DNA Replication

1-General Principles

2- Bacterial DNA Replication

A - The bacterial Replisome

B - Interference between Replication And Transcription

3- Eukaryotic DNA Replication

A-The eukaryotic Replisome

B- Dealing with Chromatin

C-Dealing with linear chromosomes: Telomeres and Telomerase

Learning outcomes:
What you need to know/understand after this unit

Understand the General principles of DNA replication (bidirectionality/semiconservative) and the need for RNA primers

Know the Organization of the replisomes in bacteria and eukaryotes

Understand the basic replication cycle and events on the lagging strand

Know how/why eukaryotic DNA replication in eukaryotes is different from replication in bacteria and how eukaryotes deal with the packaging of DNA in nucleosomes

Understand the chromosome end problem in eukaryotes and how telomerase works to limit this problem

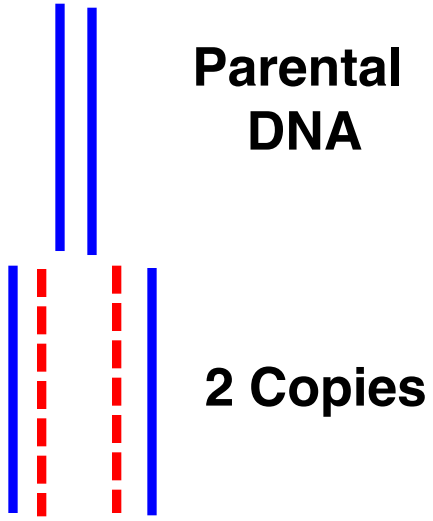
DNA Replication by DNA polymerases : Copying the genetic material to prepare for cell division

The scale of the problem:

- **The *E. coli* genome is ~5 million base pairs**
- **The human genome is ~3 BILLION base pairs**
- **If unwound, DNA in one human cell would stretch ~6 feet**
- **If all the DNA in your body was put end to end, it would reach to the sun and back over 600 times**

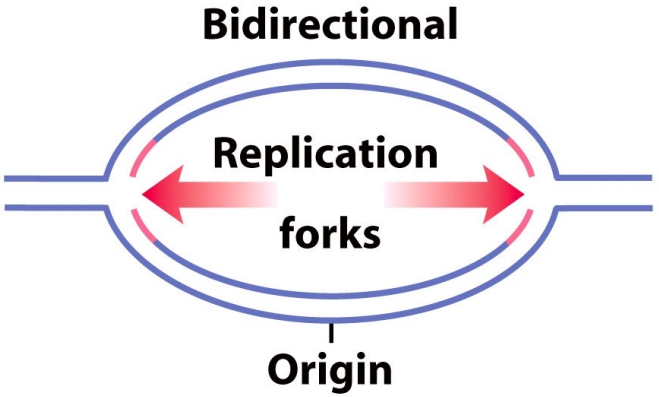
DNA Replication by DNA polymerases : Copying the genetic material to prepare for cell division

Replication is Semi-Conservative:
1 parental strand is transmitted into each daughter DNA molecule
(Meselson & Stahl experiments – you don't need to know the details)

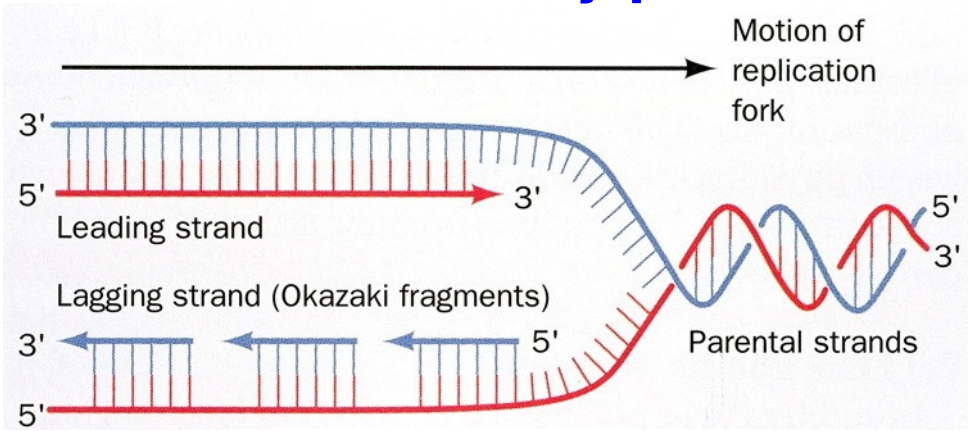


Replication is bidirectional from the Origin of Replication:
(Cairns experiments – you don't need to know the details)

- 1 origin of replication in most bacterial chromosomes
- several origins of replications in archaea, some bacteria, and in eukaryotes

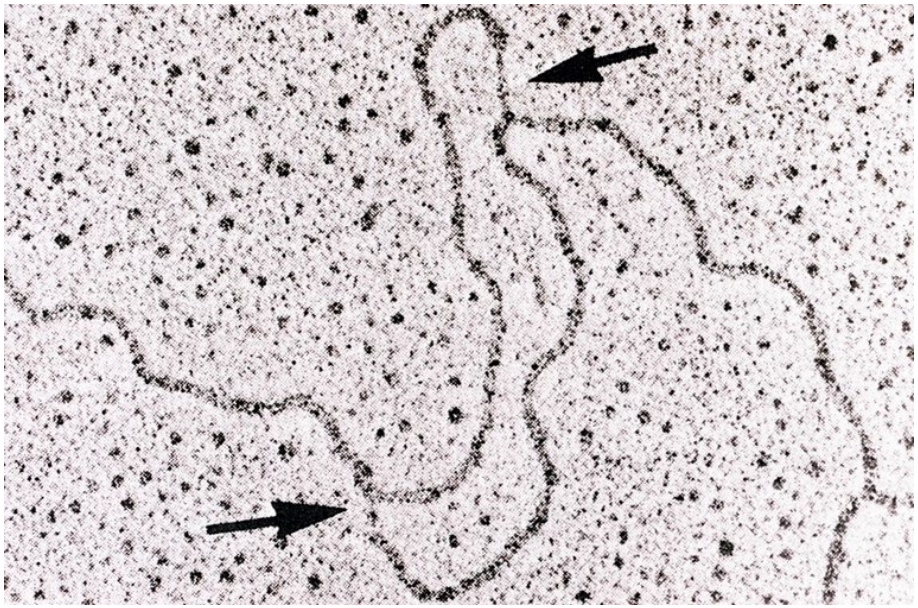


The bidirectionality problem

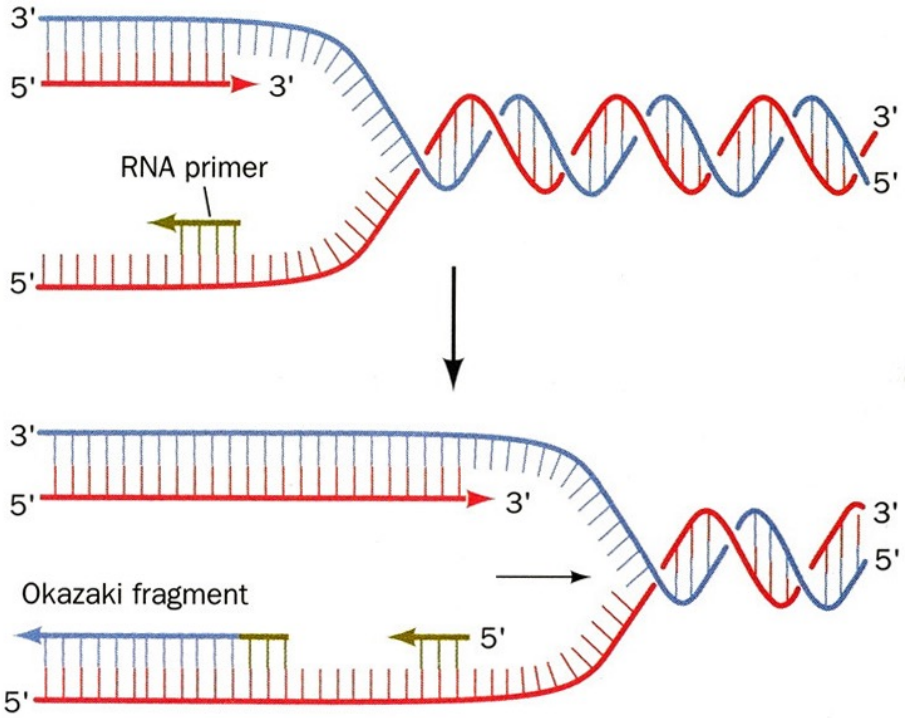


(Reminder: Polymerases work in the 5´→3´ direction)

Synthesis of DNA on the lagging strand requires continuous synthesis of primers!

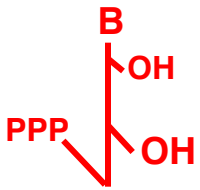


Synthesis of DNA is semi-discontinuous

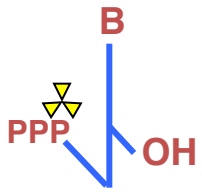


Where do the primers used during DNA replication come from ?

The primers are made of RNA since RNA polymerases do not require primers. The existence of joint RNA-DNA molecules was demonstrated by alkaline hydrolysis of Okazaki fragments.

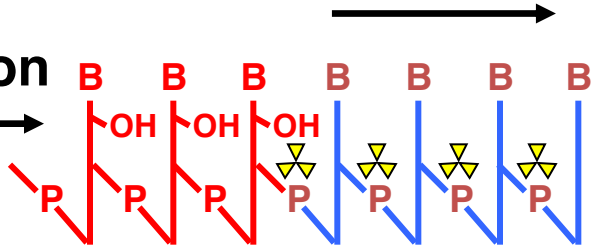


rNTPs

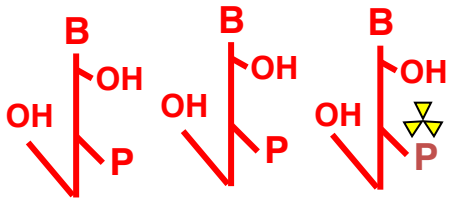


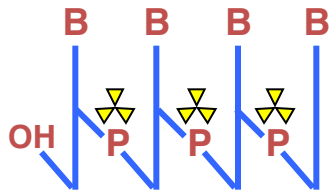
α^{32}P-dNTPs

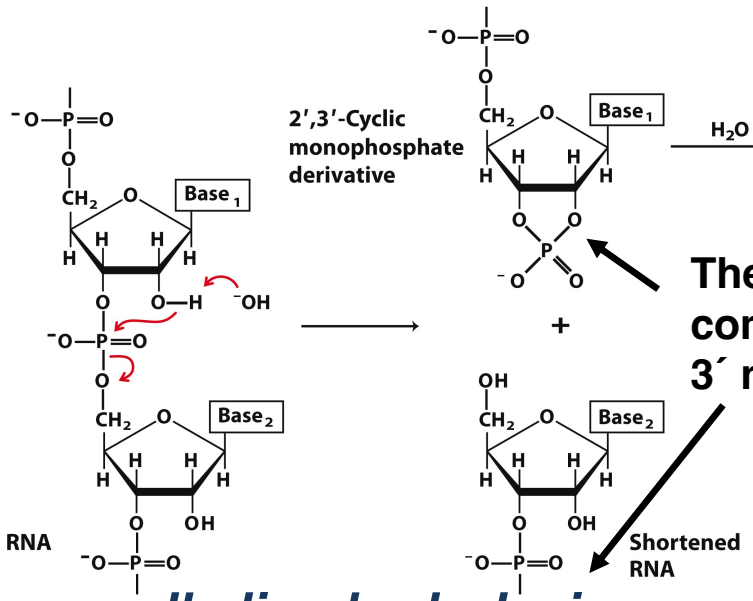
DNA replication



Alkaline hydrolysis







The phosphate comes from the 3´ nucleoside!

Ribonucleoside with a 3´P

means transfer of radiolabel from DNA-> Ribonucleotide

= indicates a 5´RNA-3´DNA junction

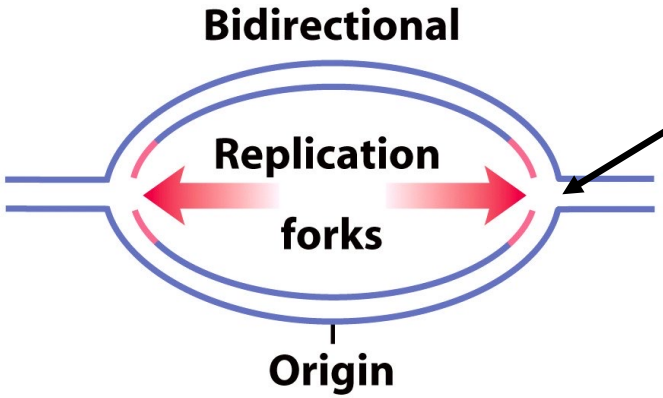
alkaline hydrolysis

(mechanism by which RNA is less stable than DNA)

Figure 8-8
Lehninger Principles of Biochemistry, Sixth Edition

Bacterial DNA replication
How does this process work in *E. coli?*

Initiation of DNA replication in *E. coli*

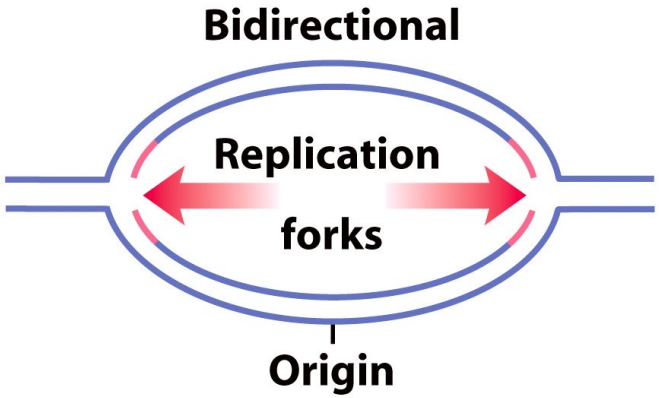


Cells need to separate or melt the 2 strands of DNA

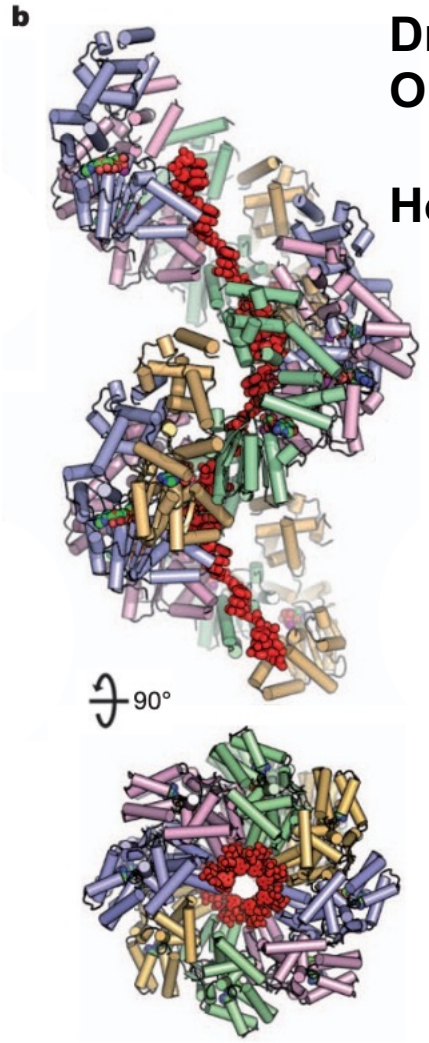
How does this process begin?

Initiation

Initiation of DNA replication in *E. coli*

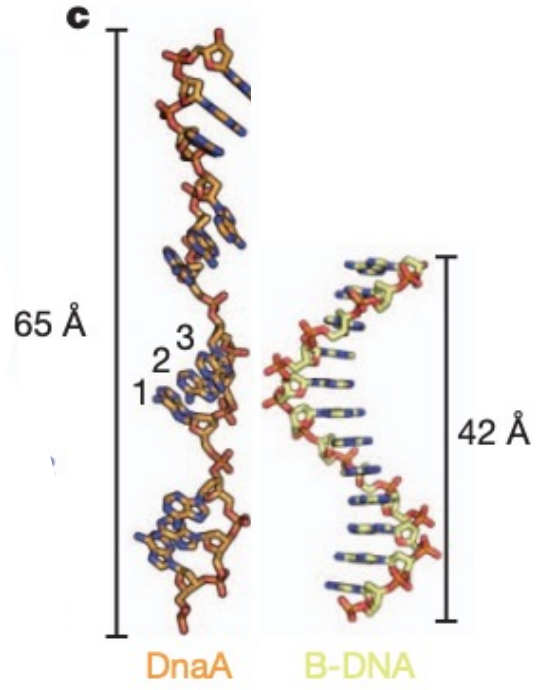


How does this process begin?



DnaA melts AT-rich DNA at Origin of Replication

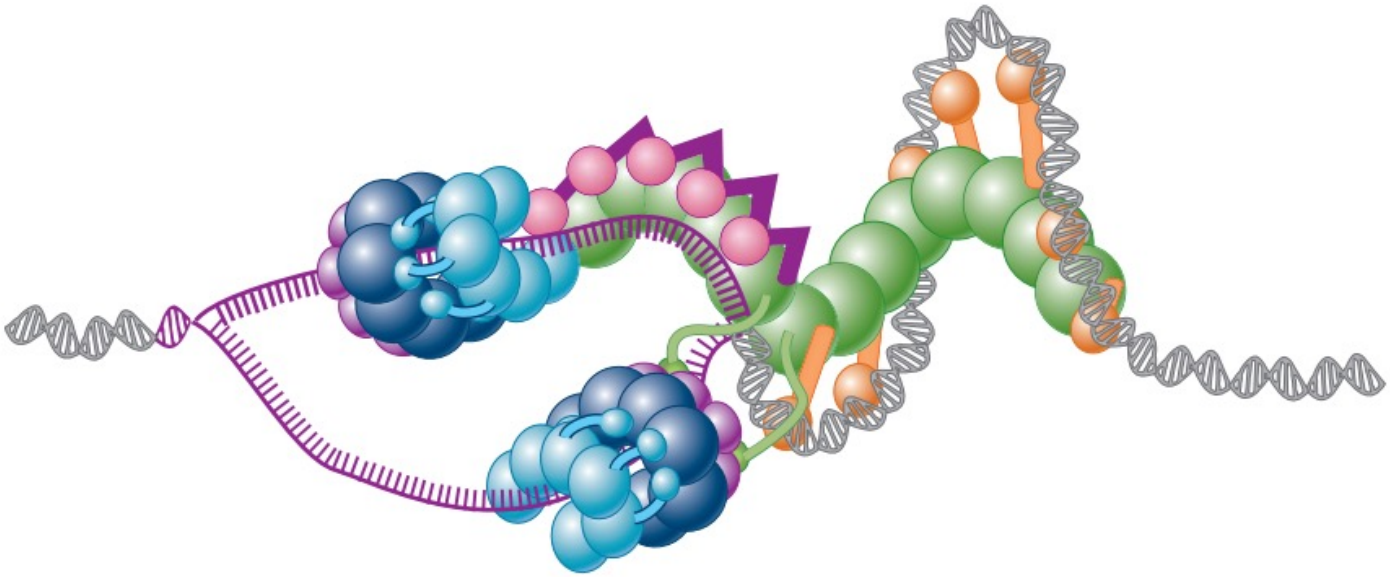
How? By stretching the DNA!



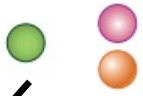
Duderstadt et al., Nature 2011

Stretching of DNA by DnaA disrupts stacking to separate strands

Initiation of DNA replication in *E. coli*



DnaA



Opens the DNA at the origin

DnaB

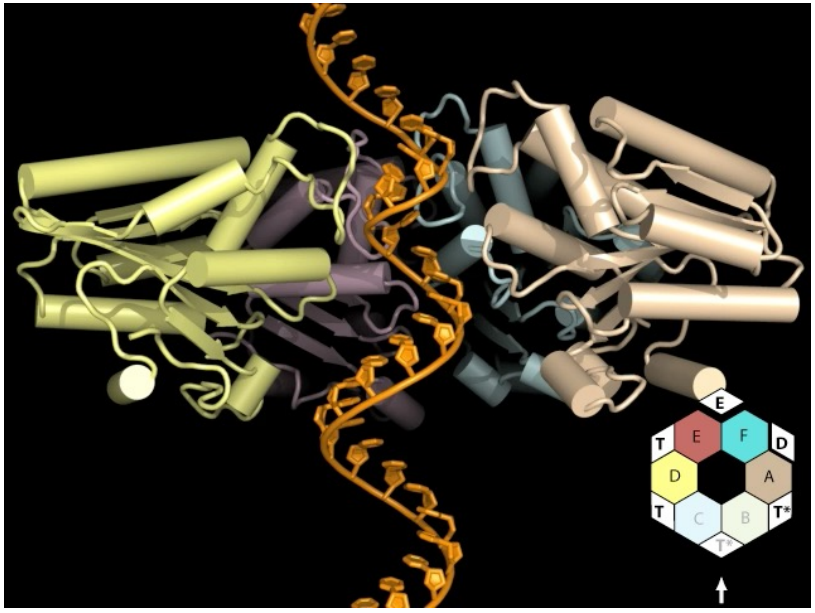
Helicase: unwinds the DNA

DnaC

Loads DnaB onto the DNA

Helicases unwind double stranded nucleic acids

Example of a helicase (similar to DnaB):



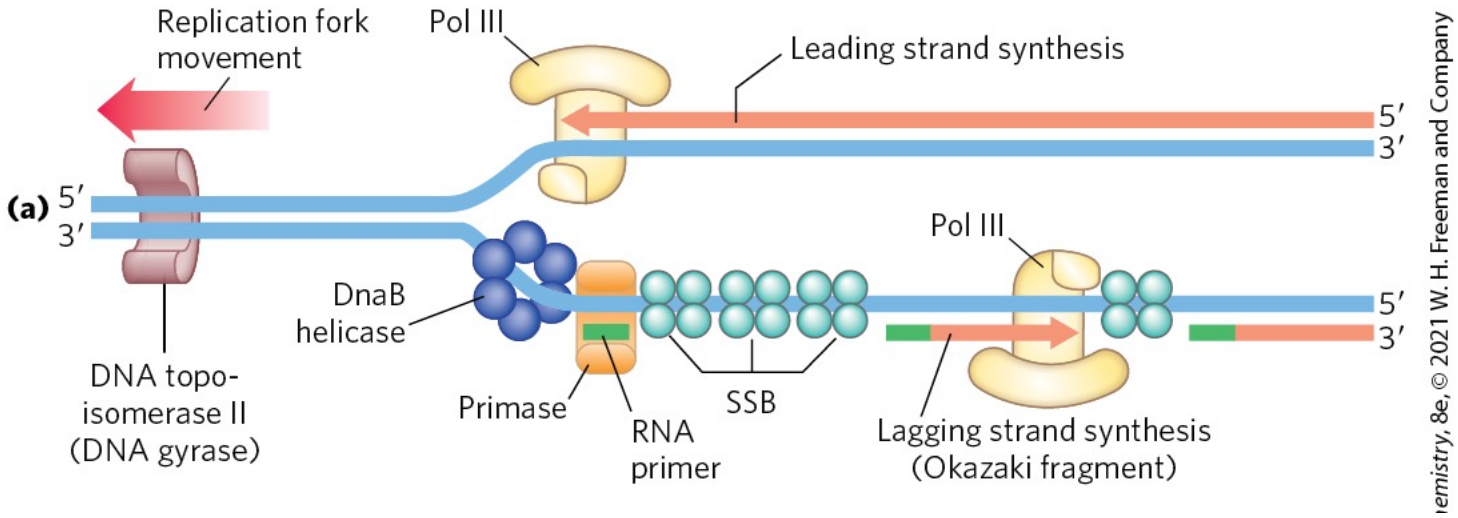
Thomsen et al., Cell 2009

- **DnaB is a homohexamer**
- **DnaC is also a homohexamer**
- **A DnaB hexamer is loaded onto each DNA strand at the origin of replication**
- **DnaB runs 5'→3' (so the two DnaB hexamers run in opposite directions)**

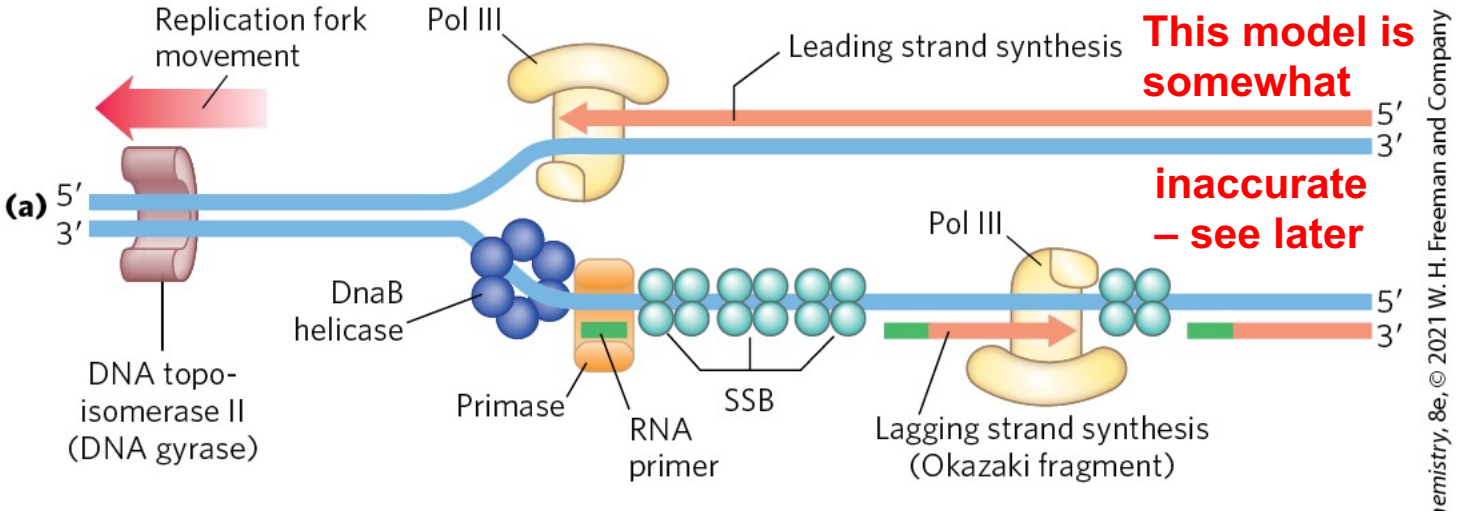
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2772833/bin/NIHMS145496-supplement-06.mov

The replisome assembles

- **Replisome = collection of proteins required for DNA replication**
- **All proteins are linked directly or indirectly to DnaB**



The replisome of *E.coli*



This model is somewhat inaccurate – see later

1) *Helicases (DnaB)*
Unwind DNA at the replication fork in a reaction coupled to ATP Hydrolysis

2) *Single-stranded DNA binding proteins (SSB)*
Bind and stabilize the DNA in a single stranded conformation after the melting by helicases

3) *Primase Enzyme (DnaG)*
Synthesizes RNA primers of the lagging strand.

4) *DNA Polymerase III (plus β subunits)*:
The polymerase which synthesizes most of the DNA during replication

6) *DNA topoisomerase II*
Relaxes supercoiled DNA that forms ahead of the replication fork. https://www.youtube.com/watch?v=HyP0cEbqKTc

7) *Rnase H*
Removes RNA primers

8) *DNA Polymerase I*
Removes RNA primers
Replaces RNA primers with DNA by nick translation

8) *DNA Ligase*
Joins the Okazaki fragments

Not shown on the model

Why is it necessary to remove the RNA primers from the Okazaki fragments?

A: Because DNA ligase cannot ligate the RNA primers to DNA

B: Because these RNA primers would be used by RNA polymerase to generate mRNAs from Okazaki fragments

C: Because the presence of RNA embedded in DNA would cause problems when this strand is replicated next, as DNA polymerases won't use a template strand with RNA

D: Because the presence of RNA embedded in DNA might lead to strand breaks

A more detailed view of the Bacterial Replisome

β Subunits = Sliding Clamp

Homodimer (β clamp)

DNA Polymerase III Core

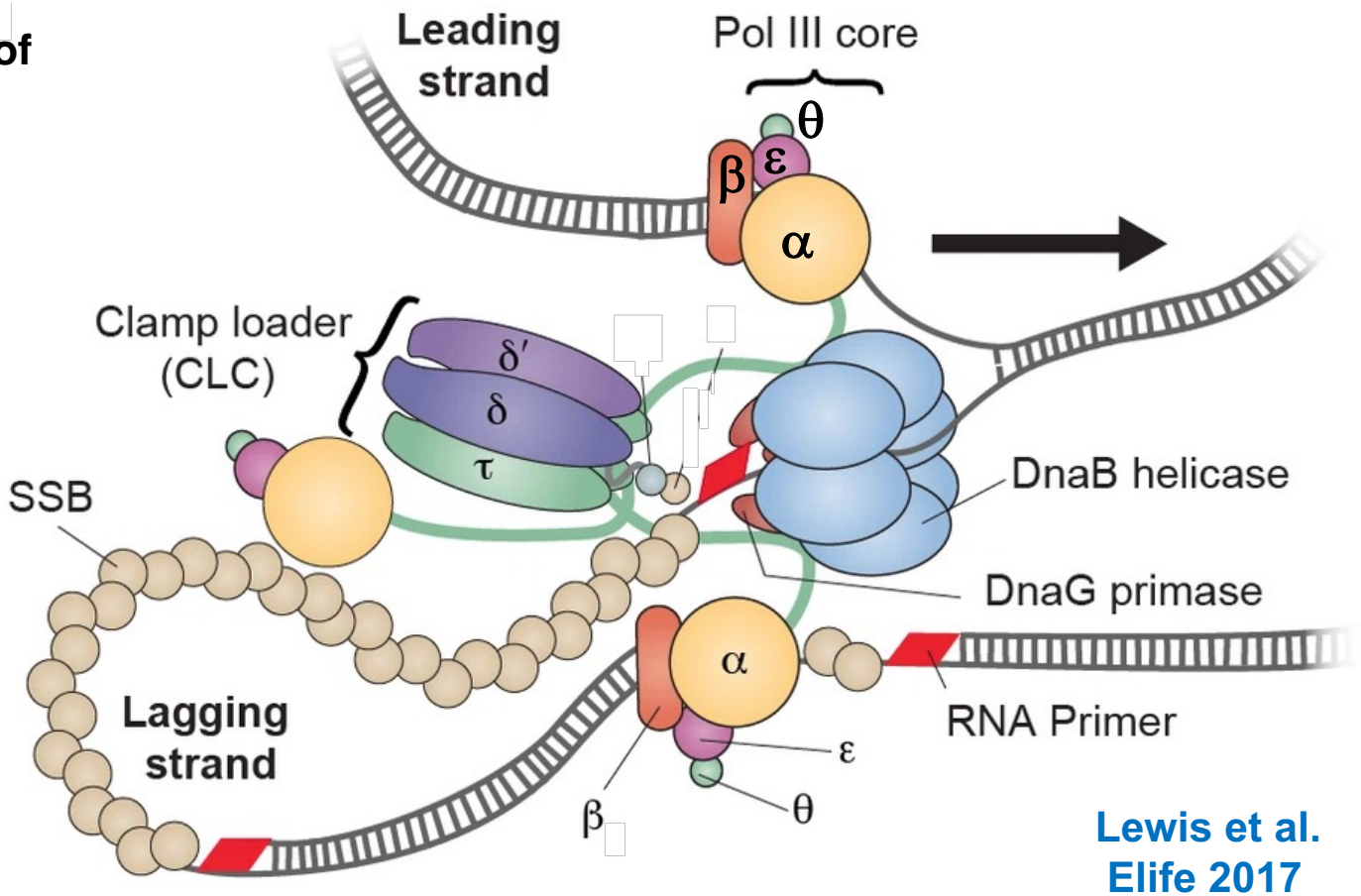
- **α subunit = Catalytic site for polymerization; no 5´→3´ or 3´→5´ exo**
- **ε subunit = 3´→5´ editing exonuclease**
- **θ subunit = structural role**

Clamp loader Complex = CLC

ATP-dependent loading of the β clamp onto DNA and unloading (δ/δ´)

τ Subunit= Dimerization factor

Holds the Pol. III cores together



Lewis et al. Elife 2017

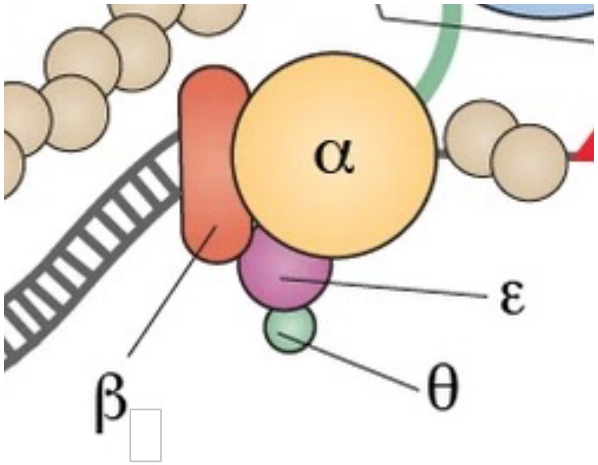
Why is there no 5´→3´ exo activity for Pol.III?

A: Because there is no need for it as it doesn't make mistakes

B: Because it is not involved in processing Okazaki fragments

C: Because it is not involved in DNA repair

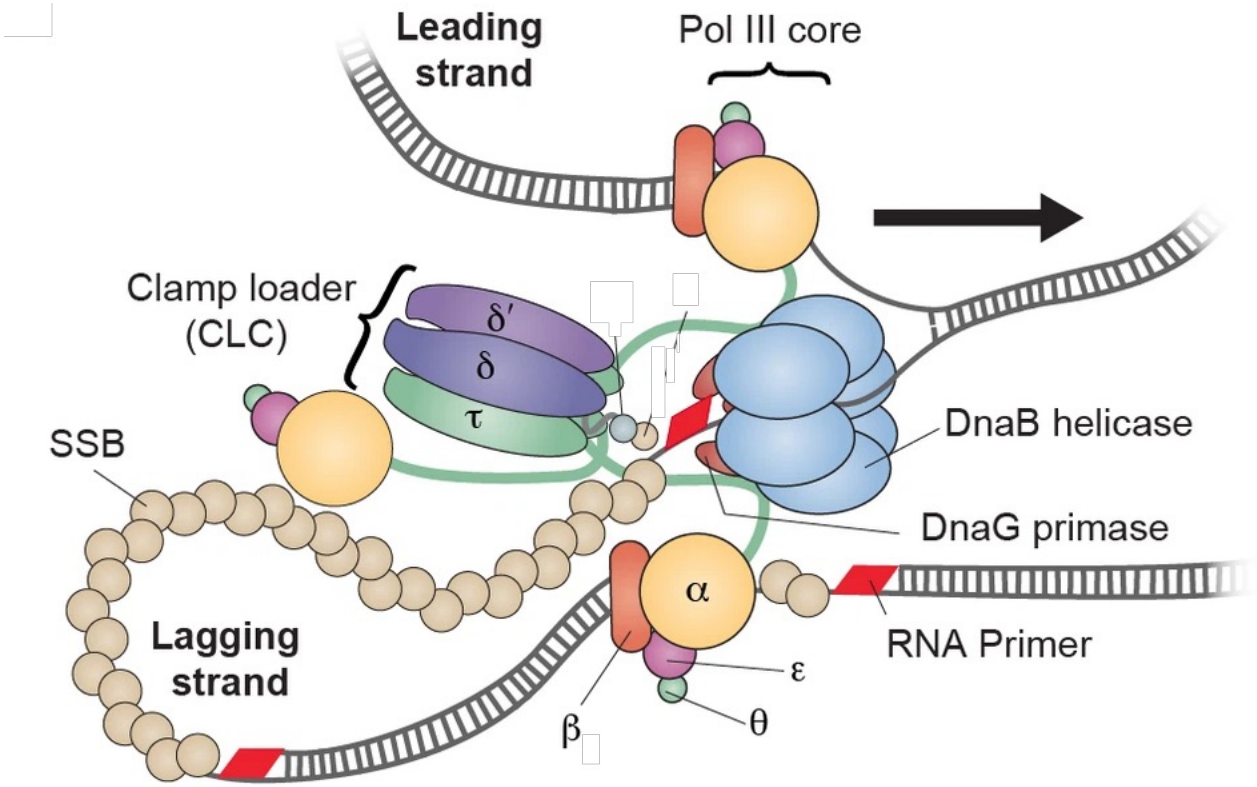
D: Because the 5´→3´exo is in the tau subunit



- α(130 kD)= Catalytic site for polymerization no 5´-3´ or 3´-5´ exo
- ε(27.5 kD) = 3´->5´ editing exonuclease
- θ - 10 kD = structural role ?

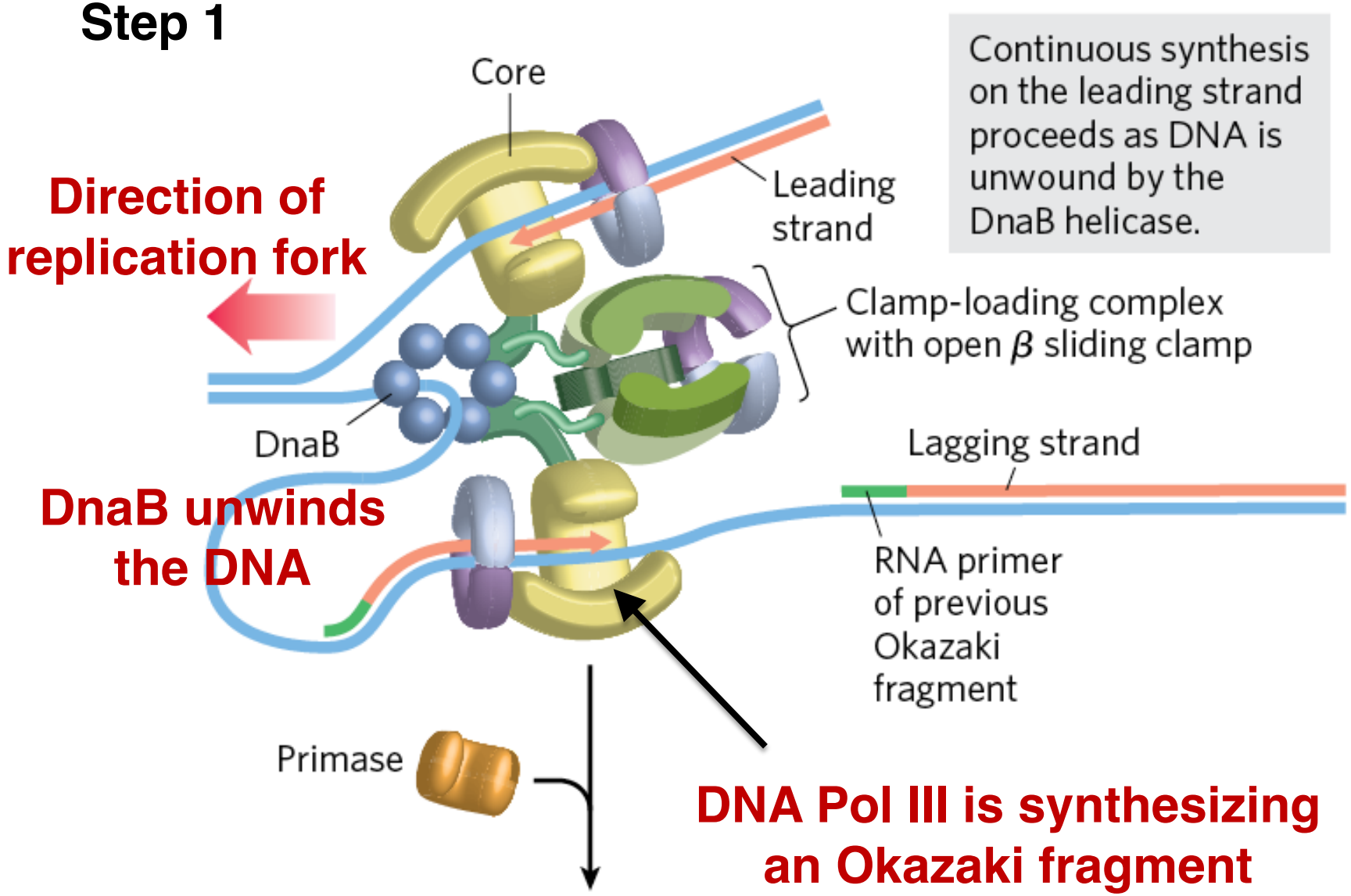
Leading and lagging strand synthesis are intricately coordinated.

Both DNA Pol III complexes are physically linked to each other!



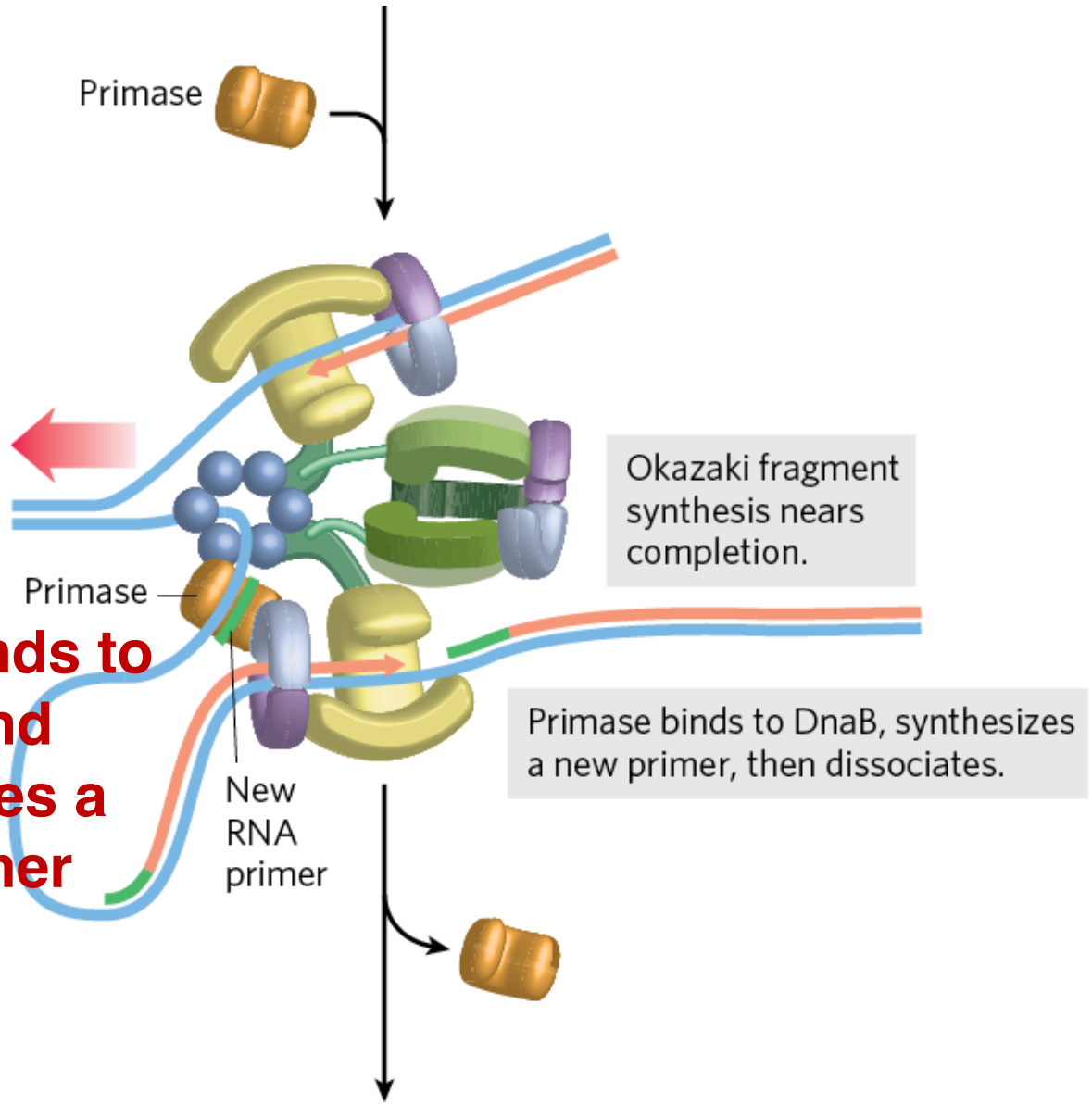
DNA synthesis on the leading and lagging strands

Step 1



DNA synthesis on the leading and lagging strands

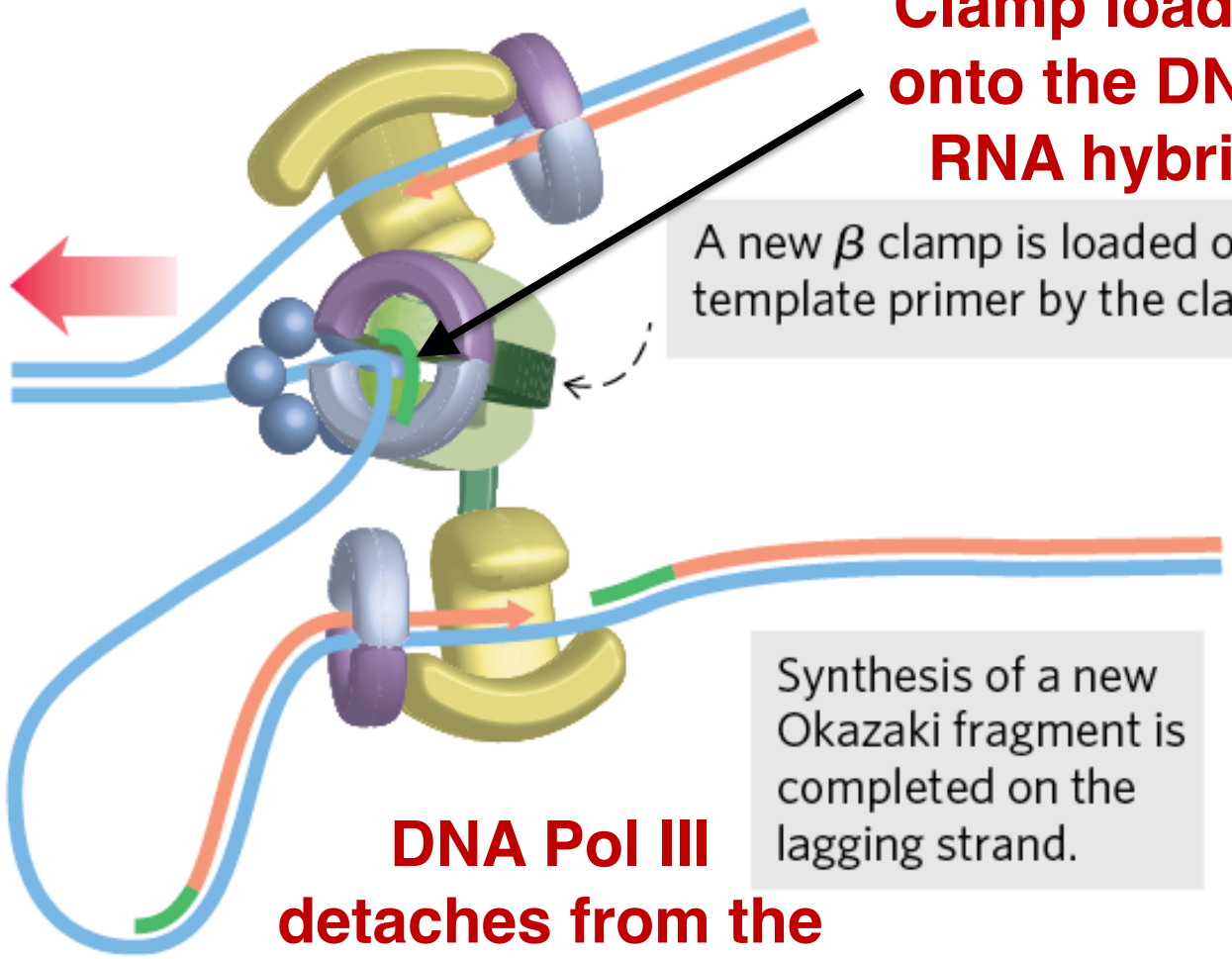
Step 2



Primase binds to DnaB and synthesizes a new primer

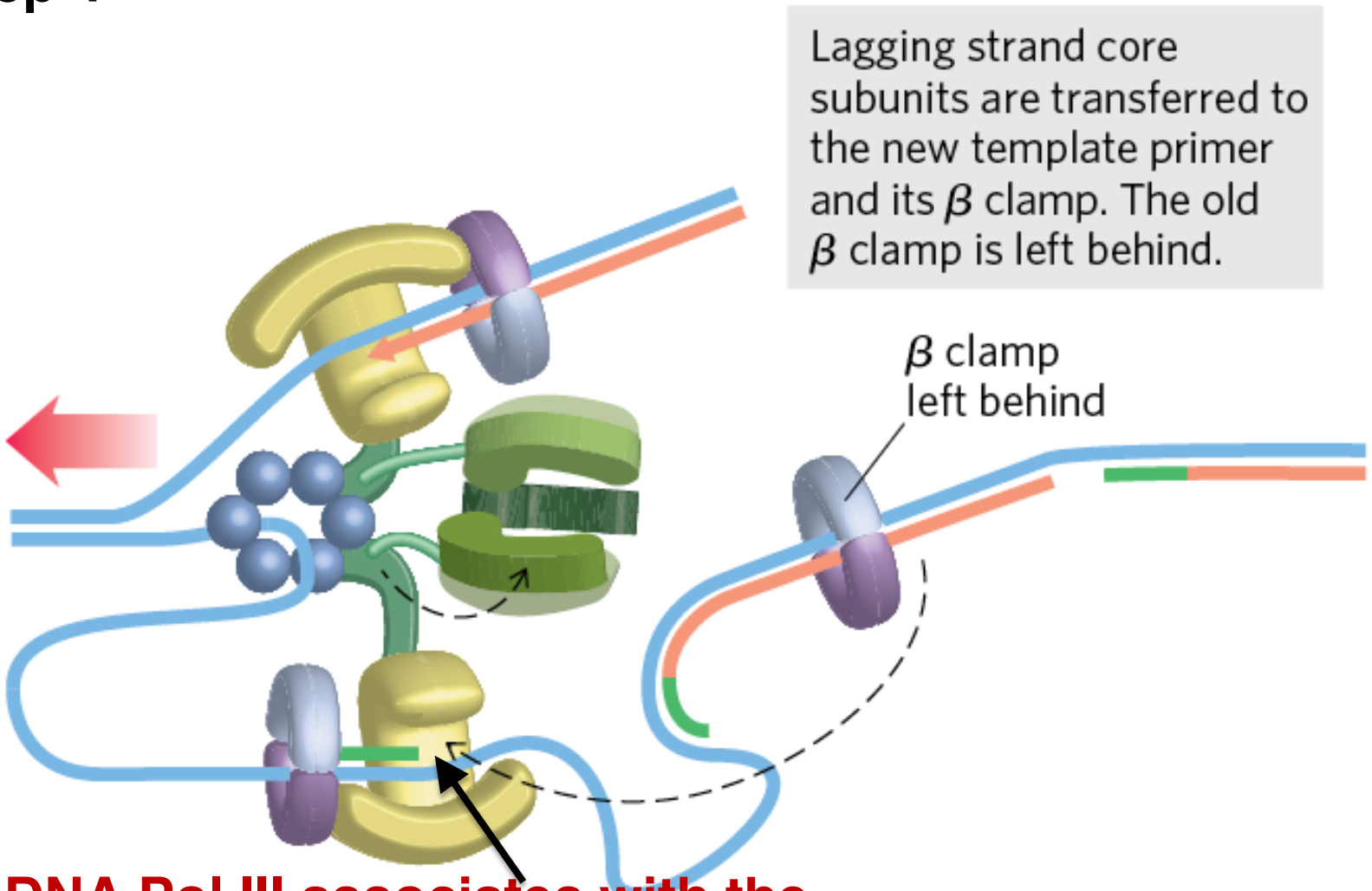
DNA synthesis on the leading and lagging strands

Step 3



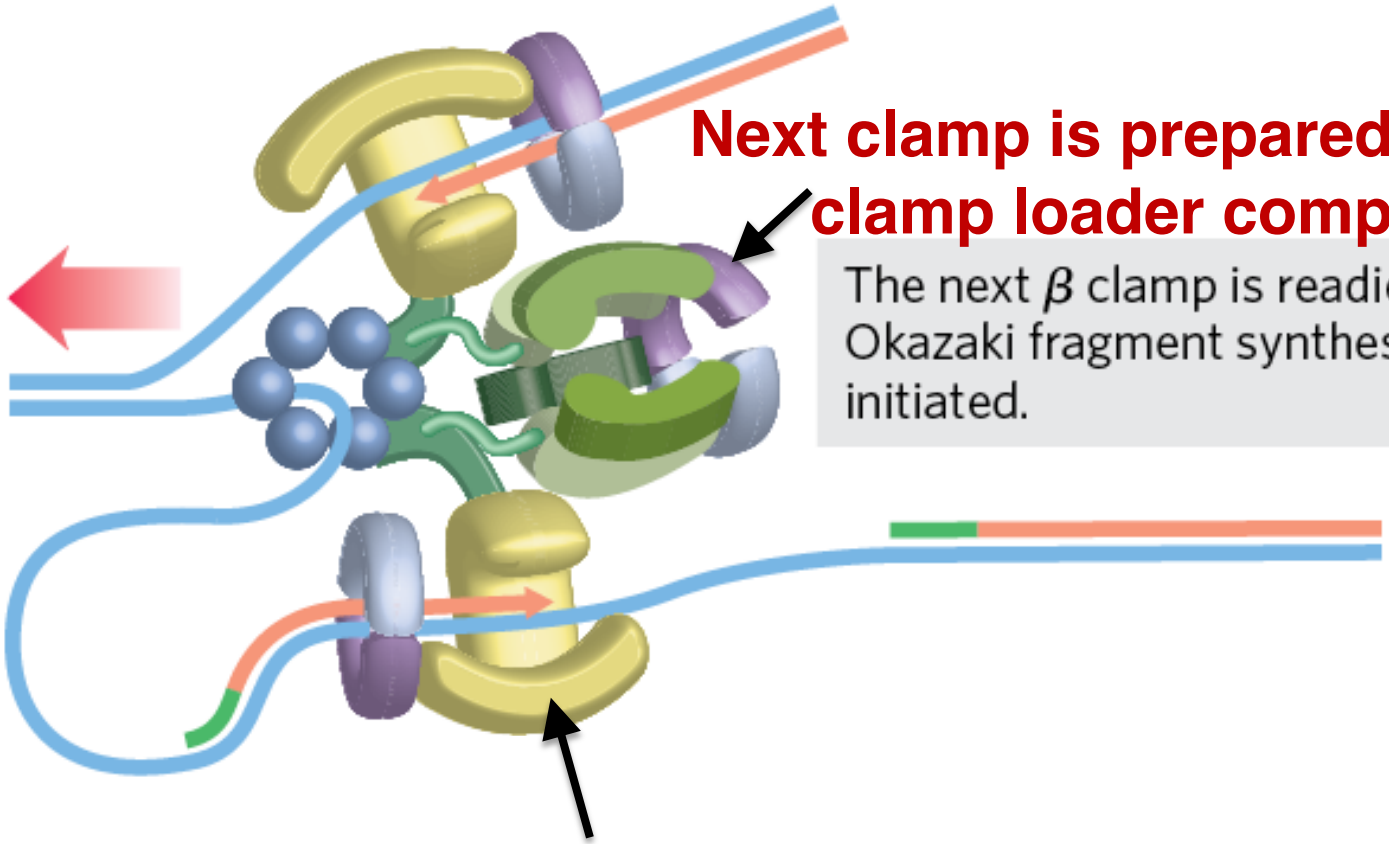
DNA synthesis on the leading and lagging strands

Step 4

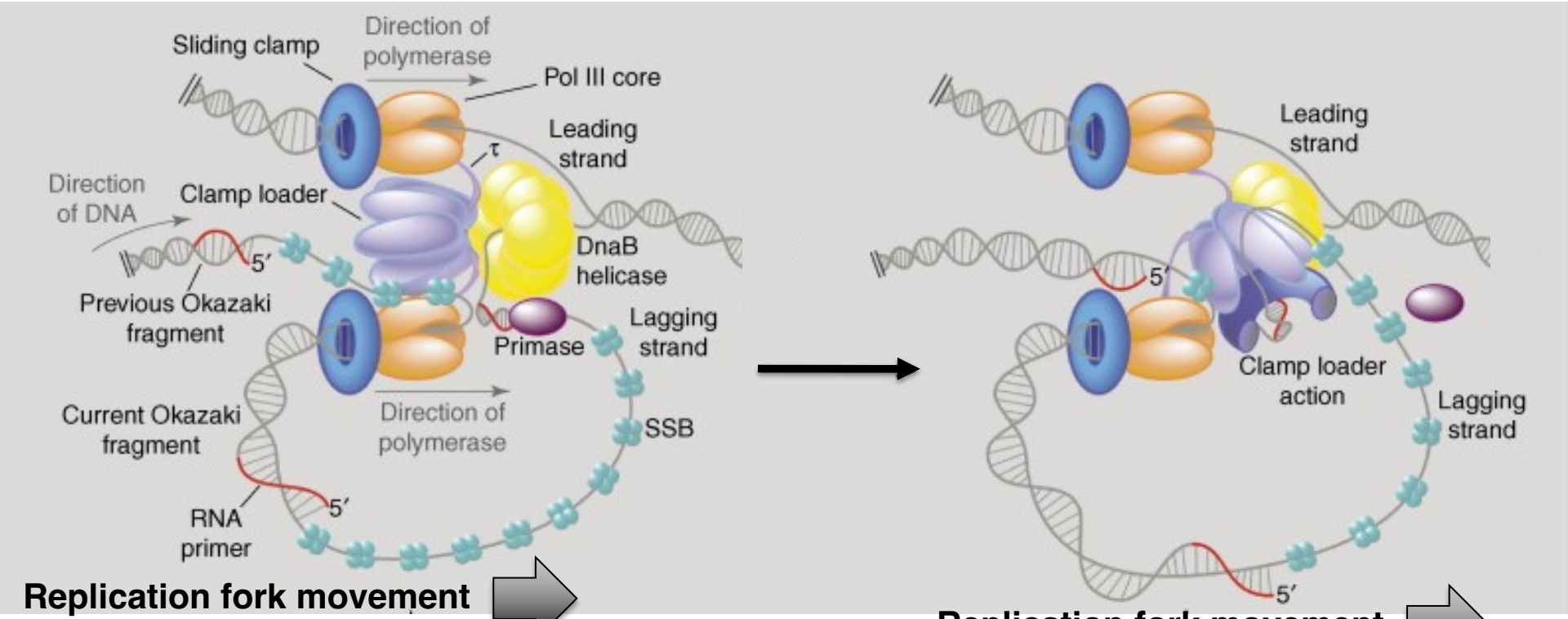


DNA synthesis on the leading and lagging strands

Step 5



The replisome of *E.coli* in action: Cycles of molecular events during Lagging Strand synthesis (1)

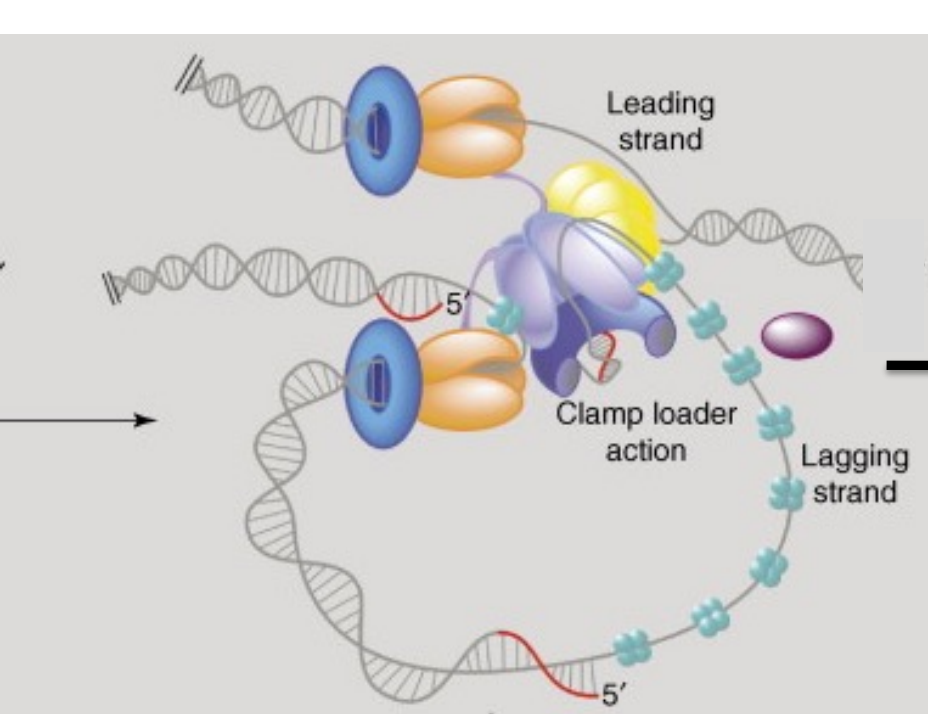


Step 1: The primase synthesizes a new RNA primer upstream in the lagging strand; the two polymerase replicate DNA

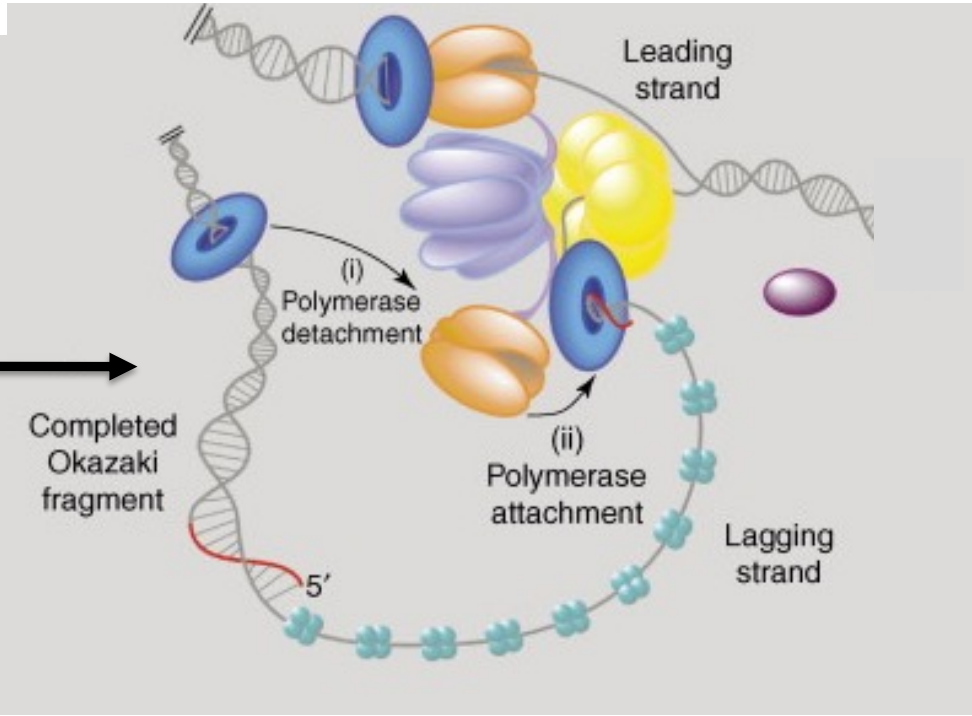
Step 2: A sliding clamp is assembled around the new RNA primer; primase dissociated

Trends in Microbiology 15, 155-64 (2007)

The replisome of *E.coli* in action: Cycles of molecular events during Lagging Strand synthesis (2)



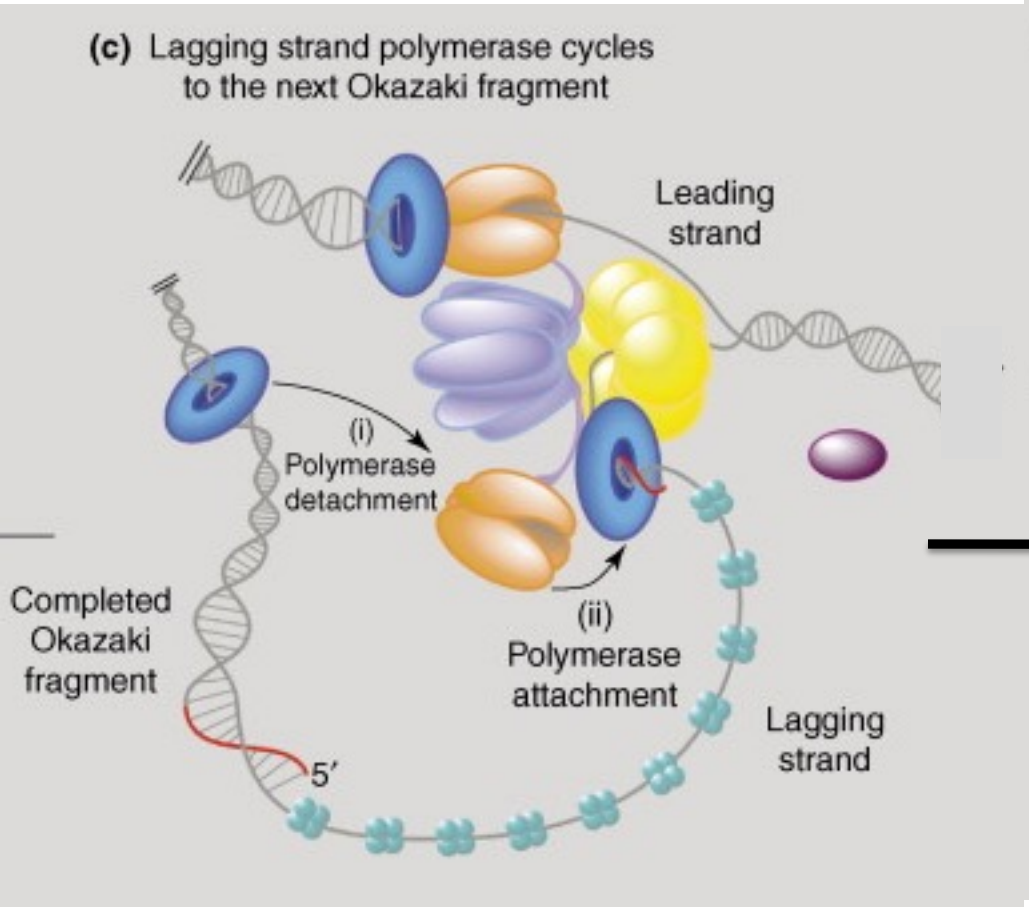
Step 2: A sliding clamp is assembled around the new RNA primer; primase dissociated

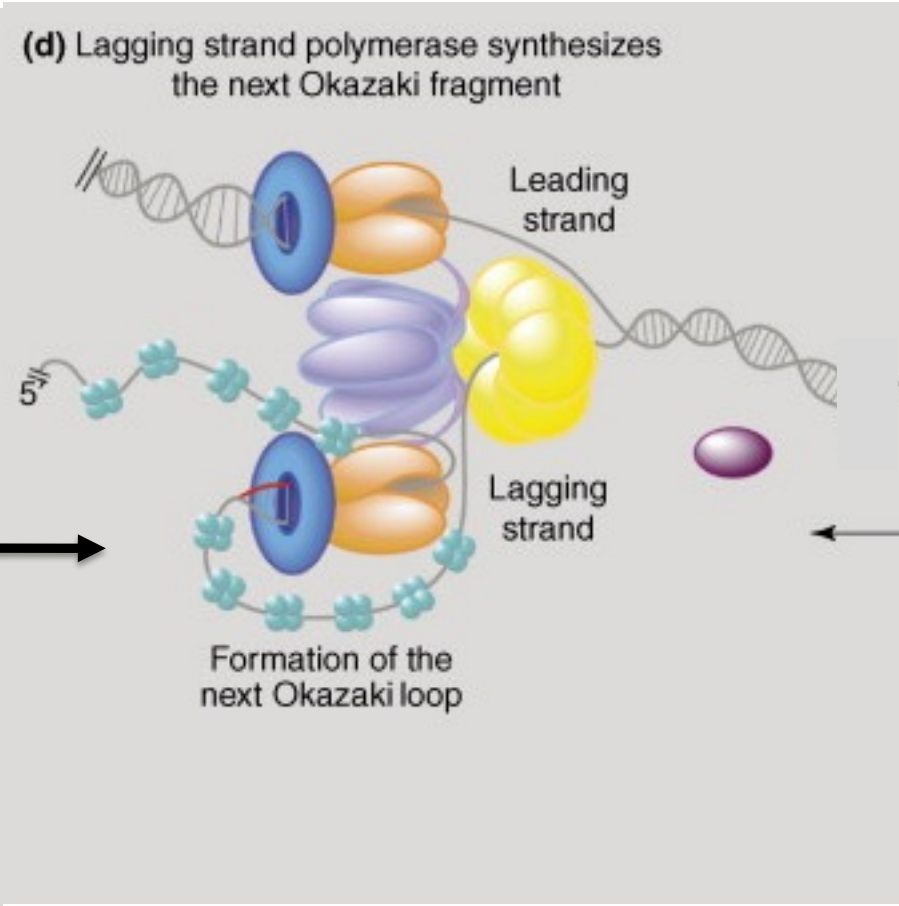


Step 3: the lagging strand polymerase detaches and associates with the newly deposited sliding clamp

Trends in Microbiology 15, 155-64 (2007)

The replisome of *E.coli* in action: Cycles of molecular events during Lagging Strand synthesis (3)





Step3: the lagging strand polymerase detaches and associates with the newly deposited sliding clamp

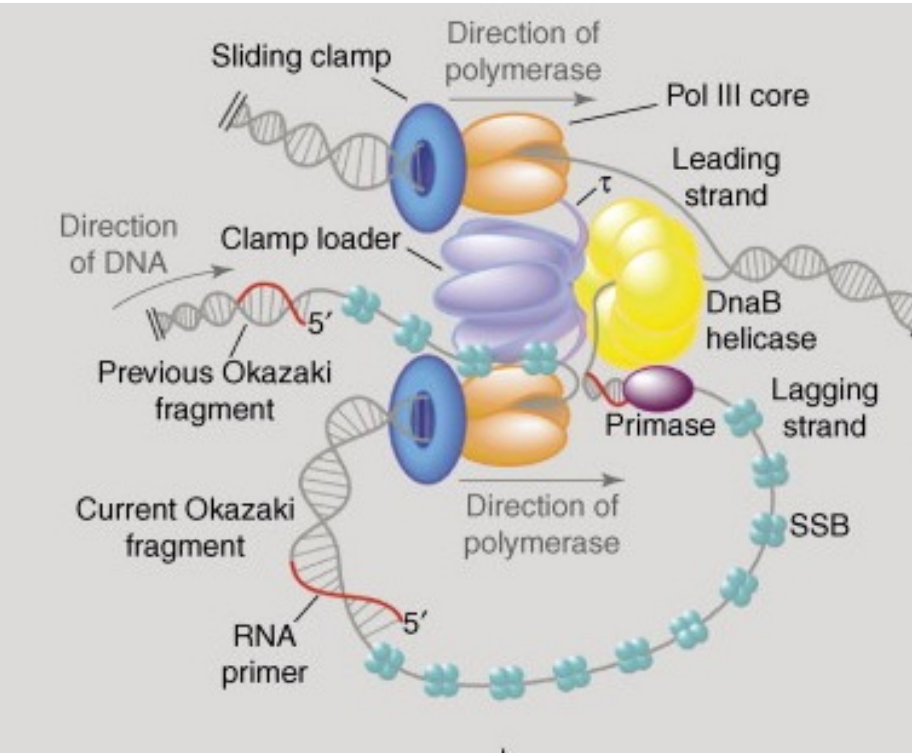
Step 4: the lagging strand polymerase start to synthesize the next Okazaki fragment; primase will reinitiate synthesis of the next RNA primer

→ ***back to Step1***

***Trends in Microbiology* 15, 155-64 (2007)**

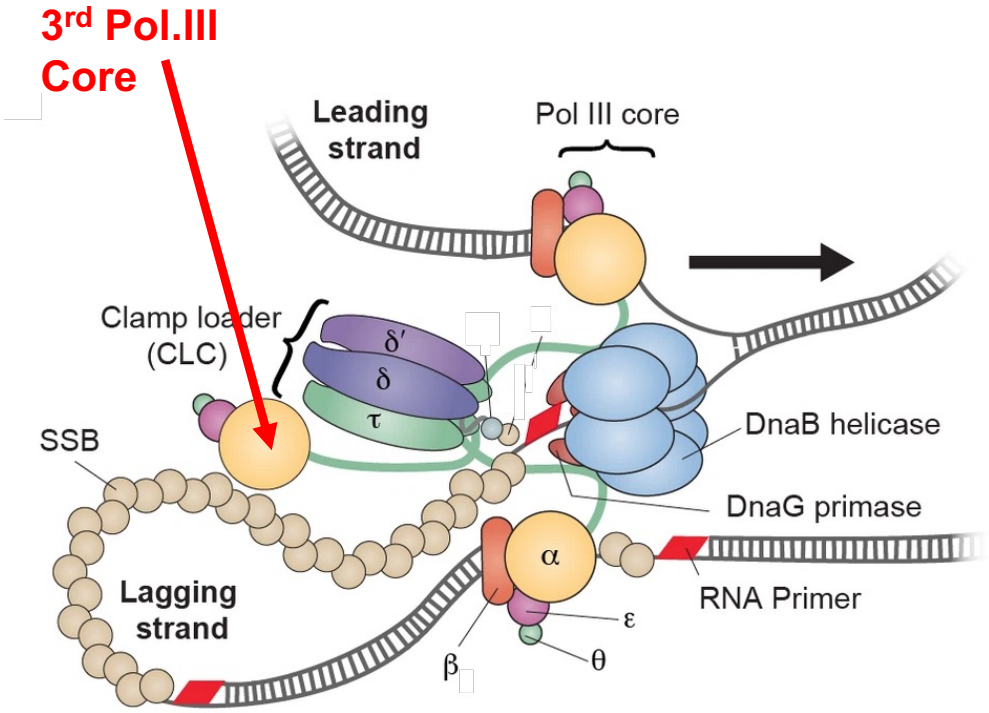
A Revision of previous models: there are 3 core DNA polymerases associated with most replisomes in vivo

Classical Model



***Trends in Microbiology* (2007)**

New Model

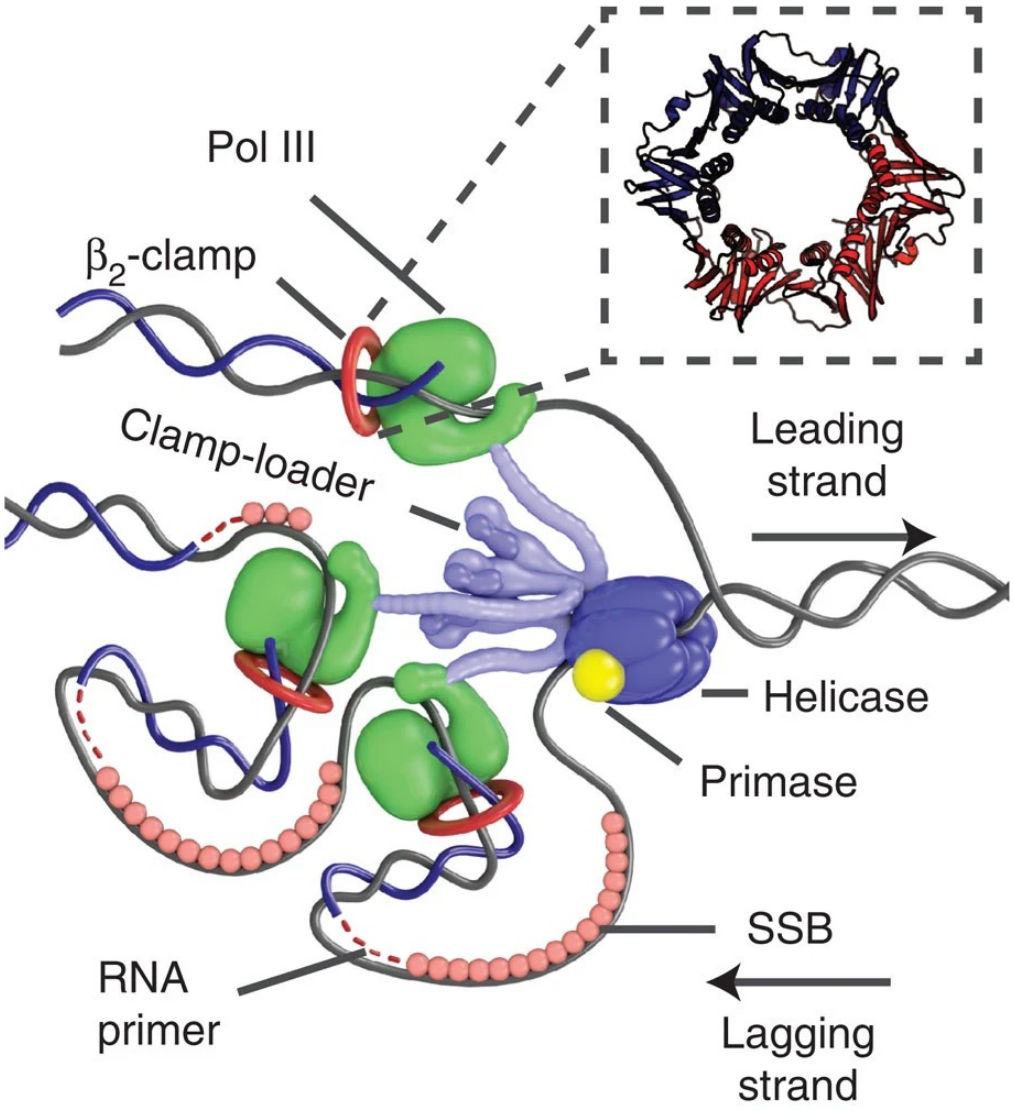


Lewis et al. Elife 2017

Great animation of this process:
https://www.youtube.com/watch?v=IjVLhoyfGAM

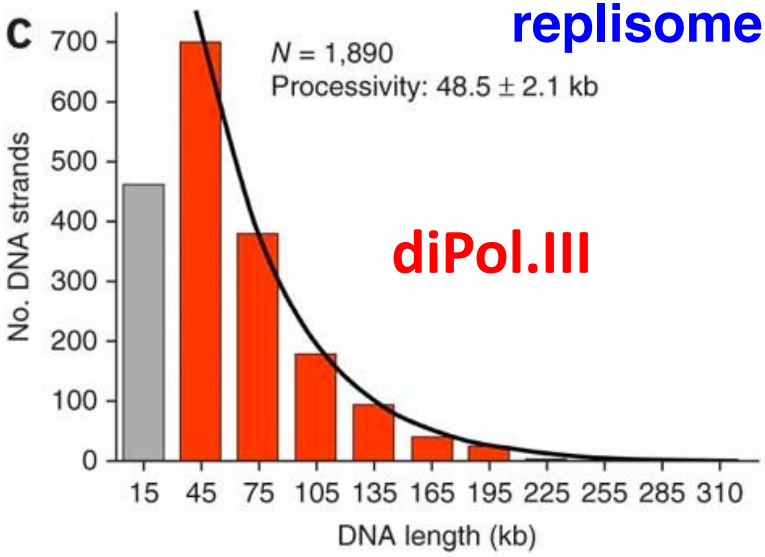
What is the advantage of 3 Polymerases vs 2 Polymerases ?

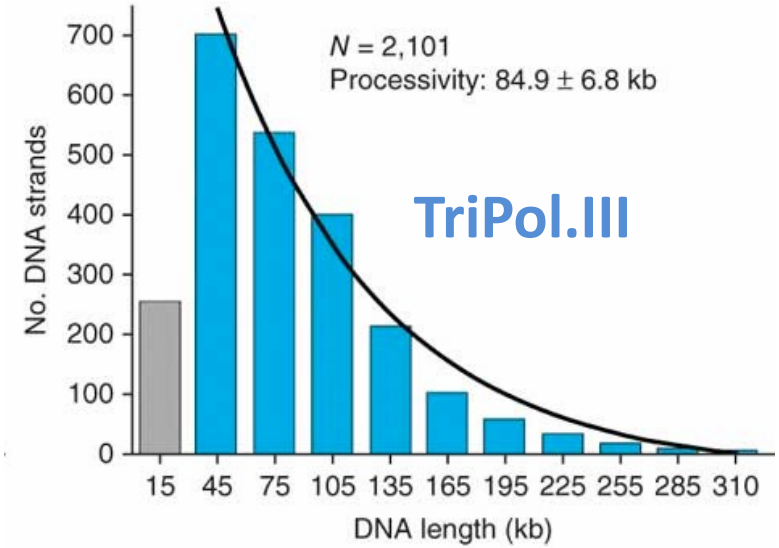
Another view of the replisome

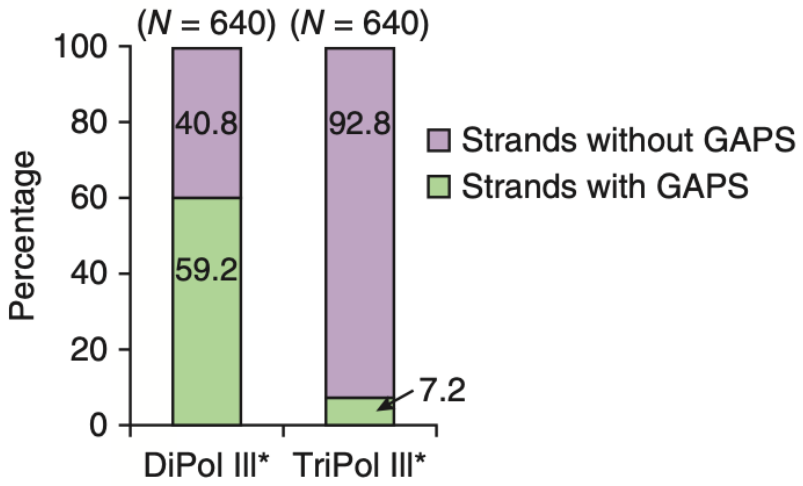


https://www.nature.com/articles/ncomms6820

Why a tri-Polymerase replisome ?







• TriPol.III replisome leaves less gaps to be filled by Pol.I in the lagging strand

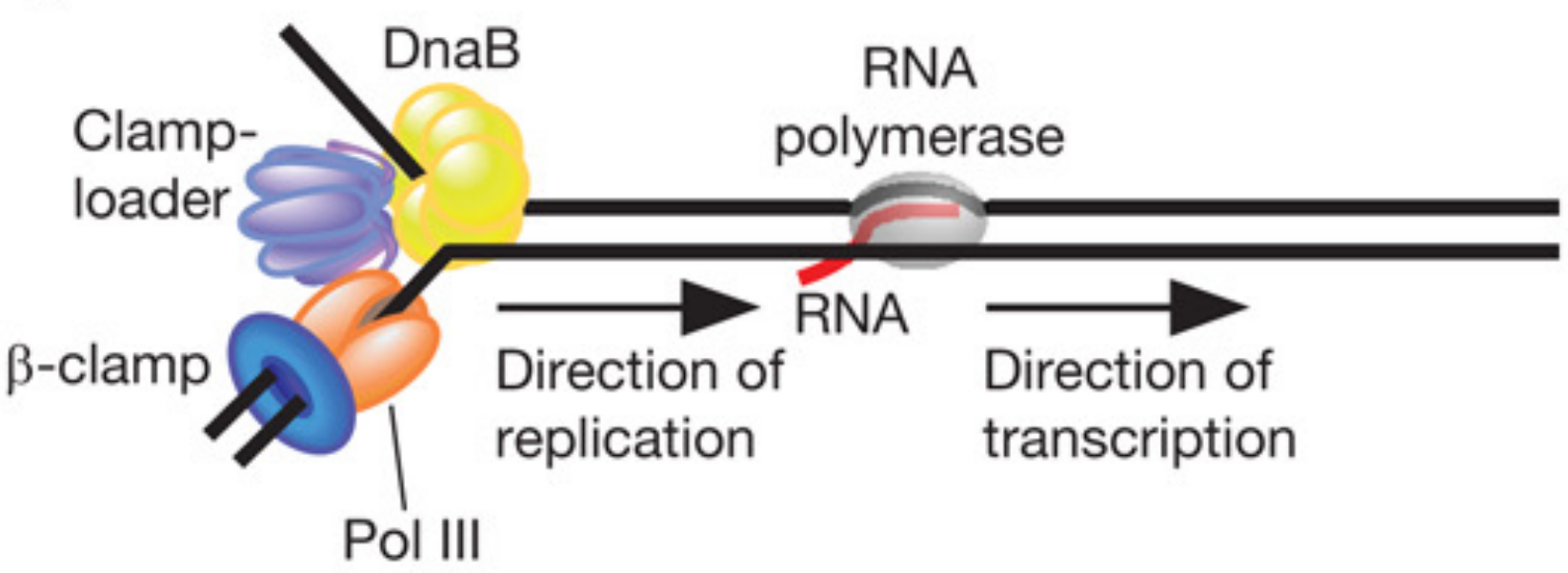
DNA Replication is overall more efficient

•Distribution of DNA fragment length shows that triPol.III synthesizes longer DNA strands -> TriPol.III replisome is more processive

Nature Structural & Molecular Biology
Volume: 19, Pages: 113–116 (2012)

Problem of Coordinated Nucleic Acids Synthesis in vivo:

What happens when DNA Polymerases and RNA polymerases collide ?
(speed of DNA polymerase >> speed of RNA polymerase)



Nature
456, 762-66
(2008)

Are DNA Polymerase<> RNA polymerases collisions problematic in bacteria?

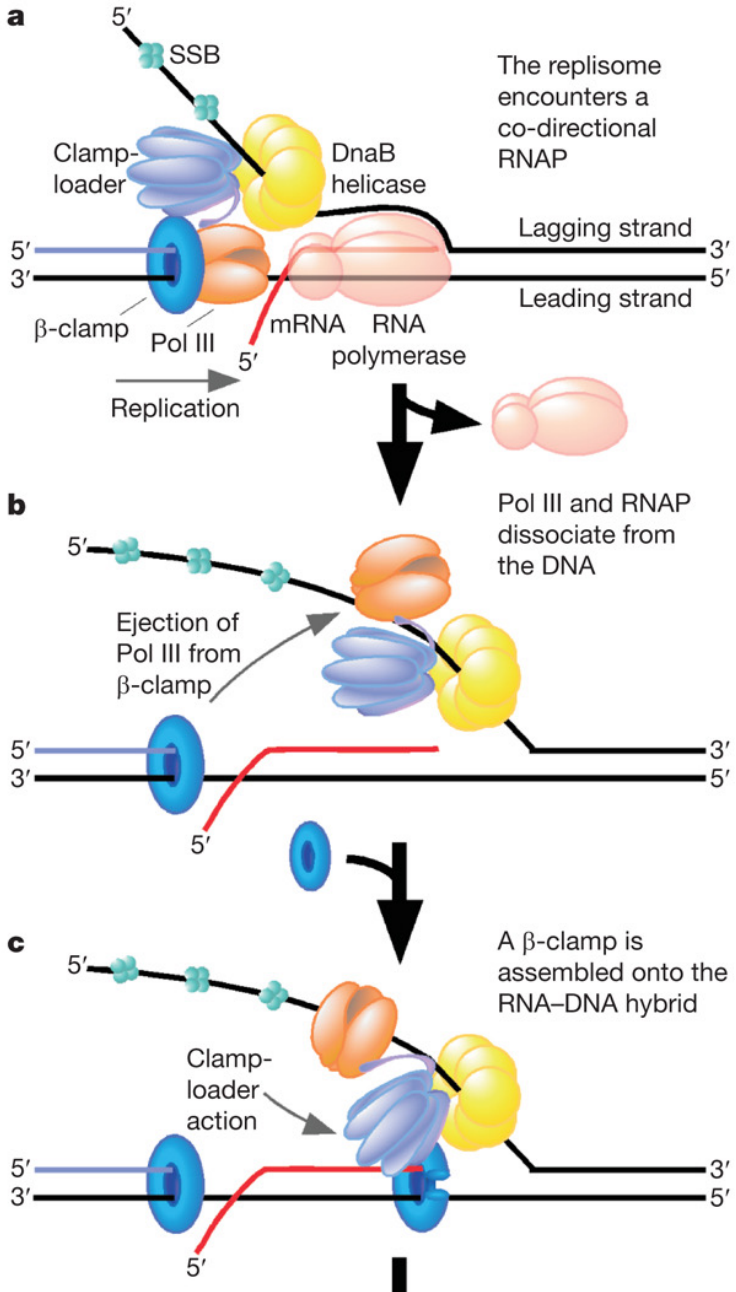
A: Yes because they would lead to replication blocks

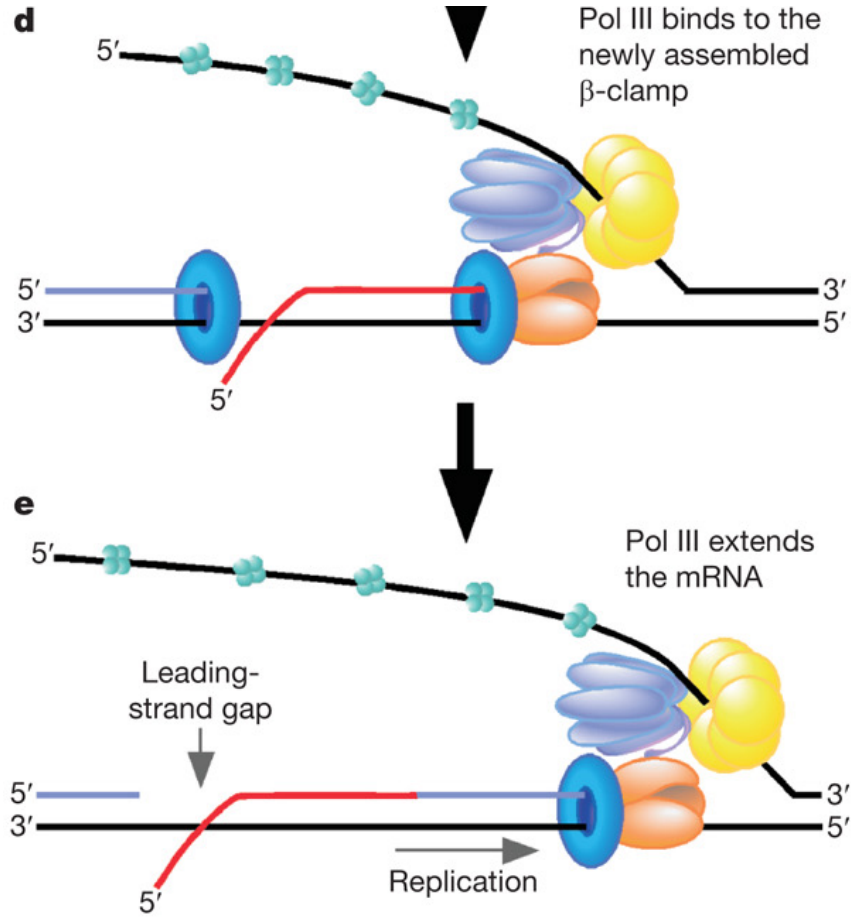
B: No because they never happen in vivo as transcription is inhibited when replication occurs

C: Yes because this leads to the release of truncated RNAs due to premature transcription termination and therefore truncated proteins

D: No because transcription and replication occur on different strands

Collision between DNA Polymerases and RNA polymerases result in polymerases dissociation and in the use of the RNA synthesized by RNA Polymerase as a primer for replication





Nature 456, 762-66 (2008)

Eukaryotic DNA replication

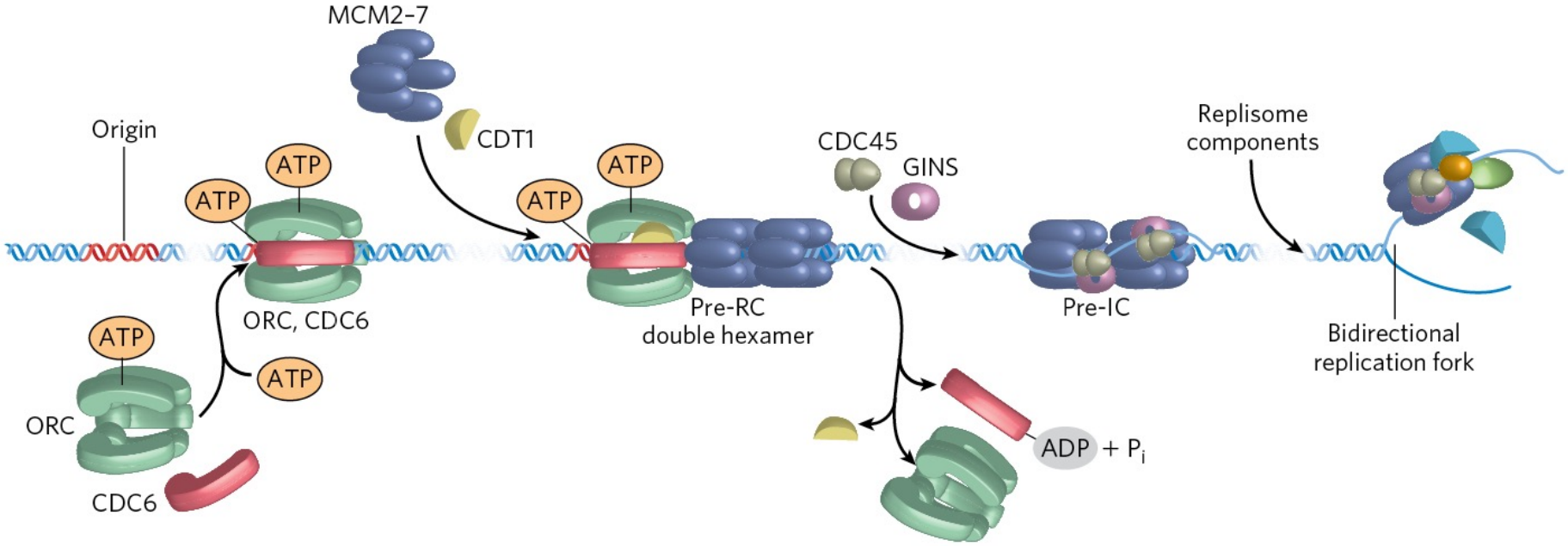
Many similarities to bacterial DNA replication, but more complicated

Eukaryotic DNA Replication

- **Machinery is overall similar to that used for bacterial DNA replication (names are different…)**

- **Idiosyncrasies of eukaryotic DNA replication are linked to unique feature of eukaryotic genomes:**
  - **Large size of eukaryotic chromosomes and limited time for DNA synthesis**
  - **Replication machinery needs to deal with nucleosome packaging of eukaryotic DNA**
  - **Problem of linear chromosomes**

Initiation of DNA replication in eukaryotes

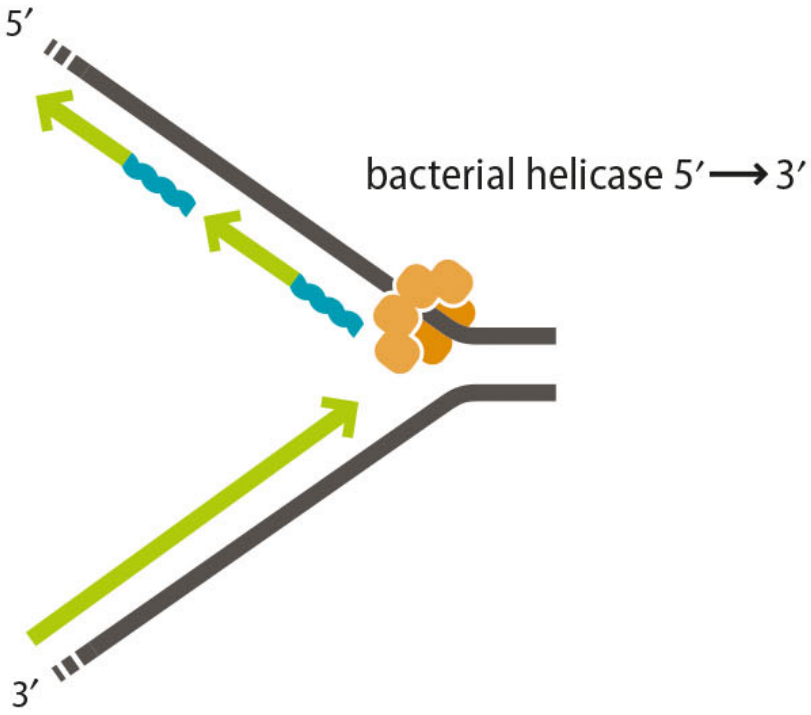


1. ORC (origin recognition complex) binds to the origin (ORC functionally similar to DnaA)
2. CDC6 binds, recruits CDT1
3. ORC, CDC6, and CDT1 load two inactive MCM2-7 helicase complexes (MCM2-7 functionally similar to DnaB) → this is the pre-replicative complex (pre-RC)
4. CDC45 and GINS bind to and activate MCM2-7
5. Replisome components bind and bidirectional replication begins

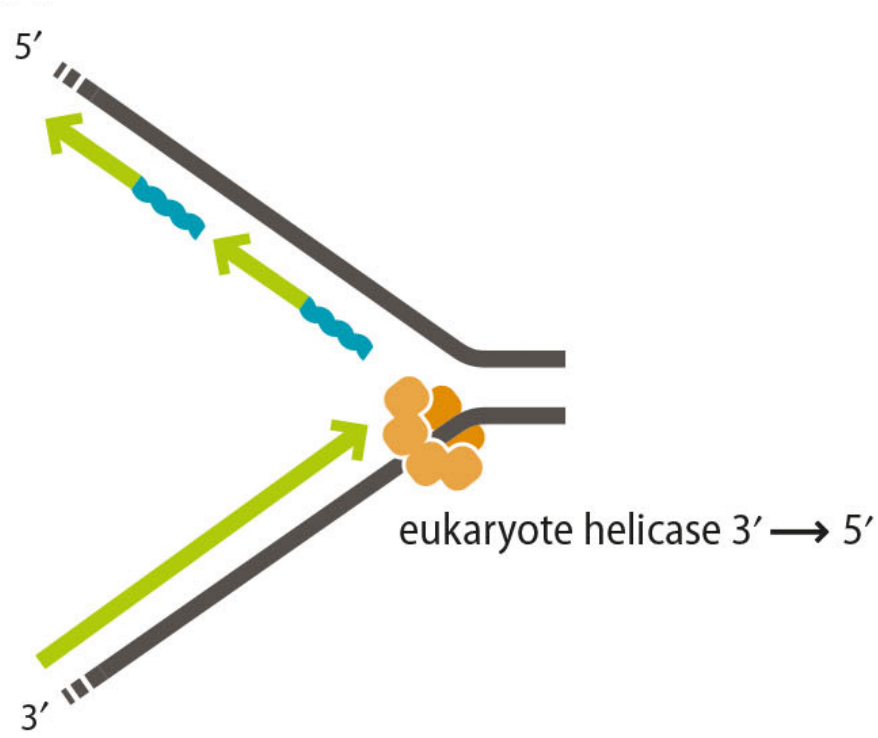
Components of the eukaryotic replisome

- **RPA (replication protein A):** stabilizes single-stranded DNA. Equivalent of bacterial SSB.
- **PCNA:** sliding clamp (trimer). Equivalent of β protein/β clamp.
- **RFC (replication factor C)**: clamp loader, consists of 5 protein subunits Rfc1-5. Equivalent of the clamp loader complex.
- **Pol α-primase complex:** synthesizes an RNA primer, then elongates it with a short stretch of DNA. Roughly equivalent to DnaG in bacteria (but in bacteria the primase is not bound to a DNA polymerase)
- **Fen1:** endonuclease involved in Okazaki fragment maturation.
- **DNA ligase:** joins nicks
- **Pol ε:** Synthesizes the leading strand
- **Pol δ:** Synthesizes the lagging strand

Bacterial and eukaryotic helicases run in opposite directions

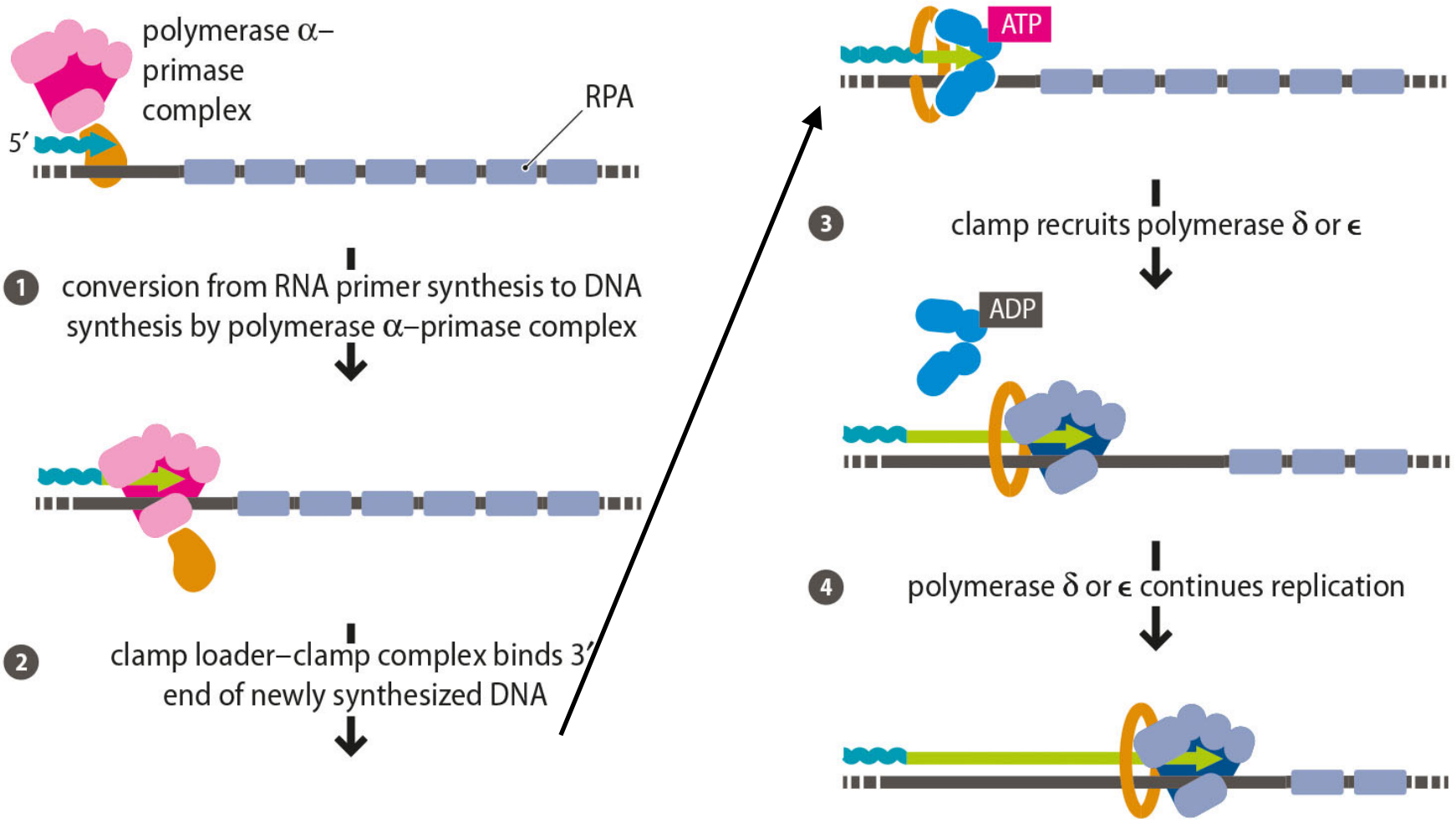


Bacterial helicase = DnaB



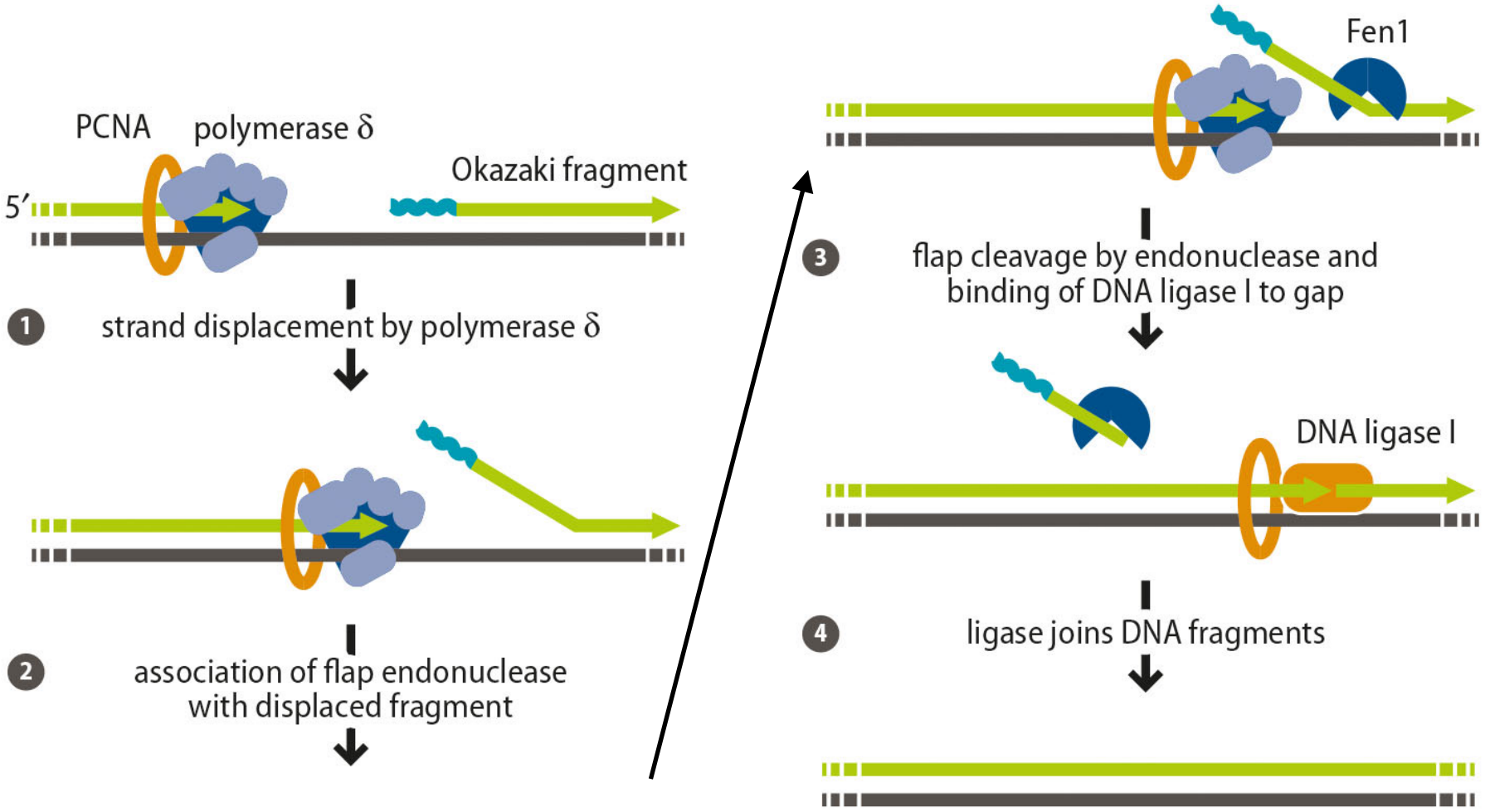
Eukaryotic helicase = MCM2-7

How do eukaryotes begin synthesizing a new strand of DNA?



<u>Polymerase switching</u> occurs each time a new DNA strand is started (on the leading or lagging strand)

Okazaki fragment maturation in eukaryotes



Architecture of the eukaryotic replisome is similar to that of the bacterial replisome

Pol ε - Replicates leading strand Pol δ - Replicates lagging strand

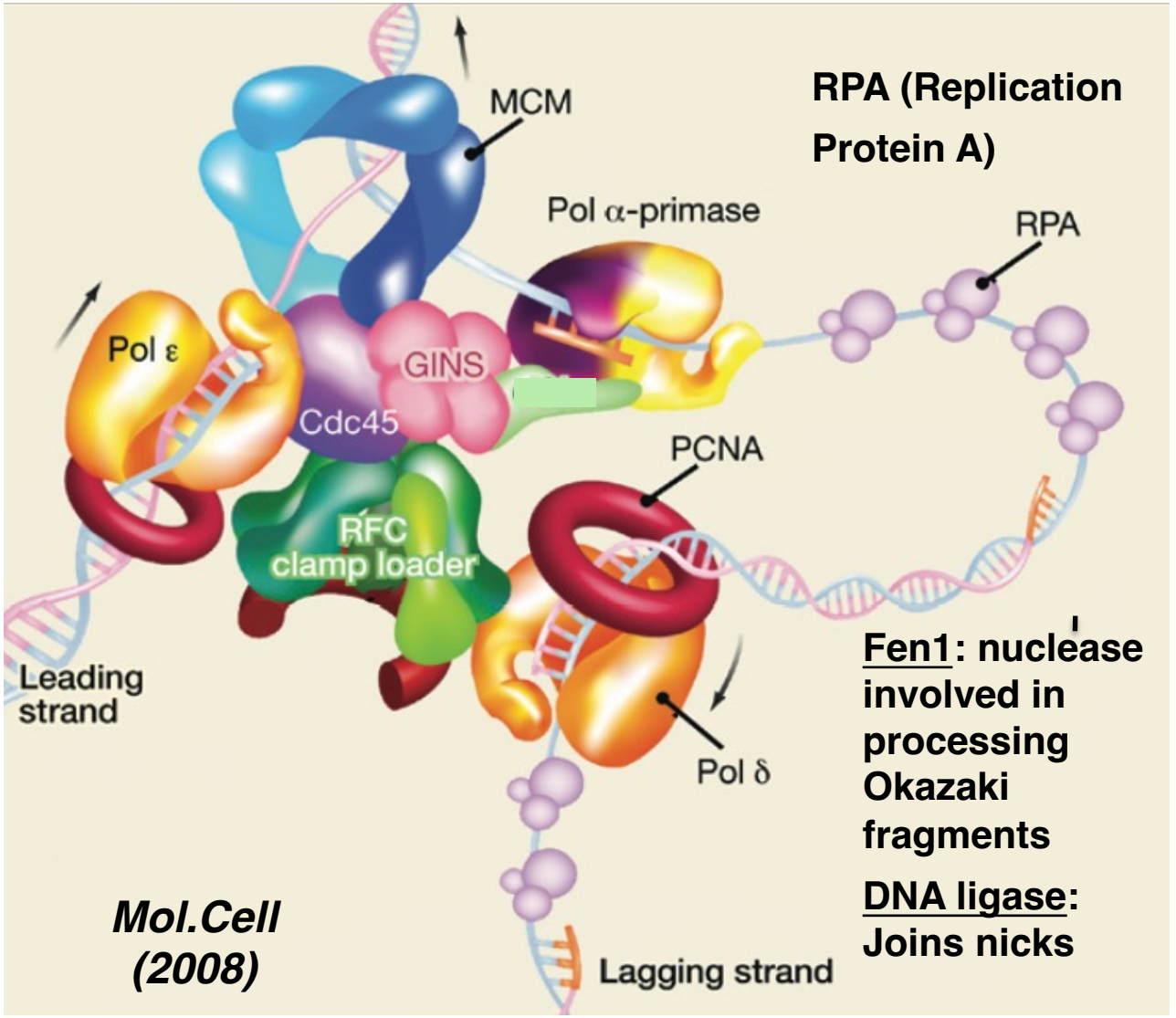
MCM2-7: hexameric ATP-dependent DNA helicase

PCNA: ***trimeric* sliding clamp**

Pol α + primase complex:

Complex contains both activities:

- Primase (RNA primer synthesis)
- DNA polymerase α (shortly extends RNA primers)



RPA (Replication Protein A)

Fen1: nuclease involved in processing Okazaki fragments

DNA ligase: Joins nicks

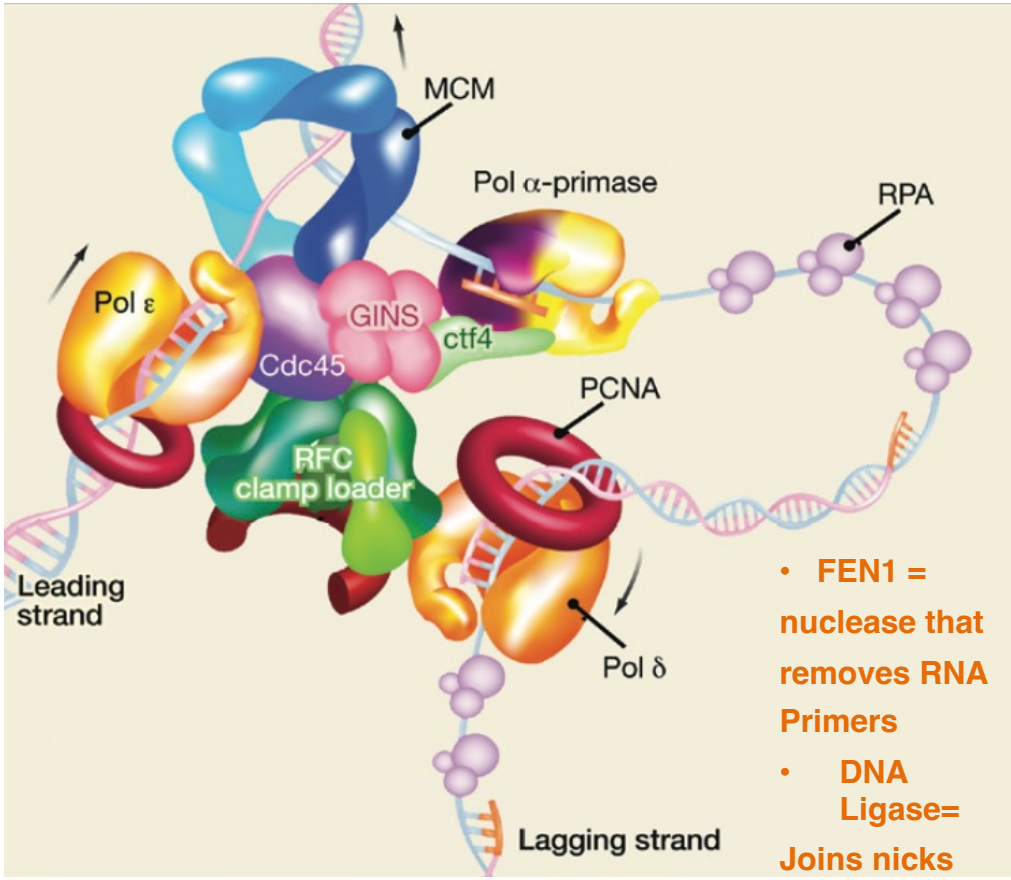
What is missing compared to a bacterial replisome?

A: A helicase that unwinds the DNA

B: A protein that holds all of the replisome components together

C: DNA ligase that ligates the Okazaki fragments after processing

D: A single-stranded DNA binding protein



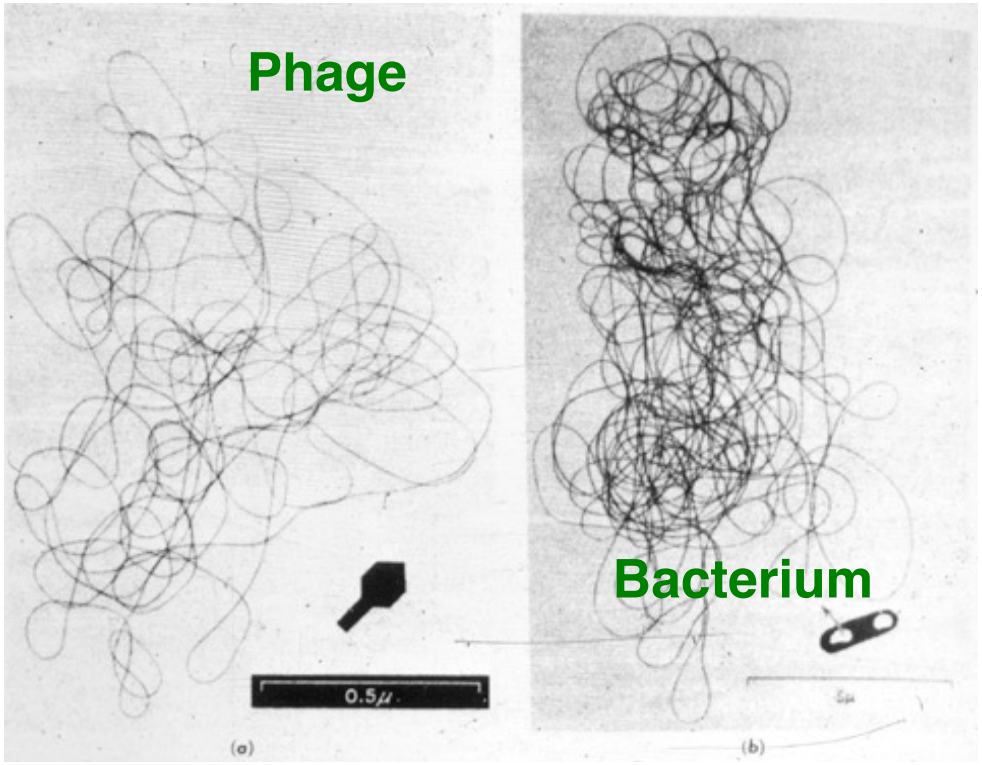
- **FEN1 = nuclease that removes RNA Primers**
- **DNA Ligase= Joins nicks**

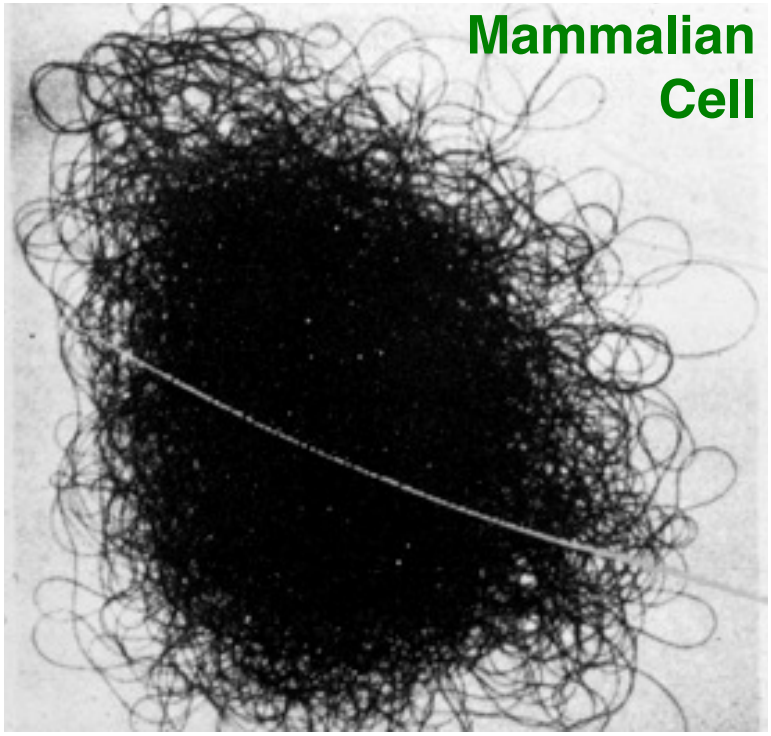
Rate of DNA synthesis is slower in eukaryotes than in bacteria

- **In eukaryotes: replication fork moves at ~50 nt/s**
- **In bacteria: replication fork moves at ~1000 nt/s**

Do eukaryotes have less DNA than bacteria?!

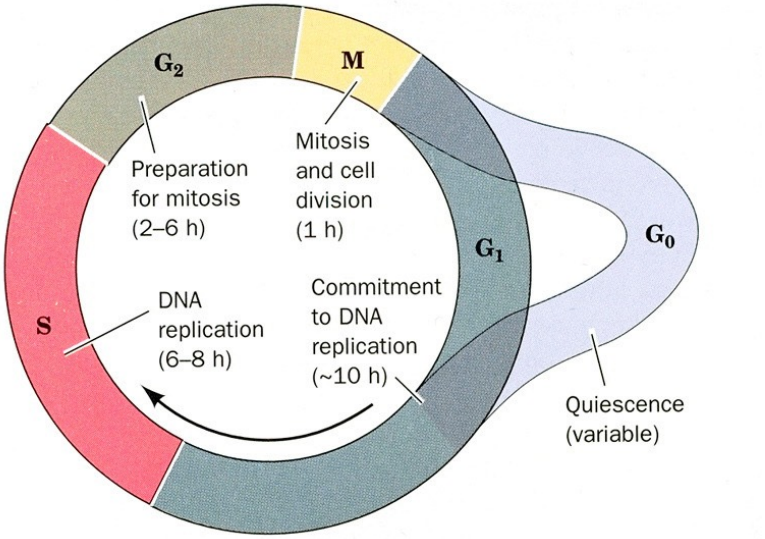
Amount of DNA in Mammalian Cells compared to Phage & Bacteria





	~10^3-10^4 base pairs	~10^6 base pairs	~10^9-10^{12} base pairs
Replication Time:	~5min.	~20-60min.	~6-8 hours

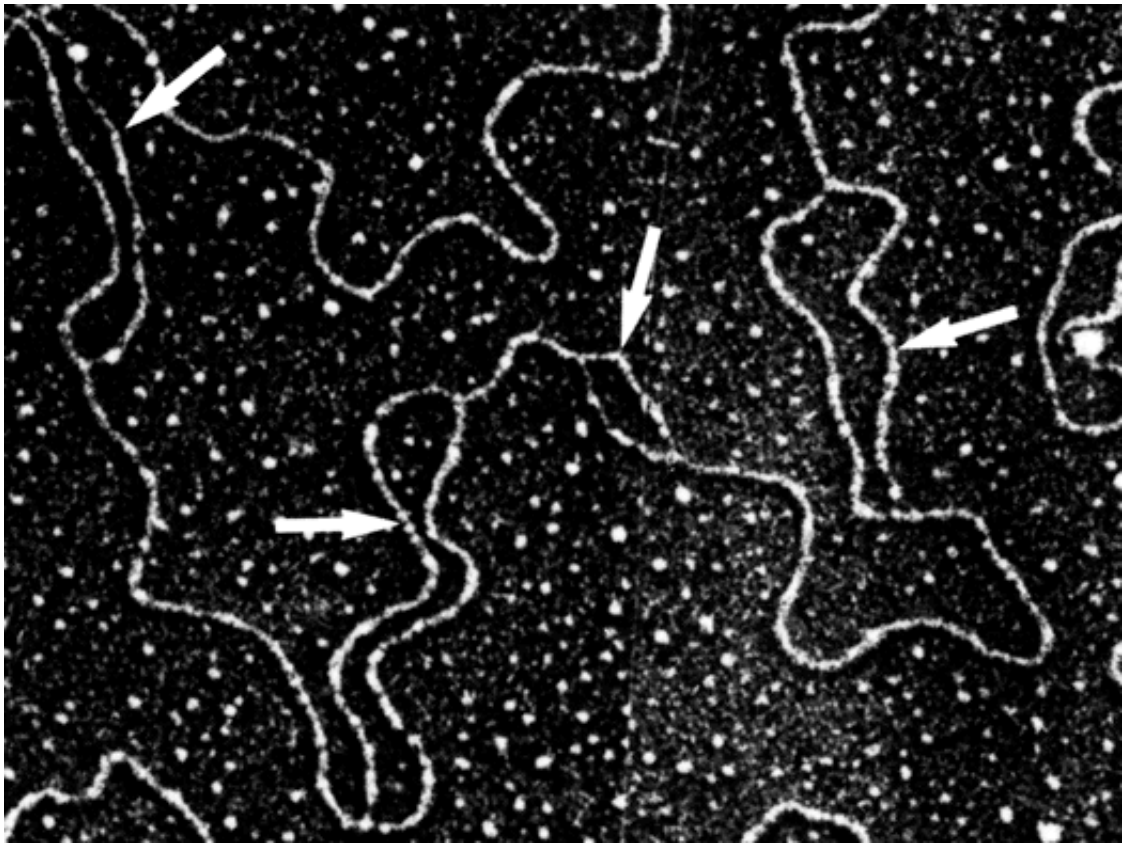
To make things even more challenging: DNA replication only occurs during part of the eukaryotic cell cycle



DNA replication → in S phase (6-8 hours!)

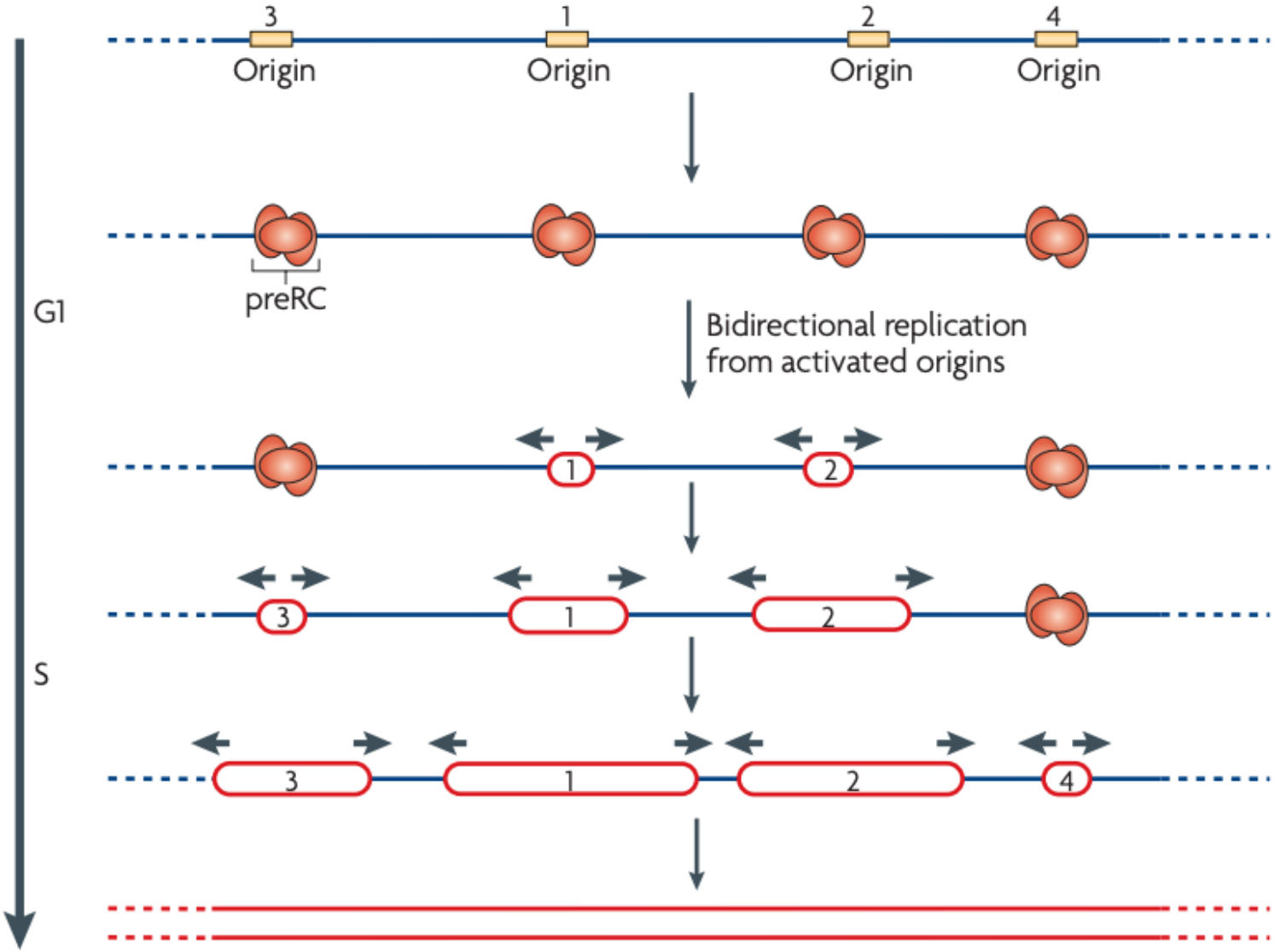
How do eukaryotes solve this?

Multiple origins of replication!



Multiple origins of replication in eukaryotes

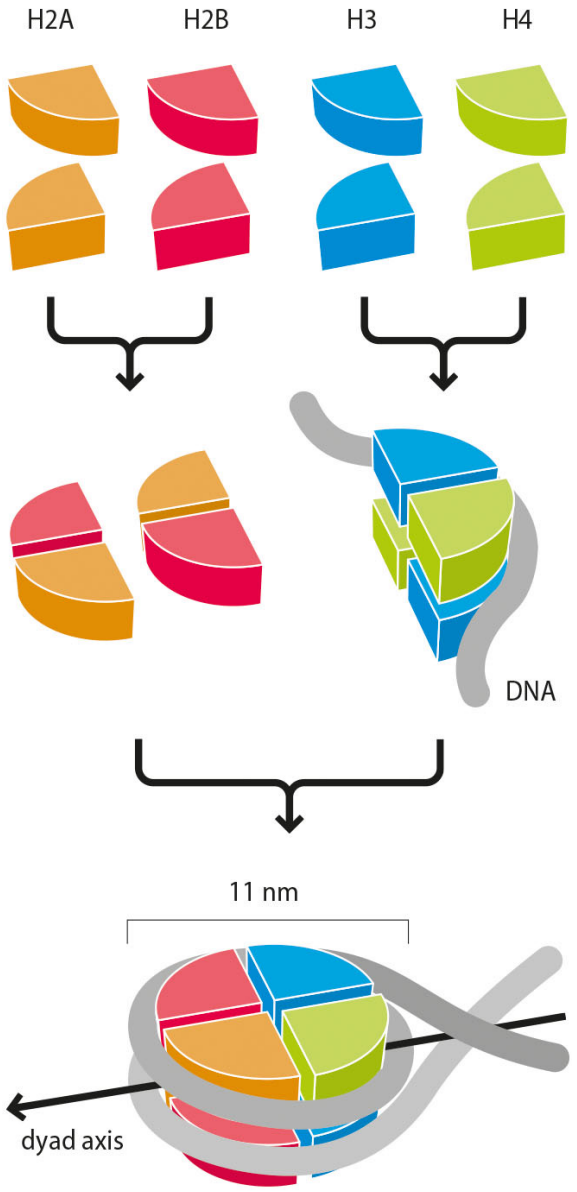
Replication is coordinated across all origins



Méchali, Nat Rev Mol Cell Bio 2010

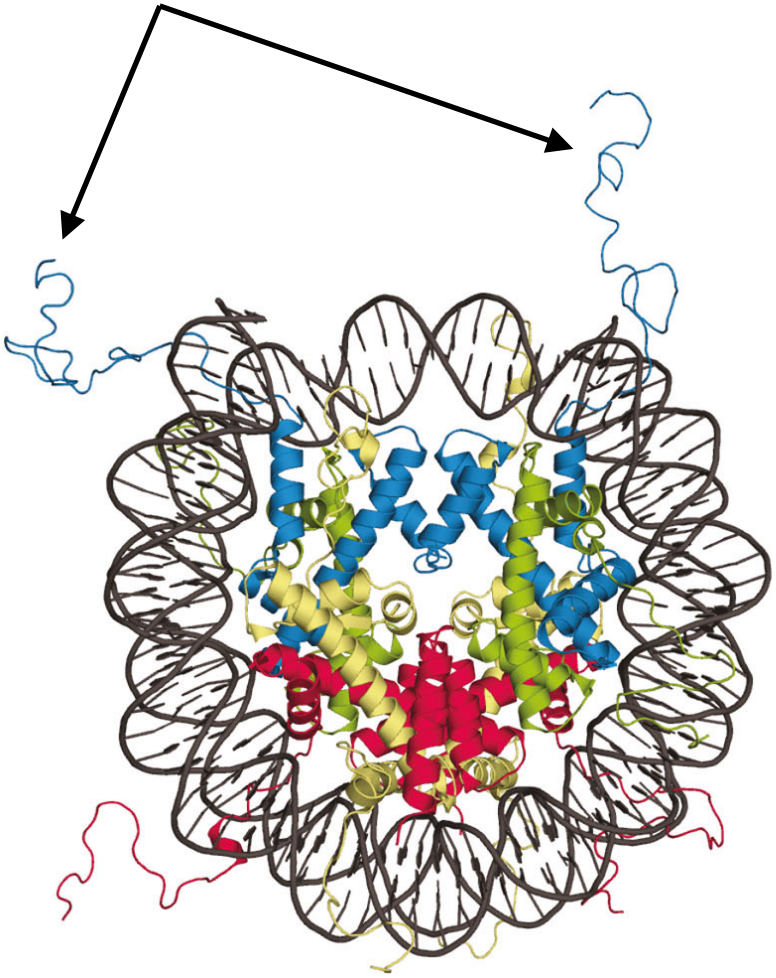
A second unique challenge for eukaryotes: Nucleosome packaging

- In a human cell: 2 meters of DNA needs to fit inside a nucleus that is 10 μm in diameter
- Histones: DNA-binding proteins that package DNA into chromatin
  - 4 core histone proteins: H2A, H2B, H3, H4
- DNA wraps around a complex of histone proteins to form a nucleosome
  - At the center of a nucleosome: a histone octamer containing 2 copies of each of the 4 core histone proteins
  - ~146 base pairs of DNA wrapped around the histone octamer: 1.76 turns around



Structure of the nucleosome core

N-terminal tails



PDB: 1KX5

Tails can be covalently modified (=epigenetic marks), which can alter chromatin structure and function

What challenges do nucleosomes cause for DNA replication?

A: Nucleosomes stop DNA ligase from sealing Okazaki fragments.

B: Newly synthesized DNA must be rapidly reassembled into nucleosomes behind the replication fork.

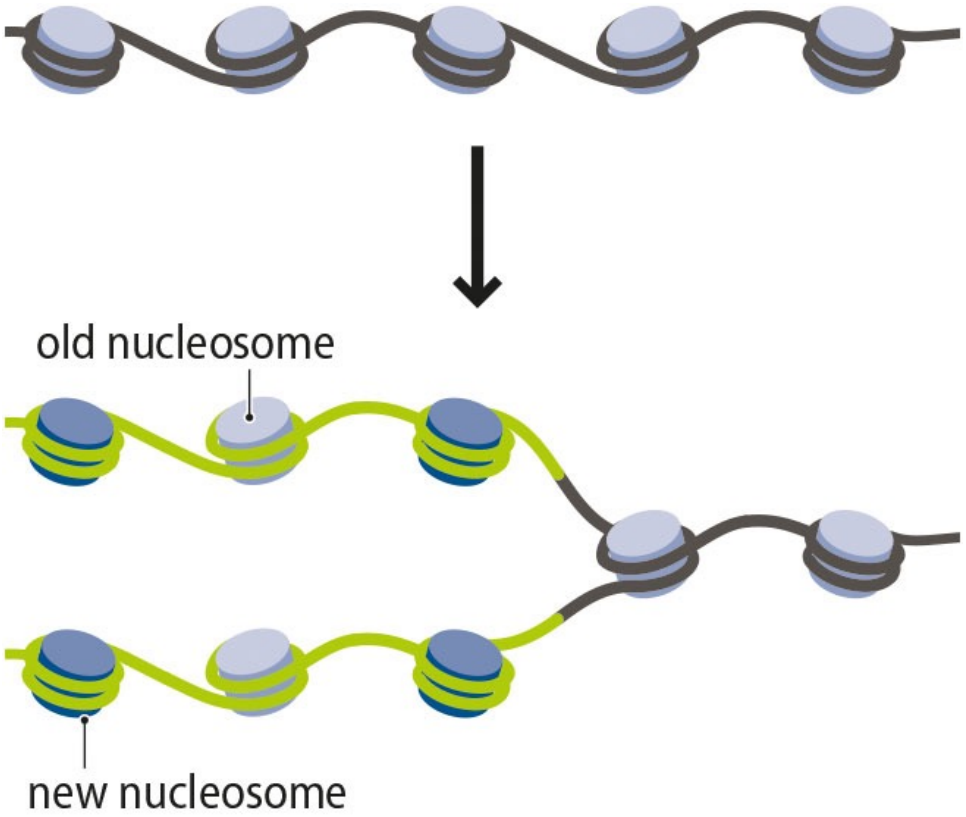
C: Nucleosomes stop DNA from being synthesized continuously on the leading strand.

D: Epigenetic information carried by histone modifications must be preserved during replication.

Eukaryotic replication machinery needs to deal with nucleosome packaging of DNA

A major challenge: The nucleosomes and associated epigenetic modifications need to be duplicated!

Parental histone segregation

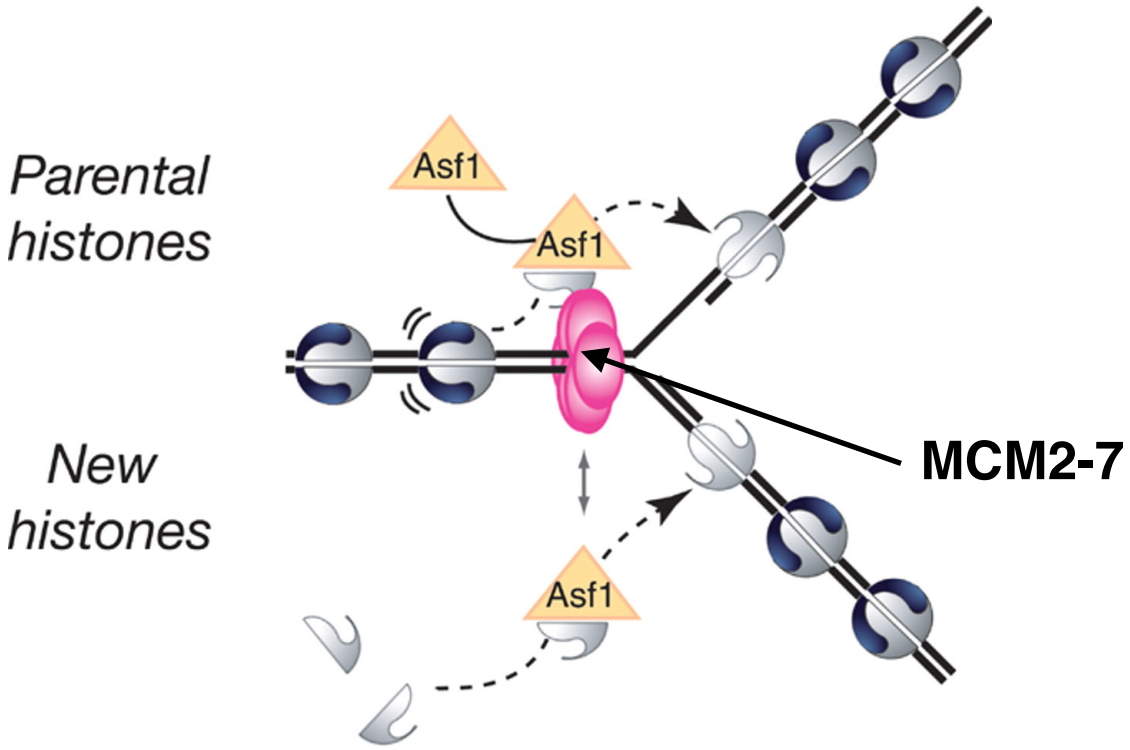


Nucleosomes reconstituted behind the replication fork as a mix of old and new

Which nucleosomes contain epigenetic modifications?

How is parental histone segregation accomplished?

- **Distribution of old histones to both new DNA molecules is carried out by active transfer of H3 and H4 by a complex that contains the histone chaperone Asf1**



New histones are added by Asf1 and Caf1

https://www.science.org/doi/10.1126/science.1148992

New histones are added by Caf1 together with Asf1

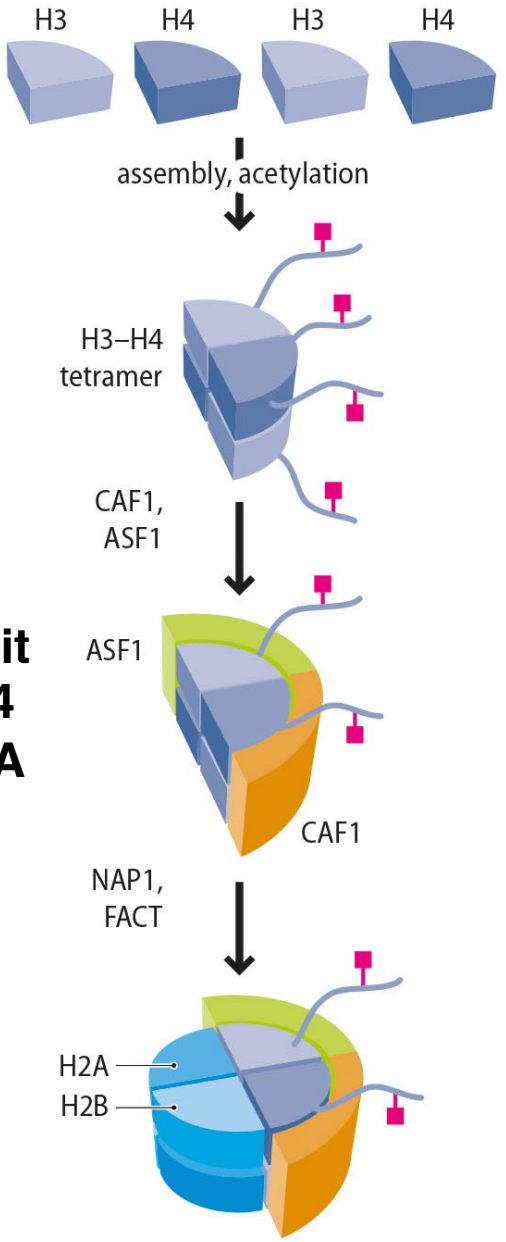
- After distributing the old histones, each daughter DNA strand has only half the number of nucleosomes required to package the DNA

New histones are synthesized during replication

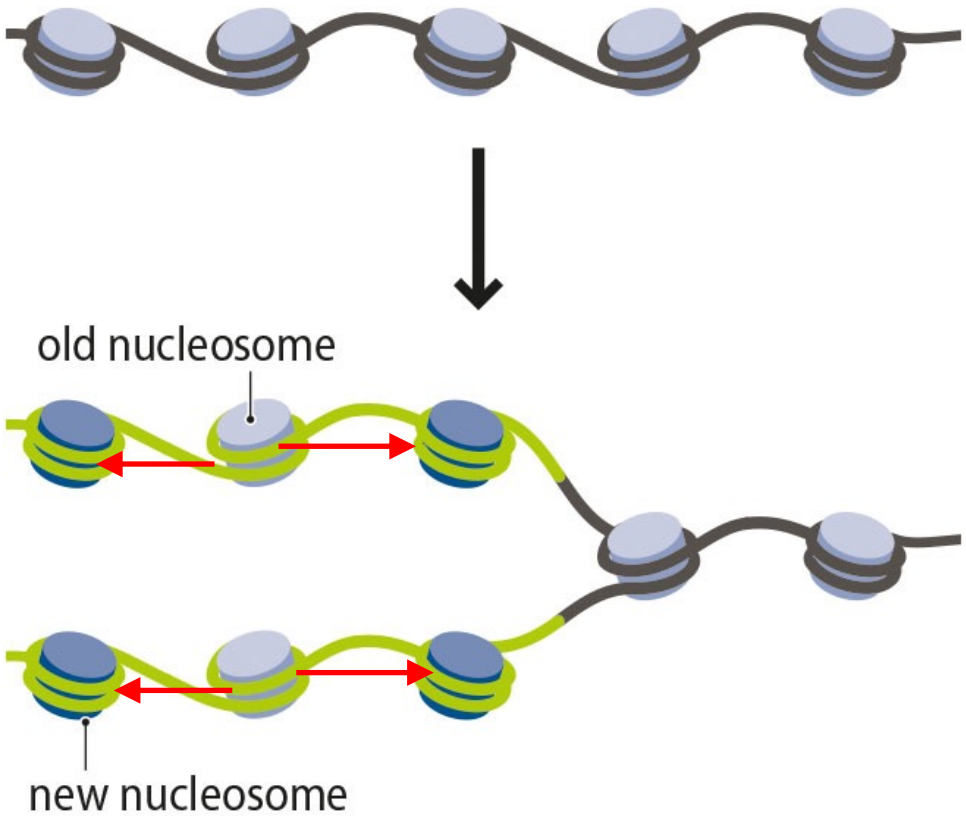
Newly synthesized histones are acetylated

Caf1 and Asf1 deposit the acetylated H3-H4 tetramers on the DNA

Other assembly factors add H2A and H2B, then the new histones are deacetylated



Modifications from old histones spread to neighbors



Modifying enzymes are recruited by old modified histones

What problem do linear chromosomes create for DNA replication?

A: Linear chromosomes can't form replication forks at their ends.

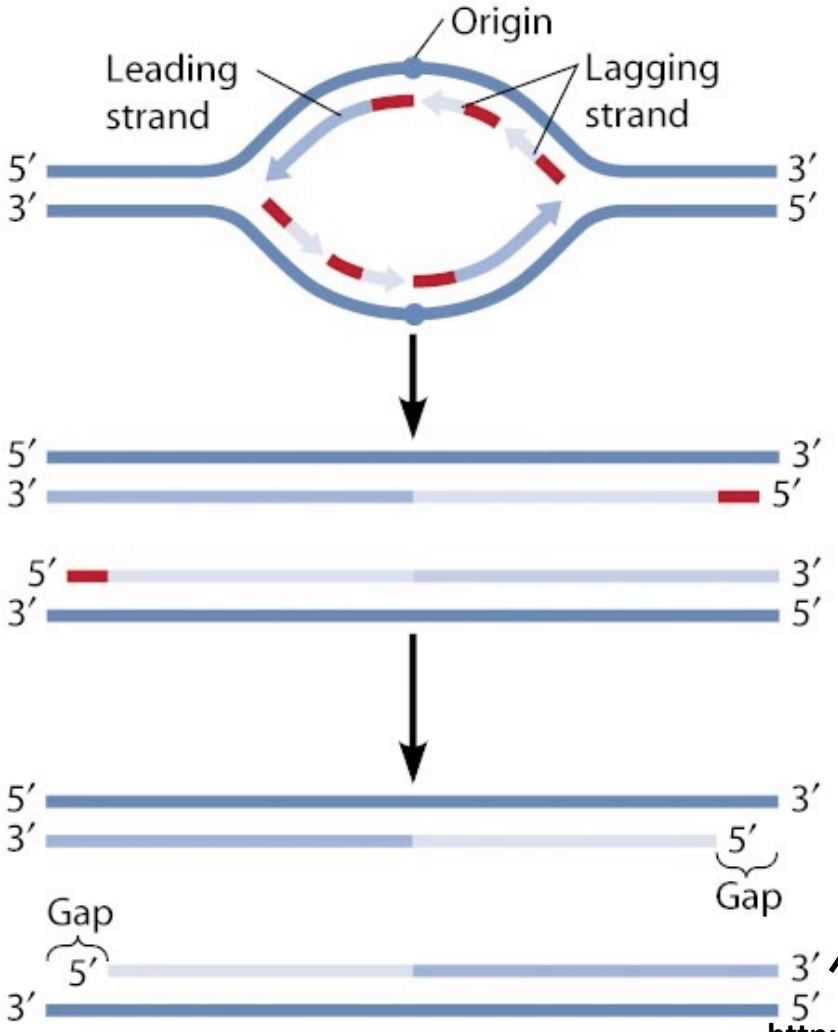
B: Linear chromosomes don't create any problems.

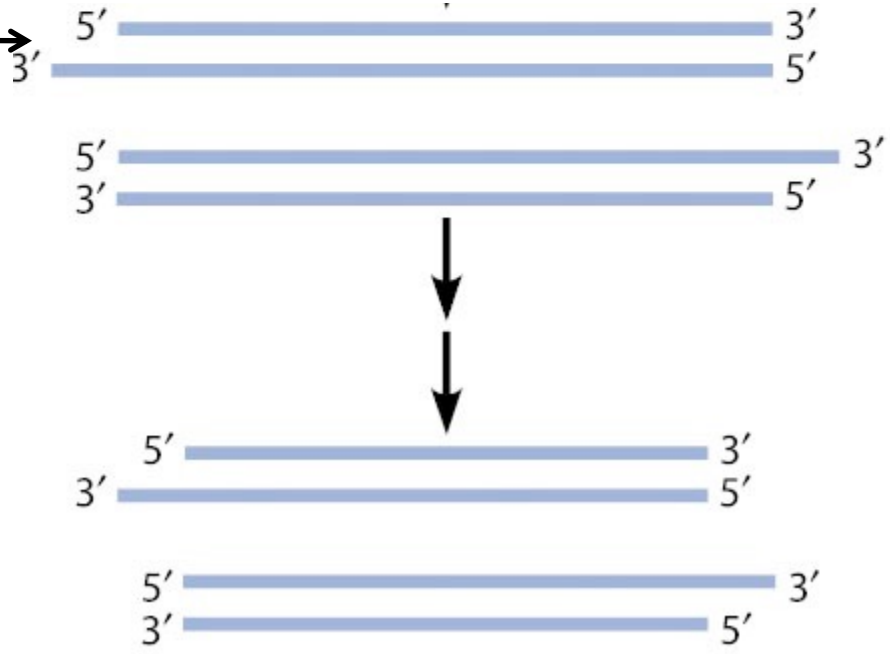
C: Removal of the final RNA primer on the lagging strand leaves a gap that can't be filled in at chromosome ends.

D: RNA primers cannot be synthesized at the ends of linear chromosomes.

Eukaryotic DNA Replication: Problem of maintaining the ends of linear chromosomes is linked to the degradation of RNA primers

Last primers on each 5´-end are removed but the gaps cannot be filled because of the lack of 3´-OH group





Each cycle of replication would result in progressive chromosome shortening

http://www.mun.ca/biology/desmid/brian/BIOL2060/BIOL2060-19/1915.jpg

Telomeres

- **The physical ends of linear chromosomes**
- **Composed of double stranded DNA with 10-1000s of G-rich repeats (on the lagging strand) ending in a single stranded 3' overhand, and associated telomere binding proteins that protect the ends (*can form G-quadruplexes*)**
- **The single strand overhang is at least 2 repeats in Oxytricha and Tetrahymena. In humans the length varies from 50-300 repeats**

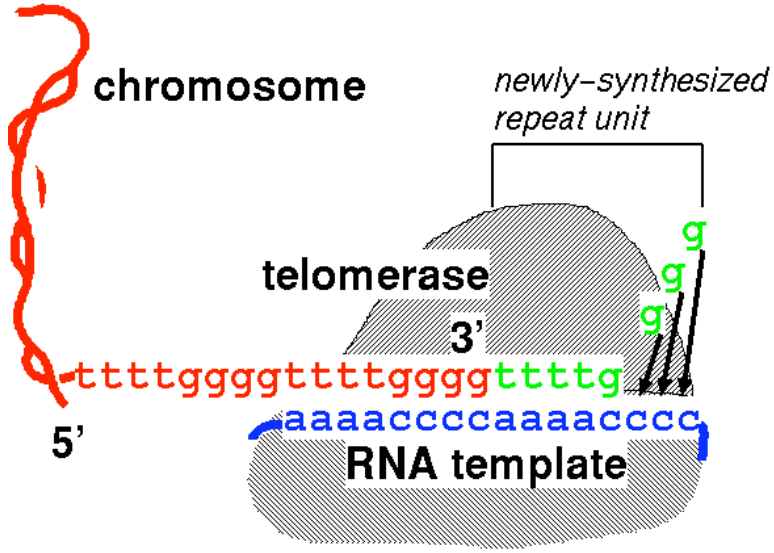
Telomere repeat sequences in some organisms

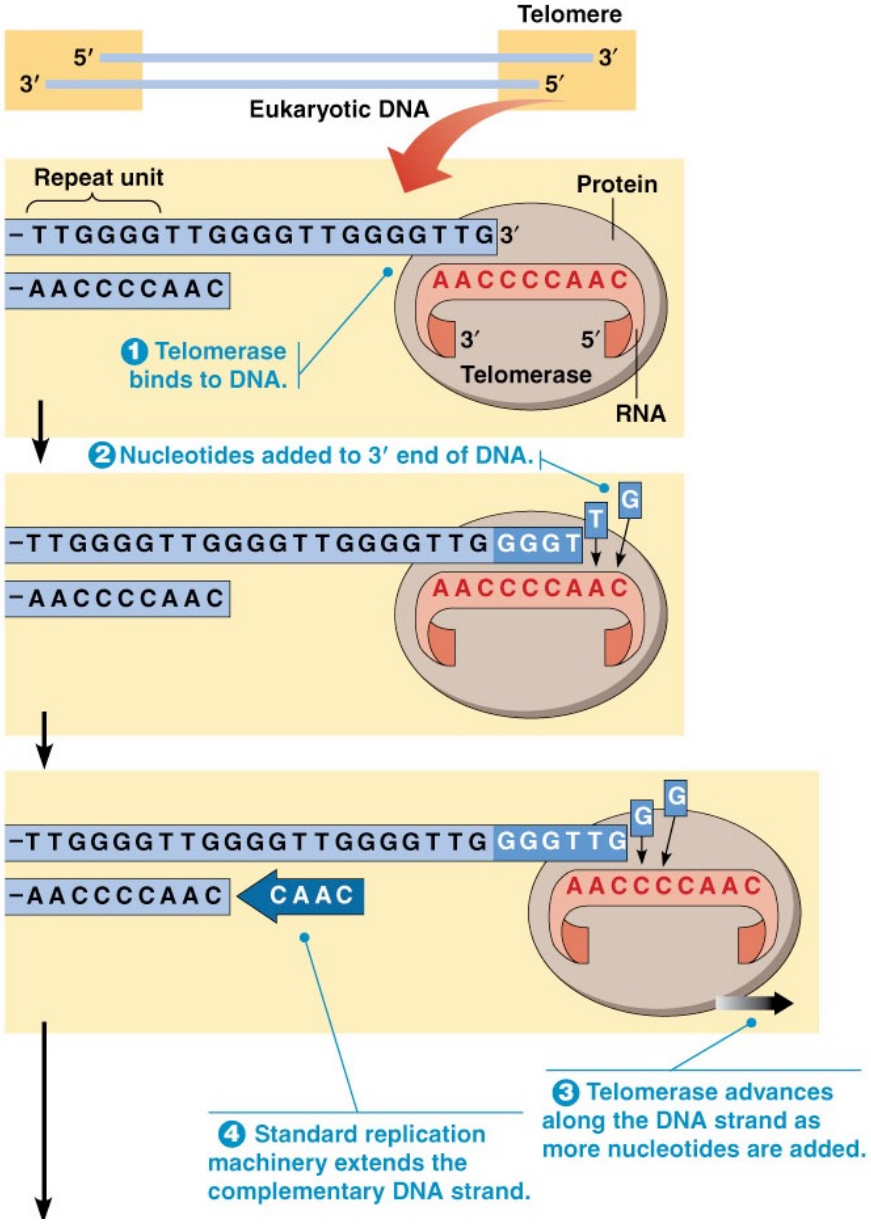
Oxytricha nova	TTTTGGGG
Tetrahymena	TTGGGG
human	TTAGGG

5′_____ $(TTTTGGGG)_n$TTTTGGGGTTTTGGGG

3′_____ $(AAAACCCC)_n$

Preservation of telomeres by telomerase





Telomerase:

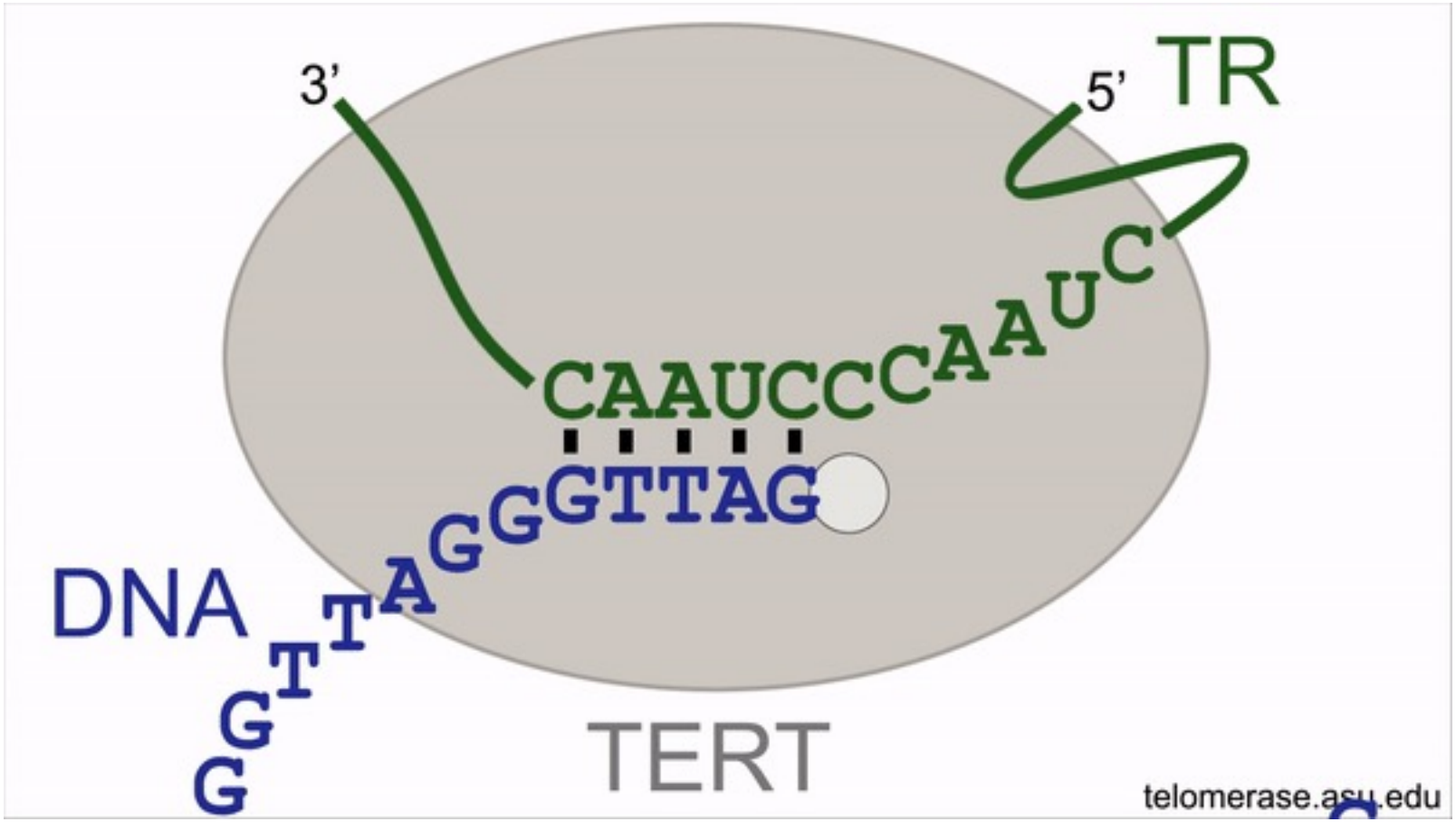
Telomerase *reverse transcriptase* (TERT):
1 RNA (TR) +
+ several proteins

Remember: Reverse transcriptases make DNA from an RNA template

Nobel Prize 2009: Greider, Blackburn, Szostak

After extension of the upper strand by telomerase, the replication machinery can now use this strand to make a new RNA primer using primase, then "elongate" this strand.

Model for telomere repeat synthesis illustrating template repositioning (translocation) for telomere repeat synthesis



From Julian Chen web site
This is human telomerase template

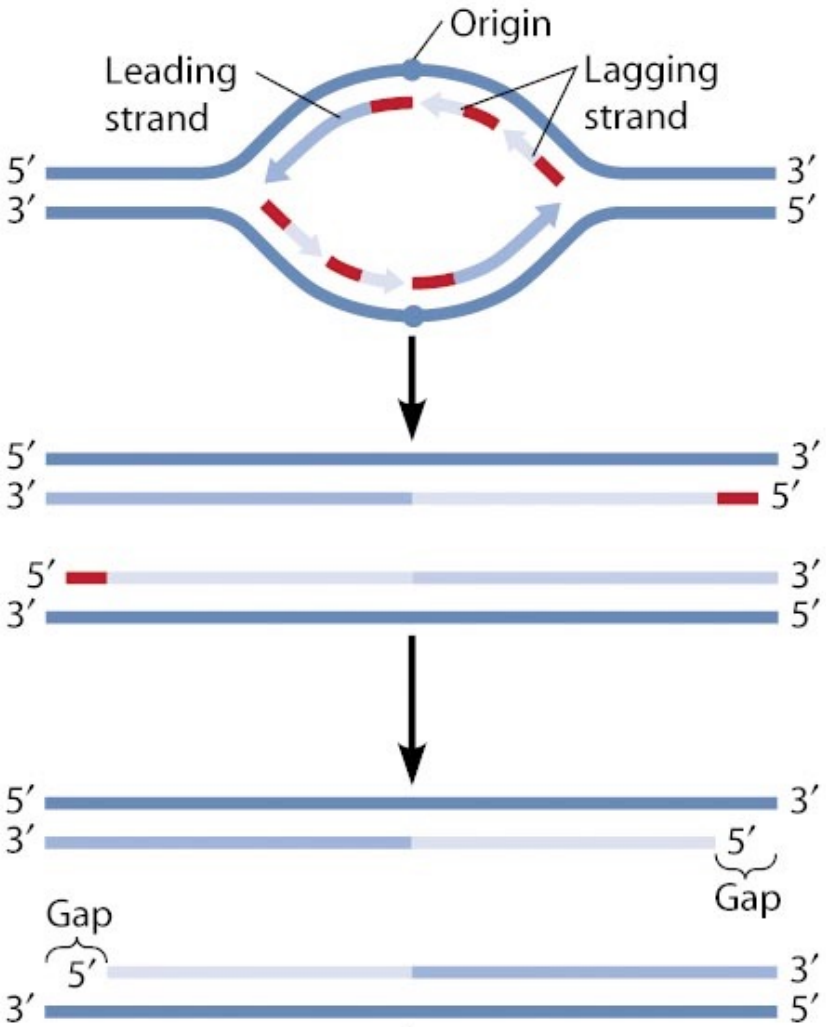
How does the 5’ to 3’ polymerase activity of telomerase solve the 5’ gap problem?

A: It directly fills in the 5’ gaps

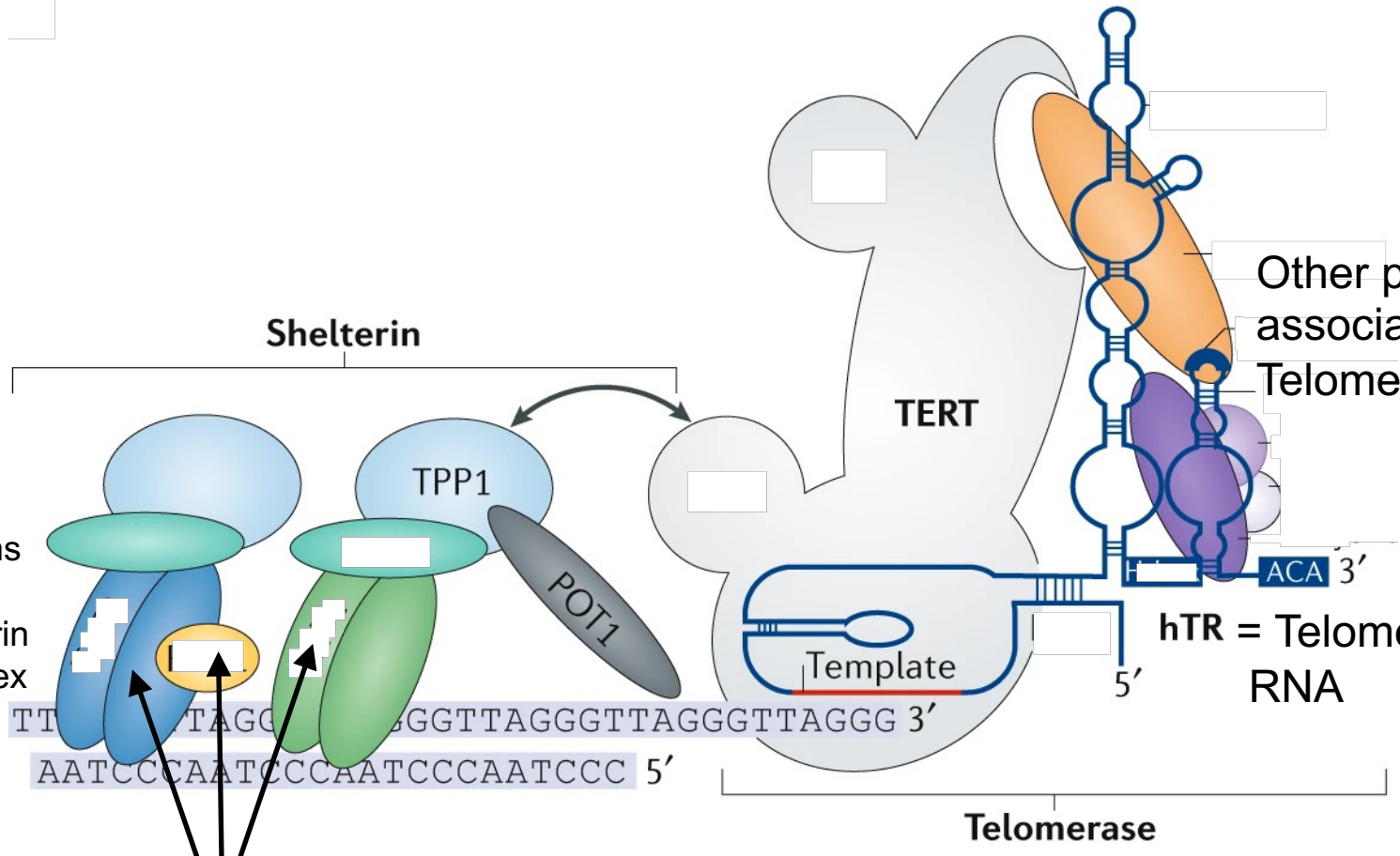
B: It doesn’t, because it is working on the wrong strand

C: It lengthens both parental strands so the overall length is maintained

D: It doesn’t matter since the ends are repetitive sequences



Telomerase is recruited to the 3´-end of chromosomes by the Shelterin protein complex: POT1, TPP1 and others



Adapted from Roake & Artandi
Nat. Rev. Mol. Cell Biol. (2020)

The problem of processivity in Telomerase

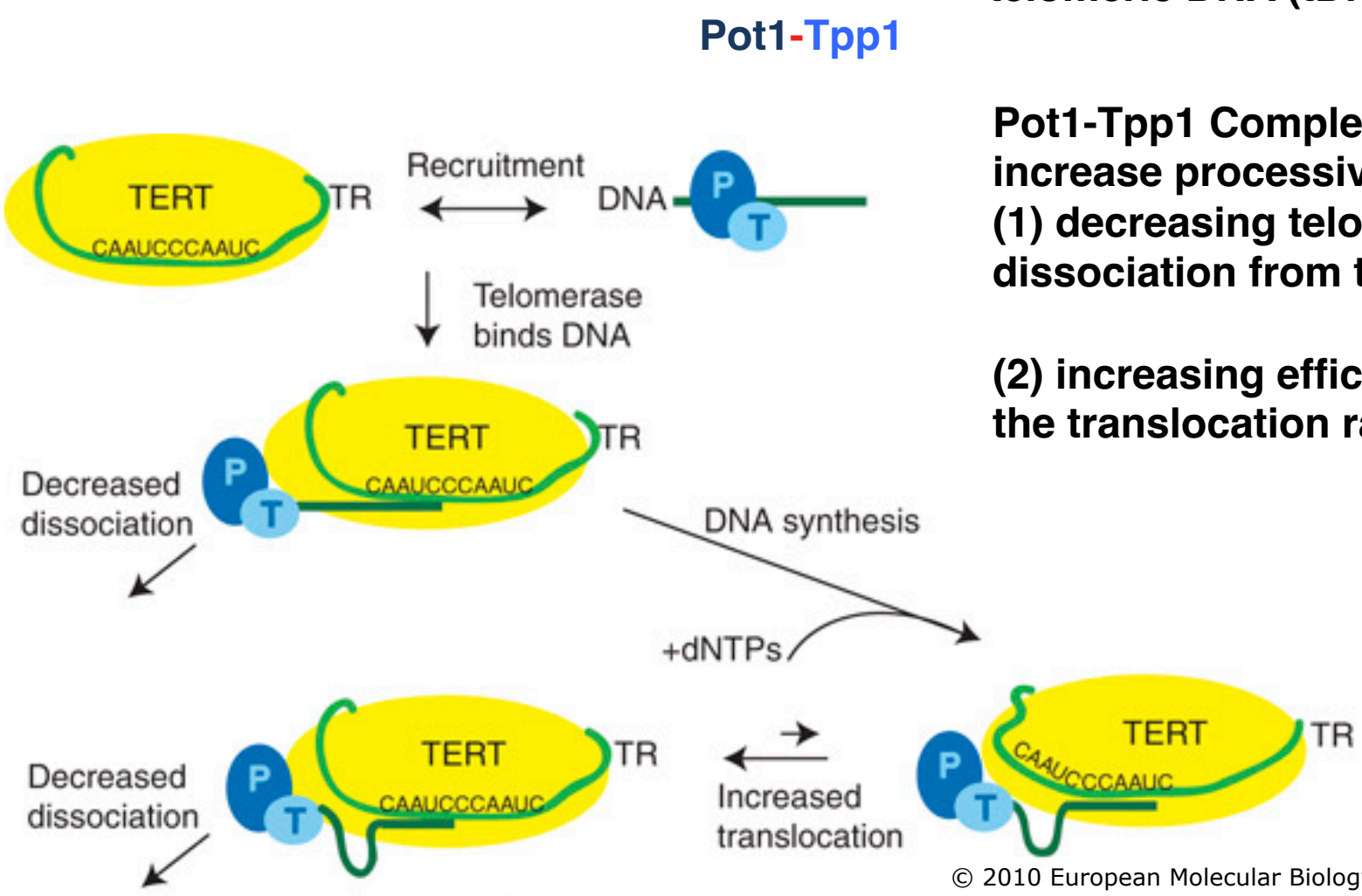
Telomerase needs to translocate after each extension cycle, it could fall off the DNA!

The Pot1-Tpp1 complex improve telomerase processivity to telomerase in vivo

Pot1-Tpp1 form a protein complex that binds to ss telomeric DNA (tDNA)

Pot1-Tpp1 Complex increase processivity by:
(1) decreasing telomerase dissociation from tDNA

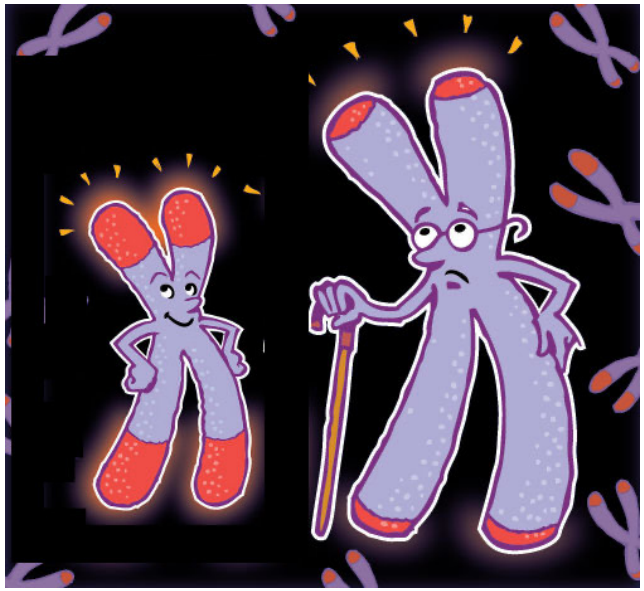
(2) increasing efficiency of the translocation rate



Telomerase: a highly regulated determinant of cellular aging, stem cell renewal, and tumorigenesis

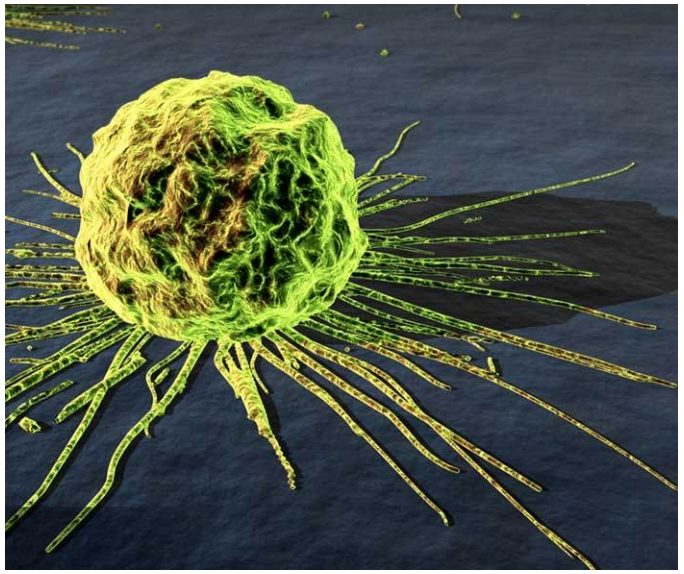
Aging:

Most somatic cells have **low to no detectable** telomerase activity.



Cancer:

~90% have **increased** telomerase activity



Stem cells and germ line cells have **active telomerase** (but less than cancer cells).

Mutations of telomerase components including telomerase RNA are linked to dyskeratosis congenita, aplastic anemia, pulmonary fibrosis, and other diseases.

Telomerase is required for telomere maintenance.

Telomerase activity is required for the immortal phenotype of cancer cells.

Correlations between telomere length, life span, aging and life style:

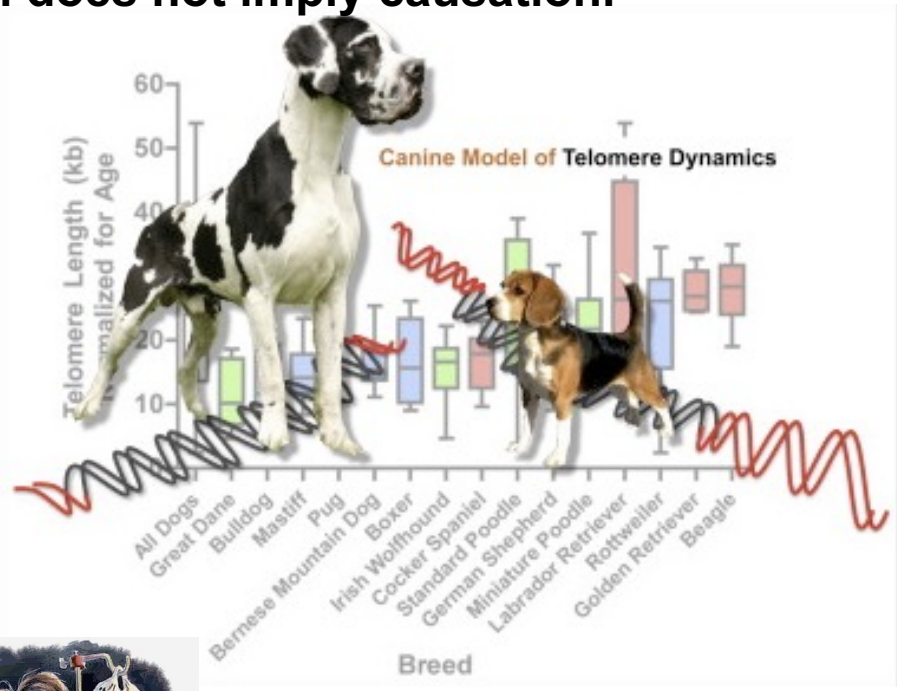
But please remember: correlation does not imply causation.

Telomere lengths and life span in dogs:

Cell Reports Dec. 2012

"telomere length is a strong predictor of average life span among 15 different breeds ($p < 0.0001$)"

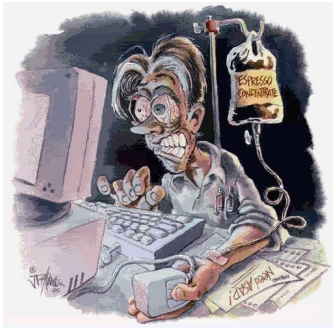
"Breeds with shorter mean telomere lengths show an increased probability of death from cardiovascular disease"



Telomere lengths, stress and aging:

Proc Natl Acad Sci U S A. 2004

"Women with the highest levels of perceived stress have telomeres shorter on average by the equivalent of at least one decade of additional aging compared to low stress women".



Longer Telomeres associated with ultra-endurance events: ultra marathon runners have longer telomeres than control subjects:

PLoS One July 2013 | Volume 8 | Issue 7 | e69377