MOLECULAR BIOLOGY

DNA Replication

How the double helix copies itself accurately, the enzymes and the proofreading, with a 10-question practice set and full solutions.

SUBJECT	LEVEL	INCLUDES
Biology	High school and early college	Practice set, answer key, revision sheet

Read this note online, with updates: gauravtiwari.org/study-notes/dna-replication

INSIDE

DNA replication is the process by which a cell makes an exact copy of its DNA before dividing. Every time a cell divides, through mitosis or meiosis, it must first duplicate its entire genome so each daughter cell receives a complete set. The process is semiconservative (each new double helix contains one parental strand and one newly synthesized strand), bidirectional (proceeds in both directions from each origin of replication), and astonishingly accurate (roughly one error per billion base pairs copied, thanks to proofreading mechanisms). Watson and Crick's 1953 DNA structure paper foreshadowed the replication mechanism in a single famous closing sentence; Meselson and Stahl proved it experimentally in 1958.

DNA replication, the double helix unzips at the replication fork; each parental strand templates a new daughter strand.

Semiconservative Replication

When DNA replicates, the two parental strands separate and each one templates the synthesis of a new complementary strand. The result: two double helices, each containing one parental strand and one daughter strand. This is what 'semiconservative' means.

Three other models were considered in the 1950s, conservative (parental helix stays together; new helix is fully new), dispersive (parental and daughter DNA is mixed throughout both new helices), and a

'mixed' hybrid model. Meselson and Stahl's 1958 experiment with nitrogen-15 isotope labeling definitively proved the semiconservative model. The experiment is regarded as one of the most elegant in the history of biology.

The Replication Fork

DNA replication initiates at specific sites called origins of replication. The DNA double helix unwinds at the origin, forming a replication bubble that expands outward in both directions. Each end of the bubble is a replication fork, a Y-shaped junction where the parental DNA splits into two single strands and new DNA is being synthesized.

A single chromosome can have hundreds or thousands of origins of replication firing simultaneously. The human genome (3 billion base pairs per haploid set) replicates in about 6-8 hours during S phase because so many forks operate in parallel.

The Key Enzymes

DNA replication requires the coordinated work of several enzymes. The major players:

- **Helicase**: unwinds the double helix at the replication fork by breaking the hydrogen bonds between complementary bases.
- **Topoisomerase**: relieves the torsional strain that builds up as the helix unwinds ahead of helicase. Without it, the DNA would tangle.
- **Single-strand binding proteins** (**SSBs**): coat the separated single strands to prevent them from re-pairing or forming secondary structure before new synthesis occurs.
- **Primase**: synthesizes short RNA primers that DNA polymerase needs to start synthesis (DNA polymerase can extend an existing strand but can't start from scratch).
- **DNA polymerase**: the main synthesis enzyme. Reads the template strand 3' → 5' and synthesizes the new strand 5' → 3', adding one complementary nucleotide at a time. Multiple types exist; in humans, Pol α, δ, and ε are the main replication polymerases.
- **DNA ligase**: joins discontinuous DNA fragments on the lagging strand into a continuous strand.

Leading vs Lagging Strand

Here's the trick that makes DNA replication mechanically interesting: DNA polymerase can only synthesize in one direction (5' → 3'). But the two parental strands run in opposite (antiparallel) directions. So at each fork, one parental strand can be copied continuously while the other must be copied in pieces.

- **Leading strand.** The strand whose template runs 3' → 5' in the direction the fork is moving. DNA polymerase follows the fork continuously, synthesizing in a single uninterrupted run.
- **Lagging strand.** The strand whose template runs 5' → 3' in the direction the fork is moving. DNA polymerase has to work backwards, synthesizing in short fragments (Okazaki fragments) that are then joined by DNA ligase. Each Okazaki fragment requires its own RNA primer to start.

Okazaki fragments are 100-200 nucleotides long in eukaryotes and 1,000-2,000 in bacteria. Reiji Okazaki discovered them in 1968.

Proofreading and Error Correction

DNA polymerase makes mistakes, pairing the wrong base about once every 10^5 to 10^6 nucleotides. Most polymerases have a 3' → 5' exonuclease activity that immediately removes incorrectly paired bases and tries again. This proofreading reduces the error rate to about 10^{-7} to 10^{-8} per nucleotide.

After replication, a separate mismatch repair (MMR) system scans new DNA for remaining errors and corrects them. The MMR machinery distinguishes the new strand (which is unmethylated, briefly) from the old strand (which is methylated). Together, polymerase proofreading and post-replication MMR reduce the final error rate to about 1 mistake per 10^9 base pairs copied, astonishingly accurate for any biological process.

When MMR fails, the result is hereditary cancer syndromes like Lynch syndrome (defective MMR genes cause runaway mutation accumulation and high colorectal cancer risk).

Telomeres, The End Replication Problem

Linear chromosomes face a unique problem: the lagging strand's last RNA primer leaves a gap at the 5' end that cannot be filled (DNA polymerase has nothing to extend from beyond it). Without compensation, chromosomes would shorten with every replication.

Telomeres are repetitive DNA sequences (TTAGGG repeats in humans) at chromosome ends that act as buffers. Each replication shortens the telomeres slightly. When telomeres get critically short, cells stop dividing (replicative senescence). The enzyme telomerase can re-extend telomeres, but most somatic cells don't express telomerase actively, which is one of the factors limiting cell division across a lifespan.

Telomerase is highly active in stem cells, germ cells, and most cancer cells. Cancer cells' ability to re-extend their telomeres is one of the hallmarks of cancer and a target for some cancer therapies.

Related study notes: Nucleic Acid, Mitosis, Meiosis, Cell Cycle.

Practice Questions

Work each question before reading its solution. The set runs from direct recall and substitution to the applied questions that exams actually use to separate grades.

Q1. What does ”semiconservative” mean for DNA replication, and what experiment proved it?

SOLUTION

Each daughter double helix contains 1 original (parental) strand and 1 newly synthesized strand. Meselson and Stahl (1958) grew bacteria in heavy nitrogen, switched them to light nitrogen, and tracked DNA density by centrifugation: after 1 generation all DNA was intermediate density, exactly the semiconservative prediction and inconsistent with the conservative or dispersive alternatives.

Q2. What does helicase do, and why does unwinding create a topological problem elsewhere on the molecule?

SOLUTION

Helicase unwinds the double helix ahead of the replication fork, breaking the hydrogen bonds between base pairs. Because DNA is a closed, twisted structure, unwinding at 1 point forces extra twisting (supercoiling) to accumulate downstream, which topoisomerase must relieve by cutting and resealing the strand, or the molecule would tangle and stall replication.

Q3. Why must DNA polymerase always synthesize in the 5’ to 3’ direction, and what problem does this create at the fork?

SOLUTION

DNA polymerase can only add a new nucleotide to a free 3’-OH group, chemically dictated by the enzyme’s active site. Since the 2 template strands run antiparallel, one new strand can be built continuously toward the fork (leading strand) while the other must be built away from the fork in short backward-stitched pieces (lagging strand, Okazaki fragments), because its template runs the wrong direction for continuous synthesis.

Q4. What is a primer, why is it required, and which enzyme makes it?

SOLUTION

A short RNA sequence laid down by primase that supplies the free 3'-OH group DNA polymerase needs to start; polymerase cannot begin a strand from nothing. Every Okazaki fragment on the lagging strand needs its own new primer, which explains why the lagging strand requires far more priming events than the leading strand.

Q5. What are Okazaki fragments, and what 2 enzymes finish stitching them into a continuous strand?

SOLUTION

Short DNA segments (100-200 nucleotides in eukaryotes) synthesized discontinuously on the lagging strand. DNA polymerase I removes each RNA primer and fills the gap with DNA, and DNA ligase seals the remaining nick between fragments with a phosphodiester bond, producing 1 continuous strand from many short pieces.

Q6. Explain proofreading: which enzyme activity catches a misincorporated base, and what does it do about it?

SOLUTION

DNA polymerase carries a built-in 3' to 5' exonuclease activity that senses a mismatched base pair by its distorted geometry, backs up, excises the wrong nucleotide, and resumes synthesis with the correct one. This immediate self-correction, on top of the initial base-pairing selectivity, is what pushes replication fidelity from roughly 1 error per 10,000 bases down toward 1 per 10 million.

Q7. DNA replication error rates land around 1 mistake per 10^9-10^{10} bases after all repair. Combine the contributing accuracy layers and their approximate individual error rates to reach that number.

SOLUTION

Base-pairing selectivity alone: about 1 in 10^4-10^5. Polymerase proofreading multiplies fidelity by roughly 100-fold, reaching 10^6-10^7. Post-replication mismatch repair adds another 100-fold, landing near 10^8-10^{10}. Three independent filters multiply their error-catching power, which is why the final rate is so much better than any single mechanism achieves alone.

Q8. What is the end-replication problem, and how do telomeres and telomerase address it?

SOLUTION

Because the lagging strand needs a primer at its very end, and that terminal primer's gap cannot be filled by ordinary polymerase (no upstream 3'-OH to extend from), each replication round shortens the chromosome slightly. Telomeres, repetitive noncoding buffer sequences at chromosome ends, absorb this loss without touching genes; telomerase, active in germ cells and stem cells, extends the telomere tem-

plate so those lineages can divide indefinitely. Most somatic cells lack active telomerase, and telomere shortening is 1 proposed clock behind cellular aging.

Q9. How many replication forks operate simultaneously on a human chromosome, and why is this necessary given the genome's size?

SOLUTION

Thousands of origins fire per chromosome, each generating 2 forks moving in opposite directions, so replication proceeds from many points at once rather than 1 continuous start-to-end pass. A single fork at typical eukaryotic speeds (~50 nucleotides/sec) would take weeks to copy a chromosome; multiple simultaneous origins compress the whole genome's replication into a manageable S-phase of a few hours.

Q10. Contrast bacterial and eukaryotic replication origins: number and consequence for genome copying time.

SOLUTION

A bacterial chromosome typically replicates from a SINGLE origin (oriC in E. coli), sufficient because the genome is small (~4.6 million bp) and the fork moves fast. A human chromosome, with genomes in the billions of base pairs, uses MANY origins per chromosome, tens of thousands genome-wide, because 1 origin would take impossibly long. Origin count scales with genome size to keep replication time roughly comparable across organisms.

Answer Key

QUICK ANSWERS

Q1 1 old strand + 1 new strand per daughter; Meselson-Stahl **Q2** unwinds the helix; forces supercoiling ahead, fixed by topoisomerase **Q3** polymerase needs a 3'-OH; forces leading/lagging asymmetry **Q4** RNA primer from primase supplies the starting 3'-OH **Q5** polymerase I replaces primers; ligase seals the nicks

Q6 3'→5' exonuclease excises mismatches, then resumes **Q7** 3 layers multiply: pairing × proofreading × mismatch repair **Q8** lagging-strand ends can't be fully primed; telomeres/telomerase buffer it **Q9** thousands of origins per chromosome; parallelism is required for speed **Q10** bacteria: 1 origin; eukaryotes: many origins, scaled to genome size

Revision Sheet

The model

Semiconservative: each daughter helix keeps 1 parental strand + 1 new strand (Meselson-Stahl).

Fidelity stack

Base pairing + proofreading exonuclease + mismatch repair multiply to ~1 error per 10^9-10^{10} bases.

Key enzymes

Helicase unwinds; topoisomerase relieves supercoiling; primase primes; polymerase extends 5→3′; ligase seals nicks.

End problem

Lagging-strand ends can't be fully primed; telomeres buffer, telomerase extends in stem/germ cells.

Leading vs lagging

Leading: continuous. Lagging: Okazaki fragments, because polymerase only extends 3′-OH and templates run antiparallel.

Scale

Bacteria: 1 origin. Eukaryotes: thousands of origins per genome, for speed.

About These Notes

This note is part of a free study-notes series on gauravtiwari.org covering mathematics, physics, chemistry, and biology: the concepts that keep deciding grades in high school, board, and entrance exams.

2008	2,000+	125+	M.Sc.
Writing and teaching online since	Articles published on gauravtiwari.org	Free study notes in this series	Mathematics, with physics and chemistry training

Gaurav Tiwari is a developer, educator, and founder of Gatilab. The study-notes series exists because most exam prep material is either a full textbook or a thin definition page, and revision needs something in between: one concept, complete, with the traps named and the practice attached.

GET THE FULL SERIES

Every note in the series is free to read online and free to download as a PDF like this one. Browse the complete collection at **gauravtiwari.org/free-study-notes-exam-prep** and pick the topics your next exam actually covers.

USE IT FREELY

This PDF is free for personal study, classrooms, and study groups. Share the file or the link with anyone it can help. The one thing not allowed is reselling it or republishing it as your own work.

Gaurav Tiwari

gauravtiwari.org