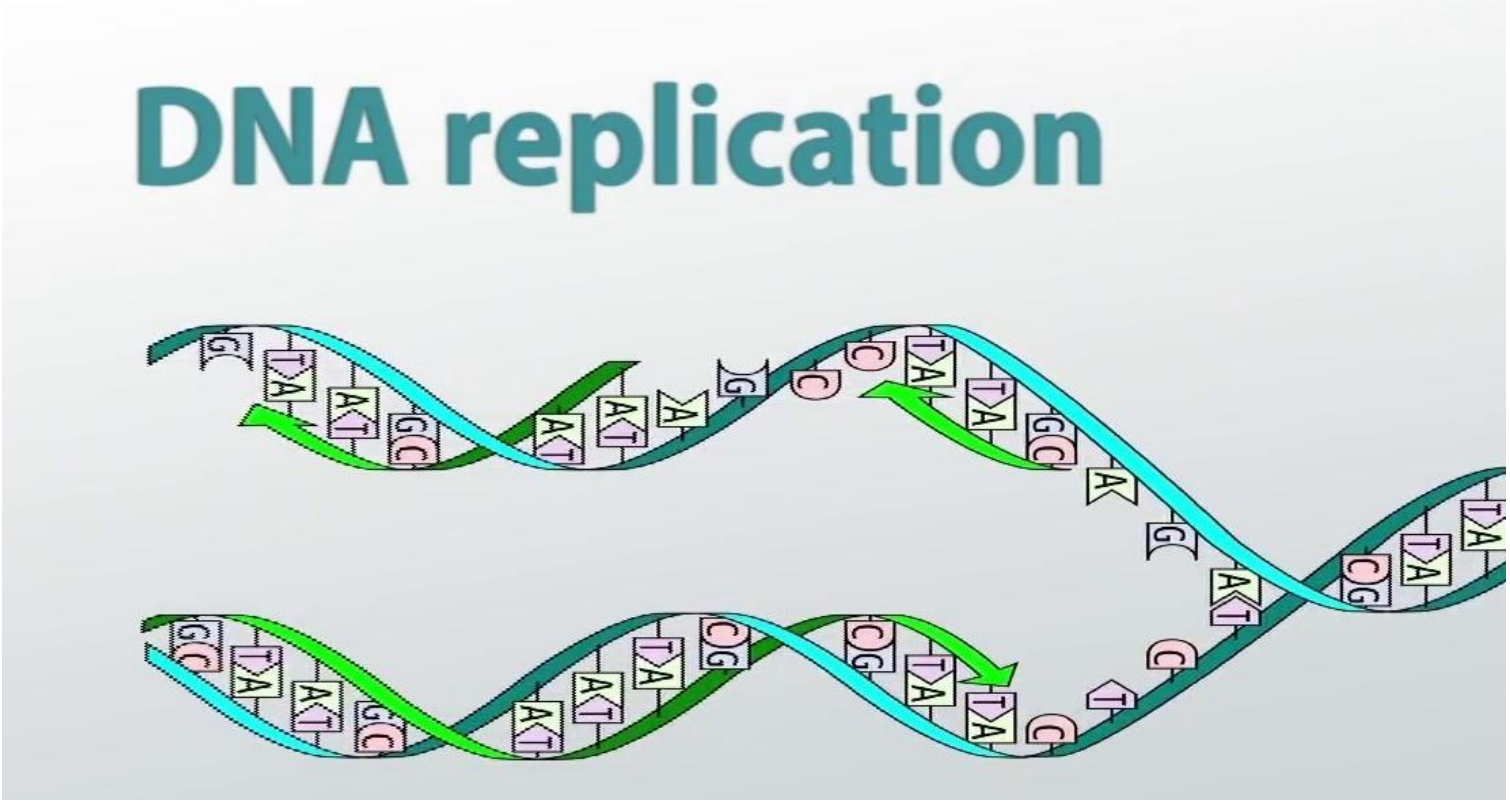
DNA replication



Dr. Animesh Mondal

Dept. of Botany

B B College, Asansol-03

Discovery of Replication

Watson and Crick's seminal paper describing the DNA double helix ended with the statement:

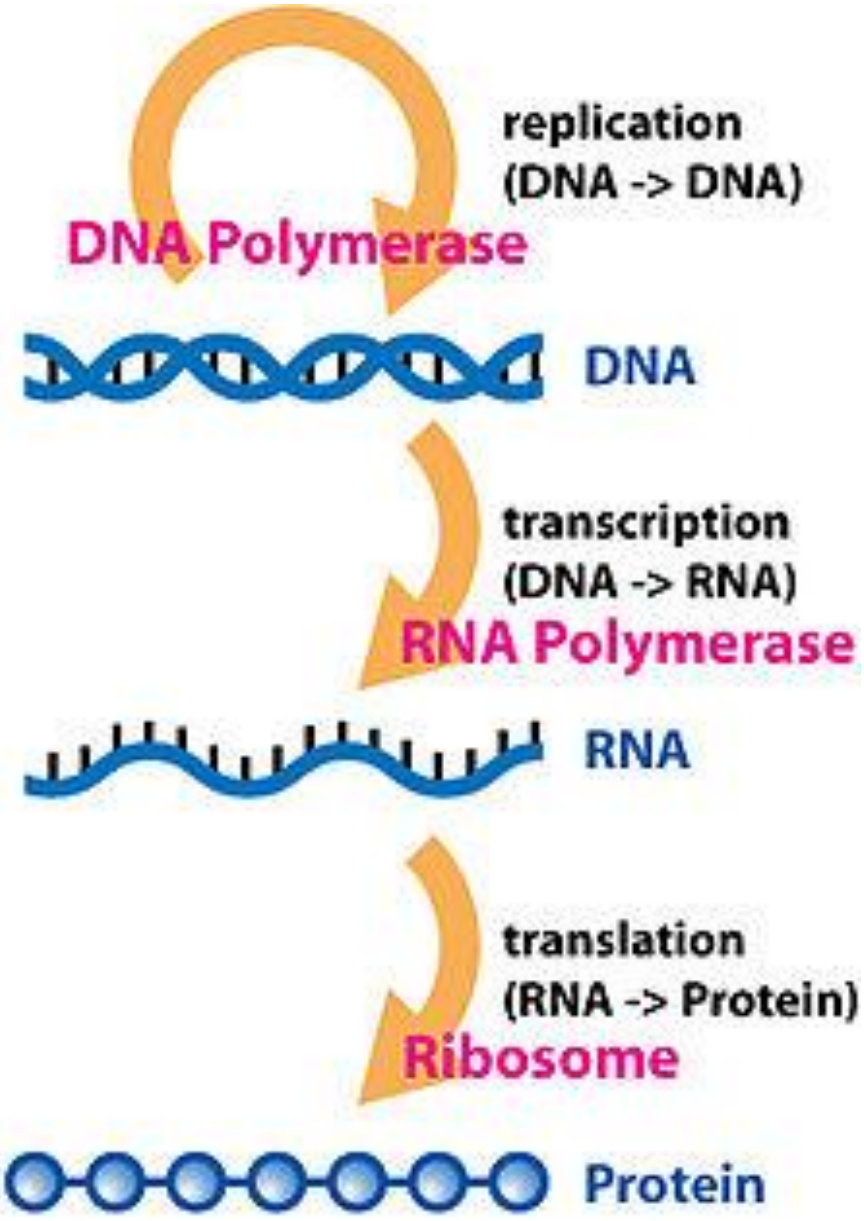
"It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material."

The Nobel Prize in Physiology or Medicine 1962

James Dewey Watson
Francis Harry Compton Crick
Maurice Hugh Frederick Wilkins
(From extreme Left to Right)
Prize share: 1/3 (each)

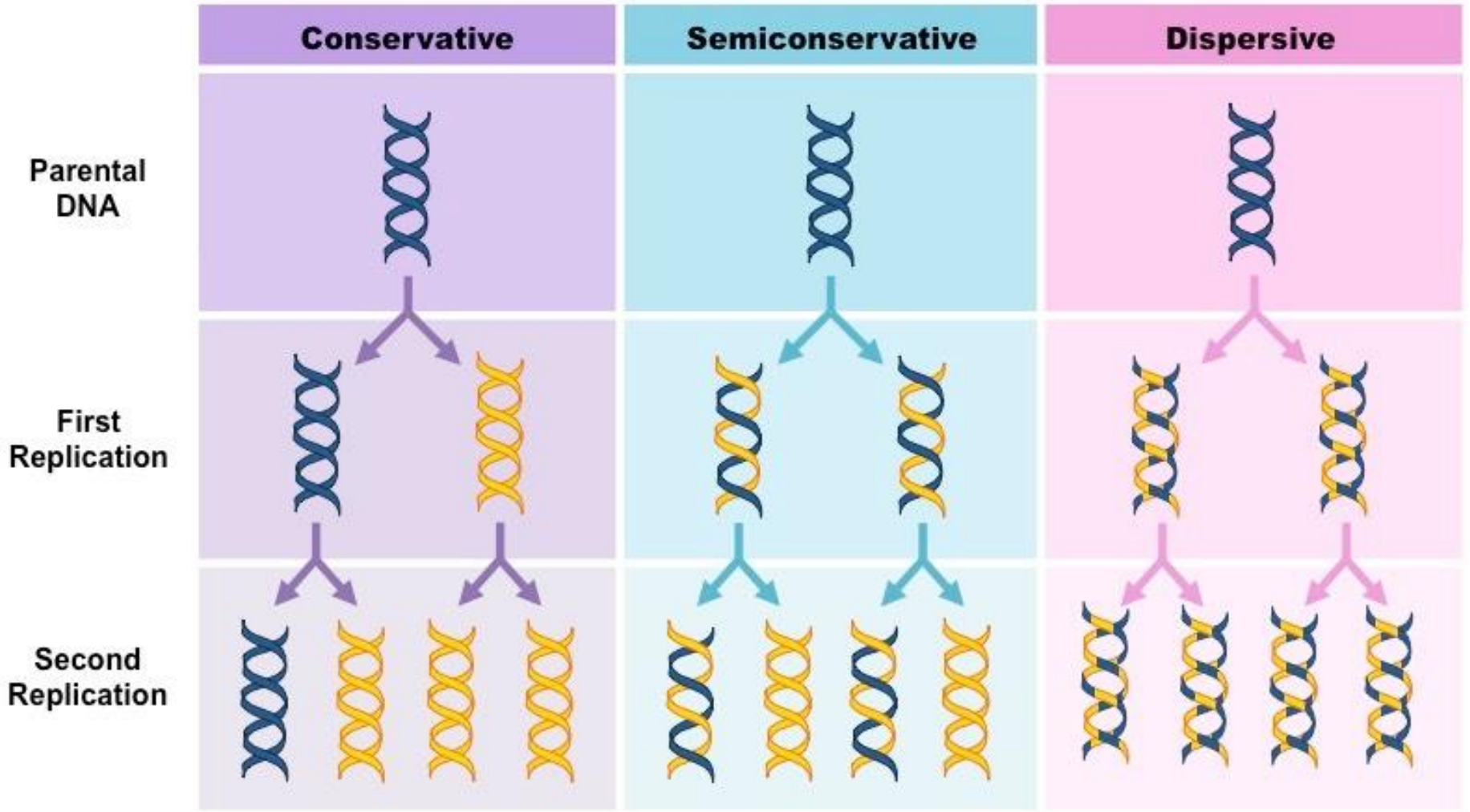


The Nobel Prize in Physiology or Medicine 1962 was awarded jointly to Francis Harry Compton Crick, James Dewey Watson and Maurice Hugh Frederick Wilkins *"for their discoveries concerning the molecular structure of nucleic acids and its significance for information transfer in living material"*.

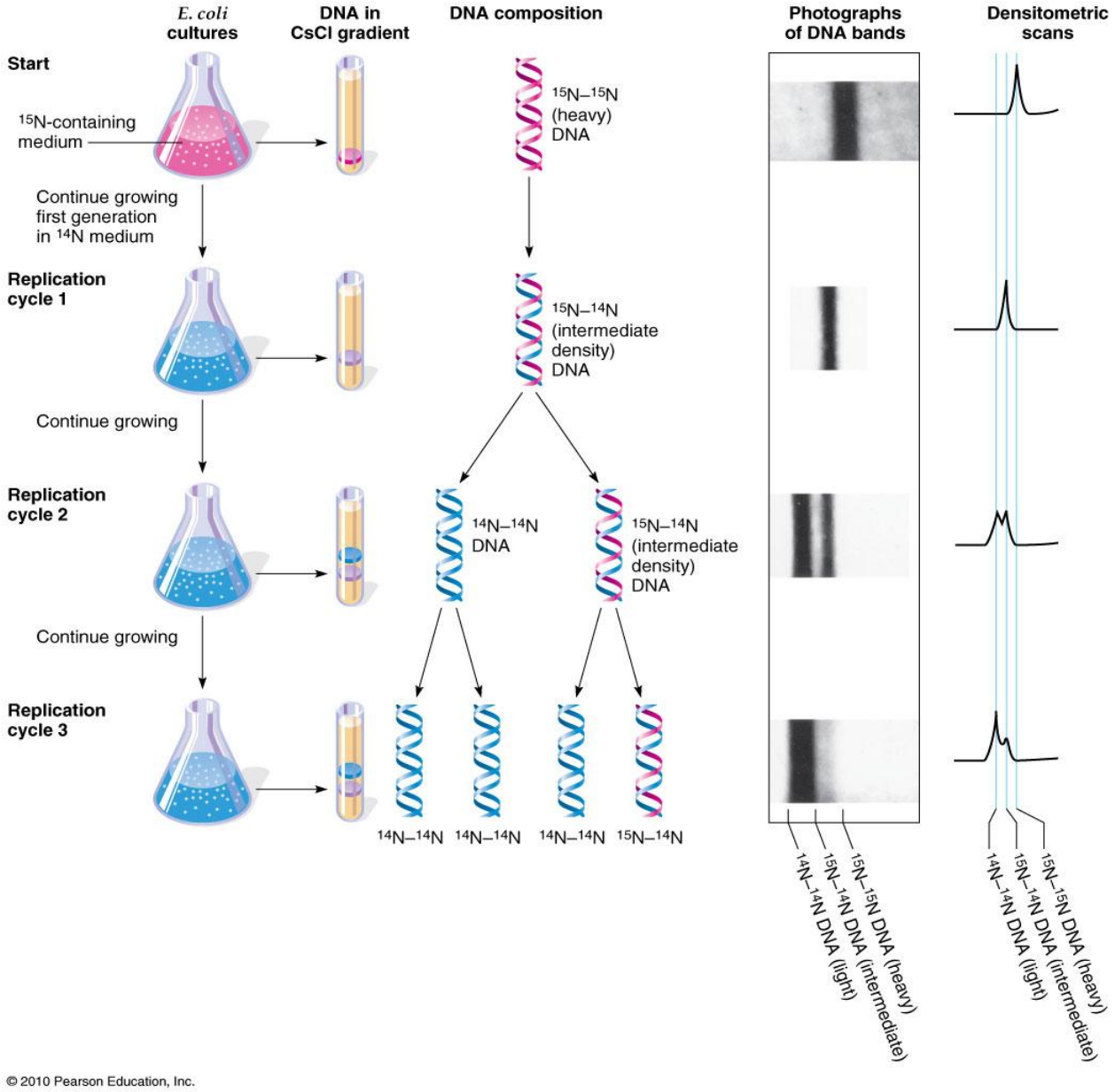

replication
(DNA -> DNA)
DNA Polymerase
DNA
transcription
(DNA -> RNA)
RNA Polymerase
RNA
translation
(RNA -> Protein)
Ribosome
Protein

Characteristics of replication

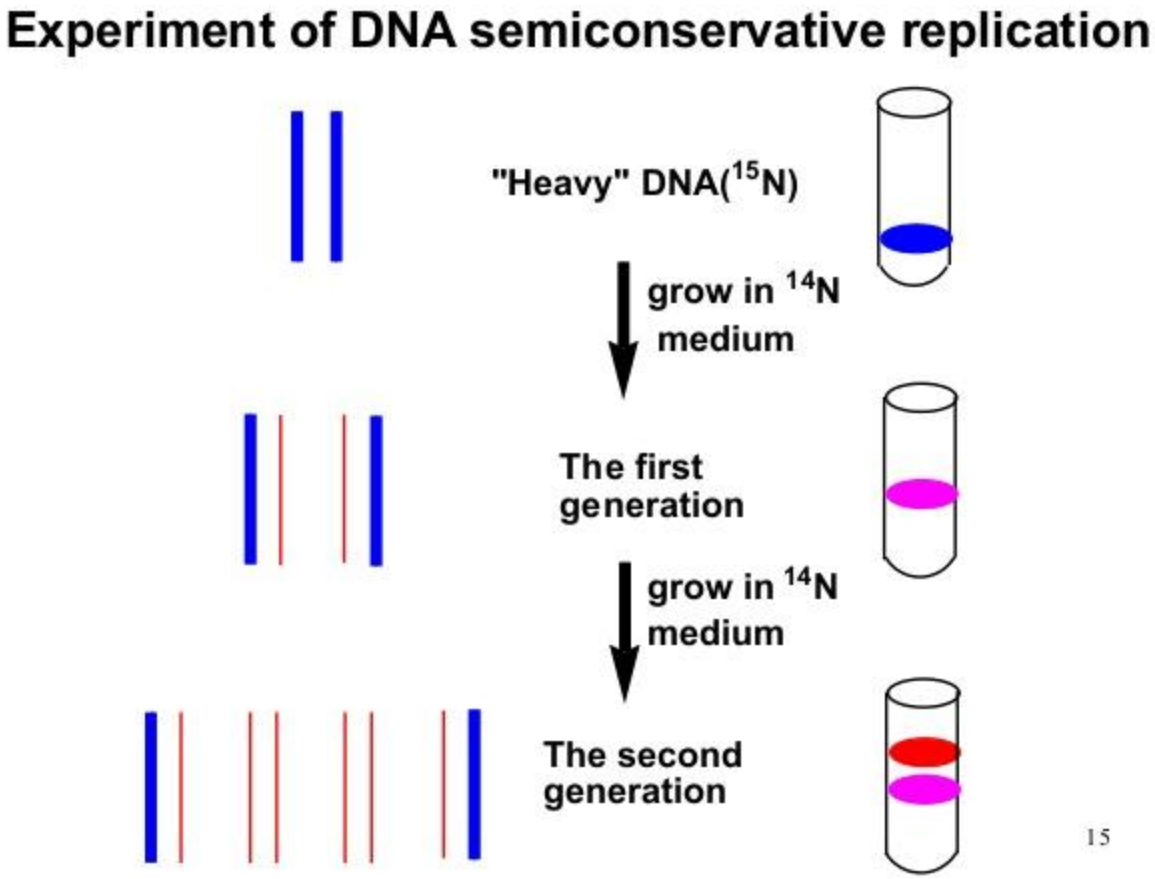
- Semi-conservative replication
- Bidirectional replication
- Semi-continuous replication
- High fidelity


Conservative
Semiconservative
Dispersive
Parental DNA
First Replication
Second Replication

CsCl equilibrium density gradient



Sucrose velocity density gradient

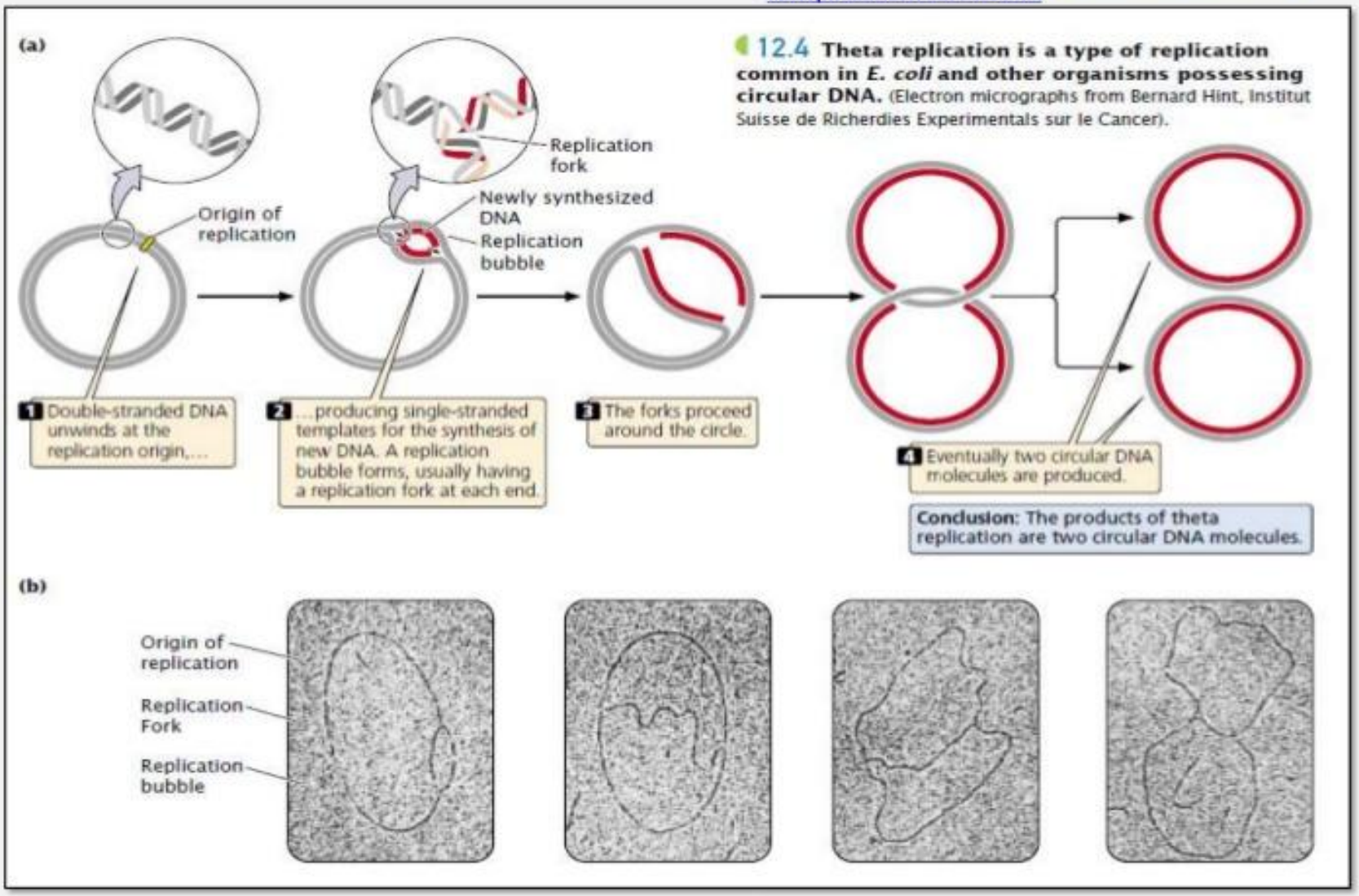


§1.2 Bidirectional Replication

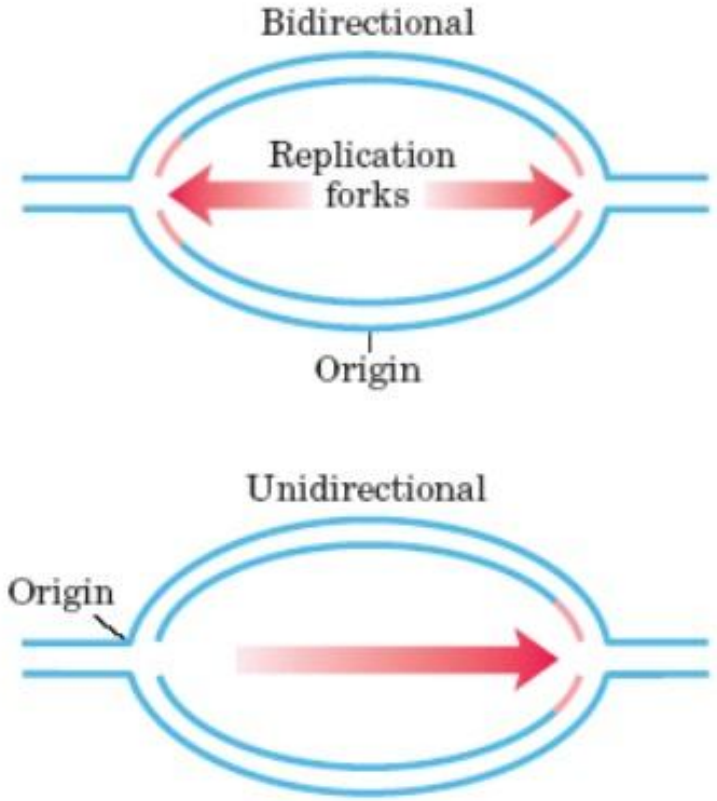
- Replication starts from unwinding the dsDNA at a particular point (called origin), followed by the synthesis on each strand.
- The parental dsDNA and two newly formed dsDNA form a Y-shape structure called replication fork.

THETA REPLICATION: *E. coli*

http://highered.mcgraw-hill.com/olcweb/cgi/pluginpop.cgi?it=swf::535::535::/sites/dl/free/0072437316/120073/micro03.swf::Bidirectional%20Replication%20of%20DNA



Bidirectional replication



§1.3 Semi-continuous Replication

The daughter strands on two template strands are synthesized differently since the replication process obeys the principle that DNA is synthesized from the 5´ end to the 3´end.

§2.7 Replication Fidelity

- Replication based on the principle of base pairing is crucial to the high accuracy of the genetic information transfer.
- Enzymes use two mechanisms to ensure the replication fidelity.
  - Proofreading and real-time correction
  - Base selection

Enzymes involved in

DNA Replication

(1) DNA polymerase.

(2) enzymes known as helicases that separate the DNA strands at the replication fork.

(3) proteins that prevent them from re-annealing before they are replicated.

(4) Enzymes that synthesize RNA primers.

(5) DNA topo-isomerases.

(6) An enzyme to remove the RNA primers, and

(7) an enzyme to covalently link successive Okazaki fragments.

DNA polymerase in *E.coli*

- There are atleast five DNA polymerases associated with *E.coli* DNA replication
- These are
  - **DNA polymerase I** – it is the first DNA pol to be isolated and purified. The enzyme is encoded by polA gene. The polymerase has 5'->3' exonuclease activity
  - **DNA polymerase II** – the enzyme is encoded by polB gene. It has no 5'->3' exonuclease activity, rather it has 3'-5' exonuclease activity(proofreading activity)
  - **DNA polymerase III –** this is the principle replication enzyme,encoded by the gene polC. It has 5'->3' proofreading activities. It lacks 3'->5' exonuclease activity. It has a very high proccessivity
  - **DNA polymerase IV and DNA polymerase V** are involved in an unusual form of DNA repair

§2.1 DNA Polymerase

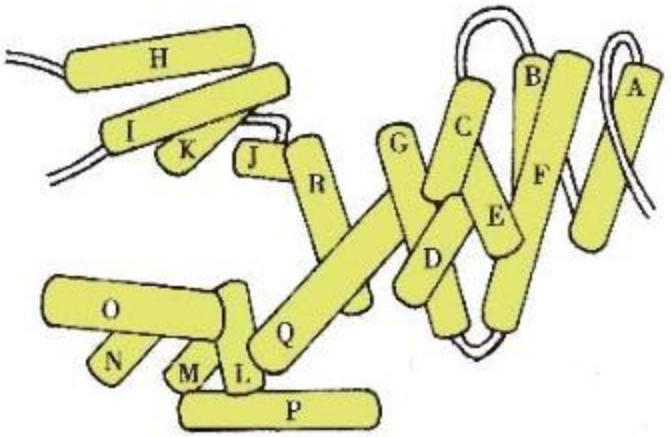
DNA-pol of prokaryotes

- **The first DNA-dependent DNA polymerase (short for DNA-pol I) was discovered in 1958 by Arthur Kornberg who received Nobel Prize in physiology or medicine in 1959.**

Arthur Kornberg

DNA-pol I

- **Mainly responsible for proofreading and filling the gaps, repairing DNA damage**



DNA Polymerases

TABLE 25–1 Comparison of DNA Polymerases of *E. coli*

	DNA polymerase		
	I	II	III
Structural gene*	*polA*	*polB*	*polC* (*dnaE*)
Subunits (number of different types)	1	7	≥10
M_r	103,000	88,000†	791,500
3′→5′ Exonuclease (proofreading)	Yes	Yes	Yes
5′→3′ Exonuclease	Yes	No	No
Polymerization rate (nucleotides/s)	16–20	40	250–1,000
Processivity (nucleotides added before polymerase dissociates)	3–200	1,500	≥500,000

DNA Polymerase-I

In 1957, Arthur Kornberg reported that he had discovered an enzyme that catalyzes the synthesis of DNA in extracts of *E. coli* through its ability to incorporate the radioactive label from [14C] thymidine triphosphate into DNA. This enzyme, which has since become known as **DNA polymerase I or Pol I, consists of a monomeric 928-residue polypeptide.**

Properties of DNA Pol-I

A. Pol I Recognizes the Incoming dNTP According to the Shape of the Base Pair It Forms with the Template DNA.

B. Pol I Can Edit Its Mistakes

In addition to its polymerase activity, Pol I has two independent hydrolytic activities:

1. It can act as a 3' → 5' exonuclease.

2. It can act as a 5' → 3' exonuclease.

C. Pol I's Polymerase and Two Exonuclease Functions Each Occupy Separate Active Sites

D. The X-Ray Structure of Klenow Fragment Indicates How It Binds DNA.

E. DNA Polymerase Distinguishes Watson–Crick Base Pairs via Sequence-Independent Interactions That Induce Domain Movements

F. The DNA Polymerase Catalytic Mechanism Involves Two Metal Ions.

Their active sites all contain two metal ions, usually Mg2+, that are liganded by two invariant Asp side chains in the palm domain. Metal ion B is liganded by all three phosphate groups of the bound dNTP, whereas metal ion A bridges the –phosphate group of this dNTP and the primer's 3'-OH group.

G. Editing Complexes Contain the Primer Strand in the 3' → 5 ' Exonuclease Site.

H. Pol- I Functions Physiologically to Repair DNA.

I. Pol- I Catalyzes Nick Translation.

J. Pol I's 5' → 3' Exonuclease Functions Physiologically to Excise RNA Primers

DNA polymerase II

- Pol II has 3'-5' exonuclease activity and participates in DNA repair.
- Pol II is also thought to be a backup to Pol III as it can interact with holoenzyme proteins and assume a high level of processivity.
- The main role of Pol II is thought to be the ability to direct polymerase activity at the replication fork and help stalled Pol III bypass terminal mismatches.

DNA Polymerase III

- **DNA polymerase III holoenzyme** is the primary enzyme complex involved in prokaryotic DNA replication
- DNA polymerase III synthesizes base pairs at a rate of around ***1000 nucleotides per second.***
- As replication progresses and the replisome moves forward, DNA polymerase III arrives at the RNA primer and begins replicating the DNA, adding onto the 3'OH of the primer
- DNA polymerase III will then synthesize a continuous or discontinuous strand of DNA
- DNA polymerase III has a ***high processivity*** and therefore, synthesizes DNA very quickly.
- This high processivity is due in part to ***the β-clamps*** that "hold" onto the DNA strands.

Table 30-1 Properties of *E. coli* DNA Polymerases

	Pol I	Pol II	Pol III
Mass (kD)	103	90	130
Molecules/cell	400	?	10–20
Turnover number[a]	600	30	9000
Structural gene	*polA*	*polB*	*polC*
Conditionally lethal mutant	+	–	+
Polymerization: 5′ → 3′	+	+	+
Exonuclease: 3′ → 5′	+	+	+
Exonuclease: 5′ → 3′	+	–	–

[a]Nucleotides polymerized $min^{-1} \cdot molecule^{-1}$ at 37°C.

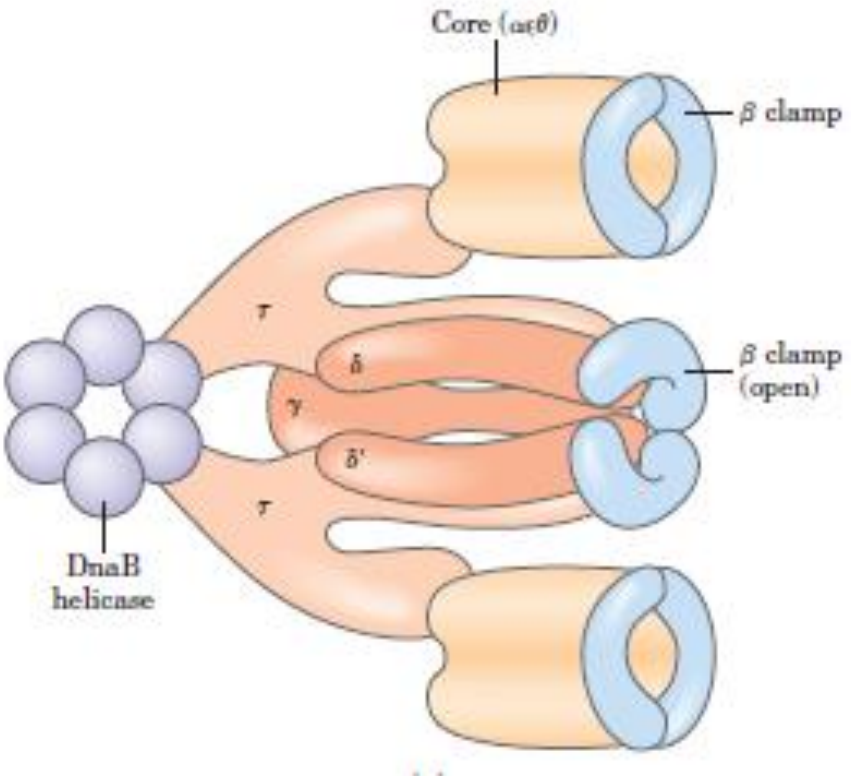
Source: Kornberg, A. and Baker, T.A., *DNA Replication* (2nd ed.), *p.* 167, Freeman (1992).

Sub units of DNA Pol-III

TABLE 25-2 Subunits of DNA Polymerase III of *E. coli*

Subunit	Number of subunits per holoenzyme	M_r of subunit	Gene	Function of subunit	
α	2	129,900	*polC* (*dnaE*)	Polymerization activity	Core polymerase
ε	2	27,500	*dnaQ* (*mutD*)	3′→5′ Proofreading exonuclease	
θ	2	8,600	*holE*		
τ	2	71,100	*dnaX*	Stable template binding; core enzyme dimerization	Clamp-loading (γ) complex that loads β subunits on lagging strand at each Okazaki fragment
γ	1	47,500	*dnaX**	Clamp loader	
δ	1	38,700	*holA*	Clamp opener	
δ'	1	36,900	*holB*	Clamp loader	
χ	1	16,600	*holC*	Interaction with SSB	
ψ	1	15,200	*holD*	Interaction with γ and χ	
β	4	40,600	*dnaN*	DNA clamp required for optimal processivity	

Structure of DNA Pol-III

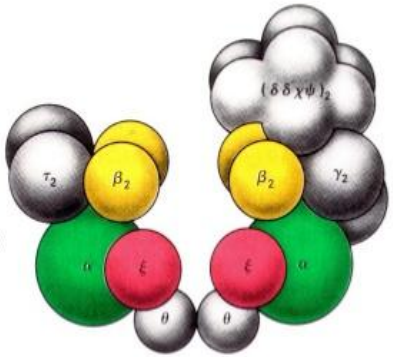


Structure of DNA-pol III

α : has 5′→ 3′ polymerizing activity

ε : has 3′→ 5′ exonuclease activity and plays a key role to ensure the replication fidelity.

θ: maintain heterodimer structure



DNA polymerase IV

- An error-prone DNA polymerase involved in non-targeted mutagenesis.
- During SOS induction, Pol IV production is increased tenfold and one of the functions during this time is to interfere with Pol III holoenzyme processivity.
- This creates a checkpoint, stops replication, and allows time to repair DNA lesions via the appropriate repair pathway.

DNA polymerase V

Pol V is a Y-family DNA polymerase that is involved in SOS response and translesion synthesis DNA repair mechanisms.

RNA Primers

- DNA polymerases' all but universal requirement for a free 3'-OH group to extend a DNA chain poses a question that was emphasized by the establishment of the semidiscontinuous model of DNA replication: How is DNA synthesis initiated? Careful analysis of Okazaki fragments revealed that *their 5' ends consist of RNA segments of 1 to 60 nt (a length that is species dependent) that are complementary to the template DNA chain (Fig. 30-7). E. coli has two enzymes* that can catalyze the formation of these **RNA primers: RNA polymerase,** the 459-kD multisubunit enzyme that mediates transcription (Section 31-2), and the much smaller **primase (60 kD),** the monomeric product of the *dnaG gene.*

§2.2 Primase

- Also called DnaG
- Primase is able to synthesize primers using free NTPs as the substrate and the ssDNA as the template.
- Primers are short RNA fragments of a several decades of nucleotides long.

Function

- **Primers provide free 3´-OH groups to react with the α-P atom of dNTP to form phosphoester bonds.**
- **Primase, DnaB, DnaC and an origin form a primosome complex at the initiation phase.**

§2.3 Helicase

- Also referred to as DnaB.
- It opens the double strand DNA with consuming ATP.
- The opening process with the assistance of DnaA and DnaC

§2.4 SSB protein

- **Stand for single strand DNA binding protein**
- **SSB protein maintains the DNA template in the single strand form in order to**
  - **prevent the dsDNA formation;**
  - **protect the vulnerable ssDNA from nucleases.**

46

Single-strand DNA-binding **protein** (**SSB**) is a **protein**, 178 amino acids long, found in *Escherichia coli* (*E. coli*) bacteria, that binds to single-stranded regions of deoxyribonucleic acid (DNA). Single-stranded DNA is produced during all aspects of DNA metabolism: replication, recombination, and repair.

§2.5 Topoisomerase

- Opening the dsDNA will create supercoil ahead of replication forks.
- The supercoil constraint needs to be released by topoisomerases.

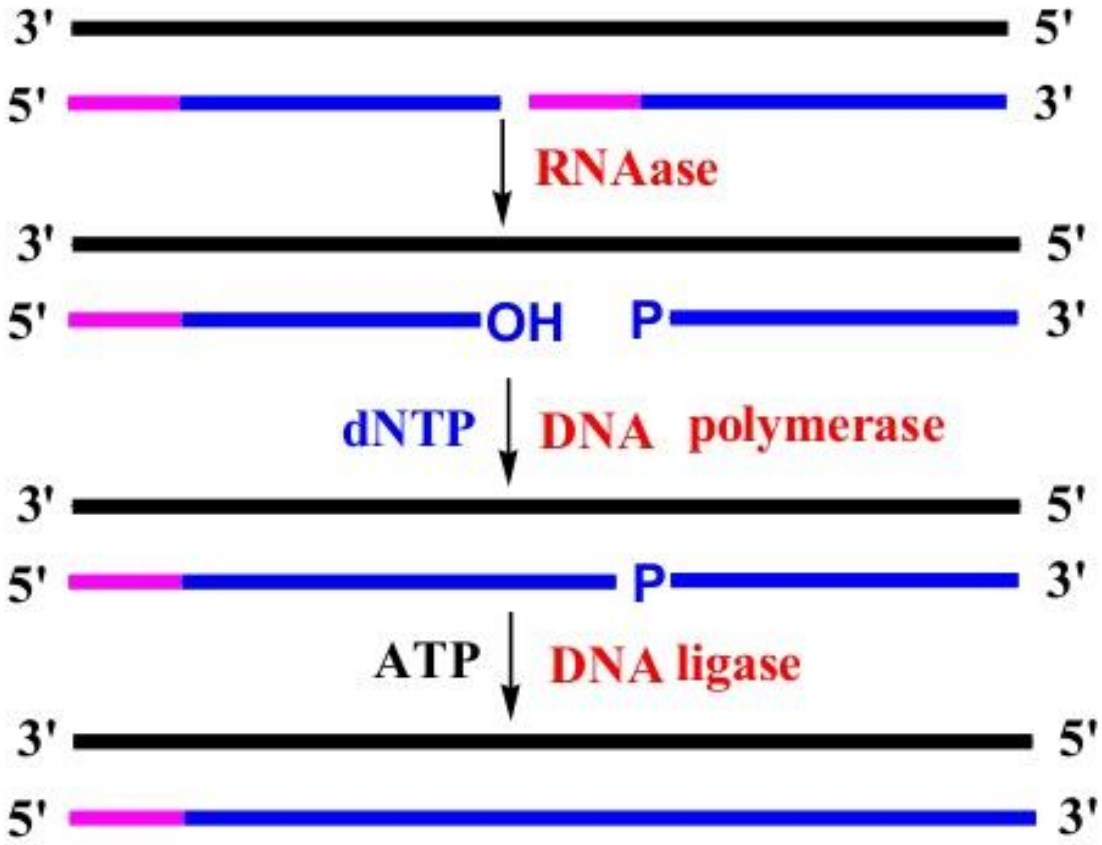
Topoisomerase I (topo I)

- Also called ω-protein in prokaryotes.
- It cuts a phosphoester bond on one DNA strand, rotates the broken DNA freely around the other strand to relax the constraint, and reseals the cut.

Topoisomerase II (topo II)

- It is named gyrase in prokaryotes.
- It cuts phosphoester bonds on both strands of dsDNA, releases the supercoil constraint, and reforms the phosphoester bonds.
- It can change dsDNA into the negative supercoil state with consumption of ATP.

§2.6 DNA Ligase



- Connect two adjacent ssDNA strands by joining the 3´-OH of one DNA strand to the 5´-P of another DNA strand.
- Sealing the nick in the process of replication, repairing, recombination, and splicing.

Short List of proteins & their *Mr.* involved in DNA Replication

TABLE 25-4 Proteins at the *E. coli* Replication Fork

Protein	M_r	*Number of subunits*	*Function*
SSB	75,600	4	Binding to single-stranded DNA
DnaB protein (helicase)	300,000	6	DNA unwinding; primosome constituent
Primase (DnaG protein)	60,000	1	RNA primer synthesis; primosome constituent
DNA polymerase III	791,500	17	New strand elongation
DNA polymerase I	103,000	1	Filling of gaps; excision of primers
DNA ligase	74,000	1	Ligation
DNA gyrase (DNA topoisomerase II)	400,000	4	Supercoiling

Modified from Kornberg, A. (1982) *Supplement to DNA Replication*, Table S11-2, W. H. Freeman and Company, New York.

Bacteriophage M13

M13 PHAGE

- Bacteriophage M13 was first isolated from wastewater in Munich (Hofschneider, 1963). Hence named as M13 phage.
- It is a filamentous phage which has **6407** nucleotides.
- It possess *single stranded circular DNA*.
- It was sequenced by ***Sanger*** in 1982.

Bacteriophage M13

M13 carries a 6408-nt single-stranded circular DNA known as its **viral or (+) strand. On infecting an** *E. coli cell, this strand directs the synthesis of its complementary* or (-) strand to form the circular duplex **replicative form (RF), which may be either nicked (RF II) or** supercoiled **(RF I). This replication process (Fig. 30-23)** may be taken as a paradigm for leading strand synthesis in duplex DNA.

As the M13 (+) strand enters the *E. coli cell, it becomes* coated with SSB except at a palindromic 57-nt segment that forms a hairpin. RNA polymerase commences primer synthesis 6 nt before the start of the hairpin and extends the RNA 20 to 30 residues to form a segment of RNA–DNA hybrid duplex. The DNA that is displaced from the hairpin becomes coated with SSB so that when RNA polymerase reaches it, primer synthesis stops. Pol III holoenzyme then extends the RNA primer around the circle to form the (–) strand. The primer is removed by Pol-I catalyzed nick translation, thereby forming RF II, which is converted to RF I by the sequential actions of DNA ligase and DNA gyrase.

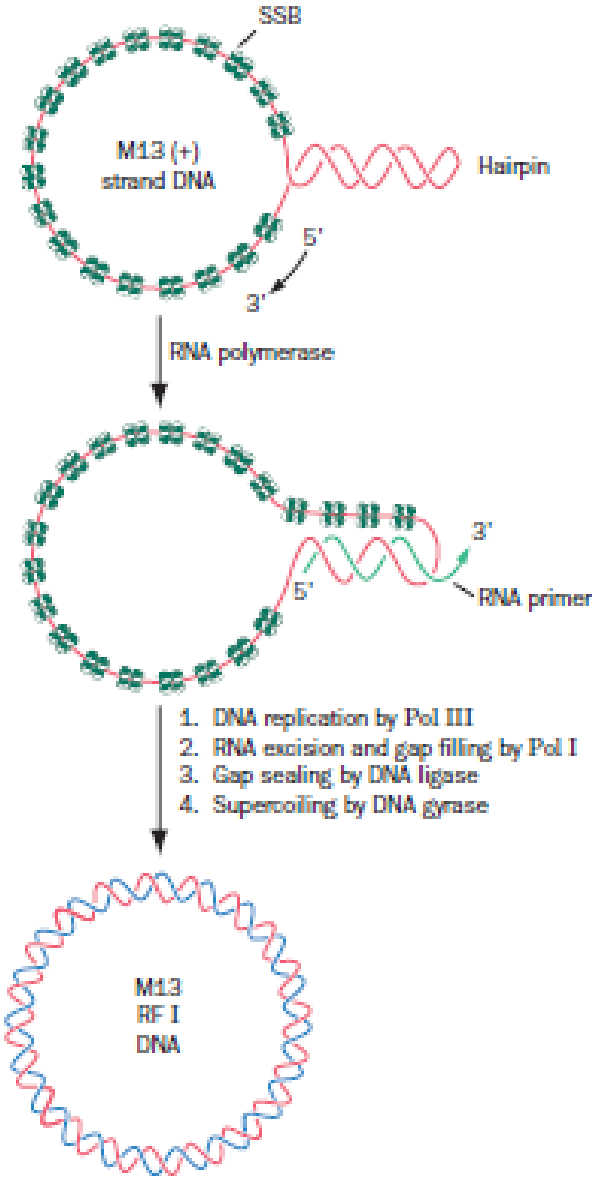


Figure 30-23 **The synthesis of the M13 (−) strand DNA on a (+) strand template to form M13 RF I DNA.**

Bacteriophage ϕX174

Phy X 174 (-) Strand Replication Is a Paradigm for Lagging Strand Synthesis

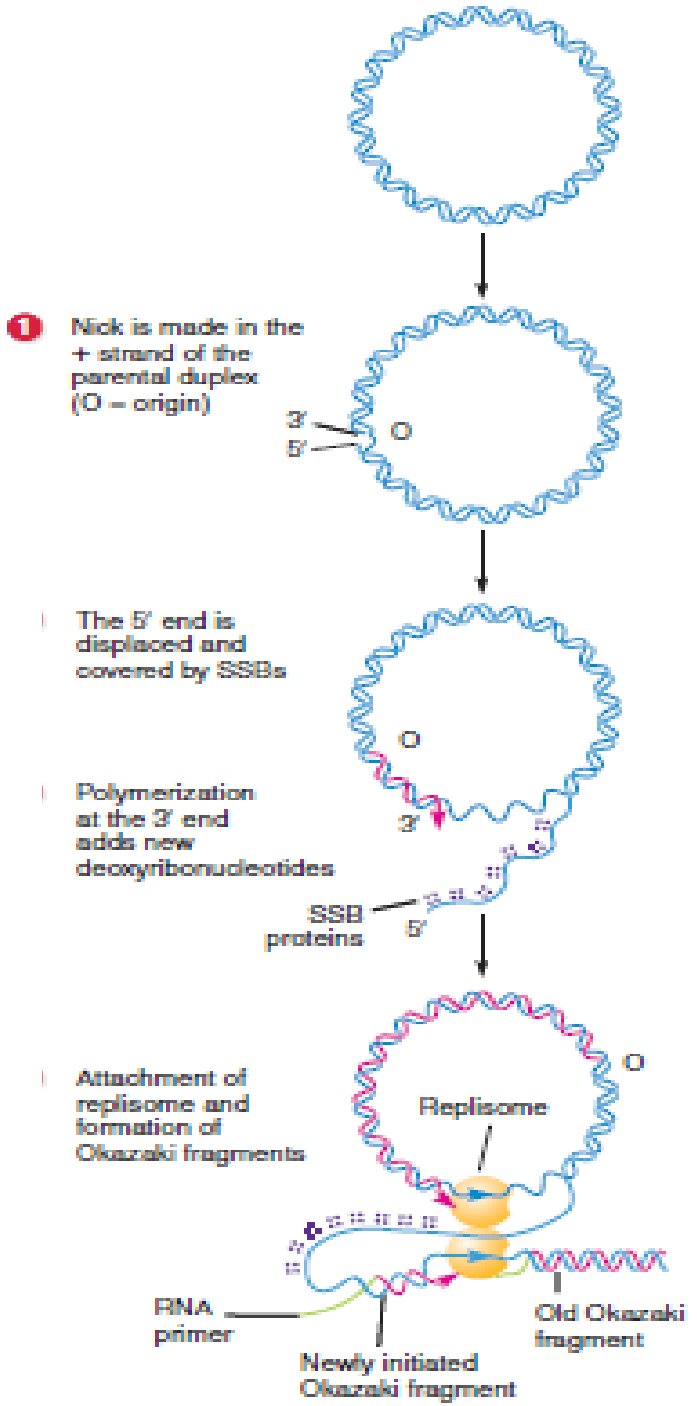


Table 30-4 **Proteins of the Primosome**[a]

Protein	Subunit Structure	Subunit Mass (kD)
PriA	Monomer	76
PriB	Dimer	11.5
PriC	Monomer	23
DnaT	Trimer	22
DnaB	Hexamer	50
DnaC[b]	Monomer	29
Primase (DnaG)	Monomer	60

[a]The complex of all primosome proteins but primase is known as the

Phi X174 (+) Strand Replication Serves as a Model for Leading Strand Synthesis

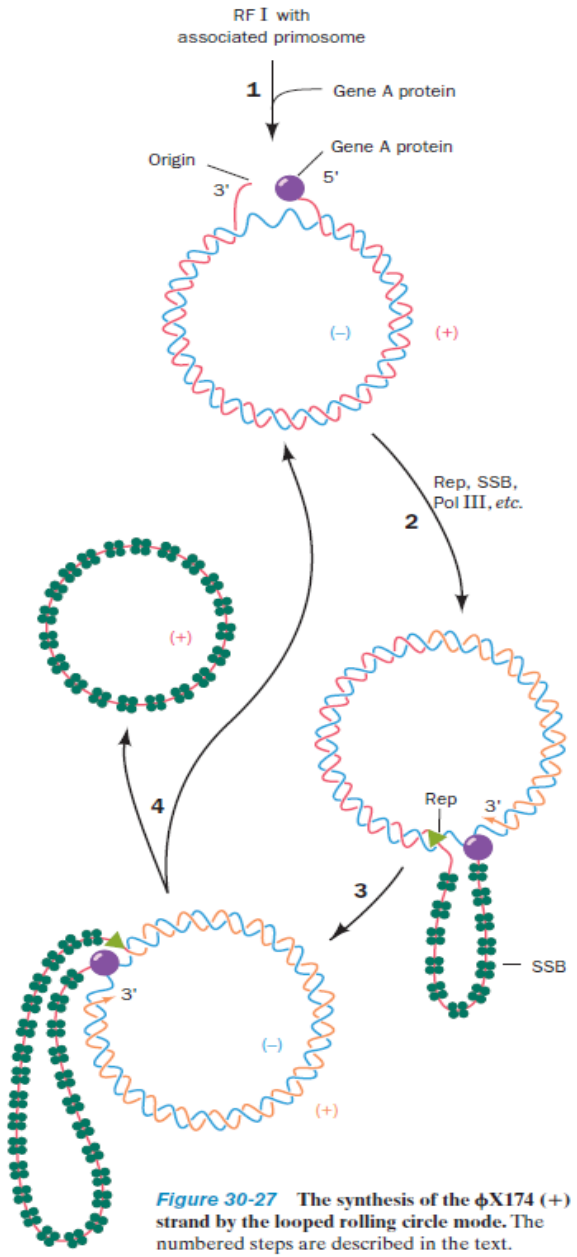


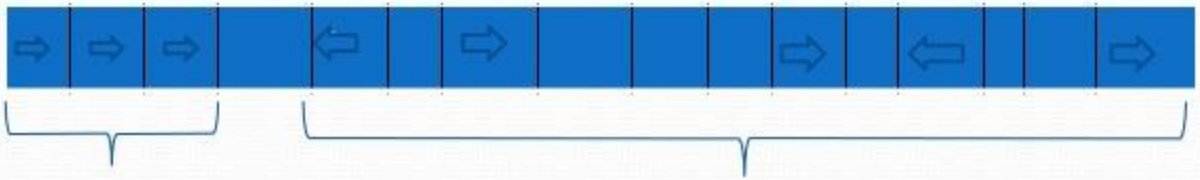
Figure 30-27 **The synthesis of the ϕX174 (+) strand by the looped rolling circle mode.** The numbered steps are described in the text.

General Features of DNA replication

- DNA replication begins with the ***unwinding*** of two anti-parallel complementary strands, resulting in the formation of two single strands.
- This unwinding produces the two ***replication forks***
- The replication proceeds in **5'-> 3' direction** and is ***semi-discontinuous***
- DNA replication is ***semi-conservative***
- DNA replication begins at the ***origin of replication (ori)***
- DNA is synthesized by the enzyme ***DNA polymerases***
- Of the two strands, one strand is synthesized **continuously** (5' to 3')in the direction of movement of the replication fork called ***leading strand***; while the other strand is synthesized **discontinuously** away from the movement of replication fork in short segments called the ***lagging strand***
- DNA replication is ***bidirectional*** from the origin of replication

DNA Replication in Prokaryote (*E.coli*)

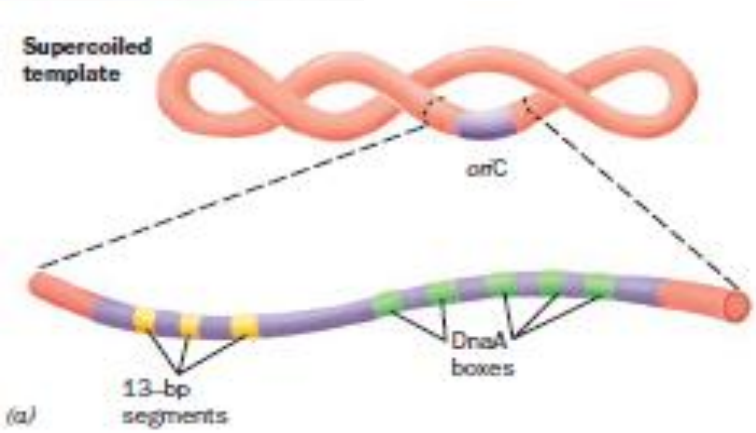
- The genome of *E.coli* is replicated bi-directionally from a single origin, oriC . *E. coli* replication is circular with no free ends. Replication of DNA in E. coli is also known as theta replication and it occurs in three steps:
  1) ***Initiation***
  2) ***Elongation***
  3) ***Termination***



THREE REPEATS OF 13 BP SEQUENCES

BINDING SITES OF DnaA PROTEIN , FIVE 9-BP SEQUENCES

FIG: ARRANGEMENT OF SEQUENCES IN THE *E. coli* REPLICATION ORIGIN. oriC



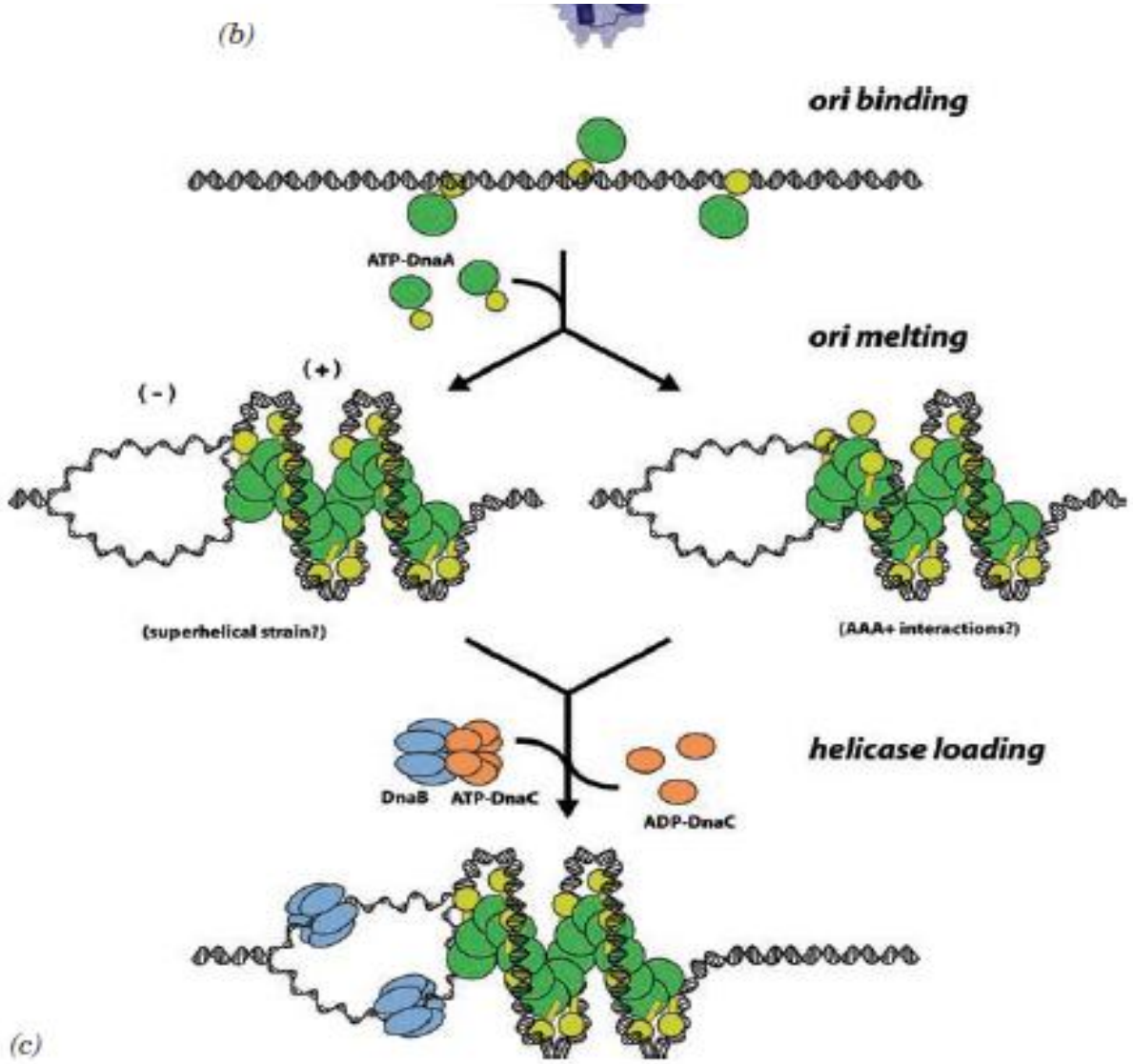

(b)
ori binding
ATP-DnaA
ori melting
(-)
(+)
(superhelical strain?)
(AAA+ interactions?)
helicase loading
DnaB
ATP-DnaC
ADP-DnaC
(c)

TABLE 25–3 Proteins Required to Initiate Replication at the *E. coli* Origin

Protein	M_r	*Number of subunits*	*Function*
DnaA protein	52,000	1	Recognizes ori sequence; opens duplex at specific sites in origin
DnaB protein (helicase)	300,000	6*	Unwinds DNA
DnaC protein	29,000	1	Required for DnaB binding at origin
HU	19,000	2	Histonelike protein; DNA-binding protein; stimulates initiation
Primase (DnaG protein)	60,000	1	Synthesizes RNA primers
Single-stranded DNA–binding protein (SSB)	75,600	4*	Binds single-stranded DNA
RNA polymerase	454,000	5	Facilitates DnaA activity
DNA gyrase (DNA topoisomerase II)	400,000	4	Relieves torsional strain generated by DNA unwinding
Dam methylase	32,000	1	Methylates (5′)GATC sequences at *oriC*

E. coli replication of origin

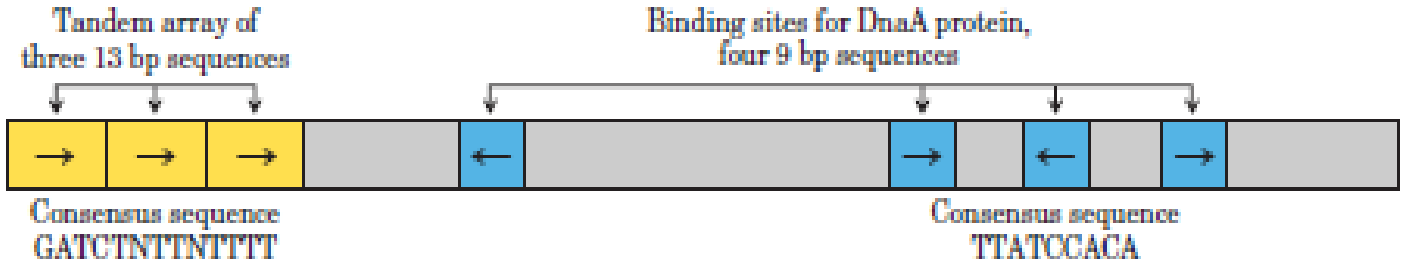


FIGURE 25–11 Arrangement of sequences in the *E. coli* replication origin, *oriC*. Although the repeated sequences (shaded in color) are not identical, certain nucleotides are particularly common in each position, forming a consensus sequence. In positions where there is no consensus, N represents any of the four nucleotides. The arrows indicate the orientations of the nucleotide sequences.

Initiation

- In *E. coli* , initiation at origin (oriC) requires several protein factors. These are:
  - >**DnaA protein** – recognizes oriC sequence, opens duplex at specific sites in origin
  - >**DnaB protein** (helicase) – unwinds DNA
  - >**DnaC protein** – required for DnaB binding at origin
  - >**Primase (DnaG protein)** – synthesizes Rna primers
  - >**SSB** – bind to ss DNA
  - >**DNA gyrase(DNA topoisomeraseII)** – relieves topological stress by DNA unwinding
  - >**Dam methylase** – methylates (5')GATC sequences at oriC

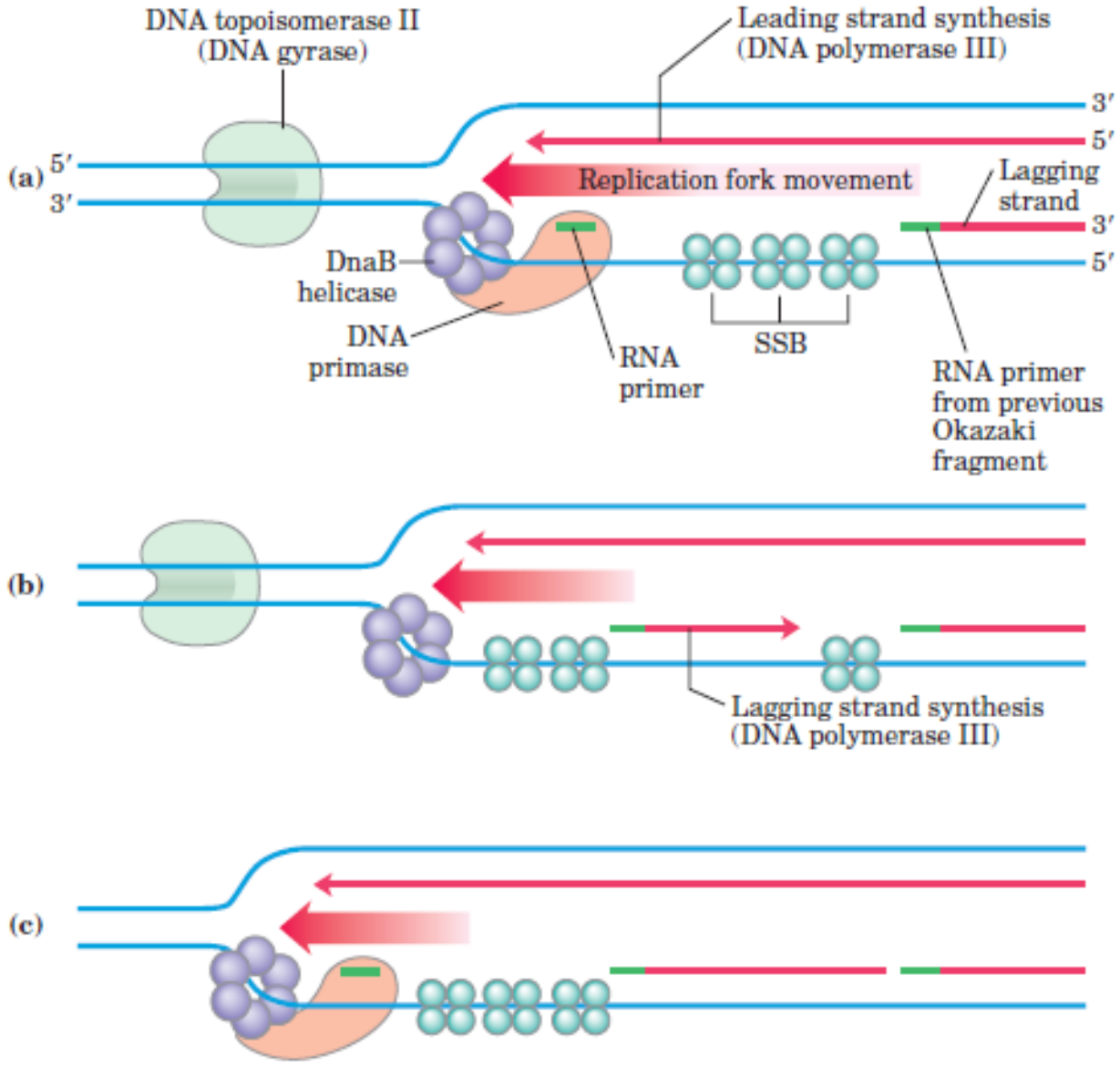
Sequence of events during Initiation

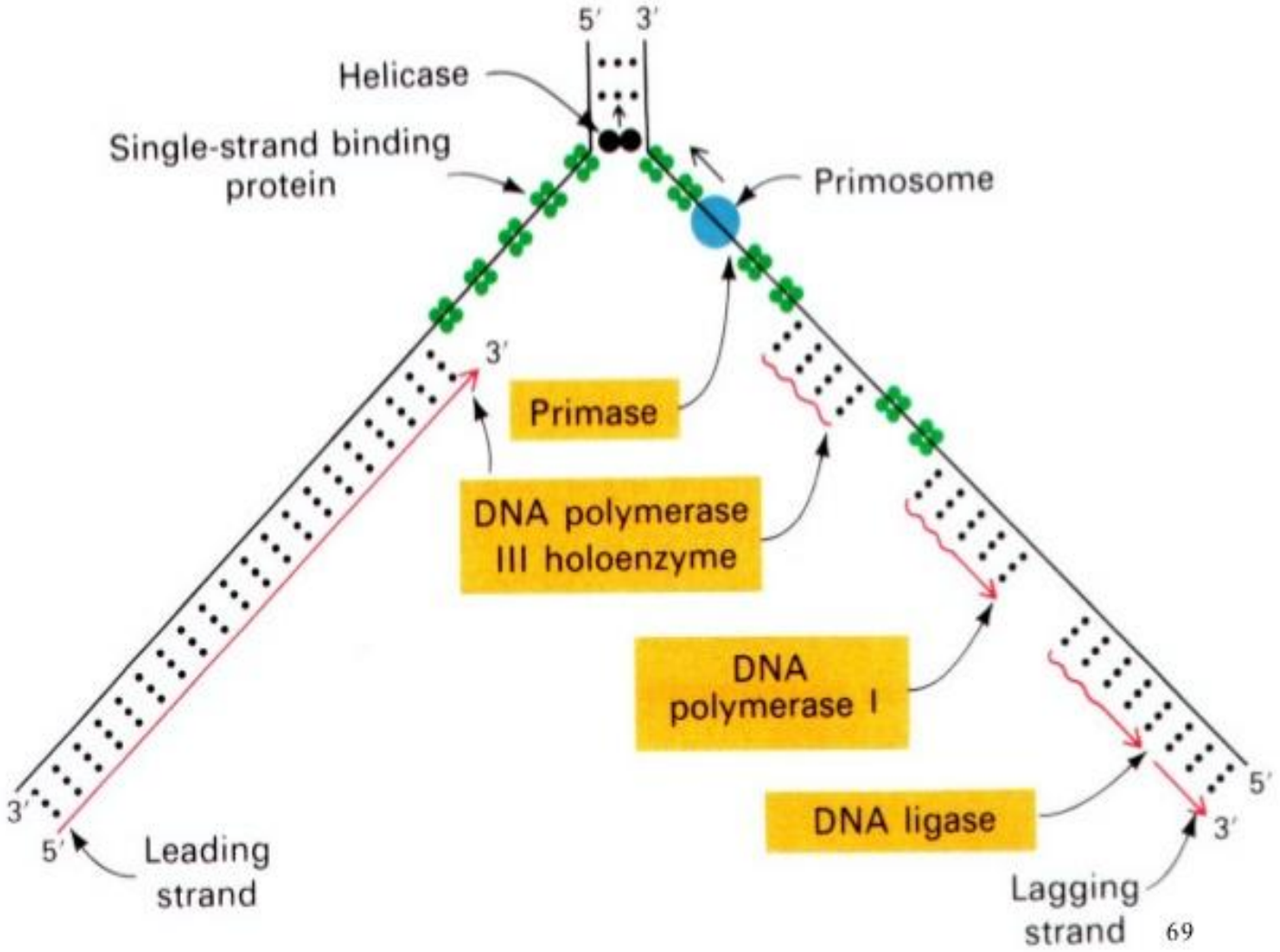
- The *E. coli* oriC , consists of **245bp of DNA** , which are highly conserved. It consists of **two short sequences:**

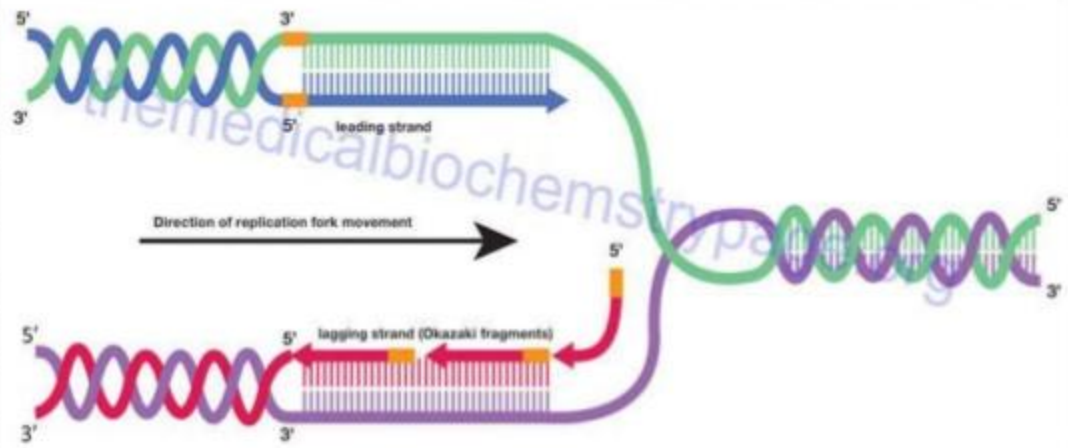
  one is of *five repeats of 9bp sequence sites*, remains dispersed , where DnaA protein binds and the other is *three repeats of 13bp sequence* , which is continuous and a region rich in A=T base pairs, called DNA unwinding elements (DUE) .
- The binding of DnaA to the 9bp requires ATP, which facilitates the initial strand separation of *E. coli* DNA duplex, which leads to the denaturation in A=T rich DUE region.
- DnaC protein binds to the DnaB protein, which results in the opening up of DnaB ring. The two DnaB ring is loaded on the denatured DUE , one on each strand
- DnaC is released and two DnaB remains bound to the two ss DNA
- ***DnaB helicase*** is then loaded on to two ss strands , which then travels along the ss DNA in the 5'->3' direction unwinding the DNA strand, creating two replication forks.
- Many SSB then binds to ss DNA to stabilize the strands separation
- Unwinding produces stress which is removed by ***DNA gyrase***
- Finally the oriC DNA is methylated by the ***Dam methylase*** , which methylates the adenine at N^6 within the palindromic sequence (5') GATC.

ELONGATION

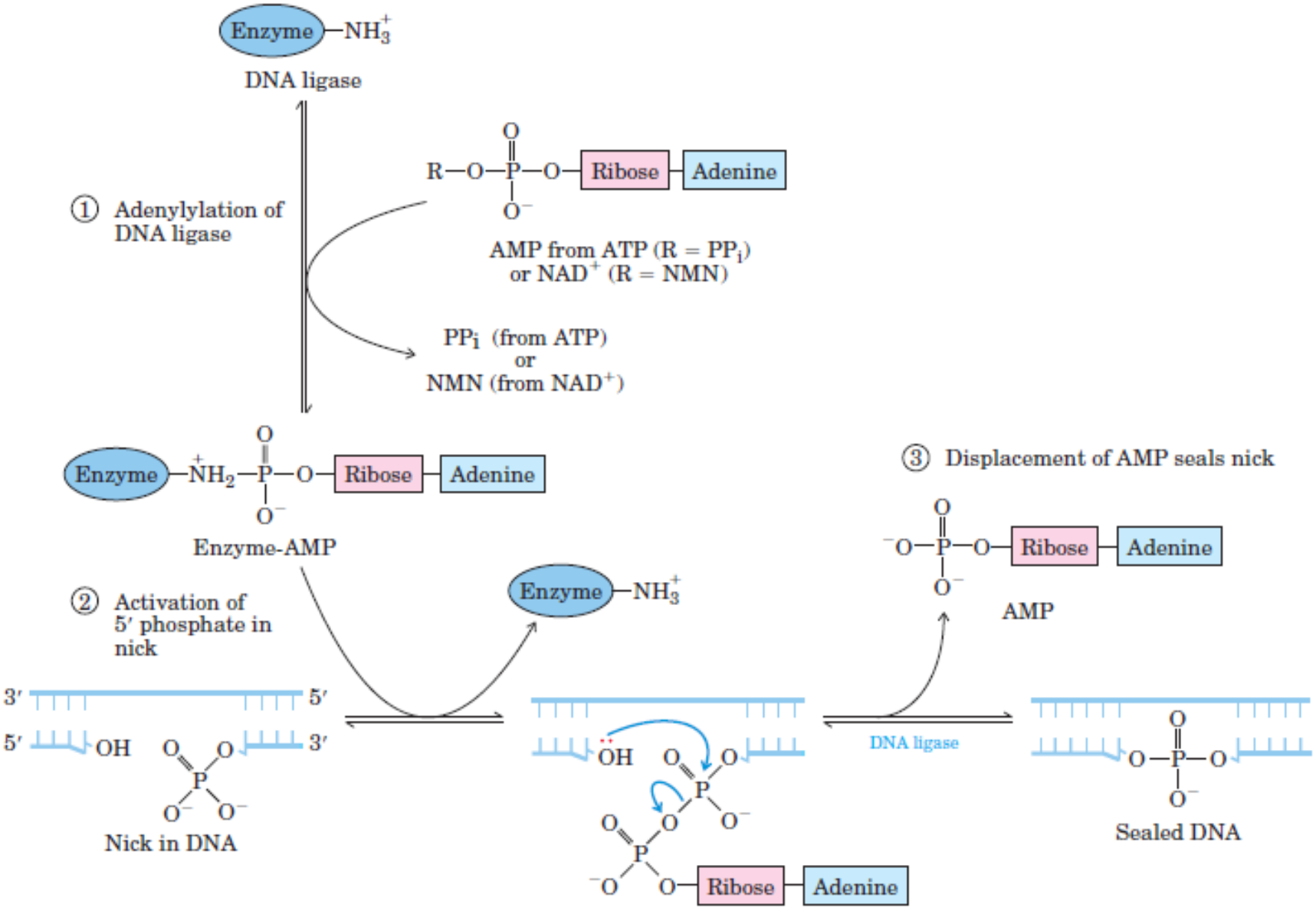
- In this phase, the synthesis of two new daughter strand takes place complementary to the template strand
- *DNA polymerase III* is the enzyme that synthesizes the daughter strands
- At this point, *a primer* is needed so that DNA polymerase III can begin to act.
- A primer is a short segment of RNA, that provides *3'-OH group* to which a nucleotide can be added
- This phase is marked by the synthesis of *leading strand and lagging strand*
- Leading strand is synthesized continuously in *5' to 3'* direction along the direction of the movement of replication fork
- Lagging strand synthesis occurs discontinuously *by loop formation* in short segments called *Okazaki fragments*.
- The lagging strand is *looped* so that DNA synthesis proceeds steadily on both the leading and lagging strand templates at the same time.
- The synthesis of Okazaki fragments on the lagging strand requires *DnaB helicase and DnaG primase* that constitute a functional unit within the replication complex, the *primosome.*
- Of the two core subunits of DNA polymerase III , one of the core subunit cycles from one Okazaki fragment to the next on the looped lagging strand.
- *DnaB helicase* first unwinds the replication fork
- *DNA primase* then associates with DnaB helicase, which synthesizes a short RNA primer
- The clamp loading complex of DNA pol III loads a *β-sliding clamp* to the primer
- The primer is then extended by the DNA pol III, which completes the synthesis of one Okazaki fragment
- When synthesis of an Okazaki fragment has been completed, replication halts, and the core subunits of DNA polymerase III dissociate from their sliding clamp (and from the completed Okazaki fragment) and associate with the new clamp This initiates synthesis of a new Okazaki fragment. Once an Okazaki fragment has been completed, its RNA primer is removed and replaced with DNA by DNA polymerase I, and the remaining nick is sealed by DNA ligase .
- DNA ligase catalyzes the formation of phosphodiester bond


DNA topoisomerase II
(DNA gyrase)
Leading strand synthesis
(DNA polymerase III)
3′
5′
(a) 5′
3′
Replication fork movement
Lagging
strand
3′
5′
DnaB
helicase
DNA
primase
RNA
primer
SSB
RNA primer
from previous
Okazaki
fragment
(b)
Lagging strand synthesis
(DNA polymerase III)
(c)


5′ 3′
Helicase
Single-strand binding protein
Primosome
3′
Primase
DNA polymerase III holoenzyme
DNA polymerase I
DNA ligase
3′
5′
5′
Leading strand
Lagging strand

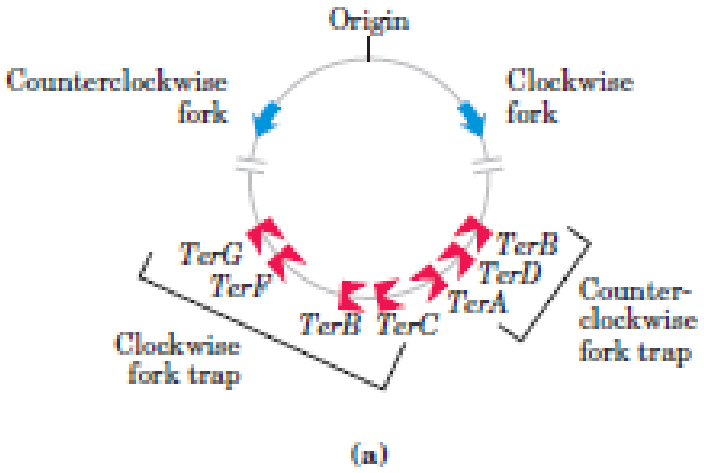

5'
3'
3'
5'
leading strand
Direction of replication fork movement
5'
3'
5'
5'
lagging strand (Okazaki fragments)
3'
3'

Elongation


Enzyme—NH3+
DNA ligase
① Adenylylation of DNA ligase
R—O—P—O—Ribose—Adenine
AMP from ATP (R = PPi) or NAD+ (R = NMN)
PPi (from ATP) or NMN (from NAD+)
Enzyme—NH2—P—O—Ribose—Adenine
Enzyme-AMP
② Activation of 5′ phosphate in nick
Enzyme—NH3+
3′
5′
5′
OH
3′
Nick in DNA
③ Displacement of AMP seals nick
O—P—O—Ribose—Adenine
AMP
DNA ligase
Ribose—Adenine
Sealed DNA

TERMINATION

- Replication of bacterial genome proceeds bi-directionally which terminates at a position diametrically opposite to the origin of replication.
- Replication terminates at the terminus region containing multiple copies of a 20bp sequence called ***Ter(terminus)*** sequences
- Ter sequence works as the binding site for protein ***Tus (terminus utilization substance)*** which stops the DnaB helicase, resulting in termination of DNA replication
- The completed chromosomes then partitioned into two daughter cells during cell division



(a)

Since a single polypeptide as small as the Pol I Klenow fragment can replicate DNA by itself, why does *E. coli* maintain a battery of 20 intricately coordinated proteins to replicate its chromosome?

THANKS

THANKS