Semester IV

17BCU402 GENE EXPRESSION AND REGULATION 4H-4C

SCOPE

To enlight the students with gene structure and their conversion to the macromolecule RNA and protein

.

OBJECTIVES

To explain gene structure, replication, transcription, translation recombination mechanism of DNA.

Unit 1

Biosynthesis of RNA in prokaryotes

RNA polymerases, transcription cycle in bacteria, sigma factor, bacterial promoters, identification of DNA binding sites by DNA footprinting, the three stages of RNA synthesis, initiation, elongation and termination, rho-dependent and rho-independent termination. Inhibitors of transcription and applications as anti-microbial drugs.

Unit 2

Biosynthesis of RNA in eukaryotes

Comparison between prokaryotic and eukaryotic transcription. Transcription by RNA polymerase II, RNA polymerase II core promoters, general transcription factors, various types of RNA processing, transcription by RNA polymerase I and III. Inhibitors of eukaryotic transcription and their applications. Comparison of fidelity of transcription and replication.**RNA splicing-** Chemistry of RNA splicing, the spliceosome machinery, splicing pathways, group I and group II introns, alternative splicing, exon shuffling, RNA editing.

Unit 3

Biosynthesis of proteins

The genetic code-Degeneracy of the genetic code, wobble in the anticodon, features of the genetic code, nearly universal code. **Biosynthesis of proteins-** Messenger RNA, transfer RNA, attachment of amino acids to tRNA, the ribosome - initiation, elongation and

termination of translation, regulation of translation. Comparison of prokaryotic and eukaryotic protein synthesis. Use of antibiotics in understanding protein synthesis and applications in medicine.**Protein targeting and degradation -** Post translational modifications, glycosylation, signal sequences for nuclear transport, bacterial signal sequences, import of proteins by receptor mediated endocytosis, specialized systems for protein degradation.

Unit 4

Regulation of gene expression in prokaryotes

Principles of gene regulation, negative and positive regulation, concept of operons, regulatory proteins, activators, repressors, DNA binding domains, regulation of lac operon and trp operon, induction of SOS response, synthesis of ribosomal proteins, regulation by genetic recombination, transcriptional regulation in λ bacteriophage.

Unit 5

Regulation of gene expression in eukaryotes

Heterochromatin, euchromatin, chromatin remodeling, regulation of galactose metabolism in yeast, regulation by phosphorylation of nuclear transcription factors, regulatory RNAs, riboswitches, RNA interference, synthesis and function of miRNA molecules, phosphorylation of nuclear transcription factors.

REFERENCES

Nelson, D.L. and Cox, M.M., (2013). Lehninger: Principles of Biochemistry 6th ed., W.H. Freeman & Company (New York), ISBN:13: 978-1-4292-3414-6 / ISBN:10-14641-0962-1.

Watson, J.D., Baker, T.A., Bell, S.P., Gann, A., Levine, M. and Losick, R., (2008). Molecular Biology of the Gene 6th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor (New York), ISBN:0-321-50781 / ISBN: 978-0-321-50781-5.

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University Established Under Section 3 of UGC Act 1956)
Coimbatore – 641 021.

LECTURE PLAN
DEPARTMENT OF BIOCHEMISTRY

STAFF NAME: Dr. Rajesh Pandiyan
SUBJECT NAME: Gene Expression and Regulation SUB.CODE:18BCU302
SEMESTER: IV CLASS: II B.Sc (BC)

Sl. No	Duration of Period	Topics to be Covered	Support material
UNIT I Biosynthesis of RNA in prokaryotes			
1	1	Introduction to Biosynthesis of RNA in prokaryotes and RNA polymerases	
2	1	Transcription cycle in bacteria	
3	1	Sigma factor, bacterial promoters, identification of DNA binding sites by DNA footprinting	
4	1	The stages of RNA synthesis-initiation, elongation and termination	
5	1	Inhibitors of transcription	
6	1	Applications as anti microbial drugs	
7	1	Revision and possible questions discussion of Unit I	
		Total no of hours planned for UNIT I = 07	
Unit II: Biosynthesis of RNA in eukaryotes			
1	1	Comparison between prokaryotic and eukaryotic transcription	
2	1	Transcription by RNA polymerase II, RNA polymerase II core promoters	
3	1	General transcription factors, various types of RNA processing	
4	1	Transcription by RNA polymerase I and III	
5	1	Inhibitors of eukaryotic transcription and their applications, Comparison of fidelity of transcription and replication	
6	1	RNA splicing-chemistry of RNA splicing, the spliceosome machinery, splicing pathways, group I and II introns	
7	1	Alternative splicing, exon shuffling, RNA editing	
8	1	Revision and possible questions discussion of Unit II	

		Total no of hours planned for UNIT II = 08	
Unit 3: Biosynthesis of proteins			
1	1	The genetic code-Degeneracy of the genetic code, wobble in the anticodon, features of the genetic code, nearly universal code.	
2	1	Biosynthesis of proteins-Messenger RNA, transfer RNA, attachment of aminoacids to tRNA, the ribosome.	
3	1	Stages of translation-Initiation and elongation and termination of translation	
4	1	Stages of translation-Termination of translation	
5	1	Regulation of translation. Comparison of prokaryotic and eukaryotic protein synthesis. Use of antibiotics in understanding protein synthesis and applications in medicine	
6	1	Protein targeting and degradation - Post translational modifications, glycosylation, signal sequences for nuclear transport,	
7	1	Bacterial signal sequences, import of proteins by receptor mediated endocytosis, specialized systems for protein degradation.	
8	1	Revision and possible questions discussion of Unit III	
		Total no of hours planned for UNIT III = 08	
Unit 4: Regulation of gene expression in prokaryotes			
1	1	Principles of gene regulation, negative and positive regulation,	
2	1	Concept of operons, regulatory proteins, activators, repressors, DNA binding domains,	
3	1	Regulation of Lac operon	
4	1	Regulation of Trp operon	
5	1	Induction of SOS response,	
6	1	Synthesis of ribosomal proteins, regulation by genetic recombination	
7	1	Transcriptional regulation in λ bacteriophage.	
8	1	Revision and possible questions discussion on Unit 4	
		Total no of hours planned for UNIT IV = 08	
Unit 5: Regulation of gene expression in eukaryotes			
1	1	Heterochromatin, euchromatin.	
2	1	chromatin remodeling.	
3	1	regulation of galactose metabolism in yeast.	
4	1	Regulation by phosphorylation of nuclear transcription factors.	
5	1	Regulatory RNAs, riboswitches.	
6	1	RNA interference, synthesis and function of miRNA molecules.	

7	1	Phosphorylation of nuclear transcription factors.	
8	1	Discussion on previous year end semester examination question paper 1	
9	1	Discussion on previous year end semester examination question paper 2	
		Total no of hours planned for UNIT V = 09	
Total number of hours planned for this course: 40			

Support Materials

R1: Nelson, D.L. and Cox, M.M., (2013). Lehninger: Principles of Biochemistry 6th ed., W.H. Freeman & Company (New York), ISBN:13: 978-1-4292-3414-6 / ISBN:10-14641-0962-1.

R2: Watson, J.D., Baker, T.A., Bell, S.P., Gann, A., Levine, M. and Losick, R., (2008). Molecular Biology of the Gene 6th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor (New York), ISBN:0-321-50781 / ISBN: 978-0-321-50781-5.

Unit 1

Biosynthesis of RNA in prokaryotes

RNA polymerases, transcription cycle in bacteria, sigma factor, bacterial promoters, identification of DNA binding sites by DNA footprinting, the three stages of RNA synthesis, initiation, elongation and termination, rho-dependent and rho-independent termination. Inhibitors of transcription and applications as anti-microbial drugs.

Introduction

Transcription is the process in which messenger RNA transcripts of genetic material in prokaryotes are produced, to be translated for the production of proteins. Prokaryotic transcription occurs in the cytoplasm alongside translation. Unlike in eukaryotes, prokaryotic transcription and translation can occur simultaneously. This is impossible in eukaryotes, where transcription occurs in a membrane-bound nucleus while translation occurs outside the nucleus in the cytoplasm. In prokaryotes genetic material is not enclosed in a membrane-enclosed nucleus and has access to ribosomes in the cytoplasm.

General feature of transcription

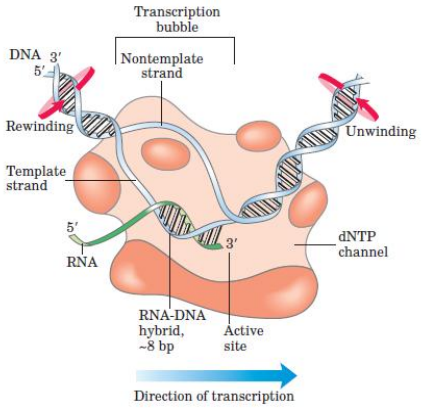
Direction : 5' → 3'

Enzyme : RNA polymerase that does not require primer but require a template

Templte strand: One of the double stranded DNA strand acts as a templte called tempalte strand or Non coding strand

Structure and functions of RNA polymerase

The RNA polymerase has two α ubunits, one β subunit and a β' subunit, a ω subunit and a σ subunit in its structure.



Functions:

- **DNA-dependent RNA polymerase** requires, in addition to a DNA template, all four ribonucleoside 5’ triphosphates (ATP, GTP, UTP, and CTP) as precursors of the nucleotide units of RNA, as well as Mg^{2+}.
- The protein also binds one Zn^{2+}.
- RNA polymerase elongates an RNA strand by adding ribonucleotide units to the 3’ hydroxyl end, building RNA in the 5’ →3’ direction. The 3’ hydroxyl group acts as a nucleophile, attacking the 5’phosphate of the incoming ribonucleoside triphosphate.
- RNA polymerase requires DNA for activity and is most active when bound to a double-stranded DNA. only one of the two DNA strands serves as a template.
- The template DNA strand is copied in the 3’ →5’ direction (antiparallel to the new RNA strand),
- RNA polymerase does not require a primer to initiate synthesis. Initiation occurs when RNA polymerase binds at specific DNA sequences called promoters.
- The moving RNA polymerase generates waves of positive supercoils ahead of the transcriptionbubble and negative supercoils behind of the transcription bubble.

Stages of prokaryotic transcription

The transcription process consist of three stages namely inititation, elongation and termination.

Initiation

All transcription starts at a point called promoter. Prokaryotic promotershave TATA Box (-10) and -35 region

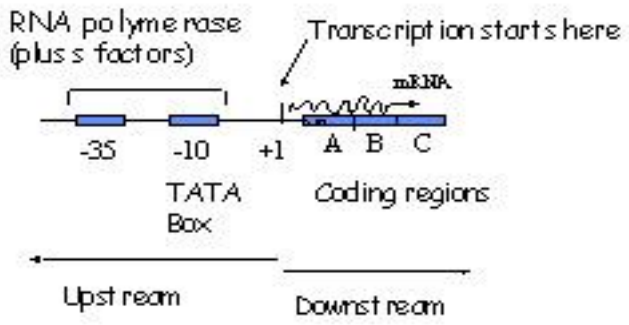
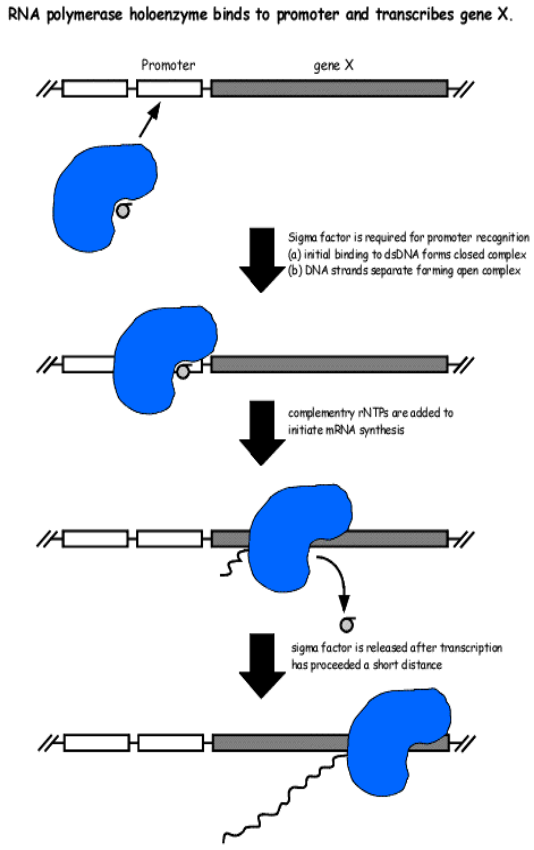


Fig: Prokaryotic promoters

The consensus -35 and -10 sequences for RNA polymerase and s^{70} binding are:

-35	-10
TTGACAT	TATAAT



The following steps occur, in order, for transcription initiation:

- RNA polymerase (RNAP) binds to one of several specificity factors, σ, to form a holoenzyme. In this form, it can recognize and bind to specific promoter regions in the DNA. The -35 region and the -10 ("Pribnow box") region comprise the basic prokaryotic promoter, and |T| stands for the terminator. The DNA on the template strand between the +1 site and the terminator is transcribed into RNA, which is then

translated into protein.At this stage, the DNA is double-stranded ("closed"). This holoenzyme/wound-DNA structure is referred to as the closed complex.

- The DNA is unwound and becomes single-stranded ("open") in the vicinity of the initiation site (defined as +1). This holoenzyme/unwound-DNA structure is called the open complex.

- The RNA polymerase transcribes the DNA (the beta subunit initiates the synthesis), but produces about 10 abortive (short, non-productive) transcripts which are unable to leave the RNA polymerase because the exit channel is blocked by the σ-factor.

- The σ-factor eventually dissociates from the holoenzyme, and elongation proceeds.

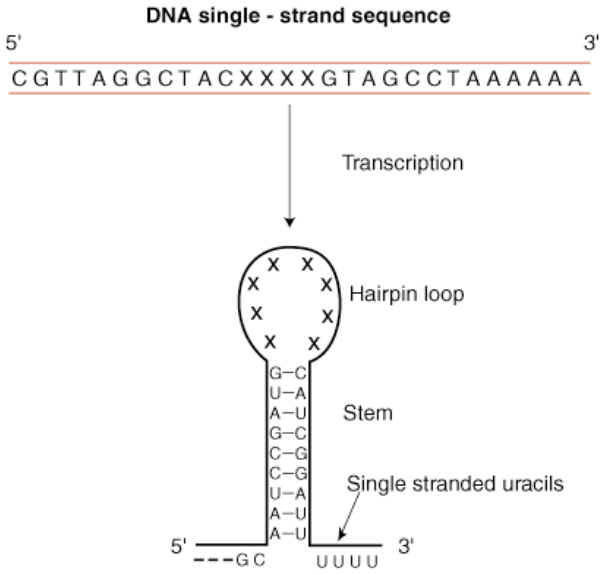
Elongation

One strand of the DNA, the template strand (or noncoding strand), is used as a template for RNA synthesis. As transcription proceeds, RNA polymerase traverses the template strand and uses base pairing complementarity with the DNA template to create an RNA copy. Although RNA polymerase traverses the template strand from 3' → 5', the coding (non-template) strand and newly-formed RNA can also be used as reference points, so transcription can be described as occurring 5' → 3'. The movement of RNA polymerase is called as inch worm movement where the back end of polymerase comprises for every base added, but front end does not move. After adding around 10 nucleotides, the front end jump several basepairs in forward direction.This produces an RNA molecule from 5' → 3', an exact copy of the coding strand (except that thymines are replaced with uracils, and the nucleotides are composed of a ribose (5-carbon) sugar where DNA has deoxyribose (one less oxygen atom) in its sugar-phosphate backbone).

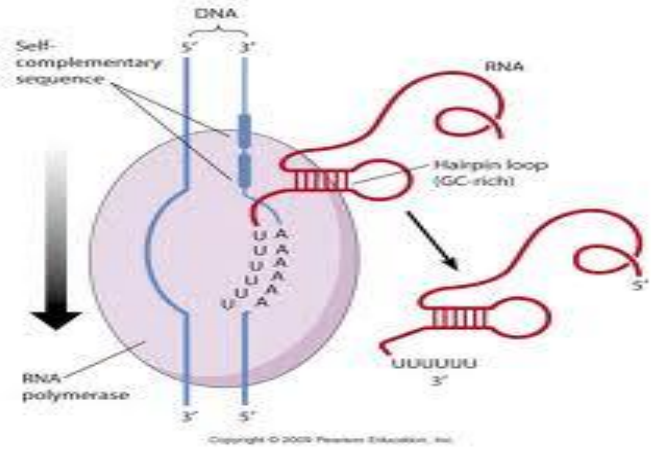
Termination

Two termination mechanisms are well known:

- Intrinsic termination (also called Rho-independent transcription termination) involves terminator sequences within the RNA that signal the RNA polymerase to stop. The terminator sequence is usually a palindromic sequence that forms a stem-loop hairpin structure that leads to the dissociation of the RNAP from the DNA template. Figures: Hairpin loop formation

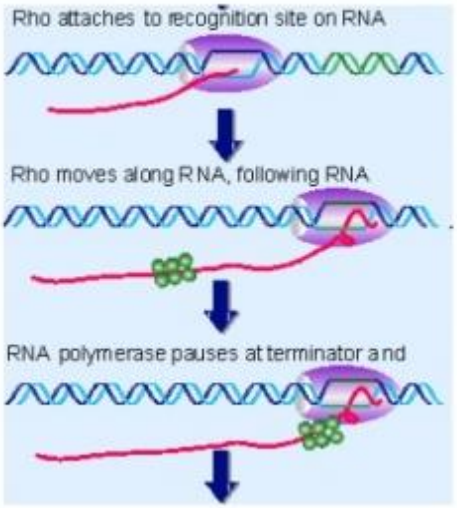


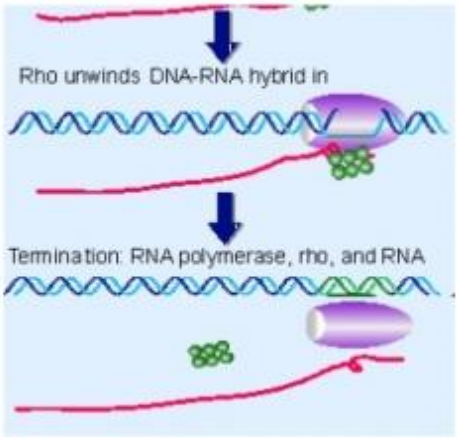
RNA Sequence



Rho-dependent termination uses a termination factor called ρ factor(rho factor) which is a protein to stop RNA synthesis at specific sites. This protein binds at a rho utilisation site on the nascent RNA strand and runs along the mRNA towards the RNAP. A stem loop structure upstream of the terminator region pauses the RNAP, when ρ-factor reaches the RNAP, it causes RNAP to dissociate from the DNA, terminating transcription.

Prokaryotic Transcription: Rho Dependent Termination





II BSc Biochemistry -Fourth Semester
GENE EXPRESSION AND REGULATION (17BCU402)
Unit I

S.No.	Questions	Opt 1	Opt 2	Opt 3	Opt 4	Answer
1	Prokaryotic transcription initiation is inhibited by _____	. Rifamycin	Rifampicin	ricin	abrin	Rifampicin
2	The strand of DNA that directs the synthesis of mRNA is	sense strand	leading strand	antisense strand	lagging strand	sense strand
3	For initiation ___subunit of RNA pol required	alpha	beta	gamma	sigma	sigma
4	Pribnow box consists of the following 6 nucleotide bases _____	TATAAT	TATATA	TTAAAT	TAATAT	TATAAT
5	Rho protein is involved in _____	transcription initiation	transcription elongation	transcription termination at some terminators	. transcription termination at all terminators	transcription termination at some terminators
6	Rho factor is a ____	Dimer	Trimer	hexamer	Tetramer	hexamer
7	Transcription charecterised by all except _____	one strand of DNA serve as template	synthesis of RNA	Synthesize proteins	Participation of RNA polymerase	Synthesize proteins
8	Termination of transcription is brought about by _____	sigma factor	RNA polymerase	rho factor	TATA box	rho factor
9	The following are associated with transcription except ____	-35 sequence	pribnow box	promoter	spacer	spacer
10	Rho factor catalyses the unwinding of RNA-DNA double helices and _____	Hydrolysis of ATP	RNA-RNA double helices	Migrates along the RNA in 3'-5' direction	Migrates along the RNA in 5'-3' direction	Hydrolysis of ATP
11	Non template strand is otherwise called as _____	template strand	coding strand	non coding strand	sense strand	coding strand
12	Promoters are recognized by _____	alpha sub unit	gamma subunit	beta subunit	sigma subunit	sigma subunit
13	When a number of genes are transcribed as one mRNA, the mRNA is said to be ___	multimeric	. polymeric	polycistronic	polyclonal	polycistronic
14	The mature 5'end of tRNA in E.coli formed as a result of endonucleolytic cleavage by	RNase P	RNAse D	RNase M5	RNase M13	RNase P
15	An E.Coli strain lacking DNA polymerase I would be deficient in DNA	Repair	Splicing	Methylation	Transcripition	Repair
16	The RNA polymerase of E.coli complex holoenzyme composed of 5 polypeptide subunits	2α,2β and one sigma factor	2α 2β and one rho factor	2α 1β,1β and one sigma factor	2α,1β,1β and one rho factor	2α 1β,1β and one sigma factor
17	During the course of transcriptuon or replication for the addition each nucleotide to the growing chain the following moiety is released	phosphate	pyrophosphate	nitrogenous base	nucleotide	pyrophosphate
18	Transcription charecterised by all except	one strand of DNA serve as template	synthesis of RNA	Synthesise proteins	Participation of RNA polymerase	Synthesise proteins
19	RNA polymerase used for the transc	rRNA	DNA	RNA	mRNA	DNA
20	Which of the following statements is not true?	Transcription but not translation is regulated in bacteria	Bacterial non coding RNAs mus t be processed by cleavage to form functional molecules	The regulation of the transcription initiation complex is a key feature of gene expression control in eukaryotes	Linking of exons must precede translation in eukaryotes	Transcription but not translation is regulated in bacteria
21	Regarding RNA polymerase in E.Coli which of the following describes the difference between the holo enyzyme and the core enzyme?	The holoenzyme is used to synthesize mRNA, while the core enzyme synthesizes noncoding RNAs	The holoenzyme consists of five subunits including σ, while the core enzyme lacks σ	The core enzyme and not the holoenzyme is required for the initation of RNA synthesis	The holoenzyme binds to DNA upstream of the promoter while the core enzyme binds to dounwstream of the promoter	The holoenzyme consists of five subunits including σ, while the core enzyme lacks σ
22	RNA polymerase II transcribes all of	rRNAs	proteins	snRNAs	miRNAs	rRNAs
23	The recognition sequence to whiich RNA polymerase binds at the initiation of transcription is found----	downstream of the promoter target sequence	b)upstream of the gene to be transcribed	withinthe first intron	downstream of the transcription bubble	upstream of the gene to be transcribed
24	Which of the following best describes a promoter?	An element that promotes termination of transcription	A regualatory protein that accelerates mRNA turnover	A specific target sequence to which RNA polymeraase binds	An extracellular environmental inducer that controls gene expression	A specific target sequence to which RNA polymeraase binds
25	Some eukaryotic promoters contain an element positioned around nucleotide +1 called----	the TATA box	the pribnow box	an initiator (Inr) sequence	an inverted repeat	an initiator Inr sequence
26	"Null" genes are characterized by their-----	transcription by RNA polymerase I and II	possession of neither TATA box nor an Inr sequence	Invariable start position of transcription	inability to be transcribed	possession of neither TATA box nor an Inr sequence
27	The conversion of a clsoed rpomoter complex to an open promoter complex in bacteria requires------	the activity of alternative promoters	hydrogen bond breakage of base pairs around the initiation site	a G-C rich sequence around to +1 site	strong interaction between the core enzyme and the -10 box	hydrogen bond breakage of base pairs around the initiation site
28	An improtant difference between the initiation of prokarytoic and eukaryotic transcription of protein coding genes is that--------	only prokaryotic genes use upstream sequence	only eukaryotic RNA polymerase possess a σ subunit	eukaryotic promoters indirectly recognize core promoter sequences	d)only prokaryotes cells alter the rate of transcription	eukaryotic promoters indirectly recognize core promoter sequences
29	Each time a ribonucleotide is added to the elongating RNA molecule during transcription a-----is released	deoxy ribonucleotide	sugar-base component	c)pyrophosphate	phosphate	pyrophosphate
30	The 5'end of mRNAs made by RNA polymerase II possess a distinct cap structure depicted in short-hand notation as-----	7-MeGppNpN	7-MeGpppNpN	pppN-7MeGppppNpN	pppNpN	7-MeGpppNpN
31	The formation of hairpin loops at the 3' end of prokaryotic RNAs during transcription results in------	stabilization of DNA:RNA pairing promoting further elongation of the transcript	activation of poly (polymerase	a reduction in base pairing between the template strand and the RNA transcript	Rho-dependent activation of transcription	a reduction in base pairing between the template strand and the RNA transcript
32	Which of the following statements regarding the poly (tail of eukaryotic RNA is incorrect?	The poly (tail is located at the 3, end	The poly (tract is transcribed from the DNA template	poly (tails are found on transcripts made by RNA polyemrase II	Tissue specific gene expression patterns soemtimes correlate with laternative polyadenylation	The poly (tract is transcribed from the DNA template
33	For initiation ___subunit of RNA pol required	alpha	beta	gamma	sigma	sigma
34	The RNA polymerase of E.coli complex holoenzyme composed of 5 polypeptide subunits	2a,2b and one sigma factor	2a 2b and one rho factor	2a 1b,1b and one sigma factor	2a,1b,1b and one rho factor	2a 1b,1b and one sigma factor
35	Pribnow box consists of the following 6 nucleotide bases	TATAAT	TATATA	TTAAAT	TAATAT	TATAAT

36	RNA polymerase utilizes the following RNA triphosphates for the formation of RNA	ATP, GTP, CTP and TTP	ATP, GTP, CTP and UTP	ATP, CTP, UTP and UTP	All	ATP, GTP, CTP and UTP
37	During the course of transcriptuon or replication for the addition each nucleotide to the growing chain the following moiety is released	phosphate	pyrophosphate	nitrogenous base	nucleotide	pyrophosphate
38	Transcription charecterised by all except	one strand of DNA serve as template	synthesis of RNA	Synthesise proteins	Participation of RNA polymerase	Synthesise proteins
39	RNA self splicing was discovered by	Watson & Crick	McClintock	Sanger	Thomas Cech	Thomas Cech
40	Termination of transcription is brought about by	sigma factor	RNA polymerase	rho factor	TATA box	rho factor
41	In eukaryotic cells, the synthesis of mRNA is carried out by	RNA polymerase I	RNA polymerase II	RNA polymerase III	DNA polymerase	RNA polymerase II
42	Ribosomal RNA perform____ all fucntion except	interact with aminoacidsequence of tRNA	Read information in mRNA	Synthesise proteins	Bring the aminoacids to the site of protein synthesis	Bring the aminoacids to the site of protein synthesis
43	The enzyme responsible for reverse transcription	DNA dependent DNA polymerase	DNA dependent RNA polymerase	RNA dependent DNA polymerase	RNA dependent RNA polymerase	RNA dependent DNA polymerase
44	The mature 5'end of tRNA in E.coli formed as a result of endonucleolytic cleavage by	RNase P	RNAse D	RNase M5	RNase M13	RNase P
45	Transfer RNA has all except	Clover leaf structure	Anticodon arm	Minor bases in addition to normal bases A, U, G and C	proteins	proteins
46	snRNAs are involved in the eukaryotic	Transcription	Splicing of mRNA	transfer of amino acids	splicing of tRNA	Splicing of mRNA
47	RNA polymerases present in eukaryotes are	RNA pol I	RNA pol II	RNA pol III	RNA pol I,II and III	RNA pol I,II and III
48	Rho factor is a	Dimer	Trimer	hexamer	Tetramer	hexamer
49	The tRNA TYC arm contains	Thiamin , pseudourindine and cysteine	Thiamin , uracil and cysteine	Thiamin , pseudourindine and cytosine	Adenine, uracil and cysteine	Thiamin , pseudourindine and cytosine
50	Transcription initiation in bacteria is inhibited by	Streptolydigin	Puromycin	Rifampicin	Tetracycline	Rifampicin
51	The hexa nucleotide AAU AAA provides the signal for	Capping	Generation of eukaryotic 3'- poly A tail	Initiation of translation	Peptide bond formation	Generation of eukaryotic 3'- poly A tail
52	Select the correct statement about spliceosomes	convert pre mRNA into mature RNA	composed of only RNA	composed of only Protein	splice RNA in prokaryotic and eukaryotic cells.	splice RNA in prokaryotic and eukaryotic cells.
53	. Rho protein is involved in	transcription initiation	transcription elongation	transcription termination at some terminators	. transcription termination at all terminators	transcription termination at some terminators
54	. tRNA specifically reacts with	mRNA	nuclear RNA	. amino acyl adenylates	ribosomal RNA	. amino acyl adenylates
55	. Pseudo uridine is found in	mRNA	. snRNA	. rRNA	. tRNA	. tRNA
56	The acceptor arm in tRNA consists of base pair stem terminates in the sequence	CCA (5'-3')	CAA (5'-3')	. CCA (3'-5')	AAC (3'-5')	CCA (5'-3')
57	sn RNA is involved in	. DNA replicatiopn	Ribosome assembly	. RNA splicing	Initiation of translation	. RNA splicing

Unit 2

Biosynthesis of RNA in eukaryotes

Comparison between prokaryotic and eukaryotic transcription. Transcription by RNA polymerase II, RNA polymerase II core promoters, general transcription factors, various types of RNA processing, transcription by RNA polymerase I and III. Inhibitors of eukaryotic transcription and their applications. Comparison of fidelity of transcription and replication.**RNA splicing-** Chemistry of RNA splicing, the spliceosome machinery, splicing pathways, group I and group II introns, alternative splicing, exon shuffling, RNA editing.

Eukaryotic transcription

Eukaryotic transcription is more complex than prokaryotic transcription and, until recently, it has seemed that every eukaryotic gene was unique requiring its own transcription machinery.

There are three distinct RNA polymerases in a **eukaryotic cell nucleus** which define the three major classes of eukaryotic transcription unit:

polymerase	location	type of RNA transcribed	sensitivity to □-amanitin§
I	nucleus/nucleolus	rRNA (except for 5S rRNA)	resistant
II	nucleus	hnRNA (i.e. pre-mRNA)	very sensitive
III	nucleus	small RNA such as tRNA and 5S rRNA	moderately sensitive

There may be as many as 14 subunits in an eukaryotic RNA polymerase; the total molecular weight is typically 500-700 kD.

Eukaryotic RNA polymerases cannot find or bind to a promoter by themselves. Each requires the binding of **assembly factors** and a **positional factor** to locate the promoter

and to orient the polymerase correctly. As we will see, the positional factor is the same in all cases.

Eukaryotic promoter

Along with **TATA box** it contain **promoter proximal region** and **enhancer** sequences

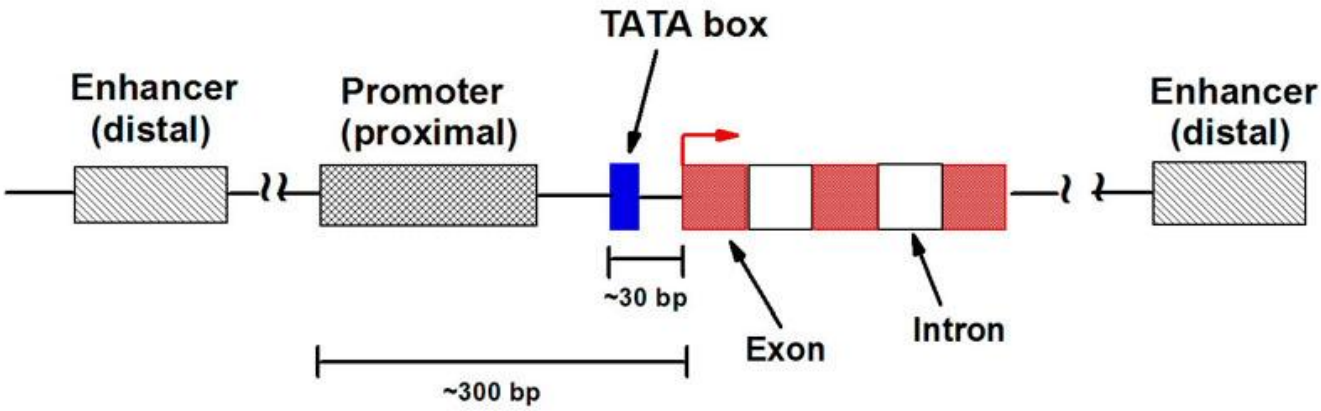


Fig: General structure of eukaryotic promoter

1. Class I Transcriptional Units

Class I genes or transcriptional units are transcribed by RNA polymerase I in the nucleolus. The best-studied examples are the rRNA transcription units:

Each transcription unit consists of 3 rRNA genes: 18S, 5.8S, and 28S; and each unit is separated by a nontranscribed spacer. Eukaryotic nucleoli typically have many hundreds of copies of these transcription units tandemly arranged.

The enzyme RNA polymerase I is a complex of 13 subunits.

The promoter

The **CORE** promoter region is located from **-31** to **+6** around the transcription startpoint. Another sequence further upstream, called the **upstream control element** (**UCE**), located from **-187** to **-107** is also required for efficient transcription.

Both elements are closely related; there is approximately 85% sequence identity between them. These elements are also unusual in that they are GC-rich. In general, sequences around the start-point of transcription tend to be AT-rich so that melting of the DNA duplex is easier.

Assembly of a transcriptional complex

Two additional transcription factors are known to be required to assist RNA polymerase I.**UBF1-**is a single polypeptide which binds to the upstream control element (**UCE**) and to the **CORE** promoter. **UBF1** recognizes a GC-rich sequence within these elements. **UBF1** is an **assembly factor**.

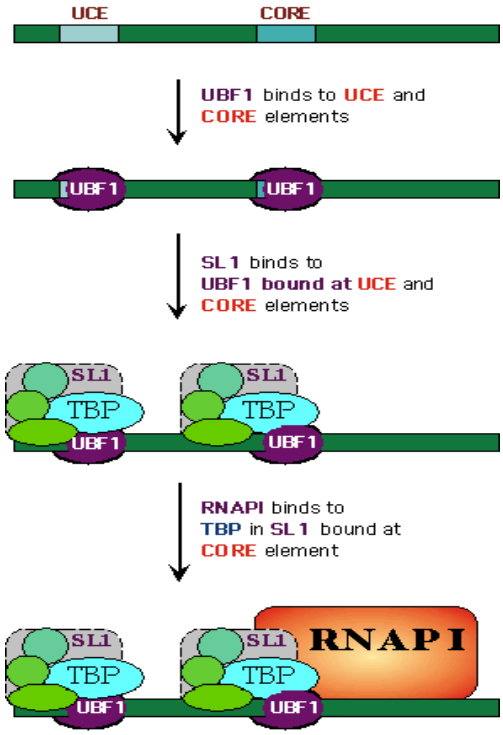


Fig: Synthesis of rRNA in eukaryotes

SL1-binds to **UBF1**. It consists of 4 proteins, one of which is TATA-box binding protein (**TBP**). **TBP** is required for the assembly of a transcriptional complex in all 3 classes of eukaryotic transcription unit. **SL1** is a **positional factor** - it targets RNA polymerase at the promoter so that it initiates transcription in the correct place.

Once **UBF1** and **SL1** have formed a complex, **RNAP I** binds to the **CORE** promoter to initiate transcription: **(Fig)**

2.**Class II Transcription Units**

All genes that are transcribed and expressed via **mRNA** are transcribed by **RNA polymerase II**. Until recently, it was common to think of eukaryotic transcription (and particularly mRNA synthesis) as taking place in discrete steps: transcription, capping, tailing, splicing and export from the nucleus for translation. The contemporary view of eukaryotic geen expression entails simultaneous transcription and processing. Recent discoveries have revealed that many of the protein factors required for these individual steps do, in fact, interact with one another. This makes sense for it allows the cell to coordinate and regulate the complete process more efficiently.

The two images below are from a recent review article. Study them carefully to see how the contemporary view of gene expression (right) contrasts with the more traditional view (left).

The enzyme

RNA polymerase II is a complex multisubunit enzyme - the yeast enzyme has 12 subunits. The largest subunit conatins the catalytic activity.

The promoter

Some of the common elements that have been described in class II eukaryotic promoters are the following:

- The **TATA Box** located approximately **25** bp upstream of the start-point of transcription is found in many promoters. The consenus sequence of this element is **TATAAAA** (so it resembles the TATAAT sequence of the prokaryotic **-10 region** but please do not mix them up). The TATA box appears to be more important for selecting the startpoint of transcription (i.e. positioning the enzyme) than for defining the promoter.

- The **Initiator** is a sequence that is found in many promoters and defines the startpoint of transcription.

- The **GC box** is a common element in eukaryotic class II promoters. Its consensus sequence is **GGGCGG**. It may be present in one or more copies which can be located between **40** and **100** bp upstream of the startpoint of transcription. The transcription factor **Sp1** binds to the **GC box**.

- The **CAAT box** - consensus sequence **CCAAT** - is also often found between **40** and **100** bp upstream of the startpoint of transcription. The transcription factor **CTF** or **NF1** binds to the **CAAT box**.

- In addition to the above elements, **Enhancers** may be required for full expression. These elements are not part of the promoter *per se*. They can be located upstream or downstream of the promoter and may be quite far away from it. The mechanism by which they work is not known. They may provide an **entry point** for RNA polymerase or they may bind other proteins that assist RNA polymerase to bind to the promoter region.

The transcriptional complex

It was found that RNA polymerase II can transcribe mRNA *in vitro* as long as a suitable template -- such as a nicked dsDNA or ssDNA -- is provided . The fact that the enzyme could not initiate transcription correctly on a dsDNA template indicated that additional transcription factors are required.At least six **general (or basal) transcription factors**

(TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH) have been characterized. In the presence of these transcription factors, the enzyme is able to initiate transcription at promoters correctly. However, even in the presence of transcription factors, the enzyme complex is unable to recognize and respond to regulatory signals.

In addition to the general transcription factors, the trsncriptional complex will also be affected by the presence of an promoter-proximal regulatory sequences and the presence of transcription factors that bind to those sequences. Such factors may be present in some cells/tissues but not in others. For example, the octamer motif (shown for the histone H2B gene above) binds two different transcription factors: Oct-1 and Oct-2. Oct-1 is ubiquitous but Oct-2 is expressed only in lymphoid cells where it activates immunoglobulin light chain gene transctiption. It has been suggested that the combination of **RNA polymerase II**, **transcription factors**, and regulatory response complexes such as the **Srb-Mediator,** is the eukaryotic equivalent of a **holoenzyme**.

Assembly of a transcriptional complex

The following image shows a model for the assembly of a Class II eukaryotic transcription complex:

The basic process is likely to include the following steps:

- TFIID recognizes and binds to the TATA box.TFIID consists of TATA box binding protein - TBP and ~10 TBP associated factors - TAFs.
- TBP is a 180 amino acid protein that consists of two very similar 66 amino acid domains separated by a short basic region. The protein has a "saddle-shaped" structure that sits astride a DNA molecule and binds to it via contacts in the minor groove. Binding also causes an 80° bend in the DNA.
- TFIID is a positional factor - it targets RNA polymerase to the promoter. In the case of class II transcriptional units, however, TBP binds directly to DNA.
- TFIIA binds and stabilizes TFIID binding.
- The RNA polymerase II holoenzyme assembles - possibly in a stepwise manner to form a preinitiation complex
- The holoenzyme consists of the RNA polymerase II complex, the regulatory complexes and the following transcription factors:
- TFIIB-TFIIB is a single polypeptide. It can bind both upstream and downstream of the TATA box (i.e. closer to the startpoint of transcription). It recruits TFIIF-RNAPII to the complex. It may interact directly with RNAP II.

- TFIIE-TFIIE is a complex of two subunits. It recruits TFIIH to the complex thereby priming the initiation complex for promoter clearance and elongation.

- TFIIF-TFIIF also has two subunits - RAP38 & RAP74. The latter has a helicase activity and may therefore be involved in melting the DNA at the promoter to expose the template strand.

- TFIIH-TFIIH is a complex of 9 subunits. One of the subunits has a kinase activity that carries out the phosphorylation that is required for promoter clearance
  The two largest subunits (**XPB** and **XPD**) have **helicase** activity; this activity of TFIIH is also required for Nucleotide Excision Repair in the cell and mutations in these subunits are associated with three genetic disorders: **Xeroderma pigmentosum**, **Cockayne's disease** (repair defects) and **Trichothiodystrophy** (a transcription defect). Another subunit is a **cyclin** (**cdk7** - **cyclinH**)

- There is some evidence that the order of assembly of transcription factors may be **TFIID** ->**TFIIA** ->**TFIIB** -> (**TFIIF** + **RNAP II**) ->**TFIIE** ->**TFIIH**

Finally, the various regulatory factors (**Srb-Mediator**, **Srb10-CDK** and **Swi-Snf**) bind to complete formation of the **pre-initiation complex**.

The carboxy terminal domain (CTD) of the largest subunit of RNA polymerase II is phosphorylated. This results in promoter clearance. RNA polymerase II dissociates from the Transcription factors and other protein complexes that were required for assembly.

The CTD consists of 52 repeats of the amino acid sequence Y-S-P-T-S-P-S. Ser5 is phosphorylated by tthe kinase activity of TFIIH.TFIIH phosphorylates Ser5. This serves to recruit the transcription elongation factor DSIF to the complex, which in turn recruits another elongation factor, NELF, which arrests transcription. This pause permits the capping enzymes to join and modify the 5' end of the transcript. A third elongation factor P-TEFb (a kinase) joins and phosphorylates CTD and NELF, neutralising them. P-TEFb phosphorylates CTD at Ser2.

There is also evidence that an additional transcription factor, **TFIIS**, participates in **transcription elongation.**

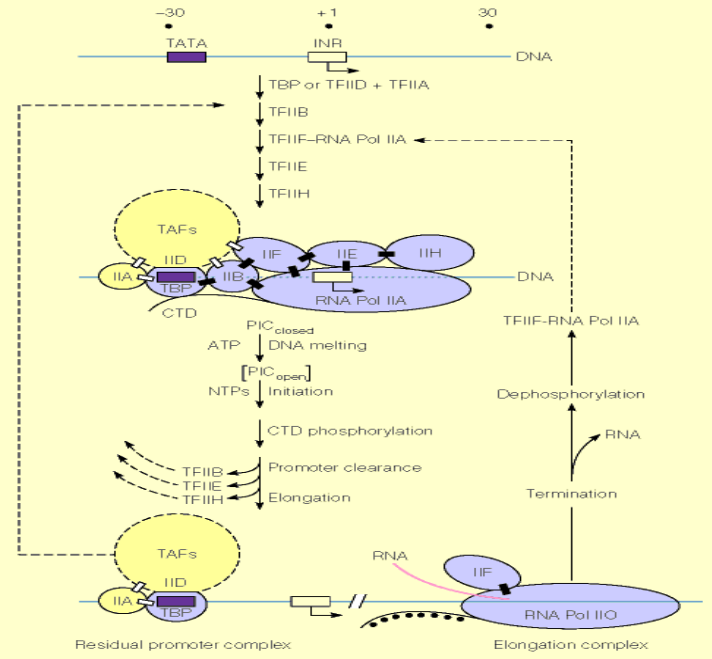
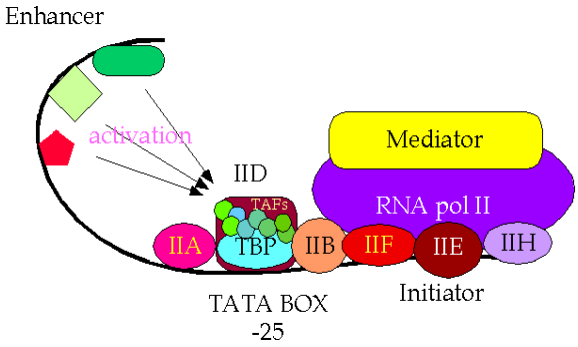


Fig: Synthesis of mRNA in eukaryotes



3.Class III Transcription Units

Class III genes are principally those for small RNA molecules in the cell. The best studied examples are the 5S rRNA gene -- which has been studied extensively in *Xenopus laevis*, and tRNA genes.

The enzyme

RNA polymerase III is the largest of the three RNA polymerases with 17 subunits and a molecular weight of over 700 kD. It is moderately sensistive to □-amanitin. It is also the most active.

The promoter

Class III promoters are distinctive because some of them are located within the gene whose transcription they direct.The promoters for 5S rRNA and tRNA genes are located within the gene. In the case of the *Xenopus laevis* 5S rRNA gene, which is 120 bp in

length, it has been found that the segment from +41 to +87 is sufficient to direct transcription and therefore defines the promoter.

The promoters for snRNA genes lie upstream of the startpoint of transcription.

Assembly of a transcriptional complex

Assembly of a functional complex requires the participation of a number of additional transcription factors. Transcription of the 5S rRNA gene requires three additional factors; transcription of tRNA genes require two.

TFIIIA-This factor is required only for the transcription of 5S rRNA genes. It contains a single polypeptide with a Zn finger DNA-binding motif. It functions as an assembly factor for some class III promoters but not for all.

TFIIIB-This factor contains three subunits, one of which is TBP - TATA-box binding protein. TFIIIB is a positional factor.

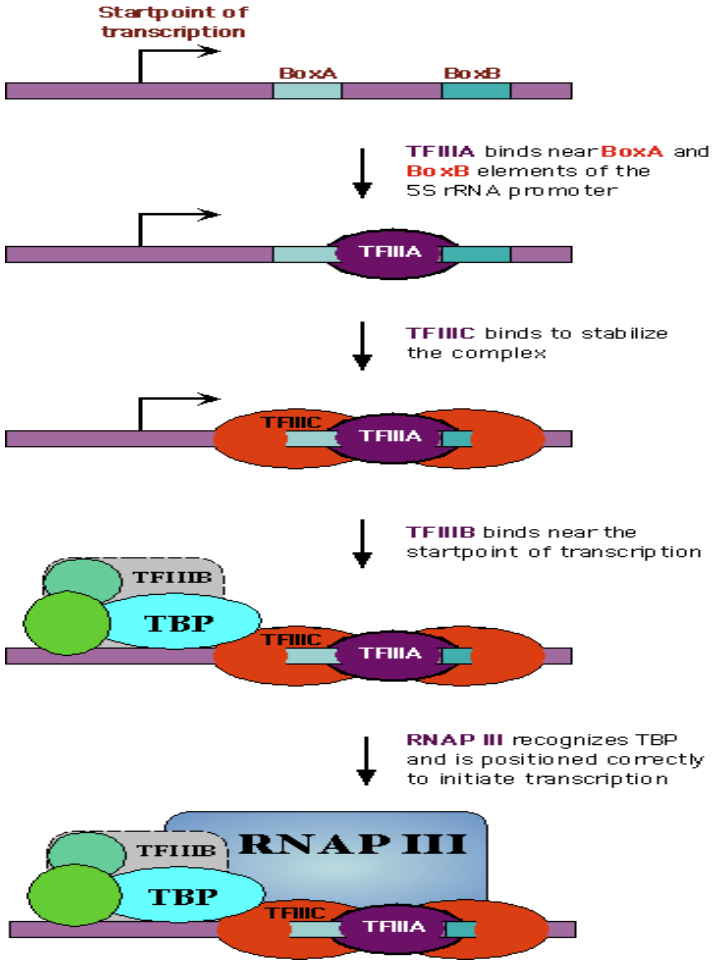
Look at Protein-DNA and Protein-Protein interactions in the TFIIIB-DNA Complex from the Hahn laboratory at the Fred Hutchinson Cancer Research Center in Seattle.

TFIIIC-TFIII consists of 6 subunits. It also functions as an assembly factor and appears to be required for all internal class III promoters.

Assembly of a transcription complex proceeds in a step-wise manner:

- TFIIIA binds to a site within the promoter region.
- TFIIIC binds to form a stable complex. TFIIIC covers the entire gene.
- TFIIB can now bind to its binding site surrounding the startpoint of transcription.
- Finally RNA polymerase III is able to bind and begin transcription of the gene.

The following diagram illustrates these steps for the 5S rRNA promoter:



Differences between transcription of prokaryotes and eukaryotes.

There are significant differences in the process of transcription in prokaryotes versus eukaryotes.

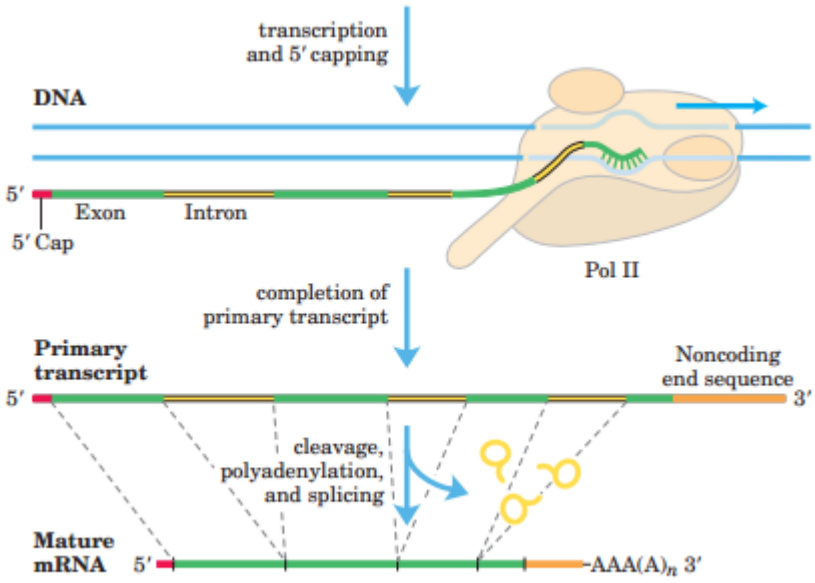
- In prokaryotes (bacteria), transcription occurs in the cytoplasm. Translation of the mRNA into proteins also occurs in the cytoplasm. In eukaryotes, transcription occurs in the cell's nucleus. mRNA then moves to the cytoplasm for translation.

- DNA in prokaryotes is much more accessible to RNA polymerase than DNA in eukaryotes. Eukaryotic DNA is wrapped around proteins called histones to form structures called nucleosomes. Eukaryotic DNA is packed to form chromatin. While RNA polymerase interacts directly with prokaryotic DNA, other proteins mediate the interation between RNA polymerase and DNA in eukaryotes.

- mRNA produced as a result of transcription is not modified in prokaryotic cells. Eukaryotic cells modify mRNA by RNA splicing, 5' end capping, and addition of a polyA tail.

Transcription and processing of mRNA

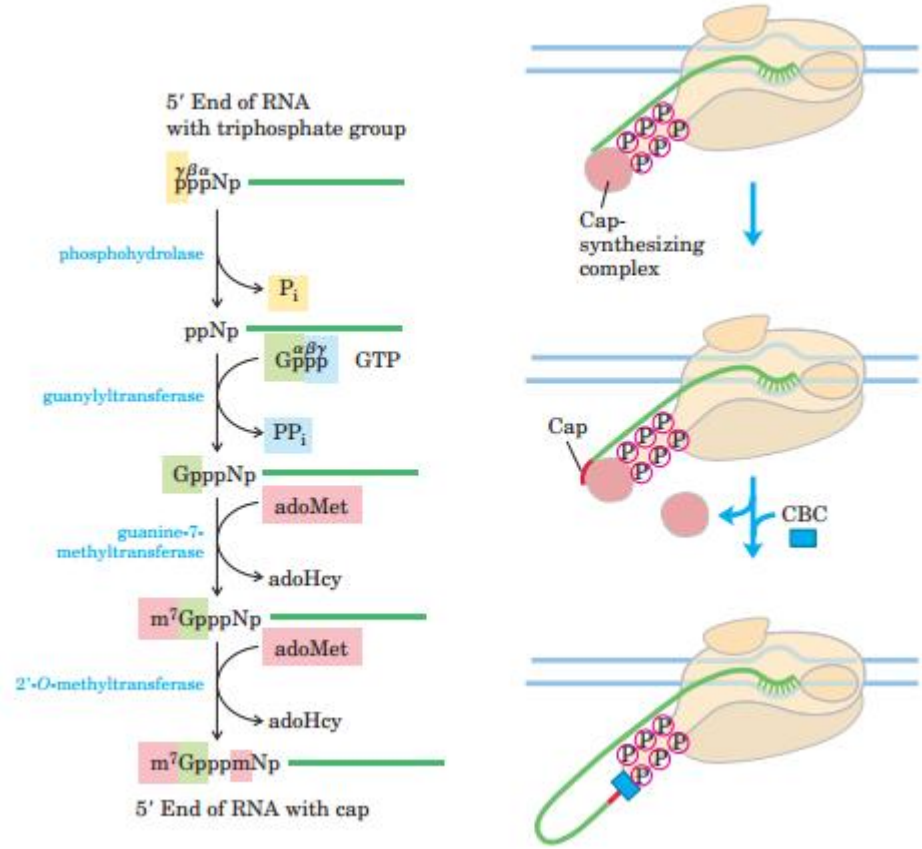
A newly synthesized RNA molecule is called a primary transcript. The most extensive processing of primary transcripts occurs in eukaryotic mRNAsand in tRNAs of both bacteria and eukaryotes. Noncoding tracts that break up the coding region of the transcript are called introns, and thecoding segments are called exons. In a processcalled splicing,the introns are removed from the primary transcript and the exons are joined to form a continuous sequence that specifies a functional polypeptide. Eukaryotic mRNAs are also modified at each end.A modified residue called a 5'cap is added atthe 5' end. The 3' end is cleaved and 80 to 250 Aresidues are added to create a poly(A) "tail." The composition of the complex changes as the primary transcript is processed,transported to the cytoplasm, and delivered to the ribosome for translation.



Capping of the eukaryotic mRNA at the 5' end:

Most eukaryotic mRNAs have a 5' cap,a residue of 7-methylguanosine linked to the 5'-terminal residue of themRNA through an unusual 5',5'-triphosphate linkage. The 5' cap helps protect mRNA fromribonucleases. The cap also binds to a specific cap-binding complex of proteins and participates in bindingof the mRNA to the ribosome to initiate translation.

The 5'cap is formed by condensation of a moleculeof GTP with the triphosphate at the 5'end of the transcript. The guanine is subsequently methylated at N-7,and additional methyl groups are often added at the 2' hydroxyls of the first and second nucleotides adjacent to the cap. The methyl groups are derivedfrom S-adenosylmethionine. All these reactions occurvery early in transcription, after the first 20 to 30 nucleotides of the transcript have been added. The capped5' end is then released from the capping enzymes andbound by the cap-binding complex.



Splicing of Introns:

There are four classes of introns. The first two, thegroup I and group II intronsare self-splicing—no protein enzymes areinvolved. Group I introns are found in some nuclear, mitochondrial, and chloroplast genes coding for rRNAs,mRNAs, and tRNAs. Group II introns are generally foundin the primary transcripts of mitochondrial or chloroplast mRNAs in fungi, algae, and plants. Neither class requires a high-energy cofactor (such as ATP) for splicing.

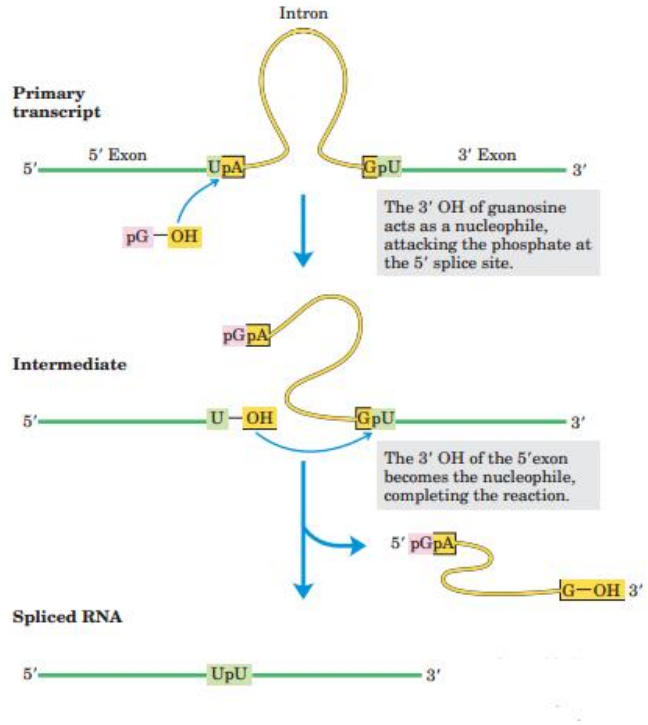
The splicingmechanisms in both groups involve two transesterification reaction steps.

Splicing of group I and II introns:

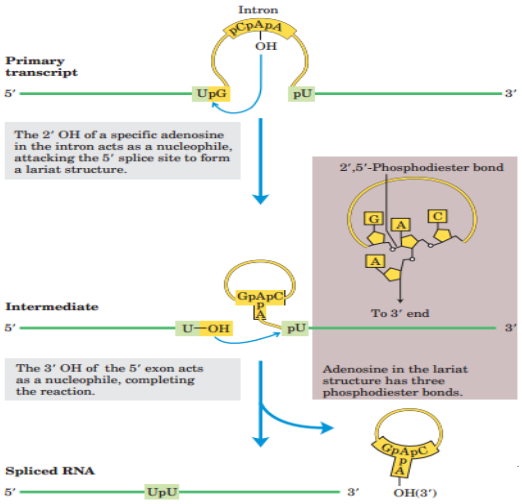
The group I splicing reaction requires a guanine nucleoside or nucleotide cofactor, but the cofactor is notused as a source of energy; instead, the 3’-hydroxylgroup of guanosine is used as a nucleophile in the firststep of the splicing pathway. The guanosine 3’-hydroxylgroup forms a normal 3’,5’-phosphodiester bond withthe 5’ end of the intron. The 3’ hydroxylof the exon that is displaced in this step then acts as anucleophile in a similar reaction at the 3’end of the introns. The result is precise excision of the intron and ligation of the exons.

In group II introns the reaction pattern is similar except for the nucleophile in the first step, which in thiscase is the 2’-hydroxyl group of an A residue within the intron. A branched lariat structure isformed as an intermediate.

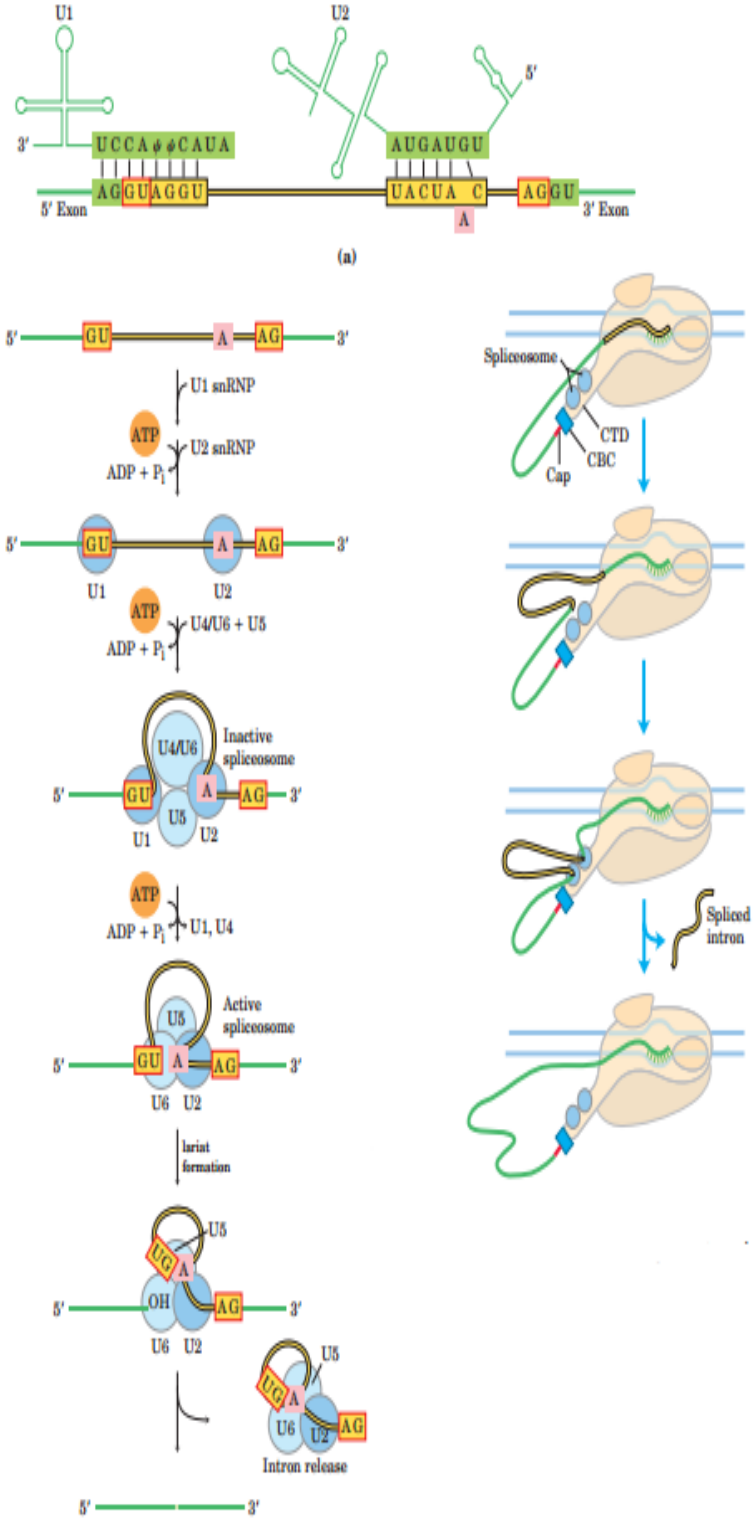
Splicing of group I introns



Splicing of group II introns



The third and largest class of introns includes those found in nuclearmRNA primary transcripts. These are called spliceosomal introns,because their removal occurs withinand is catalyzed by a large protein complex called aspliceosome.



Within the spliceosome, the introns undergo splicing by the same lariat-forming mechanism asthe group II introns. The spliceosome is made up of specialized RNA-protein complexes, small nuclear ribonucleoproteins (snRNPs, or “snurps”).Each snRNP contains one of a class of eukaryotic RNAs,100 to 200 nucleotides long, known as small nuclearRNAs (snRNAs).Five snRNAs (U1, U2, U4, U5, and U6) involved in

splicing reactions are generally found inabundance in eukaryotic nuclei. The RNAs and proteinsin snRNPs are highly conserved in eukaryotes fromyeasts to humans.

Spliceosomal introns generally have the dinucleotide sequence GU and AG at the 5'and 3'ends, respectively, and these sequences mark the sites wheresplicing occurs. The U1 snRNA contains a sequencecomplementary to sequences near the 5' splice site ofnuclear mRNA introns and the U1 snRNPbinds to this region in the primary transcript. Additionof the U2, U4, U5, and U6 snRNPs leads to formation of the spliceosome. The snRNPs togethercontribute five RNAs and about 50 proteins to thespliceosome, a supramolecular assembly nearly as complex as the ribosome. ATP isrequired for assembly of the spliceosome, but the RNAcleavage-ligation reactions do not seem to require ATP.Some mRNA introns are spliced by a less common typeof spliceosome, in which the U1 and U2 snRNPs are replaced by the U11 and U12 snRNPs. Whereas U1- andU2-containing spliceosomes remove introns with (5')GUand AG(3') terminal sequences, the U11- and U12-containing spliceosomes remove a rare class of introns that have (5')AU and AC(3')terminal sequences to mark the intronic splice sites.

The fourth class of introns, found in certain tRNAs,is distinguished from the group I and II introns in thatthe splicing reaction requires ATP and an endonuclease. The splicing endonuclease cleaves the phosphodiester bonds at both ends of the intron, and the two exons are joined by a mechanism similar to the DNA ligase reaction.

Addition of poly A tail:

At their 3' end, most eukaryotic mRNAs have a stringof 80 to 250 A residues, making up the poly(A) tail.This tail serves as a binding site for one or more specific proteins. The poly(A) tail and its associated proteins probably help protect mRNA from enzymatic destruction. Many prokaryotic mRNAs also acquirepoly(A) tails, but these tails stimulate decay of mRNA

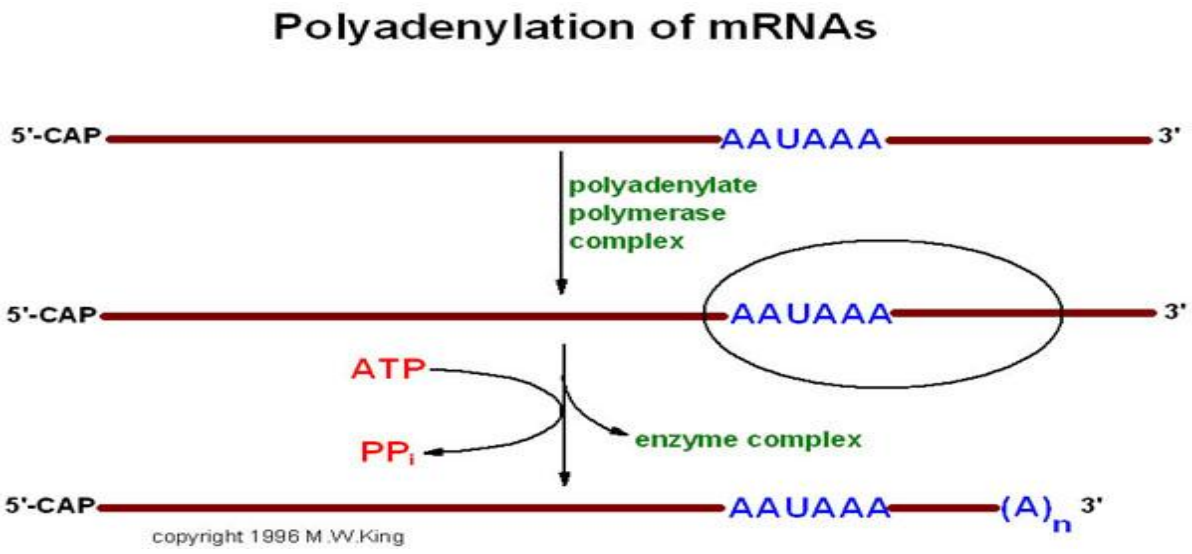
rather than protecting it from degradation.

The poly(A) tail is added in a multistep process.The transcript is extended beyond the site where thepoly(A) tail is to be added, then is cleaved at the poly(A)addition site by an endonuclease component of a largeenzyme complex, again associated with the CTD of RNA polymerase II. The mRNA site where cleavage occurs is marked by two sequence elements: thehighly conserved sequence (5')AAUAAA(3'), 10 to 30nucleotides on the 5' side (upstream) of the cleavagesite, and a less well-defined sequence rich in G and Uresidues, 20 to 40 nucleotides downstream of the cleavage site. Cleavage generates the free 3'-hydroxyl groupthat defines the end of the mRNA, to which A residuesare immediately added by polyadenylate polymerase,

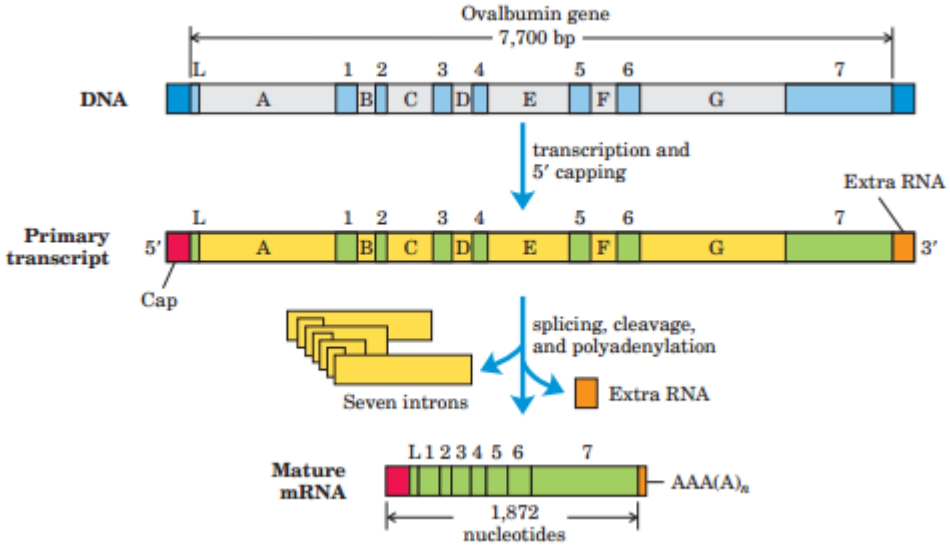
which catalyzes the reaction

$$RNA + nATP \longrightarrow RNA\text{–}(AMP)_n + nPP_i$$

where, n = 80 to 250. This enzyme does not require atemplate but does require the cleaved mRNA as a primer.



Overall processing of mRNA



Processing of rRNA and tRNA

- Ribosomal RNAs of both prokaryotic and eukaryotic cellsare made from longer precursors called preribosomalRNAs,or pre-rRNAs, synthesized by Pol I.
- In bacteria,16S, 23S, and 5S rRNAs arise from a single30S RNA precursor of about 6,500 nucleotides.
- RNA atboth ends of the 30S precursor and segments betweenthe rRNAs are removed during processing.

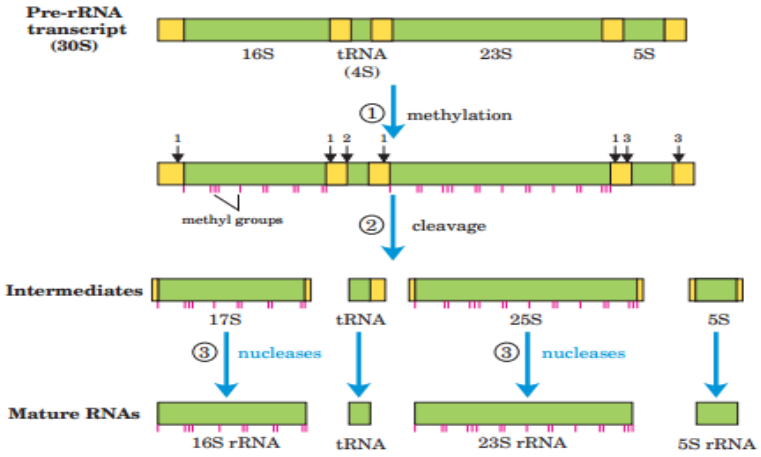


Figure: Processing of pre rRNA transcript

Fig:Processing of pre-rRNA transcripts in bacteria: Before cleavage, the 30S RNA precursor is methylated at specific bases. 2 Cleavage liberates precursors of rRNAs and tRNA(s). Cleavage at the points labeled 1, 2, and 3 is carried out by the enzymes RNase

III, RNase P, and RNase E, respectively. RNase P is a ribozyme. 3 The final 16S, 23S, and 5S rRNA products result from the action of a variety of specific nucleases.

- The genome of E. coliencodes seven pre-rRNA molecules.
- All these genes have essentially identical rRNA-coding regions, but they differ in the segments betweenthese regions.
- The segment between the 16S and 23SrRNA genes generally encodes one or two tRNAs, withdifferent tRNAs arising from different pre-rRNA transcripts.
- Coding sequences for tRNAs are also found onthe 3'side of the 5S rRNA in some precursor transcripts.

Processing of pre rRNA transcript in vertebrates

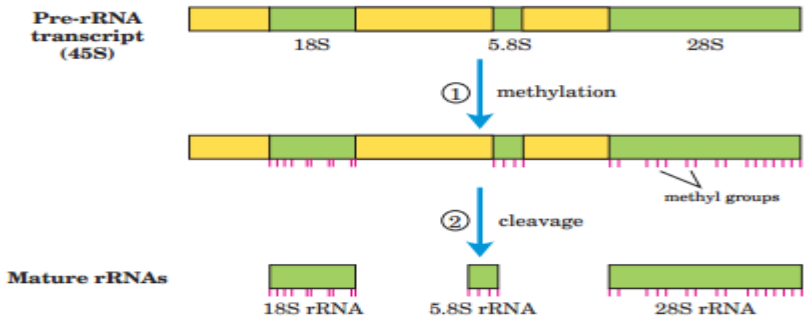


Figure:In step 1, the 45S precursor is methylated at more than 100 of its 14,000 nucleotides, mostly on the 2'-OH groups of ribose units retained in the final products. 2 A series of enzymatic cleavages produces the 18S, 5.8S, and 28S rRNAs. The cleavage reactions require RNAs found in the nucleolus, called small nucleolar RNAs (snoRNAs), within protein complexes reminiscent of spliceosomes. The 5S rRNA is produced separately.

- In eukaryotes, a 45S pre-rRNA transcript isprocessed in the nucleolus to form the 18S, 28S, and5.8S rRNAs characteristic of eukaryotic ribosomes
- The 5S rRNA of most eukaryotes is made as acompletely separate transcript by a different polymerase (Pol III instead of Pol I).
- Most cells have 40 to 50 distinct tRNAs, and eukaryotic cells have multiple copies of many of the tRNAgenes.
- Transfer RNAs are derived from longer RNA precursors by enzymatic removal of nucleotides from the5' and 3' ends.

- In eukaryotes, introns arepresent in a few tRNA transcripts and must be excised.
- Where two or more different tRNAs are contained ina single primary transcript, they are separated byenzymatic cleavage. The endonuclease RNase P, foundin all organisms, removes RNA at the 5'end of tRNAs.
- This enzyme contains both protein and RNA.
- The RNAcomponent is essential for activity, and in bacterial cellsit can carry out its processing function with precisioneven without the protein component.
- RNase P is therefore another example of a catalytic RNA.
- The 3'end of tRNAs is processedby one or more nucleases, including the exonucleaseRNase D.

3. tRNA processing

1. tRNA is transcribed by RNA polymerase III. The transcription product, the pre-tRNA, contains additional RNA sequences at both the 5' and 3'-ends. These additional sequences are removed from the transcript during processing. The additional nucleotides at the 5'-end are removed by an unusual RNA containing enzyme called **ribonuclease P** (RNase P) and **ribonuclease D** at 3' end (Ribozymes-RNA acting as enzymes).

2. Some tRNA precursors contain an intron located in the anticodon arm. These introns are spliced out during processing of the tRNA.

3. All mature tRNAs contain the trinucleotide CCA at their 3'-end. These three bases are not coded for by the tRNA gene. Instead, these nucleotides are added during processing of the pre-tRNA transcript. The enzyme responsible for the addition of the CCA-end is tRNA nucleotidyl transferase and the reaction proceeds according to the following scheme:

tRNA +CTP --> tRNA-C + PPi (pyrophosphate)

tRNA-C +CTP --> tRNA-C-C + PPi

tRNA-C-C +ATP --> tRNA-C-C-A + PPi

Mature tRNAs can contain up to 10% bases other than the usual adenine (A), guanine (G), cytidine (C) and uracil (U). These base modifications are introduced into the tRNA at the final processing step. The biological function of most of the modified bases is uncertain and the translation process seems normal in mutants lacking the enzymes responsible for modifying the bases.

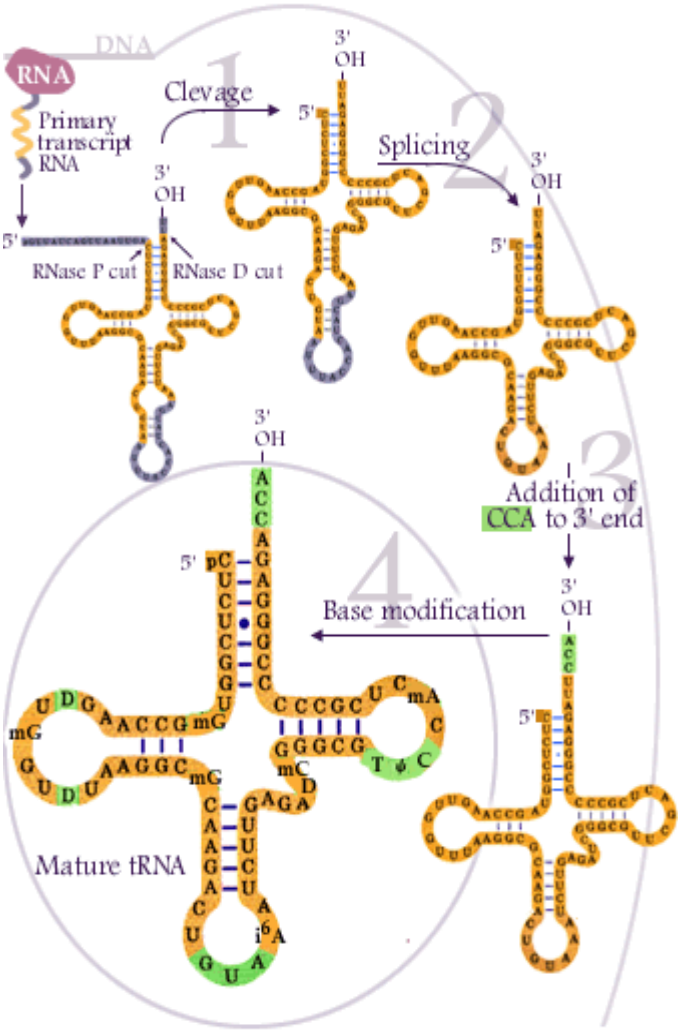
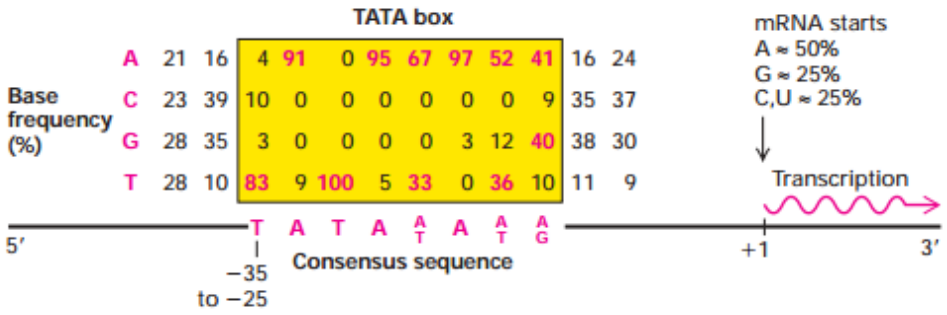


Fig: tRNA processing

Regulatory sequences in protein coding genes:

TATA box

In all the rapidly transcribing genes, a conserved sequence called the TATA boxwas found ≈25–35 base pairs upstream of the start site. Mutagenesis studieshave shown that a single-base change in this nucleotide sequence drastically decreases in vitro transcription by RNApolymerase II of genes adjacent to a TATA box. In mostcases, sequence changes between the TATA box and start sitedo not significantly affect the transcription rate. If the basepairs between the TATA box and the normal start site aredeleted, transcription of the altered, shortened template begins at a new site ≈25 base pairs downstream from the TATAbox. Consequently, the TATA box acts similarly to an E. colipromoter to position RNA polymerase II for transcriptioninitiation.



Initiators:

Instead of a TATA box, some eukaryotic genes containan alternative promoter element called an initiator. Mostnaturally occurring initiator elements have a cytosine (C) at the -1 position and an adenine (A) residue at the transcription start site (+1). Directed mutagenesis of mammalian genes with an initiator-containing promoter hasrevealed that the nucleotide sequence immediately surrounding the start site determines the strength of such promoters. Unlike the conserved TATA box sequence, however,only an extremely degenerate initiator consensus sequencehas been defined:

(5') Y-Y-A^{+1}-N-T/A-Y-Y-Y (3')

where $A^{+1}$is the base at which transcription starts, Y is apyrimidine (C or T), N is any of the four bases, and T/A is T or A at position +3.Transcription of genes with promoters containing aTATAbox or initiator element begins at a well-defined initiation site.

CpG island:

Transcription of many protein-codinggenes has been shown to begin at any one of multiple possible sites over an extended region, often 20–200 base pairsin length. As a result, such genes give rise to mRNAs withmultiple alternative 5'ends. These genes, which generally are transcribed at low rates (e.g., genes encoding the enzymes ofintermediary metabolism, often called "housekeepinggenes"), do not contain a TATA box or an initiator. Mostgenes of this type contain a CG-rich stretch of 20–50 nucleotides within ≈100 base pairs upstream of the start-site region. The dinucleotide CG is statistically underrepresented invertebrate DNAs, and the presence of a CG-rich region, orCpG island,just upstream from a start site is a distinctlynonrandom distribution. For this reason, the presence of aCpG island in genomic DNA suggests that it may contain a

transcription-initiation region.

Promoter proximal element:

Promoter-proximal elements are the control regions lying within 100–200 base pairs upstreamof the transcription start site. In some cases, promoter-proximal elementsare cell-type-specific; that is, they function only in specificdifferentiated cell types.

One approach frequently taken to determine the up-stream border of a transcription-control region for a mammalian gene involves constructing a set of 5'deletions. Once the 5'border of a transcription-control region is determined, analysis oflinker scanning mutationscan pinpoint the sequences withregulatory functions that lie between the border and thetranscription start site. One of the first usesof this type of analysis identified promoter-proximal elements of the thymidine kinase (tk) gene from herpes simplex virus (HSV). The results demonstrated that the DNAregion upstream of the HSV tkgene contains three separate transcription-control sequences: a TATA box in the interval from -32 to -16, and two other control elements farther upstream.

To test the spacing constraints on control elements inthe HSV tkpromoter region identified by analysis of linkerscanning mutations, researchers prepared and assayedconstructs containing small deletions and insertions between the elements. Changes in spacing between the promoter and promoter-proximal control elements of 20nucleotides or fewer had little effect. However, insertionsof 30 to 50 base pairs between a promoter-proximal element and the TATA box were equivalent to deleting the element. Similar analyses of other eukaryotic promotershave also indicated that considerable flexibility in thespacing between promoter-proximal elements is generallytolerated, but separations of several tens of base pairs maydecrease transcription.

Activators and repressors of transcription:

Activators are modular proteins composed of distinct functional domains. Studies with a yeast transcription activator called GAL4 provided early insight into the domain structure of transcription factors. The gene encoding the GAL4 protein, whichpromotes expression of enzymes needed to metabolize galactose, was identified by complementation analysis of gal4

mutants. Directed mutagenesis studies like thosedescribed previously identified UASs for the genes activatedby GAL4. Each of these UASs was found to contain one ormore copies of a related 17-bp sequence called UAS_{GAL}.DNase I footprinting assays with recombinant GAL4 proteinproduced in E. colifrom the yeast GAL4 gene showed thatGAL4 protein binds to UAS_{GAL}sequences. When a copy ofUAS_{GAL} was cloned upstream of a TATA box followed by alacZreporter gene, expression of lacZwas activated in galactose media in wild-type cells, but not in gal4mutants. Theseresults showed that UAS_{GAL} is a transcriptioncontrol element activated by the GAL4 protein in galactose media.

A remarkable set of experiments with gal4deletion mutants demonstrated that the GAL4 transcription factor iscomposed of separable functional domains: an N-terminalDNA-

binding domain,which binds to specific DNA sequences, and a C-terminal activation domain,which interacts with other proteins to stimulate transcription from a nearby promoter. When the N-terminalDNA-binding domain of GAL4 was fused directly to various of its C-terminal fragments, the resulting truncated proteinsretained the ability to stimulate expression of a reporter genein an in vivo assay like that depicted in Figure 11-16. Thusthe internal portion of the protein is not required for functioning of GAL4 as a transcription factor. Similar experiments with another yeast transcription factor, GCN4, whichregulates genes required for synthesis of many amino acids,indicated that it contains an ≈60-aa DNA-binding domain atits C-terminus and an ≈20-aa activation domain near themiddle of its sequence.

Further evidence for the existence of distinct activationdomains in GAL4 and GCN4 came from experiments inwhich their activation domains were fused to a DNA-bindingdomain from an entirely unrelated E. coliDNA-binding protein. When these fusion proteins were assayed in vivo, theyactivated transcription of a reporter gene containing the cognate site for the E. coliprotein. Thus functional transcriptionfactors can be constructed from entirely novel combinationsof prokaryotic and eukaryotic elements.Studies such as these have now been carried out withmany eukaryotic activators. The structural model of eukaryotic activators that has emerged from these studies is a modular one in which one or more activation domains areconnected to a sequence-specific DNA-binding domainthrough flexible protein domains. In somecases, amino acids included in the DNA-binding domain alsocontribute to transcriptional activation. Activation domains are thought to function bybinding other proteins involved in transcription. The presence of flexible domains connecting the DNA-binding domains to activation domains may explain why alterations inthe spacing between control elements are so well toleratedin eukaryotic control regions. Thus even when the positionsof transcription factors bound to DNA are shifted relativeto each other, their activation domains may still be able to interact because they are attached to their DNA-binding domains through flexible protein regions.

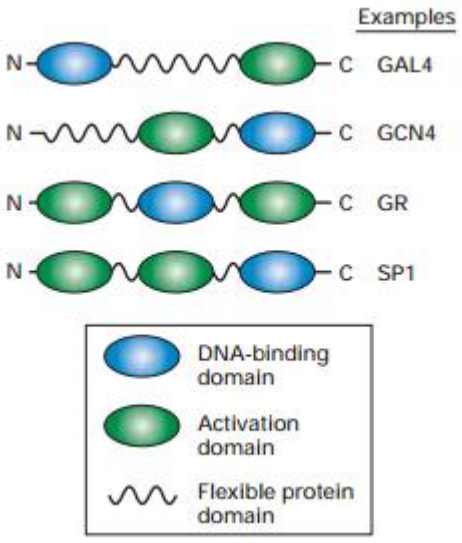


Figure: Eukaryotic transcription activators

Repressors:

Eukaryotic transcription is regulated by repressors as wellas activators. For example, geneticists have identified mutations in yeast that result in continuously high expression ofcertain genes. This type of unregulated, abnormally high expression is called constitutiveexpression and results from theinactivation of a repressor that normally inhibits the transcription of these genes. Similarly, mutants of Drosophilaand C. eleganshave been isolated that are defective in embryonic development because they express genes in embryonic cells where they are normally repressed. The mutationsin these mutants inactivate repressors, leading to abnormaldevelopment.

Repressor-binding sites in DNA have been identified bysystematic linker scanning mutation analysis. In this type of analysis, mutation ofan activator-binding site leads to decreased expression of the linked reporter gene, whereas mutation of a repressor-binding site leads to increased expression of a reporter gene.Repressor proteins that bind such sites can be purified andassayed using the same biochemical techniques describedearlier for activator proteins.

Eukaryotic transcription repressors are the functionalconverse of activators. They can inhibit transcription froma gene they do not normally regulate when their cognatebinding sites are placed within a few hundred base pairs ofthe gene's start site. Like activators, most eukaryotic repressors are modular proteins that have two functional domains:a DNA-binding domain and a repression domain.Similar toactivation domains, repression domains continue to functionwhen fused to another type of DNA-binding domain. Ifbinding sites for this second DNA-binding domain are inserted within a few hundred base pairs of a promoter, expression of the fusion protein inhibits transcription from thepromoter. Also like activation domains, repression domainsfunction by interacting with other proteins as discussed later.The absence of appropriate repressor activity canhave devastating consequences. For instance, theprotein encoded by the Wilms' tumor (WT1) geneis a repressor that is expressed preferentially in the developing kidney. Children who inherit mutations in both the maternal and paternal WT1genes, so that they produce nofunctional WT1 protein, invariably develop kidney tumorsearly in life. The WT1 protein binds to the control region ofthe gene encoding a transcription activator called EGR-1. This gene, like many other eukaryotic genes,is subject to both repression and activation. Binding by WT1represses transcription of the EGR-1gene without inhibitingbinding of the activators that normally stimulate expressionof this gene.

Multiple transcription control element:

Initially, enhancers and promoter-proximal elements werethought to be distinct types of transcription control elements. However, as more enhancers and promoter-proximalelements were analyzed, the distinctions between them became less clear. For

example, both types of element generally can stimulate transcription even when inverted, and bothtypes often are cell-type-specific. The general consensus nowis that a spectrum of control elements regulates transcription by RNA polymerase II. At one extreme are enhancers,which can stimulate transcription from a promoter tens ofthousands of base pairs away (e.g., the SV40 enhancer). Atthe other extreme are promoter-proximal elements, such asthe upstream elements controlling the HSV tkgene, whichlose their influence when moved an additional 30–50 basepairs farther from the promoter. Thestart site at which transcription initiates encodes the first (5')nucleotide of the first exon of an mRNA, the nucleotide thatis capped. For many genes, especially those encoding abundantly expressed proteins, a TATA box located approximately 25–35 base pairs upstream from the start site directsRNA polymerase II to begin transcription at the proper nucleotide. Promoter-proximal elements, which are relativelyshort (≈10–20 base pairs), are located within the first ≈200base pairs upstream of the start site. Enhancers, in contrast,usually are ≈100 base pairs long and are composed of multiple elements of ≈10–20 base pairs. Enhancers may be located up to 50 kilobases upstream or downstream from thestart site or within an intron. Many mammalian genes are

controlled by more than one enhancer region.The S. cerevisiaegenome contains regulatory elementscalled upstream activating sequences (UASs),which functionsimilarly to enhancers and promoter-proximal elements inhigher eukaryotes. Most yeast genes contain only one UAS,which generally lies within a few hundred base pairs of thestart site. In addition, S. cerevisiaegenes contain a TATA box≈90 base pairs upstream from the transcription start site

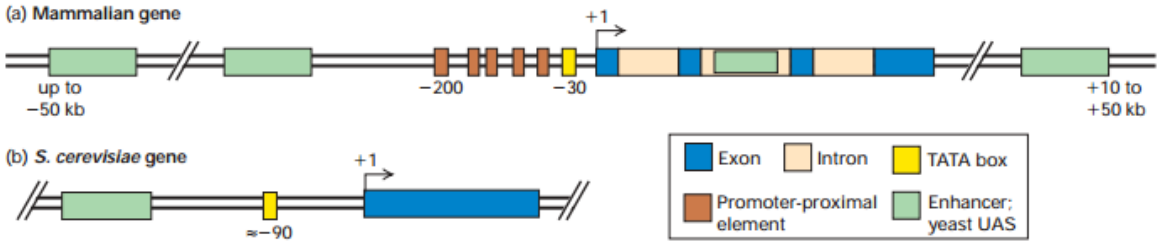


Figure: Multiple transcription control elements

Regulation of transcription factor activity by lipid-soluble hormones:

The mechanism whereby hormone binding controls the activity of nuclear receptors differs for heterodimeric and homodimeric receptors. Heterodimeric nuclear receptors (e.g.,RXR-VDR, RXR-TR, and RXR-RAR) are located exclusively in the nucleus. In the absence of their hormone ligand, they repress transcription when bound to theircognate sites in DNA. They do so by directing histonedeacetylation at nearby nucleosomes by the mechanism de-scribed earlier. As we saw earlier, in thepresence of hormone, the ligand-binding domain of theRAR monomer undergoes a dramatic conformationalchange compared with the ligand-binding domain of a nuclear receptor

in the absence of hormone. In the ligand-bound conformation, heterodimericnuclear receptors containing RXR can direct hyperacetylation of histones in nearby nucleosomes, thereby reversingthe repressing effects of the free ligand-binding domain. Inthe presence of ligand, ligand-binding domains of nuclearreceptors also bind mediator, stimulating preinitiationcomplex assembly.

In contrast to heterodimeric nuclear receptors, homodimeric receptors are found in the cytoplasm in the absence of their ligands. Hormone binding to these receptorsleads to their translocation to the nucleus. The hormone-dependent translocation of the homodimeric glucocorticoid receptor (GR) was demonstrated in the transfectionexperiments shown in Figure 11-43. The GR hormone-binding domain alone mediates this transport. Subsequentstudies showed that, in the absence of hormone, GR is anchored in the cytoplasm as a large protein aggregate complexed with inhibitor proteins, including Hsp90, a proteinrelated to Hsp70, the major heat-shock chaperone in eukaryotic cells. As long as the receptor is confined to the cytoplasm, it cannot interact with target genes and hencecannot activate transcription. Hormone binding to a ho-modimeric nuclear receptor releases the inhibitor proteins,allowing the receptor to enter the nucleus, where it canbind to response elements associated with target genes.

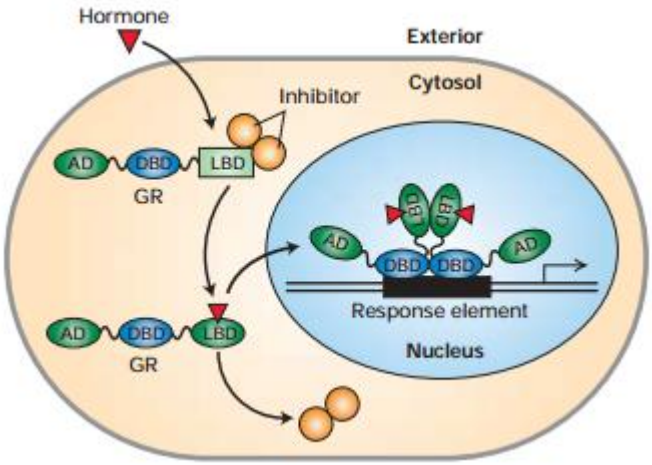


FIGURE: Model of hormone-dependent gene activation by a homodimeric nuclear receptor - In the absence of hormone, the receptor is kept in the cytoplasm by interaction between its ligand-binding domain (LBD) and inhibitor proteins. When hormone is present, it diffuses through the plasma membrane and binds to the ligand-binding domain, causing a conformational change that releases the receptor from the inhibitor proteins. The receptor with bound ligand is then translocated into the nucleus, where its DNA-binding domain (DBD) binds to response elements, allowing the ligand-binding domain and an additional activation domain (AD) at the N-terminus to stimulate transcription of target genes.

Once the receptor with bound hormonebinds to a response element, it activates transcription byinteracting with chromatin-remodeling and histone acetylase complexes and mediator.

II BSc Biochemistry -Fourth Semester
GENE EXPRESSION AND REGULATION (17BCU402)
Unit II

1	5SrRNA are synthesized by RNA polymerase	I	II	III	IV	III
2	.CpG islands are	Promotors	Operators	Silencers	Sensors	Promotors
3	rDNA	a unique gene	recombinant gene	a pseudogene	. Junk DNA	recombinant gene
4	Rho factor is	an enzyme	accessory protein	rich in U residues	rich in G-C with a palindromic sequence.	accessory protein
5	mRNAs are synthysized by RNA polymerase	. I	II	III	. all the above	II
6	The protein that bind to the TATA box in the promoter region are called	coregulators	coactivators	. enhancers	Transcriptional factors	Transcriptional factors
7	Premature chain termination in prokaryotes and eukaryotes is caused by	ricin	abrin	erythromycin	. puromycin	ricin
8	Rho dependant termination sites are rich in	only U	. Only GC	Both U and GC	. AT	Both U and GC
9	Internal promoters are present in the gene sequence of	18SrRNA gene	28SrRNA genes	mRNA genes	t RNA	t RNA
10	Transposase is encoded by	IS element	bacterial trnasposon	. rRNA	. tRNA	bacterial trnasposon
11	Prokaryotic transcription initiation is inhibited by	. Rifamycin	Rifampicin	ricin	abrin	Rifampicin
12	Who demonstrated that poly U stimulate the synthesis of poly phe	Nirenberg and Leder	. Nirenberg and Ochoa	Nirenberg and Mathai	H. Gobind Khorana	Nirenberg and Mathai
13	Shine dalgarno sequence is rich in	purines	. pyrimidines	. only A	only T	purines
14	The enzyme responsible for processing the rRNAs is	Rnase III	Rnase H	Rnase P	Rnase F	Rnase III
15	Rho factor catalyses the unwinding of RNA-DNA double helices and	Hydrolysis of ATP	RNA-RNA double helices	Migrates along the RNA in 3'-5' direction	Migrates along the RNA in 5'-3' direction	Hydrolysis of ATP
16	Mature mRNAs has the following properties except	5'Cap	Poly A tail	Exon	Intron	Intron
17	Hirpin loop like structure in transcription is related to	Initiation	Elongation	termination	Initiation&elongation	termination
18	. Snurps are involved in	Poly A tail addition	Splicing	5' cap formation	stem loop formation.	Splicing
19	Non template strand is otherwise called as	template strand	coding strand	non coding strand	sense strand	coding strand
20	____activates transcription	CRP	GMP	RNase	ribozyme	CRP
21	Eukaryotic mRNA that specifies proteins is modified before translation by all except	Addition of 5' cap	removal of exons	addition of 3' poly A tail	removal of introns	removal of introns
22	. Promoters are recognized by	alpha sub unit	gamma subunit	beta subunit	sigma subunit	sigma subunit
23	Transcription initiation in bacteria is inhibited by	Streptolydigin	Puromycin	Rifampicin	Terracycline	Rifampicin
24	Enhancers are entry point on DNA for RNA polymerase	I	II	III	I and III	II
25	. When a number of genes are transcribed as one mRNA, the mRNA is said to be	multimeric	. polymeric	polycistronic	polyclonal	polycistronic
26	The region on t RNA recognizes the appropriate base sequence m RNA is called	codon loop	anti codon loop	. P site	. A site	anti codon loop
27	. The 5' terminus of a eukaryotic mRNA molecule is	2-amino purine	5 bromo uracil	7 methyl guanosine	formyl methionine	7 methyl guanosine
28	Which of the following is not a SnRNA	U1	U2	U3	U4	U3
29	The role of tRNA is	to attach the amino acids to one another	to bring the amino acids to the correct position with respect to one another	to increase the effective concentration of amino acids	to attach m RNA to the ribosome	to bring the amino acids to the correct position with respect to one another
30	The following are associated with transcription except	-35 sequence	pribnow box	promoter	spacer	spacer
31	Proceesing of tRNA involves adding	CCA to the 3' end	CCA to the 5' end	7-methyl G to the 3' end	poly A to the 3' end	CCA to the 3' end
32	What is the characteristic form introns have after being cut from a pre-mRNA?	Linear structure	circular form	lariat shaped	theta structure	lariat shaped
33	.A tRNA molecule must be able to	recognize a codon	recognize an anticodon	distinguish one amino acid from another	recognize DNA molecules	recognize a codon
34	.Sn RNA is involved in	DNA replication	. Ribosome assembly	RNA splicing	Initiation of translation	RNA splicing
35	Telomerase is a	primase	polymerase	reverse transcriptase	helicase	reverse transcriptase
36	Translation refers to	DNA from RNA	RNA from DNA	protein from DNA	protein from RNA	protein from RNA
37	Genetic code is the dictionary of the following nucleotide bases	A, G, C and T	A, G, C and U	. A, G, T and U	. A, G, T and T	A, G, C and U
38	The initiating codon in protein synthesis	UAA	. UAG	UGA	AUG	AUG
39	The following are the termination or non sense codons except	AUG	UAA	UAG	UGA	AUG
40	The codon (of mRNA) and anticodon (of tRNA) recognize each other by pairing	5'-3' of mRNA with 5'-3' of tRNA	3'-5' of mRNA with 3'-5' of tRNA	. 5'-3' of mRNA with 3'-5' of tRNA	. 3'-5' of mRNA with 5'-3' of tRNA	. 5'-3' of mRNA with 3'-5' of tRNA
41	Wobble hypothesis explained by one of the following characteristics of the genetic code	Universality	Specificity	non overlapping	degenerate	degenerate
42	Wobble hypothesis is characterized by recognition of'	a single codon by a single tRNA	more than one codon by asingle tRNA	a single codon by more than one tRNA	more than one codon by more tRNA	more than one codon by asingle tRNA
43	. The specific information required for the sequence amino acids in protein is located on	rRNA	mRNA	tRNA	hnRNA	mRNA
44	The factories for protein synthesis	mitochondria	nucleus	ribosomes	cell	ribosomes
45	During translation mRNA is read in 5'-3' direction &polypeptide synthesis proceeds	C-terminal end to C-terminal end	C-terminal end to N terminal end	N-terminal end to C-terminal end	N-terminal end to N-terminal end	N-terminal end to C-terminal end
46	The chain initiating amino acid in protein biosynthesis in E.Coli is	Methionine	. Cysteine	N formyl methionine	. None	N formyl methionine
47	Proteins are synthesized in	Amino to carboxyl direction	carboxyl to amino direction	. Both amino and carboxyl direction	. All	Amino to carboxyl direction
48	Rifampicin inhibit Protein synthesis by binding to	beta subunit of RNA pol	delta subunit of RNA pol	gamma subunit of RNA pol	alpha subunit of RNA pol	beta subunit of RNA pol
49	Nucleic acid concerned with the protein synthesis is	DNA	mRNA	. ribosomal RNA	tRNA	. ribosomal RNA

50	The largest subunit of ribosomal RNA in prokaryotes	30S	50S	60S	40S	50S
51	one of the following compound competitively inhibits the enzyme peptidyl transferase and interfere with elongation of peptide chain	Streptomycin	Puromycin	Chloramphenicol	Terracycline	Chloramphenicol
52	Post translational modifications involve the following covalent changes	Synthesis mRNA	hydroxylation	carboxylation	Phospho diester linkage formation	Phospho diester linkage formation
53	Which of the following are termination codons	UAG, UAA, UGA	UUA, AUC, GUC	. UAG, AGG,UGA	. UAG, GAA, GUC.	UAG, UAA, UGA
54	In prokaryotes protein synthesis is carried out by ribosomes of	80s,30s, and 50s	80s,40s, and 50s	70s,30s, and 50s	70s,40s, and 60s	70s,30s, and 50s
55	The stop codon UAA is otherwise called as	Amber	Ochre	opal	none	Ochre
56	The number of 'primordial' amino acids are	15	17	18	20	20

Unit 3

Biosynthesis of proteins

The genetic code-Degeneracy of the genetic code, wobble in the anticodon, features of the genetic code, nearly universal code. **Biosynthesis of proteins-** Messenger RNA, transfer RNA, attachment of amino acids to tRNA, the ribosome - initiation, elongation and termination of translation, regulation of translation. Comparison of prokaryotic and eukaryotic protein synthesis. Use of antibiotics in understanding protein synthesis and applications in medicine.**Protein targeting and degradation -** Post translational modifications, glycosylation, signal sequences for nuclear transport, bacterial signal sequences, import of proteins by receptor mediated endocytosis, specialized systems for protein degradation.

Introduction

Genetic code

The genetic code is the set of rules by which information encoded in genetic material (DNA or mRNA sequences) is translated into proteins (amino acid sequences) by living cells. The code defines a mapping between tri-nucleotide sequences, called codons, and amino acids. With some exceptions, a triplet codon in a nucleic acid sequence specifies a single amino acid. The first elucidation of a codon was done by Marshall Nirenberg and Heinrich J. Matthaei in 1961 at the National Institutes of Health. They used a cell-free system to translate a poly-uracil RNA sequence (i.e., UUUUU...) and discovered that the polypeptide that they had synthesized consisted of only the amino acid phenylalanine. They thereby deduced that the codon UUU specified the amino acid phenylalanine. This was followed by experiments in the laboratory of Severo Ochoa demonstrating that the poly-adenine RNA sequence (AAAAA...) coded for the polypeptide, poly-lysine. The poly-cytosine RNA sequence (CCCCC...) coded for the polypeptide, poly-proline. Therefore the codon AAA specified the amino acid lysine, and the codon CCC specified the amino acid proline. Using different copolymers most of the remaining codons were then determined. Extending this work, Nirenberg and Philip Leder revealed the triplet nature of the genetic code and allowed the codons of the standard genetic code to be deciphered. In these experiments various combinations of mRNA were passed through a filter which contained ribosomes, the

components of cells that translate RNA into protein. Unique triplets promoted the binding of specific tRNAs to the ribosome. Leder and Nirenberg were able to determine the sequences of 54 out of 64 codons in their experiments.

Subsequent work by Har Gobind Khorana identified the rest of the genetic code. Shortly thereafter, Robert W. Holley determined the structure of transfer RNA (tRNA), the adapter molecule that facilitates the process of translating RNA into protein.

Wobble Hypothesis

The genetic code is a degenerate code, meaning that a given amino acid may have more than one codon.

In the genetic code, there are 43 = 64 possible codons (tri-nucleotide sequences). For translation, each of these codons requires a tRNA molecule with a complementary anticodon. If each tRNA molecule paired with its complementary mRNA codon using canonical Watson-Crick base pairing, then 64 types (species) of tRNA molecule would be required. In the standard genetic code, three of these 64 codons are stop codons, which terminate translation by binding to release factors rather than tRNA molecules, so canonical pairing would require 61 species of tRNA. Since most organisms have fewer than 45 species of tRNA, some tRNA species must pair with more than one codon. In 1966, Francis Crick proposed the Wobble hypothesis to account for this. He postulated that the 5' base on the anticodon, which binds to the 3' base on the mRNA, was not as spatially confined as the other two bases, and could, thus, have non-standard base pairing.

As an example, yeast tRNAPhe has the anticodon 5'-GmAA-3' and can recognize the codons 5'-UUC-3' and 5'-UUU-3'. It is, therefore, possible for non-Watson–Crick base pairing to occur at the third codon position, i.e., the 3' nucleotide of the mRNA codon and the 5' nucleotide of the tRNA anticodon.

tRNA Base pairing schemes

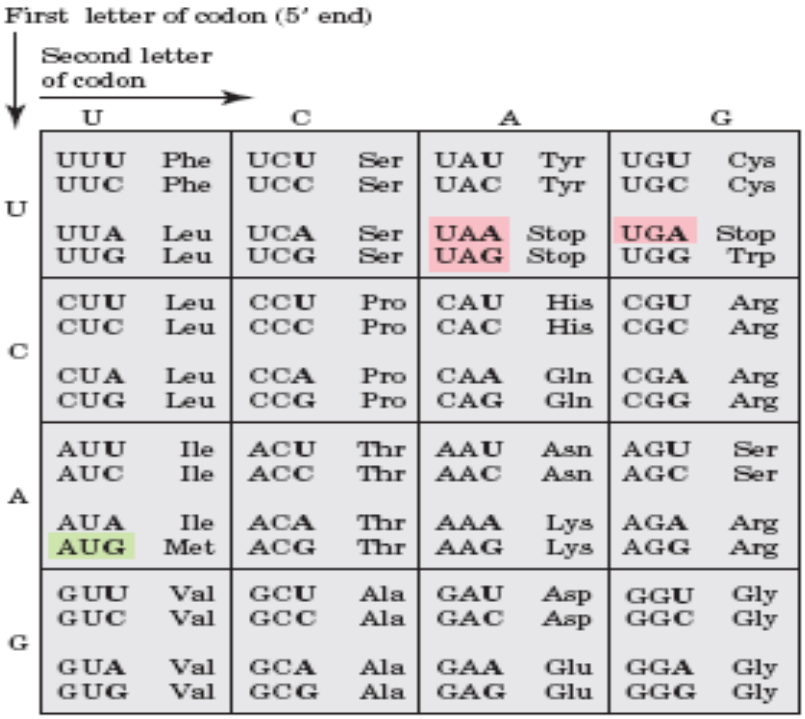
The original wobble pairing rules, as proposed by Crick. Watson-Crick base pairs are shown in **bold**, wobble base pairs in *italic*:

tRNA 5' anticodon base	**mRNA 3' codon base**
A	**U**
C	**G**
G	**C** or *U*
U	**A** or *G*

I	*A* or *C* or *U*

Revised pairing rules

tRNA 5' anticodon base	**mRNA 3' codon base**
G	U,C
C	G
k^2C	A
A	U,C,(A),G
unmodified U	U,(C),A,G
xm^5s^2U,xm^5Um,Um,xm^5U	A,(G)
xo^5U	U,A,G
I	A,C,U



The genetic code has redundancy but no ambiguity (see the codon tables above for the full correlation). For example, although codons GAA and GAG both specify glutamic acid (redundancy), neither of them specifies any other amino acid (no ambiguity). The codons encoding one amino acid may differ in any of their three positions. For example the amino acid glutamic acid is specified by GAA and GAG codons (difference in the third position), the amino acid leucine is specified by UUA, UUG, CUU, CUC, CUA,

CUG codons (difference in the first or third position), while the amino acid serine is specified by UCA, UCG, UCC, UCU, AGU, AGC (difference in the first, second or third position). Degeneracy results because there are more codons than encodable amino acids. For example, if there were two bases per codon, then only 16 amino acids could be coded for (4^2=16). Because at least 21 codes are required (20 amino acids plus stop), and the next largest number of bases is three, then 4^3 gives 64 possible codons, meaning that some degeneracy must exist.

There are 4^3 = 64 different codon combinations possible with a triplet codon of three nucleotides; all 64 codons are assigned for either amino acids or stop signals during translation. If, for example, an RNA sequence, UUUAAACCC is considered and the reading frame starts with the first U (by convention, 5' to 3'), there are three codons, namely, UUU, AAA and CCC, each of which specifies one amino acid. This RNA sequence will be translated into an amino acid sequence, three amino acids long.

Translation starts with a chain initiation codon (start codon). Unlike stop codons, the codon alone is not sufficient to begin the process. Nearby sequences (such as the Shine Dalgarno sequence in E. coli) and initiation factors are also required to start translation. The most common start codon is AUG which is read as methionine or, in bacteria, as formylmethionine. Alternative start codons (depending on the organism), include "GUG" or "UUG", which normally code for valine or leucine, respectively. However, when used as a start codon, these alternative start codons are translated as methionine or formylmethionine.

The three stop codons have been given names: UAG is amber, UGA is opal (sometimes also called umber), and UAA is ochre. "Amber" was named by discoverers Richard Epstein and Charles Steinberg after their friend Harris Bernstein, whose last name means "amber" in German. The other two stop codons were named "ochre" and "opal" in order to keep the "color names" theme. Stop codons are also called "termination" or "nonsense" codons and they signal release of the nascent polypeptide from the ribosome due to binding of release factors in the absence of cognate tRNAs with anticodons complementary to these stop signals.

The codon AUG both codes for methionine and serves as an initiation site: the first AUG in an mRNA's coding region is where translation into protein begins.

Prokaryotic translation

Translation of the genetic code involves using the genetic information to produce a polypeptide. Therefore, **translation** is synonymous with **protein synthesis**. Proteins, or to be more precise, polypeptides, are linear chains of amino acids. (You don't need to know all of the amino acids, but you should work to become more familiar with them.) As outlined in the module on the genetic code, genetic information encoded in DNA, and transferred to mRNA, is used to determine the sequence of amino acids in a polypeptide.

The following "ingredients" are required for initiation of protein synthesis:

- Aminoacylated tRNAs
- Ribosomes
- An mRNA
- Ancillary Protein "Factors"
- A special tRNA for Initiation

The actual synthesis of polypeptides is carried out by **ribosomes**.

Transfer RNA

The other player in the translation process is transfer RNA, or tRNA. tRNA serves as an adaptor or intermediary between mRNA and amino acids. tRNAs are among the best characterized RNA molecules - they are quite short (75 to 90 nucleotides long) and have nearly identical sequences in ekaryotes and prokaryotes. tRNA molecules are somewhat unique in that they contain several unusual nucleotides, such as inosine, pseudouridine, and hypoxanthine.

The sequence of each individual tRNA molecule is such that base pairing occurs between strands in different regions of the same molecule. This gives tRNA molecules a characteristic 'cloverleaf' shape. There are two main functional regions of the tRNA molecule. The middle loop of the cloverleaf contains three unpaired bases known as the **anticodon**. The anticodon base pairs with the complementary codon on mRNA during translation. Directly opposite of the anticodon is a region with no loop - it contains both ends of the linear tRNA molecule. This region, particularly the **3' end** of the tRNA is where a **specific amino acid** will bind in preparation for protein synthesis. A tRNA molecule with a particular anticodon sequence will only bind to one amino acid (for example, the tRNA with AGU as an anticodon sequence will only bind to the amino acid serine). In this way, specificity of the genetic code is maintained.

tRNA molecules are joined to their specific amino acid in a reaction known as **charging**. The 3' end of the tRNA molecule is covalently linked to the correct amino acid by an enzyme called **aminoacyl tRNA synthetase**. This enzyme recognizes the appropriate tRNA and enzyme, and uses the energy of ATP to join the two. Because the recognition of the tRNA and amino acid by the enzyme is so specific, there must be a different aminoacyl tRNA synthetase for each amino acid. Therefore, there are at least 20 different aminoacyl tRNA synthetases.

1.**Aminoacylation of Transfer RNA**

The addition of the correct amino acid to the correct tRNA is possibly the most critical step in gene expression. If this is not done correctly, then the Genetic Code would effectively be meaningless.

Once the amino acid has been added to the tRNA, it will be used for protein synthesis according to the specificity dictated by the anticodon sequence in the tRNA. This was demonstrated by the classic **Raney Nickel Experiment** carried out by **Seymour Benzer**'s group in 1962.

The aminoacylation of tRNAs is carried out by a family of enzymes called **aminoacyl tRNA synthetases** (**aaRS**). There are 21 enzymes -- one for each amino acid except lysine which has two.

There are a number of important aspects of the mechanism of action of these enzymes to consider:

- The mechanism of the addition reaction itself (aminoacylation)
- The mechanism by which the amino acyl tRNA synthetase recognizes the correct tRNA.
- The mechanism by which the amino acyl tRNA synthetase recognizes the correct amino acid.

Mechanism of Aminoacylation

Aminoacylation of the tRNA by **Aminoacyl tRNA synthetases** proceeds through a 2 step reaction mechanism that can be summarized as:

Step 1: Activation of the Amino Acid

aa + ATP <=> aa~AMP + PP_i

Step 2: Transfer of the aminoacyl group to the tRNA

aa~AMP + tRNA <=> aa-tRNA + AMP

Both steps occur in the active site of the enzyme; there is no dissociation of the aminoacyl-adenylylate intermediate from the active site during the reaction.

Mechanism of tRNA Recognition

The mechanisms by which **aaRS**s recognize their cognate tRNAs has been called a second genetic code. This is overly simplistic since the rules are clearly very complex and appear to be different for different tRNA-**aaRS** combinations.Recognition occurs through interactions with the anticodon, in which the bases of the anticodon loop are probably unstacked so that their interactions with the enzyme can be maximized.

Mechanism of Amino Acid Recognition

Recognition of the correct amino acid occurs in a manner analogous to that by which all enzymes recognize their substrates. Each amino acid will fit into an active site pocket in the **aaRS** where it will bind through a network of hydrogen bonds, electrostatic and hydrophobic interactions. Only amino acids with a sufficient number of favourable interactions will bind.

However, since some amino acids have very similar side chains, a proof-reading mechanism exists in many cases to make sure that the correct amino acid is chosen

2.Ribosomes

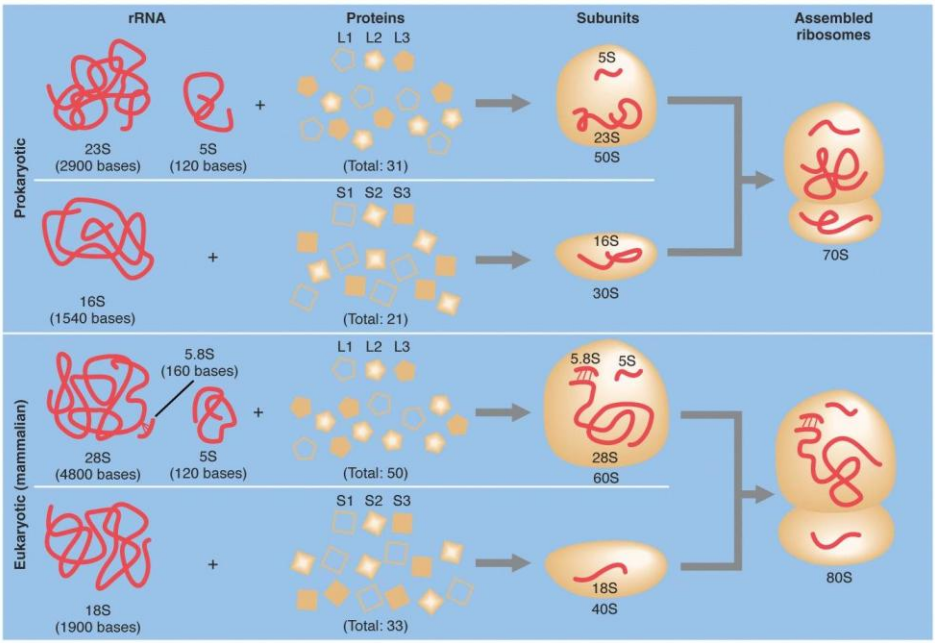
The ribosome component synthesizes protein chains. It assembles the twenty specific amino acid molecules to form the particular protein molecule determined by the nucleotide sequence of an RNA molecule. This process is known as translation; the ribosome translates the genetic information from the mRNA into proteins. Ribosomes do this by binding to an mRNA and using it as a template for determining the correct sequence of amino acids in a particular protein. The amino acids are attached to transfer RNA (tRNA) molecules, which enter one part of the ribosome and bind to the messenger RNA sequence. The attached amino acids are then joined together by another part of the ribosome. The ribosome moves along the mRNA, "reading" its sequence and producing a corresponding chain of amino acids.

The ribosomes are complex ribonucleoprotein complexes. They consist of two ribonucleoprotein subunits: a smaller subunit and a larger subunit. he smaller subunit binds to the mRNA, while the larger subunit binds to the tRNA and the amino acids. When a ribosome finishes reading a mRNA, these two subunits split apart. The sizes

and exact composition of each subunit is basically the same in all organisms though the exact details are, of course, different.

In Eukaryotes, the co-efficient of ribosomes are 80s, of which is divided into 60s for the large, and 40s for the small subunit. The 60s contain 28s rRNA,; a 5.8s, and 5sRNA. Whereas, the 40s subunit has only a single 18s rRNA along with proteins.

In prokaryotes, however, the large and small subunits are split into 50s and 30s, making a total of 70s respectively. The 50s has two types of rRNA - a 23s and a 5s Label. It also has 32 different proteins. On the other hand, the 30s contains a single 16s rRNA plus, 21 different types of proteins.



3. The messenger RNA

The mRNA must contain some feature that allows its 5' end to be recognized by and positioned correctly on the ribosome during the assembly of the protein synthesizing apparatus. In bacteria, a special **ribosome binding site** was identified by **John Shine** and **Lynn Dalgarno** in 1974. They observed that the 3' end of the **16S** rRNA is complementary to a short region just upstream of the start codon in bacterial mRNA.

The **ribosome binding site** is frequently called the **Shine-Dalgarno** sequence.

In eukaryotes, it is the 5' cap structure that is added to the mRNA which is required for correct positioning of the ribosome on the mRNA during the initiation phase of protein synthesis.

4. Ancillary Protein "Factors"

Each of the steps of protein synthesis requires the participation of a number of additional special protein factors. Each factor has a specific role to play which will be described Some of the factors are G-proteins -- they bind GTP and GTP hydrolysis is an important part of their function.

5. **A special tRNA for Initiation**

In bacteria, protein synthesis starts with a special amino acid: **N-formyl-methionine**. Addition of the formyl group to the N-terminal methionine effectively provides it with a peptide bond.

This amino acid is synthesized by modifying methionine after it has been attached to a special tRNA -- $tRNA_f^{Met}$. The same **MetRS** aminoacylates both $tRNA^{Met}$ and $tRNA_f^{Met}$.

$tRNA_f^{Met}$ is structurally different from the "regular" $tRNA^{Met}$ in a number of ways:

- It contains 3 consecutive GC base pairs in the anticodon stem.
- The terminal bases of the acceptor arm are not paired as they are in all other tRNAs.

$tRNA_f^{Met}$ is functionally different from the "regular" $tRNA^{Met}$ in two important ways:

- It is recognized by a special enzyme that will catalyse the formylation of methionyl-$tRNA_f^{Met}$ -- **transformylase**.
- **Transformylase** catalyses the formylation of the methionyl-$tRNA_f^{Met}$ but not of methionyl-$tRNA^{Met}$ or uncharged tRNA. The enzyme uses N^{10}-formyltetrahydrofolate as the formyl group donor.
- It is recognized by a special **Initiation Factor -- IF2** -- which recognizes fmet-$tRNA_f^{Met}$ but not Met-$tRNA^{Met}$ and brings it to the ribosome during the initiation phase of protein synthesis.

The presence of the formyl group on the methionine after it has been attached to $tRNA_f^{Met}$ serves two purposes:

- It ensures that this will be the only charged tRNA that is be positioned in the peptidyl site on the ribosome to start protein synthesis.
- It ensures that this tRNA will not be used for internal methionine codons.

The process of translation can be divided into three basic steps: **initiation**, **elongation**, and **termination**. Each of these steps will be considered in turn.

Translation takes place in five stages.

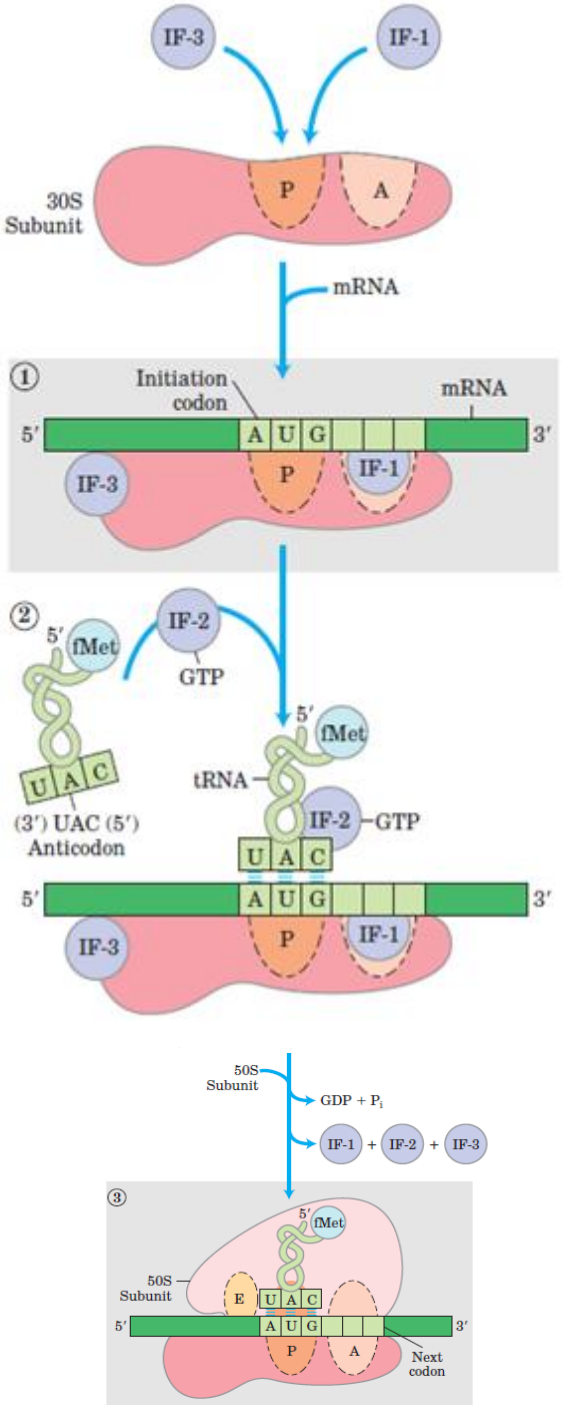
Stage 1: Activation of Amino Acids For the synthesis of a polypeptide with a defined sequence, two fundamental chemical requirements must be met: (1) the carboxyl group of each amino acid must be activated to facilitate formation of a peptide bond, and (2) a link must be established between each new amino acid and the information in the mRNA that encodes it. Both these requirements are met by attaching the amino acid to a tRNA in the first stage of protein synthesis. Attaching the right amino acid to the right tRNA is critical. This reaction takes place in the cytosol, not on the ribosome. Each of the 20 amino acids is covalently attached to a specific tRNA at the expense of ATP energy, using Mg^{2+} dependent activating enzymes known as aminoacyltRNA synthetases. When attached to their amino acid (aminoacylated) the tRNAs are said to be "charged."

$$\text{Amino acid} + \text{tRNA} + \text{ATP} \xrightarrow{Mg^{2+}} \text{aminoacyl-tRNA} + \text{AMP} + 2P_i$$

The identity of the amino acid attached to a tRNA is not checked on the ribosome, so attachment
of the correct amino acid to the tRNA is essential to the fidelity of protein synthesis.

Stage 2: Initiation

Protein synthesis begins at the amino-terminal end and proceeds by the stepwise addition of amino acids to the carboxyl-terminal end of the growing polypeptide.The AUG initiation codon thus specifies an *amino-terminal* methionine residue. Although methionine has only one codon, (5')AUG, all organisms have two tRNAs for methionine. One is used exclusively when (5')AUG



is the initiation codon for protein synthesis. The other is used to code for a Met residue in an internal position in a polypeptide.

In step 1 the 30S ribosomal subunit binds two initiation factors, IF-1 and IF-3. Factor IF-3 prevents the 30S and 50S subunits from combining prematurely. The mRNA then binds to the 30S subunit. The initiating (5')AUG is guided to its correct position by the **Shine-Dalgarno sequence** in the mRNA. This consensus sequence is an initiation signal of four to nine purine residues, 8 to 13 bp to the 5' side of the initiation codon. The sequence base-pairs with a complementary pyrimidine-rich sequence near

the 3' end of the 16S rRNA of the 30S ribosomal subunit. This mRNA-rRNA interaction positions the initiating (5')AUG sequence of the mRNA in the precise position on the 30S subunit where it is required for initiation of translation. The particular (5')AUG where fMet-tRNAfMet is to be bound is distinguished from other methionine codons by its proximity to the Shine-Dalgarno sequence in the mRNA.

Bacterial ribosomes have three sites that bind aminoacyl-tRNAs, the **aminoacyl (A) site,** the **peptidyl (P) site,** and the **exit (E) site.** Both the 30S and the 50S subunits contribute to the characteristics of the A and P sites, whereas the E site is largely confined to the 50S subunit. The initiating (5')AUG is positioned at the P site, the only site to which fMet- $tRNA^{fMet}$ can bind. The fMet-$tRNA^{fMet}$ is the only aminoacyl-tRNA that binds first to the P site; during the subsequent elongation stage, all other incoming aminoacyl-tRNAs (including the Met-$tRNA^{Met}$ that binds to interior AUG codons) bind first to the A site and only subsequently to the P and E sites. The E site is the site from which the "uncharged" tRNAs leave during elongation. Factor IF-1 binds at the A site and prevents tRNA binding at this site during initiation.

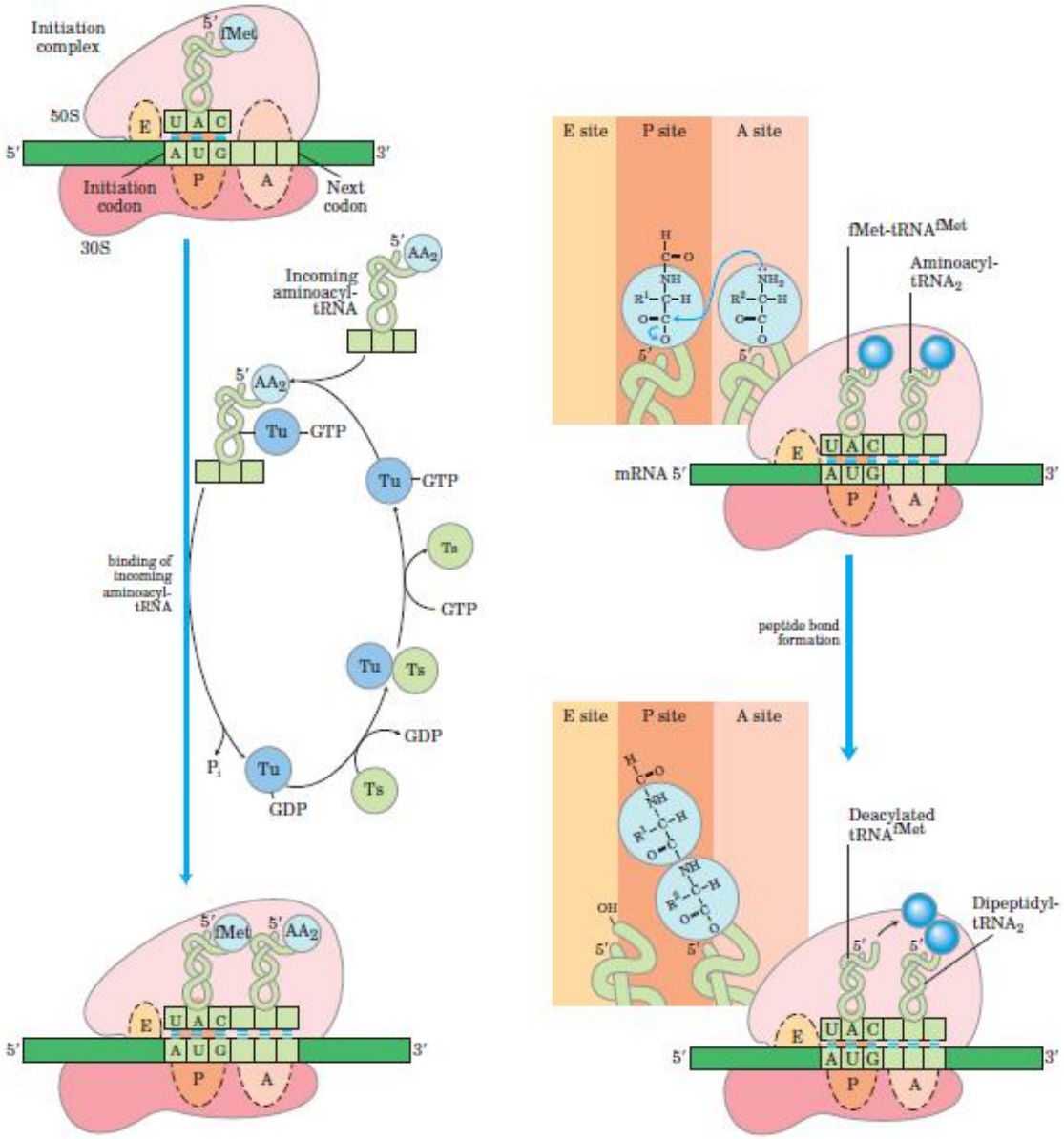
In step 2 of the initiation process, the complex consisting of the 30S ribosomal subunit, IF-3, and mRNA is joined by both GTP-bound IF-2 and the initiating fMet-$tRNA^{fMet}$. The anticodon of this tRNA now pairs correctly with the mRNA's initiation codon.

In step 3 this large complex combines with the 50S ribosomal subunit; simultaneously, the GTP bound to IF-2 is hydrolyzed to GDP and Pi, which are released from the complex. All three initiation factors depart from the ribosome at this point.

Completion of the steps produces a functional 70S ribosome called the **initiation complex,** containing the mRNA and the initiating fMettRNA^{fMet}. The correct binding of the fMet-$tRNA^{fMet}$ to the P site in the complete 70S initiation complex is assured by at least three points of recognition and attachment: the codon-anticodon interaction involving the initiation AUG fixed in the P site; interaction between the Shine-Dalgarno sequence in the mRNA and the 16SrRNA; and binding interactions between the ribosomal P site and the fMet-tRNAfMet. The initiation complex is now ready for elongation.

Stage 3: Elongation The nascent polypeptide is lengthened by covalent attachment of successive amino acid units, each carried to the ribosome and correctly positioned by its tRNA, which base-pairs to its corresponding codon in the mRNA. Elongation requires cytosolic proteins known as elongation factors. The binding of each incoming

aminoacyl-tRNA and the movement of the ribosome along the mRNA are facilitated by the hydrolysis of GTP as each residue is added to the growing polypeptide.



Elongation Step 1: In the first step of the elongation cycle, the appropriate incoming aminoacyl-tRNA binds to a complex of GTP-bound EF-Tu. The resulting aminoacyl tRNA–EF-Tu–GTP complex binds to the A site of the 70S initiation omplex. The GTP is hydrolyzed and an EF-Tu–GDP complex is released from the 70S ribosome. The EF-Tu–GTP complex is regenerated in a process involving EF-Ts and GTP.

Elongation Step 2: A peptide bond is now formed between the two amino acids bound by their tRNAs to the A and P sites on the ribosome. This occurs by the transfer of the initiating *N*-formylmethionyl group from its tRNA to the amino group of the second amino acid, now in the A site. The α amino group of the amino acid in the A site acts as

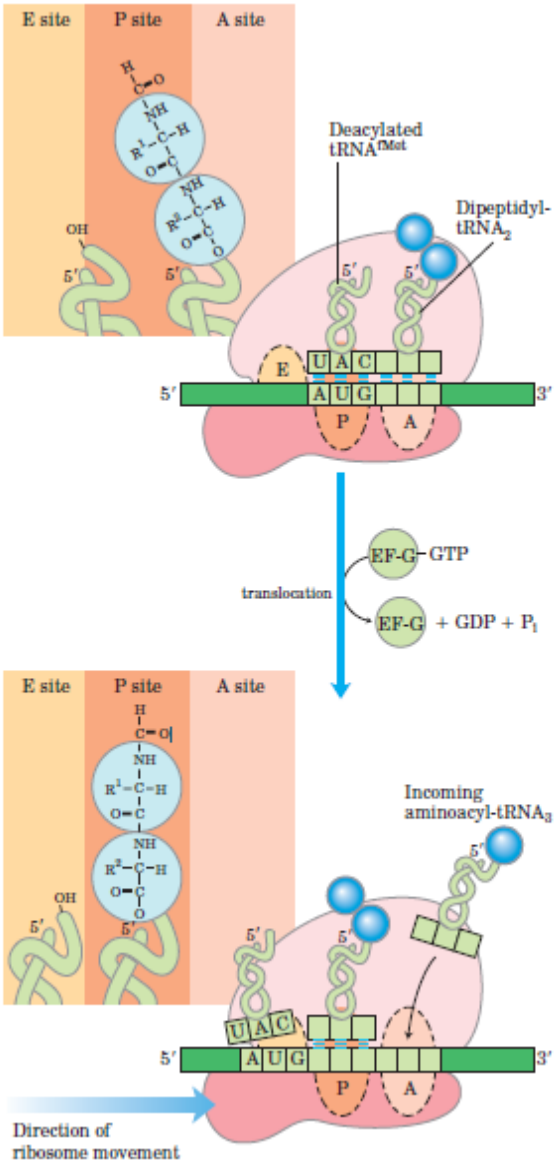
a nucleophile, displacing the tRNA in the P site to form the peptide bond. This reaction produces a dipeptidyl tRNA in the A site, and the now "uncharged" (deacylated) $tRNA^{fMet}$ remains bound to the P site. The tRNAs then shift to a hybrid binding state, with elements of each spanning two different sites on the ribosome.

Elongation Step 3: In the final step of the elongation cycle, **translocation,** the ribosome moves one codon toward the 3' end of the mRNA. This movement shifts the anticodon of the dipeptidyl tRNA, which is still attached to the second codon of the mRNA, from the A site to the P site, and shifts the deacylated tRNA from the P site to the E site, from where the tRNA is released into the cytosol. The third codon of the mRNA now lies in the A site and the second codon in the P site. Movement of the ribosome along the mRNA requires EF-G (also known as translocase) and the energy provided by hydrolysis of another molecule of GTP.

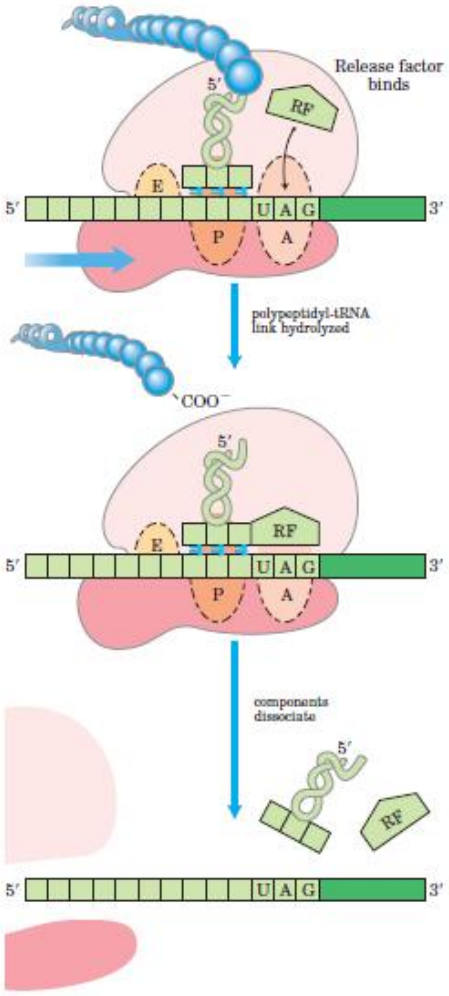
The ribosome, with its attached dipeptidyl-tRNA and mRNA, is now ready for the next elongation cycle and attachment of a third amino acid residue. This process occurs in the same way as addition of the second residue. For each amino acid residue correctly added to the growing polypeptide, two GTPs are hydrolyzed to GDP and Pi as the ribosome moves from codon to codon along the mRNA toward the 3' end.

Stage 4: Termination and Release Completion of the polypeptide chain is signaled by a termination codon in the mRNA. The new polypeptide is released from the ribosome, aided by proteins called release factors.

Elongation continues until the ribosome adds the last amino acid coded by the mRNA. **Termination,** the fourth stage of polypeptide synthesis, is signaled by the presence of one of three termination codons in the mRNA (UAA, UAG, UGA), immediately following the final coded amino acid. In bacteria, once a termination codon occupies the ribosomal A site, three **termination factors,** or **release factors**— the proteins RF-1, RF-2, and RF-3— contribute to (1) hydrolysis of the terminal peptidyltRNA bond; (2) release of the free polypeptide and the last tRNA, now uncharged, from the P site; and (3) dissociation of the 70S ribosome into its 30S and 50S subunits, ready to start a new cycle of polypeptide synthesis. RF-1 recognizes the termination codons UAG and UAA, and RF-2 recognizes UGA and UAA. Either RF-1 or RF-2



(depending on which codon is present) binds at a termination codon and induces peptidyl transferase to transfer the growing polypeptide to a water molecule rather than to another amino acid.



Fidelity in protein synthesis.

- On average, the energy derived from the hydrolysis of more than four NTPs to NDPs is required
- for the formation of each peptide bond of a polypeptide.
- This energy permits very high fidelity in the biological translation of the genetic message of mRNA into the amino acid sequence of proteins.
- The degree of fidelity in protein synthesis is sufficient to ensure that most proteins contain no mistakes and that the large amount of energy required to synthesize a protein is rarely wasted. One defective protein molecule is usually unimportant when many correct copies of the same protein are present.

Eukaryotic Translation

As mentioned previously, eukaryotic translation is very similar overall to prokaryotic translation. There are a few notable differences, however. These include the following:

Eukaryotic mRNAs do not contain a Shine-Delgarno sequence. Instead, ribosomal subunits recognize and bind to the 5' cap of eukaryotic mRNAs. In other words, the 5' cap takes the place of the Shine-Delgarno sequence.

Eukaryotes do not use formyl methionine as the first amino acid in every polypeptide; ordinary methionine is used. Eukaryotes do have a specific initiator tRNA, however.

Eukaryotic translation involves many more protein factors than prokaryotic translation (For example, eukaryotic initiation involves at least 10 factors, instead of the 3 in prokaryotes.)

Post translational modification of proteins

Protein post-translational modification (PTM) increases the functional diversity of the proteome by the covalent addition of functional groups or proteins, proteolytic cleavage of regulatory subunits or degradation of entire proteins. These modifications include phosphorylation, glycosylation, ubiquitination, nitrosylation, methylation, acetylation, lipidation and proteolysis and influence almost all aspects of normal cell biology and pathogenesis. Therefore, identifying and understanding PTMs is critical in the study of cell biology and disease treatment and prevention.

Post-translational modification can occur at any step in the "life cycle" of a protein. For example, many proteins are modified shortly after translation is completed to mediate proper protein folding or stability or to direct the nascent protein to distinct cellular compartments (e.g., nucleus, membrane). Other modifications occur after folding and localization are completed to activate or inactivate catalytic activity or to otherwise influence the biological activity of the protein. Proteins are also covalently linked to tags that target a protein for degradation. Besides single modifications, proteins are often modified through a combination of post-translational cleavage and the addition of functional groups through a step-wise mechanism of protein maturation or activation.

Protein PTMs can also be reversible depending on the nature of the modification. For example, kinases phosphorylate proteins at specific amino acid side chains, which is a common method of catalytic activation or inactivation. Conversely, phosphatases hydrolyze the phosphate group to remove it from the protein and reverse

the biological activity. Proteolytic cleavage of peptide bonds is a thermodynamically favorable reaction and therefore permanently removes peptide sequences or regulatory domains.

Consequently, the analysis of proteins and their post-translational modifications is particularly important for the study of heart disease, cancer, neurodegenerative diseases and diabetes. The characterization of PTMs, although challenging, provides invaluable insight into the cellular functions underlying etiological processes. Technically, the main challenges in studying post-translationally modified proteins are the development of specific detection and purification methods. Fortunately, these technical obstacles are being overcome with a variety of new and refined proteomics technologies.

As noted above, the large number of different PTMs precludes a thorough review of all possible protein modifications. Therefore, this overview only touches on a small number of the most common types of PTMs studied in protein research today. They are

Phosphorylation

Reversible protein phosphorylation, principally on serine, threonine or tyrosine residues, is one of the most important and well-studied post-translational modifications. Phosphorylation plays critical roles in the regulation of many cellular processes including cell cycle, growth, apoptosis and signal transduction pathways.

Glycosylation

Protein glycosylation is acknowledged as one of the major post-translational modifications, with significant effects on protein folding, conformation, distribution, stability and activity. Glycosylation encompasses a diverse selection of sugar-moiety additions to proteins that ranges from simple monosaccharide modifications of nuclear transcription factors to highly complex branched polysaccharide changes of cell surface receptors. Carbohydrates in the form of aspargine-linked (N-linked) or serine/threonine-linked (O-linked) oligosaccharides are major structural components of many cell surface and secreted proteins.

Ubiquitination

Ubiquitin is an 8-kDa polypeptide consisting of 76 amino acids that is appended to the Îµ-NH2 of lysine in target proteins via the C-terminal glycine of ubiquitin. Following an initial monoubiquitination event, the formation of a ubiquitin polymer may occur,

and polyubiquitinated proteins are then recognized by the 26S proteasome that catalyzes the degradation of the ubiquitinated protein and the recycling of ubiquitin.

S-Nitrosylation

Nitric oxide (NO) is produced by three isoforms of nitric oxide synthase (NOS) and is a chemical messenger that reacts with free cysteine residues to form S-nitrothiols (SNOs). S-nitrosylation is a critical PTM used by cells to stabilize proteins, regulate gene expression and provide NO donors, and the generation, localization, activation and catabolism of SNOs are tightly regulated.

S-nitrosylation is a reversible reaction, and SNOs have a short half life in the cytoplasm because of the host of reducing enzymes, including glutathione (GSH) and thioredoxin, that denitrosylate proteins. Therefore, SNOs are often stored in membranes, vesicles, the interstitial space and lipophilic protein folds to protect them from denitrosylation. For example, caspases, which mediate apoptosis, are stored in the mitochondrial intermembrane space as SNOs. In response to extra- or intracellular cues, the caspases are released into the cytoplasm, and the highly reducing environment rapidly denitrosylates the proteins, resulting in caspase activation and the induction of apoptosis.

S-nitrosylation is not a random event, and only specific cysteine residues are S-nitrosylated. Because proteins may contain multiple cysteines and due to the labile nature of SNOs, S-nitrosylated cysteines can be difficult to detect and distinguish from non-S-nitrosylated amino acids. The biotin switch assay, developed by Jaffrey et al., is a common method of detecting SNOs, and the steps of the assay are listed below:

- All free cysteines are blocked.
- All remaining cysteines (presumably only those that are denitrosylated) are denitrosylated.
- The now-free thiol groups are then biotinylated.
- Biotinylated proteins are detected by SDS-PAGE and Western blot analysis or mass spectrometry.

Methylation

The transfer of one-carbon methyl groups to nitrogen or oxygen (N- and O-methylation, respectively) to amino acid side chains increases the hydrophobicity of the protein and can neutralize a negative amino acid charge when bound to carboxylic acids. Methylation is mediated by methyltransferases, and S-adenosyl methionine (SAM) is the primary methyl group donor.

Methylation occurs so often that SAM has been suggested to be the most-used substrate in enzymatic reactions after ATP. Additionally, while N-methylation is irreversible, O-methylation is potentially reversible. Methylation is a well-known mechanism of epigenetic regulation, as histone methylation and demethylation influences the availability of DNA for transcription. Amino acid residues can be conjugated to a single methyl group or multiple methyl groups to increase the effects of modification.

N-Acetylation

N-acetylation, or the transfer of an acetyl group to nitrogen, occurs in almost all eukaryotic proteins through both irreversible and reversible mechanisms. N-terminal acetylation requires the cleavage of the N-terminal methionine by methionine aminopeptidase (MAP) before replacing the amino acid with an acetyl group from acetyl-CoA by N-acetyltransferase (NAT) enzymes. This type of acetylation is co-translational, in that N-terminus is acetylated on growing polypeptide chains that are still attached to the ribosome. While 80-90% of eukaryotic proteins are acetylated in this manner, the exact biological significance is still unclear.

Acetylation at the Îµ-NH2 of lysine (termed lysine acetylation) on histone N-termini is a common method of regulating gene transcription. Histone acetylation is a reversible event that reduces chromosomal condensation to promote transcription, and the acetylation of these lysine residues is regulated by transcription factors that contain histone acetyletransferase (HAT) activity. While transcription factors with HAT activity act as transcription co-activators, histone deacetylase (HDAC) enzymes are co-repressors that reverse the effects of acetylation by reducing the level of lysine acetylation and increasing chromosomal condensation.

Sirtuins (silent information regulator) are a group of NAD-dependent deacetylases that target histones. As their name implies, they maintain gene silencing by hypoacetylating histones and have been reported to aid in maintaining genomic stability.

While acetylation was first detected in histones, cytoplasmic proteins have been reported to also be acetylated, and therefore acetylation seems to play a greater role in cell biology than simply transcriptional regulation. Furthermore, crosstalk between acetylation and other post-translational modifications, including phosphorylation, ubiquitination and methylation, can modify the biological function of the acetylated protein.

Protein acetylation can be detected by chromosome immunoprecipitation (ChIP) using acetyllysine-specific antibodies or by mass spectrometry, where an increase in histone by 42 mass units represents a single acetylation.

Lipidation

Lipidation is a method to target proteins to membranes in organelles (endoplasmic reticulum [ER], Golgi apparatus, mitochondria), vesicles (endosomes, lysosomes) and the plasma membrane. The four types of lipidation are:

- C-terminal glycosyl phosphatidylinositol (GPI) anchor
- N-terminal myristoylation
- S-myristoylation
- S-prenylation

Each type of modification gives proteins distinct membrane affinities, although all types of lipidation increase the hydrophobicity of a protein and thus its affinity for membranes. The different types of lipidation are also not mutually exclusive, in that two or more lipids can be attached to a given protein.

GPI anchors tether cell surface proteins to the plasma membrane. These hydrophobic moieties are prepared in the ER, where they are then added to the nascent protein en bloc. GPI-anchored proteins are often localized to cholesterol- and sphingolipid-rich lipid rafts, which act as signaling platforms on the plasma membrane. This type of modification is reversible, as the GPI anchor can be released from the protein by phosphoinositol-specific phospholipase C. Indeed, this lipase is used in the detection of GPI-anchored proteins to release GPI-anchored proteins from membranes for gel separation and analysis by mass spectrometry.

N-myristoylation is a method to give proteins a hydrophobic handle for membrane localization. The myristoyl group is a 14-carbon saturated fatty acid (C14), which gives the protein sufficient hydrophobicity and affinity for membranes, but not enough to permanently anchor the protein in the membrane. N-myristoylation can therefore act as a conformational localization switch, in which protein conformational changes influence the availability of the handle for membrane attachment. Because of this conditional localization, signal proteins that selectively localize to membrane, such as Src-family kinases, are N-myristoylated.

N-myristoylation is facilitated specifically by N-myristoyltransferase (NMT) and uses myristoyl-CoA as the substrate to attach the myristoyl group to the N-terminal glycine. Because methionine is the N-terminal amino acid of all eukaryotic proteins, this PTM

requires methionine cleavage by the above-mentioned MAP prior to addition of the myristoyl group; this represents one example of multiple PTMs on a single protein.

S-palmitoylation adds a C16 palmitoyl group from palmitoyl-CoA to the thiolate side chain of cysteine residues via palmitoyl acyl transferases (PATs). Because of the longer hydrophobic group, this anchor can permanently anchor the protein to the membrane. This localization can be reversed, though, by thioesterases that break the link between the protein and the anchor; thus, S-palmitoylation is used as an on/off switch to regulate membrane localization. S-palmitoylation is often used to strengthen other types of lipidation, such as myristoylation or farnesylation. S-palmitoylated proteins also selectively concentrate at lipid rafts.

S-prenylation covalently adds a farnesyl (C15) or geranylgeranyl (C20) group to specific cysteine residues within 5 amino acids from the C-terminus via farnesyl transferase (FT) or geranylgeranyl transferases (GGT I and II). Unlike S-palmitoylation, S-prenylation is hydrolytically stable. Approximately 2% of all proteins are prenylated, including all members of the Ras superfamily. This group of molecular switches is farnesylated, geranylgeranylated or a combination of both. Additionally, these proteins have specific 4-amino acid motifs at the C-terminus that determine the type of prenylation at single or dual cysteines. Prenylation occurs in the ER and is often part of a stepwise process of PTMs that is followed by proteolytic cleavage by Rce1 and methylation by isoprenyl cysteine methyltransferase (ICMT).

Proteolysis

Peptide bonds are indefinitely stable under physiological conditions, and therefore cells require some mechanism to break these bonds. Proteases comprise a family of enzymes that cleave the peptide bonds of proteins and are critical in antigen processing, apoptosis, surface protein shedding and cell signaling.

The family of over 11,000 proteases varies in substrate specificity, mechanism of peptide cleavage, location in the cell and the length of activity. While this variation suggests a wide array of functionalities, proteases can generally be separated into groups based on the type of proteolysis. Degradative proteolysis is critical to remove unassembled protein subunits and misfolded proteins and to maintain protein concentrations at homeostatic concentrations by reducing a given protein to the level of small peptides and single amino acids. Proteases also play a biosynthetic role in cell biology that includes cleaving signal peptides from nascent proteins and activating zymogens, which are inactive enzyme precursors that require cleavage at specific sites

for enzyme function. In this respect, proteases act as molecular switches to regulate enzyme activity.

Proteolysis is a thermodynamically favorable and irreversible reaction. Therefore, protease activity is tightly regulated to avoid uncontrolled proteolysis through temporal and/or spatial control mechanisms including regulation by cleavage in cis or trans and compartmentalization (e.g., proteasomes, lysosomes).

The diverse family of proteases can be classified by the site of action, such as aminopeptidases and carboxypeptidase, which cleave at the amino or carboxy terminus of a protein, respectively. Another type of classification is based on the active site groups of a given protease that are involved in proteolysis. Based on this classification strategy, greater than 90% of known proteases fall into one of four categories as follows:

- Serine proteases
- Cysteine proteases
- Aspartic acid proteases
- Zinc metalloproteases

Inhibitors of protein synthesis

It is a substance that stops or slows the growth or proliferation of cells by disrupting the processes that lead directly to the generation of new proteins.

While a broad interpretation of this definition could be used to describe nearly any antibiotic, in practice, it usually refers to substances that act at the ribosome level (either the ribosome itself or the translation factor), taking advantages of the major differences between prokaryotic and eukaryotic ribosome structures.

Toxins such as ricin also function via protein synthesis inhibition. Ricin acts at the eukaryotic 60S.Examples:

- **Neomycin**
- **Geneticin, also called G418**

Mechanism

In general, protein synthesis inhibitors work at different stages of prokaryotic mRNA translation into proteins, like initiation, elongation (including aminoacyl tRNA entry, proofreading, peptidyl transfer, and ribosomal translocation) and termination:

Earlier stages

- **Rifampicin inhibits prokaryotic DNA transcription into mRNA by inhibiting DNA-dependent RNA polymerase by binding its beta-subunit.**

Initiation

Linezolid acts at the initiation stage, probably by preventing the formation of the initiation complex, although the mechanism is not fully understood.

Aminoacyl tRNA entry

- **Tetracyclines and Tigecycline (a glycylcycline related to tetracyclines) block the A site on the ribosome, preventing the binding of aminoacyl tRNAs.**

Proofreading

- **Aminoglycosides, among other potential mechanisms of action, interfere with the proofreading process, causing increased rate of error in synthesis with premature termination.**

Peptidyl transfer

- **Chloramphenicol blocks the peptidyl transfer step of elongation on the 50S ribosomal subunit in both bacteria and mitochondria.**
- **Macrolides (as well as inhibiting ribosomal translocation and other potential mechanisms) bind to the 50s ribosomal subunits, inhibiting peptidyl transfer.**
- **Quinupristin/dalfopristin act synergistically, with dalfopristin, enhancing the binding of quinupristin, as well as inhibiting peptidyl transfer. Quinupristin binds to a nearby site on the 50S ribosomal subunit and prevents elongation of the polypeptide, as well as causing incomplete chains to be released.**

Ribosomal translocation

- Macrolides, clindamycin and aminoglycosides (with all these three having other potential mechanisms of action as well), have evidence of inhibition of ribosomal translocation.
- Fusidic acid prevents the turnover of elongation factor G (EF-G) from the ribosome.

Termination

- Macrolides and clindamycin (both also having other potential mechanisms) cause premature dissociation of the peptidyl-tRNA from the ribosome.
- Puromycin has a structure similar to that of the tyrosinyl aminoacyl-tRNA. Thus, it binds to the ribosomal A site and participates in peptide bond formation, producing peptidyl-puromycin. However, it does not engage in translocation and quickly dissociates from the ribosome, causing a premature termination of polypeptide synthesis.
- Streptogramins also cause premature release of the peptide chain.

Protein synthesis inhibitors of unspecified mechanism

- Retapamulin

Binding site

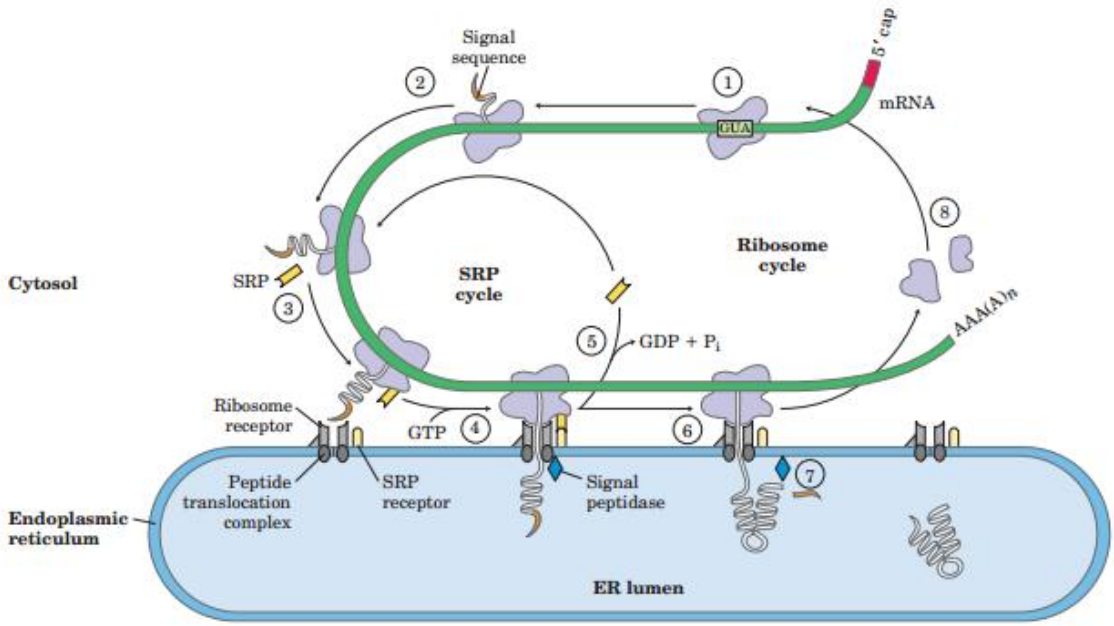
The following antibiotics bind to the 30S subunit of the ribosome:

- Aminoglycosides
- Tetracyclines

The following antibiotics bind to the 50S ribosomal subunit:

- Chloramphenicol
- Erythromycin
- Clindamycin
- Linezolid
- Telithromycin
- Streptogramins
- Retapamulin

Directing eukaryotic proteins with the appropriate signals to the endoplasmic reticulum



Posttranslational mdification of many eukaryotic proteins begins in the endoplasmic reticulum. Most lysosomal, membrane, or secreted proteins have an amino-terminal signal sequence that marks them for translocation into the lumen of the ER; The carboxyl terminus of the signal sequence is defined by a cleavage site, where protease action removes the sequence after the protein is imported into the ER. Signal sequences vary in length from 13 to 36 amino acid residues, but all have the following features: (1) about 10 to 15 hydrophobic amino acid residues; (2) one or more positively charged residues, usually near the amino terminus, preceding the hydrophobic sequence; and (3) a short sequence at the carboxyl terminus (near the cleavage site) that is relatively polar, typically having amino acid residues with short side chains (especially Ala) at the positions closest to the cleavage site. Proteins with these signal sequences are synthesized on ribosomes attached to the ER. The signal sequence itself helps to direct the ribosome to the ER.

1. The targeting pathway begins with initiation of protein synthesis on free ribosomes.

2. The signal sequence appears early in the synthetic process, because it is at the amino terminus,

which is synthesized first.

3. As it emerges from the ribosome, the signal sequence and the ribosome itself are bound by the

large signal recognition particle (SRP); SRP then binds GTP and halts elongation of the

polypeptide when it is about 70 amino acids long and the signal sequence has completely

emerged from the ribosome.

4. The GTP-bound SRP now directs the ribosome (still bound to the mRNA) and the incomplete

polypeptide to GTP-bound SRP receptors in the cytosolic face of the ER; the nascent

polypeptide is delivered to a peptide translocation complex in the ER, which may interact

directly with the ribosome.

5. SRP dissociates from the ribosome, accompanied by hydrolysis of GTP in both SRP and the

SRP receptor.

6. Elongation of the polypeptide now resumes, with the ATP-driven translocation complex

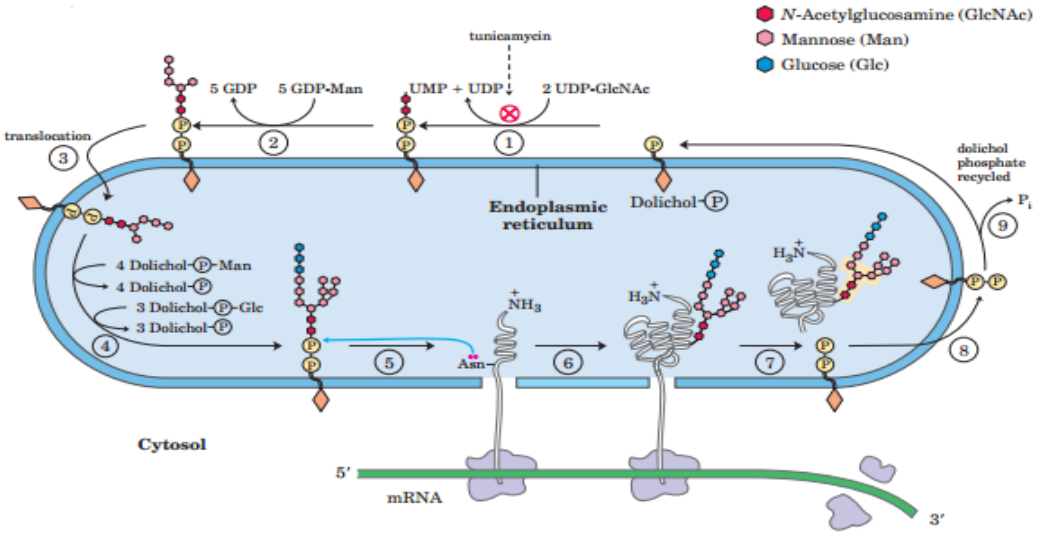
feeding the growing polypeptide into the ER lumen until the complete protein has been

synthesized.

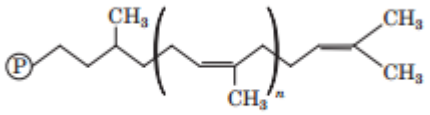
7. The signal sequence is removed by a signal peptidase within the ER lumen

8. The ribosome dissociates and is recycled.

Role of glycosylation in protein targeting



In the ER lumen, following the removal of signal sequences, polypeptides are folded, disulfide bonds formed, and many proteins glycosylated to form glycoproteins. In many glycoproteins the linkage to their oligosaccharides is through Asn residues. A 14 residue core oligosaccharide is built up in a step-wise fashion, and then transferred from a dolichol phosphate donor molecule to certain Asn residues in the protein. The transferase is on the luminal face of the ER and thus cannot catalyze glycosylation of cytosolic proteins. After transfer, the core oligosaccharide is trimmed and elaborated in different ways on different proteins, but all N-linked oligosaccharides retain a pentasaccharide core derived from the original 14 residue oligosaccharide.



Dolichol phosphate
(n = 9–22)

Antibiotics such as tunicamycin act by interfering with one or more steps in this process. It mimics the structure of UDP-N-acetylglucosamine and blocks the first step of the process. A few proteins are O-glycosylated in the ER, but most O-glycosylation occurs in the Golgi complex or in the cytosol (for proteins that do not enter the ER). Suitably modified proteins can now be moved to a variety of intracellular destinations. Proteins travel from the ER to the Golgi complex in transport vesicles. In the Golgi complex, oligosaccharides are O-linked to some proteins, and N-linked oligosaccharides are further modified. By mechanisms not yet fully understood, the Golgi complex also sorts proteins and sends them to their final destinations. The processes that segregate proteins targeted for secretion from those targeted for the plasma membrane or lysosomes must distinguish among these proteins on the basis of structural features other than signal sequences, which were removed in the ER lumen.

This sorting process is best understood in the caseof hydrolases destined for transport to lysosomes. The three-dimensional structure of the hydrolase is recognized by a phosphotransferase, which phosphorylates certain mannose residues in the oligosaccharide. The presence of one or more mannose 6-phosphate residues in its N-linked oligosaccharide is the structural signal that targets the protein to lysosomes. A receptor protein in the membrane of the Golgi complex recognizes the mannose 6-phosphate signal and binds the hydrolase so marked. Vesicles containing these receptor-hydrolase complexes bud from the trans side of the Golgi complex and make their way

to sorting vesicles. Here, the receptor-hydrolase complex dissociates in a process facilitated by the lower pH in the vesicle and by phosphatase-catalyzed removal of phosphate groups from the mannose 6-phosphate residues. The receptor is then recycled to the Golgi complex, and vesicles containing the hydrolases bud from the sorting vesicles and move to the lysosomes. In cells treated with tunicamycin hydrolases that should be targeted for lysosomes are instead secreted, confirming that the N-linked oligosaccharide plays a key role in targeting these enzymes to lysosomes.

The pathways that target proteins to mitochondria and chloroplasts also rely on amino-terminal signal sequences. Although mitochondria and chloroplasts contain DNA, most of their proteins are encoded by nuclear DNA and must be targeted to the appropriate organelle. Unlike other targeting pathways, however, the mitochondrial and chloroplast pathways begin only after a precursor protein has been completely synthesized and released from the ribosome. Precursor proteins destined for mitochondria or chloroplasts are bound by cytosolic chaperone proteins and delivered to receptors on the exterior surface of the target organelle. Specialized translocation mechanisms then transport the protein to its final destination in the organelle, after which the signal sequence is removed.

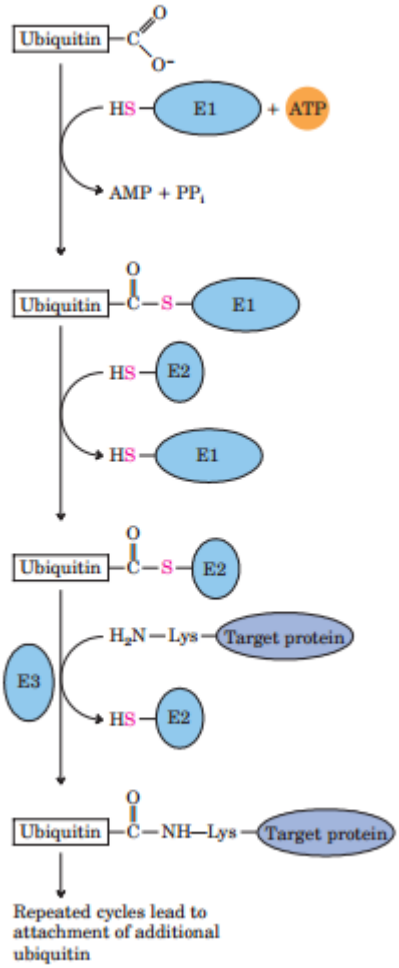
Protein degradation-Ubiquitin mediated pathway:

Protein degradation prevents the buildup of abnormal or unwanted proteins and permits the recycling of amino acids. The half-lives of eukaryotic proteins vary from 30 seconds to many days. Most proteins turn over rapidly relative to the lifetime of a cell, although a few (such as hemoglobin) can last for the life of the cell (about 110 days for an erythrocyte). Rapidly degraded proteins include those that are defective because of incorrectly inserted amino acids or because of damage accumulated during normal functioning. And enzymes that act at key regulatory points in metabolic pathways often turn over rapidly.

Defective proteins and those with characteristically short half-lives are generally degraded in both bacterial and eukaryotic cells by selective ATP-dependent cytosolic systems. A second system in vertebrates, operating in lysosomes, recycles the amino acids of mem-brane proteins, extracellular proteins, and proteins with characteristically long half-lives. In E. coli, many proteins are degraded by an ATP-dependent protease called Lon (the name refers to the “long form” of proteins, observed only when this pro-tease is absent). The protease is activated in the presence of defective proteins or those

slated for rapid turnover; two ATP molecules are hydrolyzed for every peptide bond cleaved. The precise role of this ATP hydrolysis is not yet clear. Once a protein has been reduced to small inactive peptides, other ATP-independent proteases complete the degradation process.

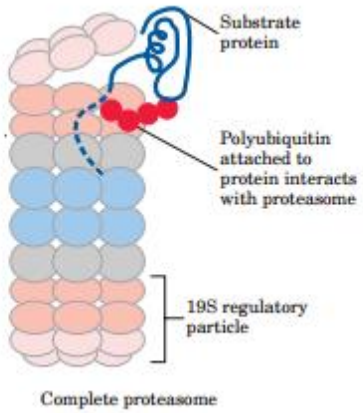
The ATP-dependent pathway in eukaryotic cells is quite different, involving the protein ubiquitin,which, as its name suggests, occurs throughout the eukaryotic kingdoms. One of the most highly conserved proteins known, ubiquitin (76 amino acid residues) is essentially identical in organisms as different as yeasts and humans. Ubiquitin is covalently linked to proteins slated for de-struction via an ATP-dependent pathway involving three separate enzymes (E1, E2, and E3).



Ubiquitinated proteins are degraded by a large complex known as the 26S proteasome. The proteasome consists of two copies each of at least 32 different subunits, most of which are highly conserved from yeasts to humans. The proteasome contains two main types of subcomplexes, a barrel-like core particle and regulatory

particles on either end of the barrel. The 20S core particle consists of four rings; the outer rings are formed from seven α subunits, and the inner rings from seven β subunits. Three of the seven subunits in each β ring have protease activities, each with different substrate specificities. The stacked rings of the core particle form the barrel-like structure within which target proteins are degraded. The 19S regulatory particle on each end of the core particle con-tains 18 subunits, including some that recognize and bind to ubiquitinated proteins. Six of the subunits are ATPases that probably function in unfolding the ubiquitinated proteins and translocating the unfolded polypeptide into the core particle for degradation.

Although we do not yet understand all the signals that trigger ubiquitination, one simple signal has been found. For many proteins, the identity of the first residue that remains after removal of the amino-terminal Met residue, and any other posttranslational proteolytic processing of the amino-terminal end, has a profound influence on half-life. These amino-terminal signals have been conserved over billions of years of evolution, and are the same in bacterial protein degradation systems and in the human ubiquitination pathway. More complex signals, are also being identified. Ubiquitin-dependent proteolysis is as important for the regulation of cellular processes as for the elimination of defective proteins. Many proteins required at only one stage of the eukaryotic cell cycle are rapidly degraded by the ubiquitin-dependent pathway after completing their function. The same pathway also processes and presents class I MHC antigens. Ubiquitin dependent destruction of cyclin is critical to cell-cycle regulation. The E2 and E3 components of the ubiquitination cascade pathway are in fact two large families of proteins. Different E2 and E3 enzymes exhibit different specificities for target proteins and thus regulate different cellular processes. Some E2 and E3 enzymes are highly localized in certain cellular compartments, reflecting a specialized function.



Complete proteasome

Not surprisingly, defects in the ubiquitination pathway have been implicated in a wide range of disease states. An inability to degrade certain proteins that activate cell division (the products of oncogenes) can lead to tumor formation, whereas a too-rapid degradation of proteins that acts as tumor suppressors can have the same effect. The ineffective or overly rapid degradation of cellular proteins also appears to play a role in a range of other conditions: renal diseases, asthma, neurodegenerative disorders such as Alzheimer's and Parkinson's diseases (associated with the formation of characteristic proteinaceous structures in neurons), cystic fibrosis (caused in some cases by a too-rapid degradation of a chloride ion channel, with resultant loss of function; Liddle's syndrome (in which a sodium channel in the kidney is not degraded, leading to excessive Na^+ absorption and early-onset hypertension)—and many other disorders. Drugs designed to inhibit proteasome function are being developed as potential treatments for some of these conditions. In a changing metabolic environment, protein degradation is as important to a cell's survival as is protein synthesis, and much remains to be learned about these interesting pathways.

POSSIBLE QUESTIONS

UNIT IV

Part-B

1. Explain the role of ribosomal RNA in translation
2. Describe the action of any three inhibitors in protein synthesis
3. Explain the wobble hypothesis
4. How does amino acyl tRNA recognize its corresponding tRNA?
5. Write the composition of eukaryotic and prokaryotic ribosomes
6. Explain the proof reading activity of amino acyl tRNA synthetase
7. What are the post translational modifications of proteins?
8. How are aminoacid activated?
9. Write in brief about the termination process of translation
10. How proteins of endoplasmic reticulum reach their destination after the synthesis.

Part-C

1. Describe the protein synthesis in prokaryotes
2. Give an account on inhibitors of protein synthesis
3. Explain the targeting of proteins to

a) Nucleus

b) Endoplasmic reticulum

c) Golgi complex

4. List out the post translational modification of proteins

5. Explain about the fidelity of transcription

II BSc Biochemistry -Fourth Semester
GENE EXPRESSION AND REGULATION (17BCU402)
Unit III

S.No.	Questions	Opt 1	Opt 2	Opt 3	Opt 4	Answer
1	Which one of the following bases if present in the third position of an anticodon cannot form wobbl base pairing	G	C	U	A	A
2	. The binding of $tRNA^{fmet}$ to the P site is inhibited by	Streptomycin	. Chloramphenicol	Erythromycin	Lincomycin	Streptomycin
3	Translocase is an enzyme required in the process of	.DNA replication	RNA synthesis	Initiation of protein synthesis	. Elongation of peptides	. Elongation of peptides
4	. which one of the following post translational modifications occurs in clotting factors?	hydroxylation	methylation	carboxylation	phosphorylation	carboxylation
5	. Diphtheria toxin inhibits elongation step in translation through	Glycosylation	Methylation	ADP- ribosylation	Ubiquitinylation	ADP- ribosylation
6	The Shine Dalgarno sequence responsible for initiation of protein synthesis is found in	. bacterial mRNA	bacterial & eukaryotic mRNA	Bacterial rRNA	. bacterial & eukaryotic rRNA	. bacterial mRNA
7	Selection of translational initiation site is carried out by	Base pairing between mRNA and 16rRNA	initiation codon	termination codon	promoter	Base pairing between mRNA and 16rRNA
8	How many polypeptide chains can be formed simultaneously by a given ribosome?	one	. about a dozen	. up to 30	variable depending on mRNA	one
9	. RF1 recognises	UAA	. GUG	UGA	AUG	UAA
10	The antibiotic cycloheximide	blocks chain elongation during protein synthesis	blocks glycosylation of proteins	resembles aminoacyl tRNA	inhibits the formation of peptide bond.	inhibits the formation of peptide bond.
11	Large subunit of ribosome is present in eukaryotes	60s	. 40s	30s	80s	60s
12	Large subunit of ribosome is present in prokaryotes	50s	. 40s	20s	30s	50s
13	Which of the following inhibits the peptidyl transferase activity of the eukaryotic 60S ribosomal subunit	Chloroamphenical	Cycloheximide	Erythromycin	Puromycin	Cycloheximide
14	in protein aminoacids arew joined by _____ bond	peptide	phosphodiester	disulfide	phosphodiester and disulfide	peptide
15	Which of the following Binds to 50S subunit and inhibits translocation	Chloroamphenical	Cycloheximide	Erythromycin	Puromycin	Erythromycin
16	Eukaryotic cells have _____ kinds of RNA polymerase	1	2	3	4	3
17	Eukaryotic mRNA was synthesised b	DNA Pol I	RNA Pol I	RNA Pol II	RNA Pol III	RNA Pol II
18	Principle function of RNA pol II in Eukaryote is	Synthesis mRNA	Synthesis tRNA	Synthesis rRNA	splicing of tRNA	Synthesis mRNA
19	chloramphenicol inhibits	Peptidyl transferase	Binding of amino acyl tRNA	translocation	Initiation of translation	Peptidyl transferase
20	The association of 50S and 30Ssubunit is preveted by	IF 1	IF 2	IF 3	EF2	IF 3
21	Premature termination of poly peptide is caused by	Erythromycin	puromycin	Streptomycin	Ricin	puromycin
22	The codon that terminates protein synthesis is	UGG	UGA	UGC	UGU	UGA
23	Which of the following antibiotics mimics amino acyl tRNA	. Ricin	Puromycin	Streptomycin	Erythromycin	Puromycin
24	The drug chloramphenical blocks	cellwall formation	transcription	translation termination release factors	polypeptide chain elongation.	polypeptide chain elongation.
25	The Wobble hypothesis accounts for the	triple nature of codon	degeneracy of codon	universality of the codon	doublet nature of the codon	degeneracy of codon
26	Shine Dalgarno sequence is	rich in pyrimindine	. rich in purine	centered 100 nucleotide upstream from the start codon	. centered 100 nucleotide downstream from the start codon	. rich in purine
27	Which of the following amino acids is specified by only a single codon?	. Glutamine	. tryptophan	asparagine	isoleucine	isoleucine
28	The enzyme responsible the cleavage of the bond connecting amino acid and tRNA is	amino acyl-tRNA synthetase	peptidyl transferase	aminoacyl transferase	tRNA deacylase	tRNA deacylase
29	Post dimer initiation is responsible for	recombination repair	SOS repair	excision repair	photo reactivation	recombination repair
30	Coupled transcription and translation occurs	only in eucaryoptes	only in prokaryotes	both in prokaryotes and in eukaryotes	neither in prokaryotes nor in eukaryotes	only in prokaryotes
31	Which of the following antibiotics is an inhibitor of protein synthesis	Chloramphenicol	Penicillin	Bacitracin	Cephalosporin	Chloramphenicol
32	The amino acyl synthetase must	recognize the codon	recognize the anticodon	be able to distinguish one amino acid from another d	recognize the DNA	be able to distinguish one amino acid from another d
33	Mannose 6 phospahte target a protein to	plasma membrane	golgicomplex	lysosomes	endoplasmic reticulam	lysosomes
34	_____are proteins that block the synthesis of RNA at gene level	Initiators	Repressors	stimulators	retarders	Repressors
35	tRNA, 5S RNA, Special RNA s are synthesised by RNA Pol_____	I	II	III	IV	III
36	Catalytic RNA is	Abzyme	ribozyme	enzyme	protein	ribozyme
37	One of the following is distinctive Structure of eukarytic mRNA	poly A tail	poly T tail	polu U tail	poly G tail	poly A tail
38	Large protein complex that involved in intron removal is	replisome	primosome	spliceosome	none	spliceosome
39	5' cap in mRNA is done by	ATP	GTP	CTP	UTP	GTP
40	Transcription is controlled by all except	Repressor	Activators	Specificity factors	release factors	release factors
41	Repressor binding site on DNA is	Activator	Operator	enhancer	stimulators.	Operator
42	Operator binds _____ and regulate gene expression	Activator	repressor	primer	stimulator	repressor
43	House keeping gene is the gene that expressed	rarely	constantly	over expressed	under expressed	constantly
44	RNA polymerase binds to DNA at	activator	Operator	Promoter	repressor	Promoter

45	The Holliday model provides a molecular basis for	the association between aberrant segregation and crossing over	the activation of transcription	the repression of early transcription	DNA repair	the association between aberrant segregation and crossing over
46	Rec BCD protein is a multifunctional protein with an	endonuclease activity	exonuclease activity	polymerase activity	ligase activity	endonuclease activity
47	Rec A performs all of the following except	it coats the ss region and polymerizes	it catalyses the invitro formation of Holliday structures	It participates in DNA repair by regulating SOS response	It is involved in the DNA replication	It is involved in the DNA replication
48	Rec BCD is also known as	Exonuclease N	Exonuclease B	Exonuclease V	Exonuclease D	Exonuclease V
49	Recombination in Meselson-Radding model is initiated by	ss nicks	ds nicks	exchange process	mispairing	ss nicks
50	Rotation of the holliday structure at the crossover site forms a	rotational isomer	Circular isomer	Crosslinking isomer	exchange isomer	rotational isomer
51	Site specific recombination require the recognition of unique nucleotide sequences in both DNA molecules by enzymes	recombinase	Helicase	Isomerase	ligase	recombinase
52	Rotational isomer is also called as	Holliday stucture	Radding structure	circular structure	None	Holliday stucture
53	Who gave the basis of the most popular current models for the molecular events of recombination	Robin holliday	Matthew Meselson and charles Radding	Arthur Kornberg	Messelson Stahl	Robin holliday
54	Who suggested mechanism for creating cross-strand Holliday structure?	Matthew Meselson and charles Radding	Robin holliday	Arthur Kornberg	Messelson Stahl	Robin holliday
55	A molecular signal that binds to repressor and control the gene expression is called	activator	operator	Promoter	effector	operator
56	Effector molecule cause _____regulation	Positive	Negative	Positive and negative	feed back inhibition	Negative
57	Enhancers binds _____ and regulate gene expression	Activator	repressor	primer	stimulator	Activator
58	Lac repressor is a	Monomer	Dimer	Trimer	Tetramer	Tetramer
59	Binding of _____ by Lac repressor cause its dissociation from the operator site	Arabinose	glucose	allolactose	Allontoin	allolactose
60	Structural analog of allolactose is	IPTG	ITGP	IBGT	IBTG	IPTG
	IPTG acts as _____on Lac operon	inducer	operator	Promoter	Effector	inducer

Unit 4

Regulation of gene expression in prokaryotes

Principles of gene regulation, negative and positive regulation, concept of operons, regulatory proteins, activators, repressors, DNA binding domains, regulation of lac operon and trp operon, induction of SOS response, synthesis of ribosomal proteins, regulation by genetic recombination, transcriptional regulation in λ bacteriophage.

Introduction

In genetics, an operon is a functioning unit of genomic material containing a cluster of genes under the control of a single regulatory signal or promoter[1]. The genes are transcribed together into an mRNA strand and either translated together in the cytoplasm, or undergo trans-splicing to create monocistronic mRNAs that are translated separately, i.e. several strands of mRNA that each encode a single gene product. The result of this is that the genes contained in the operon are either expressed together or not at all. Originally operons were thought to exist solely in prokaryotes but since the discovery of the first operons in eukaryotes in the early 1990s[2][3], more evidence has arisen to suggest they are more common than previously assumed[4]. Several genes must be both co-transcribed and co-regulated to define an operon.

Structure

This is the general structure of an operon:

- Promoter – a nucleotide sequence that enables a gene to be transcribed. The promoter is recognized by RNA polymerase, which then initiates transcription. In RNA synthesis, promoters indicate which genes should be used for messenger RNA creation – and, by extension, control which proteins the cell manufactures.

- Operator – a segment of DNA that a regulator binds to. It is classically defined in the lac operon as a segment between the promoter and the genes of the opero. In the case of a repressor, the repressor protein physically obstructs the RNA polymerase from transcribing the genes.

- Structural genes – the genes that are co-regulated by the operon.

Regulation

Control of an operon is a type of gene regulation that enables organisms to regulate the expression of various genes depending on environmental conditions. Operon regulation can be either negative or positive by induction or repression.[6]

Negative control involves the binding of a repressor to the operator to prevent transcription.

- In negative inducible operons, a regulatory repressor protein is normally bound to the operator and it prevents the transcription of the genes on the operon. If an inducer molecule is present, it binds to the repressor and changes its conformation so that it is unable to bind to the operator. This allows for expression of the operon.

- In negative repressible operons, transcription of the operon normally takes place. Repressor proteins are produced by a regulator gene but they are unable to bind to the operator in their normal conformation. However certain molecules called corepressors are bound by the repressor protein, causing a conformational change to the active state. The activated repressor protein binds to the operator and prevents transcription.

Operons can also be positively controlled. With positive control, an activator protein stimulates transcription by binding to DNA (usually at a site other than the operator).

- In positive inducible operons, activator proteins are normally unable to bind to the pertinent DNA. When an inducer is bound by the activator protein, it undergoes a change in conformation so that it can bind to the DNA and activate transcription.

- In positive repressible operons, the activator proteins are normally bound to the pertinent DNA segment. However, when a corepressor is bound by the activator, it is prevented from binding the DNA. This stops activation and transcription of the system.

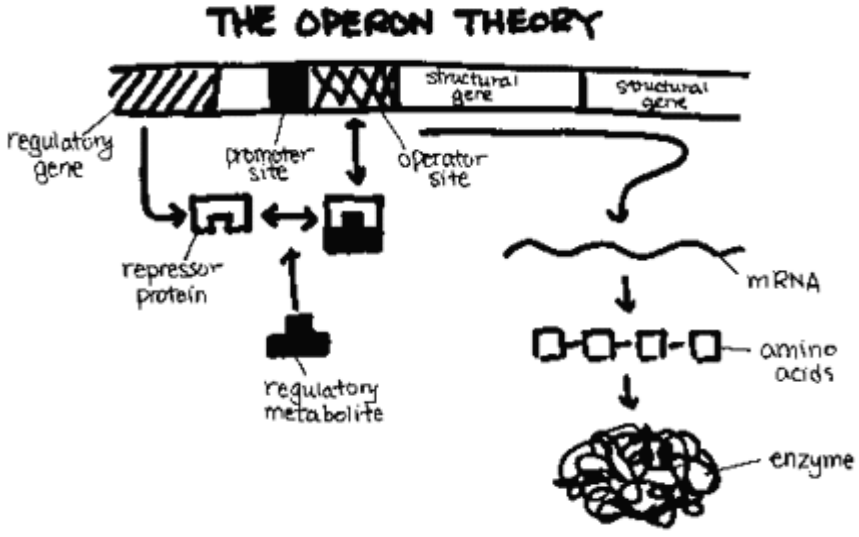


Figure 5.1 The operon theory

Inducible genes - The operon model

Definition

An inducible gene is a gene that is expressed in the presence of a substance (an inducer) in the environment. This substance can control the expression of one or more genes (structural genes) involved in the metabolism of that substance. For example, lactose induces the expression of the *lac* genes that are involved in lactose metabolism. An certain antibiotic may induce the expression of a gene that leads to resistance to that antibiotic.

Induction is common in metabolic pathways that result in the catabolism of a substance and the inducer is normally the substrate for the pathway

Lactose Operon

Structural genes

The lactose operon (figure 1) contains three structural genes that code for enzymes involved in lactose metabolism.

- The *lac z* gene codes for β-galactosidase, an enzyme that breaks down lactose into glucose and galactose
- The *lac y* gene codes for a permease, which is involved in uptake of lactose
- The *lac a* gene codes for a galactose transacetylase.

These genes are transcribed from a common promoter into a polycistronic mRNA, which is translated to yield the three enzymes.

Regulatory gene

The expression of the structural genes is not only influenced by the presence or absence of the inducer, it is also controlled by a specific regulatory gene. The regulatory gene may be next to or far from the genes that are being regulated. The regulatory gene codes for a specific protein product called a REPRESSOR.

Operator

The repressor acts by binding to a specific region of the DNA called the operator which is adjacent to the structural genes being regulated. The structural genes together with the operator region and the promoter are called an OPERON. However, the binding of the repressor to the operator is prevented by the inducer and the inducer can also remove repressor that has already bound to the operator. Thus, in the presence of the inducer the repressor is inactive and does not bind to the operator, resulting in transcription of the structural genes. In contrast, in the absence of inducer the repressor is active and binds to the operator, resulting in inhibition of transcription of the

structural genes. This kind of control is referred to a NEGATIVE CONTROL since the function of the regulatory gene product (repressor) is to turn off transcription of the structural genes.

Inducer

Transcription of the lac genes is influenced by the presence or absence of an inducer (lactose or other β-galactosides) (Figure 2).

e.g	+ inducer	expression
	- inducer	no expression

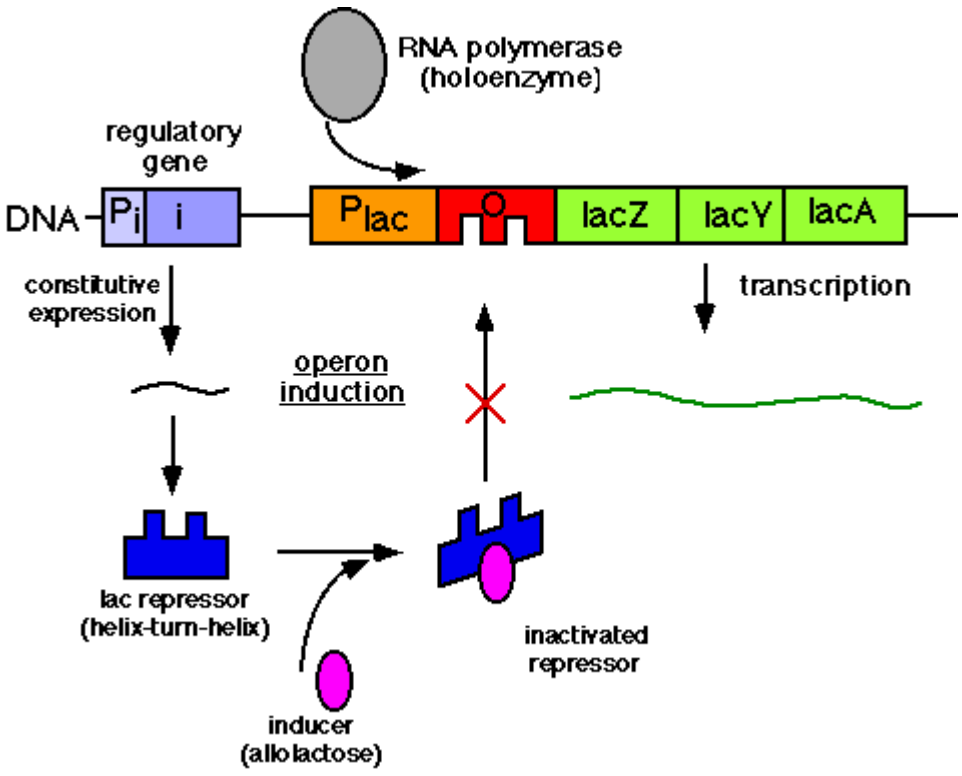


Figure 5.2 Lac operon

Catabolite repression (Glucose Effect)

Many inducible operons are not only controlled by their respective inducers and regulatory genes, but they are also controlled by the level of glucose in the environment. The ability of glucose to control the expression of a number of different inducible operons is called

Catabolite Repression

Catabolite repression is generally seen in those operons which are involved in the degradation of compounds used as a source of energy. Since glucose is the preferred energy source in bacteria, the ability of glucose to regulate the expression of other operons ensures that bacteria will utilize glucose before any other carbon source as a source of energy.

Mechanism

There is an inverse relationship between glucose levels and cyclic AMP (cAMP) levels in bacteria. When glucose levels are high cAMP levels are low and when glucose levels are low cAMP levels are high. This relationship exists because the transport of glucose into the cell inhibits the enzyme adenyl cyclase which produces cAMP. In the bacterial cell cAMP binds to a cAMP binding protein called CAP or CRP. The cAMP-CAP complex, but not free CAP protein, binds to a site in the promoters of catabolite repression-sensitive operons. The binding of the complex results in a more efficient promoter and thus more initiations of transcriptions from that promoter as illustrated in Figures 4 and 5.Since the role of the CAP-cAMP complex is to turn on transcription this type of control is said to be POSITIVE CONTROL. The consequences of this type of control is that to achieve maximal expression of a catabolite repression sensitive operon glucose must be absent from the environment and the inducer of the operon must be present. If both are present, the operon will not be maximally expressed until glucose is metabolized. Obviously, no expression of the operon will occur unless the inducer is present.

The Tryptophan Operon

Repressible genes - The operon model

Definition

Repressible genes are those in which the presence of a substance (a co-repressor) in the environment turns off the expression of those genes (structural genes) involved in the metabolism of that substance.

e.g., Tryptophan represses the expression of the trp genes.

Repression is common in metabolic pathways that result in the biosynthesis of a substance and the co-repressor is normally the end product of the pathway being regulated.

Structural genes

The tryptophan operon (figure 6) contains five structural genes that code for enzymes involved in the synthesis of tryptophan. These genes are transcribed from a common promoter into a polycistronic mRNA, which is translated to yield the five enzymes.

Regulatory gene

The expression of the structural genes is not only influenced by the presence or absence of the co-repressor, it is also controlled by a specific regulatory gene. The regulatory gene may be next to or far from the genes that are being regulated. The regulatory gene codes for a specific protein product called an REPRESSOR (sometimes called an apo-repressor). When the repressor is synthesized it is inactive. However, it can be activated by complexing with the co-repressor (*i.e.* tryptophan).

Operator

The active repressor/co-repressor complex acts by binding to a specific region of the DNA called the operator which is adjacent to the structural genes being regulated. The structural genes together with the operator region and the promoter is called an OPERON. Thus, in the presence of the co-repressor the repressor is active and binds to the operator, resulting in repression of transcription of the structural genes. In contrast, in the absence of co-repressor the repressor is inactive and does not bind to the operator, resulting in transcription of the structural genes. This kind of control is referred to a NEGATIVE CONTROL since the function of the regulatory gene product (repressor) is to turn off transcription of the structural genes.

Co-repressor

Transcription of the tryptophan genes is influenced by the presence or absence of a co-repressor (tryptophan) (Figure 7).

e.g.	+co-repressor	no expression
	-co-	expression

repressor

Attenuation

In many repressible operons, transcription that initiates at the promoter can terminate prematurely in a leader region that precedes the first structural gene. (*i.e.* the polymerase terminates transcription before it gets to the first gene in the operon). This phenomenon is called ATTENUATION; the premature termination of transcription. Although attenuation is seen in a number of operons, the mechanism is best understood in those repressible operons involved in amino acid biosynthesis. In these instances attenuation is regulated by the availability of the cognate aminoacylated t-RNA.

Mechanism

When transcription is initiated at the promoter, it actually starts before the first structural gene and a leader transcript is made. This leader region contains a start and a stop signal for protein synthesis. Since bacteria do not have a nuclear membrane, transcription and translation can occur simultaneously. Thus, a short peptide can be made while the RNA polymerase is transcribing the leader region. The test peptide contains several tryptophan residues in the middle of the peptide. Thus, if there is a sufficient amount of tryptophanyl-t-RNA to translate that test peptide, the entire peptide will be made and the ribosome will reach the stop signal. If, on the other hand, there is not enough tryptophanyl-t-RNA to translate the peptide, the ribosome will be arrested at the two tryptophan codons before it gets to the stop signal.

The sequence in the leader m-RNA contains four regions, which have complementary sequences (Figure 9). Thus, several different secondary stem and loop structures can be formed. Region 1 can only form base pairs with region 2; region 2 can form base pairs with either region 1 or 3; region 3 can form base pairs with region 2 or 4; and region 4 can only form base pairs with region 3. Thus three possible stem/loop structures can be formed in the RNA.

region 1:region 2

region 2:region 3

region 3:region 4

One of the possible structures (region 3 base pairing with region 4) generates a signal for RNA polymerase to terminate transcription (*i.e.* to attenuate transcription). However, the formation of one stem and loop structure can preclude the formation of others. If region 2 forms base pairs with region 1 it is not available to base pair with region 3. Similarly if region 3 forms base pairs with region 2 it is not available to base pair with region 4.

The ability of the ribosomes to translate the test peptide will affect the formation of the various stem and loop structures Figure 10. If the ribosome reaches the stop signal for translation it will be covering up region 2 and thus region 2 will not available for forming base pairs with other regions. This allows the generation of the transcription termination signal because region 3 will be available to pair with region 4. Thus, when there is enough tryptophanyl-t-RNA to translate the test peptide attenuation will occur and the structural genes will not be transcribed. In contrast, when there is an insufficient amount of tryptophanyl-t-RNA to translate the test peptide no attenuation will occur. This is because the ribosome will stop at the two tryptophan codons in region 1, thereby allowing region 2 to base pair with region 3 and preventing the formation of the attenuation signal (*i.e.* region 3 base paired with region 4). Thus, the structural genes will be transcribed.

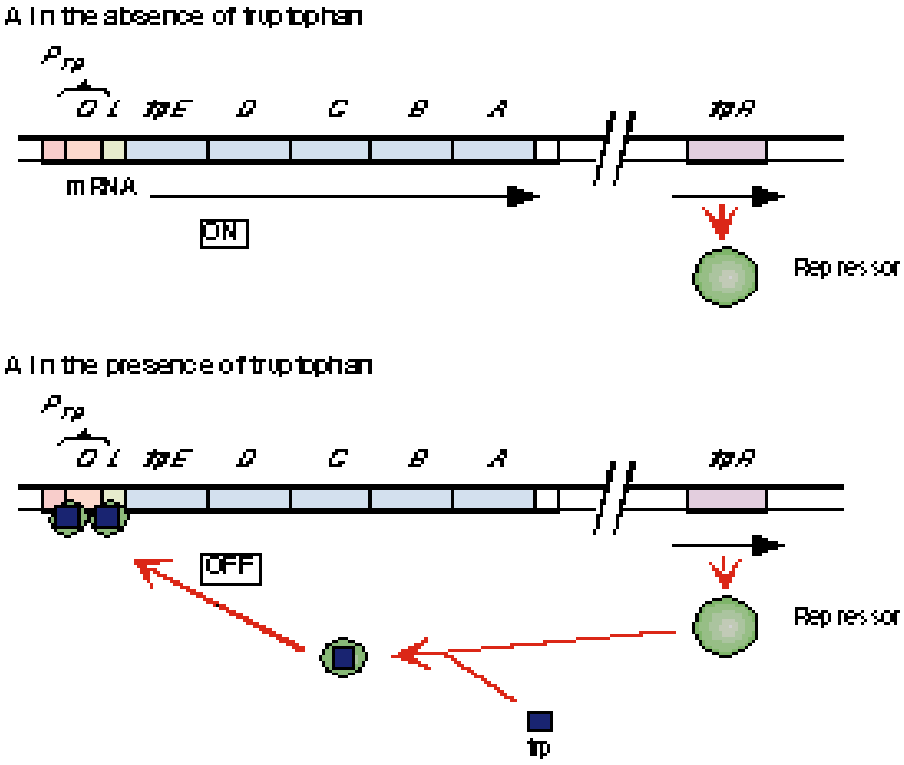
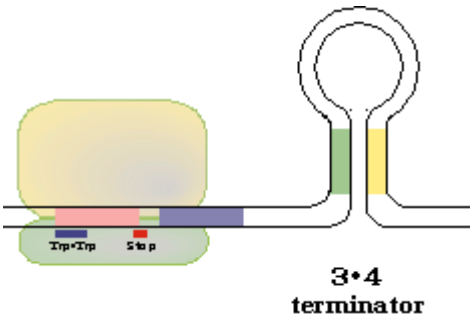


Figure 5.3 Trp operon



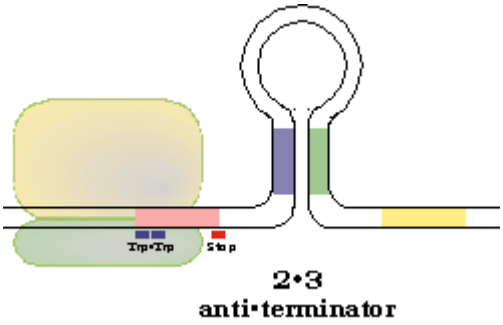


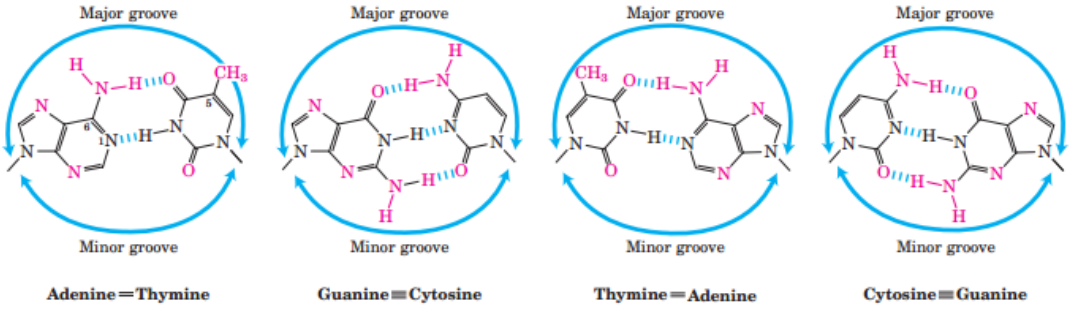
Figure 5.4 Termination process

Regulatory proteins

DNA binding domain

Regulatory proteins generally bind to specific DNA sequences. Their affinity for these target sequences is roughly 10^4 to 10^6 times higher than their affinity for any other DNA sequences. Most regulatory proteins have discrete DNA-binding domains containing sub-structures that interact closely and specifically with the DNA. These binding domains usually include one or more of a relatively small group of recognizable and characteristic structural motifs.

To bind specifically to DNA sequences, regulatory proteins must recognize surface features on the DNA. Most of the chemical groups that differ among the four bases and thus permit discrimination between base pairs are hydrogen-bond donor and acceptor groups exposed in the major groove of DNA , and most of the protein-DNA contacts that impart specificity are hydrogen bonds.



A notable exception is the nonpolar surface near C-5 of pyrimidines, where thymine is readily distinguished from cytosine by its protruding methyl group.

Protein-DNA contacts are also possible in the minor groove of the DNA, but the hydrogen-bonding patterns here generally do not allow ready discrimination between base pairs. Within regulatory proteins, the amino acid side chains most often hydrogen-bonding to bases in the DNA are those of Asn, Gln, Glu, Lys, and Arg residues. Is there

a simple recognition code in which a particular amino acid always pairs with a particular base? The two hydrogen bonds that can form between Gln or Asn and the N6 and N-7 positions of adenine cannot form with any other base. And an Arg residue can form two hydrogen bonds with N-7 and O6 of guanine.

Examination of the structures of many DNA-binding proteins, however, has shown that a protein can recognize each base pair in more than one way, leading to the conclusion that there is no simple amino acid–base code. For some proteins, the Gln-adenine interaction can specify AUT base pairs, but in others a van der Waals pocket for the methyl group of thymine can recognize AUT base pairs. To interact with bases in the major groove of DNA, a protein requires a relatively small structure that can stably protrude from the protein surface. The DNA-binding domains of regulatory proteins tend to be small (60 to 90 amino acid residues), and the structural mo-tifs within these domains that are actually in contact with the DNA are smaller still. Many small proteins are unstable because of their limited capacity to form layers of structure to bury hydrophobic groups.

The DNA-binding motifs provide either a very compact stable structure or a way of allowing a segment of protein to protrude from the protein surface. The DNA-binding sites for regulatory proteins are often inverted repeats of a short DNA sequence (a palindrome) at which multiple (usually two) subunits of a regulatory protein bind cooperatively. The Lac repressor is unusual in that it functions as a tetramer, with two dimers tethered together at the end distant from the DNA-binding sites. An E. coli cell normally contains about 20 tetramers of the Lac repressor. Each of the tethered dimers separately binds to a palindromic operator sequence, in contact with 17 bp of a 22 bp region in the lac operon. And each of the tethered dimers can independently bind to an operator sequence, with one generally binding to $O_1$and the other to O_2or O_3. The symmetry of the O1 operator sequence corresponds to the twofold axis of symmetry of two paired Lac repressor subunits. The tetrameric Lac repressor binds to its operator sequences in vivo with an estimated dissociation constant of about 10^{-10} M. The repressor discriminates between the operators and other sequences by a factor of about 106, so binding to these few base pairs among the 4.6 million or so of the E. coli chromosome is highly specific. Several DNA-binding motifs have been described, but here we focus on two that play prominent roles in the binding of DNA by regulatory proteins: the helix-turn-helix and the zinc finger. We also consider a type of DNA-binding domain—the homeodomain—found in some eukaryotic proteins.

Helix-Turn-Helix

This DNA-binding motif is crucial to the interaction of many prokaryotic regulatory proteins with DNA, and similar motifs occur in some eukaryotic regulatory proteins. The helix-turn-helix motif comprises about 20 amino acids in two short α-helical segments, each seven to nine amino acid residues long, separated by a β turn. This structure generally is not stable by itself; it is simply the reactive portion of a somewhat

larger DNA-binding domain. One of the two α-helical segments is called the recognition helix, be-cause it usually contains many of the amino acids that interact with the DNA in a sequence-specific way. This α helix is stacked on other segments of the protein structure so that it protrudes from the protein surface. When bound to DNA, the recognition helix is positioned in or nearly in the major groove. The Lac repressor has this DNA-binding motif

Zinc Finger

In a zinc finger, about 30 amino acid residues form an elongated loop held together at the base by a single Zn2+ ion, which is coordinated to four of the residues (four Cys, or two Cys and two His). The zinc does not itself interact with DNA; rather, the coordination of zinc with the amino acid residues stabilizes this small structural motif. Several hydrophobic side chains in the core of the structure also lend stability.

Many eukaryotic DNA-binding proteins contain zinc fingers. The interaction of a single zinc finger with DNA is typically weak, and many DNA-binding proteins, like Zif268, have multiple zinc fingers that substantially enhance binding by interacting simultaneously with the DNA. One DNA-binding protein of the frog Xenopus has 37 zinc fingers. There are few known examples of the zinc finger motif in prokaryotic proteins.

Some zinc fingers contain the amino acid residues that are important in sequence discrimination, whereas others appear to bind DNA nonspecifically (the amino acids required for specificity are located elsewhere in the protein). Zinc fingers can also function as RNA-binding motifs—for example, in certain proteins that bind eukaryotic mRNAs and act as translational repressors.

Homeodomain

Another type of DNA-binding domain has been identified in a number of proteins that function as transcriptional regulators, especially during eukaryotic development. This domain of 60 amino acids—called the homeodomain, because it was discovered in homeotic genes (genes that regulate the development of body pat-terns)—is highly conserved and has now been identified in proteins from a wide variety of organisms, including humans DNA-binding segment of the domain is related to the helix-turn-helix motif. The DNA sequence that encodes this domain is known as the homeobox.

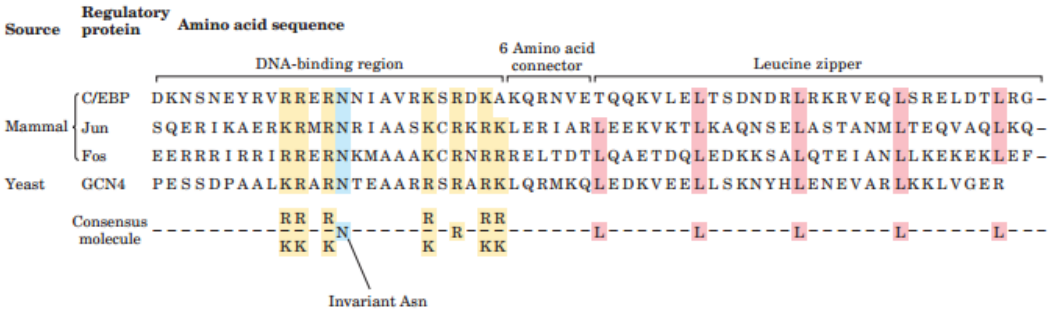
Protein protein interaction domain

Regulatory proteins contain domains not only for DNA binding but also for protein-protein interactions—with RNA polymerase, other regulatory proteins, or other sub-

units of the same regulatory protein. Examples include many eukaryotic transcription factors that function as gene activators, which often bind as dimers to the DNA, using DNA-binding domains that contain zinc fingers. Some structural domains are devoted to the interactions required for dimer formation, which is generally a pre-requisite for DNA binding. Like DNA-binding motifs, the structural motifs that mediate protein-protein interactions tend to fall within one of a few common categories.

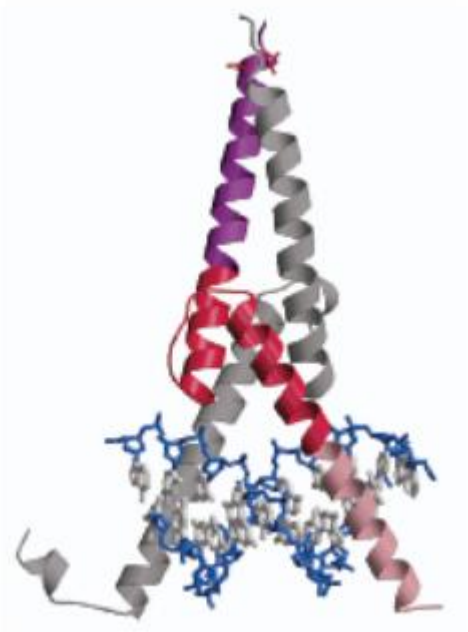
Two important examples are the leucine zipper and the basic helix-loop-helix. Structural motifs such as these are the basis for classifying some regulatory proteins into structural families.

Leucine Zipper This motif is an amphipathic α helix with a series of hydrophobic amino acid residues concentrated on one side, with the hydrophobic surface forming the area of contact between the two polypeptides of a dimer.



A striking feature of these α helices is the occurrence of Leu residues at every seventh position, forming a straight line along the hydrophobic surface. Although researchers initially thought the Leu residues interdigitated (hence the name "zipper"), we now know that they line up side by side as the interacting α helices coil around each other forming a coiled coil. Regulatory proteins with leucine zippers often have a separate DNA-binding domain with a high concentration of basic (Lys or Arg) residues that can interact with the negatively charged phosphates of the DNA backbone. Leucine zippers have been found in many eukaryotic and a few prokaryotic proteins.

Basic Helix-Loop-Helix Another common structural motif occurs in some eukaryotic regulatory proteins implicated in the control of gene expression during the development of multicellular organisms. These proteins share a conserved region of about 50 amino acid residues important in both DNA binding and protein dimerization. This region can form two short amphipathic α helices linked by a loop of variable length, the helix-loop-helix (distinct from the helix-turn-helix motif associated with DNA binding). The helix-loop-helix motifs of two polypeptides interact to form dimmers.



In these proteins, DNA binding is mediated by an adjacent short amino acid sequence rich in basic residues, similar to the separate DNA-binding region in proteins containing leucine zippers.

II BSc Biochemistry -Fourth Semester
GENE EXPRESSION AND REGULATION (17BCU402)
Unit IV

S.No.	Questions	Opt 1	Opt 2	Opt 3	Opt 4	Answer
1	In lac operon the effect of glucose is mediated by the activator protein	CRP	CPR	PCR	PRC	CRP
2	Which of the following aminoacid synthesis s controlled by transcription attenuation	Glycine	tryptophan	gly	cys	tryptophan
3	The leader region of try operon contains---------- nucleotides	150	245	176	162	162
4	Attenuator structure in trp operon is due to joining ___& ____sequence	1&2	2&3	3&4	1&4	3&4
5	Trp repressor is	Hetero dimer	homodimer	Heterotetramer	Homotetramer	homodimer
6	The effect of gluose on CRP is mediated by	AMP	cAMP	GMP	cGMP	cAMP
7	The leader region of trp operon contains ____Trp residues	2	4	5	6	2
8	Leu operon leader peptide has_____ Leu residues	2	4	5	6	4
9	His operon leader peptide has_____ His residues	2	4	5	7	7
10	Ara operon is controlled by _____ regulatory protein	AraA	Ara B	Ara C	Ara D	Ara C
11	Hormone receptor complex binds to _____element in DNA	HRE	HER	RHE	REH	HRE
12	Which of the following RNA cause gene silencing	snRNA	hnRNA	siRNA	ssRNA	siRNA
13	Expression of gene is shut down by one of the following method	PCR	RNAi	Southern blot	Western blot	RNAi
14	Antisnse RNA is related	PCR	RNAi	Southern blot	Western blot	RNAi
15	Trancriptionally active chromatins are	hetero chramatin	Euchromatin	chromospme	polytene chromosome	Euchromatin
16	Trancriptionally inactive chromatins are	hetero chramatin	Euchromatin	chromospme	polytene chromosome	hetero chramatin
17	Double strand RNA is degraded by	endouclease	protease	RNase	Dnase	RNase
18	mRNA that result in gene product is	Antisense RNA	Sense RNA	hnRNA	snRNA	Sense RNA
19	mRNA that result in destruction of gene product is	Antisense RNA	Sense RNA	hnRNA	snRNA	Antisense RNA
20	Holliday model of recombination requires	Sequence homology and same length of DNA	no sequence homology	Different length	no sequence homologyand same length of DNA	Sequence homology and same length of DNA
21	Mesolson radding model of recombination occur if DNA have	Sequence homology and same length of DNA	no sequence homology	Sequence homology and different length	no sequence homology andsame length of DNA	Sequence homology and different length
22	Site specific recombination requires	Sequence homology and same length of DNA	no sequence homology	Sequence homology and different length	same length of DNA	no sequence homology
23	Enzyme involved in Site specific recombination is	helicase	integrase	polymerase	topoisomerase	integrase
24	tRNA acts as _____ molcule	intergrator	Adaptor	regulator	promoter	Adaptor
25	The macromolecule regarded as the reserve bank of genetic information	Protein	DNA	RNA	Polysaccharide	DNA
26	The biological information flows from DNA to RNA and from RNA to	lipids	carbohydrates	proteins	nucleotides	proteins
27	The total genetic information contained in a DNA is referred to as	gene	genome	Okazaki piece	ribozyme	genome
28	The DNA base pairing is based on _____ rules	Watson & Crick	Arther Kornberg	Stahl & Meselson	McClintock	Watson & Crick
29	Pseudogenes are	Related to non functional genes	Transcribed into mRNA	Translated in to functional proteins	Transcribed into tRNA	Related to non functional genes
30	Mobile genetic elements were visualized by	T.H Morgan	Barbara McClintock	G Khorana	C.B Bridge	Barbara McClintock
31	Fundamental unit of DNA organization	Replisome	Nucleosome	Primosome	Chromatin	Nucleosome
32	Uptake of naked DNA is called	Transduction	Transformation	Conjugation	Recombination	Transformation
33	Transduction experiment was conducted by	Frederick Griffith	Zinder and Ledenberg	Hershey and Chase	Avery and Mc Carty	Hershey and Chase
34	Histones are rich in	Arg and Lys	Lys & Gly	Arg & Glu	Arg & Gly	Arg and Lys
35	Which histone protein is not a part of core particle of nucleosome?	H1	H2a	H2b	H4	H1
36	Which histone protein is involved in the transition between the solenoid form and the extended nucleosome form?	H1	H2a	H2b	H4	H1
37	Transfer of genetic material through bacteriophage is called	Transduction	Transformation	Conjugation	Recombination	Transduction
38	Blendor experiment is involved in	Transduction	Transformation	Conjugation	Recombination	Transduction
39	Smaller blocks occur at the end of chromosome arm is called _____	centromere	telomere	blastomeres	blastocyst	telomere
40	The functional unit of DNA is	genome	gene	nucleotide	chromosome	nucleotide
41	The coiling that cannot be separated except by unwinding is called	supercoiling	negative super coiling	plectonemic coiling	anti parellel coiling	plectonemic coiling
42	Self integrating DNA fragments are known as	Transposons	Self posons	Transducers	Transfragments	Transposons
43	Strain used for conjugation is	Pneumococci	Streptococci	Staphylococci	Bacillus	Pneumococci
44	Transformation experiment was conducted by	Griffith & Avery et al	Hershey and Chase	Watson & Crick	Zinder and Ledenberg	Griffith & Avery et al
45	P^{32} labelling was done to bacteriophage	DNA	Protein coat	Lipid	polysaccharide	DNA

46	S^{35} labelling was done to bacteriophage	DNA	Protein coat	Lipid	Polysaccharide	Protein coat
47	The molecular chaperone which causes the nucleosome assembly is	Nucleoplasmin	Histone	Hu protein	Ubiquitin	Histone
48	Simple sequence DNA is concentrated in	Centromere	Telomere	Blastomeres	Blastocyst	Centromere
49	LTRs are absent in	LINES	SINES	LINES AND SINES	Viral retrotransposons	LINES AND SINES
50	Chicken lysozyme gene is a good example of	Single copy DNA	Moderately repetitive DNA	Simple sequence DNA	Highly repetitive DNA	Single copy DNA
51	Difference in length of simple sequence tandem arrays helps to develop a technique called	Foot printing	Northern blotting	Western	Finger printing	Finger printing
52	Important hallmark of IS element is	Short direct repeats	Long direct repeats	Inverted repeats	Tandem repeats	Short direct repeats
53	IS element contain _____ enzyme	Helicase	primase	Transposase	topoisomerase	transposase
54	The length of the DNA segment present in the nucleosome core particle is	140 bp	200 bp	166 bp	114 bp	166 bp
55	The housekeeping genes have	GC box	TATA box	Pribnow box	CAAT box	GC box
56	Negative supercoils are removed by	Topoisomerase I	Gyrase	Helicase	Rep protein	Topoisomerase I
57	DNA chains differ from each other by one nucleotide can be revolved using	20% agarose	20% polyacrylamide	1% agarose	1% polyacrylamide	1% agarose
58	Satellite DNAs are found in the region of	Euchromatin	Heterochromatin	Hypervariable regions	Functional elements	Heterochromatin
59	Polytene chromosomes are produced by	Gene inversion	Gene conversion	DNA amplification	Retrotransposons	DNA amplification
60	In conjugation DNA is transformed from	Conjugation tube	Bacteriophage	Dead organism	Mice	Dead organism

Unit 5

Regulation of gene expression in eukaryotes

Heterochromatin, euchromatin, chromatin remodeling, regulation of galactose metabolism in yeast, regulation by phosphorylation of nuclear transcription factors, regulatory RNAs, riboswitches, RNA interference, synthesis and function of miRNA molecules, phosphorylation of nuclear transcription factors.

Regulation of gene expression in eukaryotes

Transcriptionally active chromatin

About 10% of the chromatin in a typical eukaryotic cell is in a more condensed form than the rest of the chromatin. This form, heterochromatin, is transcriptionally inactive. Heterochromatin is generally associated with particular chromosome structures—the centromeres, for example. The remaining, less condensed chromatin is called euchromatin.

Transcription of a eukaryotic gene is strongly repressed when its DNA is condensed within heterochromatin. Some, but not all, of the euchromatin is transcriptionally active. Transcriptionally active chromosomal regions can be detected based on their increased sensitivity to nuclease-mediated degradation. Nucleases such as DNase I tend to cleave the DNA of carefully isolated chromatin into fragments of multiples of about 200 bp, reflecting the regular repeating structure of the nucleosome. In actively transcribed regions, the fragments produced by nuclease activity are smaller and more heterogeneous in size. These regions contain hypersensitive sites, sequences especially sensitive to DNase I, which consist of about 100 to 200 bp within the 1,000 bp flanking the 5’ ends of transcribed genes. In some genes, hypersensitive sites are found farther from the 5’ end, near the 3’ end, or even within the gene itself. Many hypersensitive sites correspond to binding sites for known regulatory proteins, and the relative absence of nucleosomes in these regions may allow the binding of these proteins. Nucleosomes are entirely absent in some regions that are very active in transcription, such as the rRNA genes.

Transcriptionally active chromatin tends to be deficient in histone H1, which binds to the linker DNA between nucleosome particles. Histones within transcriptionally active chromatin and heterochromatin also differ in their patterns of covalent modification. The core histones of nucleosome particles (H2A, H2B, H3, H4) are modified by irreversible methylation of Lys residues, phosphorylation of Ser or Thr residues, acetylation , or attachment of ubiquitin . Each of the core histones has two distinct structural domains. A central domain is involved in histone-histone interaction and the wrapping of DNA around the nucleosome. A second, lysine-rich amino-terminal domain is generally positioned near the exterior of the assembled

nucleosome particle; the covalent modifications occur at specific residues concentrated in this amino-terminal domain. The patterns of modification have led some researchers to propose the existence of a histone code, in which modification patterns are recognized by enzymes that alter the structure of chromatin. Modifications associated with transcriptional activation would be recognized by enzymes that make the chromatin more accessible to the transcription machinery.

5-Methylation of cytosine residues of CpG sequences is common in eukaryotic DNA but DNA in transcriptionally active chromatin tends to be undermethylated. Furthermore, CpG sites in particular genes are more often undermethylated in cells from tissues where the genes are expressed than in those where the genes are not expressed. The overall pattern sug-gests that active chromatin is prepared for transcription by the removal of potential structural barriers.

Chromatin remodeling

The detailed mechanisms for transcription-associated structural changes in chromatin, called chromatin remodeling, are now coming to light, including identification of a variety of enzymes directly implicated in the process. These include enzymes that covalently modify the core histones of the nucleosome and others that use the chemical energy of ATP to remodel nucleosomes on the DNA. The acetylation and deacetylation of histones figure prominently in the processes that activate chromatin for transcription. As noted above, the amino-terminal domains of the core histones are generally rich in Lys residues. Particular Lys residues are acetylated by histone acetyltransferases (HATs). Cytosolic (type B) HATs acetylate newly synthesized histones before the histones are imported into the nucleus. The subsequent assembly of the histones into chromatin is facilitated by additional proteins: CAF1 for H3 and H4, and NAP1 for H2A and H2B.

Where chromatin is being activated for transcription, the nucleosomal histones are further acetylated by nuclear (type A) HATs. The acetylation of multiple Lys residues in the amino-terminal domains of histones H3 and H4 can reduce the affinity of the entire nucleosome for DNA. Acetylation may also prevent or promote interactions with other proteins involved in transcription or its regulation. When transcription of a gene is no longer required, the acetylation of nucleosomes in that vicinity is reduced by the activity of histone deacetylases, as part of a general gene-silencing process that restores the chromatin to a transcriptionally inactive state. In addition to the removal of certain acetyl groups, new covalent modification of histones marks chromatin as transcriptionally inactive. As an example, the Lys residue at position 9 in histone H3 is often methylated in heterochromatin.

Chromatin remodeling also requires protein complexes that actively move or displace nucleosomes, hydrolyzing ATP in the process. The enzyme complex SWI/SNF found in all eukaryotic cells, contains 11 polypeptides that together create hypersensitive sites in the chromatin and stimulate the binding of transcription factors. SWI/SNF is not required for the transcription of every gene. NURF is another ATP-dependent enzyme complex that remodels

chromatin in ways that complement and overlap the activity of SWI/SNF. These enzyme complexes play an important role in preparing a region of chromatin for active transcription.

DNA binding transactivator and coactivator

To continue our exploration of the regulation of gene expression in eukaryotes, we return to the interactions between promoters and RNA polymerase II (Pol II), the enzyme responsible for the synthesis of eukaryotic mRNAs. Although most (but not all) Pol II promoters include the TATA box and Inr (initiator) sequences, with their standard spacing, they vary greatly in both the number and the location of additional sequences required for the regulation of transcription. These additional regulatory sequences are usually called enhancers in higher eukaryotes and upstream activator sequences (UASs) in yeast. A typical enhancer may be found hundreds or even thousands of base pairs upstream from the transcription start site, or may even be downstream, within the gene itself. When bound by the appropriate regulatory proteins, an enhancer increases transcription at nearby promoters regardless of its orientation in the DNA. The UASs of yeast function in a similar way, although generally they must be positioned upstream and within a few hundred base pairs of the transcription start site. An average Pol II promoter may be affected by a half-dozen regulatory sequences of this type, and even more complex promoters are quite common.

Successful binding of active RNA polymerase II holoenzyme at one of its promoters usually requires the action of other proteins. These are of three types:

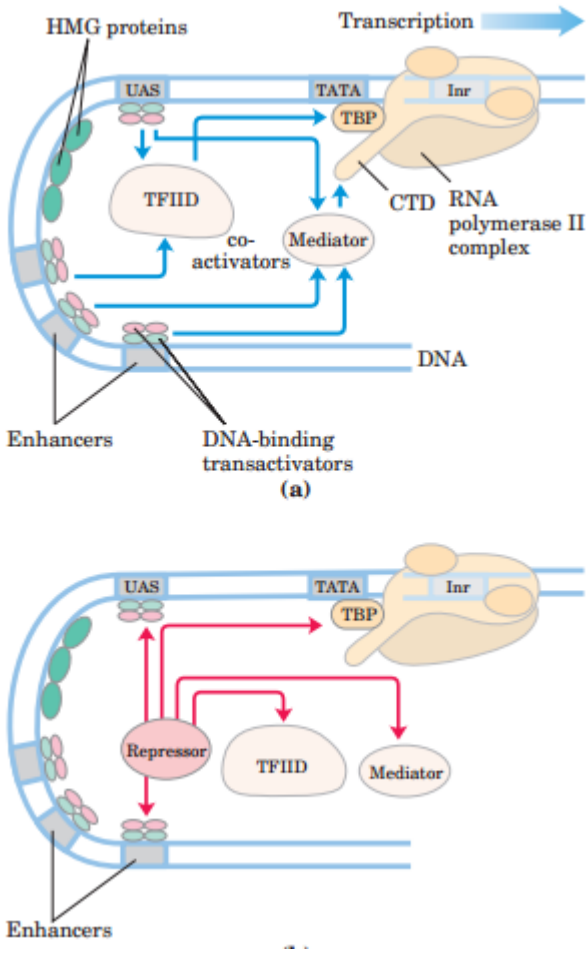


Fig: Eukaryotic promoters and regulatory proteins

(1)basal transcription factors required at every Pol II promoter;

(2) DNA-binding transactivators, which bind to enhancers or UASs and facilitate transcription; and (3) coactivators.The latter group act indirectly—not by binding to the DNA—and are required for essential communication between the DNA-binding transactivators and the complex composed of Pol II and the general transcription factors.

Furthermore, a variety of repressor proteins can interfere with communication between the RNA polymerase and the DNA-binding transactivators, resulting in repression of transcription.

TATA-Binding Protein - The first component to bind in the assembly of a preinitiation complex at the TATA box of a typical Pol II promoter is the TATA-binding protein (TBP). The complete complex includes the basal (or general) transcription factors TFIIB, TFIIE, TFIIF, TFIIH; Pol II; and perhaps TFIIA (not all of the factors are shown in Fig.). This minimal preinitiation complex, however, is often insufficient for the initiation of transcription

and generally does not form at all if the promoter is obscured within chromatin. Positive regu-lation leading to transcription is imposed by the trans-activators and coactivators.

DNA-Binding Transactivators - The requirements for trans-activators vary greatly from one promoter to another. A few transactivators are known to facilitate transcription at hundreds of promoters, whereas others are specific for a few promoters. Many transactivators are sensitive to the binding of signal molecules, providing the capacity to activate or deactivate transcription in response to a changing cellular environment. Some enhancers bound by DNA-binding transactivators are quite distant from the promoter's TATA box. How do the transactivators function at a distance? The answer in most cases seems to be that, as indicated earlier, the intervening DNA is looped so that the various protein complexes can interact directly. The looping is promoted by certain non-histone proteins that are abundant in chromatin and bind nonspecifically to DNA. This high mobility group (HMG) proteins (Fig "high mobility" refers to their electrophoretic mobility in polyacrylamide gels) play an important structural role in chromatin remodeling and transcriptional activation.

Coactivator Protein Complexes - Most transcription requires the presence of additional protein complexes. Some major regulatory protein complexes that interact with Pol II have been defined both genetically and biochemically. These coactivator complexes act as intermediaries between the DNA-binding transactivators and the Pol II complex.

The best-characterized coactivator is the transcription factor TFIID. In eukaryotes, TFIID is a large complex that includes TBP and ten or more TBP-associated factors (TAFs). Some TAFs resemble histones and may play a role in displacing nucleosomes during the activation of transcription. Many DNA-binding transactivators aid in transcription initiation by interacting with one or more TAFs. The requirement for TAFs to initiate transcription can vary greatly from one gene to another. Some promoters require TFIID, some do not, and some require only subsets of the TFIID TAF subunits.

Another important coactivator consists of 20 or more polypeptides in a protein complex called mediator(Fig); the 20 core polypeptides are highly conserved from fungi to humans. Mediator binds tightly to the carboxyl-terminal domain (CTD) of the largest subunit of Pol II. The mediator complex is required for both basal and regulated transcription at promoters used by Pol II, and it also stimulates the phosphorylation of the CTD by TFIIH. Both mediator and TFIID are required at some promoters. As with TFIID, some DNA-binding transactivators interact with one or more components of the mediator complex. Coactivator complexes function at or near the promoter's TATA box.

Choreography of Transcriptional Activation - We can now begin to piece together the sequence of transcriptional activation events at a typical Pol II promoter. First, crucial remodeling of the chromatin takes place in stages. Some DNA-binding transactivators have

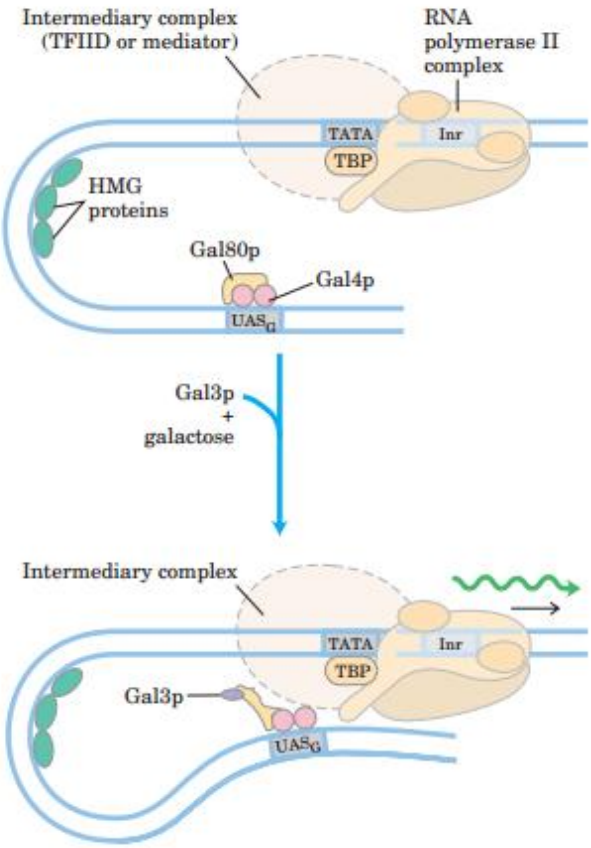
significant affinity for their binding sites even when the sites are within condensed chromatin. Binding of one transacti-vator may facilitate the binding of others, gradually dis-placing some nucleosomes. The bound transactivators can then interact directly with HATs or enzyme complexes such as SWI/SNF (or both), accelerating the remodeling of the surrounding chromatin. In this way a bound transactivator can draw in other components necessary for further chromatin remodeling to permit transcription of specific genes. The bound transactivators, generally acting through complexes such as TFIID or mediator (or both), stabilize the binding of Pol II and its associated transcription factors and greatly facilitate formation of the preinitiation transcription complex. Complexity in these regulatory circuits is the rule rather than the exception, with multiple DNA-bound transactivators promoting transcription.

The script can change from one promoter to another, but most promoters seem to require a precisely ordered assembly of components to initiate transcription. The assembly process is not always fast. At some genes it may take minutes; at certain genes in higher eukaryotes the process can take days.

Reversible Transcriptional Activation - Although rarer, some eukaryotic regulatory proteins that bind to Pol II promoters can act as repressors, inhibiting the formation of active preinitiation complexes. Some transactivators can adopt different conformations, enabling them to serve as transcriptional activators or repressors. For example, some steroid hormone receptors function in the nucleus as DNA-binding transactivators, stimulating the transcription of certain genes when a particular steroid hormone signal is present. When the hormone is absent, the receptor proteins revert to a repressor conformation, prevent-ing the formation of preinitiation complexes. In some cases this repression involves interaction with histone deacetylases and other proteins that help restore the surrounding chromatin to its transcriptionally inactive state.

Both Positive and Negative Regulation Of The Genes of Galactose Metabolism in Yeast

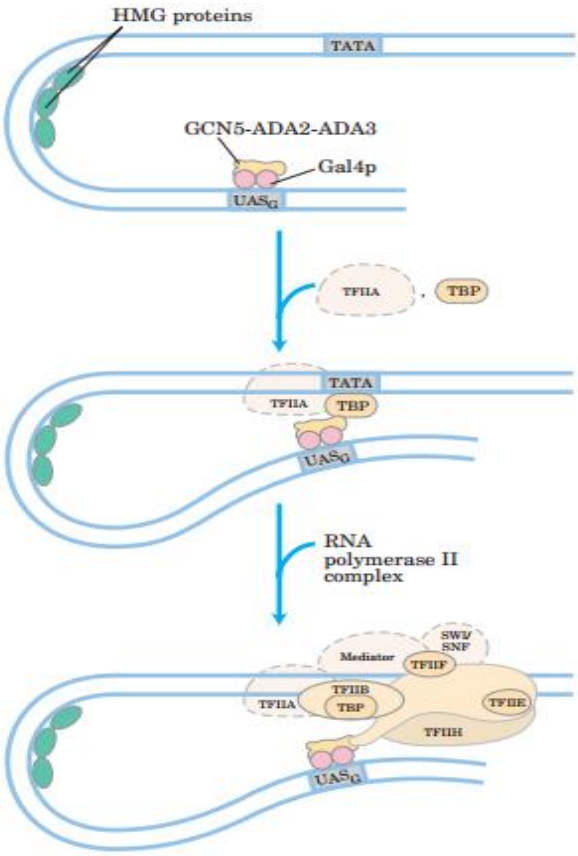
The enzymes required for the importation and metabolism of galactose in yeast are encoded by genes scattered over several chromosomes. Each of theGALgenes is transcribed separately, and yeast cells have no operons like those in bacteria. However, all the GALgenes have similar pro-moters and are regulated coordinately by a common set of proteins. The promoters for the GALgenes consist of the TATA box and Inr sequences, as well as an up-stream activator sequence (UASG) recognized by a DNA-binding transcriptional activator known as Gal4 protein (Gal4p). Regulation of gene expression by galac-tose entails an interplay between Gal4p and two other proteins, Gal80p and Gal3p (Fig. 28–28). Gal80p forms a complex with Gal4p, preventing Gal4p



from functioning as an activator of the GALpromoters. When galactose is present, it binds Gal3p, which then interacts with Gal80p, allowing Gal4p to function as an activator at the various GALpromoters.

Other protein complexes also have a role in activating transcription of the GALgenes. These may include the SAGA complex for histone acetylation, the SWI/SNF complex for nucleosome remodeling, and the mediator complex. The following figure provides an idea of the complexity of protein interactions in the overall process of transcriptional activation in eukaryotic cells.

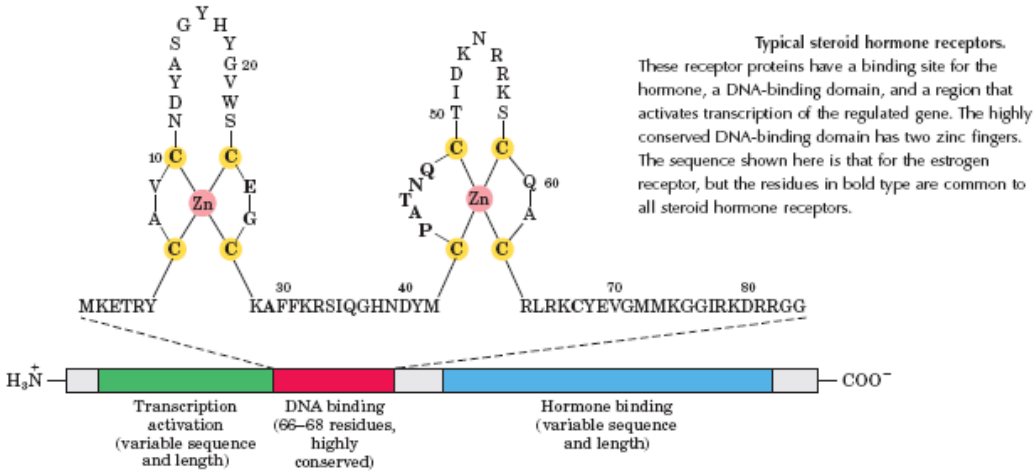
Glucose is the preferred carbon source for yeast, as it is for bacteria. When glucose is present, most of the GALgenes are repressed—whether galactose is present or not. The GALregulatory system described above is effectively overridden by a complex catabolite repression system that includes several proteins (not depicted in the fig).



Regulation of gene expression by intracellular and intercellular signal

The effects of steroid hormones (and of thyroid and retinoid hormones, which have the same mode of action) provide additional well studied examples of the modulation of eukaryotic regulatory proteins by direct interaction with molecular signals. Unlike other types of hormones, steroid hormones do not have to bind to plasma membrane receptors. Instead, they can interact with intracellular receptors that are themselves transcriptional transactivators. Steroid hormones too hydrophobic to dissolve readily in the blood (estrogen, progesterone, and cortisol, for example) travel on specific carrier proteins from their point of release to their target tissues. In the target tissue, the hormone passes through the plasma membrane by simple diffusion and binds to its specific receptor protein in the nucleus. The hormone-receptor complex acts by binding to highly specific DNA sequences called **hormone response elements (HREs),** thereby altering gene expression. Hormone binding triggers changes in the conformation of the receptor proteins so that they become capable of interacting with additional transcription factors. The bound hormone-receptor complex can either enhance or suppress the expression of adjacent genes. The DNA sequences (HREs) to which hormonereceptor complexes bind are similar in length and arrangement, but differ in sequence, for the various steroid hormones. Each receptor has a consensus HRE sequence to

which the hormone-receptor complex binds well, with each consensus consisting of two six-nucleotide sequences, either contiguous or separated by three nucleotides, in tandem or in a palindromic arrangement. The hormone receptors have a highly conserved DNA-binding domain with two zinc fingers.



Typical steroid hormone receptors. These receptor proteins have a binding site for the hormone, a DNA-binding domain, and a region that activates transcription of the regulated gene. The highly conserved DNA-binding domain has two zinc fingers. The sequence shown here is that for the estrogen receptor, but the residues in bold type are common to all steroid hormone receptors.

The hormone-receptor complex binds to the DNA as a dimer, with the zinc finger domains of each monomer recognizing one of the six-nucleotide sequences. The ability of a given hormone to act through the hormone-receptor complex to alter the expression of a specific gene depends on the exact sequence of the HRE, its position relative to the gene, and the number of HREs associated with the gene.

Unlike the DNA-binding domain, the ligand-binding region of the receptor protein—always at the carboxyl terminus—is quite specific to the particular receptor. In the ligand-binding region, the glucocorticoid receptor is only 30% similar to the estrogen receptor and 17% similar to the thyroid hormone receptor. The size of the ligand- binding region varies dramatically; in the vitamin D receptor it has only 25 amino acid residues, whereas in the mineralocorticoid receptor it has 603 residues. Mutations that change one amino acid in these regions can result in loss of responsiveness to a specific hormone. Some humans unable to respond to cortisol, testosterone, vitamin D, or thyroxine have mutations of this type.

Regulation Can Result from Phosphorylation of Nuclear Transcription Factors

We noted in Chapter 12 that the effects of insulin on gene expression are mediated by a series of steps leading ultimately to the activation of a protein kinase in the nucleus that phosphorylates specific DNA-binding proteins and thereby alters their ability to act as transcription factors (see Fig. 12–6). This general mechanism mediates the effects of many nonsteroid hormones. For example, the _-adrenergic pathway that leads to elevated

levels of cytosolic cAMP, which acts as a second messenger in eukaryotes as well as in prokaryotes (see Figs 12–12, 28–18), also affects the transcription of a set of genes, each of

which is located near a specific DNA sequence called a cAMP response element (CRE). The catalytic subunit of protein kinase A, released when cAMP levels rise (see Fig. 12–15), enters the nucleus and phosphorylates a nuclear protein, the CRE-binding protein (CREB). When phosphorylated, CREB binds to CREs near certain genes and acts as a transcription factor, turning on the expression of these genes.

RNAi.

Posttranscriptional Gene Silencing Is Mediated by RNA Interference

In higher eukaryotes, including nematodes, fruit flies, plants, and mammals, a class of small RNAs has been discovered that mediates the silencing of particular genes. The RNAs function by interacting with mRNAs, often in the 3UTR, resulting in either mRNA degradation or translation inhibition. In either case, the mRNA, and thus the gene that produces it, is silenced. This form of gene regulation controls developmental timing in at least some organisms. It is also used as a mechanism to protect against invading RNA viruses (particularly important in plants, which lack an immune system) and to control the activity of transposons. In addition, small RNA molecules may play a critical (but still undefined) role in the formation of heterochromatin. The small RNAs are sometimes called micro-RNAs (miRNAs). Many are present only transiently during development, and these are sometimes referred to as small temporal RNAs (stRNAs). Hundreds of different miRNAs have been identified in higher eukaryotes. They are transcribed as precursor RNAs about 70 nucleotides long, with internally complementary sequences that form hairpinlike structures. The precursors are cleaved by endonucleases to form short duplexes about 20 to 25 nucleotides long. The best-characterized nuclease goes by the delightfully suggestive name Dicer; endonucleases in the Dicer family are widely distributed in higher eukaryotes. One strand of the processed miRNA is transferred to the target mRNA (or to a viral or transposon RNA), leading to inhibition of translation or degradation of the RNA. This gene regulation mechanism has an interesting and very useful practical side. If an investigator intro-duces into an organism a duplex RNA molecule corre-sponding in sequence to virtually any mRNA, the Dicer endonuclease cleaves the duplex into short segments, called small interfering RNAs (siRNAs). These bind to the mRNA and silence it. The process is known as RNA interference(RNAi).In plants, virtu-ally any gene can be effectively shut down in this way.

In nematodes, simply introducing the duplex RNA into the worm's diet produces very effective suppression of the target gene. The technique has rapidly become an important tool in the ongoing efforts to study gene func-tion, because it can disrupt gene function without cre-ating a mutant organism. The procedure can be applied to humans as well. Laboratory-produced siRNAs have already been used to block HIV and poliovirus infections in cultured human cells for a week or so at a time. Al-though this work is in its infancy, the rapid progress makes RNA interference a field to watch for future med-ical advances.

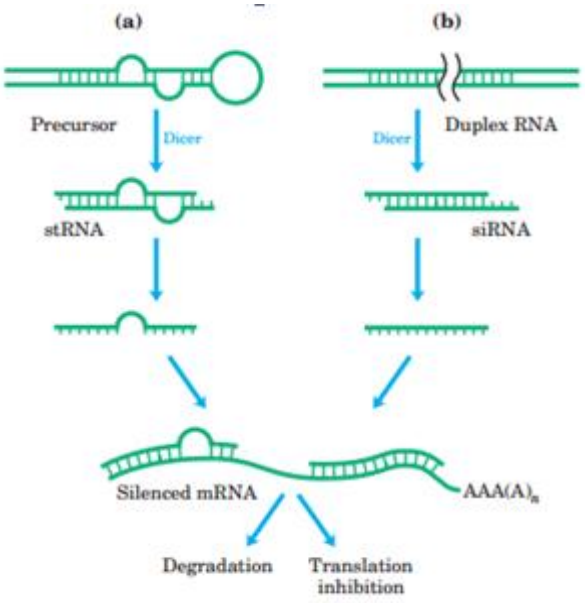


Figure Gene Silencing by RNA interference. (a) Small temporal RNAs (SiRNAs) are generated by Dicer-mediated cleavage of longer precursors that fold to create duplex regions. The SiRNAs then bind to mRNAs leading to degradation of mRNA or inhibition of translation. (b) Double-stranded RNAs can be constructed and introduced into a cell. Dicer processes the duplex RNAs into small interfering RNAs, which interact with the target mRNA. Again, the mRNA is either degraded or its translation inhibited.

II BSc Biochemistry -Fourth Semester
GENE EXPRESSION AND REGULATION (17BCU402)
Unit V

S.No.	Questions	Opt 1	Opt 2	Opt 3	Opt 4	Answer
1	Dead organism of pneumococci was completely inactivated by _____ enzymes	DNAase	Amylase	Lipase	Protease	DNAase
2	Nonvirulent in pneumococci is imparted by	Polysaccharide	Protein	Lipid	Lipopolysaccharide	Lipopolysaccharide
3	In DNA, the genetic information residues in	Purine bases	Pyrimidine bases	Purine and Pyrimidine bases	Sugar	Purine and Pyrimidine bases
4	Important characteristic of satellite DNA is, they contain	Repetitive base sequence	GC rich	Unique sequence	AT rich	GC rich
5	LINES and SINES come under _____ DNA	Single copy DNA	Moderatively repetitive DNA	Simple sequence DNA	RNA	Moderatively repetitive DNA
6	What is the difference between II S and II R strains of bacteria used in Griffith's experiment?	III S strain covers itself with a polysaccharide coat	III S strain covers itself with a protein coat	II R strain covers itself with a protein coat	II R strain covers itself with a polysaccharide coat	III S strain covers itself with a polysaccharide coat
7	Avery, Macleod and Mc Carty's experiment was conducted in the year	1928	1974	1964	1938	1974
8	To explore the genetic properties of DNA Hershey and Chase used	bacteria	bacterial viruses	fungi	dog	bacterial viruses
9	Conjugation was discovered by	Zinder and Ledenberg	Lederberg and Tatum	Frederick Griffith	Hershey and Chase	Lederberg and Tatum
10	In conjugation the integration of F factor is mediated by	transposon	IS elements	viruses	simple sequence DNA	IS elements
11	The DNA in the chromatin is very tightly associated with proteins called	albumin	globulin	myosin	histones	histones
12	The molecular weight of histones is	between 11,000 and 21,000	between 1,000 and 11,000	between 8,000 and 20,000	between 1,000 and 31,000	between 11,000 and 21,000
13	Histones are very rich in	acidic amino acids	basic amino acids	neutral amino acids	aromatic amino acids	basic amino acids
14	Number of amino acids in H4 histones is	104	108	102	120	102
15	The bead of each nucleosome contains	eight histone molecules	ten histone molecules	two histone molecules	eight histone molecules	eight histone molecules
16	The number of base pairs serves as linker DNA between nucleosome beads is	200 bp	146bp	54 bp	45 bp	54 bp
17	The spacing of the nucleosome beads provides a repeating unit typically of about	200 bp	146bp	54bp	45bp	200 bp
18	The total number of base pairs bound tightly around the eight-part histone core are	200 bp	146 bp	54 bp	45 bp	146 bp
19	X-ray diffraction analysis of nucleosome reveals that it is a	left-handed solenoidal supercoil	right-handed solenoidal supercoil	left handed double helical supercoil	right handed double helical supercoil	left-handed solenoidal supercoil
20	Specific binding sites of histone core to DNA is	A=T rich base pairs	G≡C rich base pairs	equal ratio of both base pairs	1:2 ratio of both the base pairs	A=T rich base pairs
21	The SMC protein play a substantial role in linking together sister chromatids is	cohesins	condensins	albumins	globulins	cohesins
22	A well studied example of a solitary protein coding gene is	lactate dehydrogenase gene	protein kinase gene	chicken lysozyme gene	Immunoglobulin gene	protein kinase gene
23	The bacterial strain used in Griffith's experiment is	*Staphylococcus aureus*	*Streptococcus pneumoniae*	*E Coli*	*Streptococcus pyogenes*	*Streptococcus pneumoniae*
24	The activity of the enzyme that helps to keep the surface of the eye and the chicken egg sterile is	lactate dehydrogenase	protein kinase	lysozyme	Immunoglobulin	protein kinase