ribosomes in **mitochondria** and **chloroplasts**. Ribosomes dissociate into large and small subunits when they are not synthesizing proteins and reassociate during the **initiation of translation**. In *E. coli*, the **small subunit** is described as 30S (which contains the 16S rRNA subunit), and the **large subunit** is 50S (which contains the 5S and 23S rRNA subunits), for a total of 70S (Svedberg units are not additive). Eukaryote ribosomes have a small 40S subunit (which contains the 18S rRNA subunit) and a large 60S subunit (which contains the 5S, 5.8S and 28S rRNA subunits), for a total of 80S. The small subunit is responsible for binding the mRNA template, whereas the large subunit binds tRNAs (discussed in the next subsection).

Each mRNA molecule is simultaneously translated by many ribosomes, all synthesizing protein in the same direction: reading the mRNA from 5′ to 3′ and synthesizing the polypeptide from the N terminus to the C terminus. The complete structure containing an mRNA with multiple associated ribosomes is called a **polyribosome** (or **polysome**). In both bacteria and **archaea**, before transcriptional termination occurs, each protein-encoding transcript is already being used to begin synthesis of numerous copies of the encoded polypeptide(s) because the processes of transcription and translation can occur concurrently, forming polyribosomes (Figure 2). The reason why transcription and translation can occur simultaneously is because both of these processes occur in the same 5′ to 3′ direction, they both occur in the cytoplasm of the cell, and because the RNA transcript is not processed once it is transcribed. This allows a prokaryotic cell to respond to an environmental signal

requiring new proteins very quickly. In contrast, in eukaryotic cells, simultaneous

transcription and translation is not possible. Although polyribosomes also form in eukaryotes, they cannot do so until RNA synthesis is complete and the RNA molecule has been modified and transported out of the nucleus.

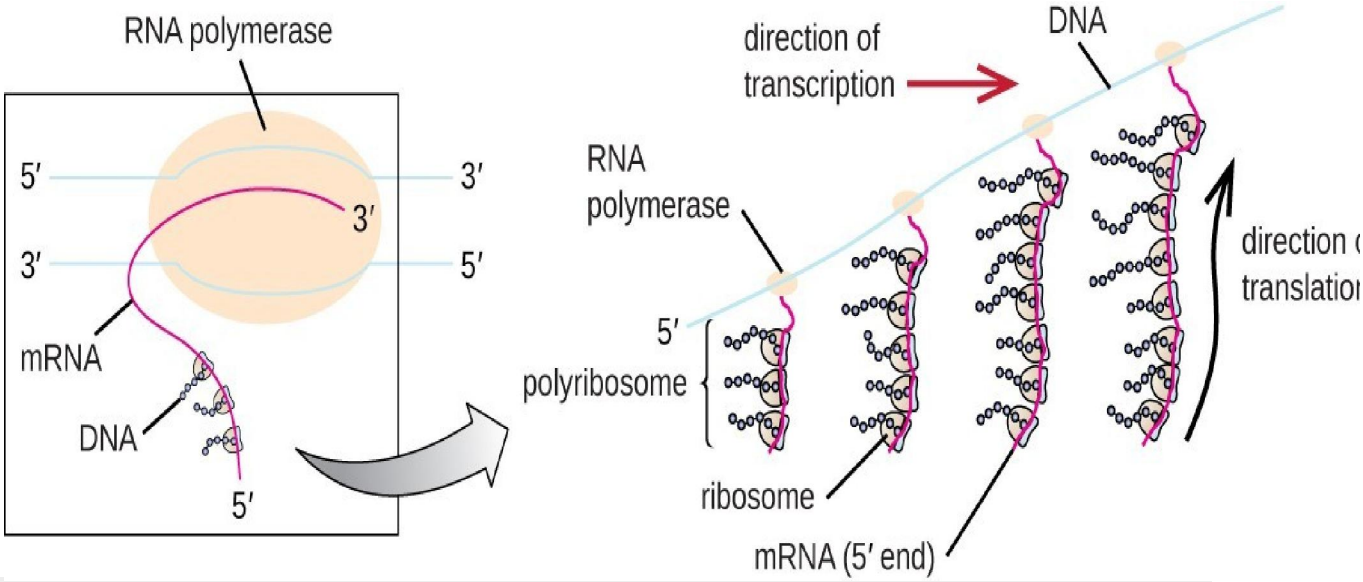


Figure 2. In prokaryotes, multiple RNA polymerases can transcribe a single bacterial gene while numerous ribosomes concurrently translate the mRNA transcripts into polypeptides. In this way, a specific protein can rapidly reach a high concentration in the bacterial cell.

Transfer RNAs

Transfer RNAs (tRNAs) are structural RNA molecules and, depending on the species, many different types of tRNAs exist in the cytoplasm. Bacterial species typically have between 60 and 90 types. Serving as adaptors, each **tRNA** type binds to a specific codon on the mRNA template and adds the corresponding amino acid to the polypeptide chain. Therefore, tRNAs are the molecules that actually "translate" the language of RNA into the language of proteins. As the adaptor molecules of translation, it is surprising that tRNAs can fit so much specificity into such a small package. The tRNA molecule interacts with three factors: aminoacyl tRNA synthetases, ribosomes, and mRNA.

Mature tRNAs take on a three-dimensional structure when complementary bases exposed in the single-stranded RNA molecule hydrogen bond with each other (Figure 3). This shape positions the amino-acid binding site, called the **CCA amino acid binding end**, which is a cytosine-cytosine-adenine sequence at the 3′ end of the tRNA, and the **anticodon** at the other end. The anticodon is a three-nucleotide sequence that bonds with an mRNA codon through complementary base pairing.

An amino acid is added to the end of a tRNA molecule through the process of tRNA "charging," during which each tRNA molecule is linked to its correct or **cognate amino**

acid by a group of enzymes called **aminoacyl tRNA synthetases**. At least one type of aminoacyl tRNA synthetase exists for each of the 20 amino acids. During this process, the amino acid is first activated by the addition of adenosine monophosphate (AMP) and then transferred to the tRNA, making it a **charged tRNA**, and AMP is released.

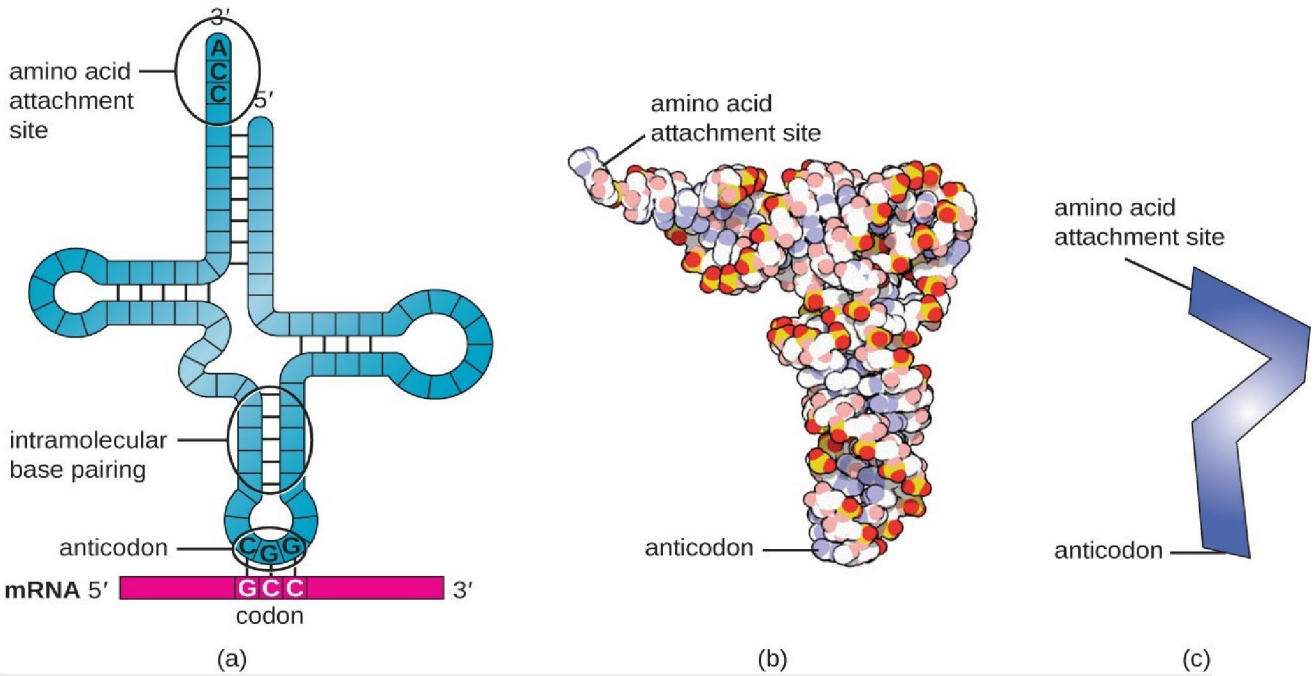


Figure 3. (a) After folding caused by intramolecular base pairing, a tRNA molecule has one end that contains the anticodon, which interacts with the mRNA codon, and the CCA amino acid binding end. (b) A space-filling model is helpful for visualizing the three-dimensional shape of tRNA. (c) Simplified models are useful when drawing complex processes such as protein synthesis.

Aminoacyl TRNA synthetases

Aminoacyl-tRNA synthetases and translation factors are key enzymes required for pro tein biosynthesis. *Escherichia coli* and *Bacillus subtilis* often use different strategies to regulate the expression of the genes encoding these enzymes. Synthesis of several *E. coli* aminoacyl-tRNA synthetases is controlled by different mechanisms acting at the transcriptional or translational level. By contrast, in *B. subtilis*, expression of the majority of these proteins is regulated by a common, yet specific transcriptional antitermination mechanism. However, all of these controls share a common effector, the tRNA. In the case of the *E. coli* translation factors, the primary enzyme function is often exploited for autoregulating their own synthesis at the translational level. Here we will focus on the gene organization and the multiple types of gene regulation governing prokaryotic aminoacyl-tRNA synthetase and translation factor expression.

Introduction

Decoding the genetic message is the major and most energy consuming process in the cell. In addition to ribosomes, mRNA and tRNA, translation requires amino acids, nucleotides and specialized proteins. The latter include aminoacyl-tRNA synthetases and translation factors transiently associated with the ribosomes. These proteins catalyze sequential steps during translation, starting with the charging of tRNA, ribosome-dependent polypeptide synthesis, final release of the protein and ribosome recycling. Here, we will consider our current knowledge of the gene organization and the expression of aminoacyl-tRNA synthetases and translation factors in prokaryotes. We focus essentially on two bacterial systems, the Gram negative bacterium *Escherichia coli* and *Bacillus subtilis*, the best-studied Gram positive organism. Translation factors are only considered for *E. coli*. Structural and mechanistic aspects of these enzymes are treated in accompanying chapters.

Aminoacyl-tRNA Synthetases

Aminoacyl-tRNA synthetases (aaRS) play a central role in protein biosynthesis by catalyzing the attachment of a given amino acid to the 3' end of its cognate tRNA. They do this by forming an energy-rich aminoacyl-adenylate intermediate of the cognate amino acid, which serves to transfer the amino acid to the tRNA. The intrinsic proofreading capacities of the aaRS and their balanced expression contribute greatly to the accuracy of translation of the genetic code. In addition to their crucial role in protein biosynthesis, aaRS are involved in a number of regulatory processes via their product, the charged tRNA. The control of the expression of amino acid biosynthetic operons by the level of tRNA aminoacylation in vivo is well documented. Another phenomenon, the pleiotropic stringent response, operates when tRNA aminoacylation is diminished in wild-type (relA) strains and inhibits rRNA and tRNA synthesis as well as the synthesis of some other macromolecules involved in translation.[1] As a by-product of the amino acid activation step aaRS can form diadenosine 5'-5'''-P1, P4-tetraphosphate (AppppA), considered as a pleiotropically acting alarmone that has been associated with oxidative stress or timing of cell division.[2] Eucaryotes exploit yet other functions of aaRS, e.g., some aaRS are essential factors in certain splicing activities.[3] It is obvious that control of the cellular levels of the aaRS is important for any organism.

The Mechanism of Protein Synthesis

Translation is similar in prokaryotes and eukaryotes. Here we will explore how translation occurs in *E. coli*, a representative prokaryote, and specify any differences between bacterial and eukaryotic translation.

Initiation

The **initiation of protein synthesis** begins with the formation of an initiation complex. In ***E. coli***, this complex involves the small 30S ribosome, the mRNA template, three **initiation factors** that help the ribosome assemble correctly, guanosine triphosphate (GTP) that acts as an energy source, and a special initiator tRNA carrying ***N*-formyl-methionine** (fMet-$tRNA^{fMet}$) (Figure 4). The initiator tRNA interacts with the start codon AUG of the mRNA

and carries a formylated methionine (fMet). Because of its involvement in initiation, fMet is inserted at the beginning (N terminus) of every polypeptide chain synthesized by *E. coli*. In *E. coli* mRNA, a leader sequence upstream of the first AUG codon, called the **Shine-Dalgarno sequence** (also known as the ribosomal binding site AGGAGG), interacts through complementary base pairing with the rRNA molecules that compose the ribosome. This interaction anchors the 30S ribosomal subunit at the correct location on the mRNA template. At this point, the 50S ribosomal subunit then binds to the initiation complex, forming an intact ribosome.

In eukaryotes, initiation complex formation is similar, with the following differences:

- The initiator tRNA is a different specialized tRNA carrying methionine, called Met-tRNAi
- Instead of binding to the mRNA at the Shine-Dalgarno sequence, the eukaryotic initiation complex recognizes the 5′ cap of the eukaryotic mRNA, then tracks along the mRNA in the 5′ to 3′ direction until the AUG start codon is recognized. At this point, the 60S subunit binds to the complex of Met-tRNAi, mRNA, and the 40S subunit.

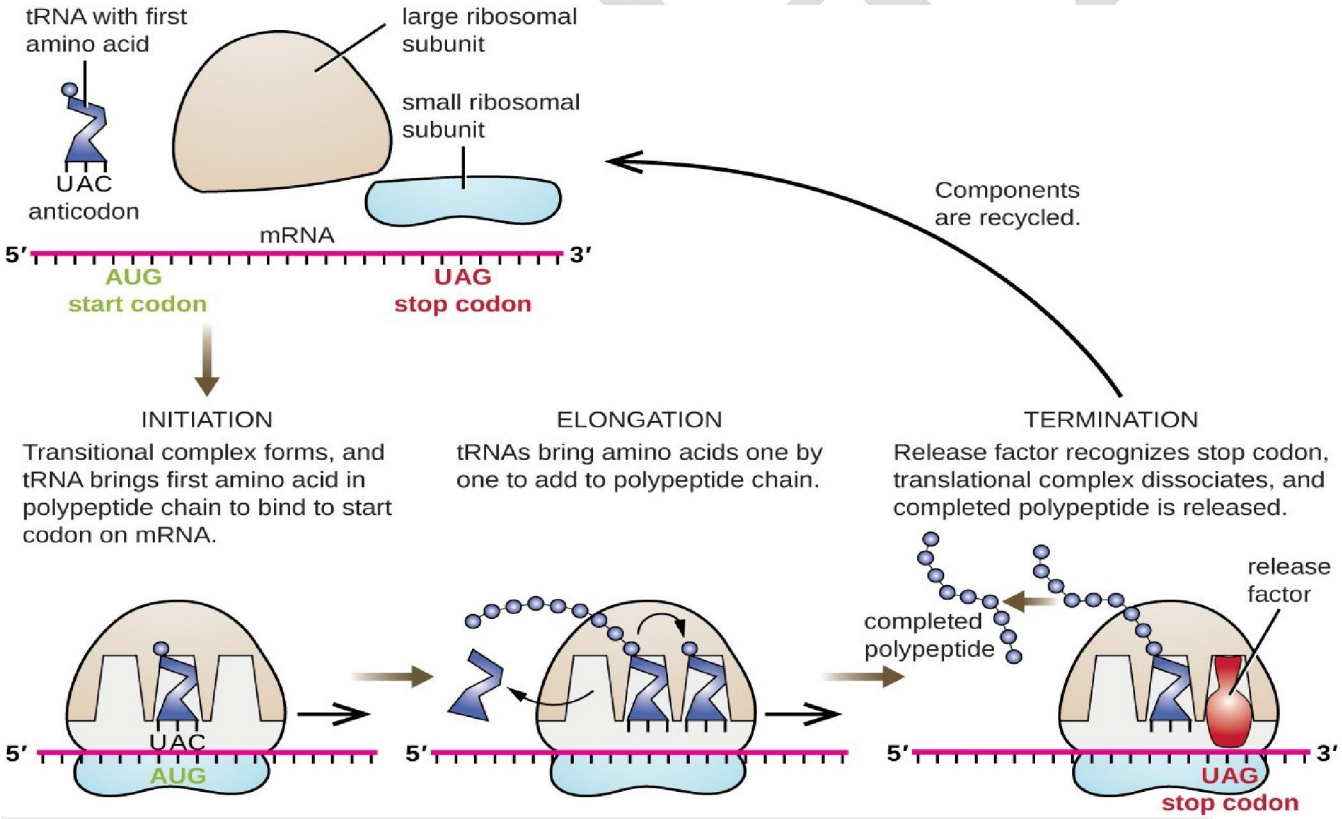


Figure 4. Translation in bacteria begins with the formation of the initiation complex, which includes the small ribosomal subunit, the mRNA, the initiator tRNA carrying N-formyl-methionine, and initiation factors. Then the 50S subunit binds, forming an intact ribosome.

Elongation

In prokaryotes and eukaryotes, the basics of **elongation of translation** are the same. In ***E. coli***, the binding of the 50S ribosomal subunit to produce the intact ribosome forms three functionally important ribosomal sites: The **A (aminoacyl) site** binds incoming charged aminoacyl tRNAs. The **P (peptidyl) site** binds charged tRNAs carrying amino acids that have formed peptide bonds with the growing polypeptide chain but have not yet dissociated from their corresponding tRNA. The **E (exit) site** releases dissociated tRNAs so that they can be recharged with free amino acids. There is one notable exception to this assembly line of tRNAs: During initiation complex formation, bacterial fMet–tRNA^{fMet} or eukaryotic Met-tRNAi enters the P site directly without first entering the A site, providing a free A site ready to accept the tRNA corresponding to the first codon after the AUG.

Elongation proceeds with single-codon movements of the ribosome each called a translocation event. During each translocation event, the charged tRNAs enter at the A site, then shift to the P site, and then finally to the E site for removal. Ribosomal movements, or steps, are induced by conformational changes that advance the ribosome by three bases in the 3′ direction. Peptide bonds form between the amino group of the amino acid attached to the A-site tRNA and the carboxyl group of the amino acid attached to the P-site tRNA. The formation of each peptide bond is catalyzed by **peptidyl transferase**, an RNA-based ribozyme that is integrated into the 50S ribosomal subunit. The amino acid bound to the P-site tRNA is also linked to the growing polypeptide chain. As the ribosome steps across the mRNA, the former P-site tRNA enters the E site, detaches from the amino acid, and is expelled. Several of the steps during elongation, including binding of a charged aminoacyl tRNA to the A site and translocation, require energy derived from GTP hydrolysis, which is catalyzed by specific elongation factors. Amazingly, the *E. coli* translation apparatus takes only 0.05 seconds to add each amino acid, meaning that a 200 amino-acid protein can be translated in just 10 seconds.

Termination

The **termination of translation** occurs when a **nonsense codon** (UAA, UAG, or UGA) is encountered for which there is no complementary tRNA. On aligning with the A site, these nonsense codons are recognized by release factors in prokaryotes and eukaryotes that result in the P-site amino acid detaching from its tRNA, releasing the newly made polypeptide. The small and large ribosomal subunits dissociate from the mRNA and from each other; they are recruited almost immediately into another translation init iation complex.

In summary, there are several key features that distinguish prokaryotic gene expression from that seen in eukaryotes. These are illustrated in Figure 6 and listed in Table 1.

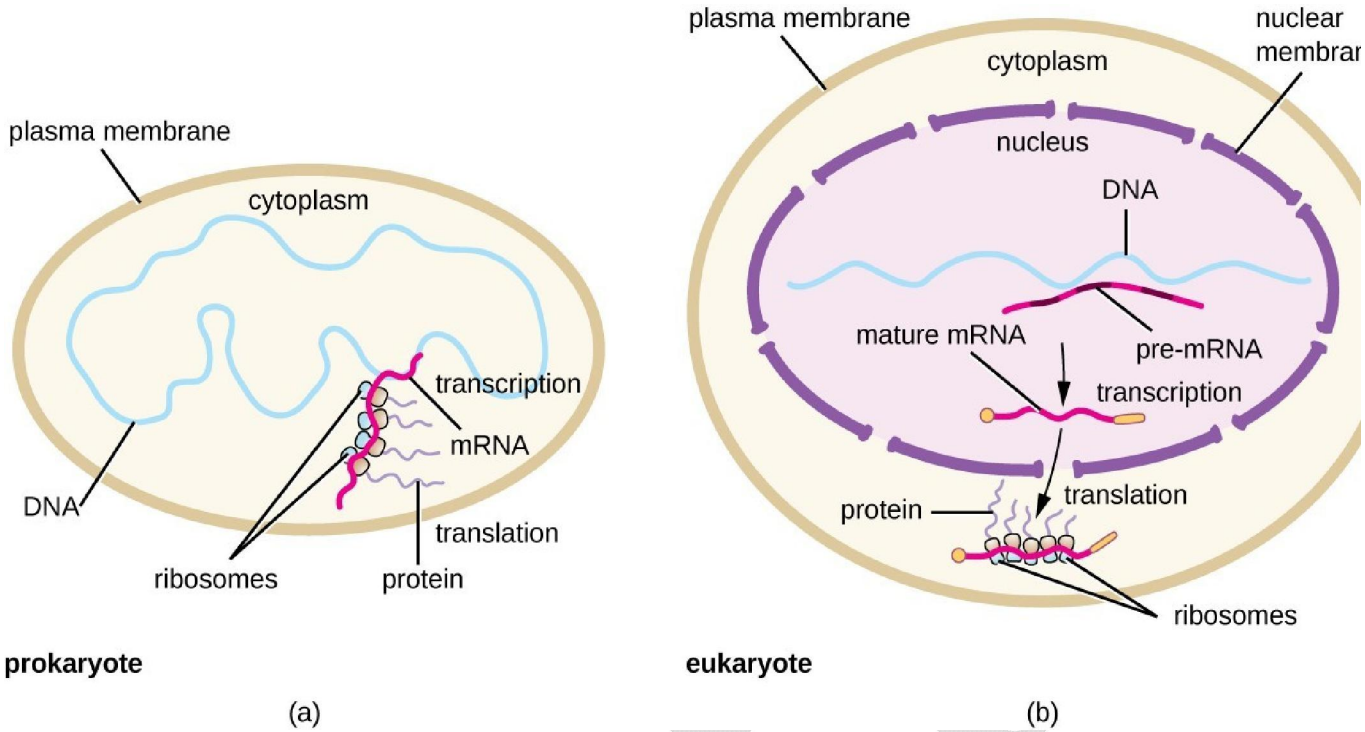


Figure 6. (a) In prokaryotes, the processes of transcription and translation occur simultaneously in the cytoplasm, allowing for a rapid cellular response to an environmental cue. (b) In eukaryotes, transcription is localized to the nucleus and translation is localized to the cytoplasm, separating these processes and necessitating RNA processing for stability.

Fidelity of translation

Fidelity and transparency are two factors that, for thousands of years, have been regarded as the highest ideals to be endeavored for in human translation(particularly literary translation); even up until now, when translation jobs are farmed out to different professional translation groups and translation service freelancers, these twin qualities are still considered top-priority guidelines to better achieve successful translation work with clear messages. Localization and globalization may have helped a lot in aiding different audiences to better understand the gist of any given human translation project, but the balance between transparency and fidelity remains the best way for a client to convey the ultimate message of their source text with little to no misunderstanding on the part of their target demographic.

With that said, these ideals are often at odds with each other. Ergo, a seventeenth-century French critic once humorously observed that translations, like women, could either be beautiful and faithful, but never both simultaneously. In other words, a translation could have more fidelity and less transparency or vice-versa, but never both at equally high amounts. More to the point, as many a translation service agency or professional translation company knows, fidelity refers to the faithfulness of the translation to the source text, while

transparency refers to the comprehensibility of the translation in terms of the target audience's cultural perspective.

In-Depth Definition

Fidelity refers to the limits to which a given human translation work precisely depicts the underlying message or meaning of the source text without distorting it, without intensifying or weakening any part of its context, and otherwise without subtracting or adding to it at all. Transliteration a la with machine translation comes closest to this school of thought, with the caveat that it usually fails to properly convey the message because of its rigid faithfulness to the original document.

Meanwhile, transparency pertains to the degree to which a translation caters to native speakers and the target audience, such that idiomatic, syntactic, and grammatical conventions are followed while cultural, political, and social context is kept in mind at all times. Adaptation and localization comes closest to this school of thought, with the caveat that a bit of sacrifice in terms of the intended message will inevitably happen whenever translators use this approach in their translation.

Translations that have high fidelity are classified as "faithful" translations; in turn, translations that meet the second standard are referred to as "idiomatic" translations. Then again, the two ideals aren't necessarily mutually exclusive from each other; keeping the context intact, for example, can be seen in both a high-fidelity translation and a high-transparency translation. A veritable multitude of paradigms and measures could also be used to review the faithfulness of a given translation, such as social or historical context, its literary qualities, function and use of the text, the type, the precision of the original contents, and the subject.

Important Inhibitors of Protein Synthesis

Many substances are known to act as inhibitors of various stages of protein synthesis. Included among these are a number of antibiotics produced by one strain of microorganism and lethal to other strains of the same or a different species.

Some of the best understood inhibitors of protein synthesis are listed in Table 22-10. Because the actions of many of these inhibitors are quite specific, they have proved to be extremely useful tools in the step-by-step elucidation of the mechanism of protein synthesis.

TABLE 22–10 INHIBITORS OF PROTEIN SYNTHESIS IN PROKARYOTIC AND EUKARYOTIC CELLS

Inhibitor	Effective in	
	Prokaryotes	Eukaryotes
Anisomycin	–	+
Aurintricarboxylic acid	+	+
Chloramphenicol	+	–
Colicin E3	+	–
Cycloheximide	–	+
Diphtheria toxin	–	+
Edeine	+	+
Erythromycin	+	–
Fusidic acid	+	+
Pactamycin	–	+
Puromycin	+	+
Ricin	–	+
Sodium fluoride	–	+
Sparsomycin	–	+
Streptomycin	+	–
Tetracycline	+	+
Trichodermin	–	+

Inhibitors of both Prokaryotic and Eukaryotic Protein Synthesis:

Aurintricarboocylic acid inhibits formation of the initiation complex by preventing the association of mRNA with the small ribosomal subunit. Inhibitors of initiation are readily distinguished from inhibitor's blocking other stages of protein synthesis because of the delay effect that follows their administration, that is, protein synthesis continues for a short time after administration of the inhibitor, because peptide chains whose growth had already begun are unaffected and grow to completion.

Edeine, a polypeptide isolated from Bacillus brevis, inhibits the binding of aminoacyl-tRNA and N-formylmet-tRNA$^{M}_{f}{}^{et}$ (in prokaryotes) to the small subunit. Fusidic acid is a steroidal antibiotic; in prokaryotes, it inhibits the binding of aminoacyl-tRNA to the ribosome, whereas in eukaryotes, it inhibits translocation by reacting with elongation factor.

Puromycin was one of the first inhibitors of protein synthesis to have its specific effect determined. This antibiotic mimics aminoacyl-tRNA and binds to the free A site of ribosomes engaged in protein synthesis. Catalytic formation of a bond between the nascent polypeptide and puromycin is followed by the release of the peptidyl-puromycin from the ribosome, as no further elongation is possible.

:

The specific effects of puromycin have been used to advantage for studies of nascent chain length, the kinetics of chain elongation, and the identification of the effects of other antibiotics. Tetracycline inhibits protein synthesis by blocking aminoacyl-tRNA binding to the small subunit.

2. Inhibitors Specific for Prokaryotes:

Chloramphenicol (Chloromycetin) binds to the large subunit of prokaryotic ribosomes and interferes with the functioning of peptide synthetase, thereby inhibiting chain elongation. Colicin E3 inhibits protein synthesis in prokaryotes by interfering in some manner with the functioning of the small subunit. Erythromycin binds to ribosomes that are not engaged in protein synthesis, preventing their potential participation, but does not bind to ribosomes containing nascent chains (i.e., ribosomes that are part of a functioning polysome). Streptomycin was one of the earliest discovered antibiotics and was employed as an agent against bacterial infection for many years before its specific chemical actions were known.

Streptomycin binds to protein S12 of the small ribosome subunit, causing release of N-formylmet-tRNA$^{M}_{f}{}^{et}$ from initiation complexes (thereby preventing initiation of chain growth) and also causing misreading of the codons of mRNA by ribosomes already involved in chain elongation.

3. Inhibitors Specific for Eukaryotes:

Anisomycin is an antibiotic produced by Streptomyces that inhibits peptide bond formation when bound to the small ribosomal subunit. Cycloheximide binds to the large subunit, preventing the translocation of tRNA in the A site to the P site. Diphtheria toxin (produced by a strain of Corynebacterium diphtherial) inhibits protein synthesis through its action on EF-2 (translocase). EF-2 exists in cells in two forms— ribosome-bound and free. Diphtheria toxin acts enzymatically to alter free EF-2, rendering the factor inactive. Ribosome-bound EF-2 is not susceptible to inactivation by the toxin.

Pactamydn (produced by a strain of Streptomyces) binds to free small subunits (not to small subunits already part of polysomes), where it prevents initiation by inhibiting binding of met-tRNA j and formation of the initiation complex. The toxic effects of ricin (a protein present in the castor bean) have been known for nearly a century.

Ricin consists of two polypeptide chains (linked by disulfide bridges), one of which acts as the inhibitor once incorporated into the cell. Ricin acts on the large subunit, preventing formation of the 80 S initiation complex. Like ricin, sodium fluoride acts as an inhibitor or initiation; NaF blocks addition of the large subunit to mRNA.

Sparsomycin, another antibiotic produced by Strep- tomyces, inhibits the association of the amino acid moiety of aminoacyl-tRNA from binding to the large sub- unit and, in so doing, blocks peptide synthetase. THchodermin is the only chemical compound so far identified as a specific inhibitor of the termination stage of polypeptide synthesis.

Mating type switching in budding yeast

Homothallic yeast cells can, remarkably, switch their mating type as often as every generation by a highly regulated, site-specific homologous recombination event that replaces one MAT allele with a different DNA sequence that encodes the opposite MAT allele. This replacement process involves the participation of two intact but unexpressed copies of mating-type information at the heterochromatic loci, HMLα and HMRa, which are located at opposite ends (**L**eft and **R**ight, respectively) of the same chromosome that encodes MAT.

Saccharomyces cerevisiae is heterothallic but a clone of haploid cells of the same mating type frequently sporulates, and there will be equal numbers of a and α cells amongst the progeny. This results from mating type switching controlled by the gene *HO* (*HO*mothallic) that exists in two allelic forms (dominant *HO* and recessive *ho*), and encodes an endonuclease. On either side of the *MAT* locus, and on the same chromosome, there are silent storage loci for each mating type, called *HML* and *HMR*. The *HO*/*ho* endonuclease creates a double-strand break at the *MAT* locus that initiates switching of information, by an homologous recombination event between the two parts of the same chromosome, at the *MAT* locus with that at either *HML* or *HMR* (Fig. 4).

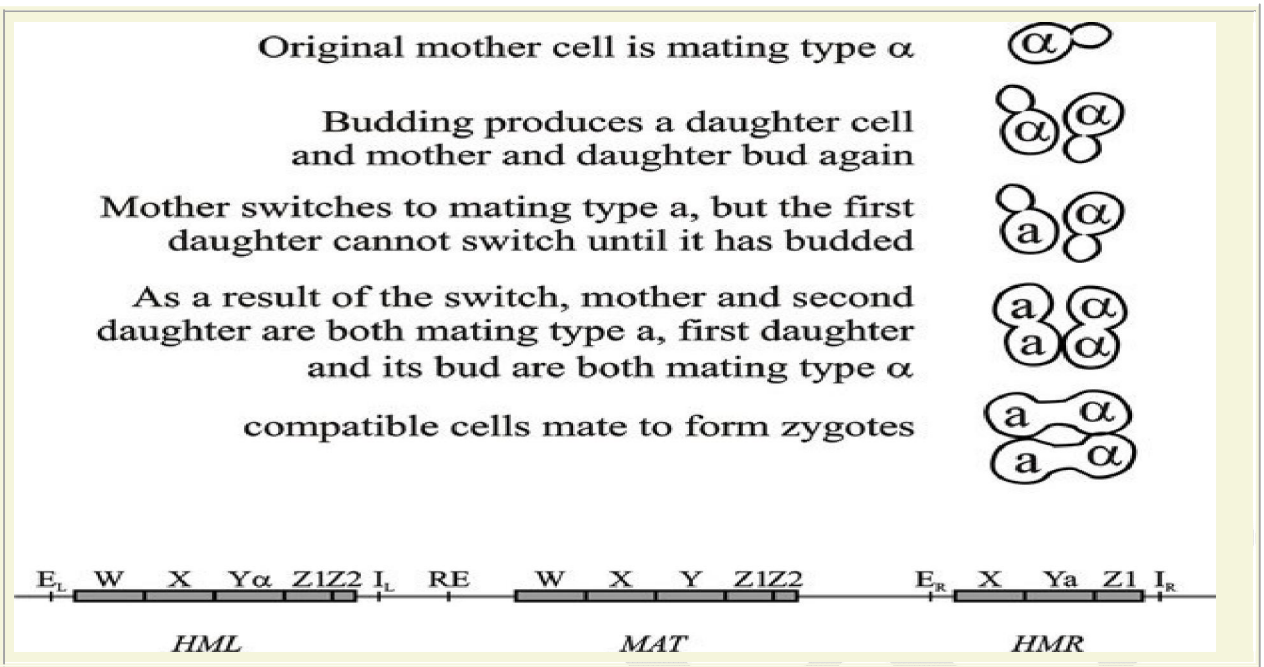


Fig. 4. Top: pattern of mating type switching in *Saccharomyces cerevisiae* showing the consequences of a mating type switch in one mother cell. Bottom: the three loci involved in mating type switching, *HML*, *MAT* and *HMR*, are located on the same chromosome (***not*** drawn to scale). *HML* is about 180 kb from *MAT*, and *HMR* about 120 kb from *MAT*; the centromere is located between RE and the *MAT* locus. A double strand break at the *MAT* locus, caused by the HO endonuclease, initiates a gene conversion event that replaces the Y region of the *MAT*locus with Y sequences from one of the storage loci. *HML* and *HMR* contain complete copies of the mating type genes but are not expressed because they have a repressed heterochromatin structure imposed by the E and I silencer sequences. *HML* shares more of the *MAT* sequences (W, X, Z1 and Z2) than does *HMR*. RE is a recombination enhancer that controls preferential recombination between *MAT*a and *HML*, or between *MAT*a and *HMR*. Modified from Chapter 2 in Moore & Novak Frazer, 2002.

Again, MAT heterozygosity plays a key role in controlling the switching of mating-type genes. Switching occurs in homothallic strains expressing the HO endonuclease gene, HO, which are able to switch from MATa to MATα and vice versa, but once cells of opposite mating type conjugate to form a diploid, HO expression is repressed by the a1-α2 repressor. The Mata1, Matα1, and Matα2 transcription regulators that organise haploid cell specificity are quite rapidly turned over, being degraded by ubiquitin-mediated proteolysis by the proteasome; in contrast, the a1-α2 corepressor is much more stable.

Since yeasts can live in very small habitats, like flower nectaries and surfaces of individual fruits, yeast populations can be very isolated from one another in nature, so the rare mating type switching will give isolated populations the opportunity to undergo sexual reproduction;

this is presumably its selective advantage. Mating type switching occurs about once in 10^5 divisions in cultures carrying allele *ho*, whereas in strains carrying *HO* the switch occurs at every cell division. However, there is an asymmetry in the cell division in that a new daughter bud is not able to switch mating types until it has itself budded. In *S. cerevisiae*, this is achieved by actively transporting the mRNA of a gene called *Ash1* into the budding daughter cell. This mRNA encodes an inhibitor of the HO-endonuclease. Consequently, immediately after each division switching by the daughter is inhibited and only the mother cell is switchable. This means that even if there is only one cell to start with, a single division cycle will produce two cells of opposite mating type.

Saccharomyces cerevisiae has evolved an elaborate set of mechanisms to enable cells to switch their mating types. MAT switching depends on four phenomena:

1. Presence of two unexpressed (silenced or cryptic) copies of mating-type sequences that act as donors during MAT switching. This implies that there is a mechanism which influences chromatin structure to maintain it in a silent configuration. The mechanism involves silencer sequences surrounding HML and HMR that interact, directly or indirectly, with several protein factors to repress the transcription of these genes. Among these are four **S**ilent **I**nformation**R**egulator (**Sir**) proteins, a set of silencer binding proteins, histone proteins, the Rap1 protein, as well as several chromatin modifiers. Together, these create short regions (about 3 kb) of heterochromatin, in which the DNA sequences of HML and HMR are in a highly ordered nucleosome structure, known as **heterochromatin**, which is not transcribed by either RNA polymerase II or RNA polymerase III, and is resistant to cleavage by several endonucleases, including the HO endonuclease.
2. Programmed creation of a site-specific double-strand break at MAT that results in the replacement of Ya or Yα sequences. HO is a site-specific endonuclease that recognises a 24-bp sequence that spans the MAT-Y/Z border. Haploid yeast has three possible targets for HO: the MAT locus, HMLα, and HMRa, but only the MAT locus is accessible under normal conditions because of gene silencing of HML and HMR. Normally the HO gene is tightly regulated and is expressed only in haploid mother cells and only at the G1 stage of the cell cycle. The result in mother cells is a single programmed double-strand break at the MAT locus prior to the initiation of DNA replication.
3. A cell lineage rule ensuring that any cell that has previously divided once can switch MAT, while new daughter cells cannot (so only half of the cells in a population switch at any one time). The way this works is that a germinating haploid spore

grows, produces a bud, and divides without changing mating type. Then, in the next cell division cycle, the older mother cell and its second daughter change mating type while the first daughter buds and divides without changing. This ensures that there will be cells of both mating types in close proximity; they readily conjugate, forming MATa/MATα diploids in which the HO endonuclease gene is turned off so that further mating-type switching is repressed. Control of this lineage pattern depends on expression of the HO gene being restricted to mother cells that have divided by the Swi5 transcription factor, which is localised to the mother cell nucleus and not that of her daughter. Lack of Swi5 expression in daughters is caused by the Ash1 repressor protein that is found only in the daughter cell. Ash1 may directly repress SWI5 gene transcription, thus restricting HO expression to the mother cell in the next G1 stage of the cell cycle. Ash1 mRNA is localised to the daughter prior to cell division by a myosin-like protein, actin, cargo binding proteins, translation repressors and a nuclear localising protein. Since Ash1 mRNA localisation was first discovered, many other mRNAs have been found to show similar localisation in *Saccharomyces*.

4. A mechanism regulates the selective use of the two donors (called donor preference); study of which has yielded much of what we know about double strand break-induced recombination during eukaryote mitosis. Switching one mating type to the other involves the replacement at the MAT locus of Ya or Yα by a gene conversion induced by the site-specific double-strand break at MAT caused by HO endonuclease; it is a DNA-damage repair process. The process is highly directional, in that the sequences at MAT are replaced by copying new sequences from either HMLα or HMRa (whichever is the alternate to the resident sequence), while the two donor gene loci remain unchanged. HO endonuclease cannot cleave its recognition sequence at either HML or HMR, as these sites are occluded by nucleosomes in silenced DNA and this prevents crossing over. Any DNA-single strand exchanges (Holliday junctions) that might otherwise become crossovers are removed by two helicase enzymes, Sgs1 and Mph1, working with their partner proteins. Thus the (resident) MAT locus is cleaved and becomes the recipient in this gene conversion process. Overall, following HO cleavage of MATa, the end of the broken DNA molecules is excised in a 5′ to 3′ direction, creating a 3′-ended single-strand DNA tail on which assembles a filament of the Rad51 recombinase protein. Rad51 is essential for repairing damaged DNA and is highly conserved in most eukaryotes; this family of proteins interact with several other single-strand DNA-binding proteins to form a helical nucleoprotein filament on the DNA. This protein-DNA complex engages in a search for a homologous DNA sequence (since we started with MATa, in this case it would search for HMLα) to effect the repair. This search culminates in strand exchange in which the single

stranded DNA base pairs with the complementary sequence of the donor, creating a displacement loop (D-loop). The 3′ end of the invading strand is then used as a primer to initiate copying of one strand of the donor locus, and the newly copied strand is displaced until it can anneal with homologous sequences on the opposite end of the double strand break. The 3′-ended nonhomologous tail is clipped off and the new 3′ end is used to prime a second strand of DNA synthesis, completing the replacement of MATa with MATα.

These important mechanisms (reviewed by Haber, 2012) are more examples of how research on fungi has informed our fundamental knowledge of the molecular genetics of eukaryotes.

The switching system of the related yeast, *Kluyveromyces lactis*, is similar to that of *Saccharomyces cerevisiae* to some extent, and includes sequences recognisably similar to MATa1, MATα1, and MATα2, but the shared flanking sequences are not closely related to the W, X, or Z1/Z2 sequences of *Saccharomyces cerevisiae* and there is ***no functional HO gene***. While the *Kluyveromyces lactis* HMR has a1 and a2; HML includes a novel gene, α3, in addition to α1 and α2. Both HML and HMR are silenced by a mechanism dependent on Sir-proteins as in *S. cerevisiae*. Switching is dependent on a protein (Mts1) that is the homologueof the *S. cerevisiae* RME1 repressor; but in *K. lactis* it is required to activate switching and is turned off in MATa/MATα cells by an a1-α2 repressor. The major difference between the two yeasts, at least for MATα to MATa switching, is that it is the α3 gene, which is a **transposable element** that can excise from the DNA as a circle and somehow catalyse switching (Barsoum *et al*. 2010). MATa to MATα switching seems to be under the control of a different transposable element (Haber, 2012).

Mating type switching also occurs in the more distantly related fission yeast *Schizosaccharomyces pombe*, but the switching system differs in almost every detail from *Saccharomyces cerevisiae*. Mating-type switching in *Schizosaccharomyces pombe* involves replacing genetic information at the expressed mat1 locus with sequences copied from ***one*** of ***two*** silent donor loci, mat2-P or mat3-M, which are close together and located only a short distance away from mat1 on the same chromosome arm in a 20-kb length of heterochromatin. There is no HO-like enzyme; instead, a persistent single-strand break is created in the DNA at mat1, which is converted to a double strand break when cells enter S phase. Only one of the two daughter cells can switch. Donor selection is dictated by cell type: mat2 is the preferred donor in M cells, and mat3 is the preferred donor in P cells. Donor preference involves major changes in chromatin modification and structure and the silencing system is different from that in *Saccharomyces cerevisiae* (Jia *et al*. 2004).

Yeasts in a third clade of the Saccharomycotina, the **methylotrophs**, have a simpler two-locus switching system based on reversible inversion of a section of chromosome with MATa genes at one end and MATα genes at the other end. In *Hansenula polymorpha* the invertible region, which is 19 kb long, lies beside a ***centromere*** so that, depending on the orientation, either MATa or MATα is silenced by centromeric heterochromatin. In *Pichia pastoris*, the orientation of a 138 kb invertible region puts either MATa or MATα beside a ***telomere*** where heterochromatin silences MATa2 or MATα2. Both species are homothallic, and inversion of their MAT regions can be induced by crossing two strains of the same mating type. The three-locus *Saccharomyces cerevisiae* system may have been derived from mating-type switching by chromosomal inversion as seen in methylotrophic yeasts; the increased complexity of the *S. cerevisiae* switching apparatus, with three loci, donor bias, and cell lineage tracking, resulting from continuous evolutionary selection to increase sporulation ability in young colonies (Hanson *et al*. 2014).

Mating type switching, which is often referred to as stochastic (meaning randomised) mating type determination, has evolved independently in a number of organisms other than the yeasts; for example, in the ciliates, *Tetrahymena thermophila*, *Paramecium* spp., and *Euplotes crassus*. There is also evidence for some degree of randomised mating type identity during vegetative growth in the green algae *Chlamydomonas monoica* and *Closterium ehrenbergii*, and the dinoflagellate *Gymnodinium catenatum*, although the switching mechanism in these species is not known. Mating types in filamentous fungi tend to be far more stable. Oddly enough, switching does not occur in any of the best-studied organisms like *Neurospora*, *Aspergillus*, or *Podospora*, but has been claimed in *Chromocrea spinulosa*, *Sclerotinia trifoliorum*, *Glomerella cingulata* and *Ophiostoma ulmi* (Tsui *et al*., 2013), and the basidiomycete *Agrocybe aegerita*(Hadjivasiliou *et al*. 2016).

Chromatin structure and modification

The genetic information of eukaryotic cells is packaged in the form of chromatin. The fundamental unit of this packaging is the nucleosome, comprising two copies of each of four different histone proteins, around which is wrapped the double-stranded DNA. The nucleosome "beads" are packed together into higher orders of structure, so that the entire length of the eukaryotic chromosomes can fit into the confines of the cellular nucleus. This packaging creates a barrier for the molecular machinery that needs access to the information encoded in DNA for gene expression, replication, recombination, and chromosome stability. Thus, an important area of modern molecular biology focuses on the structure of chromatin and how the various machines gain access to the DNA sites at which they exert their function.

Research at MSU explores many features of chromatin structure and function. Covalent modifications to histone proteins, including acetylation, phosphorylation, and methylation, may affect the higher-order structure of chromatin and may also serve to recruit transcription or replication machinery to specific locations. Genetic studies of histone acetylation in yeast and in plants seeks to define the acetylation enzymes themselves, the proteins that recognize acetylated histones, and additional pathways that work in conjunction with acetylation in regulating chromatin function. Deacetylation of histones represents the other aspect of this control, and is frequently associated with gene silencing. Plant pathogens that secrete inhibitors of histone deacetylases are being studied to define the impact of that inhibition on both the host and the pathogen. The developmental impact of changes in chromatin at specific genes is being explored in mammalian white blood cells and in plants. The ways in which animal viruses employ or bypass chromatin in regulating viral gene expression are also being examined. The novel structure of telomeres, at the ends of chromosomes, and the mechanisms for maintaining that structure are crucial to chromosome stability, with profound implications for cancer and aging. As befits the wide-ranging biological impact of chromatin structure and function, many of these laboratories are also associated with other focus groups in transcriptional regulation, signal transduction, or cancer cell biology.

Histone Acetylation

Acetylation of the lysine residues at the N terminus of histone proteins removes positive charges, thereby reducing the affinity between histones and DNA. This makes RNA polymerase and transcription factors easier to access the promoter region. Therefore, in most cases, histone acetylation enhances transcription while histone deacetylation represses transcription.

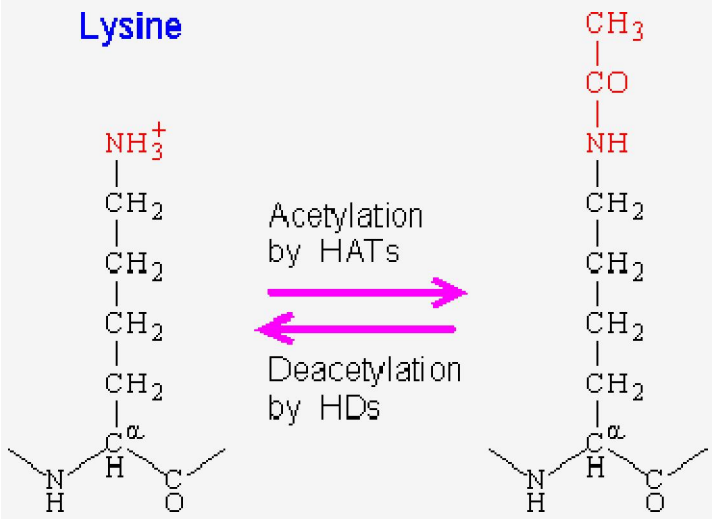


Figure 4-G-1. Acetylation and deacetylation of the lysine residue.

Histone acetylation is catalyzed by histone acetyltransferases (HATs) and histone deacetylation is catalyzed by histone deacetylases (denoted by HDs or HDACs). Several different forms of HATs and HDs have been identified. Among them, CBP/p300 is probably the most important, since it can interact with numerous transcription regulators.

DNA Methylation

DNA methylation is the addition of a methyl group (CH3) to the DNA's cytosine base. It may affect gene transcription through several different mechanisms (Illustration). The methylation pattern is heritable after cell division (Illustration). Therefore, DNA methylation plays an important role in cell differentiation during development.

POSSIBLE QUESTIONS

PART B

1) What is lac operon?
2) What is trp operon?
3) What is DNA methylation?
4) Define histone acetylation?
5) Define Chromatin structure

PART C

1) Write short notes on Principles of transcriptional regulation?
2) Explain in detail about the mechanism of Protein synthesis?
3) Explain in detail about Mating type switching in budding yeast?
4) Write short notes on Important Inhibitors of Protein Synthesis?
5) Write short notes on Sporulation in Bacillus?

KARPAGAM ACADEMY OF HIGHER EDUCATION

CLASS: III B.Sc MB COURSE NAME: MOLECULAR BIOLOGY

COURSE CODE: 17MBU603B UNIT: V BATCH-2017-2020

Sl. No	Question	Option A	Option B	Option C	Option D	Correct Ans
1	Petite yeast mutants are unable to grow on media containing	Only nitrogen sources	Only Mineral sources	Only Lipid Source	Only fermentable carbon sources	Only non-fermentable carbon sources
2	Meiosis in yeast life cycle leads in forming	Sporangium	Fragmented mycelium	Ascus spores	Endospores	Ascus spores
3	In yeast, 22% of the genome is made up of _____ DNA	Polygenic	Monogenic	Intragenic	Intergenic	Intergenic
4	Genetic maps of chromosome are based on the frequencies of _____	Reproducible mitotic product	Reproducible meiotic product	Recombinant mitotic product	Recombinant meiotic product	Recombinant meiotic product
5	Gene conversion can be either allelic or _____	Allergic	Ectopic	Endemic	Endopic	Ectopic
6	Gene conversion in yeast may be important in maintaining_____within families	Sequence homogeneity	Sequence heterogenecity	Sequence array	Sequence hologenecity	Sequence homogeneity
7	The leu2- strain carries a mutation that inactivates leu2 gene which codes for _____	Endonuclease	Beta isoprophylmalate dehydrogenase	Helicase	Ligase	Beta isoprophylmalate dehydrogenase
8	The percentage of recombinant meiotic product is one half of the percentage of _____	Second division	Fourth division	Primary division	Tetrad	Tetrad
9	A new genetic nomenclature for S. cerevisiae transposon	jumping element	Ty elements	Tx elements	Ta element	Ty elements

10	Among haploid and diploid vegetative cells of yeast ,which is mainly used for genetic mapping	Haploid and diploid	Diploid	Haploid	Triploid	Haploid
11	His 3 is an _____ protein.	Transducer	Indicator	Selector	Repressor	Indicator
12	If single crossover occurs between a- & b- then tetra type ___ results	A	B	T	U	T
13	Pleiotrophy is common in _____	No organism	All organism	Higher organism	Lower organism	Higher organism
14	Process by which one DNA sequence replaces a homologous sequence	Gene mutation	Gene repulsion	Gene transtition	Gene conversion	Gene conversion
15	The term genetic linkage was given by ----------	Morgan	Meischer	Wilkins	Mendel	Morgan
16	The_____ to a particular chromosome is the first step in genetic mapping	Elemination of mutated type	Localization of wild type	Localization of mutation	Elimination of wild type	Localization of mutation
17	Yeast are _____	Multicellular fungi	Dicellular	Acellular	Unicellular fungi	Unicellular fungi
18	Yeast genome has _____ introns	233	236	323	326	233
19	Yeast has two mating types, _____ , which show primitive aspects of sex differentiation	a & b	a & α	b & β	α & β	a & α
20	Yeasts fail to grow on	Glucose	Lactose	Maltose	Trehalose	Lactose
21	Sample of ascospores is spread on to the agar medium without leucine and survival was tested	ELISA	PCR	Random spore analysis	RPR	Random spore analysis

	using					
22	_____ test is used to determine which gene(s) are defective in petite yeast mutants	Complementation test	Complement fixation test	Completed test	Conjugation test	Complementation test
23	_____ contains two types of spores of same parental genotype	Parental haploid	Parental diploid	Non-parental ditype	Parental ditype	Parental diploid
24	_____determines the number of crossover events and give correct map distance	Genetic analysis	Physical analysis	Random spore analysis	Tetrad analysis	Tetrad analysis
25	_____developed an algebriac method to determine the consequence of various number of exchanging	Klebs & Loeffler	Pastuer & Winogradsky	Shult & Lindegrin	Klug & Cumming	Shult & Lindegrin
26	Which contains four different ascospores , one of each genotypes	Tetratype	Parental type	Parental ditype	Non-parental ditype	Tetratype
27	The developing haploid spores are enclosed in a membranous structure called_____	Spores	Pycus	Zygote	Ascus	Ascus
28	Method available for locating mutation in *Neurospora crassa* is	Co-segregation	Co-opression	Independent assortment	Tetrad analysis	Co-segregation
29	In Transcription of yeast genome,_____ makes all RNA to serve as mRNA	DNA polymerase	RNA polymerase	RNA pol I	RNA pol II	RNA pol II
30	Intermediate compound responsible for red color of adenine-requiring yeast petite mutants	Aminoimidazole	Aminoimidazole ribonucleotide	Aminoimidazoleribotide	Aminoimmunoribotide	Aminoimidazoleribotide

31	Arg^+ is	A strain not requiring aspargine	A strain requiring aspargine	A strain not requiring arginine	A strain requiring arginine	A strain not requiring arginine
32	Cells which contains single copy of chromosome is called _____	Triploid	Haploid and diploid	Diploid	Haploid	Haploid
33	_____ intial products of meiosis forms two identical spore	4	8	16	32	4
34	A feature of petite is the occurrence of _____ from the circular mitochondrial genome	Insertion	Excertion	Addition	Deletion	Deletion
35	Generation time of yeast takes place at _____	2hrs 30 min	3 hrs	60 min	30 min	2hrs 30 min
36	Give full form for NPD	Non-parental dikaryon	Non-performing data	Non-parental data	Non-parental ditype	Non-parental ditype
37	HFT is	High frequency transducing	High frequency transcribing	Height frequency transducing	Heavy frequency transducing	High frequency transducing
38	Significant feature of sex determination in Drosophila is the presence of abnormal flies called _____	Gyno variants	Gynocoid	Gyrates	Gynondromorphs	Gynondromorphs
39	Tetrad showing second division segregation has _____ products	Four meiotic	Two Meiotic	Four mitotic	Two mitotic	four meiotic
40	_____ refers to a genetic changes in different genomes of same cell	Trans type	Same type	Cis type	Different type	Cis type

41	_____distance in map unit is only half the percentage of second division segregated tetrads	Tetromere	Telomere	Centromer	Primer	Telomere
42	An estimation of map between a- & b- can be obtained from number of recombinant ascospore detected by using _____	Genetic analysis	Physical analysis	Random spore analysis	Tetrad analysis	Random spore analysis
43	Common model organism in studying unicellular eukaryotes/budding yeast	*Saccharomyces cerevisiae*	*Cryptococcus Neoformans*	*Candida albicans*	*Pitchia pastoris*	*Saccharomyces cerevisiae*
44	Gene conversion is the _____ transfer of information from one DNA duplex to another	Reciprocal	Non-reciprocal	Direct	Indirect	Non-reciprocal
45	Haploid to diploid phase in yeast is brought about by	Fusion of opposite mating types	Fusion of positive mating types	Fusion of negative mating types	Fusion of opposite genes	Fusion of opposite mating types
46	In *Drosophila*, the meiotic recombination occurs only in _____	Both Female and Male	Female	Male	None	Female
47	Map distance is equal to the percentage of	Reproducible mitotic product	Recombinant mitotic product	Reproducible meiotic product	Recombinant meiotic product	Recombinant meiotic product
48	Petite phenotypes caused by the absence of, or mutations in, mitochondrial DNA are termed as	Cytoplasmic petites	Energy deficient petites	Mitochondrial petites	Chrosomal petites	Cytoplasmic petites
49	Random spore analysis gives _____ values	Approximate	Accurate	Null	Partial	Approximate
50	Sex linkage was explained by _____	Morgan	Mendel	Primrose	Pastuer	Morgan

51	The ascus burst releasing the ascospores, each of which germinates and divides by mitosis to produce new _____	Vegetative cells	Spores	Ascus	Zygote	Vegetative cells
52	The general mapping function of Haldane is based on _____	Haldane distribution	Poisson distribution	Twart	Switz	Poisson distribution
53	The repair of double-strand gaps is an efficient process in yeast known to be	Seggregation	Assortment	SOS repair	Crossing over	Crossing over
54	Well characterized Baker's yeast has __ cell	Irregular	Round	Spheroid	Disc shaped	Spheroid
55	Which of the following is used in density gradient centrifugation?	Glucose	Sucrose	Fructose	Agarose	Sucrose
56	Yeast genome is _____	12,520 kb	1,252 kb	1,02,520 kb	15,052 kb	12,520 kb
57	When mutation in single gene affect more than one trait is called _____	Parental genes	Pleiomorphic genes	Priogenic genes	Pleiotrophic genes	Pleiotrophic genes
58	The binding of two DNA helices through X-shaped junction called _____	Polytron	Cholistron	Diptron	Cistron	Cistron
59	Small pieces of DNA that can insert themselves into chromosomes are known	Plasmid	Transposon	Cosmid	Artificial chrosome	Transposon
60	Recombination does not only occur during meiosis, but also as a mechanism for	Repair of single-strand breaks	Repair of double-strand breaks	Repair of proteins	Repair of RNA	Repair of double-strand breaks