BMS8103

Advanced Cell and Molecular Biology _ Lecture 1

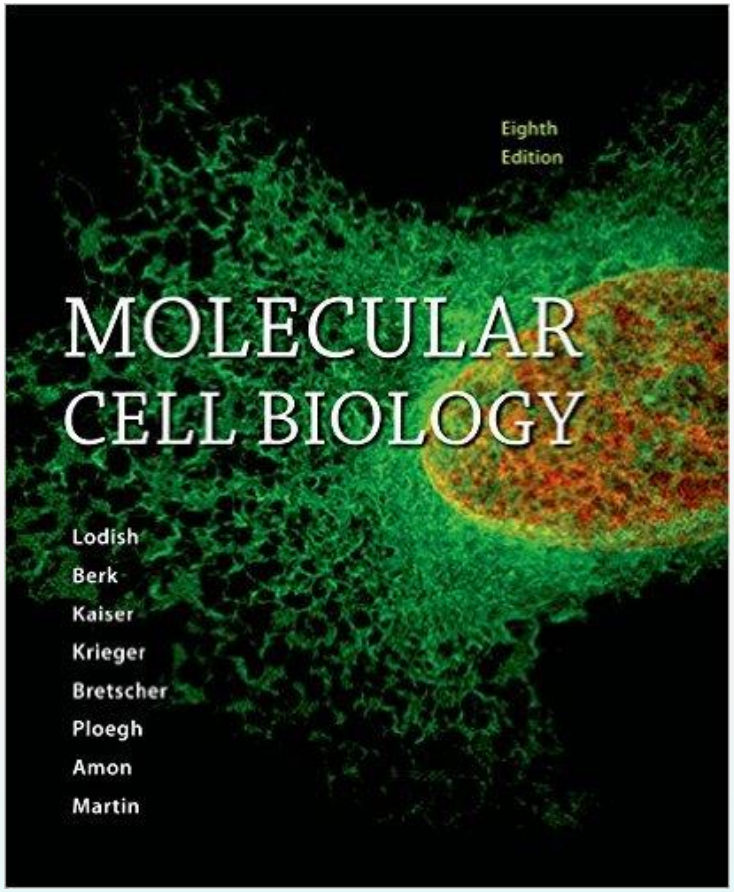
Fundamental Molecular Biology - DNA Replication

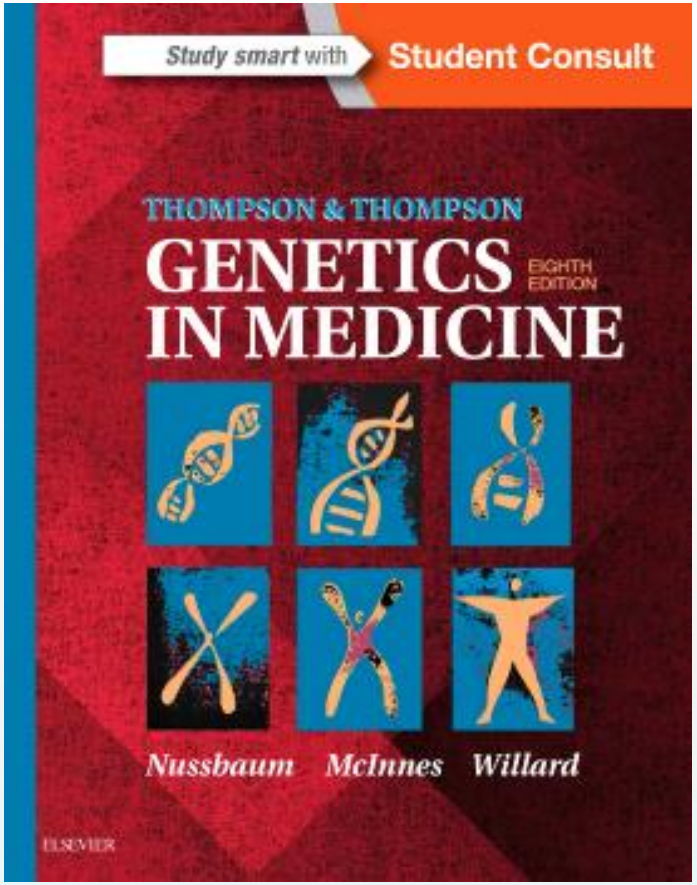
Dr Kui Ming CHAN
Assistant Professor
Department of Biomedical Sciences (BMS)

5 Sept 2019

BMS8103
Advanced Cell and Molecular Biology

Week	Date	Tutorial (1hr) / Lecture (1hr)	Teacher
1	5 Sept	Lecture 1: Fundamental Molecular Biology - DNA Replication and damage repair	Dr CHAN
2	12 Sept	Lecture 2: Molecular Genetics Techniques Student Presentation: DNA Replication	Dr CHAN
3	19 Sept	Lecture 3: Genes, Genomics and Chromosomes Student Presentation: CRISPR/Cas9 mediated Gene Knockout	Dr CHAN
4	26 Sept	Lecture 4: Transcriptional control of Gene expression Student Presentation: Cellular Organization of Genome Function	Dr CHAN
5	3 Oct	Lecture 5: Chromatin and Epigenetic Gene Regulation Student Presentation: Transcription-Replication Conflicts	Dr CHAN
6	10 Oct	Lecture 6: Student Presentation: Epigenetics in Gene regulation and diseases (2 papers for 2 groups)	Dr CHAN





Chapter 5

Molecular Cell Biology 8th edition (Global edition)

Chapter 2

Genetics in Medicine 8th edition

Fundamental Molecular Biology

1.1 Structure of Nucleic Acids (DNA & RNA)

1.2 DNA Replication

1.3 DNA Repair and Recombination

Fundamental Molecular Biology

1.1 Structure of Nucleic Acids

- DNA and RNA structure and function

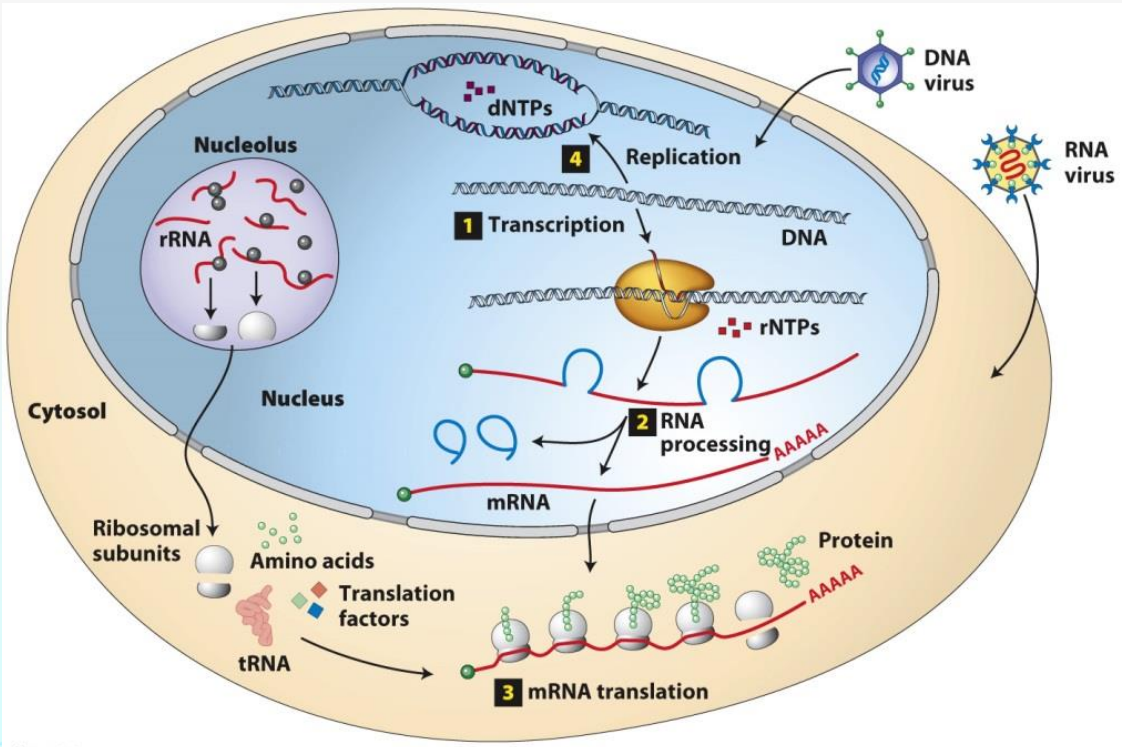


Figure 5-1
Molecular Cell Biology, Eighth Edition

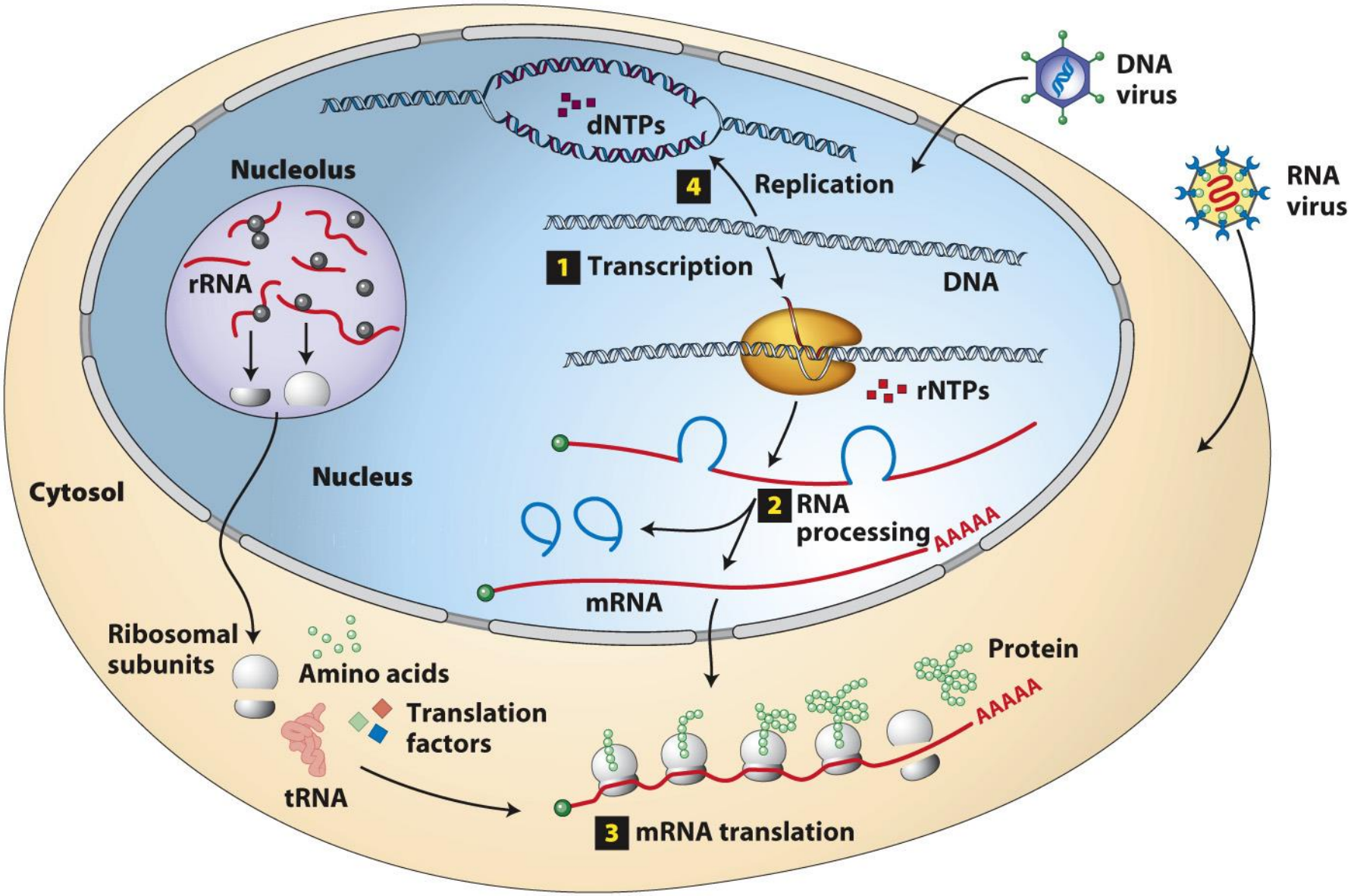
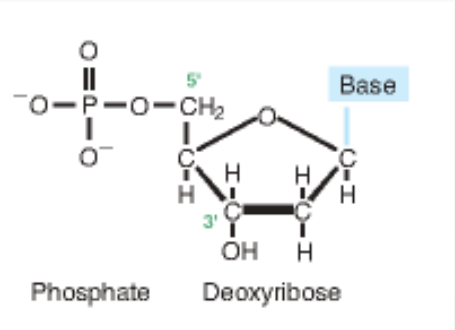


Figure 5-1
Molecular Cell Biology, Eighth Edition

DNA

- Deoxyribonucleic Acid makes up our genome
- Polymeric nucleic acid macromolecule composed of 3 subunits
  - A five-carbon sugar deoxyribose
  - A nitrogen-containing base
  - A phosphate group



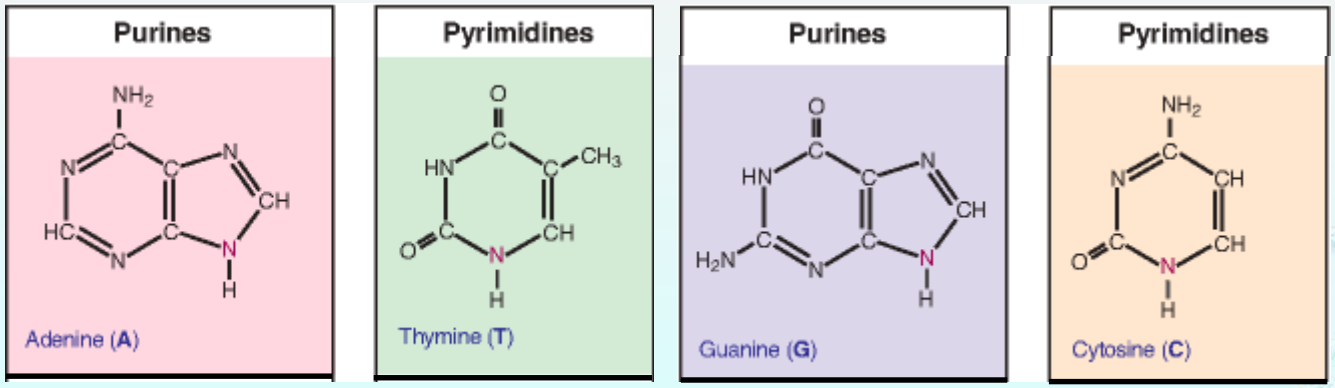


Figure 2-2 The four bases of DNA and the general structure of a nucleotide in DNA. Each of the four bases bonds with deoxyribose (through the nitrogen shown in *magenta*) and a phosphate group to form the corresponding nucleotides.

A & **G** are purines with two fused rings

C & **T** are pyrimidines with one ring only

RNA

**Ribonucleic Acid (RNA) is the intermediate between the genetic code (DNA) and the functional unit (Protein)

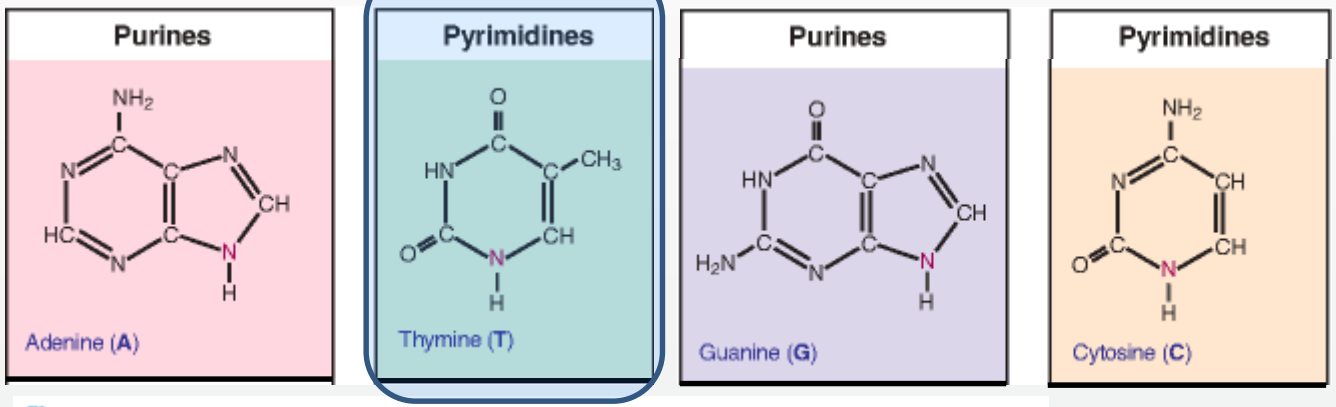
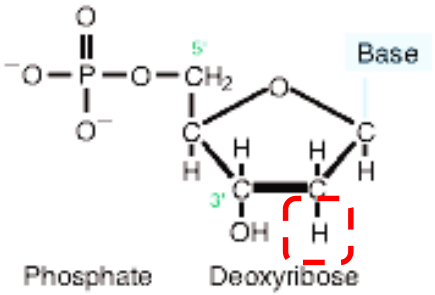


Figure 2-2 The four bases of DNA and the general structure of a nucleotide in DNA. Each of the four bases bonds with deoxyribose (through the nitrogen shown in *magenta*) and a phosphate group to form the corresponding nucleotides.

RNA:
***Ribose**
***Uracil**
***Single or double stranded**



DNA:
***Deoxyribose**
***Thymine**
***Double stranded**

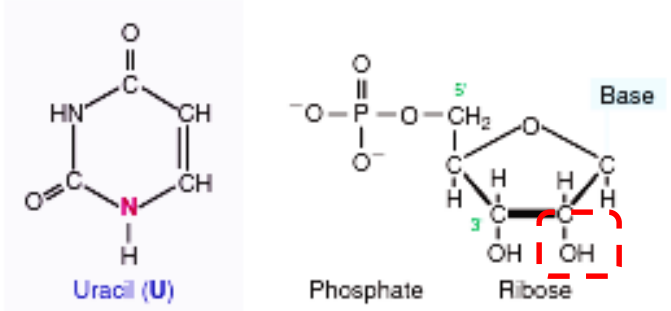


Figure 3-3 The pyrimidine uracil and the structure of a nucleotide in RNA. Note that the sugar ribose replaces the sugar deoxyribose of DNA. Compare with Figure 2-2.

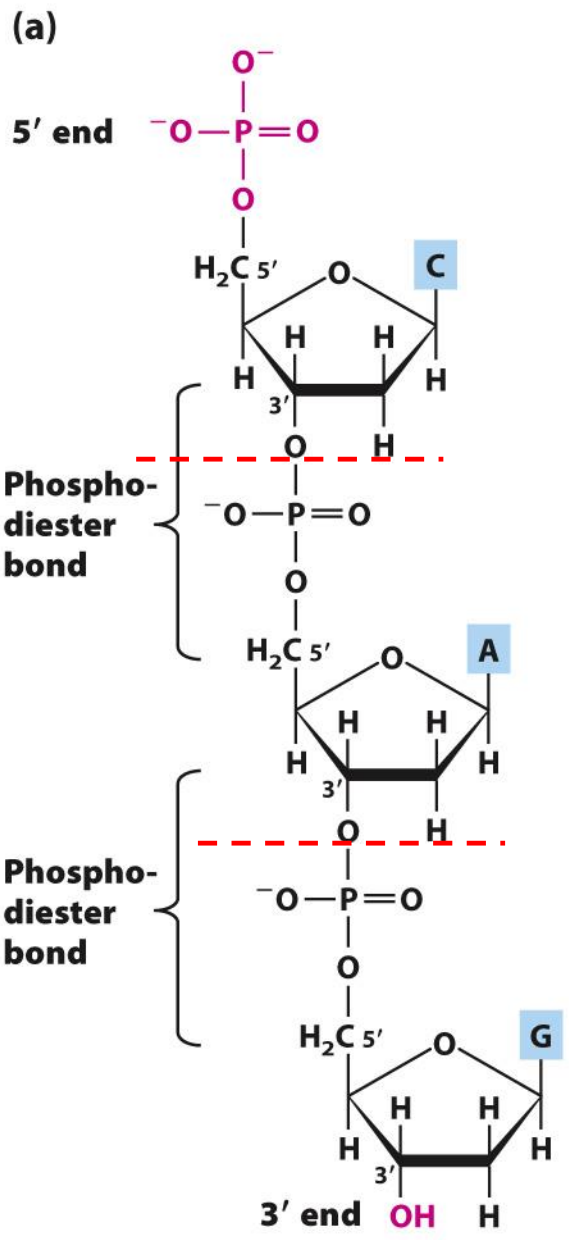
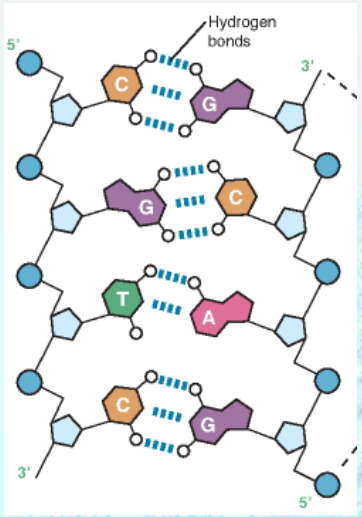


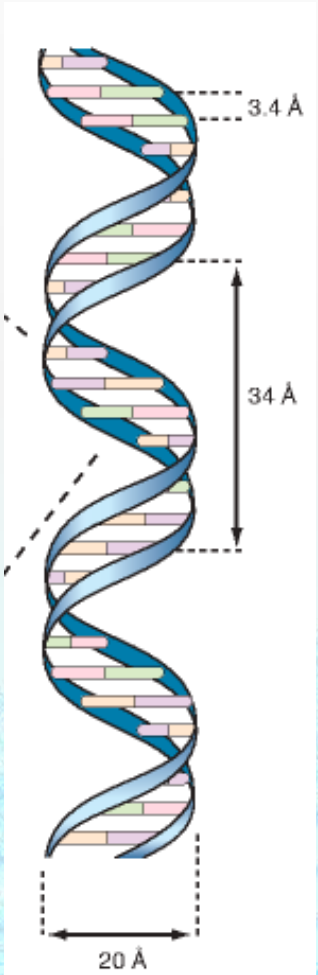
Figure 5-2
Molecular Cell Biology, Eighth Edition

Directionality of DNA:
DNA sequences are written and read in the 5' → 3' direction

DNA double helix

- James Watson and Francis Crick identified the double helix structural property of DNA in 1953
- Right-handed spiral staircase
- <u>Two polynucleotide chains run in opposite directions</u>
- Held together by hydrogen bonds between pairs of bases
  - A pairs with T (2 hydrogen bonds)
  - C pairs with G (3 hydrogen bonds)
  - **Complementary in nature
- → Allows the DNA to be replicated precisely by separation of two strands





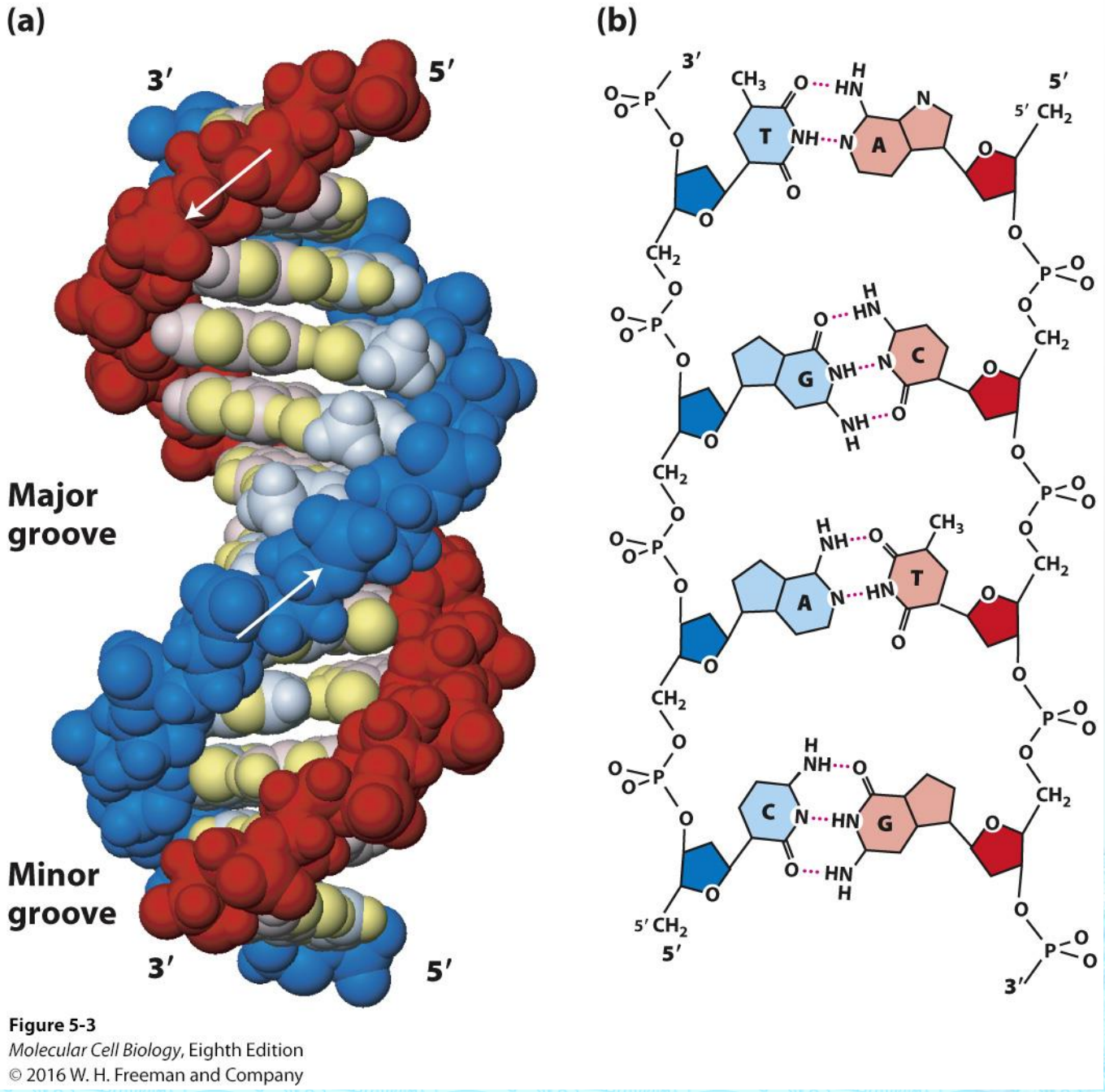
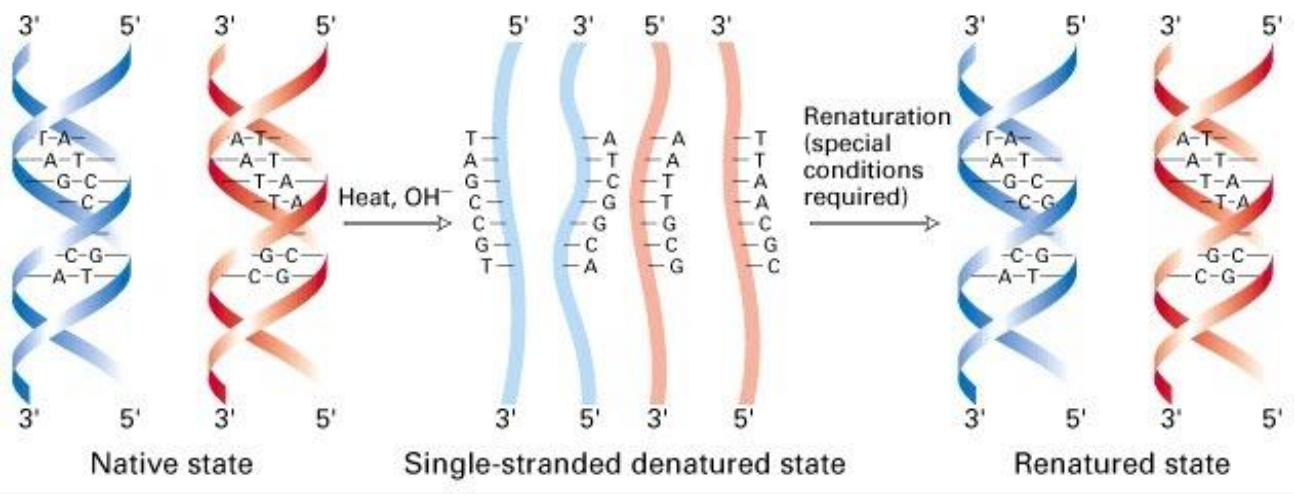


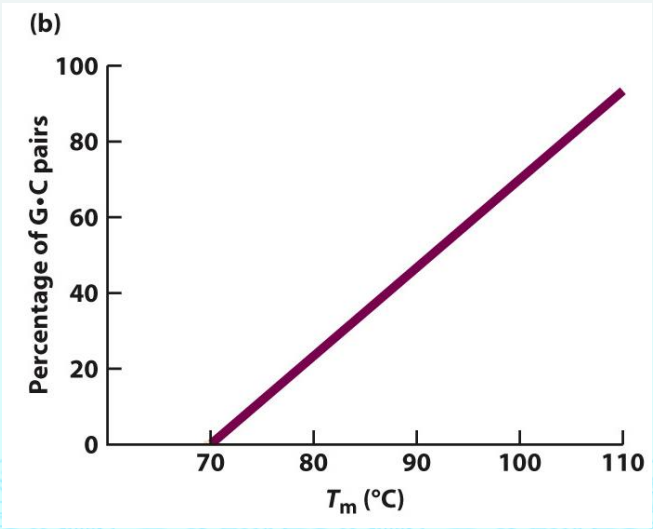
Figure 5-3
Molecular Cell Biology, Eighth Edition

DNA denaturation



http://www.majordifferences.com/2013/02/difference-between-denaturation-and.html

Strand separation of dsDNA is reversible e.g. during replication and transcription



→ By breaking the hydrogen bonds between bases of the two single strands (high temp, pH……)

T_m : melting temperature

→ The temperature at which half the bases in a double stranded DNA have denatured

→ Dependent on CG content

→ Higher CG content, higher T_m

Secondary and Tertiary structures of RNA

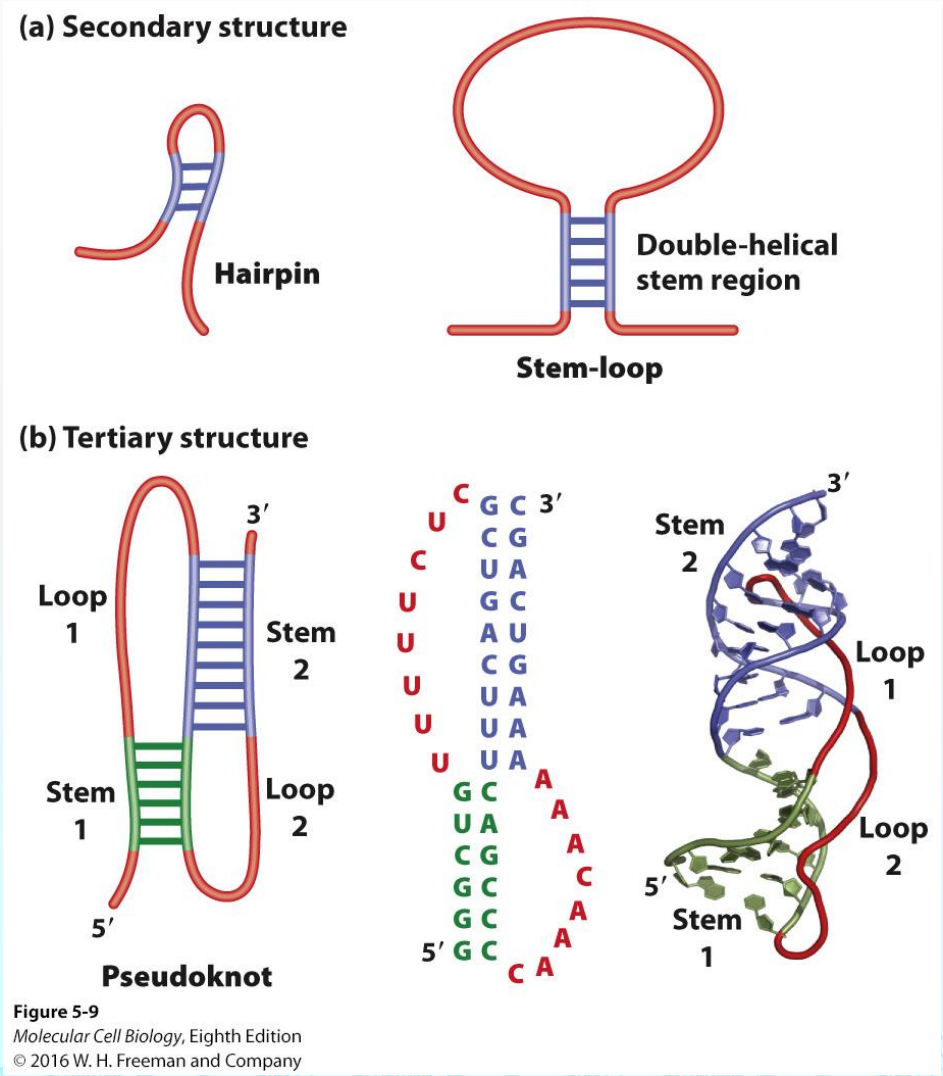


Figure 5-9
Molecular Cell Biology, Eighth Edition

<u>RNA:</u>

- Usually single stranded
- Exhibits a variety of structural conformations
- e.g. by pairing of complementary bases

→ Hairpin structure
  → Pairing of bases within about 5 to 10 nucleotides

→ Stem-loops structure
  → Pairing of bases that are separated by 11 to several hundred nucleotides

Simple folds can cooperate to form more complicated structure (pseudoknot)

Examples of RNA higher order structure
→ tRNAs and rRNAs

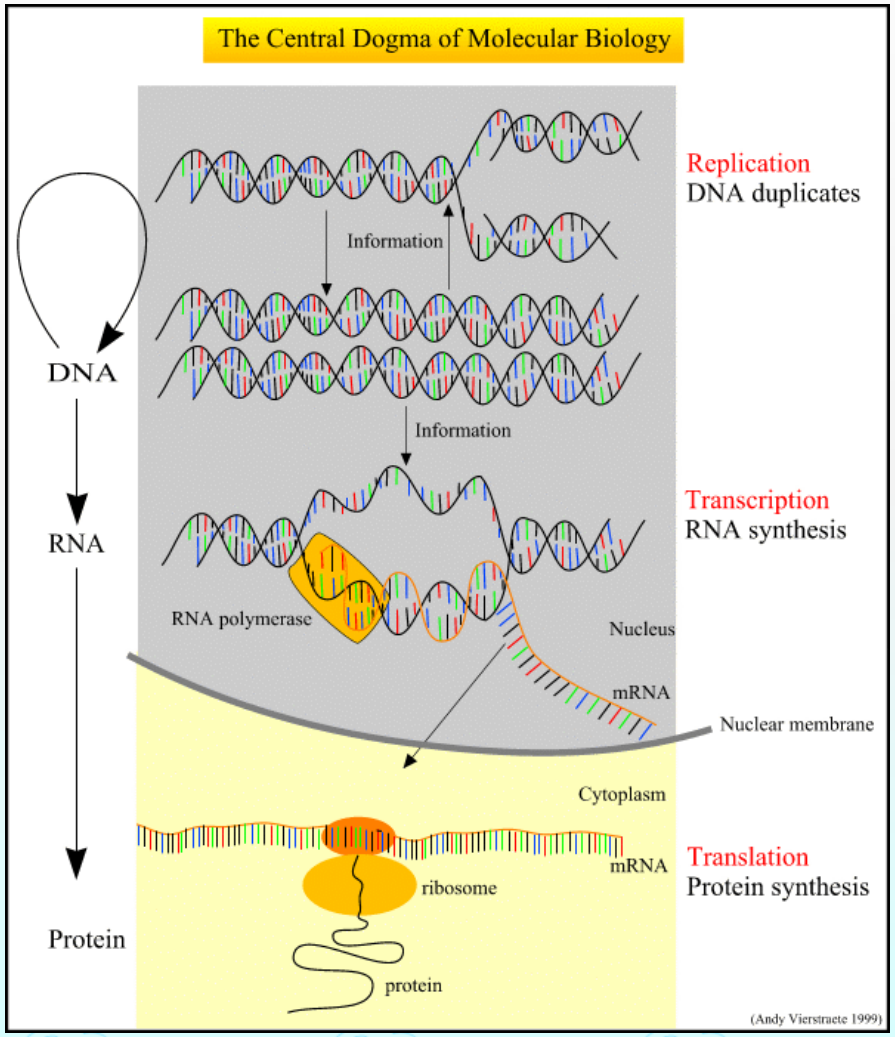
Fundamental Molecular Biology

1.2 DNA Replication

- Semiconservative replication mechanism

Importance of DNA replication??

The central dogma of Molecular Biology



http://users.ugent.be/~avierstr/principles/centraldogma.html

- Genetic info is stored in the DNA by means of a code
- Genetic info is replicated before cell division via DNA replication
- RNA is synthesized from the DNA template via transcription
- RNA carrying the genetic info is called mRNA (messenger RNA)
- mRNA is transported from the nucleus to cytoplasm
- mRNA is decoded/translated to amino acids to form proteins via protein translation by ribosomes

DNA replication – semiconservative

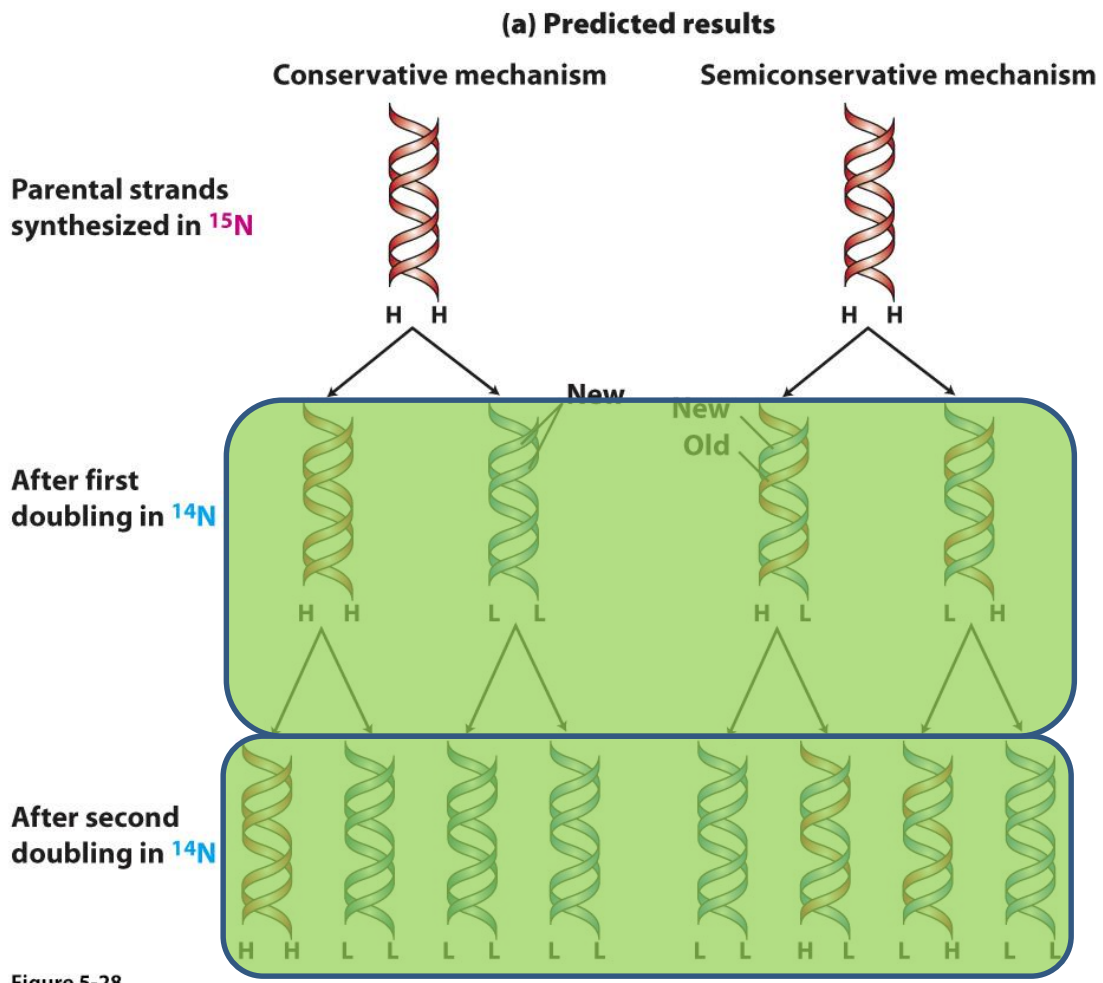


Figure 5-28
Molecular Cell Biology, Eighth Edition

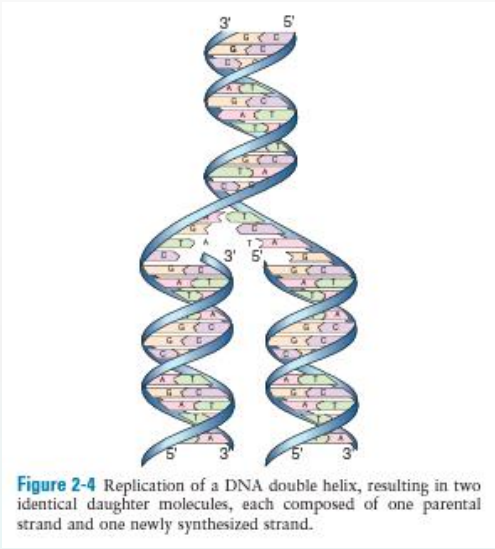
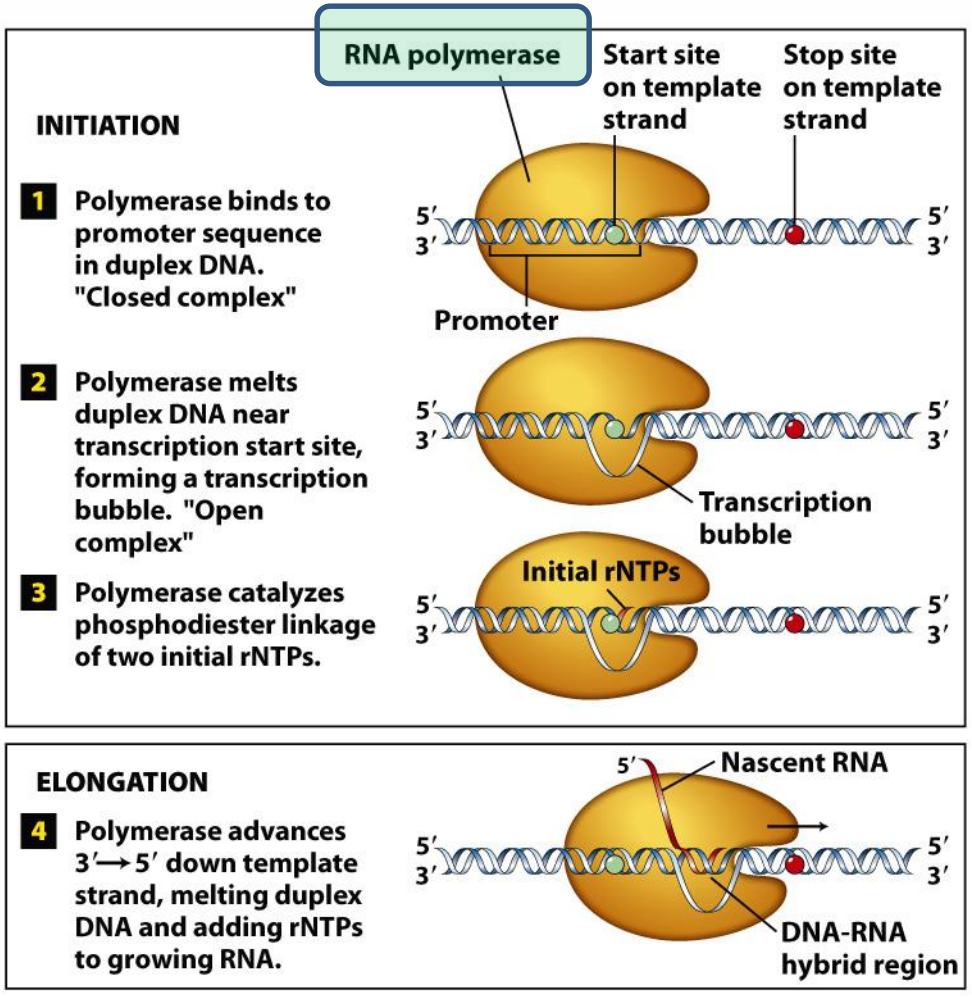
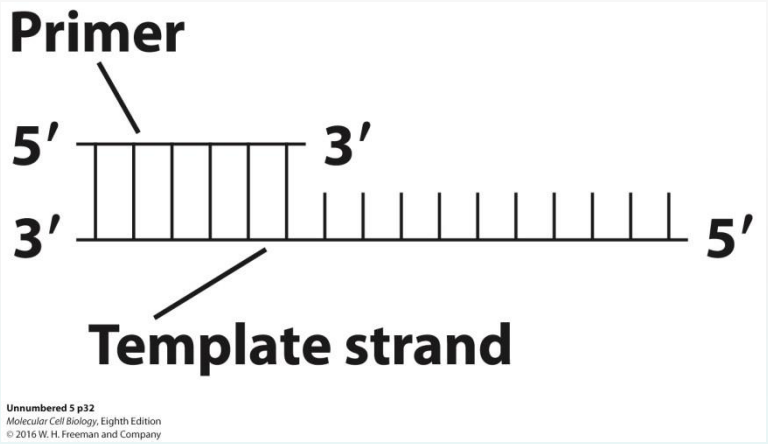


Figure 2-4 Replication of a DNA double helix, resulting in two identical daughter molecules, each composed of one parental strand and one newly synthesized strand.

Transcription →

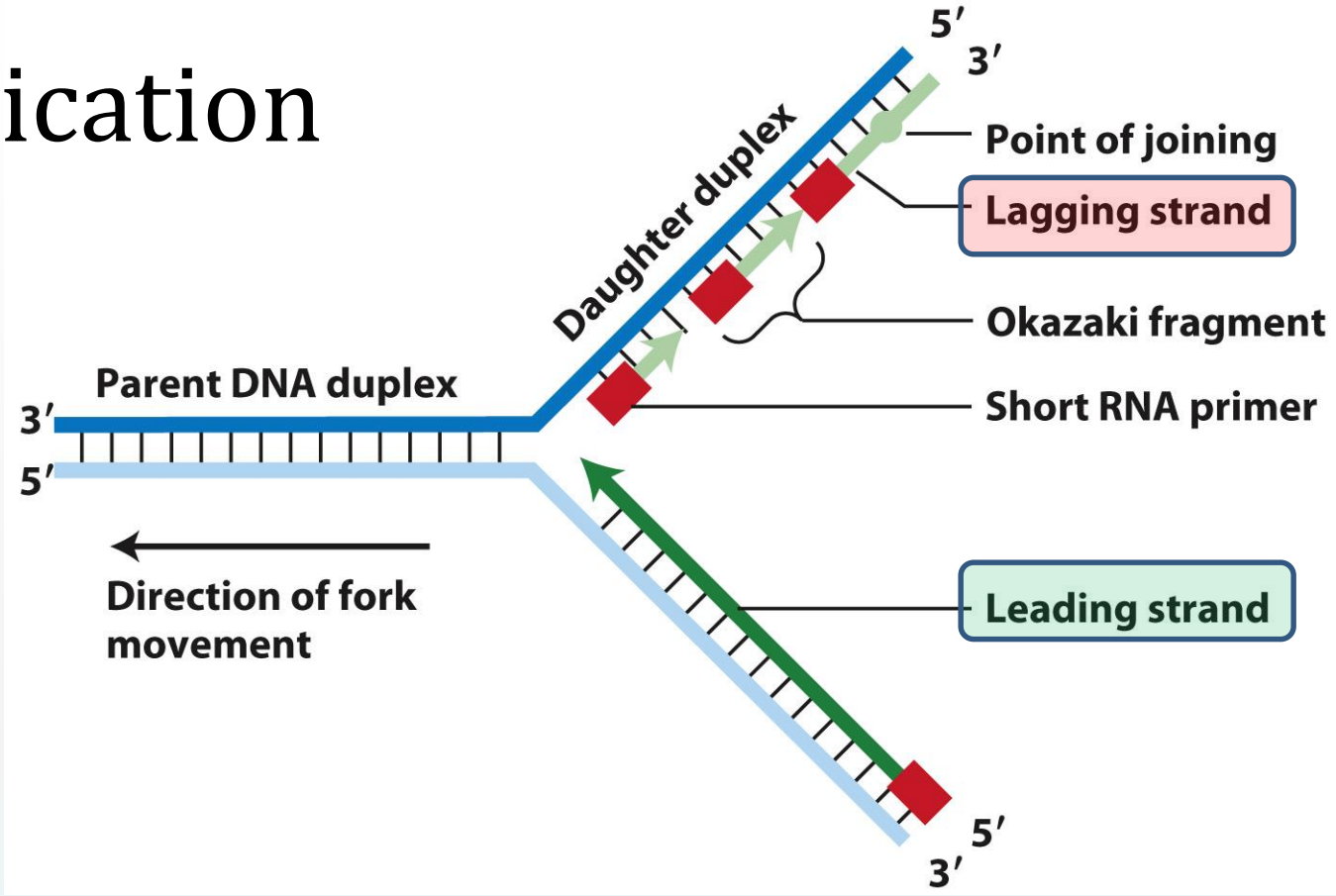


DNA replication



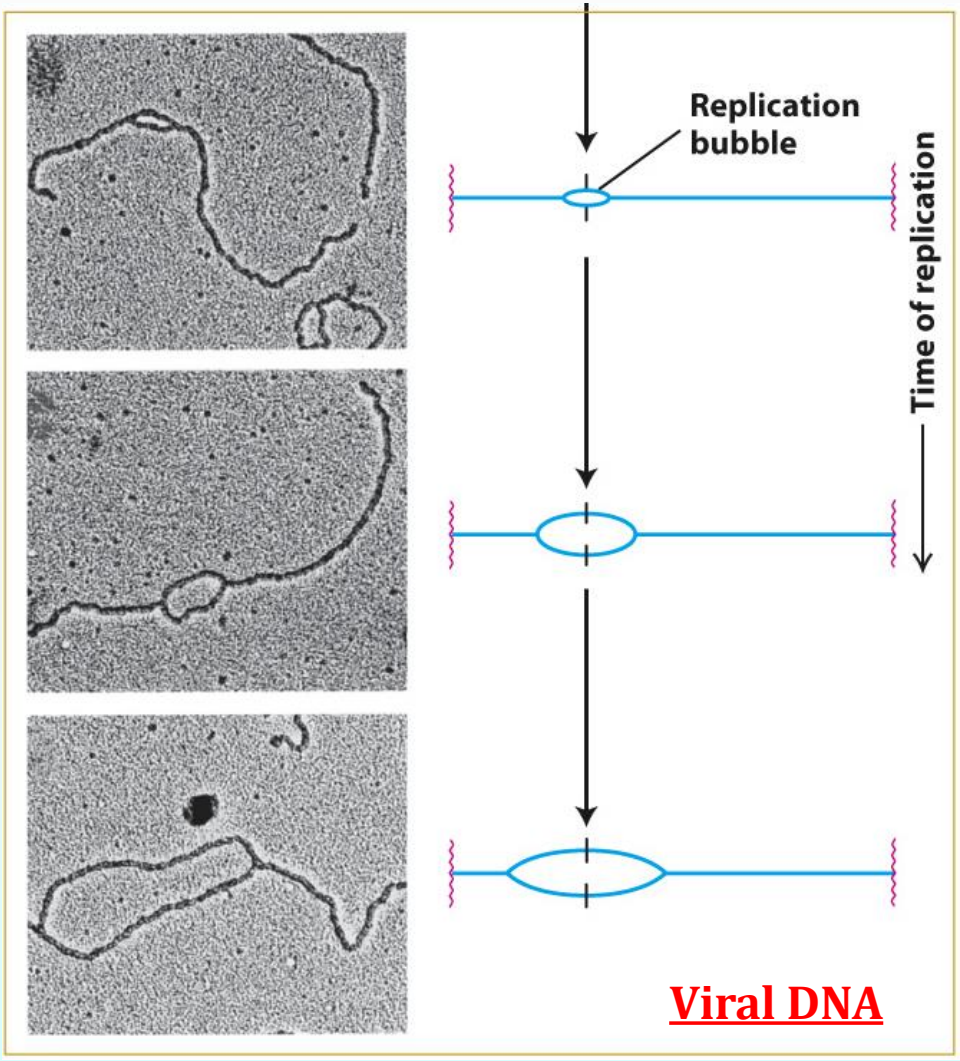
- DNA synthesis always proceeds in the 5′→3′ direction
- Unlike RNA polymerase, DNA polymerases CANNOT initiate chain synthesis *de novo*
- It requires a short, preexisting RNA or DNA primer strand (with free 3' hydroxyl) that is base-paired to the template strand to begin second strand growth.

DNA replication



- Duplex DNA is unwound, and daughter strands are formed at the DNA replication fork.
- DNA polymerase adds nucleotides to a growing daughter strand in the 5′→3′ direction.
- Leading strand: synthesized continuously from a single RNA primer (red) at its 5′ end.
- Lagging strand:
  - synthesized discontinuously from multiple RNA primers.
  - Elongation of lagging strand primers initially produces Okazaki fragments.
  - Growing Okazaki fragments displace the previous primer, and the elongated fragments are ligated into a continuous strand.

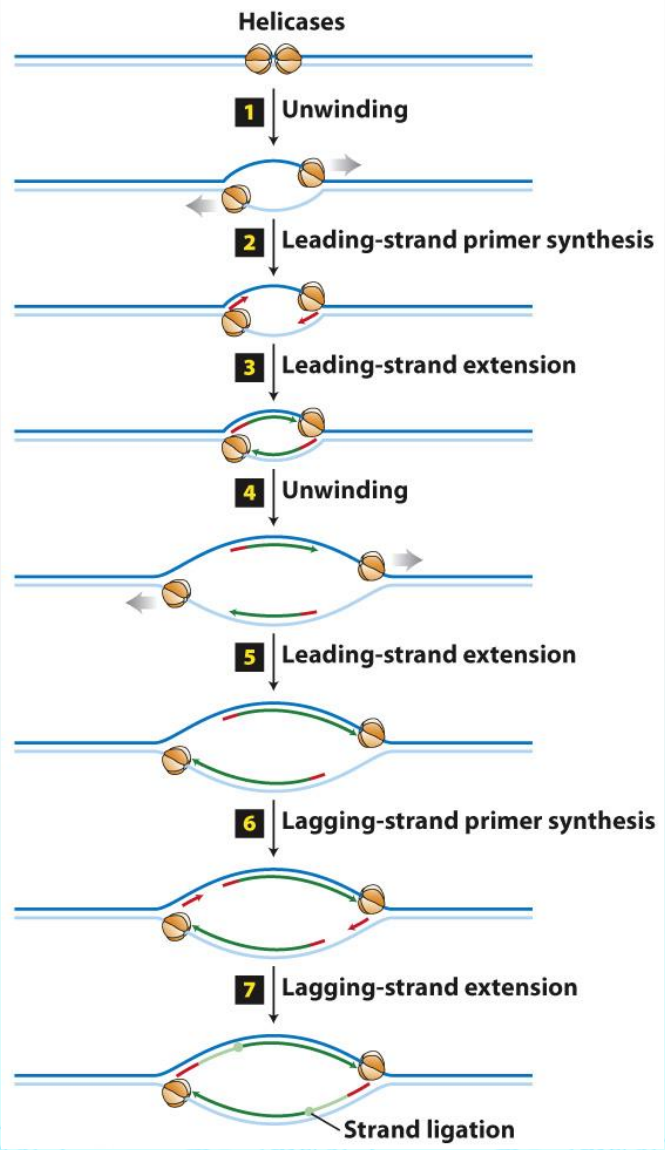
DNA replication



Bidirectional replication of DNA

- All bacterial, archaeal, and eukaryotic cells replicate DNA bidirectionally.
- Replication is initiated by binding two large T-antigen hexameric helicases to the single SV40 origin and formation of two oppositely oriented replication forks.

DNA replication



REVIEWS

DNA replication origin activation in space and time

Michalis Fragkos, Olivier Ganier, Philippe Coulombe* and Marcel Méchali*

Abstract | DNA replication begins with the assembly of pre-replication complexes (pre-RCs) at thousands of DNA replication origins during the G1 phase of the cell cycle. At the G1–S-phase transition, pre-RCs are converted into pre-initiation complexes, in which the replicative helicase is activated, leading to DNA unwinding and initiation of DNA synthesis. However, only a subset of origins are activated during any S phase. Recent insights into the mechanisms underlying this choice reveal how flexibility in origin usage and temporal activation are linked to chromosome structure and organization, cell growth and differentiation, and replication stress.

ARTICLE

doi:10.1038/nature14285

Regulated eukaryotic DNA replication origin firing with purified proteins

Joseph T. P. Yeeles[1], Tom D. Deegan[1], Agnieszka Janska[1], Anne Early[1] & John F. X. Diffley[1]

Eukaryotic cells initiate DNA replication from multiple origins, which must be tightly regulated to promote precise genome duplication in every cell cycle. To accomplish this, initiation is partitioned into two temporally discrete steps: a double hexameric minichromosome maintenance (MCM) complex is first loaded at replication origins during G1 phase, and then converted to the active CMG (Cdc45–MCM–GINS) helicase during S phase. Here we describe the reconstitution of budding yeast DNA replication initiation with 16 purified replication factors, made from 42 polypeptides. Origin-dependent initiation recapitulates regulation seen *in vivo*. Cyclin-dependent kinase (CDK) inhibits MCM loading by phosphorylating the origin recognition complex (ORC) and promotes CMG formation by phosphorylating Sld2 and Sld3. Dbf4-dependent kinase (DDK) promotes replication by phosphorylating MCM, and can act either before or after CDK. These experiments define the minimum complement of proteins, protein kinase substrates and co-factors required for regulated eukaryotic DNA replication.

DNA replication errors

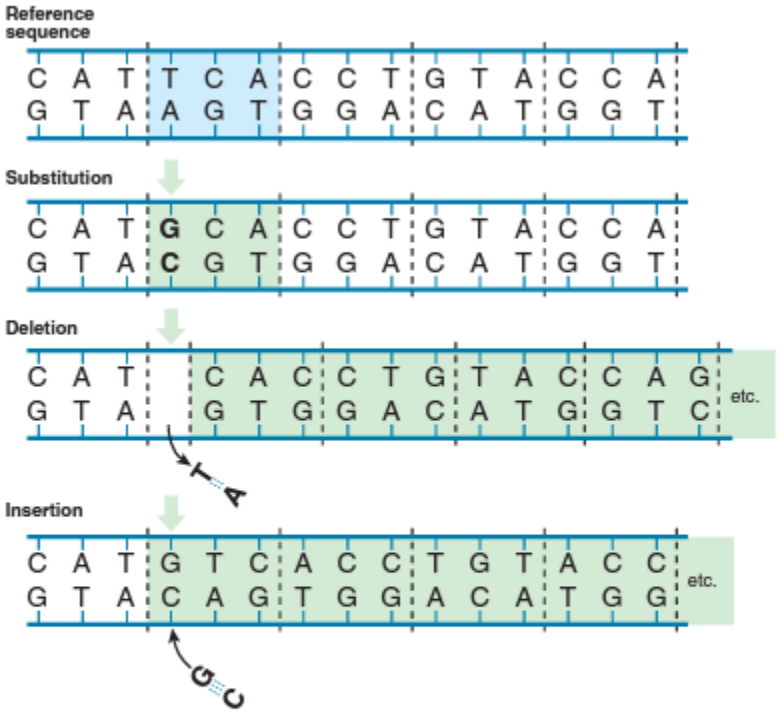


Figure 4-4 Examples of mutations in a portion of a hypothetical gene with five codons shown (delimited by the *dotted lines*). The first base pair of the second codon in the reference sequence (shaded in *blue*) is mutated by a base substitution, deletion, or insertion. The base substitution of a G for the T at this position leads to a codon change (shaded in *green*) and, assuming that the upper strand is the sense or coding strand, a predicted nonsynonymous change from a serine to an alanine in the encoded protein (see genetic code in Table 3-1); all other codons remain unchanged. Both the single base pair deletion and insertion lead to a frameshift mutation in which the translational reading frame is altered for all subsequent codons (shaded in *green*), until a termination codon is reached.

- DNA replication is typically highly accurate
- The majority of replication errors are rapidly removed from the DNA
- Subsequently corrected by a series of DNA repair enzymes
  - Recognize the error, replace with the proper nucleotide
  - "DNA proofreading"
- DNA polymerase introduces 1 error per 10 million bp
- "Proofreading" correct more than 99.9% of the errors

- Overall mutation rate per base pair as a result of replication errors is very low : 1×10^{-10} per cell division, less that one mutation per genome per cell division

“Proofreading” of DNA polymerase

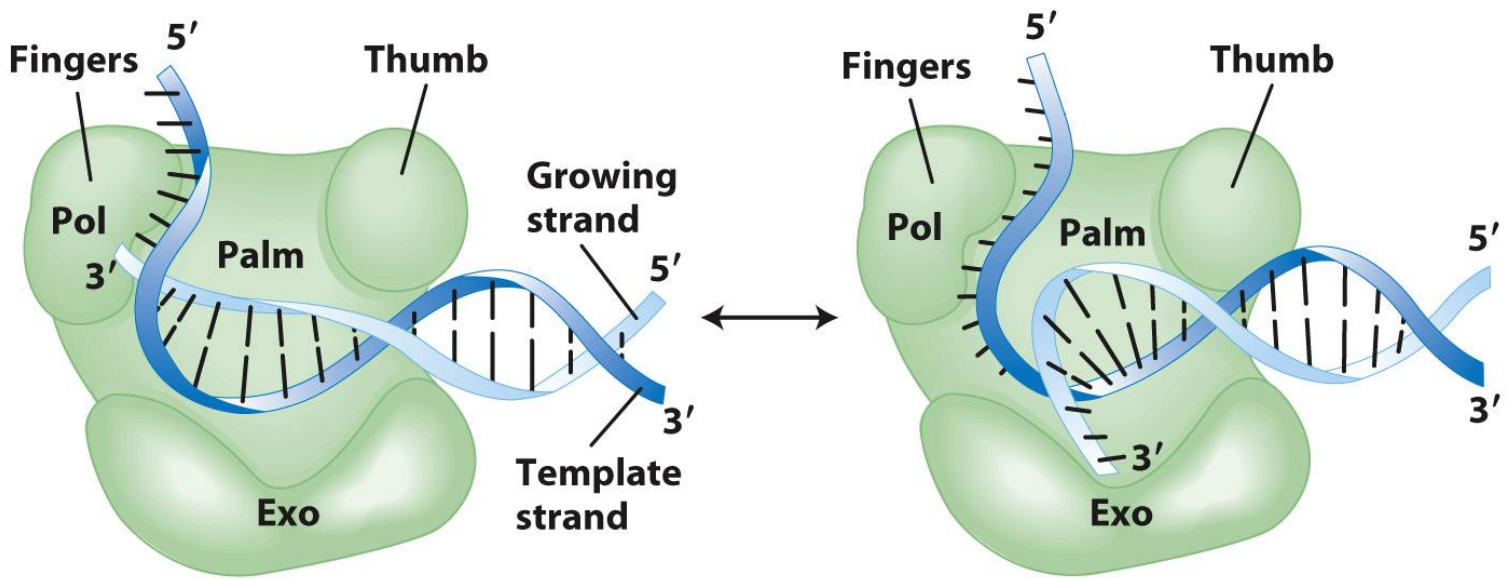


Figure 5-33
Molecular Cell Biology, Eighth Edition

- DNA polymerases have a 3D structure resembling a half-opened right hand. The “fingers” bind the single-stranded segment of the template strand, and the polymerase catalytic activity (Pol) lies in the junction between the fingers and palm.
- Incorrect base addition at the 3′ end causes melting of the newly formed end of the duplex and polymerase pausing.
- The growing strand 3′ end is transferred to the 3′ → 5′ exonuclease site (Exo), where the mis-paired base(s) are removed.
- The 3′ end flips back into the polymerase site and elongation resumes.

Repair damaged DNA

- ** ~10,000 to 1 M nucleotides are damaged per human cell per day by spontaneous chemical processes
  - E.g. deamination, de-purination, demethylation and
  - By interaction with chemical mutagens from the environment
  - Or by exposure to UV or ionizing radiation
- Some but not all of these damages are repaired
- Of those repaired, mutations may be introduced during the repair process by repair proteins/enzymes
- E.g. Substitution of T for C is a very common spontaneous mutation

Fundamental Molecular Biology

1.3 DNA Repair and Recombination

- DNA sequence changes – copying errors and the effects of various physical and chemical agents.
- DNA repair systems:
  - Base excision repair
  - Mismatch excision repair
  - Nucleotide excision repair
  - Non-homologous end joining
  - Homologous recombination
  - Repair of a collapse replication fork
- Defects in DNA repair are associated with cancers

Deamination leads to point mutations

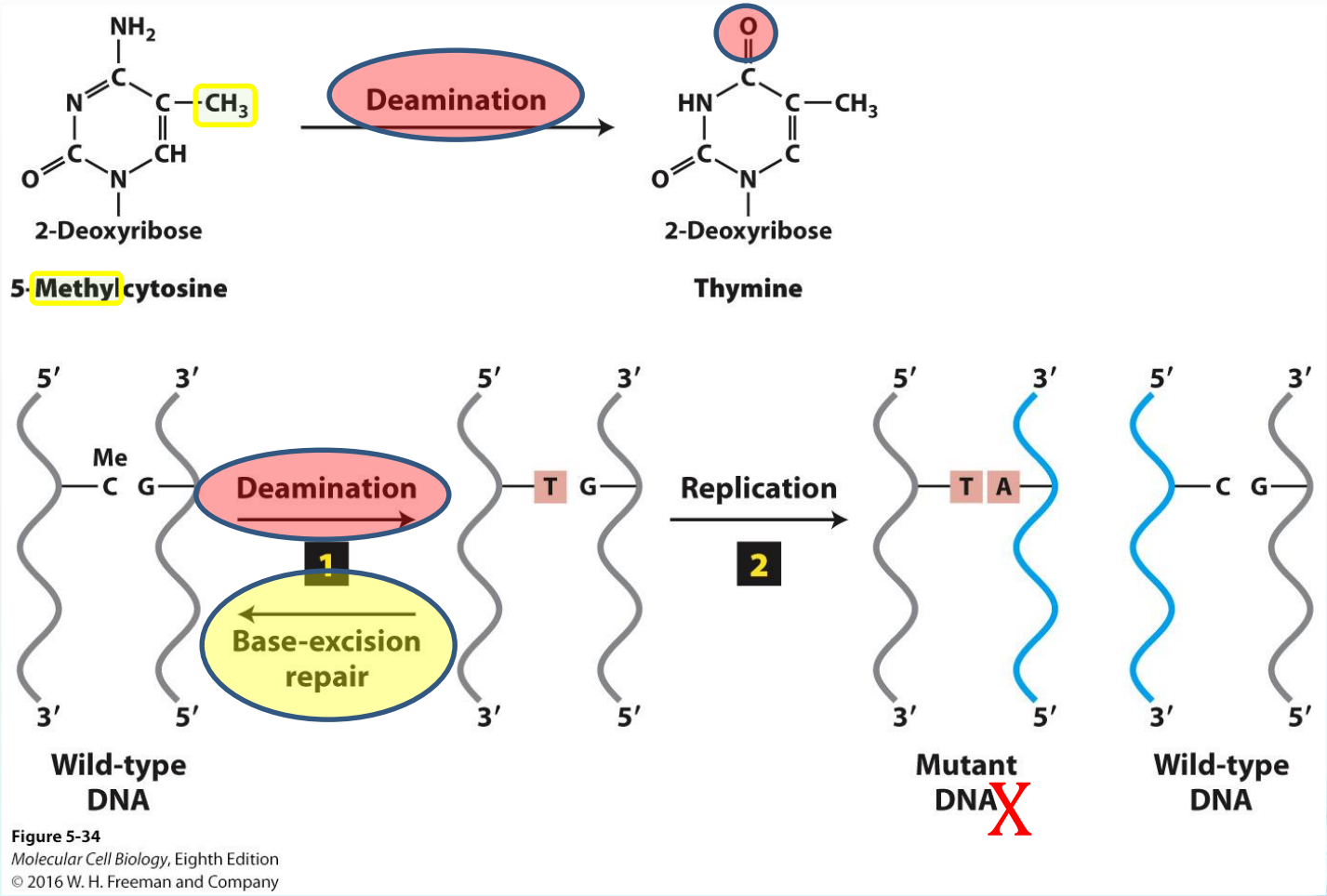


Figure 5-34
Molecular Cell Biology, Eighth Edition

**5-methylcytosine (C) deamination to thymine (T) is a frequent cause of point mutations; if C is not restored, the normal C·G base pair will become a T-A mutation during replication of one of the DNA strands.

Base excision repair of a T*G mismatch

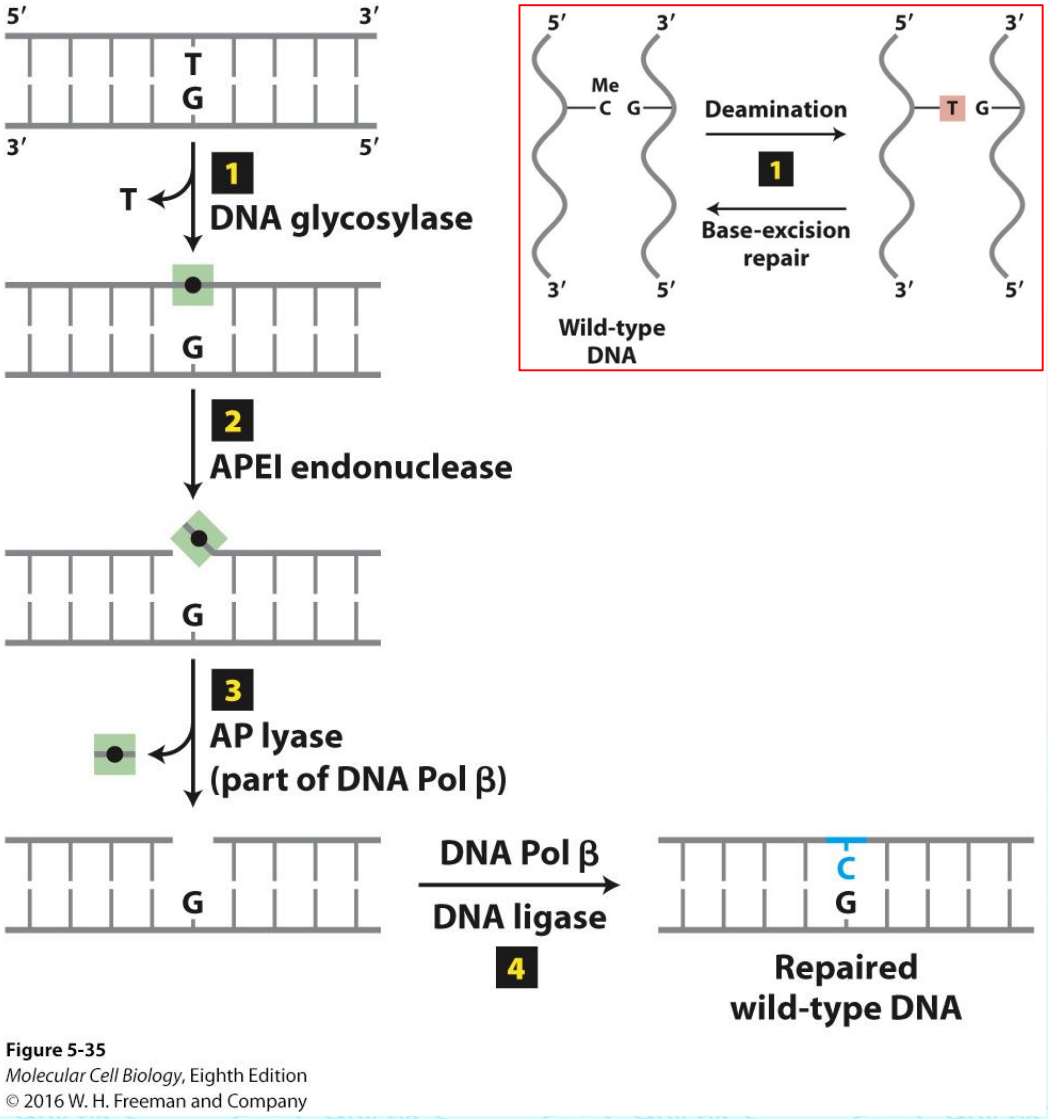


Figure 5-35
Molecular Cell Biology, Eighth Edition

High-fidelity DNA excision-repair systems recognize and repair damage

→repair system evolved to remove a mismatched T and replace it with a C

- Step 1: DNA glycosylase specific for G·T mismatches flips the thymine base out of the helix and cuts the base off the sugar-phosphate DNA backbone, leaving just the deoxyribose phosphate (black dot).
- Step 2: Apurinic endonuclease I (APE1) cuts the DNA backbone at the abasic site.
- Step 3: Apurinic lyase (AP lyase) associated with DNA polymerase β (a specialized DNA polymerase used in repair) removes the deoxyribose phosphate.
- Step 4: DNA Pol β inserts the single-base and DNA ligase links it to the backbone.

Mismatch excision repair in eukaryotes

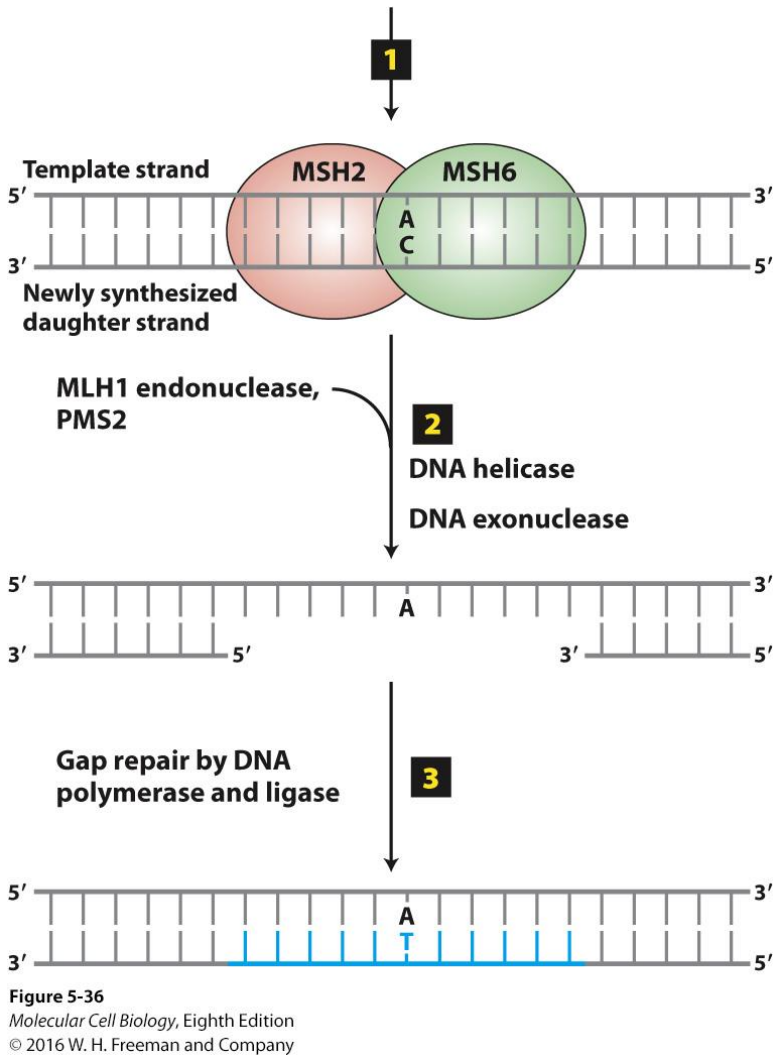


Figure 5-36
Molecular Cell Biology, Eighth Edition

- A mismatch excision-repair pathway corrects errors introduced during replication.
  - Step 1: A MSH2 and MSH6 protein complex binds to a mis-paired segment of DNA and distinguishes between the template and the newly synthesized daughter strand.
  - Step 2: Recruit other proteins MLH1 and PMS2.
    - The DNA-protein complex binds an endonuclease that cuts the newly synthesized daughter strand.
    - DNA helicase unwinds the helix.
    - An exonuclease removes several nucleotides from the cut end of the daughter strand, including the mismatched base.
  - Step 3: DNA polymerase Pol δ fills the gap DNA ligase connects the new section of DNA to the backbone.

Formation of thymine-thymine dimers

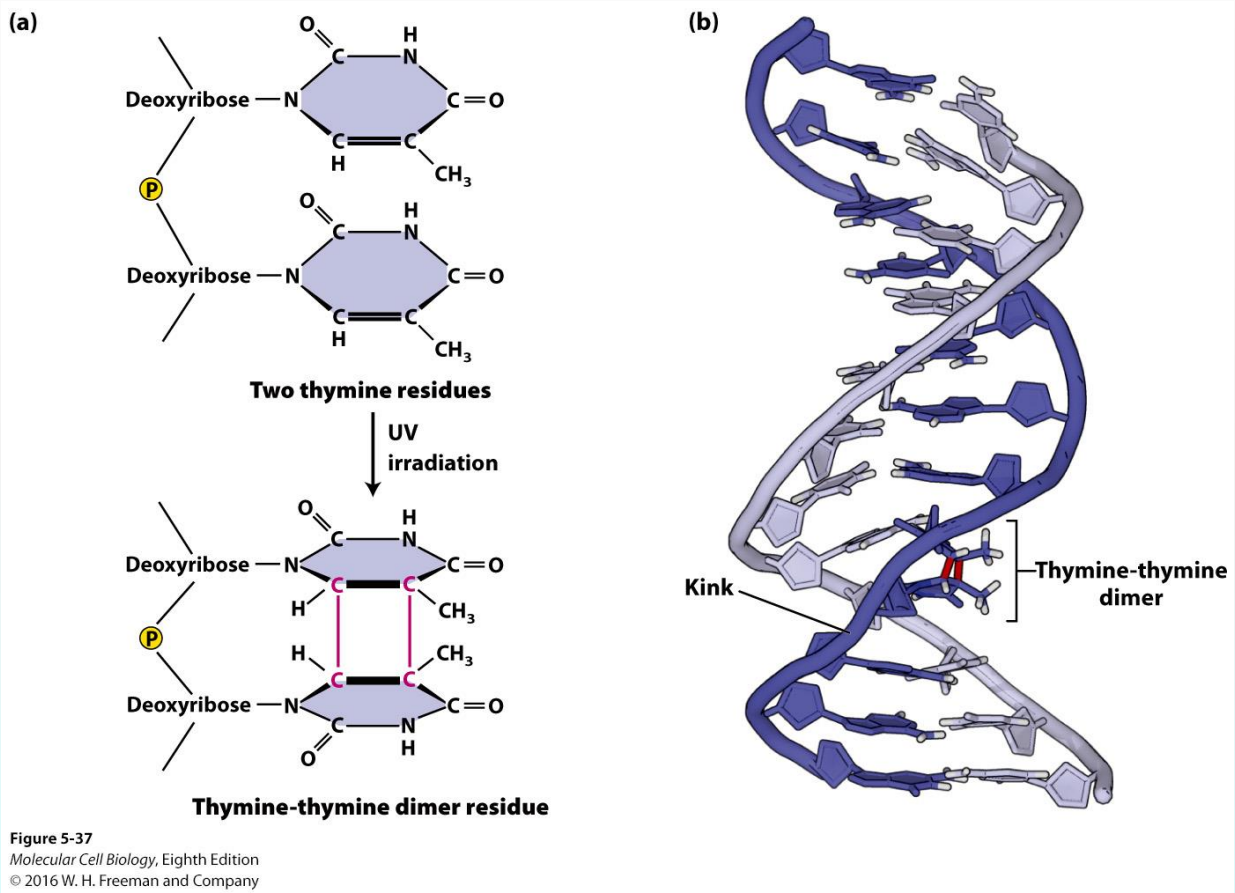
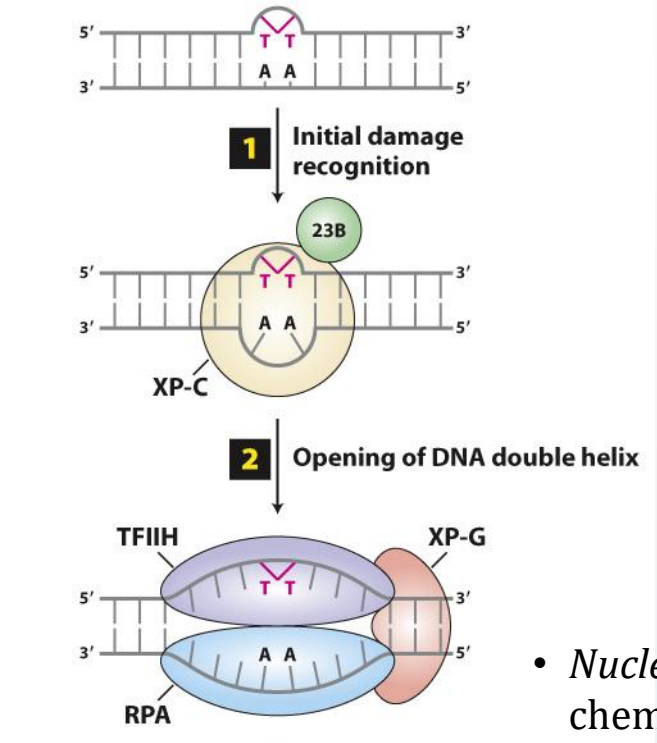


Figure 5-37
Molecular Cell Biology, Eighth Edition

- <u>UV irradiation</u> causes formation of C-C bonds (red lines) between adjacent thymines in DNA (thymine-thymine dimer).
- T-T dimers interfere with replication and transcription.

Nucleotide excision repair in eukaryotes



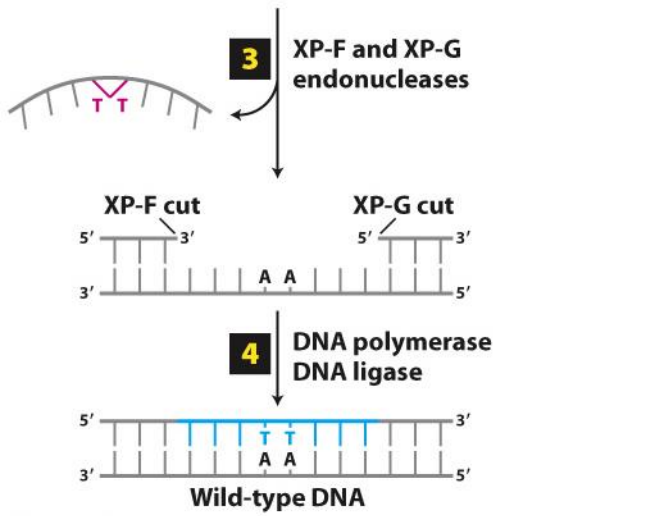


Figure 5-38
Molecular Cell Biology, Eighth Edition

- *Nucleotide excision repair* mechanisms fix DNA regions containing chemically modified bases (*chemical adducts*) that distort the normal shape of DNA locally.

Step 1: An 23B protein complex recognizes double helix distortion, such as that caused by a T-T dimer.
Step 2: The 23B protein complex recruits transcription factor TFIIH, whose ATP-powered helicase subunits partially unwind the double helix. XP-G and RPA proteins bind and further unwind and stabilize the helix, forming a bubble of about 25 bases.
Step 3: Endonucleases XP-F and XP-G (now acting as an endonuclease) cut out the damaged strand at points 24–32 bases apart on each side of the lesion. (The damaged DNA fragment is degraded to mononucleotides.)
Step 4: The gap is filled by DNA polymerase as in DNA replication, and the remaining nick is sealed by DNA ligase.

Non-Homologous End Joining (NHEJ)

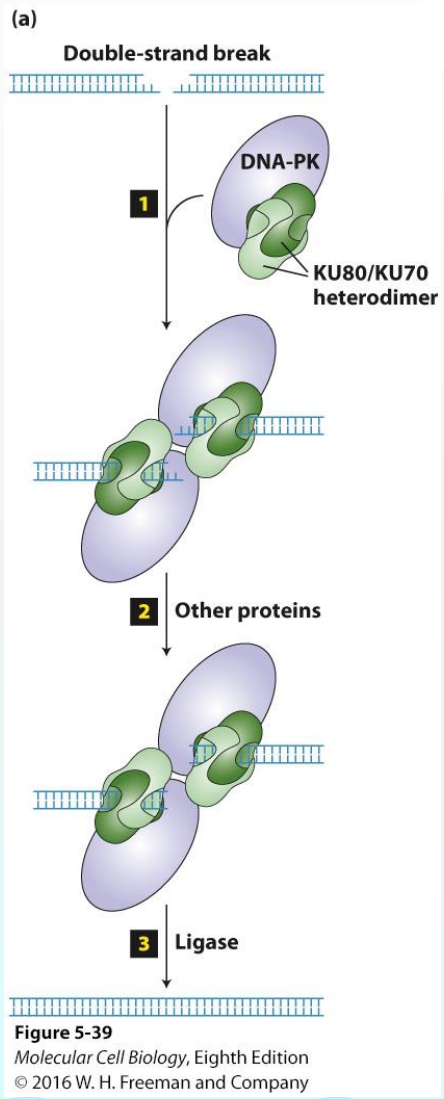
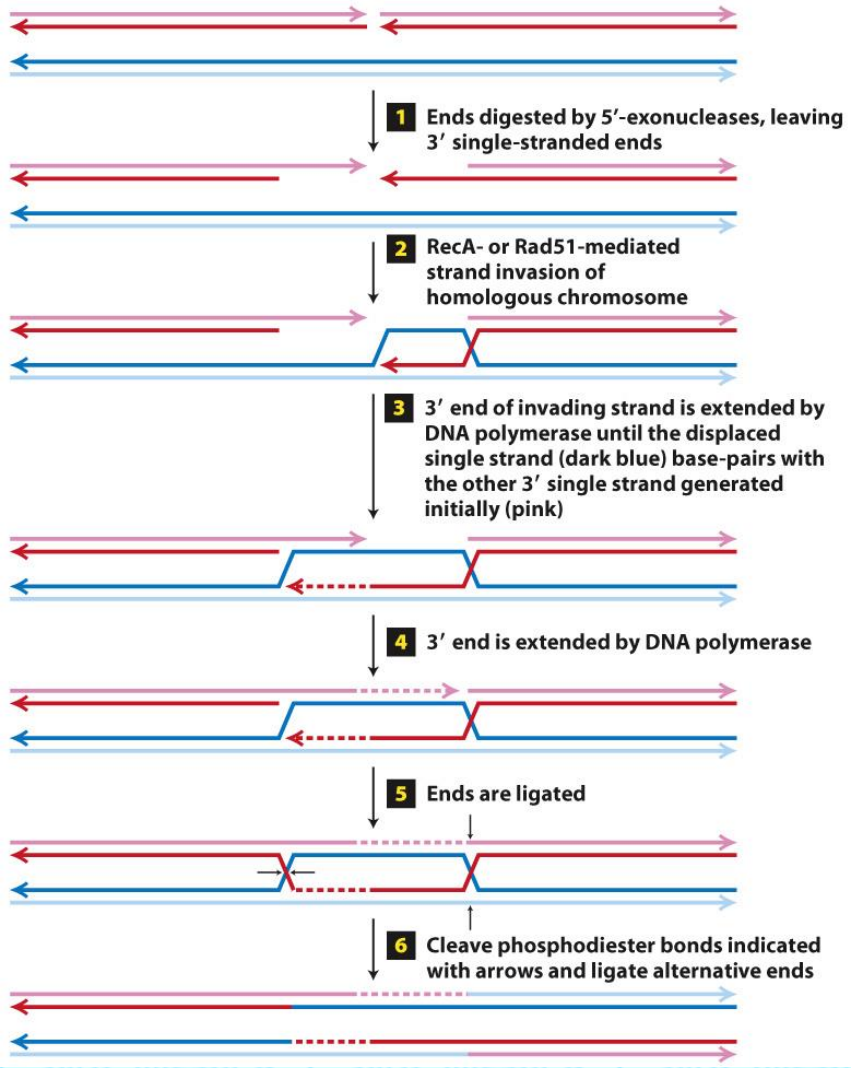


Figure 5-39
Molecular Cell Biology, Eighth Edition

- *Nonhomologous end joining* (*NHEJ*) is one of two systems using recombination to repair DNA double-strand breaks, which can be caused by ionizing radiation and anticancer drugs.
- When sister chromatids are not available to help repair double-strand breaks:
  - Step 1: A Ku-DNA-dependent protein kinase (DNAPK) complex binds to the ends of a double-strand break forming a "synapse".
  - Step 2: In the synapse, nucleases remove bases from the DNA ends.
  - Step 3: The two double-stranded molecules are ligated together, with several base pairs missing.
- NHEJ is error prone: base loss causes gene mutation, and incorrect rejoining can cause gross chromosomal rearrangement mutations that affect the expression of genes or create a "hybrid" gene that encodes the N-terminal portion of one amino acid sequence fused to the C-terminal portion of a completely different protein.

Homologous Recombination (HR)

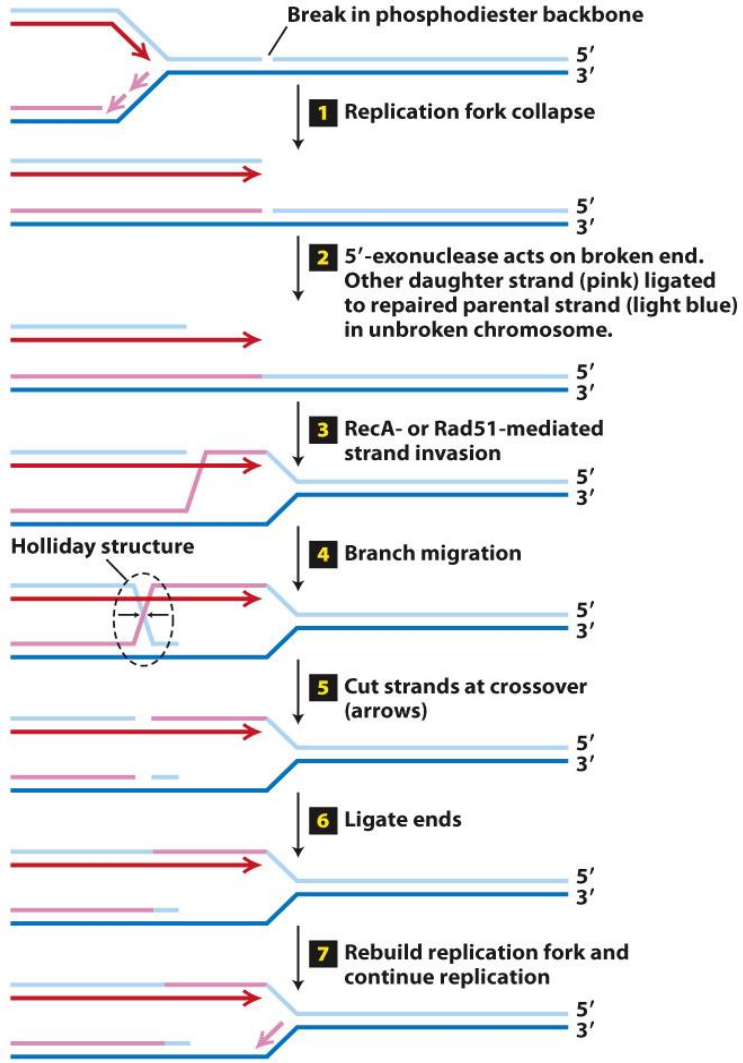
- *Homologous recombination* can repair a double-strand break in a chromosome or exchange large segments of two double-stranded DNA molecules.



- Step 1: 5′-exonucleases digest both broken ends, leaving each with a single-stranded region of DNA with a 3′ end.
- Step 2: RecA (bacteria)/Rad51 (eukaryotes) catalyzes invasion of one 3′ end into the homologous region of the homologous chromosome.
- Step 3: DNA polymerase extends the 3′ end of the invading DNA strand, displacing the parent strand as an enlarging single-stranded loop of DNA (dark blue).
- Step 4: When the loop extends to a sequence that is complementary to the other 5′-exonuclease–digested end of DNA, the complementary sequences base-pair, and DNA polymerase extends this 3′ end, using the displaced single-stranded loop of parent DNA (dark blue) as a template.
- Step 5: The new 3′ ends are ligated to the exonuclease-digested 5′ ends, generating two Holliday structures in the paired molecules. Holliday structure branch migration can occur in either direction (not shown).
- Step 6: Cleavage of the strands and ligation of the alternative 5′ and 3′ ends at each cleaved Holliday structure generates two *recombinant* chromosomes that contain the DNA of one *parent* DNA molecule on one side of the initial break point (pink and red strands) and the other parent DNA molecule on the other side of the break point (light and dark blue).

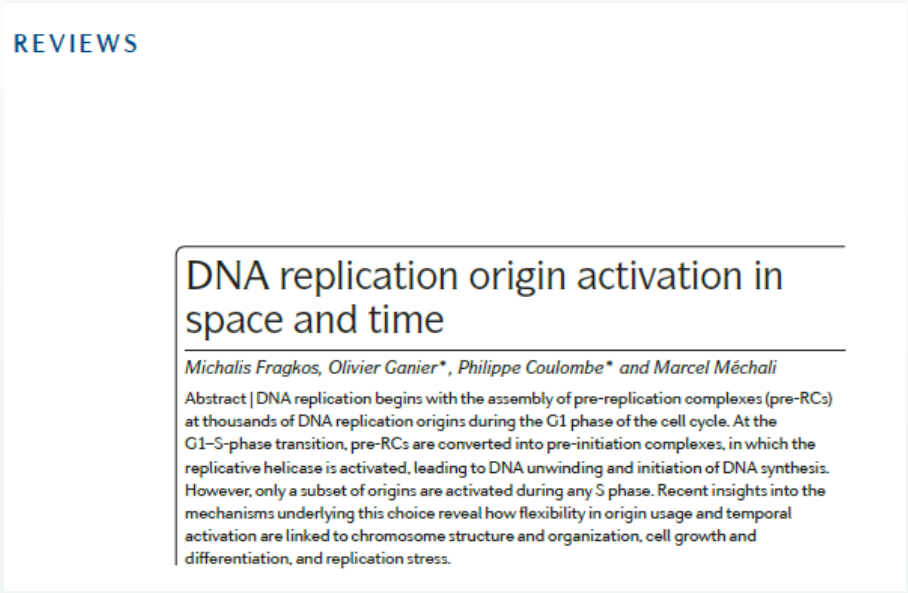
Repair of a collapse replication fork

- Homologous recombination can repair DNA damage and generate genetic diversity by causing the exchange of large regions of chromosomes between the maternal and paternal pair of homologous chromosomes during *meiosis*.



- Step 1: A nick in one strand causes *replication fork collapse.*
- Step 2: 5’ exonuclease digests the strand with its 5′ end at the broken end of the DNA, leaving the strand with its 3′ end at the break single-stranded. The lagging nascent strand (pink) base-paired to the unbroken parent strand (dark blue) is ligated to the unreplicated portion of the parent chromosome (light blue).
- Step 3: Multiple RecA (bacteria)/Rad51 (eukaryotes) proteins bind to the 3’ end of the single-stranded DNA and catalyze its *strand invasion* into another homologous, double-stranded DNA molecule and base-pairing with its complementary sequence, displacing the other strand.
- Step 4: The target DNA-invading strand hybrid region is extended in the direction away from the break (*branch migration*) forming a Holliday structure, in which all bases are base-paired to complementary bases in the parent strands (the diagonal lines represent single phosphodiester bonds).
- Step 5: Cleavage of the phosphodiester bonds that *cross over* from one parent strand to the other.
- Step 6: Ligation of the 5′ and 3′ ends base-paired to the same parent strands generates a structure similar to a replication fork.
- Step 7: Rebinding of replication fork proteins extends the leading strand past the point of the original strand break and re-initiates lagging-strand synthesis.

Week 2 (12 Sept 2019)

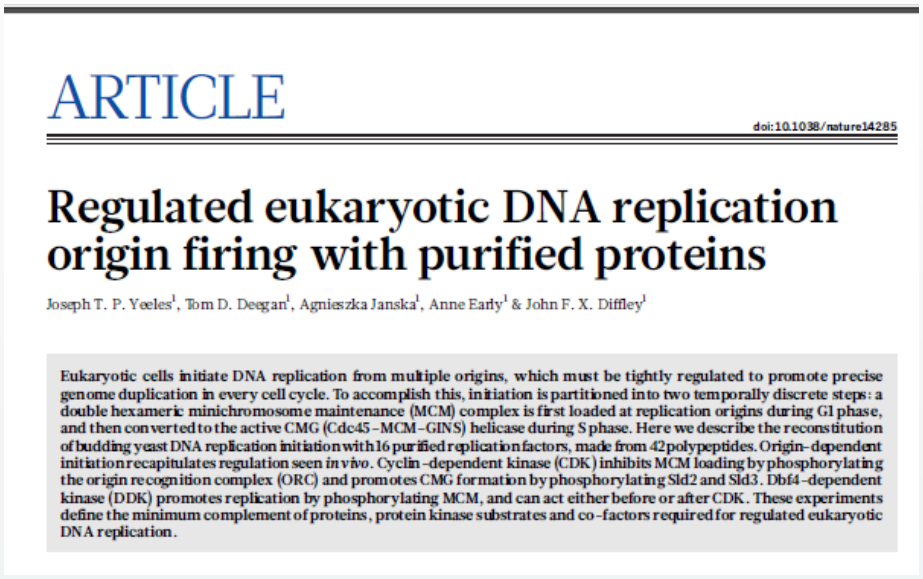


REVIEWS

DNA replication origin activation in space and time

Michalis Fragkos, Olivier Ganier, Philippe Coulombe* and Marcel Méchali*

Abstract | DNA replication begins with the assembly of pre-replication complexes (pre-RCs) at thousands of DNA replication origins during the G1 phase of the cell cycle. At the G1–S-phase transition, pre-RCs are converted into pre-initiation complexes, in which the replicative helicase is activated, leading to DNA unwinding and initiation of DNA synthesis. However, only a subset of origins are activated during any S phase. Recent insights into the mechanisms underlying this choice reveal how flexibility in origin usage and temporal activation are linked to chromosome structure and organization, cell growth and differentiation, and replication stress.



ARTICLE

doi:10.1038/nature14285

Regulated eukaryotic DNA replication origin firing with purified proteins

Joseph T. P. Yeeles[1], Tom D. Deegan[1], Agnieszka Janska[1], Anne Early[1] & John F. X. Diffley[1]

Eukaryotic cells initiate DNA replication from multiple origins, which must be tightly regulated to promote precise genome duplication in every cell cycle. To accomplish this, initiation is partitioned into two temporally discrete steps: a double hexameric minichromosome maintenance (MCM) complex is first loaded at replication origins during G1 phase, and then converted to the active CMG (Cdc45–MCM–GINS) helicase during S phase. Here we describe the reconstitution of budding yeast DNA replication initiation with 16 purified replication factors, made from 42 polypeptides. Origin-dependent initiation recapitulates regulation seen *in vivo*. Cyclin-dependent kinase (CDK) inhibits MCM loading by phosphorylating the origin recognition complex (ORC) and promotes CMG formation by phosphorylating Sld2 and Sld3. Dbf4-dependent kinase (DDK) promotes replication by phosphorylating MCM, and can act either before or after CDK. These experiments define the minimum complement of proteins, protein kinase substrates and co-factors required for regulated eukaryotic DNA replication.

<u>40 min</u> group presentation on two papers about DNA replication
<u>5-10 min</u> Q & A (questions from teacher and other students)

- Background
- Methods and materials
- Results
- Discussion