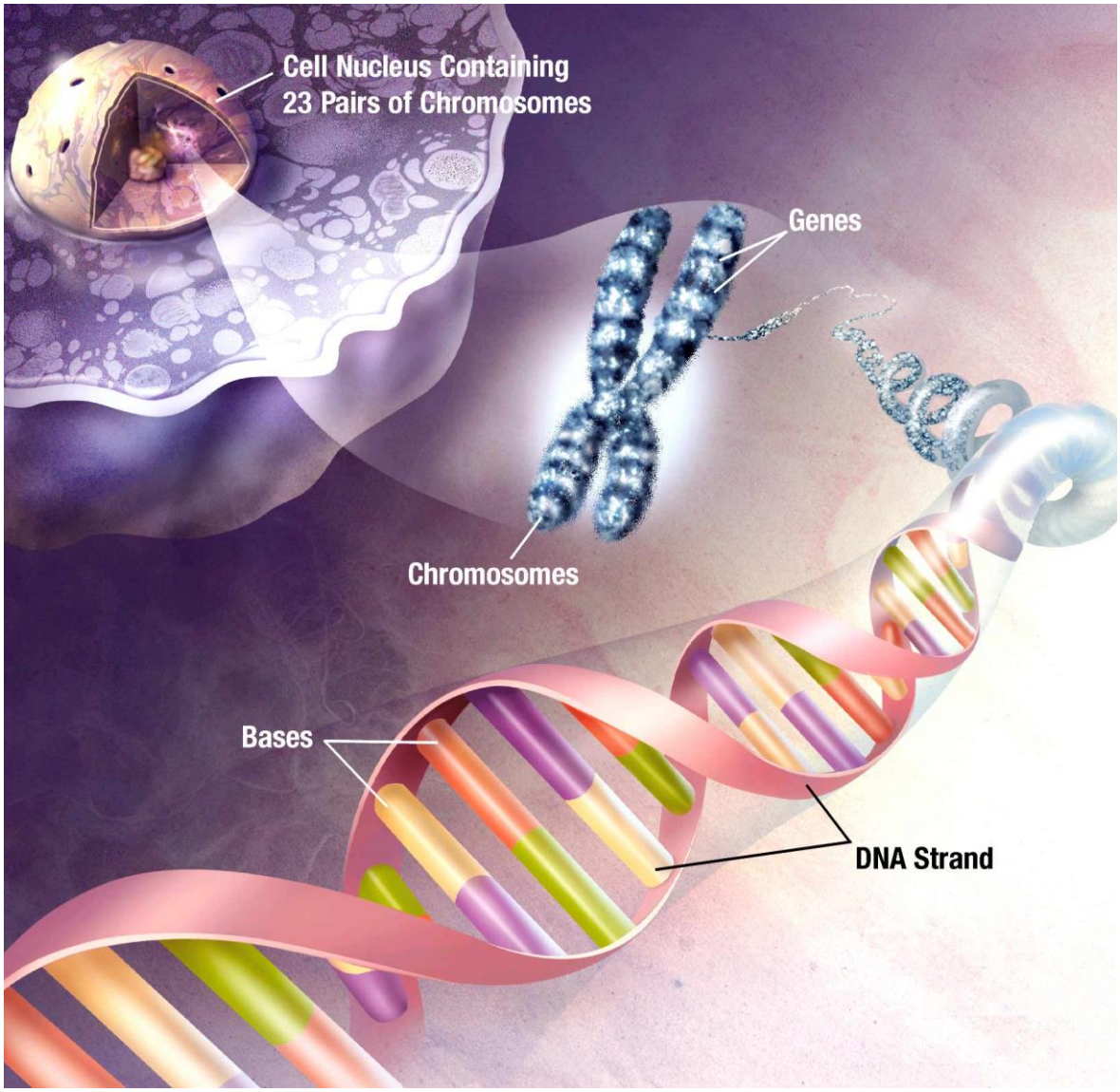
《Molecular Biology》双语教学简明教程

朱国萍　　王鹏　　主编



二零二二年三月

课 程 简 介

课程性质：专业基础课/必修课程

所属一级学科：生物学

所属二级学科：生物化学与分子生物学

授课对象：生命科学学院本科生、生命科学及其相关领域研究生

开课系及教科组：生命科学学院《分子生物学》课程组

学分数：3 学分

总学时数：每周 3 学时，共 51 学时

预备知识：遗传学、生物化学、细胞生物学基础知识。

课程描述

本门课程作为生命科学领域的一门新兴学科，是以学科的相互交叉和相互渗透为基础发展起来的，其发展反过来又成为其它学科从分子水平揭示生命现象实质的强有力工具。

本课程的特点是涉及面广，知识更新快，因此本课程着重讲解分子生物学基本概念和基本理论，同时力求反映分子生物学最新进展，使学生掌握坚实宽广的基础理论知识，了解发展前沿，培养创新精神和实践能力，以适应科学技术突飞猛进的时代要求。

本课程旨在引领学生理解生命活动尤其遗传与发育的分子本质，并培养学生运用分子生物学的方法观察、辩证分析和综合应用能力，解释生命的一般遗传与发育的规律。

Chapters

Supplementary Parts

第一章　绪论

Chapter 1　Introduction

【教学目的】

通过本章教学，使学生明确分子生物学的学科性质、基本内容和学习意义，了解本门课程的教学要求和学习方法。

【重点难点】

明确分子生物学的概念及研究内容。

【教学方法】

多媒体教学

English Animation

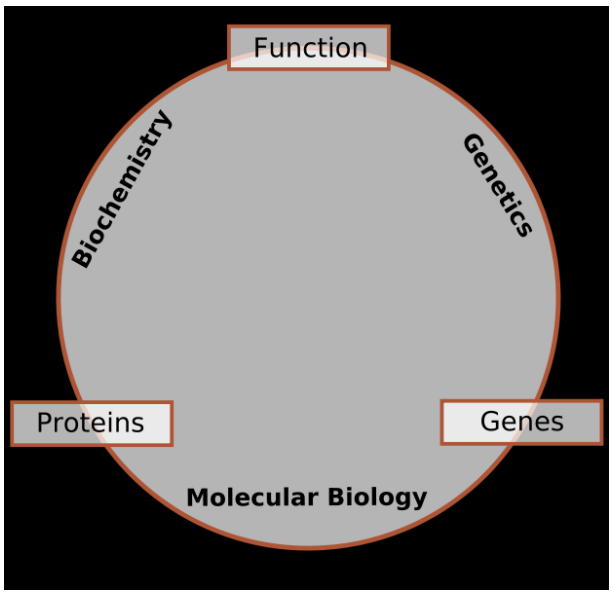
【课时安排】

2 课时

主要教学内容

1. Introduction

Molecular biology is the study of biology at a molecular level. The field overlaps with other areas of biology and chemistry, particularly genetics and biochemistry. Molecular biology chiefly concerns itself with understanding the interactions between the various systems of a cell, including the interrelationship of DNA, RNA and protein synthesis and learning how these interactions are regulated. The following figure is a schematic that depicts one possible view of the relationship between the fields:



Schematic relationship between biochemistry, genetics and molecular biology

2. History of Molecular Biology

(1) The Early Years of Genetics

1) 1859, Charles Darwin and Evolutionary Theory
2) 1865, Gregor Mendel and Genetics
3) 1869, Discovery of "Nuclein"
4) 1909, Gene, Genotype, Phenotype
5) 1910, Golden Age of Genetics - Thomas Hunt Morgan
6) 1924, it was proposed by Hsien Wu that the protein denaturation is related to its structure. The native protein is folded into a proper structure. But proteins can be denatured through exposure to heat or chemicals. Denatured proteins lose its ordered structure and become loose and disordered.
7) 1941, George Beadle and Edward Tatum: One gene controls one enzyme or one protein. Their work from the 1930s to 1960 marked the transition from classical genetics to the molecular era. 1958, Nobel Prize in Physiology or Medicine.
8) 1943, DNA - Transforming Principle
9) 1952, DNA -- Inherited Substance
10) 1953, Unraveling the DNA Double Helix
11) 1967, Cracking the Genetic Code

(2) The Birth of Biotechnology

1) 1968, Plasmid
2) 1970, Restriction Enzymes Discovered
3) 1972, Recombinant DNA Technology
4) 1975, DNA Sequencing
5) 1975, Monoclonal Antibody Technology
6) 1978, Human Insulin Cloned
7) 1983, American biologist Kevin M. Ulmer published a paper on Science, and the Protein Engineering was first proposed.

(3) The Revolution of the Biotechnology

1) 1986, The Polymerase Chain Reaction (PCR)
2) 1987, Victor McKusick and Frank Ruddle added “genomics” to the scientific lexicon as the title for the new journal they cofounded in 1987, with emphasis on linear gene mapping, DNA sequencing, and comparison of genomes from different species.
3) 1989, The Human Genome Project (HGP)
4) 1990, First Case of Gene Therapy
5) 1994, Brave New Foods
6) 1996, Announcing Dolly**,** First Mammal Cloned from Adult Cells
7) 1996, Development of the Gene Chip
8) 1997, Three Cloned Mice
9) 1997, First Human Artificial Chromosome

(4) Welcome to the Genomics and Post-Genomics Era

1) 1998, Race for the Genome
2) 1998, Breakthrough of Stem Cell
3) 1999, Shotgun sequencing
4) 2001, Draft of Human Genome
5) 2001, Cloning Ban
6) 2001-2018, Some cloned animals: cat, pig, horse, dog, monkey……
7) 2006, The Nobel Prize in Physiology or Medicine was awarded jointly to Andrew Fire and Craig C. Mello for their work on *RNA interference in C. elegans*, which they published in 1998.
8) 2007, The Nobel Prize in Physiology or Medicine was awarded jointly to Mario R. Capecchi, Sir Martin J. Evans and Oliver Smithies *"for their discoveries of principles for introducing specific gene modifications in mice by the use of embryonic stem cells"*.
9) 2008, The Nobel Prize in Physiology or Medicine was awarded jointly to Harald zur Hausen for his discovery of "*human papilloma viruses causing cervical cancer*" and the other half jointly to Françoise Barré-Sinoussi and Luc Montagnier for their discovery of "*human immunodeficiency virus*"

10) 2008, The Nobel Prize in Chemistry was awarded jointly to Osamu Shimomura, Martin Chalfie and Roger Y. Tsien *"for the discovery and development of the green fluorescent protein, GFP"*.

11) 2009, The Nobel Prize in Physiology or Medicine was awarded jointly to Elizabeth H. Blackburn, Carol W. Greider and Jack W. Szostak *"for the discovery of how chromosomes are protected by telomeres and the enzyme telomerase"*.

12) 2009, The Nobel Prize in Chemistry was awarded jointly to Venkatraman Ramakrishnan, Thomas A. Steitz and Ada E. Yonath *"for studies of the structure and function of the ribosome"*.

13) 2010, The Nobel Prize in Physiology or Medicine was awarded to Robert G. Edwards *"for the development of in vitro fertilization"*.

14) 2010, The Nobel Prize in Chemistry was awarded jointly to Richard F. Heck, Ei-ichi Negishi and Akira Suzuki *"for palladium-catalyzed cross couplings in organic synthesis"*.

15) 2011, The Nobel Prize in Physiology or Medicine 2011 was divided, one half jointly to Bruce A. Beutler and Jules A. Hoffmann *"for their discoveries concerning the activation of innate immunity"* and the other half to Ralph M. Steinman *"for his discovery of the dendritic cell and its role in adaptive immunity"*.

16) 2012, The Nobel Prize in Physiology or Medicine was awarded jointly to Sir John B. Gurdon and Shinya Yamanaka *"for the discovery that mature cells can be reprogrammed to become pluripotent"*

17) 2012, The Nobel Prize in Chemistry was awarded jointly to Robert J. Lefkowitz and Brian K. Kobilka *"for studies of G-protein-coupled receptors"*

18) 2013, The Nobel Prize in Physiology or Medicine was awarded jointly to James E. Rothman, Randy W. Schekman and Thomas C. Südhof *"for their discoveries of machinery regulating vesicle traffic, a major transport system in our cells"*.

19) 2014, The Nobel Prize in Physiology or Medicine was awarded jointly to John O'Keefe, May Britt Moser and Edvard Moser "*for their discoveries of cells that constitute a positioning system in the brain*".

20) 2015, The Nobel Prize in Physiology or Medicine was awarded jointly to William C. Campbell and Satoshi Ōmura "*for their discoveries concerning a novel therapy against infections caused by roundworm parasites*" and the other half to Youyou Tu "*for her discoveries concerning a novel therapy against Malaria*".

21) 2015, The Nobel Prize in Chemistry was awarded jointly to Tomas Lindahl, Paul Modrich and Aziz Sancar "*for mechanistic studies of DNA repair*".

22) 2016, The Nobel Prize in Physiology or Medicine was awarded to Yoshinori Ohsumi "for his discoveries of mechanism*s for autophagy".*

23) 2017, The Nobel Prize in Physiology or Medicine was awarded jointly to Jeffrey C. Hall, Michael Rosbash and Michael W. Young *"for their discoveries of molecular mechanisms controlling the circadian rhythm".*

24) 2017, The Nobel Prize in Chemistry was awarded to Jacques Dubochet, Joachim Frank and Richard Henderson *"for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution".*

25) 2018, The Nobel Prize in Physiology or Medicine was awarded jointly to James P. Allison and Tasuku Honjo *"for their discovery of cancer therapy by inhibition of negative immune regulation".*

26) 2019, The Nobel Prize in Physiology or Medicine was awarded jointly to William G. Kaelin Jr, Sir Peter J. Ratcliffe and Gregg L. Semenza *"for their discoveries of how cells sense and adapt to oxygen availability".*

27) 2020, The Nobel Prize in Chemistry was awarded jointly to Emmanuelle Charpentier and Jennifer A. Doudna *"for the development of a method for genome editing".*

28) 2020, The Nobel Prize in Physiology or Medicine was awarded jointly to Harvey J. Alter, Michael Houghton and Charles M. Rice *"for the discovery of Hepatitis C virus".*

29) 2021, The Nobel Prize in Physiology or Medicine was awarded jointly to David Julius and Ardem Patapoutian *"for their discoveries of receptors for temperature and touch".*

3. What is Molecular Biology?

Molecular biology is the study of biology at a molecular level. It chiefly concerns itself with understanding the interactions between the various systems of a cell, including the interrelationship of DNA, RNA and protein synthesis and learning how these interactions are regulated.

4. Contents of Molecular Biology

(1) Central dogma

(2) Contents of molecular biology

- ❖ Recombinant DNA Technology
- ❖ Expression and Regulation of Genes
- ❖ Structure and Function of Bio-macromolecular
- ❖ Genome, Transcriptome, Proteome, Bioinformatics

5. Progress and Future of Molecular Biology

第二章　基因和染色体

Chapter 2　DNA and Chromosomes

【教学目的】

本章要求学生掌握原核生物和真核生物基因组的区别、染色体的组成、真核生物基因组的复杂性、DNA 的结构等；掌握 DNA 复制的复杂性、几种 DNA 的复制方式、原核生物 DNA 复制的酶、真核生物 DNA 复制的酶、DNA 的修复、重组和转座等。

【重点难点】

掌握基本概念、原核生物和真核生物基因组特点、DNA 复制、修复、重组及转座机制。

【教学方法】

讲述式和启发式教学

多媒体教学

English Animation

【课时安排】

10 课时

主要教学内容

Overview

Basic material: DNA, RNA, Protein (composition, classes, structures, functions)

Transcription: DNA → RNA

- Transcriptional factors
- RNA processing: RNA splicing, RNA editing

Translation: RNA → Protein

Posttranslational modification: Intein splicing, Phosphorylation, Modification……

Regulation: Cis-acting elements, Trans-acting factors

Central dogma: Old version, New content, Progress……

2.1　DNA and DNA structure

1.　Nucleoside & Nucleotide

2. Nucleic Acid Chain

3. DNA Structure

(1) DNA Primary Structure: nucleotide acid sequence

(2) DNA Secondary Structure: double helix

- Right-handed helix -- A-DNA, B-DNA
- Left-handed -- Z-DNA

(3) DNA Supercoiling:

(4) DNA Denaturation and Renaturation

2.2 Gene and Chromosomes

1. Concepts of Gene

(1) Gene **-** In molecular terms, it is the entire DNA sequence including **exons, introns,** and **noncoding transcription-control regions** necessary for production of a functional protein or RNA.

In molecular terms, a gene commonly is defined as the entire nucleic acid sequence that is necessary for the synthesis of a functional polypeptide or RNA molecule.

(2) Regulatory Gene - A DNA sequence or a gene that functions to control the expression of other genes.

(3) Structural Gene - A DNA sequence or a gene that codes the production of RNA, a specific protein or peptide.

(4) Gene Cluster - It is a group of adjacent genes that are identical or related.

(5) Gene Family - A group of related genes having similar DNA sequence evolved from a single ancester. These genes make similar products and may or may not be located in the same region of a chromosome.

2. Genome

(1) Genome - In modern molecular biology the genome of an organism is its total hereditary information encoded in DNA (or, for some viruses, RNA).

(2) Genomics – It is a discipline in genetics that applies recombinant DNA, DNA sequencing methods, and bioinformatics to sequence, assemble, and analyze the function and structure of genomes.

(3) Structural Genomics - The branch of genomics that attempts to make use of the vast wealth of data produced by genomic projects (such as genome sequencing projects) to describe gene (and protein) functions and interactions.

As its name suggests, the aim of structural genomics is to characterize the structure of the genome. Knowledge of the structure of an individual genome can be useful in manipulating genes and DNA segments in that particular species. Structural genomics proceeds through increasing levels of analytic resolution, starting with the assignment of genes and markers to individual chromosomes, then the mapping of these genes and markers within a chromosome, and finally the preparation of a physical map culminating in sequencing.

When a number of genomes have been characterized at the structural level, the hope is that, through comparative genomics, it will become possible to deduce the general rules that govern the overall structural organization of all genomes.

(4) Functional Genomics – It is a field of molecular biology that attempts to make use of the vast wealth of data produced by genomic projects (such as genome sequencing projects) to describe gene (and protein) functions and interactions.

Unlike genomics, functional genomics focuses on the dynamic aspects such as gene transcription, translation, and protein–protein interactions, as opposed to the static aspects of the genomic information such as DNA sequence or structures. Functional genomics attempts to answer questions about the function of DNA at the levels of genes, RNA transcripts, and protein products. A key characteristic of functional genomics studies is their genome-wide approach to these questions, generally involving high-throughput methods rather than a more traditional “gene-by-gene” approach.

3. Size of Genome

(1) Prokaryotic cell vs Eukaryotic cell

(2) Relationship of genomic size and evolution

4. Model Organisms

(1) Bacteria (*E. coli*, several others)

(2) Yeast (*Saccharomyces cerevisiae*)

(3) Plant (Arabidopsis thaliana)

(4) Caenorhabditis elegans

(5) Fruit fly

(6) Zebrafish

(7) Mouse

(8) Human

2.3 Features of Genomic Organization

1. Features of prokaryotic genomes

(1) Relatively small

(2) Simple structure

(3) Transcription unit - A segment of DNA that contains signals for the initiation and termination of transcription, and is transcribed into one primary RNA molecule.

(4) Overlapping gene - Different genes whose nucleotide coding sequences overlap to some extent. The common nucleotide sequence is read in two or three different reading frames thus specifying different polypeptides.

2. Genomes of Prokaryotes

(1) *E.coli* Genome

(2) Phage genome

(3) Mu phage

3. Features of Eukaryotic Genomes

(1) Big size, big content

C value - A measure of the amount of DNA in the haploid genome of an organism, which can be by mass or by molecular weight.

C value paradox - The observation that the amount of deoxyribonucleic acid in the haploid genome is not related to its evolutionary complexity.

(2) Repetitive sequence

- Single copy
- Moderately repetitive sequence
- Highly repetitive sequence

Satellite DNA - A portion of DNA in eukaryotes whose density differs from that of the majority of DNA and that consists of short, repeating sequences of nucleotide pairs, often found near the region of the centromere.

Minisatellites - The size of a minisatellite ranges from 1 kb to 20 kb. One type of minisatellites is called variable number of tandem repeats (VNTR). Its repeat unit ranges from 10-100 bp. They are located in non-coding regions. The number of repeats for a given minisatellite may differ between individuals. This feature is the basis of **DNA fingerprinting**.

Another type of minisatellites is the telomere. In a human germ cell, the size of a telomere is about 15 kb. In an aging somatic cell, the telomere is shorter. The telomere contains tandemly repeated sequence GGGTTA.

Microsatellites - Microsatellites are also known as **short tandem repeats** (STR), because a repeat unit consists of only 1-6 bp and the whole repetitive region spans <150 bp. Similar to minisatellites, the number of repeats for a given microsatellite may differ between individuals. Therefore, microsatellites can also be used for DNA fingerprinting.

(3) >90% of genome is Non-coding sequence, lots of cis-acting elements

(4) Single cistron

(5) Interrupted gene - exon and intron

GT-AG rule (GU-AG rule)- Almost all introns have constant dinucleotide GT (GU) at the first two position and AG at the last two position.

(6) Pseudogene - An inactive gene derived from an ancestral active gene.

A “duplicated” pseudogene arises when a cell is replicating its own DNA and inserts an extra copy of a gene into the genome in a new location.

A “processed” pseudogene is formed during gene expression, when a gene is transcribed into RNA, then that transcript is processed into a mRNA. Normally, the

mRNA is destined for translation into a protein—but sometimes it can instead be reverse-transcribed back into DNA form and inserted in the genome.

(7) Organelle genomes (mitochondrial DNA, mtDNA; chloroplast DNA, ctDNA)

(8) DNA polymorphism

(9) Telomere

4. Genome of Eukaryotes

(1) Human genome

(2) Other genomes of Eukaryotes

- Mouse genome
- Yeast
- *Drosophila*
- *C. elegans*
- *Arabidopsis thaliana*

2.4 Nucleosome & Chromosome

1. Histones and Non-histone

2. Nucleosome

(1) Concept -

(2) Packing

Nucleosomes form the primary level of nuclear DNA compaction in almost all eukaryotes. In all cases, they contain essentially the same four histones, H2A, H2B, H3 and H4, always arranged as (H2A+H2B)-(H3+H4)2-(H2A+H2B) heterodimers in an octamer core, with ≈200 bp of DNA.

3. DNA Coiling into a Chromosome

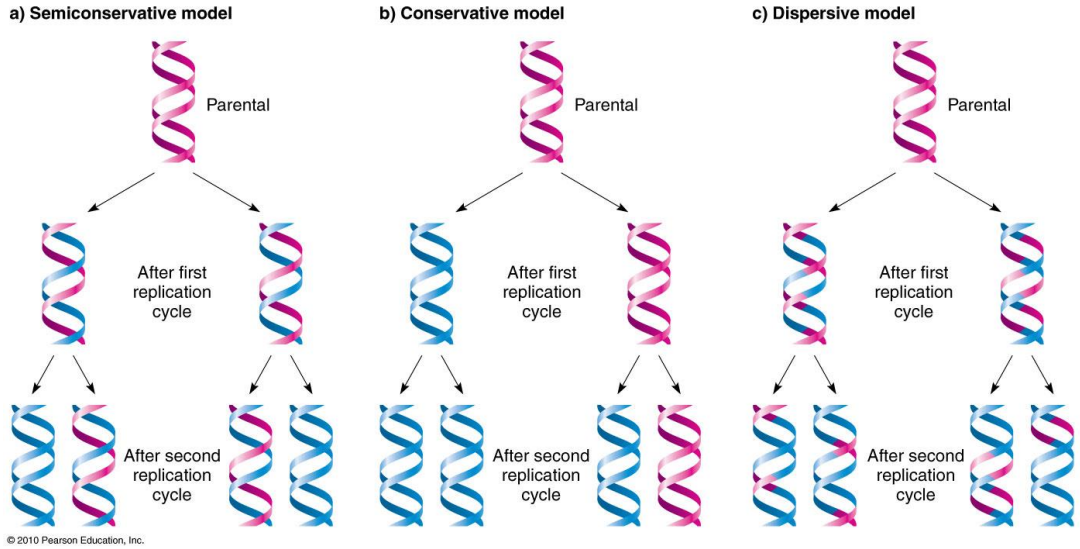
DNA (2 nm) + histones → nucleosome (11 nm) → solenoid (30 nm)
→ chromatin fiber (300 nm) → chromatid (700 nm) → chromosome (1400 nm)

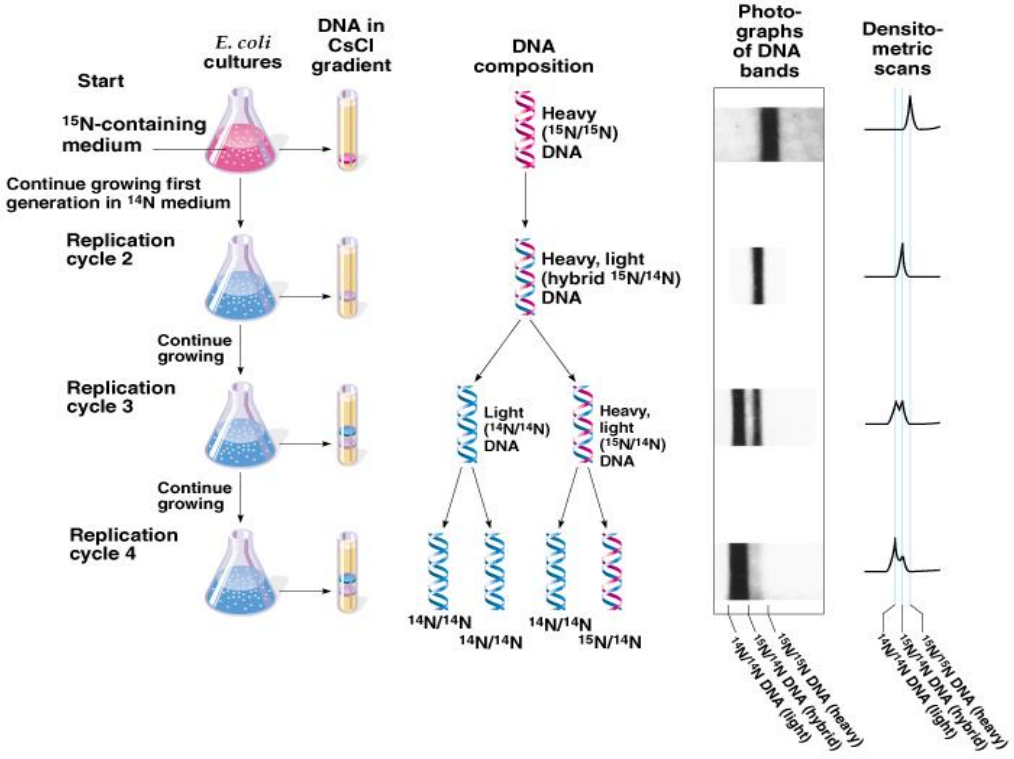
2.5 DNA Replication

1. Semiconservative replication

(1) Models of DNA replication

- Semiconservaticve model - Semiconservative replication (Each strand acts as a template for a new double helix, in which each double-stranded molecule is composed of one parental strand and one newly polymerized strand.
- Conservative model
- Dispersive model





(2) Replication elements

- Replicon - A replicon is a DNA molecule or RNA molecule, or a region of DNA or RNA, that replicates from a single origin of replication.
- Replication fork - The replication fork is a structure that forms within the nucleus during DNA replication. It is created by helicases, which break the hydrogen bonds holding the two DNA strands together.
- Origin (ori) - The origin is a sequence of DNA at which replication is initiated. Prokaryotes have a single circular molecule of DNA, and only a single origin of replication per circular chromosome. Eukaryotes often have multiple origins of replication on each linear chromosome.
- Terminus: A terminus is a segment of DNA at which replication ends.
- Direction: unidirectional, bidirectional replication
- Replication speed

2. Enzymes involved in DNA replication

(1) Topoisomerase I - removes supercoils

(2) Helicase - untwists DNA

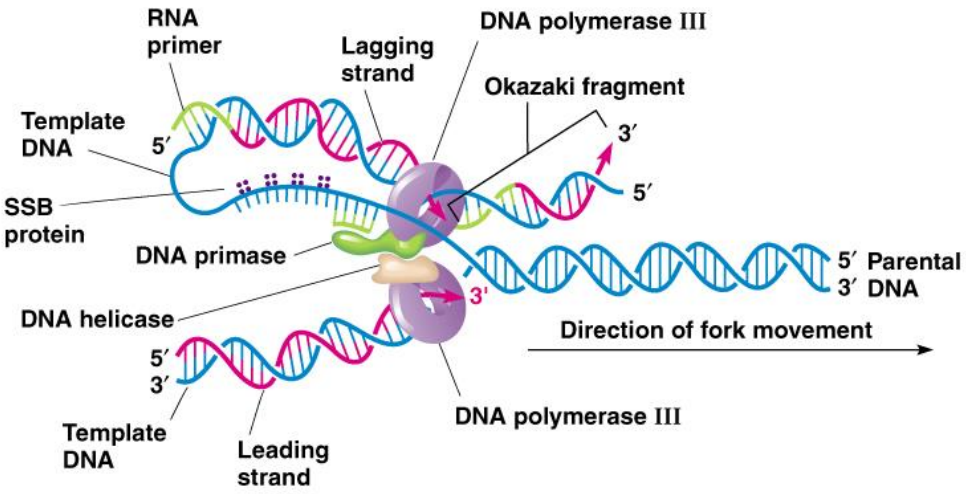
(3) Single strand binding proteins - stabilize replication fork

(4) Primase - makes RNA primer

(5) DNA Polymerase III - synthesizes new DNA strands

(6) DNA polymerase I - removes RNA primer, adds DNA bases

(7) DNA ligase - seals lagging strand 3' open holes



3. DNA Replication in Prokaryotes

(1) Ori in *E.coli*

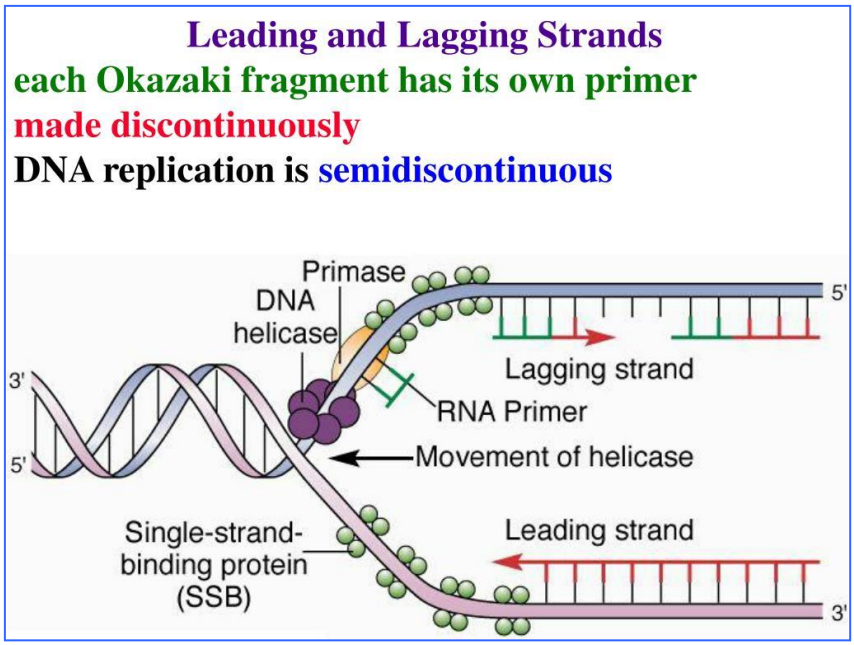
(2) Initiation of DNA replication in *E.coli*

(3) DNA polymerases

- Pol I: implicated in DNA repair
- Pol II: involved in replication of damaged DNA
- Pol III: elongates in DNA replication
- Pol IV: a Y-family DNA polymerase
- Pol V: a Y-family DNA polymerase; participates in bypassing DNA damage.

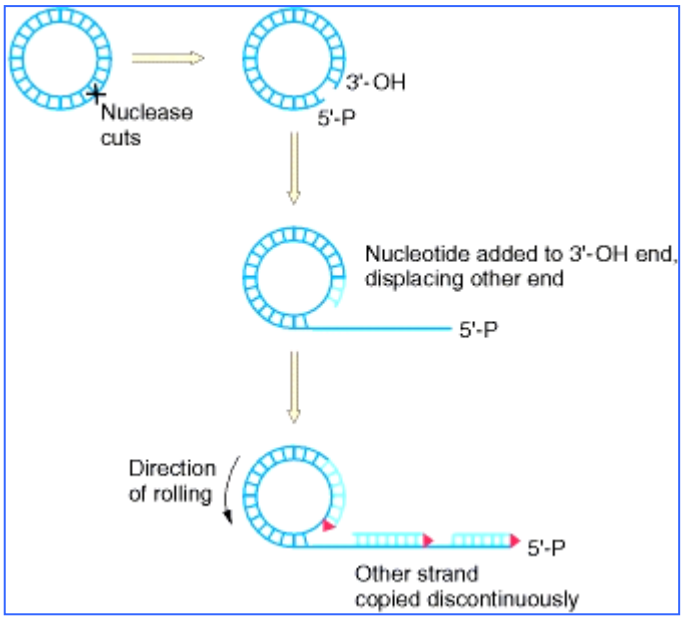
(4) Semidiscontinuous replication

This method of replication for the lagging strand is a little less foolproof, the lagging strand must only replicate in tiny little segments. The segments, referred to as Okazaki fragments, are synthesized in the 5' to 3' direction away from the replication fork. These fragments are then stitched together by DNA ligase, creating a continuous strand. This type of replication is called semidiscontinuous replication.



(5) Termination of Replication

(6) Rolling Circle Replication



4. DNA Replication in Eukaryotes

(1) Multiple replicons

(2) DNA polymerase

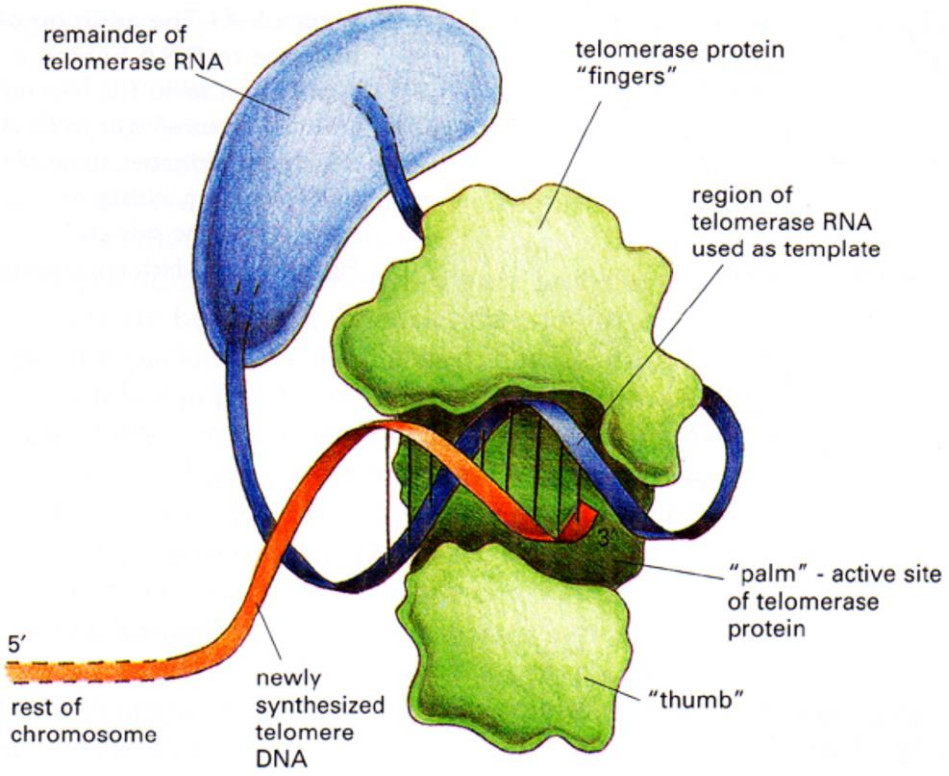
- Pol α: Acts as a primase (synthesizing a RNA primer), and then as a DNA Pol elongating that primer with DNA nucleotides. After around 20 nucleotides elongation is taken over by Pol δ (on the lagging strand) and ε (on the leading strand).
- Pol β: Implicated in repairing DNA, in base excision repair and gap-filling synthesis.
- Pol γ: Replicates mitochondrial DNA.
- Pol δ: Highly processive and has proofreading 3'->5' exonuclease activity. Thought to be the main polymerase involved in lagging strand synthesis, though there is still debate about its role.
- Pol ε: Also highly processive and has proofreading 3'->5' exonuclease activity. Highly related to Pol δ, and thought to be the main polymerase involved in leading strand synthesis……still debate about its role.

(3) Replication of Telomere

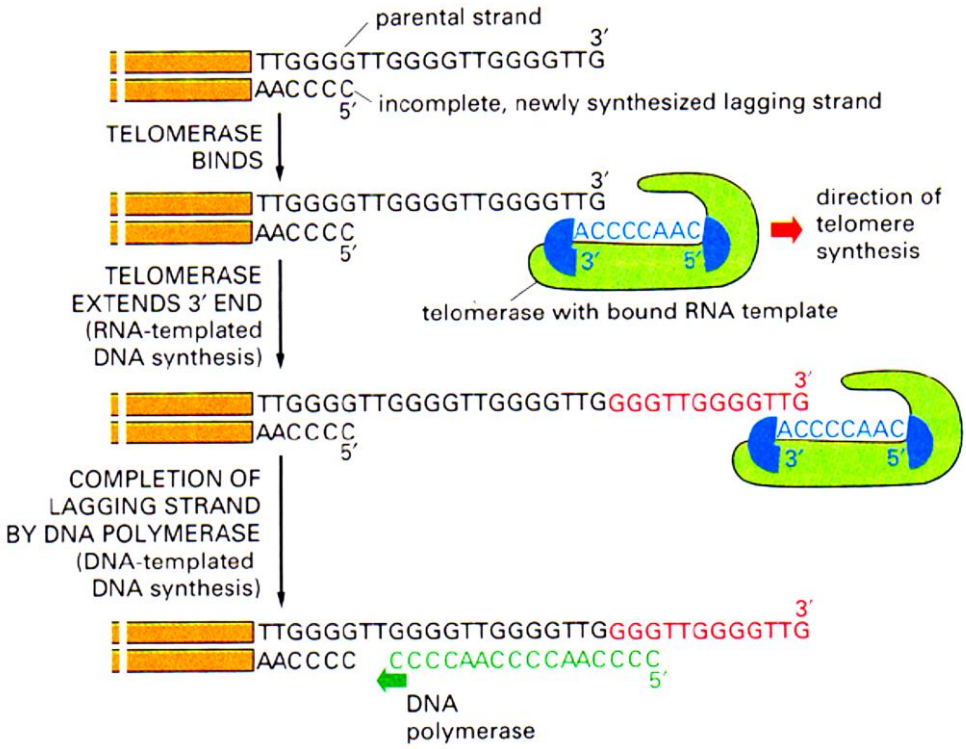
A telomere is a region of repetitive DNA at the end of chromosomes, which protects the end of the chromosome from destruction.

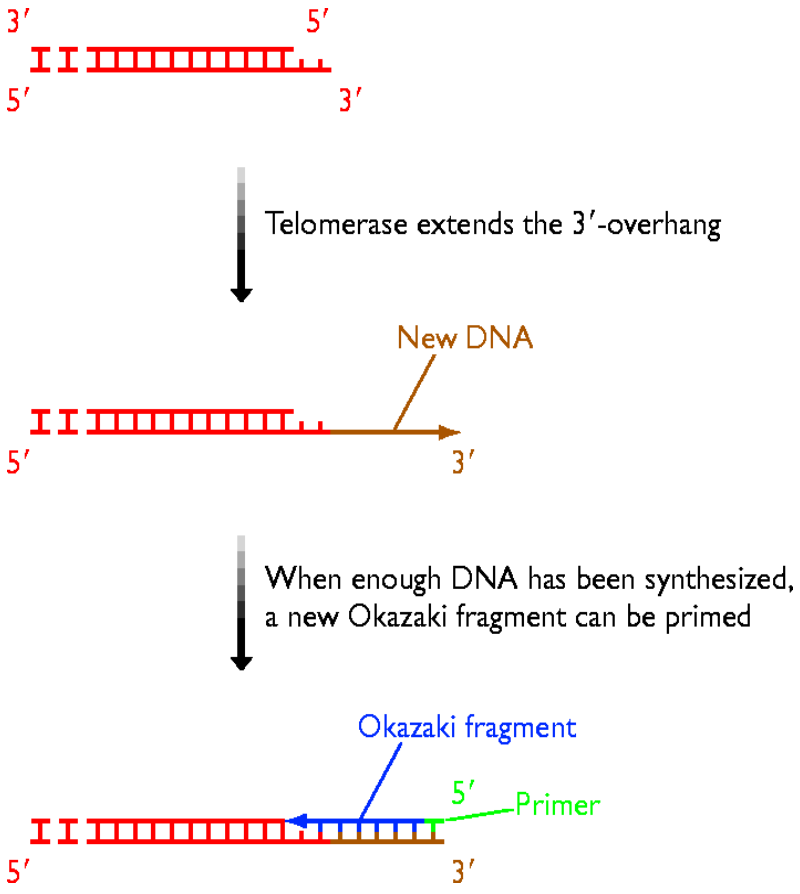
- Telomerase - Telomerase is a ribonucleoprotein that is an enzyme made of protein and RNA subunits. Its role is to elongate chromosomes by adding telomeric sequences to the end of existing chromosomes.

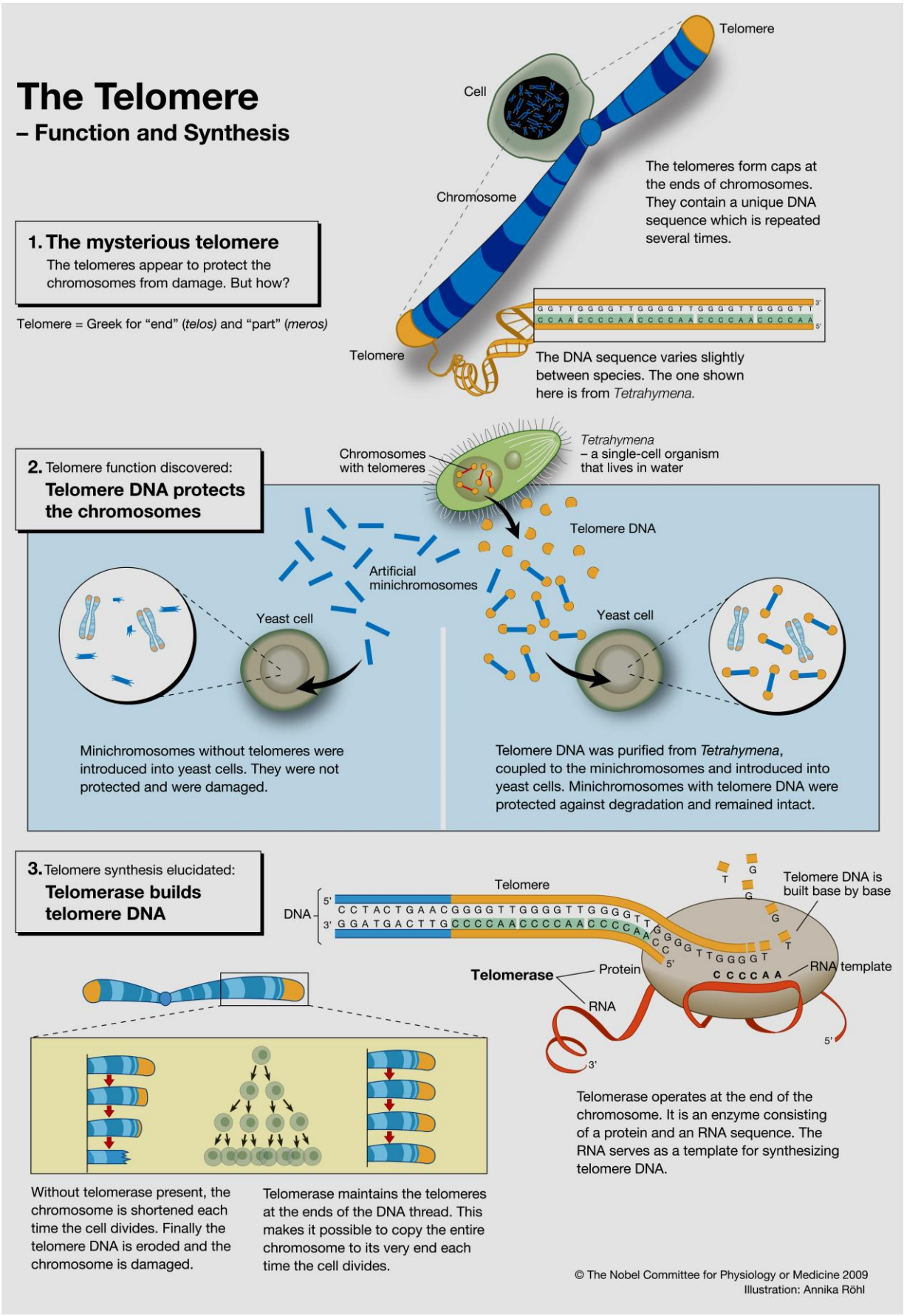
  As a result, every time the chromosome is copied, only 100–200 nucleotides are lost, which causes no damage to the organism's DNA. Telomerase is a reverse transcriptase that carries its own RNA molecule, which is used as a template when it elongates telomeres, which are shortened after each replication cycle.



- Replication of Telomere


parental strand
3′
TTGGGGTTGGGGTTGGGGTTG
AACCCC
5′
incomplete, newly synthesized lagging strand
TELOMERASE BINDS
ACCCCAAC
3′
5′
direction of telomere synthesis
telomerase with bound RNA template
TELOMERASE EXTENDS 3′ END (RNA-templated DNA synthesis)
TTGGGGTTGGGGTTGGGGTTGGGGTTGGGGTTG
COMPLETION OF LAGGING STRAND BY DNA POLYMERASE (DNA-templated DNA synthesis)
CCCCAACCCCAACCCC
DNA polymerase


3′
5′
Telomerase extends the 3′-overhang
New DNA
When enough DNA has been synthesized, a new Okazaki fragment can be primed
Okazaki fragment
Primer



(4) mtDNA Replication - D-loop replication

5. Proofreading of Replication

Error correction is a property of some, but not all, DNA polymerases. This process corrects mistakes in newly-synthesized DNA. When an incorrect base pair is recognized, DNA polymerase reverses its direction by one base pair of DNA. The 3'->5' exonuclease activity of the enzyme allows the incorrect base pair to be excised (this activity is known as *proofreading*). Following base excision, the polymerase can re-insert the correct base and replication can continue.

2.6 DNA Mutagenesis

1. Types of Mutagenesis

(1) Base Substitution

(2) Frameshift Mutation - A frameshift mutation (also called a framing error) is a genetic mutation caused by insertion or deletion of a number of nucleotides which can disrupt the reading frame, resulting in a completely different translation from the original.

(3) Exon skipping

Examples

2. Mutagens

3. Significance of mutation

2.7 DNA Repair

1. Excision Repair -- Dark Repair

(1) Mismatch Repair

(2) Base Excision Repair – AP Repair

(3) Nucleotide Excision Repair -- UV Damage Repair

2. Directed Repair – Photoreactivation (Light Repair)

3. Recombinational Repair

4. SOS Response - Error-prone repair

(1) Concept

(2) Involved Genes and Proteins

- Repressor: LexA
- Inducer: RecA
- Repair proteins UvrABC
- SOS genes
- SOS box

(3) Mechanism

- In nomal cell: SOS system→close
- Inducing signal: SOS system→triggered
- RecA activated: SOS system → open
- After repair: SOS system→close

2.8 DNA Recombination

1. Overview

(1) Important Contributions of Genetic Recombination

(2) Types of DNA recombination

- Homologous recombination
- Site-specific recombination
- Transposition and Retro-transposition

2. Homologous recombination

3. Site-specific Recombination

4. DNA Transposition and Retrotransposition

(1) Definition

- Transposition

- Retrotransposon

(2) Types and Structural Features

- Insertion Sequences (IS elements)
- Transposons
- Composite Transposons
- Retrotransposon

(3) Mechanism of Transposition and Retrotransposition

- Conservative Transposition
- Replicative Transposition
- Retrotransposition

(4) The difference between Transposition and Retrotransposon

(5) Use of Transposons

(6) Transposons in eukaryotes

2.9 SNPs

1. What are SNPs？

Single nucleotide polymorphisms (SNPs) are a type of polymorphism involving variation of a single base pair. Scientists are studying how single nucleotide polymorphisms, or SNPs (pronounced "snips"), in the human genome correlate with disease, drug response, and other phenotypes.

2. Importance

- Biomedical research
- Forensics
- Pharmacogenetics
- Disease

Summary

1. Concepts:

Gene	Exon	Substitution
Regulatory Gene	Intron	Mismatch Repair
Structural Gene	Pseudogene	Excision Repair
Gene Cluster	DNA Primary Structure	Directed Repair
Gene Family	DNA Secondary Structure	Recombination Repair
Genome	DNA Supercoiling	Error-prone Repair (SOS repair)
Genomics	DNA Denaturation	Homologous Recombination
Functional Genomics	DNA Renaturation	Site-specific Recombination
Structural Genomics	Semi-conservative Replication	Transposition
C value	Replication Fork	Transposon
C value paradox	Semidiscontinous Replication	Retrotransposon

2. Genomes of prokaryotes and eukaryotes
3. Genomic structure of prokaryotes and eukaryotes
4. DNA polymerase of prokaryotes and eukaryotes
5. Mechanism of DNA replication in prokaryotes and eukaryotes
6. Mechanism of DNA mutation, repair and recombination
7. Mechanism of DNA transposition and retrotransposition

Further reading

Questions and Problems

第三章　生物信息的传递(上)——From DNA to RNA

Chapter 3　Transcription — From DNA to RNA

教学目的：

本章要求学生掌握中心法则及其进展、原核生物的 RNA 聚合酶和启动子及其结构、真核生物的 RNA 聚合酶和启动子及其结构、真核生物的 RNA 聚合酶及其启动子、真核基因转录产物的修饰和剪切等。

重点与难点：

1、原核生物和真核生物的转录机器

2、启动子的结构与转录起始

3、RNA 转录的延伸与终止

4、RNA 的加工

教学方法：

讲述式和启发式教学 (适当结合讨论法)

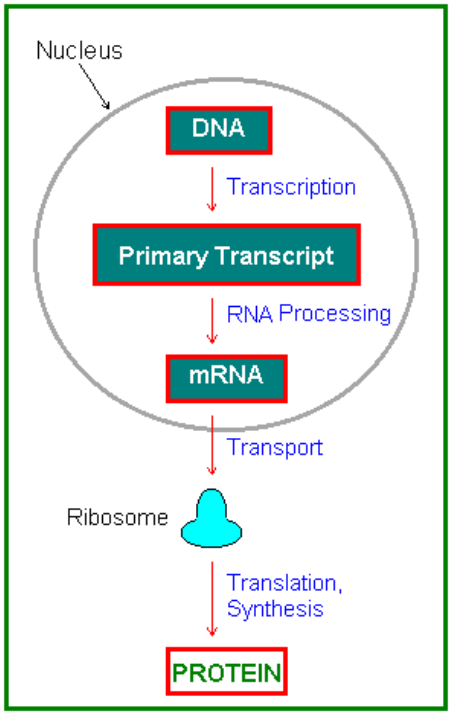
多媒体教学

English Animation

课时安排：

8 课时

主要教学内容：



3.1　Transcription Overview

1.　Gene Expression Processes

(1)　Gene Expression

(2)　Steps

- ◆ Transcription
- ◆ RNA processing
- ◆ Nuclear transport
- ◆ Translation

2.　Coding Strand & Antisense Strand

5’ ------ATTGCCTGAATGGTCCC----3’ (+)
coding strand, non-template strand, sense strand, plus strand……

3’ ------TAACGGACTTACCAGGG----5’(–)
template strand, antisense strand, minus strand……

5’ ------AUUGCCUGAAUGGUCCC----3’ sense RNA

3’ ------UAACGGACUUACCAGGG----5’ antisense RNA

3. Steps are required for the transcription

(1) Initiation:
- Template recognition
- RNA polymerase binding

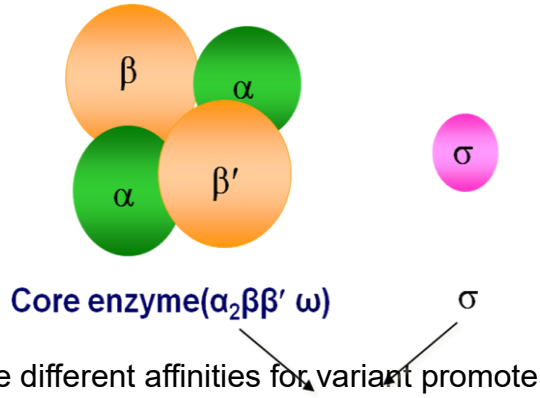
(2) Elongation

(3) Termination

3.2 Main Component of Transcription

1. RNA Pol in Prokaryotes - *E.coli*

- αI αII - required for DNA binding and assembly
- β: this has the polymerase activity (catalyzes the synthesis of RNA), which includes chain initiation and elongation.
- β': binds to DNA (nonspecifically).
- ω: stabilises β′ binding, restores denatured RNA polymerase to its functional form in vitro. It has been observed to offer a protective/chaperone function to the β' subunit in *Mycobacterium smegmatis*. It is now known to promote assembly.
- σ - forms holoenzyme, promoter recognition activity



Different σfactors have different affinities for variant promoters.

❖σ^{70} - general purpose.

❖σ^{54} - nitrogen metabolism.

❖σ^{32} - heat shock proteins.

❖σ^{S} - stationary phase.

❖σ^{F} - flagella.

2. RNA Polymerases in Eukaryotes

Polymerase	Location	Copies per cell	Products	Polymerase activity of cell
RNA polymerase I	Nucleolus	40000	35–47S pre-rRNA	50–70%
RNA polymerase II	Nucleoplasm	40000	mRNA precursors U1, U2, U4, and U5 snRNA	20–40%
RNA polymerase III	Nucleoplasm	20000	5S rRNA	10%
			tRNA	
			U6 snRNA	
			7S RNA	
			Other small	
			RNA molecules	
Mitochondrial RNA polymerase	Mitochondrion	?	Products of all mitochondrial genes	<1%
Chloroplast RNA polymerase	Chloroplast	?	Products of all chloroplast genes	<1%

3.3 Promoter and Transcription Initiation

1. DNA Elements for Transcription in Prokaryotes

(1) Promoter

(2) Structure of Promoter

-35 region

-10 region

INR (Inr)----startpoint (+1)

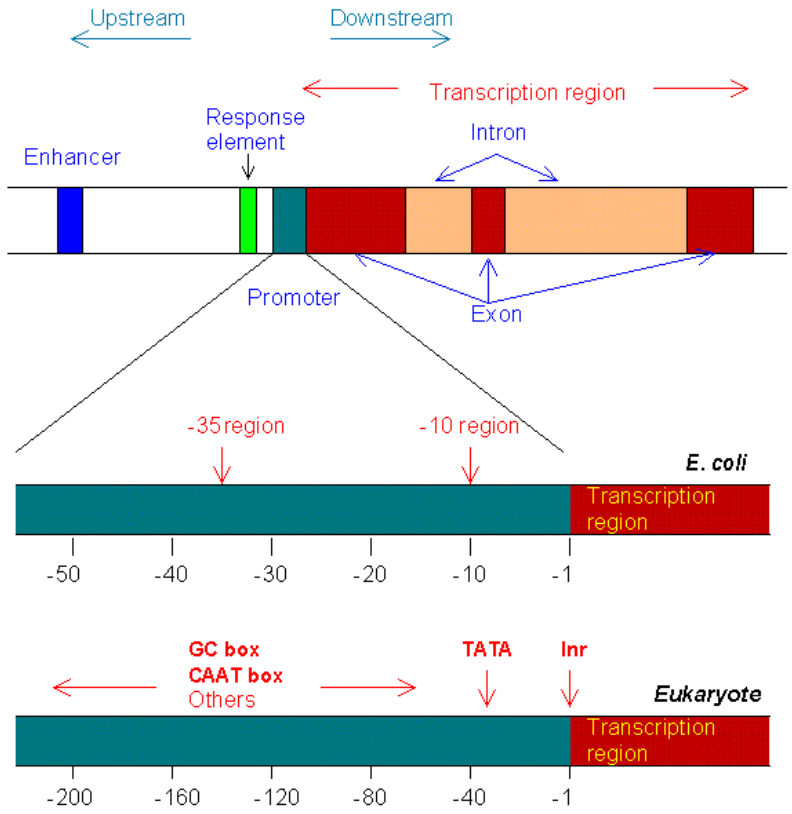
2. DNA Elements for Transcription in Eukaryotes

(1) Core Element of Promoter

TATA box, Inr

(2) UPE or UAS

CAAT box, GC box



3. Initiation of Transcription in Prokaryotes

- formation of closed binary complex
- formation of open binary complex
- DNA-RNA Polymerase-RNA tertiary complex
- σ released, new RNA chain is elongated

4. Initiation of Transcription in Eukaryotes

(1) Transcription factor (TF)

① Concept

② Types

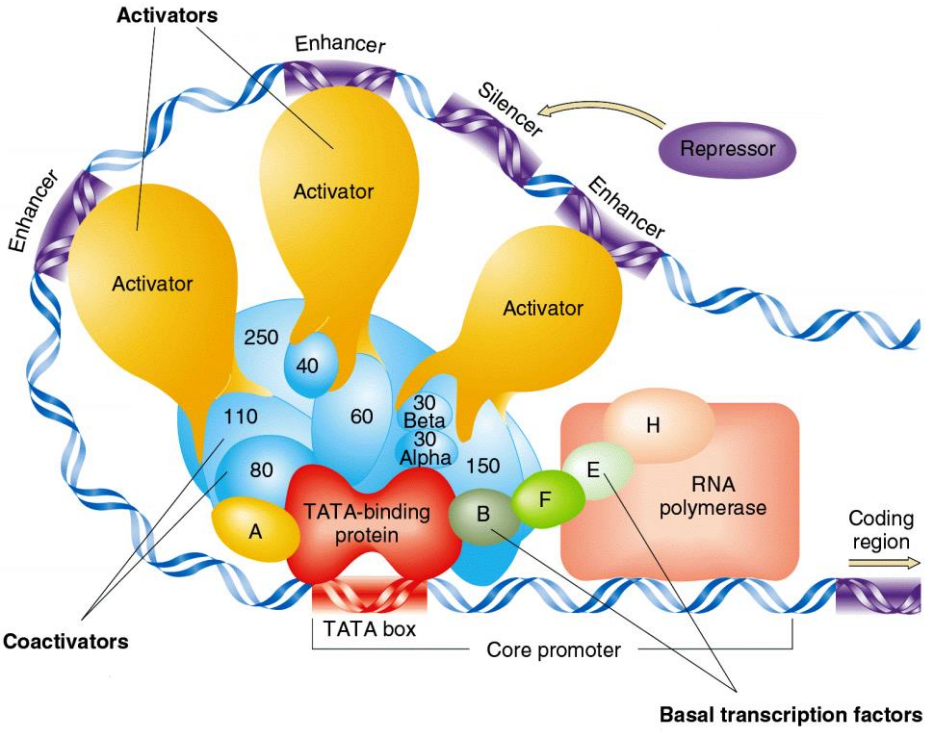
- General TFs or basal TFs
- Specific TFs

③ Structure

- **DBD**, which attach to specific sequences of DNA (enhancer or promoter sequences) adjacent to regulated genes.
- **SSD** (*e.g.*, a ligand binding domain), which senses external signals and in response transmit these signals to the rest of the transcription complex, resulting in up or down regulation of gene expression.
- **TAD**, which contain binding sites for other proteins such as transcription coregulators.

④ Functions

- **activation, upregulation, or promotion**
- **repression, downregulation, or suppression**



(2) Initiation of transcription by RNA Pol II

① DNA elements

❖ Core Element of Promoter

- TATA box (hogness box): A conserved A·T-rich septamer found about 25 bp before the startpoint; the protein binding to this TATA box called TATA-box binding protein (TBP), which is an important subunit of the transcription factor TF II D.
- Inr (initiator sequence)

❖ UPE or UAS

UPE (or UAS): the regulation elements locate upstream of Inr. Such as CAAT box, GC box, -35-like box.

② **TFs**

RNA Pol II does not contain a subunit similar to σ factor, which can recognize the promoter and unwind the DNA double helix. In eukaryotes, these two functions are carried out by TFs. RNA Pol II is associated with six general transcription factors, designated as TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH. **TFIID** consists of **TBP** and **TAFs**. TBP binds the core promoter. TAFs assist TBP in this process.

Factor	Number of Subunits	Mw. (kD)	Function
TFIID - TBP	1	38	Recognize core promoter; Recruit TFIIB
TFIID - TAFs	12	15 - 250	Assist transcription activation; Assist promoter recognition
TFIIA	3	12, 19, 35	Stablize TFIID and promoter binding
TFIIB	1	35	Recruit RNA Pol II and TFIIF
TFIIF	2	30, 74	Assist RNA Pol II to reach promoter
TFIIE	2	34, 57	Recruit TFIIH; Modulate TFIIH helicase, ATPase and kinase activities
TFIIH	9	89,80,62,52,44 34,32,38,40	Promoter melting using helicase activity; DNA repair

* TFIIE is recruited by RNA Pol II.

†Modified from: Roeder, R.G. 1996. *Trends in Biochem. Sci.* **21**:327-334.

③ **Initiation**

TFIID + TATA box → TFIIA → TFIIB → TFIIF (RNA Pol) → TFIIE → TFIIH → PIC → CTD domain phosphorylated, RNA Pol keeps moving → RNA new chain

(3) Initiation of transcription by RNA Pol I

① **DNA elements**

The regulatory region of pre-rRNA gene contains a core element and an upstream control element (UCE). Binding of two upstream binding factors (UBF) to both elements may include DNA looping, and subsequently recruiting TATA-box binding protein (TBP) and TBP-associated factors (TAFs). Finally, RNA Polymerase I joins the complex and completes the assembly process.

② **TFs**

③ **Initiation**

(4) Initiation of transcription by RNA Pol III

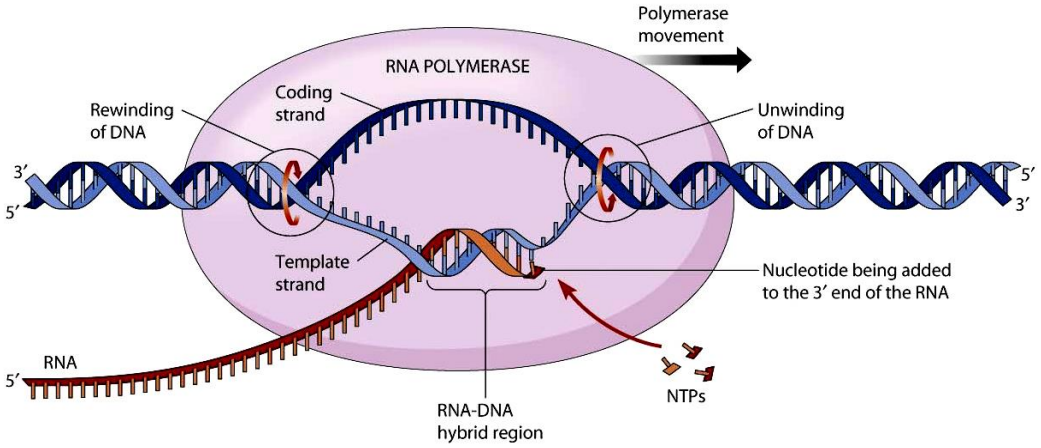
① **Initiation of transcription of tRNA gene**

The regulatory region of tRNA gene contains A box and B box located inside the transcription unit. The PIC assembly begins with the binding of TFIIIC to both elements.

② **Initiation of Transcription of 5S rRNA gene**

The regulatory region of 5S-rRNA gene contains a C box, also located inside the transcription unit. The PIC assembly begins with the binding of TFIIIA to the C box.

3.4 Elongation and Termination of Transcription in Eukaryotes



1. Elongation of Transcription

After PIC is assembled at the promoter, TFIIH can use its helicase activity to unwind DNA. This requires energy released from ATP hydrolysis. The DNA melting starts from about -10 bp. Then, RNA Pol II uses nucleoside triphosphates (NTPs) to synthesize a RNA transcript. During RNA elongation, TFIIF remains attached to the RNA polymerase, but all of the other transcription factors have dissociated from PIC.

The carboxyl-terminal domain (CTD) of the largest subunit of RNA Pol II is critical for elongation. In the initiation phase, CTD is unphosphorylated, but during elongation it has to be phosphorylated. This domain contains many proline, serine and threonine residues.

2. Termination of Transcription

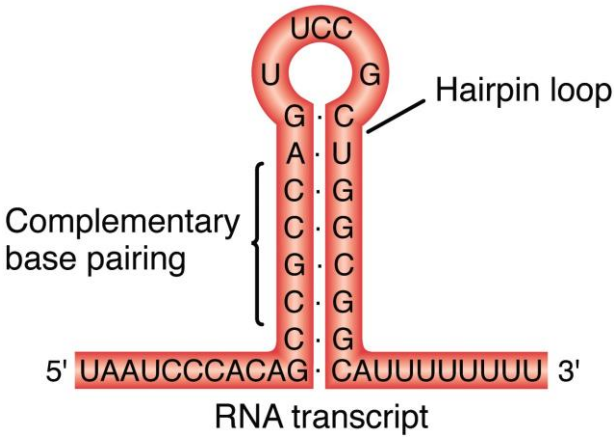
Terminator ---- A terminator, or transcription terminator is a section of genetic sequence that marks the end of gene or operon on genomic DNA for transcription.

(1) Termination of transcription in prokaryotes

Bacteria use two different strategies for transcription termination:

① **ρ-independent termination (strong)**

Intrinsic transcription terminators where a hairpin structure forms within the nascent transcript that disrupts the mRNA-DNA-RNA polymerase ternary complex. The ρ-independent termination signal is a stretch of 30-40 bp sequence, consisting of many GC residues followed by a series of T ("U" in the transcribed RNA). The resulting RNA transcript will form a stem-loop structure to terminate transcription.



② **ρ-dependent termination (weak)**

Rho-dependent transcription terminators that require Rho factor, an RNA helicase protein complex, to disrupt the nascent mRNA-DNA-RNA polymerase ternary complex.

ρ-factor functions solely at the stage of termination. In *E. coli*, it is a **~275 kD hexamer** of identical subunits. The subunit has an RNA-binding domain and an ATP hydrolysis domain. ρ is a member of the family of ATP-dependent hexameric **helicases** that function by passing nucleic acid through the hole in the middle of the hexamer formed from the RNA-binding domains of the subunits. ρ functions as an ancillary factor for RNA Polymerase. It also has **NTPase activity**. ρ-dependent terminators account for about half *E. coli* terminators.

(2) Termination of transcription in eukaryotes

Transcription termination in eukaryotes is less understood.

- RNA polymerase I terminates when it comes to a polymerase-specific DNA binding protein attached at the termination site.

- RNA polymerase III terminates at a series of U residues but does not require an upstream stem-loop be present in the mRNA.
- RNA polymerase II transcripts are essentially terminated by the cleavage near the polyadenylation site followed by the addition of the poly(A) tail. The cleaved 3' transcriptional product is rapidly degraded as are un-polyadenylated transcripts.

3. Antitermination of Transcription

(1) Antiterminator

Antiterminator in transcription is a protein that, when bound at its normal attachment sites in DNA, lets RNA polymerase read through normal terminator sequences.

Example, the N- and Q- gene products of λ phage.

(2) Antitermination (Read through)

Transcription or translation beyond the normal termination signals in DNA or mRNA, respectively.

(3) Mechanisms

① **Damaged stem-loop structure**

trp operon (attenuation)

② **Functions of antiterminator**

◆ **N protein :** prevent ρ- independent termination

◆ **S4 protein** : prevent ρ- dependent termination

(4) Applications

3.5 The Features of mRNA

1. The Features of mRNA in Prokaryotes

(1) short half life

(2) polycistronic mRNA

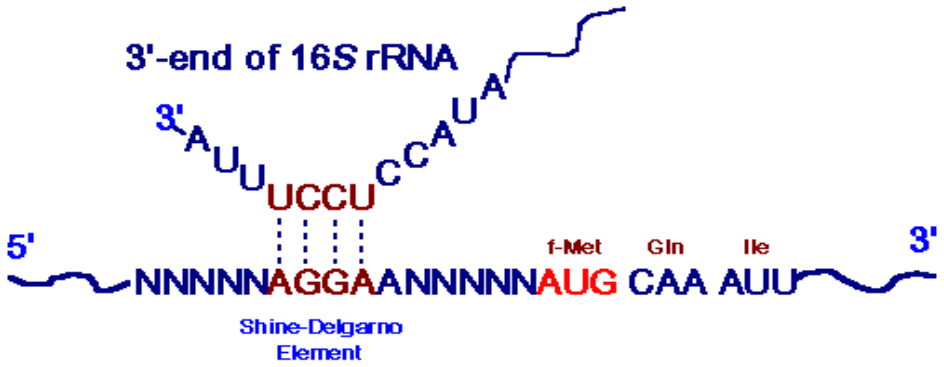
The mRNAs of prokaryotes are polycistronic i.e. it contains several sites for initiating and terminating polypeptide synthesis while eukaryote mRNAs are monocistronic. *lac* operon, *trp* operon……

(3) 5'-SD sequence

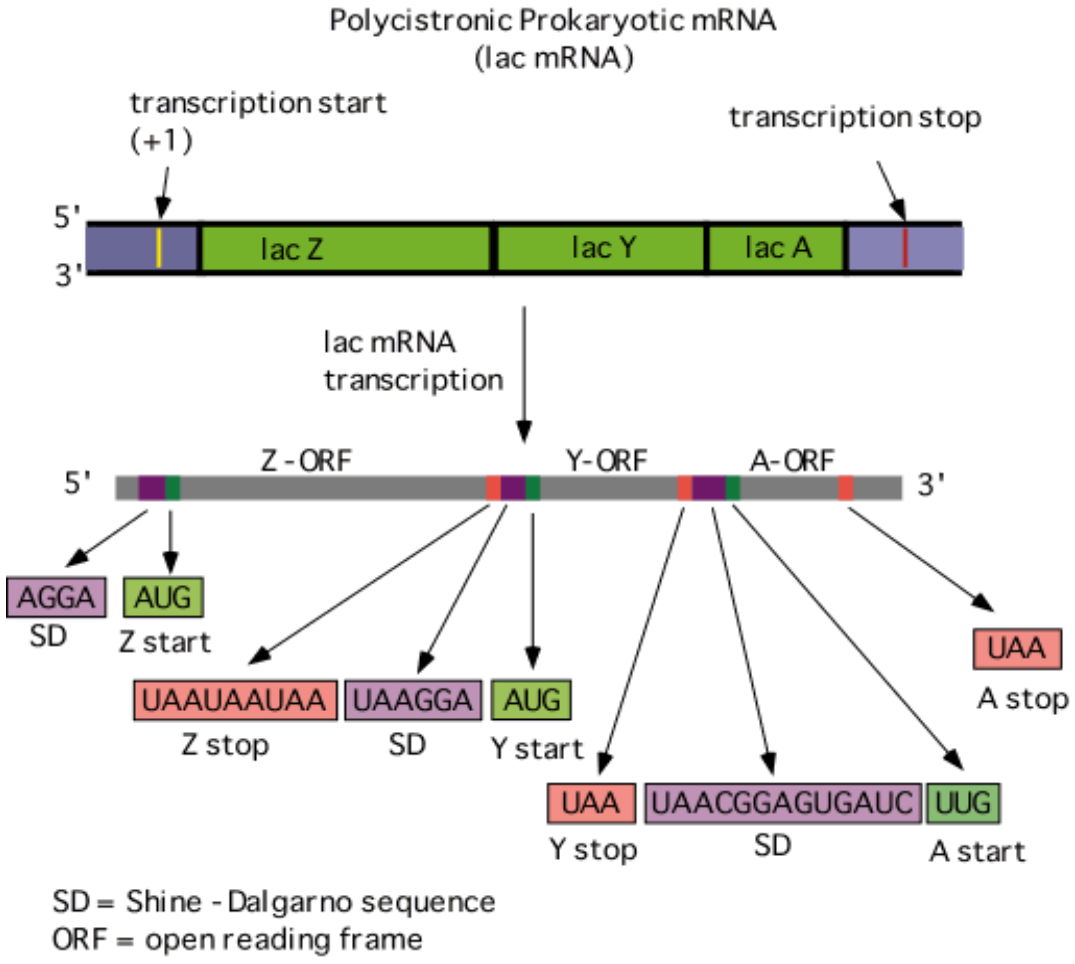
The Shine-Dalgarno sequence (or Shine-Dalgarno box), proposed by Australian scientists John Shine and Lynn Dalgarno in 1974, is a RBS(rbs) in the mRNA, generally located upstream of the start codon AUG.

The consensus sequence is AGGAGG. The complementary sequence

(CCUCCU), is called the anti-SD sequence and is located at the 3' end of the 16S rRNA in the ribosome.



(4) No 5'-cap structure and no or short 3'-p
(5) AUG, GUG, UUG
(6) Undergoing very little processing after transcription.
(7) Translation begins while the mRNA is still being transcribed on DNA.



2. The Features of mRNA in Eukaryotes

(1) long half life

(2) monocistronic mRNA

(3) Kozak consensus sequence

(4) 5’-cap: In eukaryotes, the mRNA is modified at both ends - the 5' end is "Capped" with a modified GTP, and to the end is added a long "tail" of poly(A) (usually greater than 150 A's are added).

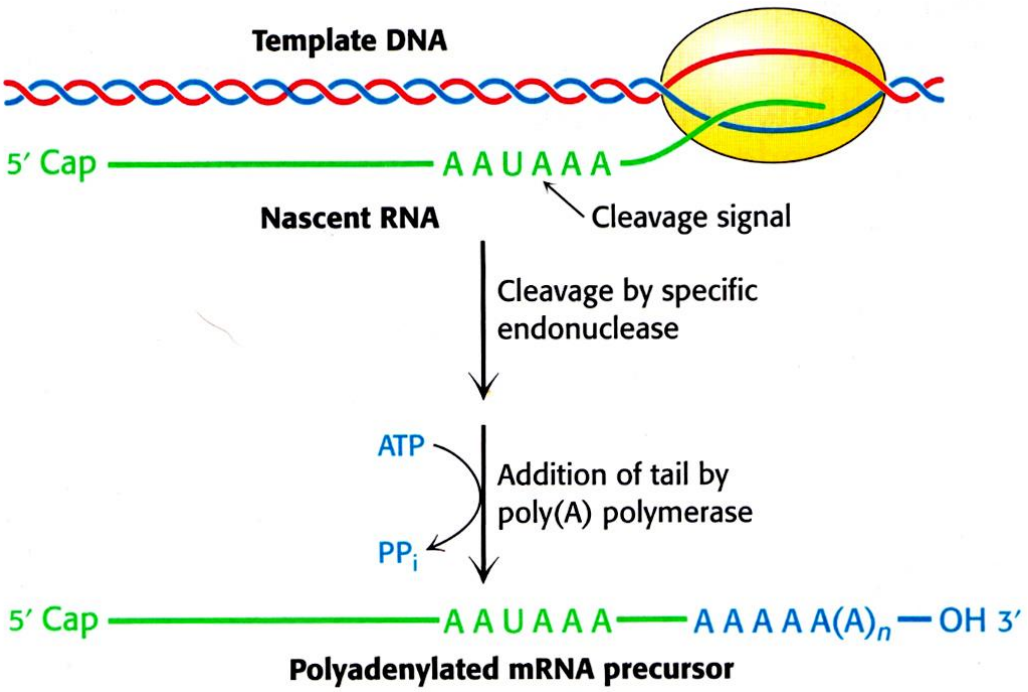
① Structure

② The 5’-cap has 4 main functions:

- Regulation of nuclear export
- Prevention of degradation
- Promotion of translation
- Promotion of 5‘ proximal intron excision

(5) 3’-polyA

① AAUAAA sequence



② Polyadenylate polymerase

- adds 80 – 250 A to the end of mRNA
- requires the cleaved mRNA as a primer
- does NOT require a template for synthesis
- increases mRNA stability

③ Proteins involved

④ Functions of 3'-poly(A)

- promotes export from the nucleus and translation
- Signal for RNA processing
- protects the mRNA from degradation.

(6) AUG

(7) RNA processing → mature mRNA

(8) Translation begins only after transcription

3.6 RNA Processing

1. RNA components in cells

Major Types of RNA	Function of RNA
ribosomal - rRNA	mRNA translation
transfer - tRNA	mRNA translation
messenger - mRNA	template of protein translation/regulatory
heterogeneous nuclear - hnRNA	precursor of mRNAs

(1) Coding RNA – mRNA

hnRNA---heterogeneous nuclear RNA, a diverse group of long primary transcripts formed in the eukaryotic nucleus, many of which will be processed to mRNA molecules by splicing.

(2) Non-coding RNAs(ncRNA)

ncRNA - is any RNA molecule that is not translated into a protein.

- rRNA(ribosomal RNA)
- tRNA(transfer RNA)
- Small RNA molecules

Small nuclear RNA (snRNA) - a class of eukaryotic small RNA molecules found in the nucleus, usually as ribonucleoproteins, and apparently involved in processing heterogeneous nuclear RNA.

<u>Functions of Non-Coding RNA:</u>

- Regulation of transcription

- Replication of eukaryotic chromosome
- RNA processing
- RNA modification
- RNA editing
- mRNA stability and degradation
- Regulation of translation
- Protein translocation

2. Types of RNA processing

RNA processing---is to generate a mature mRNA (for protein genes) or a functional tRNA or rRNA from the primary transcript.

RNA splicing---is a process that removes introns and joins exons in a primary transcript.

RNA editing---describes those molecular processes in which the information content in an RNA molecule is altered through a chemical change in the base makeup.

- ❖ End-modification: 5'-cap, 3'-poly(A)
- ❖ Splicing : pre-mRNA, pre-rRNA, pre-tRNA
- ❖ Cutting events: pre-rRNA, pre-tRNA
- ❖ Chemical modification: mRNA, rRNA, tRNA

(1) End-modification of RNA Primary Transcript

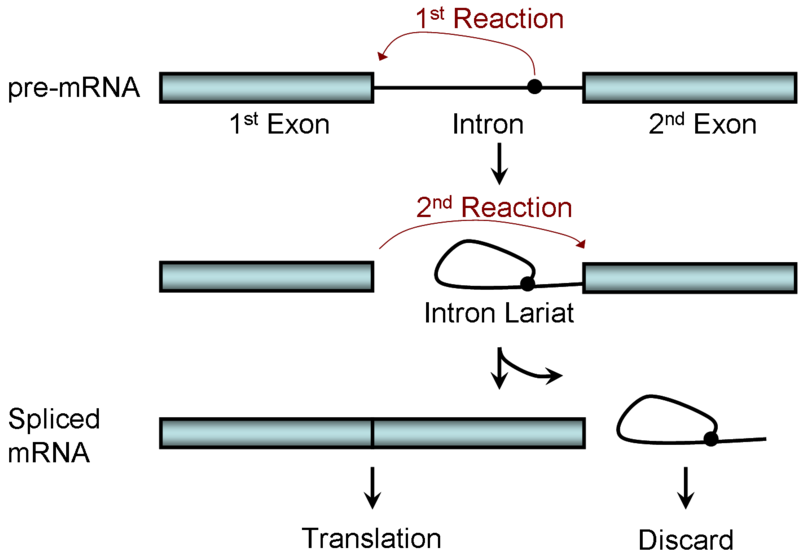
5'-Cap, Poly-(A) Tail

(2) RNA Splicing

- ◆pre-mRNA splicing
- ◆tRNA splicing
- ◆Group I intron splicing (Self-splicing)
- ◆Group II intron splicing (Self-splicing)
- ◆ Alternative splicing
- ◆ Cis- or Trans-splicing

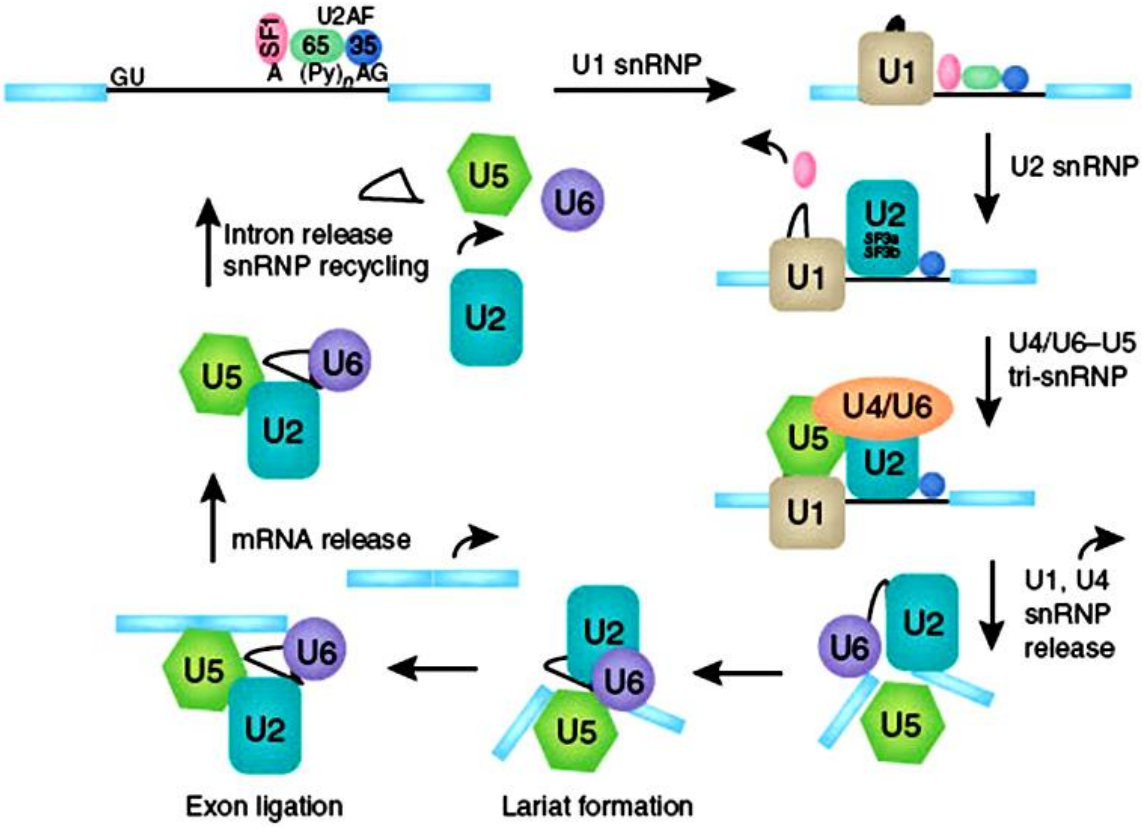
① Pre-mRNA Splicing (GU-AG, AU-AC)

- ❖ Conserved sequence - 5'-splice site, 3'-splice site, branch site
- ❖ Steps (two transesterification reactions)



❖ **snRNPs and spliceosome**

- U1-snRNP binds 5'-splice site, U2AF binds pyrimidine tract
- U2-snRNP binds branch site
- U4/U6 and U5 snRNPs enter complex mixtures, forming splicesome
- U1snRNP is released, U6-snRNP binds at 5'-splice site
- U4-snRNP is released, U2/U5/U6-snRNP remain bound
- Spliced RNA is released and lariat intron is debranched and degraded



mRNA transport (5'-cap, intron-free, 3'-poly(A))

mRNA degradation

wipe off poly(A) → disengage (5'-cap) → exonuclease digest → mRNA degradation

② tRNA splicing

Cleavage; Splicing; Addition of CCA to 3'-end; Base modification

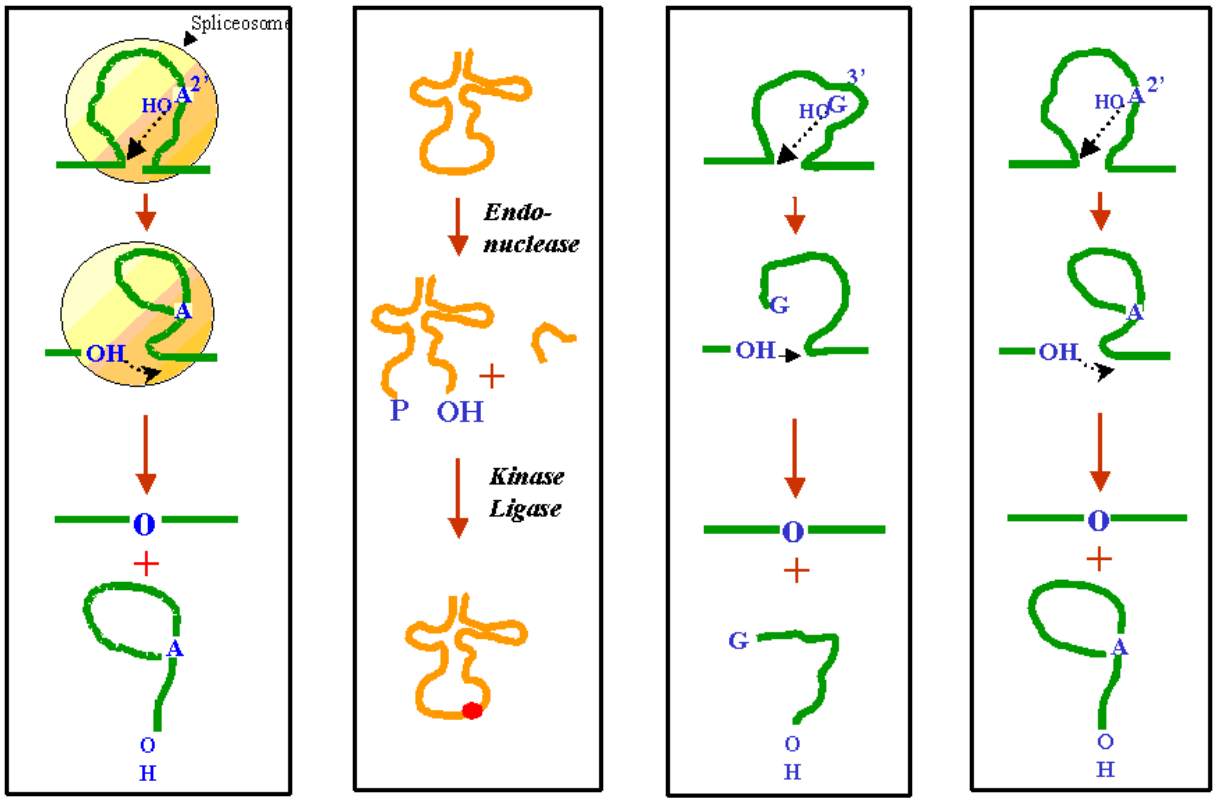
③ Self-splicing

(a) Group I intron splicing

- conserved secondary structure
- self- catalyse
- linear intron released

(b) Group II intron splicing

- conserved secondary structure
- self- catalyse
- lariat intron released
- similar of pre-mRNA splicing in splicing mechanism



④ **Alternative splicing**

⑤ **Cis-splicing and Trans-splicing**

(3) Cutting Events

① Pre-rRNA processing

② Pre-tRNA processing

(4) Chemical modification of RNA

3. RNA Editing

(1) Types

U-insertion/deletion

C-insertion

C→ U (and U to C?) editing

tRNA editing

A → I editing

SnoRNA-mediated nucleotide modification of rRNAs

(2) Editing mechanism

① Editing by deaminations

② Editing by insertion or deletion of U

▲gRNAs are small and complementary to portions of the edited mRNA

▲base-pairing of gRNA with unedited RNA gives mismatched regions, which are recognized by the editing machinery

▲machinery includes an endonuclease, a terminal uridylyl transferase (TUTase), and a RNA ligase

(3) Biological significances

3.7 Ribozymes

(1) What is a ribozyme?

A ribozyme (also called RNA enzyme or catalytic RNA) is an RNA molecule that catalyzes a chemical reaction.

hydrolysis of phosphodiester bonds

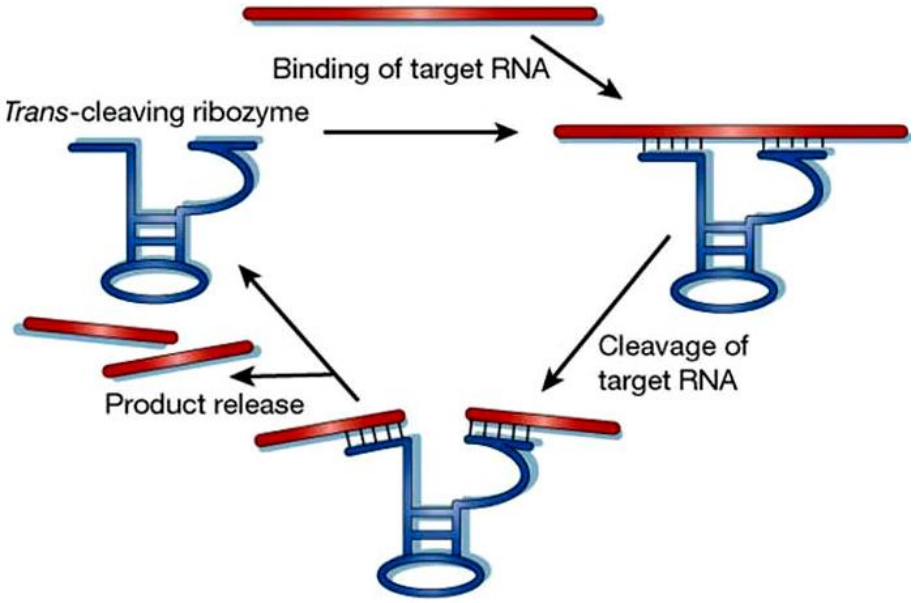
aminotransferase activity of the ribosome.

(2) Structure of a ribozyme

(3) Types of Ribozyme Page 104

- the hammerhead ribozyme (plant virus)
- the hairpin ribozyme (plan virus)
- hepatitis delta ribozyme (human virus)
- group I and group II intron ribozyme
- RNAse P (tRNA maturation)

(4) Mechanism of Ribozyme



3.8 Reverse Transcription

1. What is reverse transcription?

--------The process by which DNA is synthesized from an RNA template by means of the enzyme reverse transcriptase.

2. Reverse Transcriptase

---- is a RNA-dependent DNA polymerase that permits the retroviruses to use their viral RNA as a template to make DNA.

3. Processes of Reverse Transcription

(1) Sequence features of retrovirus genomes

R Region: A short (18-250nt) repeat sequence, 'terminally redundant'.

U5: A unique, non-coding region of 75-250nt

PBS (Primer Binding Site): 18nt complementary to the 3' end of the specific tRNA primer.

Leader: A relatively long (90-500nt) non-translated region.

Polypurine Tract: A short (~10) run of A/G residues responsible for initiating (+)strand synthesis.

U3: A unique non-coding region of 200-1,200nt which contains the promoter elements responsible for transcription of the provirus.

(2) Mechanism

◆A retrovirus-specific cellular tRNA hybridizes with a complementary region called the primer-binding site (PBS).

◆A DNA segment is extended from tRNA based on the sequence of the retroviral genomic RNA.

◆The viral R and U5 sequences are removed by RNase H.

◆First jump: DNA hybridizes with the remaining R sequence at the 3' end.

◆A DNA strand is extended from the 3' end.

◆Most viral RNA is removed by RNase H.

◆A second DNA strand is extended from the viral RNA.

◆Both tRNA and the remaining viral RNA are removed by RNase H.

◆Second jump: The PBS region of the second strand hybridizes with the PBS region of the first strand.

◆Extension on both DNA strands. LTR stands for "long terminal repeat".

Summary

1. Concepts

Gene expression	Antitermination	RNA splicing
Promoter	Read through	snRNA
σ factor	Coding RNA	RNA editing
Terminator	Non-coding RNA	gRNA
Antiterminator	hnRNA	Ribozyme
Termination	RNA prodessing	Reverse transcription
General TF	Specific TF	

2. Regulatory elements of prokaryotes and eukaryotes for RNA transcription
3. RNA polymerase of prokaryotes and eukaryotes
4. Procedures of RNA transcription
5. mRNA structure of prokaryotes and eukaryotes
6. Types of RNA processing
7. Types of RNA splicing and their mechanisms
8. RNA editing mechanism
9. Mechanism of reverse transcription
10. Ribozyme and silencing mechanism

Further reading

Questions and Problems

第四章　生物信息的传递(下)——From mRNA to Protein

Chapter 4　Translation: From RNA to Protein

教学目的：

本章要求学生掌握遗传密码、tRNA 的结构、核糖体的结构、蛋白质合成过程及相关因子、分子伴侣的功能、蛋白质转运机制、蛋白质的降解机制等。

重点与难点：

1、核糖体的结构

2、分子伴侣的功能

3、蛋白质的降解机制

教学方法：

讲述式和启发式教学 (适当结合讨论法)

多媒体教学

English Animation

课时安排：

8 课时

主要教学内容：

4.1　Genetic Codon

1　Discovery

What is genetic code? ——The genetic code is the set of rules by which information encoded in genetic material (DNA or RNA sequences) is translated into proteins (amino acid sequences) by living cells.

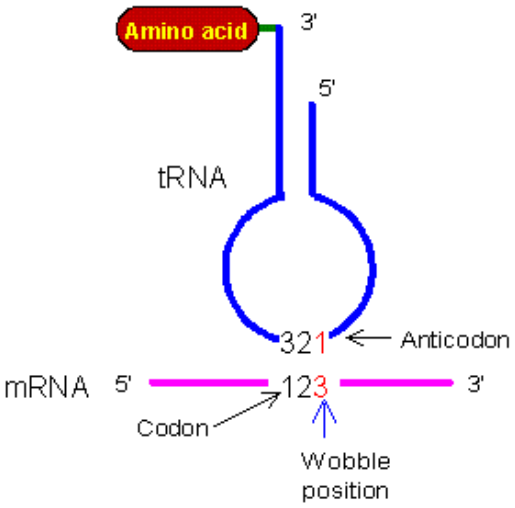
2　Features

(1) Continuity; (2) Degeneracy; (3) Generality; (4) Particularity

(5) Codon / Anticodon

3　Wobble hypothesis

Wobble pairing



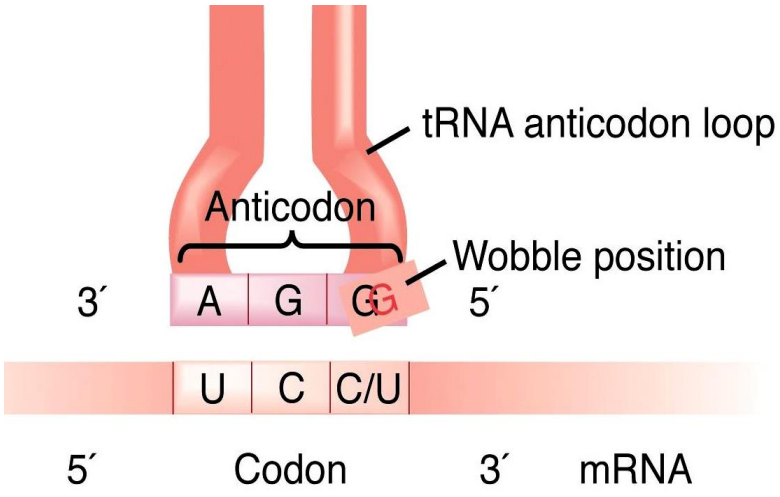
Wobble bases

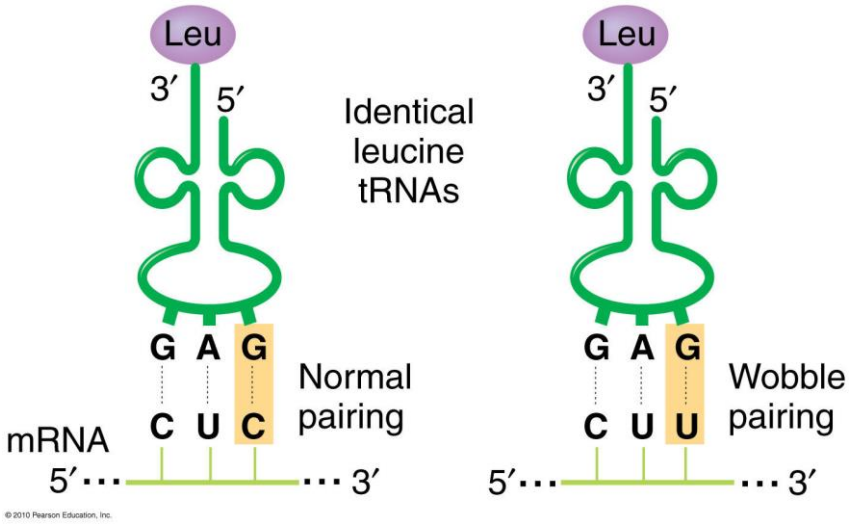
tRNA	C	A	G	U	I
mRNA	G	U	C U	A G	C A U

Wobble bases

mRNA	C	A	G	U
tRNA	G I	U I	C U	A G I

Wobble hypothesis





4.2 tRNA —— Carrier in translation System

1 tRNA structure

(1) Secondary Structure – Cloverleaf

(2) Tertiary structure –Upside-down “L”

(3) Features of tRNA structure

---Each is a single chain containing between 73 ~ 93 ribonucleotides (~25 kd).

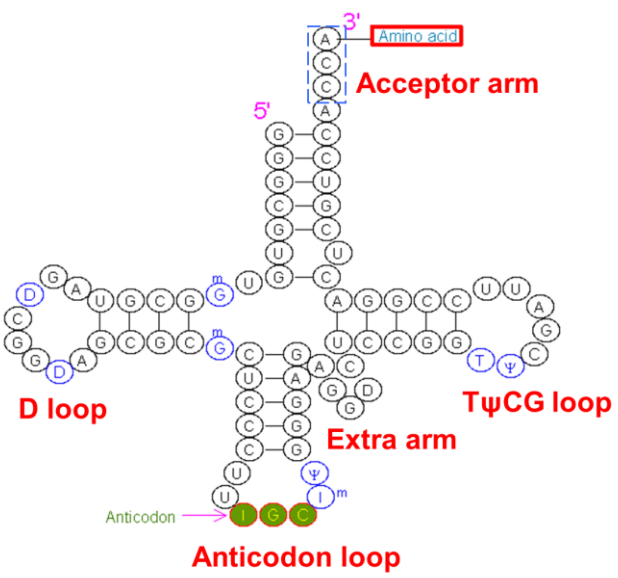
---They contain many unusual bases, typically between 7 and 15 per molecule.

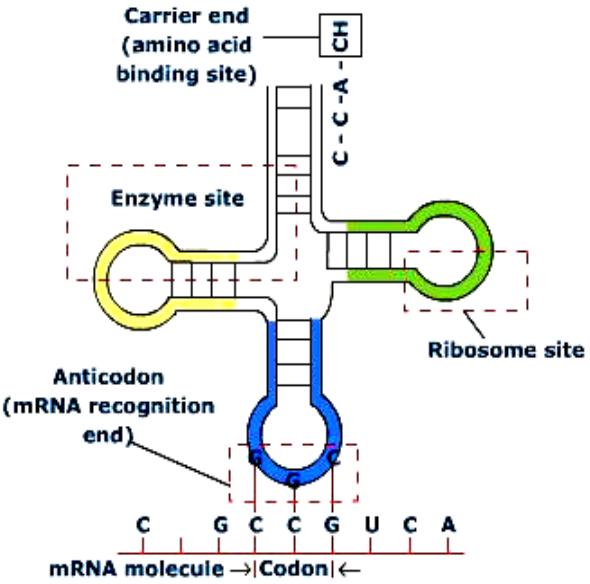
---About half the nucleotides in tRNAs are base-paired to form double helices.

---The 5′end of a tRNA is phosphorylated. The 5′terminal residue is usually pG.

---The activated amino acid is attached to a hydroxyl group of the A located at the end of the 3′CCA component of the acceptor stem.

---The anticodon is present in a loop near the center of the sequence.





2 tRNA Functions

tRNA serves as the adapter molecule that binds to a specific codon and brings an amino acid for incorporation into the polypeptide chain.

3 tRNA Types

(1) tRNA in initiation and elongation

(2) Cognate tRNA

(3) Supressor tRNA

4 Aminoacyl tRNA synthetase

(1) Concept ---An aminoacyl tRNA synthetase is an enzyme that catalyzes the esterification of a specific amino acid or its precursor to one of all its compatible cognate tRNAs to form an aminoacyl-tRNA.

(2) Function

(3) Types

---Class I has two highly conserved sequence motifs. It aminoacylates at the 2'-OH of an adenosine nucleotide, and is usually monomeric or dimeric (one or two subunits, respectively).

---Class II has three highly conserved sequence motifs. It aminoacylates at the 3'-OH of the same adenosine, and is usually dimeric or tetrameric (two or four subunits, respectively).

Although phenylalanine-tRNA synthetase is class II, it aminoacylates at the 2'-OH.

4.3 Ribosome

1 Discovery

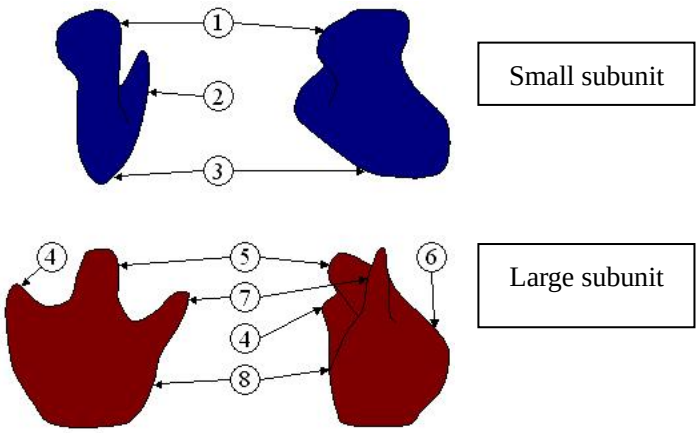
2 Components

	size of ribosome	Size of subunits	Ratio (RNA : protein)
Prokaryotes	70S	50S, 30S	2 : 1
Mammalian	80S	60S, 40S	3 : 2

3 Structure

(1) Two subunits

① Head.
② Platform.
③ Base.
④ Ridge.
⑤ Central protuberance.
⑥ Back.
⑦ Stalk.
⑧ Front.



(2) Ribosomal protein (r-protein)

(3) Ribosomal RNA

(4) tRNA Binding sites in Ribosome

mRNA binding site, A site, P site, E site

4 Functions of ribosome

5 Autoregulation of ribosome synthesis

6 Cycle of ribosome

4.4 Protein Synthesis

1 Activation of amino acid

---Each amino acids are recognized by its specific aminoacyl-tRNA synthetase.

---Fidelity of translation depends on this step.

2 Protein synthesis in prokaryotes

(1) Initiation

① Components: Ribosome 30S subunit, Template mRNA, fMet-tRNAfMet, IF-1, IF-2, IF-3, GTP, Ribosome 50S subunit, $Mg2^{+}$

② Initiation steps

IF-3 binds to the 30S ribosomal subunit, freeing it from its complex with the 50S subunit. IF-1 assists binding of IF-3 to the 30S ribosomal subunit. Binding of IF-1 also occludes the A site domain of the small subunit, helping to insure that the initiation aminoacyl-tRNA, fMet-$tRNA^{fMet}$, can bind only in the P site and that no other aminoacyl-tRNA can bind in the A site during initiation. IF-2 is a small GTP-binding protein. IF-2-GTP binds the initiator fMet-$tRNA^{fMet}$ and helps it to dock with the small ribosome subunit. A region of the mRNA upstream of the initiation codon, the SD sequence, base pairs with the 3' end of the 16S rRNA. This positions the small ribosomal subunit in relation to the initiation codon.

(2) Elongation of peptide chain

AA-tRNA →A site: EF-Tu-GTP binds and delivers an aminoacyl-tRNA to the A site. Formation of peptide bond: The peptide attached to the peptidyl-tRNA at the P site is transferred to the new aminoacyl-tRNA at the A site, generating a peptidyl-tRNA with a longer peptide.

Translocation: The empty tRNA at the P site is ejected from the ribosome and the peptidyl-tRNA generated at the A site takes over the vacant P site. The A to P switch is catalyzed by the elongation factor EF-G in bacteria.

(3) Termination of Protein Biosynthesis in Prokaryotes

RF-1 and RF-2 recognize and bind to STOP codons. One or the other binds when a stop codon is reached. RF-3-GTP facilitates binding of RF-1 or RF-2 to the ribosome.

3 Protein Biosynthesis in Eukaryotes

(1) Initiation

① Cap-dependent initiation

43S Pre-initiation complex

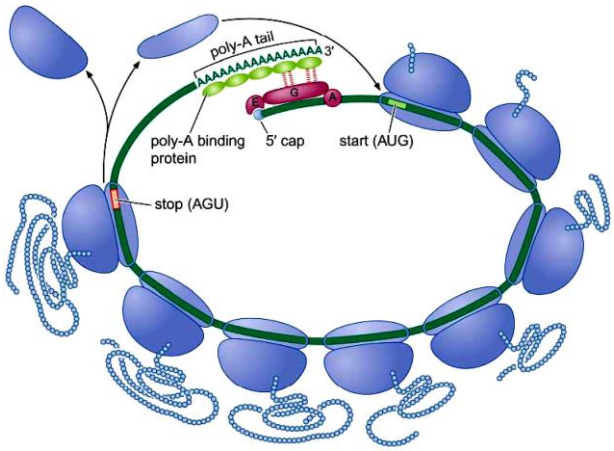
eIF-3: The eukaryotic Initiation Factor 3 (eIF3) is associated with 40S small ribosomal subunit, and plays a role in keeping the large ribosomal subunit from prematurely binding.

eIF-1A: (formerly known as eIF-4C) eIF1A is an essential eukaryotic translation initiation factor. The protein is required for the binding of the 43S complex to the 5' end of capped RNA.

eIF-2 mediates the binding of Met-tRNA^{Met} to the ribosome. α-subunit: regulatory subunit; β-subunit : interaction with eIF2B; γ-subunit : docking site for GTP/GDP, Met-tRNA^{Met}.

40S initiation complex

eIF-4F: The eukaryotic translation initiation factor 4F (eIF-4F) mediates 40S ribosomal subunit binding to the 5'-end of capped mRNA. eIF-4F is a complex containing three proteins: eIF-4E, the cap-binding subunit; eIF-4A, an RNA-dependent ATPase/ATP-dependent RNA helicase; and eIF-4G, a high-molecular weight protein that acts as a scaffold for binding eIF-4E and eIF-4A. eIF-4G interacts with the 40S ribosome binding factor eIF3 and the poly(A)-binding protein, thereby establishing a critical link between mRNA and the ribosome.



80S initiation complex (eIF-5)

② Cap-independent initiation --- Internal Ribosome Entry Site (IRES)

An IRES, is a nucleotide sequence that allows for translation initiation in the middle of a mRNA sequence as part of the greater process of protein synthesis. Commonly, IRES is located in the 5'UTR of RNA viruses and allow translation of the RNAs in a cap-independent manner. The virus accomplishes this feat by cleaving eIF-4G so that it cannot interact with eIF-4E. Interaction between these two initiation factors is necessary for mRNA 5'-cap to 3'-poly-A-tail loop formation, which is usually a necessary event for initiation of translation. The virus may even use the eIF-4G to aid in intitiation of IRES-mediated translation.

(2) Elongation

(3) Termination

4.5 Protein Processing and Modification

1. Types

(1) Removing fMet or Met

(2) Formation of S-S bond

(3) Chemical modification of amino acid

(4) Intein splicing

2. Inhibitors of Protein Biosynthesis

4.6 Protein Folding & Molecular Chaperone

1. Molecular Chaperones

(1) Concept

----A large group of unrelated protein families whose role is to stabilize unfolded proteins, unfold them for translocation across membranes or for degradation, and/ or to assist in their correct folding and assembly.

(2) Functions of Chaperones

- protein folding;
- translocation;
- stabilization;
- prevent aggregation;
- degradation

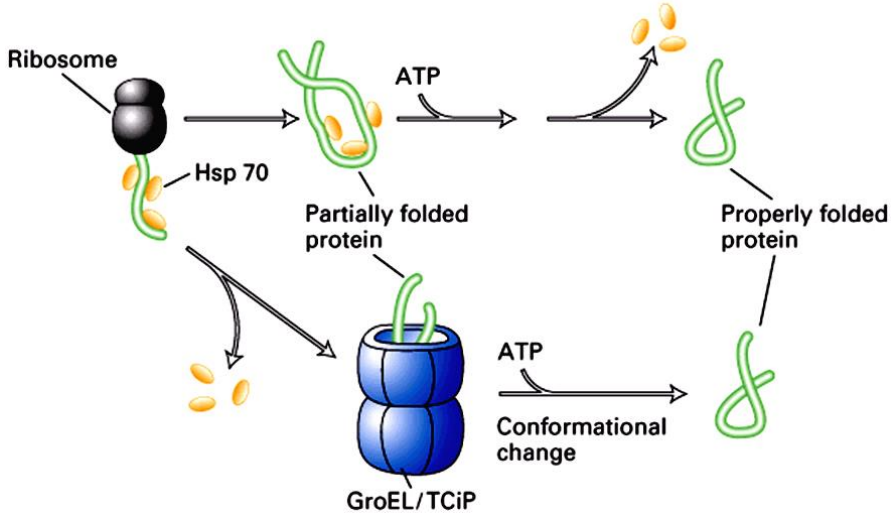
(3) Main Types of Molecular Chaperone

① Stress protein

All of the Hsp70 proteins have three major functional domains: N-terminal ATPase domain; substrate binding domain; C-terminal domain rich inα-helix.

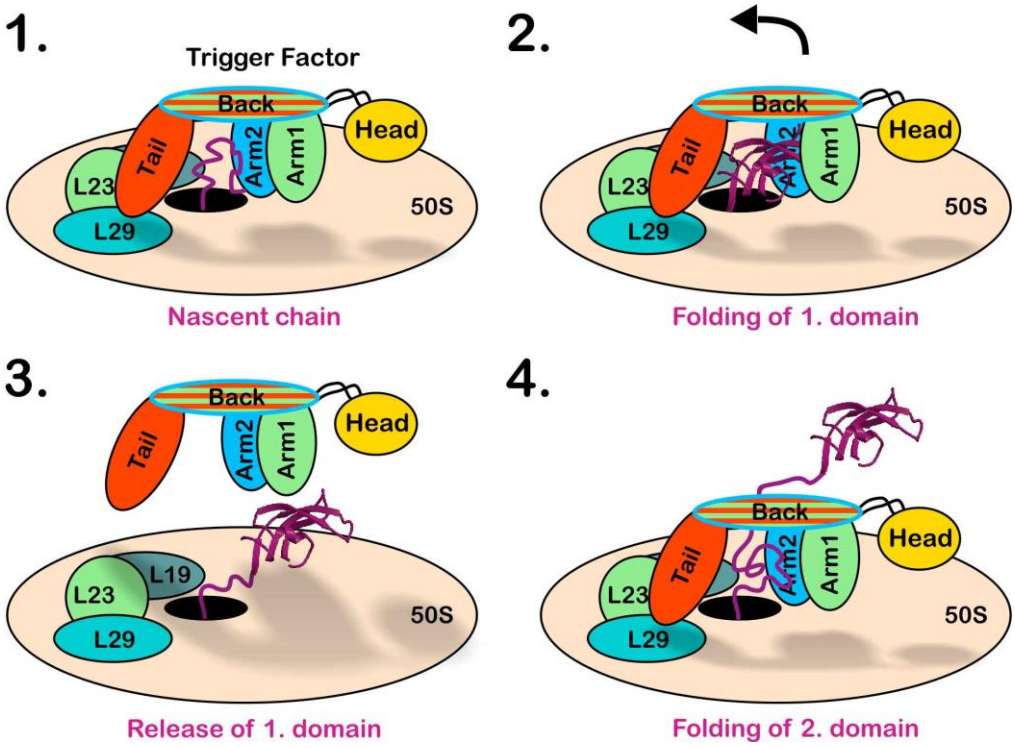
② Chaperonin

Structure: resemble two donuts stacked on top of one another to create a barrel; each ring is composed of either 7, 8 or 9 subunits.



③ Trigger Factor (TF)

Model for the Trigger Factor mechanism



2. Protein folding

(1) Concept ---Protein folding is the physical process by which a polypeptide folds into its characteristic and functional three-dimensional structure.

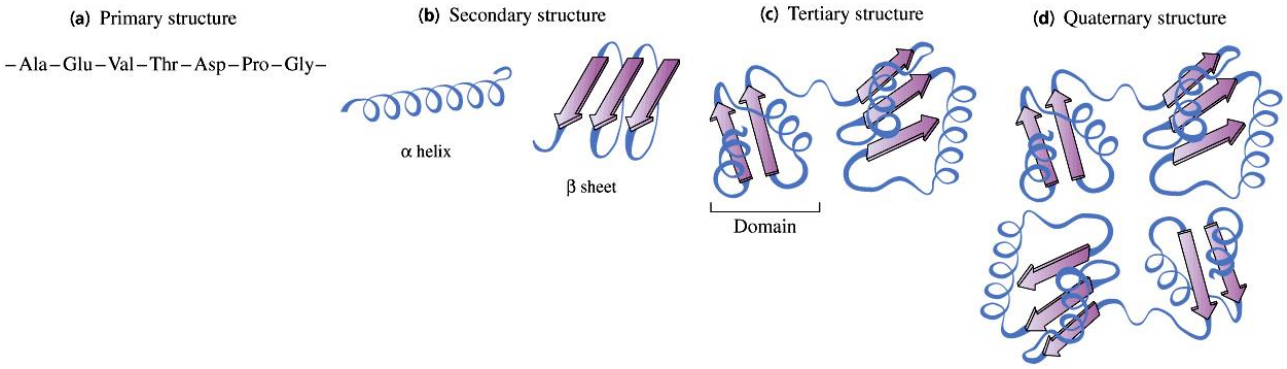
(2) Forces Involved in the Protein Folding---Electrostatic interactions; van der Waals interactions; Hydrogen bonds; Hydrophobic interactions.

(3) Models for Protein Folding--- Framework Model; Hydrophobic Collapse Model; Nuclear-Condensation Model; Diffusion-Collision model; Funnel or Energy-landscape models.

(4) Significance of Protein Folding: Structural: unfolding→3D-protein; Functional: inactive→active protein; Medical: therapy.

3. Protein Structure

- Primary Structure
- Secondary Structure
- Supersecondary Structure
  - Motif
  - Domain
- Tertiary Structure
- Quaternary Structure



4.7 Protein translocation

1. Cotranslational translocation

(1) Signal peptide

(2) Signal hypothesis

(3) How do proteins get into the ER?

2. Posttranslational Transport

(1) Translocation of protein into mitochondria

(2) Translocation of protein into chloroplast

(3) Nuclear localization sequence

(4) Protein translocation in bacteria

4.8 Protein Degradation

Some protein degradation pathways are nonspecific - randomly cleaved proteins seem to be rapidly degraded. There is also a selective, ATP-dependent pathway for protein degradation - the ubiquitin-mediated pathway.

1. Ubiquitin

(1) Concept--- Ubiquitin is a small protein that occurs in all eukaryotic cells. Its main known function is to mark other proteins for destruction, known as proteolysis.

► Ubiquitin is a highly-conserved, 76 residue (8.5 kD) protein found widely in eukaryotes

► Proteins are committed to degradation by conjugation with ubiquitin

(2) Components of Ubiquitin

→ ubiquitin- activating enzyme (E1): it forms a thioester bond with C-terminal Gly of ubiquitin. This is the ATP-dependent step.

→ ubiquitin-conjugating enzyme (E2): Ubiquitin is then transferred to the ubiquitin-carrier protein E2.

→ ubiquitin ligase (E3): E3 selects proteins for degradation. the E2-S~ubiquitin complex transfers ubiquitin to these selected proteins.

2. Proteasomes

→ 19S regulatory particle - a regulatory cap of 16-18 proteins (6 with ATPase activity). It has ATPase active sites and ubiquitin binding sites. The 19S regulatory cap complex recognizes multi-ubiquitinated proteins, unfolds them, removes ubiquitin chains, and provides a passageway for threading unfolded proteins into the core complex.

→ 20S core particle - a barrel core of 4 stacked protein rings with protease activities. 7 α-type proteins form each of the two α-rings, at the ends of the cylindrical structure. 7β-type proteins form each of the 2 central β-rings.

3. Ubiquitin-dependent protein degradation pathways

(1) Ubiquitination and targeting: → Activation of ubiquitin by E1 (ATP); → Transfer of ubiquitin from E1 to E2; → E3 interacts with E2 and substrate.

(2) Unfolding (ATP) and translocation: → The ubiquitinated protein is recognized by the 19S regulatory particle in an ATP-dependent binding step. The substrate protein then enters the 20S particle.

(3) Proteolysis: → The protein is degraded to the short polypeptides typically 7-9 residues long.

Summary

1. Concepts

Genetic code	Ribosome	Protein folding
Wobble hypothesis	r-protein	Supersecondary Structure
Aminoacyl tRNA synthetase	Intein splicing	Domain
Cognate tRNA	Molecular Chaperone	Quaternary Structure
Supressor tRNA	Ubiquitin	Proteasomes

2. Structure and function of ribosome
3. Protein biosynthesis in prokaryotes (initiation)
4. Protein biosynthesis in eukaryotes (initiation)
5. Functions of molecular chaperone
6. Functions of ubiquitin
7. Structure and function of proteasome
8. ATP-dependent protein degradation

Further reading

Questions and Problems

参考教材：

1.《现代分子生物学》第 5 版，朱玉贤等主编，2019，高教出版社

2.《Molecular Biology》双语教学简明教程，朱国萍、王鹏主编，2010- 2022

国外参考教材：

1.《Genes XII》, Benjamin Lewin, 2018, Oxford University Press

其他参考资料：

1. Biochemistry. Eighth Edition, Jeremy M. Berg, 2015

2. The Cell. Eighth Edition, Geoffrey M. Cooper et al. 2018

3. 分子克隆实验指南(上、下册)，第四版，2017，科学出版社

4. 精编分子生物学实验指南(第五版)，2021，科学出版社