DNA polymerase I

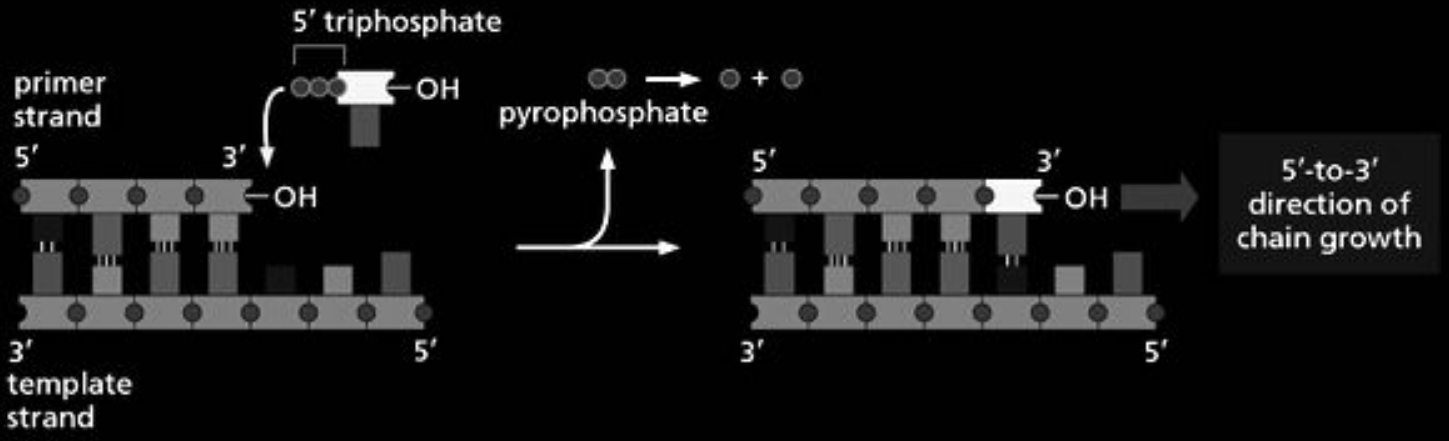
Pedro Fortes, Sergi Garcia, Nuria Güil, Alba Ponce

Structural Biology - Course 2016-2017

Index

Introduction
Polymerases' generalities

$$(DNA)_{n\ residues} + dNTP \rightarrow (DNA)_{n+1\ residues} + PP_i$$



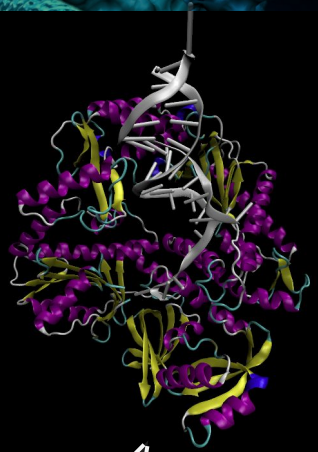
Addition of a deoxyribonucleotide to the 3'-OH end of a polynucleotide chain, catalyzed by a DNA polymerase.
Modified from Molecular Biology of the Cell, 6th edition by B. Alberts et al.

Introduction
Prokaryotic polymerases

E. Coli

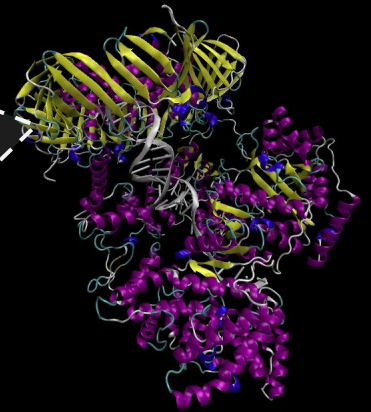
DNA Pol I

Delete RNA primers and replace the strand with DNA



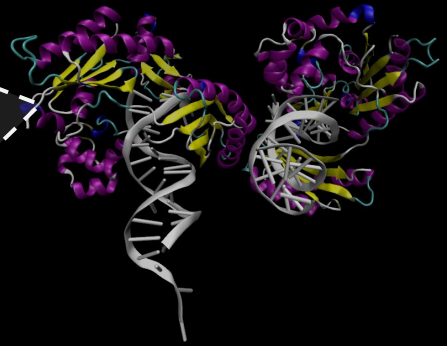
DNA Pol III

DNA synthesis



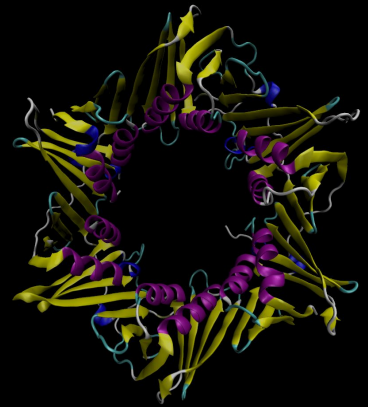
DNA Pol IV

Untargeted mutagenesis



DNA Pol II

Reparation



DNA Pol V

Reparation

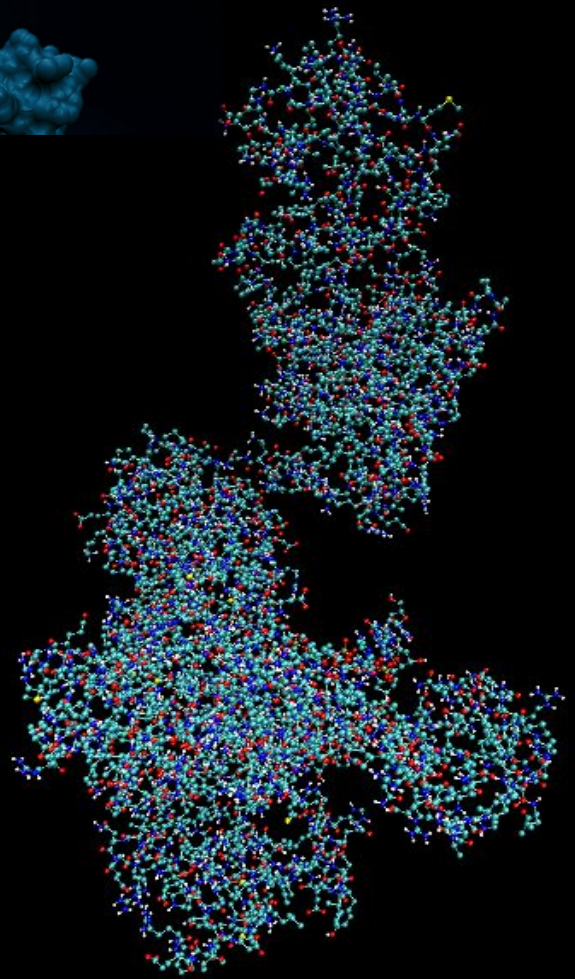
DNA polymerase I
SCOP Classification

Class
Multi-domain proteins (alpha and beta)

Fold
DNA/RNA polymerases

Superfamily
DNA/RNA polymerases

Family
PolymeraseA

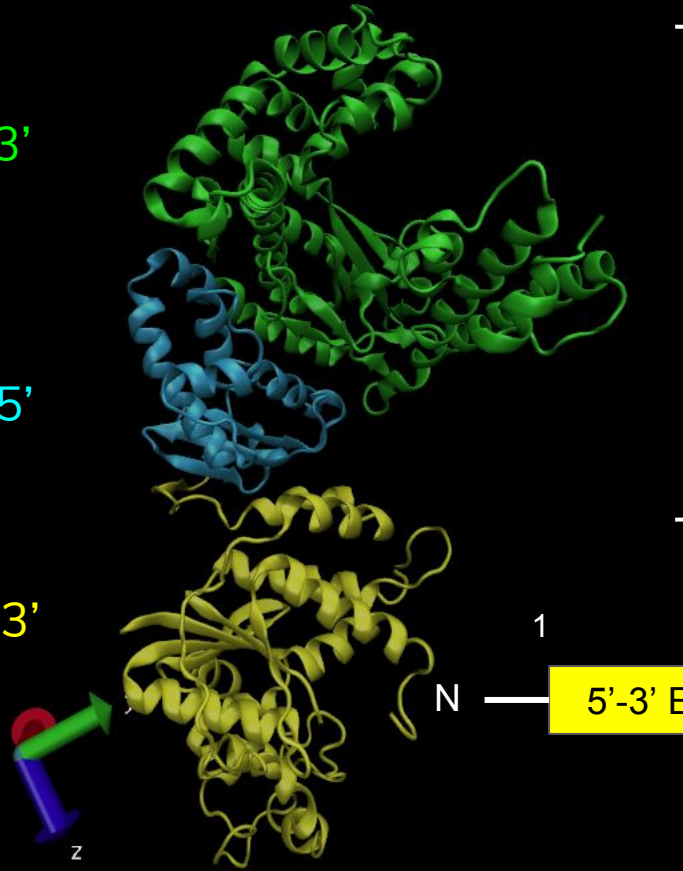


DNA polymerase I
Structure

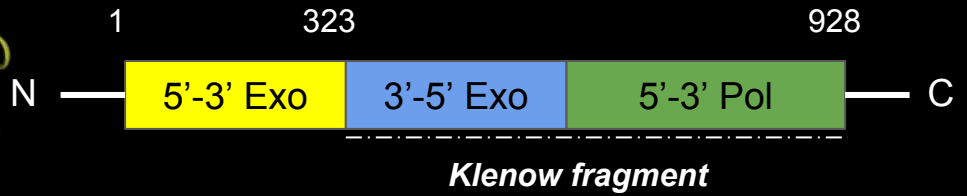
Polymerase 5’ - 3’

Exonuclease 3’ -5’

Exonuclease 5’ - 3’



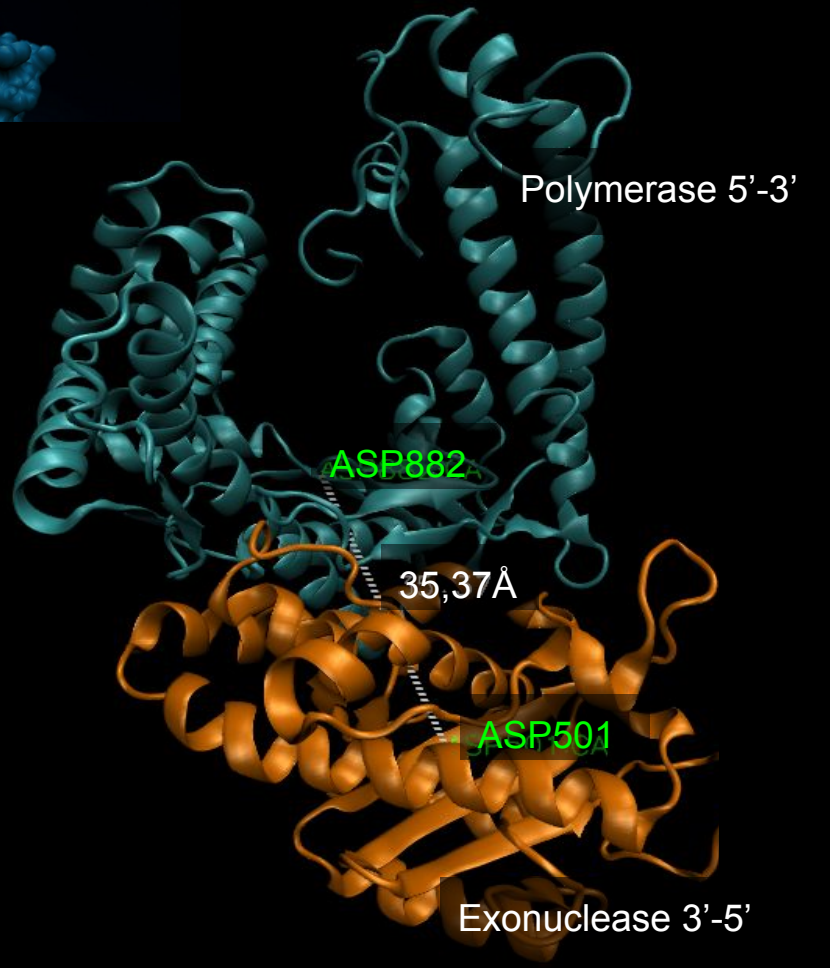
Klenow fragment



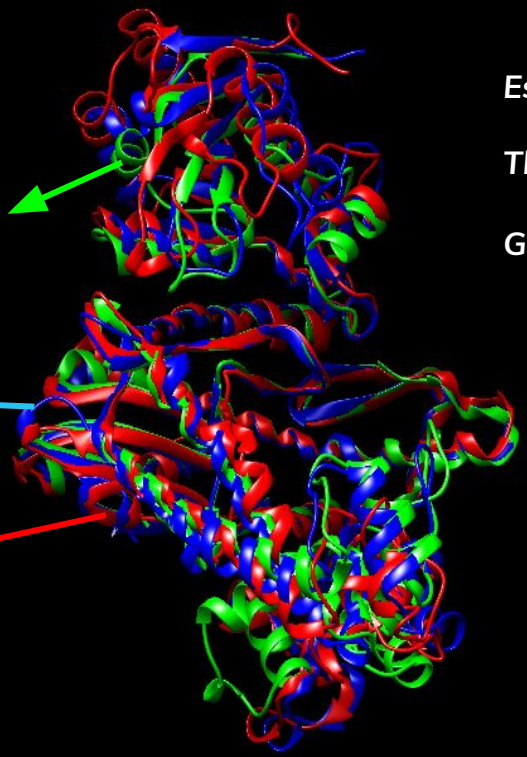
Klenow fragment

Distance between the two catalytic sites

The two catalytic sites of the enzyme are separated by approximately 35 Å



Klenow fragment
Conservation



1L3S *Geobacillus stearothermophilus*

1QTM *Thermus aquaticus*

2KZM *Escherichia coli*

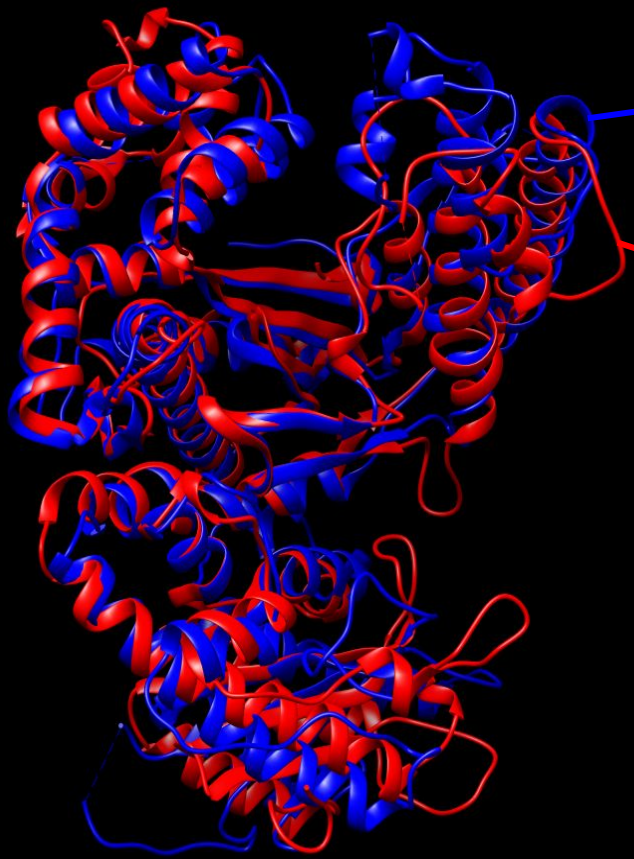
Escherichia coli: *Protobacterium*

Thermus aquaticus: *Dienococcus-thermus*

Geobacillus stearothermophilus: *Firmicutes*

SCORE = 8,28
RMS = 1,89

Klenow fragment
Conservation



DNA polymerase nu *Homo Sapiens*

DNA polymerase I *Escherichia coli*

SCORE = 4,94
RMS = 1,80

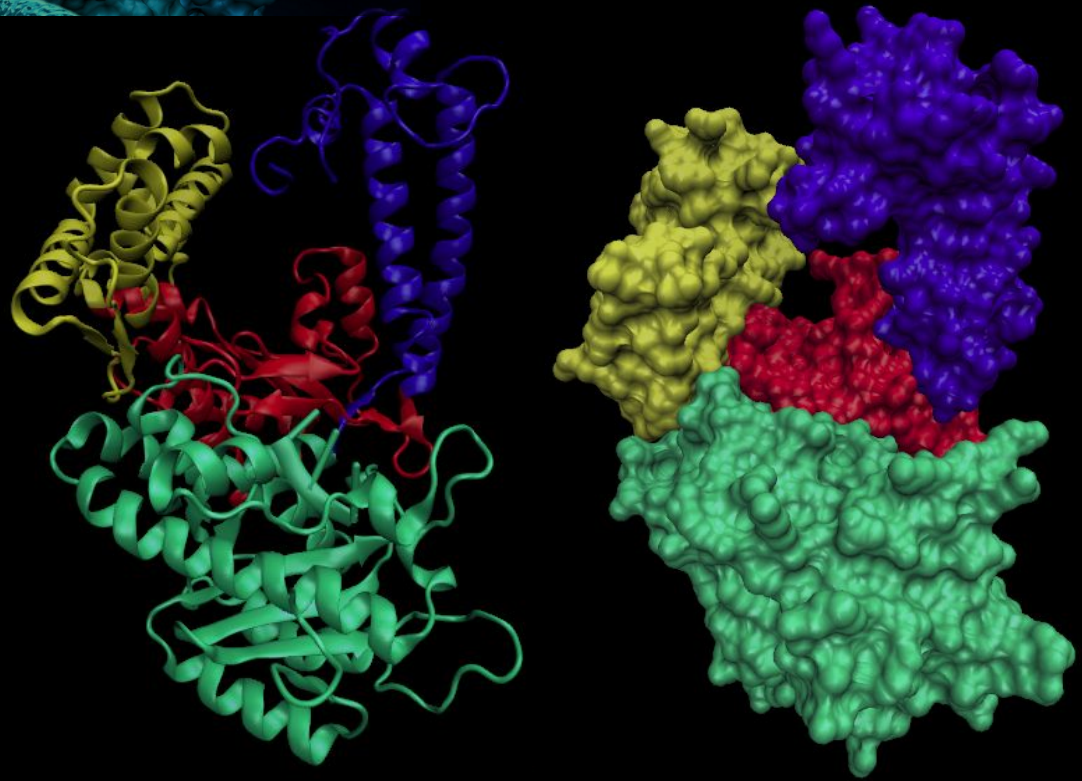
Klenow fragment
Domains

2 main regions in Klenow fragment:

Exonuclease 3’-5’ (324 - 547)

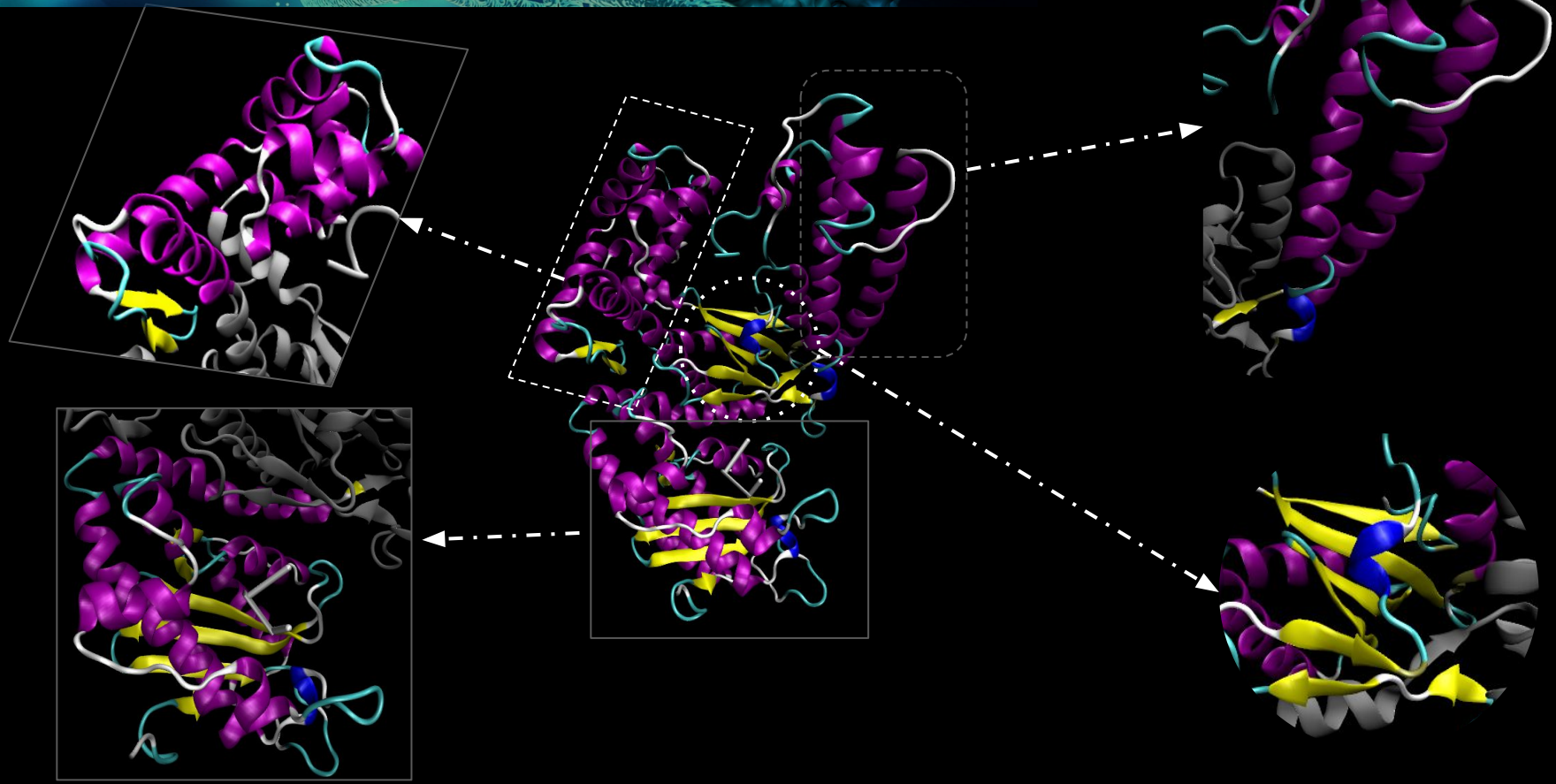
Polymerase 5’-3’ domain (548 - 928)

- **Palm** (648 - 717, 848 - 928)
- **Fingers** (718-847)
- **Thumb** (548-647)

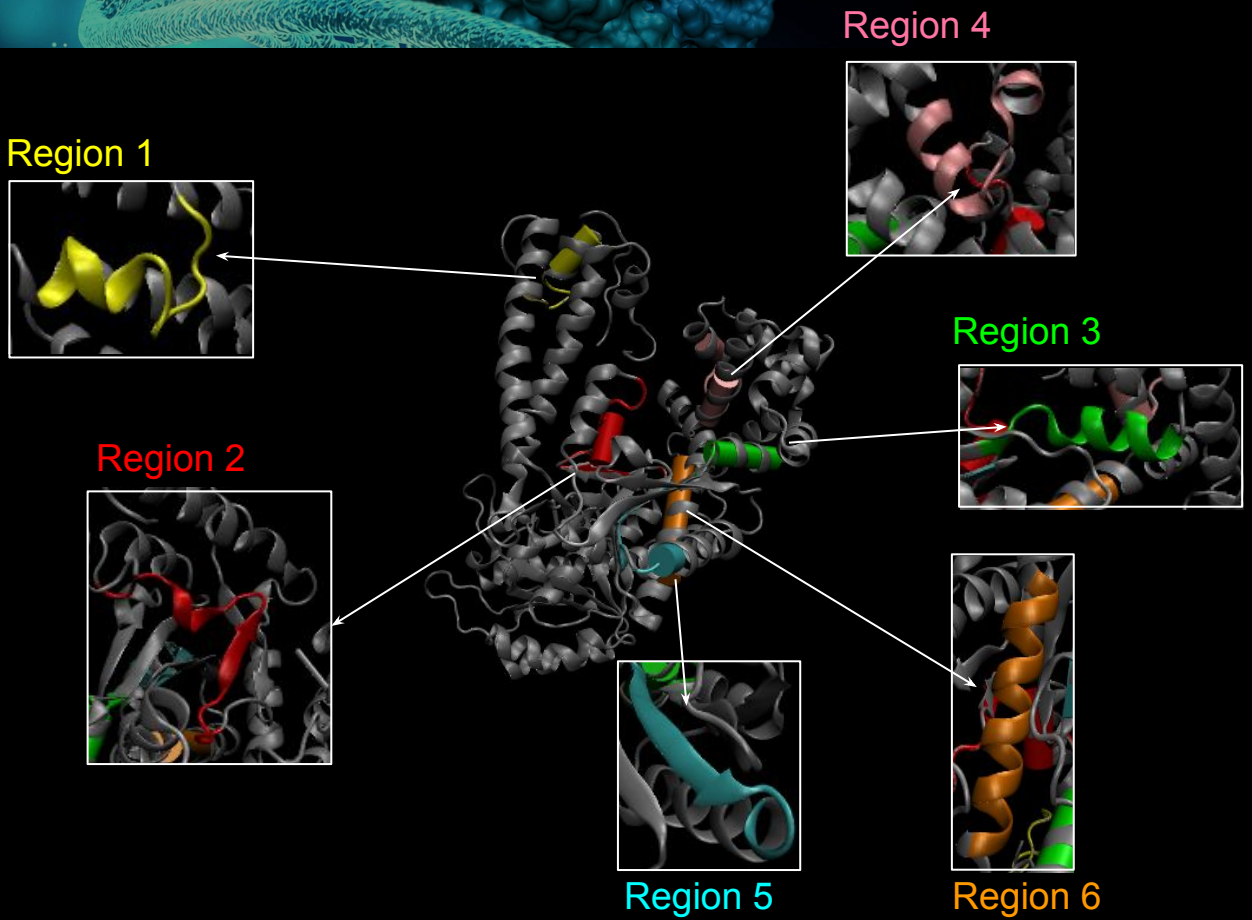


Klenow Fragment

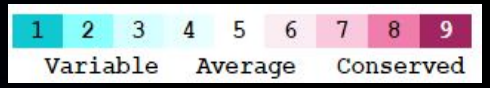
Secondary structure



Polymerase domain
Regions



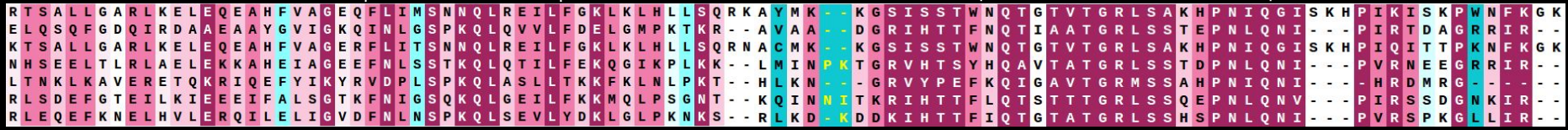
Polymerase domain
Regions



1. *M. musculus*
2. *M. tuberculosis*
3. *H. sapiens*
4. *E. coli*
5. *A. aeolicus*
6. *R. prowazekii*
7. *H. pylori*

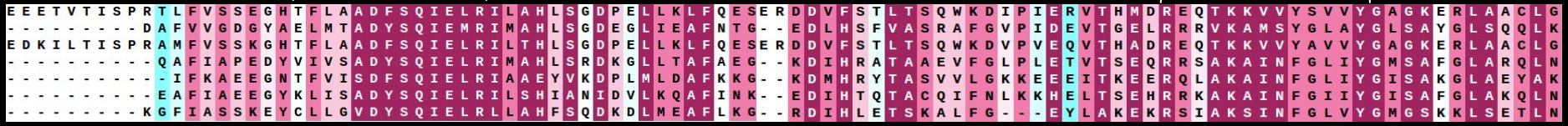
Region 1: hxhxSxpQLxxhL

Region 2: xQxxxhTGRhSxxpPNhQ



Region 3: xhhxxDxSQIEhR

Region 4: RpxxKxhxxxhhY

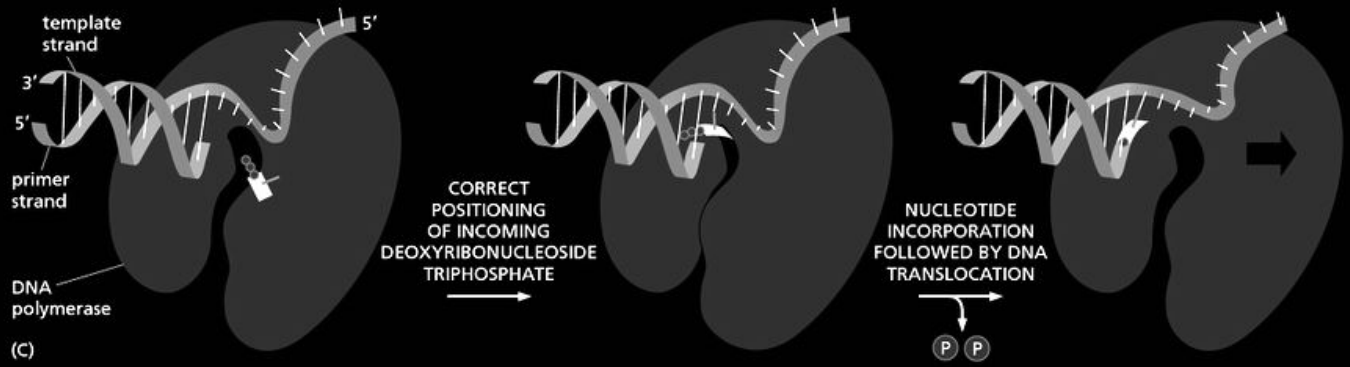


Region 6: xxhNxhhQGpxxDhxKx

Region 5: hxxhHDEhhhE

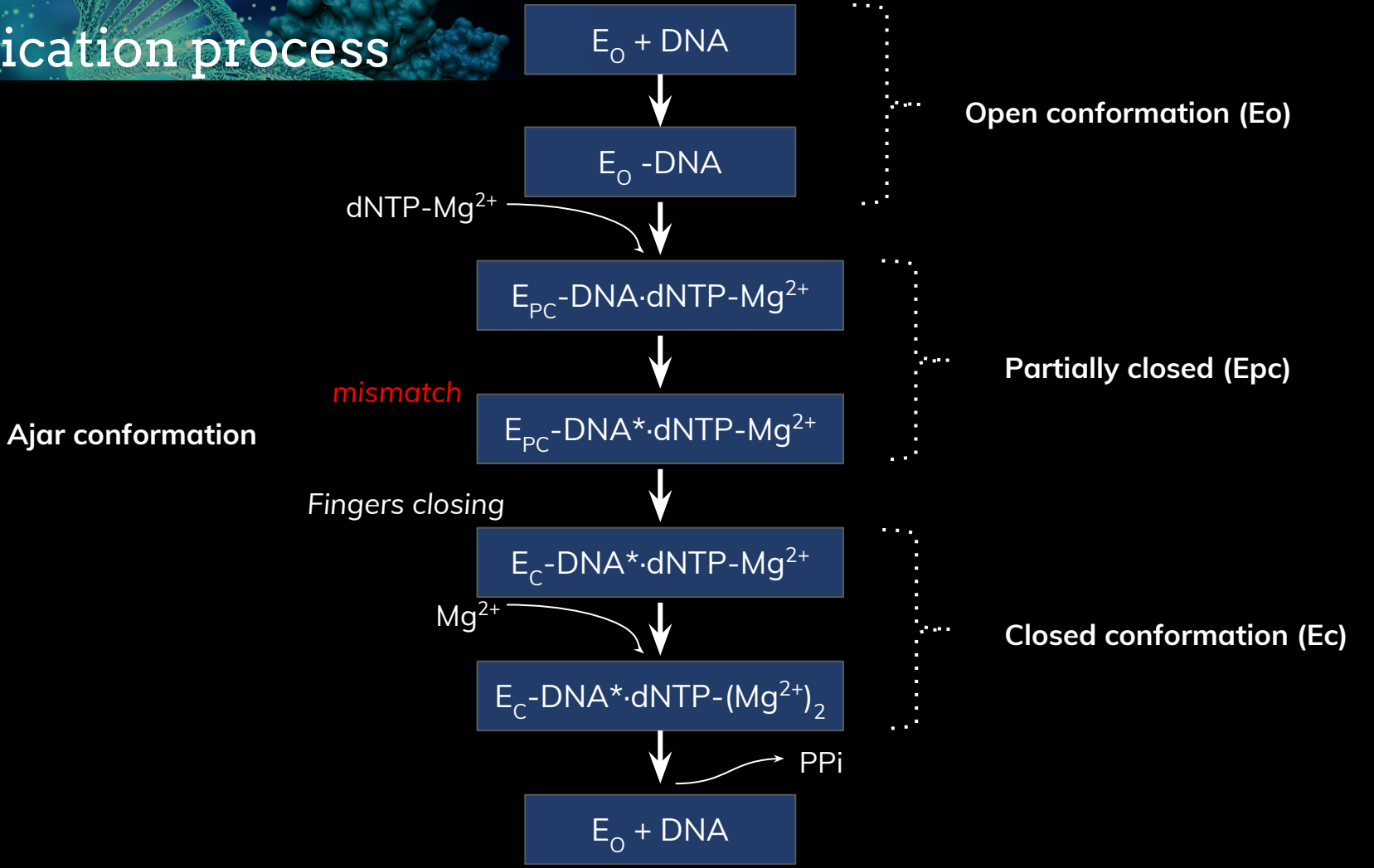
Replication process

1. Binding of the DNA template and primer
2. Binding to an incoming dNTP
3. Phosphodiester bond formation
4. Release of the pyrophosphate
5. Translocation to the next 3'-OH primer terminus



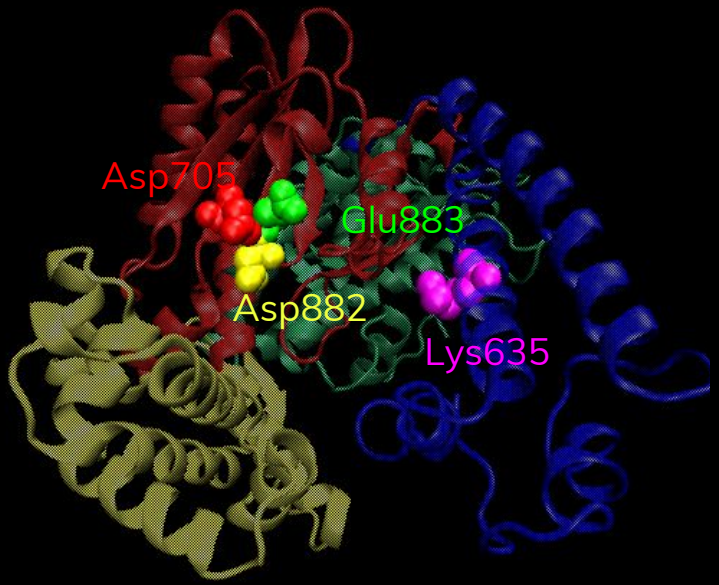
Schematic diagram of DNA polymerase. Modified from Molecular Biology of the Cell, 6th edition by B. Alberts et al.

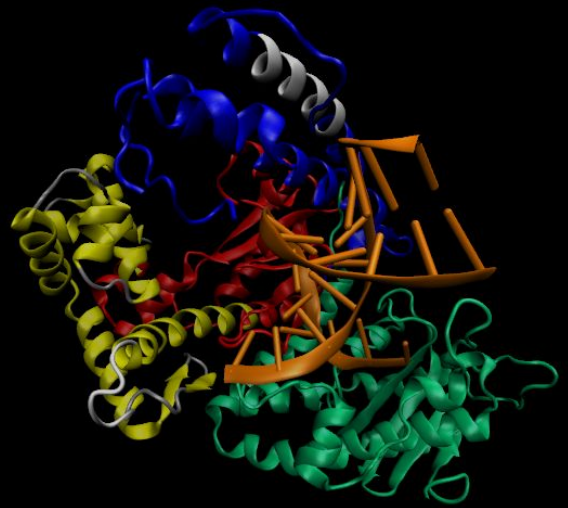
Replication process

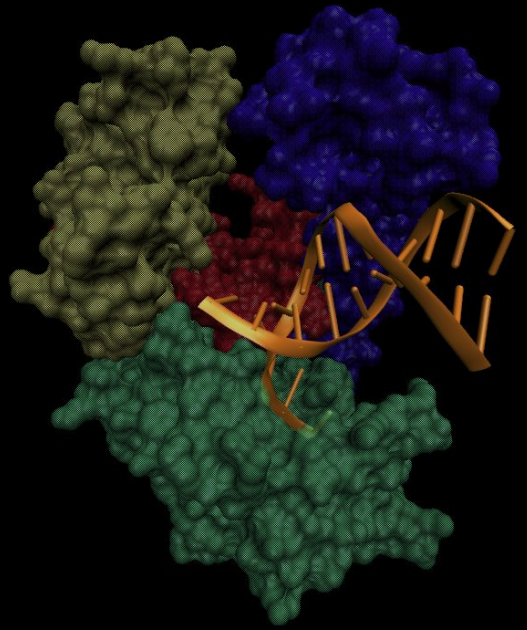


Binding of the DNA template and primer

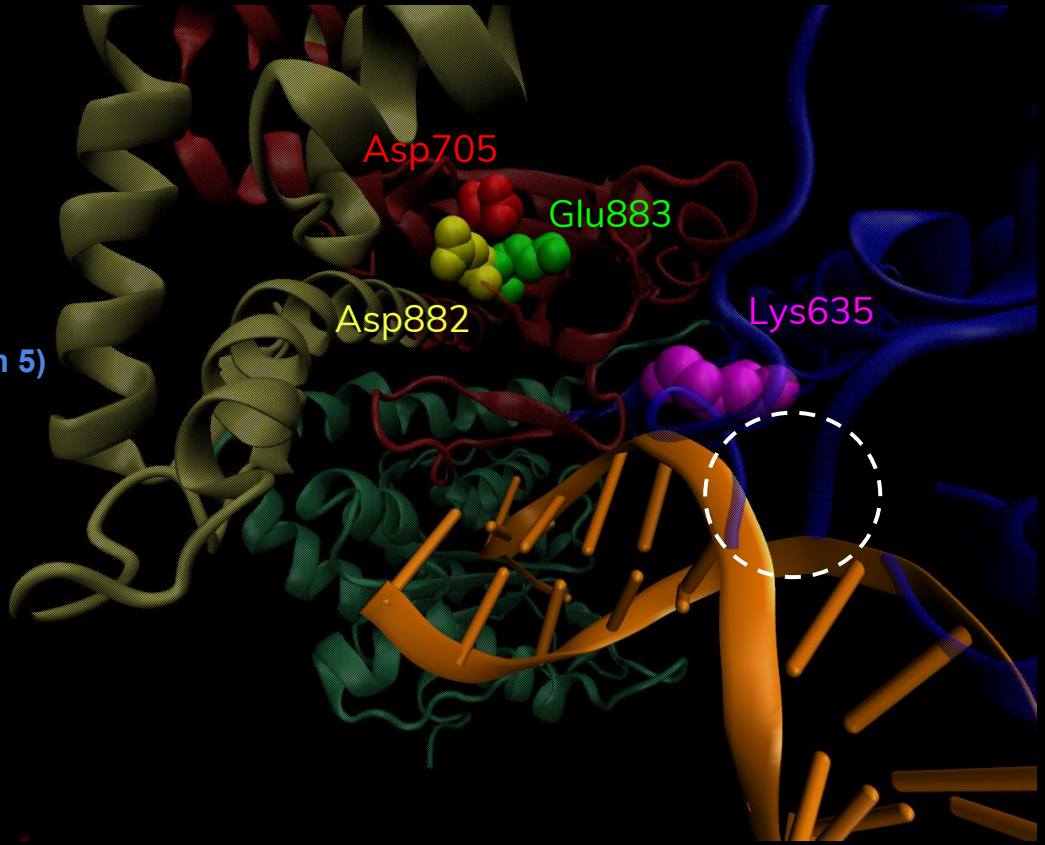
Main regions of primer binding:







Binding of the DNA template and primer



Asp705 (Region 3)

T	F	L	A	A	D	F	S	Q	I	E	L	R		
E	L	M	T	A	D	Y	S	Q	I	E	M	R		
T	F	L	A	A	D	F	S	Q	I	E	L	R		
V	I	V	S	A	D	Y	S	Q	I	E	L	R		
T	F	V	I	S	D	F	S	Q	I	E	L	R		
K	L	I	S	A	D	Y	S	Q	I	E	L	R		
C	L	L	G	V	D	Y	S	Q	I	E	L	R		

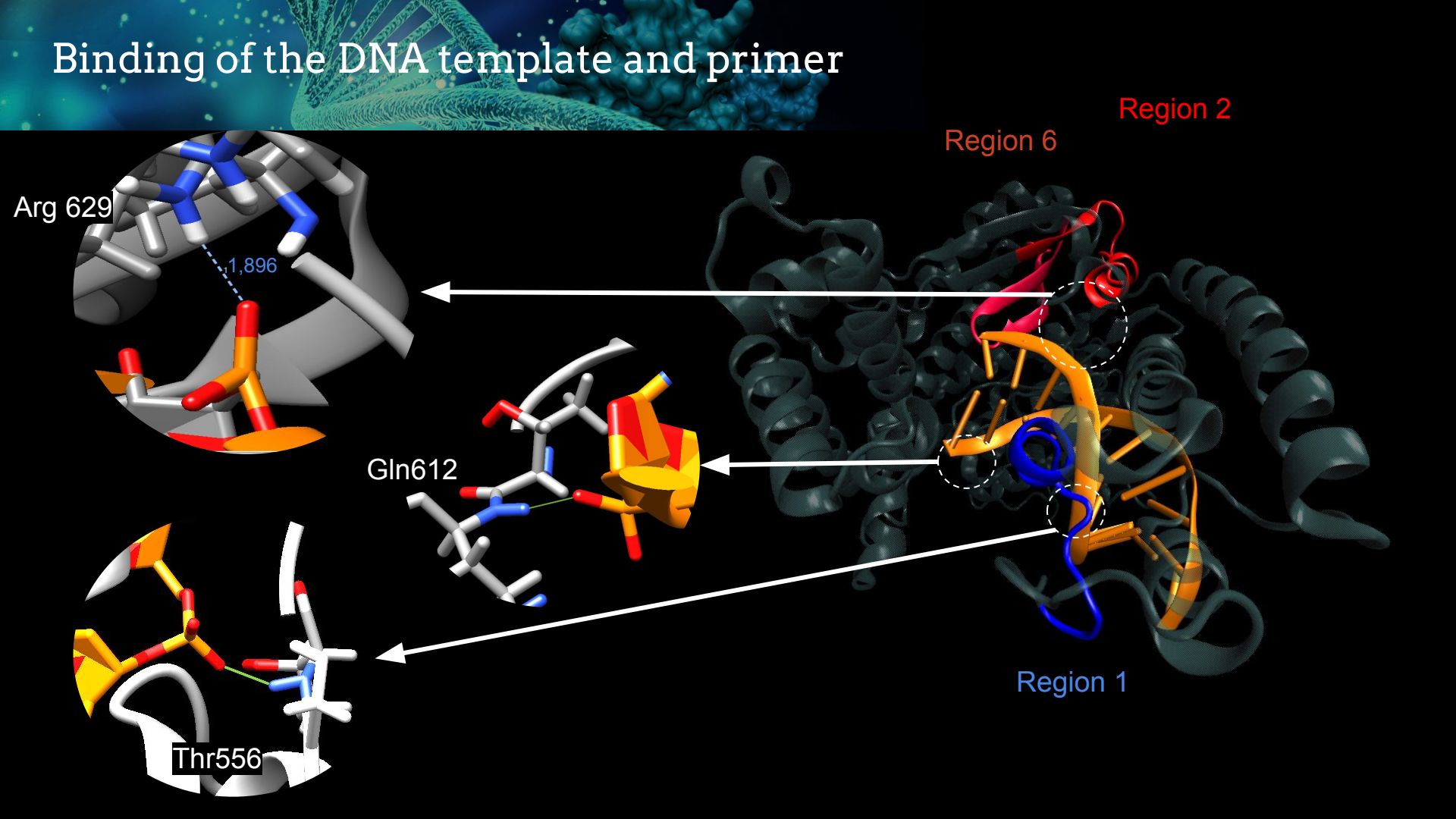
Asp882 (Region 5)
Glu883

A	Q	I	H	D	E	L	L	F	E					
L	Q	V	H	D	E	L	L	F	E					
A	Q	I	H	D	E	L	L	F	E					
M	Q	V	H	D	E	L	V	F	E					
N	L	V	H	D	E	I	V	V	E					
L	Q	I	H	D	E	L	L	F	E					
L	Q	V	H	D	E	L	I	F	E					

Lys635

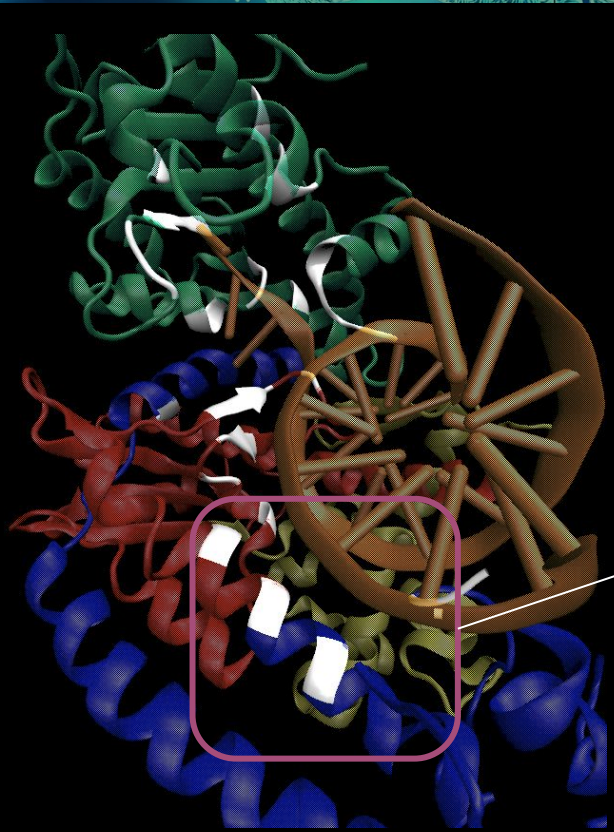
L	E	Y	R	Q	V	H	K	I	K	S	T			
L	A	H	R	D	V	T	R	L	K	V	T			
L	E	Y	R	Q	V	H	K	I	K	S	T			
L	E	Y	R	G	L	A	K	L	K	S	T			
L	E	I	R	K	L	K	K	I	A	D	K			
L	R	W	R	Q	L	T	K	L	K	N	T			
L	E	Y	R	E	L	N	K	L	F	N	T			

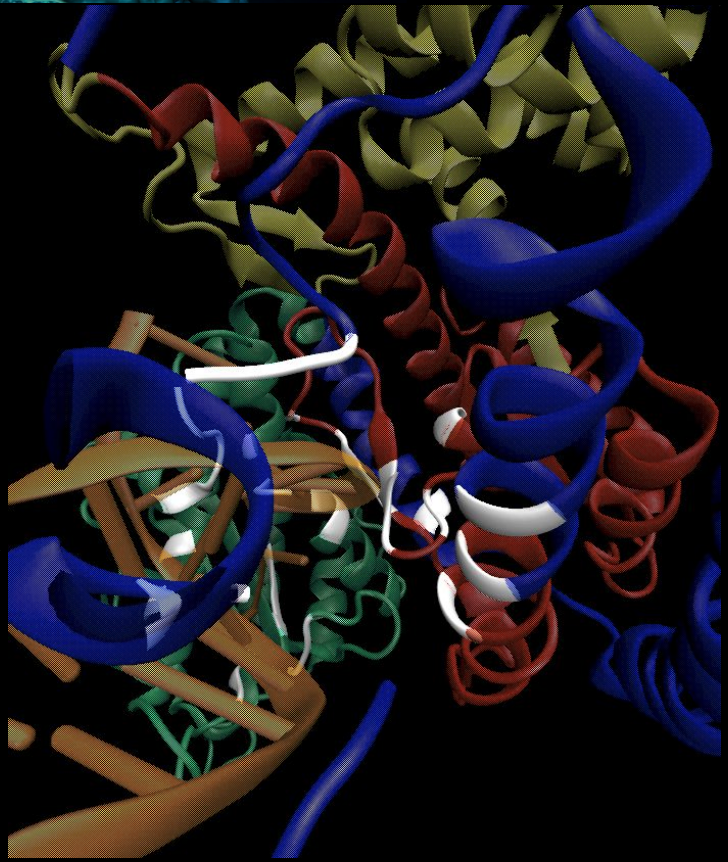
Binding of the DNA template and primer



DNA binding site

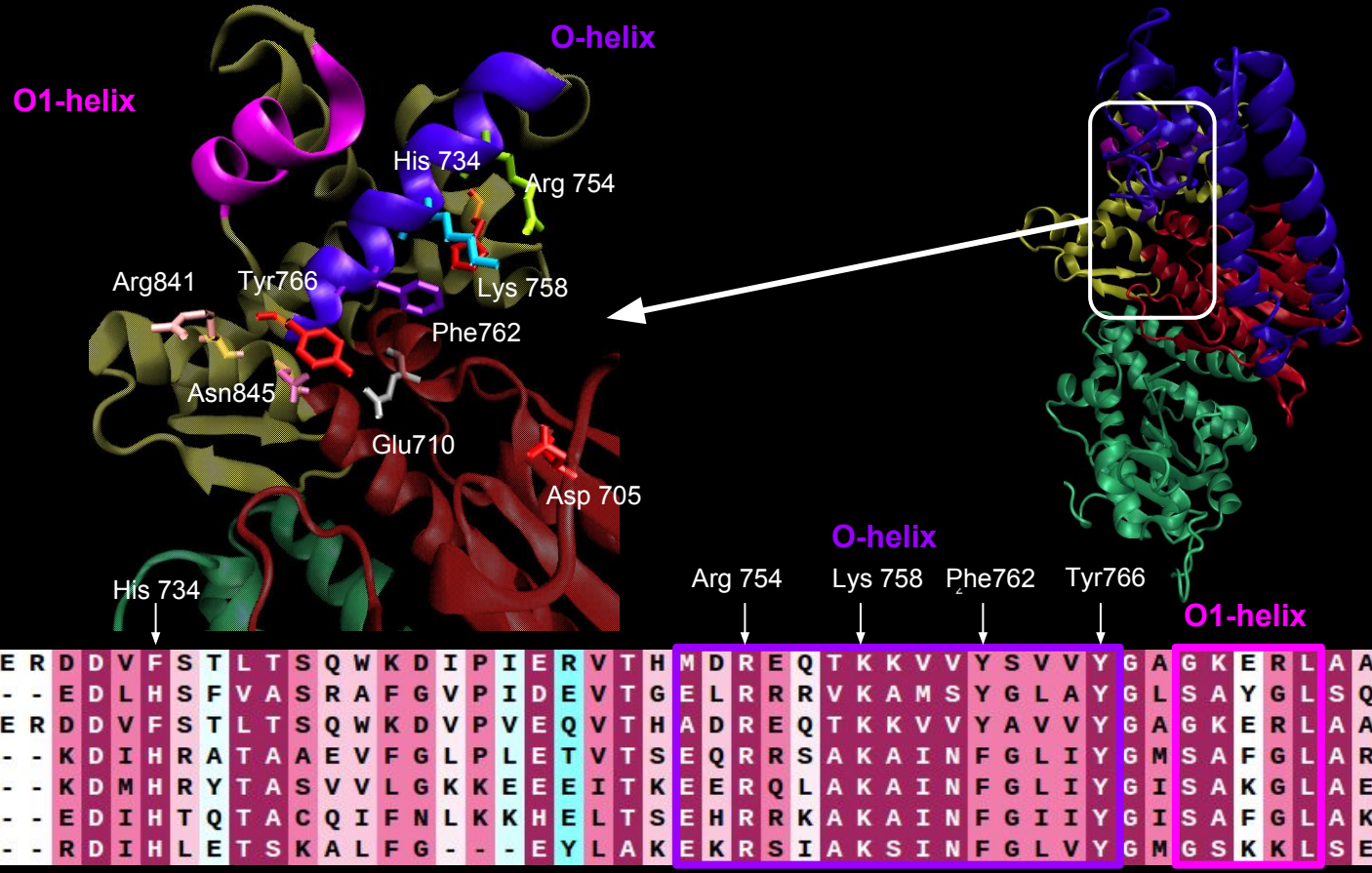
DISPLAR+ prediction





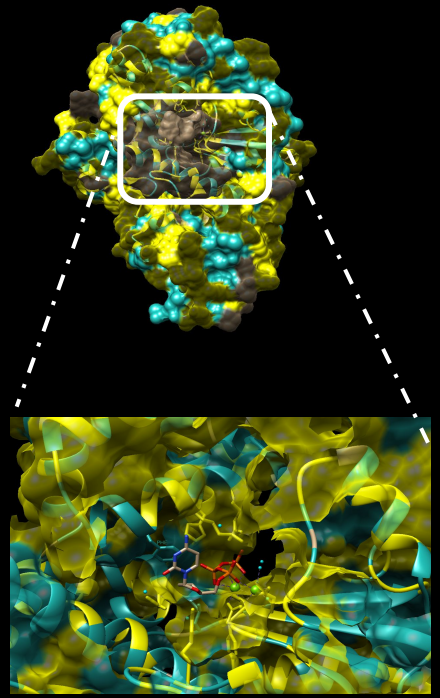
T356, E357, T358, S360, L361, Q419, N420, K422, Y423, M443, R455, H456, D457, M458, D459, F473, E474, F486, Y497, D501, E541, Q585, T609, S610, E611, R631, K635, T639, S658, Y659, H660, Q661, V663, T672, D673, N675, N678

Deoxynucleotides binding



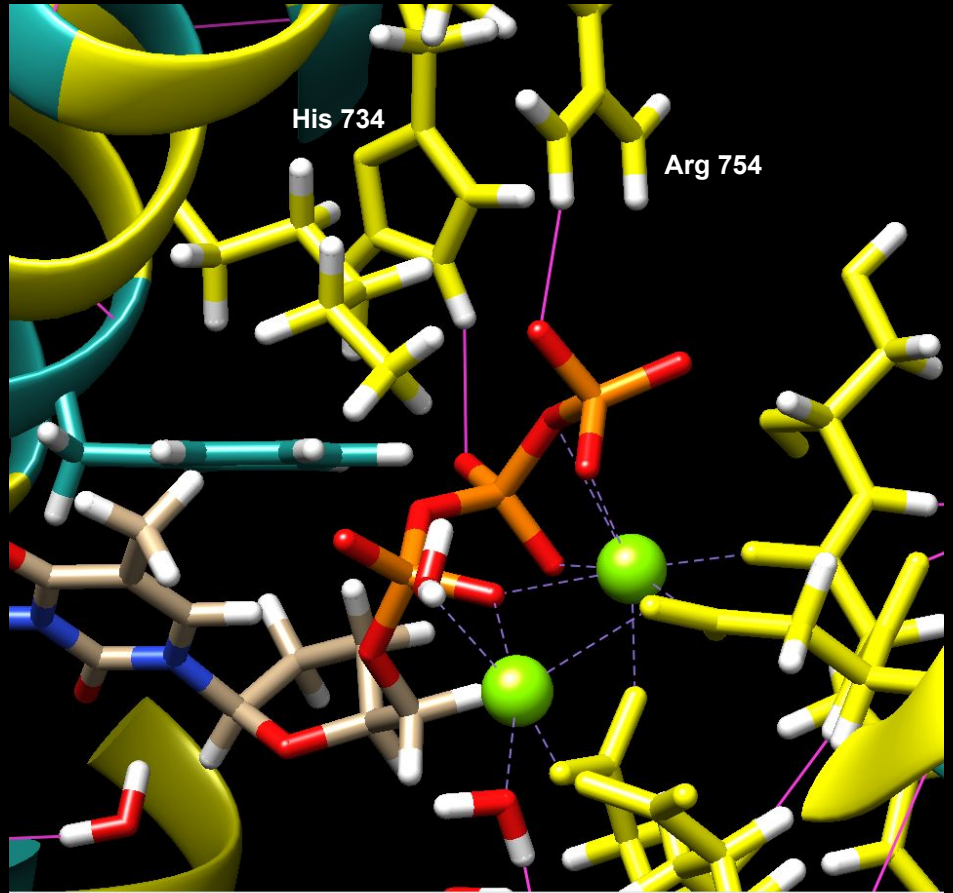
Deoxynucleotides binding

Hydrophobic pocket

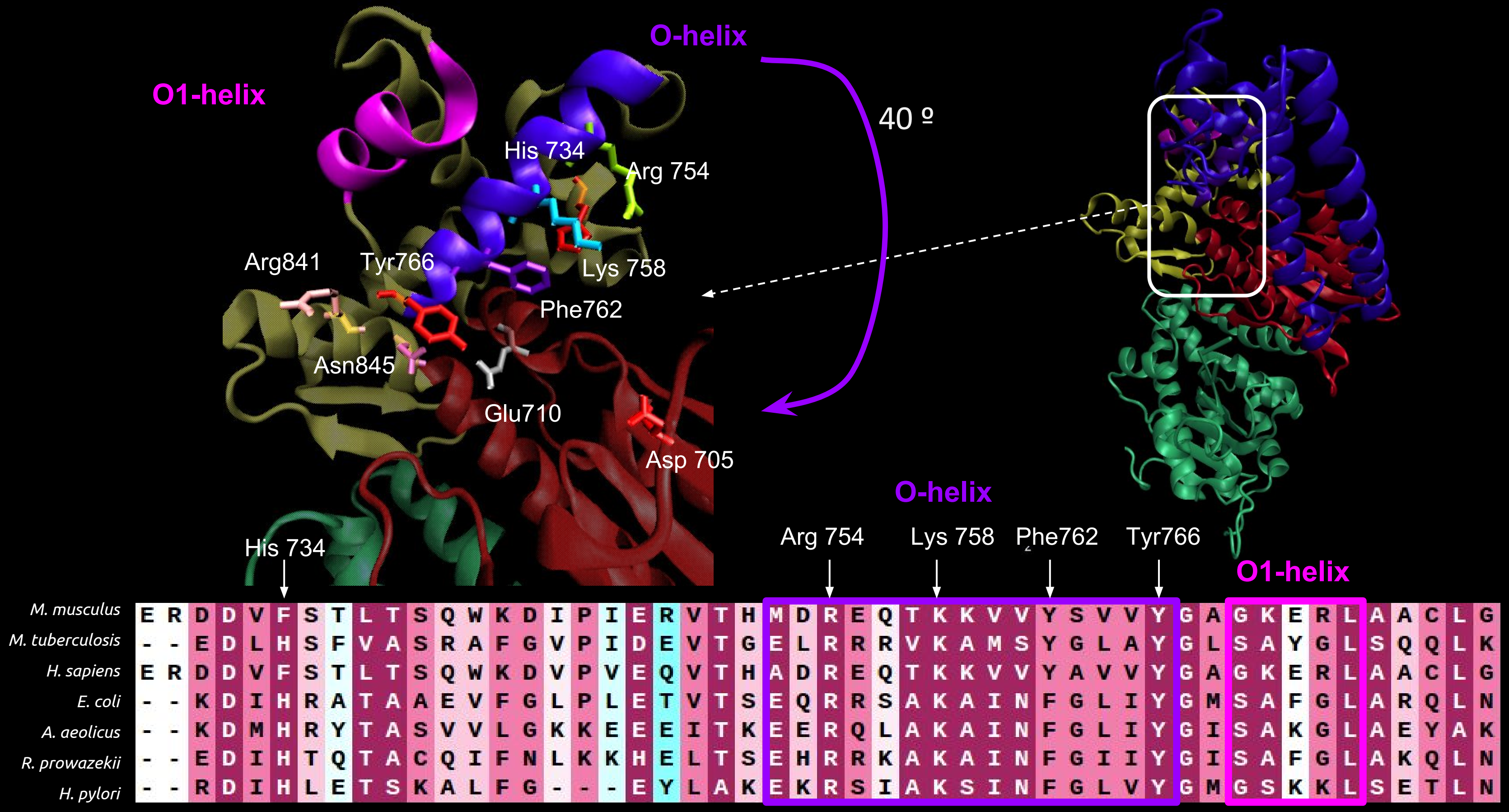


Deoxynucleotides binding

Hydrophilic pocket



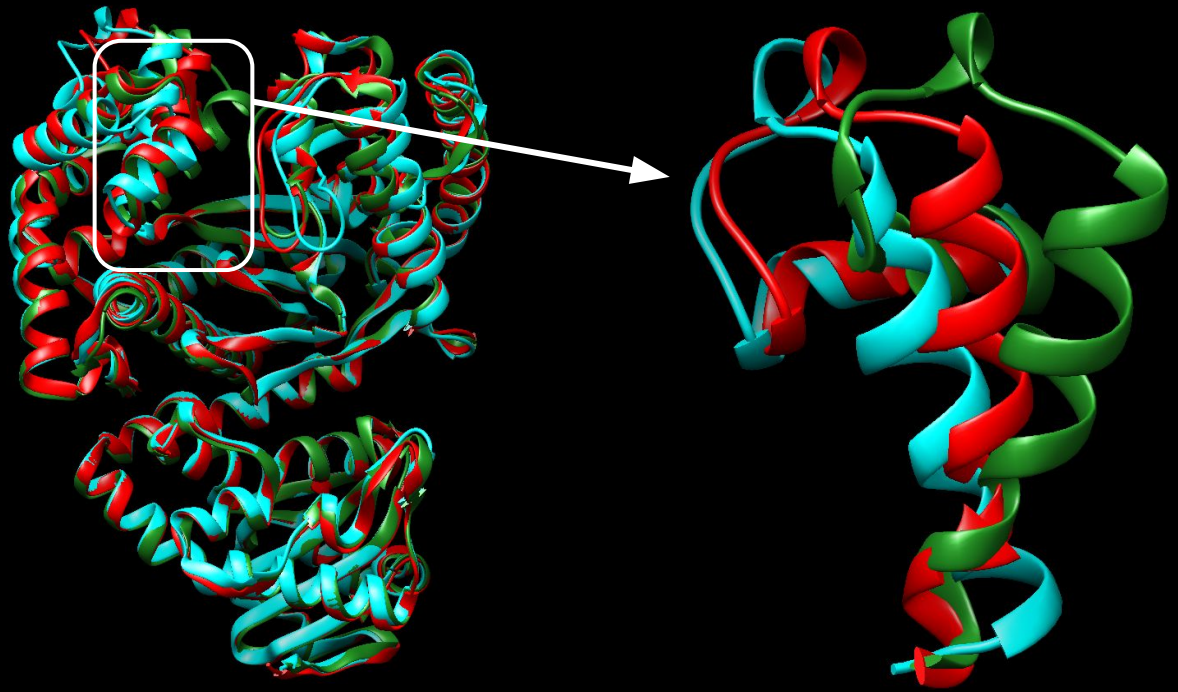
Deoxynucleotides binding

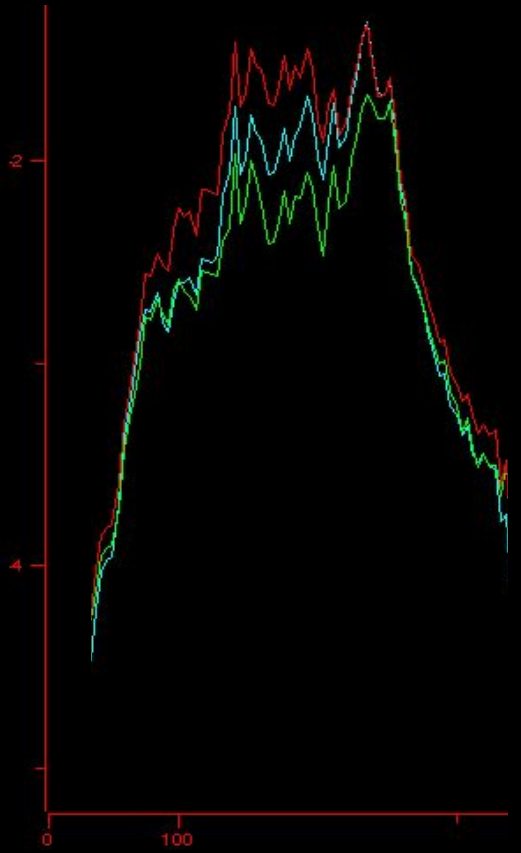


Deoxynucleotides incorporation

Conformational states

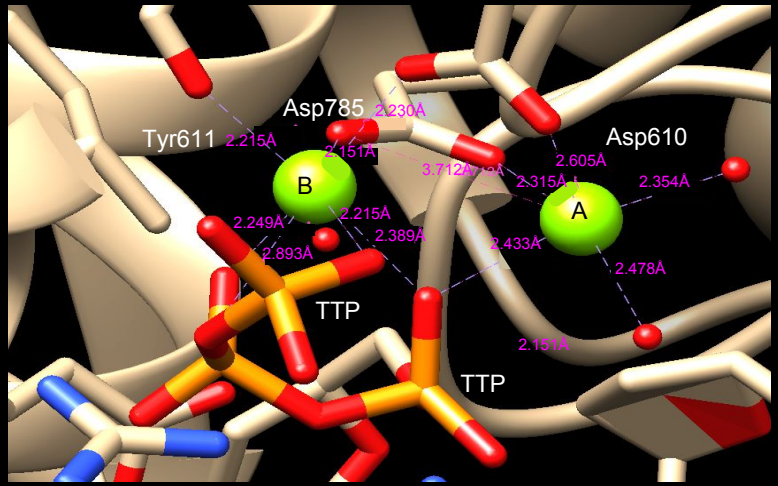
Open binary (1L3S)
Closed ternary (1LV5)
"Ajar" (3HP6)

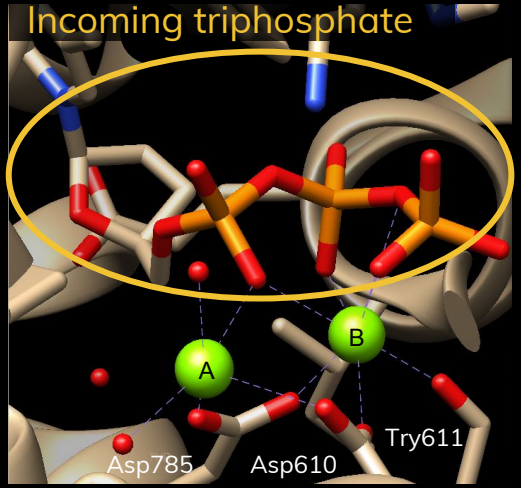




Deoxynucleotides incorporation

Nucleophilic attack





M. musculus	E E E T V T I S P R T L F V S S E G H T F L A A D F S Q I E L R I L A H L S G D P E L L K L F Q E S
M. tuberculosis	- - - - - - - - - D A F V V G D G Y A E L M T A D Y S Q I E M R I M A H L S G D E G L I E A F N T G
H. sapiens	E D K I L T I S P R A M F V S S K G H T F L A A D F S Q I E L R I L T H L S G D P E L L K L F Q E S
E. coli	- - - - - - - - - Q A F I A P E D Y V I V S A D Y S Q I E L R I M A H L S R D K G L L T A F A E G
A. aeolicus	- - - - - - - - - - I F K A E E G N T F V I S D F S Q I E L R I A A E Y V K D P L M L D A F K K G
R. prowazekii	- - - - - - - - - E A F I A E E G Y K L I S A D Y S Q I E L R I L S H I A N I D V L K Q A F I N K
H. pylori	- - - - - - - - K G F I A S S K E Y C L L G V D Y S Q I E L R L L A H F S Q D K D L M E A F L K G

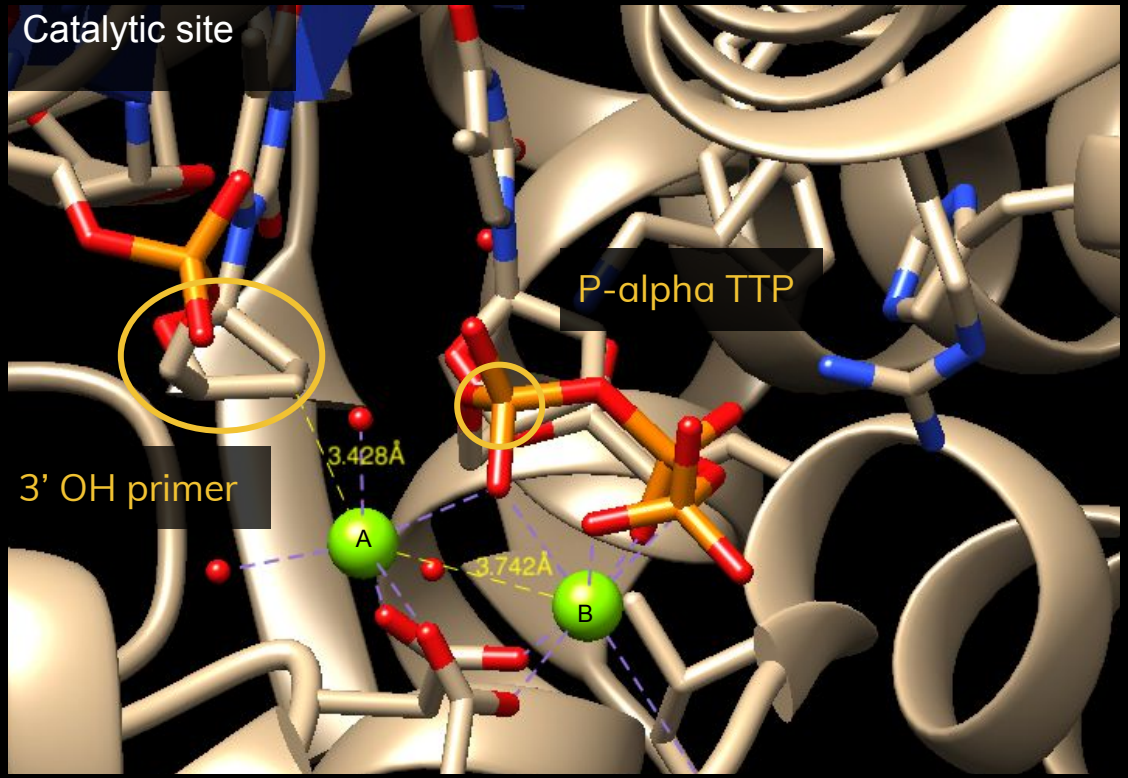
Asp610
Region 3

M. musculus	A Q I H D E L L F E V E D T Q V P E F A A L V R R I M E S L Q Q V Q T L E L Q L Q V P L K V N L S V
M. tuberculosis	L Q V H D E L L F E I A P G E R E R V E A L V R D K M G - - - - G - - - A Y P L D V P L E V S V G Y
H. sapiens	A Q I H D E L L F E V E D P Q I P E C A A L V R R T M E S L E Q V Q A L E L Q L Q V P L K V S L S A
E. coli	M Q V H D E L V F E V H K D D V D A V A K Q I H Q L M E N - - - - - - - C T R L D V P L L V E V G S
A. aeolicus	N L V H D E I V V E C E K E K A E E V K E I L E K S M K T A G - - - - K I I L K E V P V E V E S V I
R. prowazekii	L Q I H D E L L F E V P E I E V E I V I P I I K K I M E H - - - - - - - S T N M N V P I I T E I K A
H. pylori	L Q V H D E L I F E I E E K N A P E L Q Q E I Q R I L N D - - - - - - E V Y P L R V P L E T S A F I

Asp785
Region 5

Deoxynucleotides incorporation

Nucleophilic attack

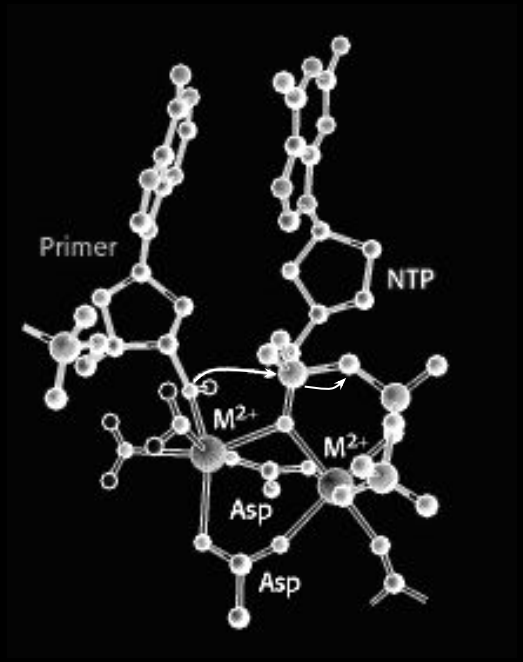


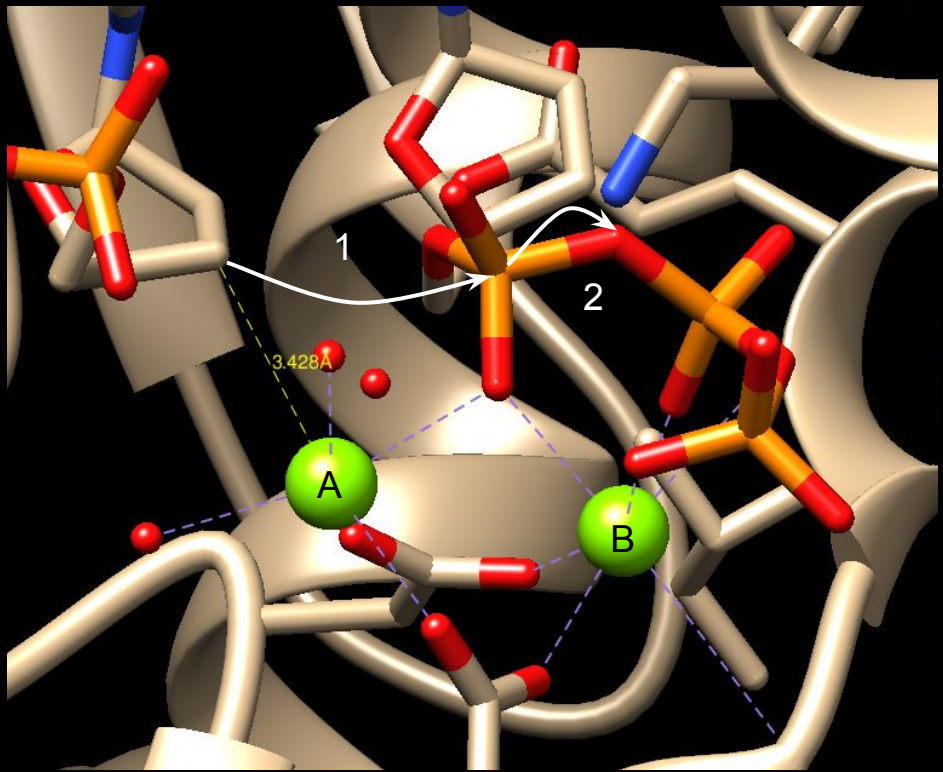
3'OH ➡ Nucleophile

P-alpha ➡ Electrophile

Deoxynucleotides incorporation

Nucleophilic attack



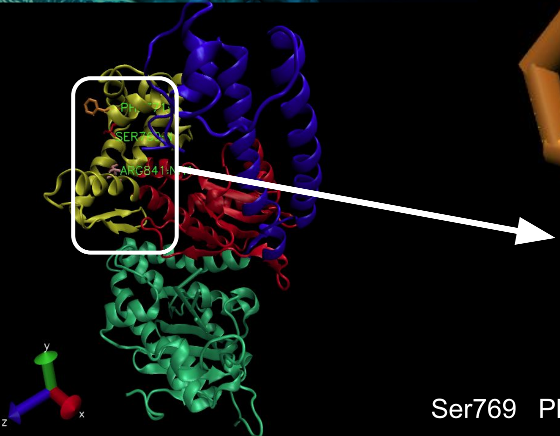


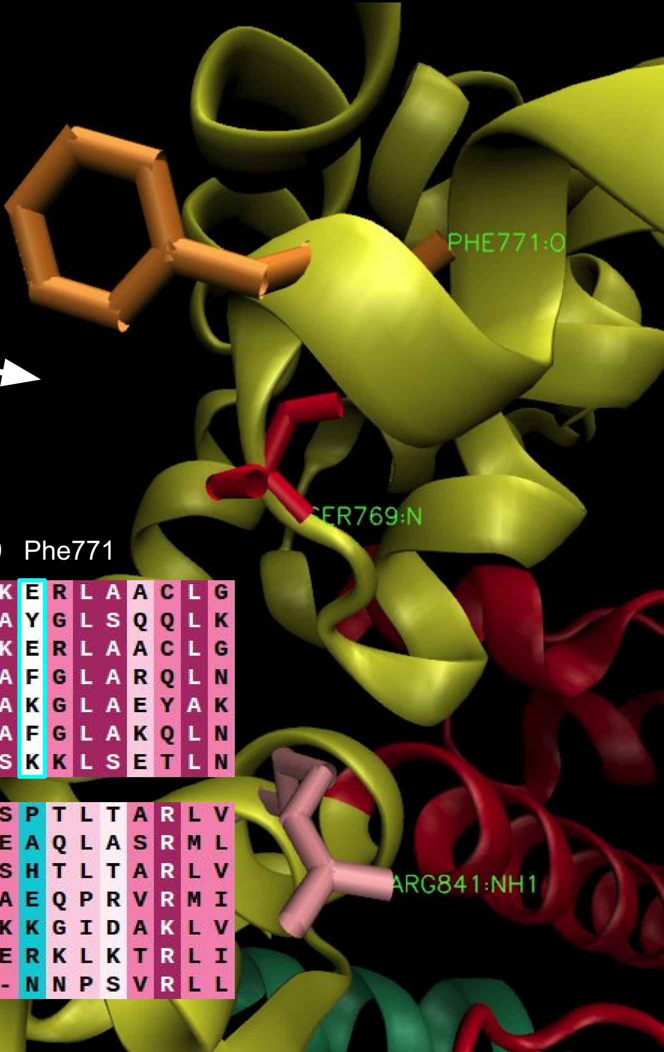
$$(DNA)_{n\ residues} + dNTP \rightarrow (DNA)_{n+1\ residues} + PP_i$$

Berg, Jeremy M, John L Tymoczko, and Lubert Stryer. *Biochemistry*. 1st ed. New York: W.H. Freeman and Co., 2002. Print.

Displacement

The presence of Ser769, Phe771 and Arg841 in the Pol I may be prerequisite for the expression of strand displacement synthesis.





Ser769 Phe771

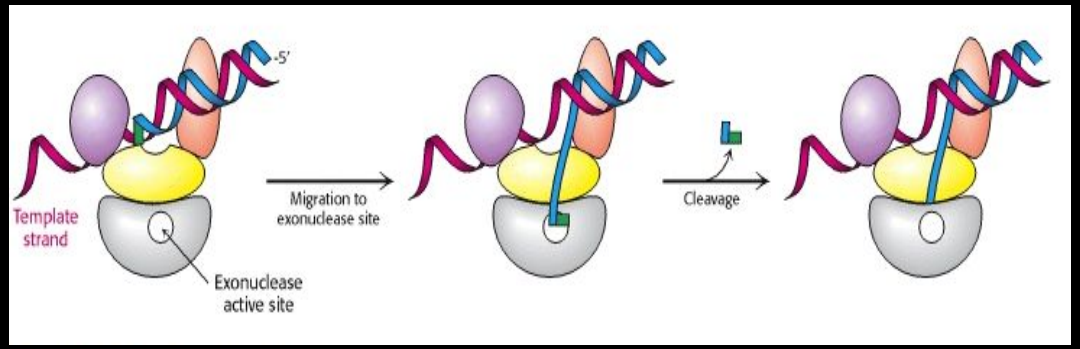
Species	Sequence
M. musculus	E R D D V F S T L T S Q W K D I P I E R V T H M D R E Q T K K V V Y S V V Y G A G K E R L A A C L G
M. tuberculosis	- - E D L H S F V A S R A F G V P I D E V T G E L R R R V K A M S Y G L A Y G L S A Y G L S Q Q L K
H. sapiens	E R D D V F S T L T S Q W K D V P V E Q V T H A D R E Q T K K V V Y A V V Y G A G K E R L A A C L G
E. coli	- - K D I H R A T A A E V F G L P L E T V T S E Q R R S A K A I N F G L I Y G M S A F G L A R Q L N
A. aeolicus	- - K D M H R Y T A S V V L G K K E E E I T K E E R Q L A K A I N F G L I Y G I S A K G L A E Y A K
R. prowazekii	- - E D I H T Q T A C Q I F N L K K H E L T S E H R R K A K A I N F G I I Y G I S A F G L A K Q L N
H. pylori	- - R D I H L E T S K A L F G - - - E Y L A K E K R S I A K S I N F G L V Y G M G S K K L S E T L N

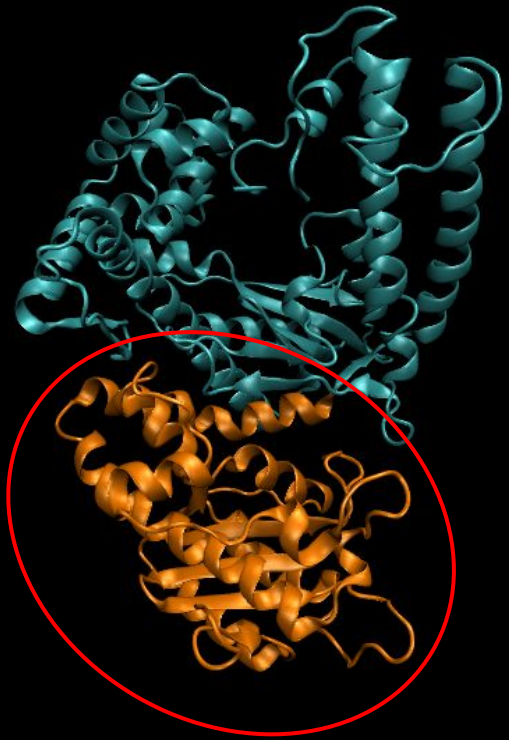
Species	Sequence
M. musculus	C A Q D Q Q L R A Q A E R Q A V N F V V Q G S A A D L C K L A M I R I S T A V A T S P T L T A R L V
M. tuberculosis	D S S N R Q V R E A A E R A A L N A P I Q G S A A D I I K V A M I Q V D K A L N - E A Q L A S R M L
H. sapiens	H A H D Q Q L R A Q A E R Q A V N F V V Q G S A A D L C K L A M I H V F T A V A A S H T L T A R L V
E. coli	K S S N G A R R A A A E R A A I N A P M Q G T A A D I I K R A M I A V D A W L Q - A E Q P R V R M I
A. aeolicus	G - - - R R F S A N T F N D A V N Y P I Q G T G A D L L K L A V L L F D A N L Q - K K G I D A K L V
R. prowazekii	H - - D K K L K Q F A E R A A I N A P I Q G T N A D I I K I A M I K L D K E I E - E R K L K T R L I
H. pylori	G - A N D Y V K G N Y L R E G V N A I F Q G S A S D L L K L G M L K V S E R F K - - N N P S V R L L

Arg841

3′ → 5′ exonuclease domain

- **Location:** N-terminus
- **Length:** 225 residues
- **Catalytic function:** 3’ → 5’ nucleotide excision (*proofreading*)

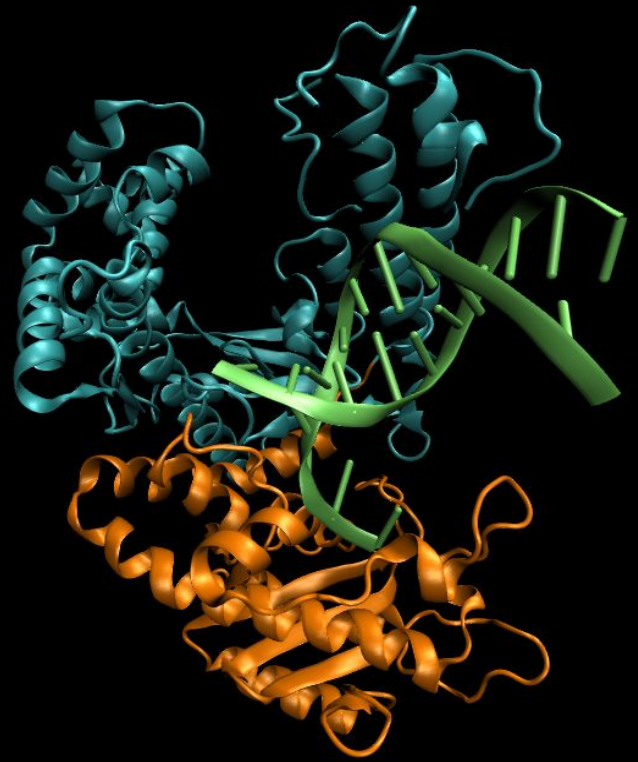


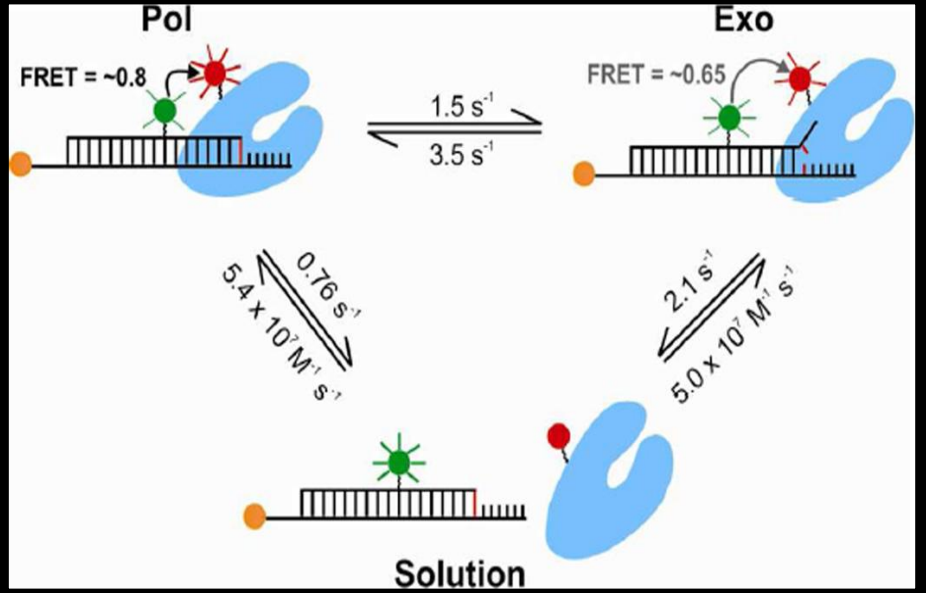


Source: Berg JM, Tymoczko JL, Stryer L. Biochemistry, 5th edition. New York: W H Freeman; 2002.

3′ → 5′ exonuclease domain

Fidelity and site switching

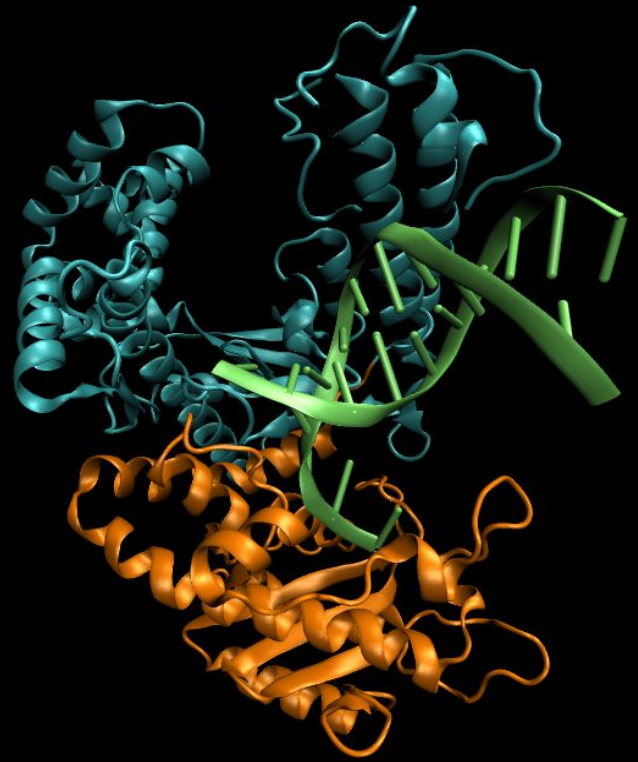


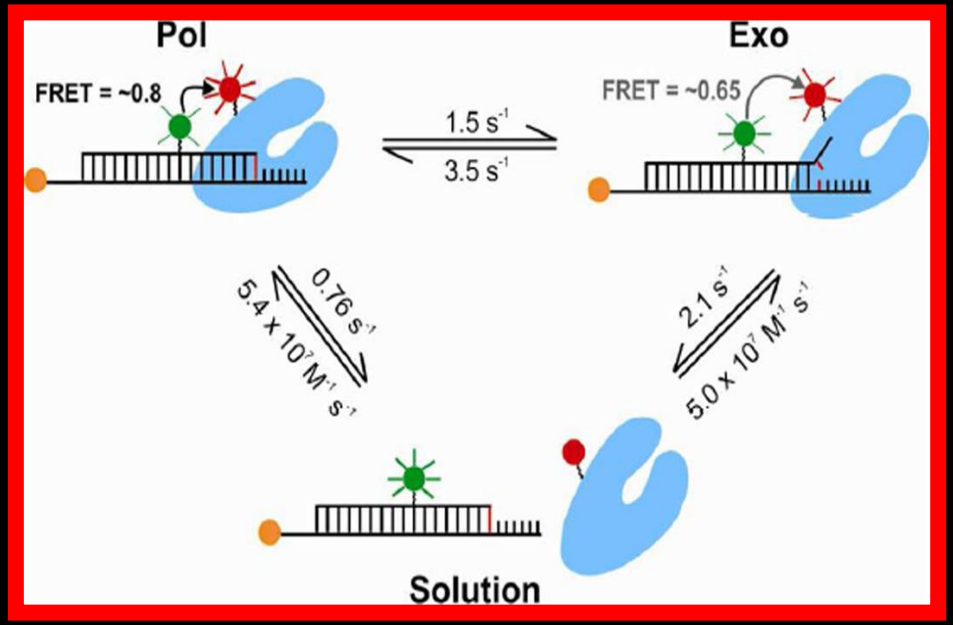


Source: Lamichhane *et al.* Dynamics of site switching in DNA polymerase. J Am Chem Soc. 2013; 135 (12): 4735-42.

3′ → 5′ exonuclease domain

Fidelity and site switching

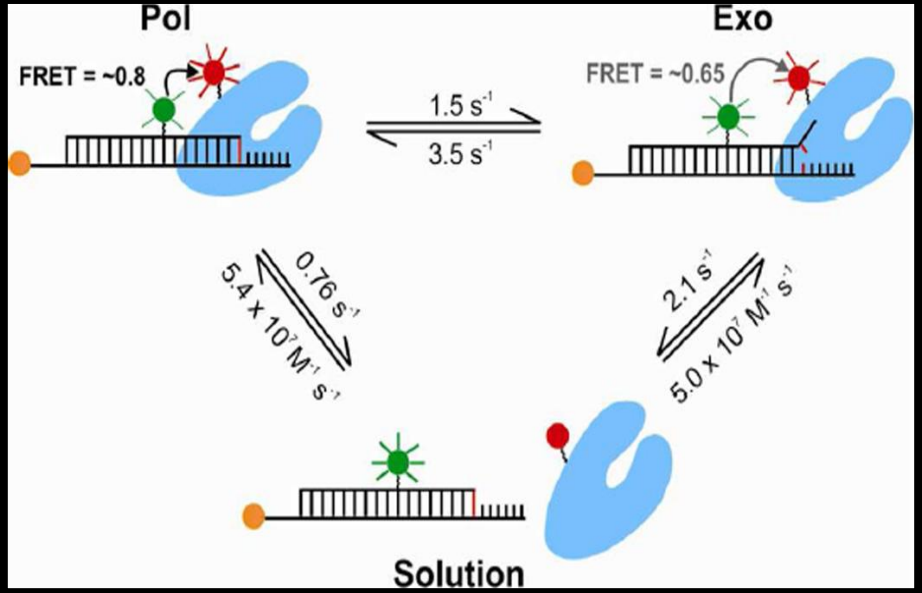




Source: Lamichhane *et al.* Dynamics of site switching in DNA polymerase. J Am Chem Soc. 2013; 135 (12): 4735-42.

3′ → 5′ exonuclease domain

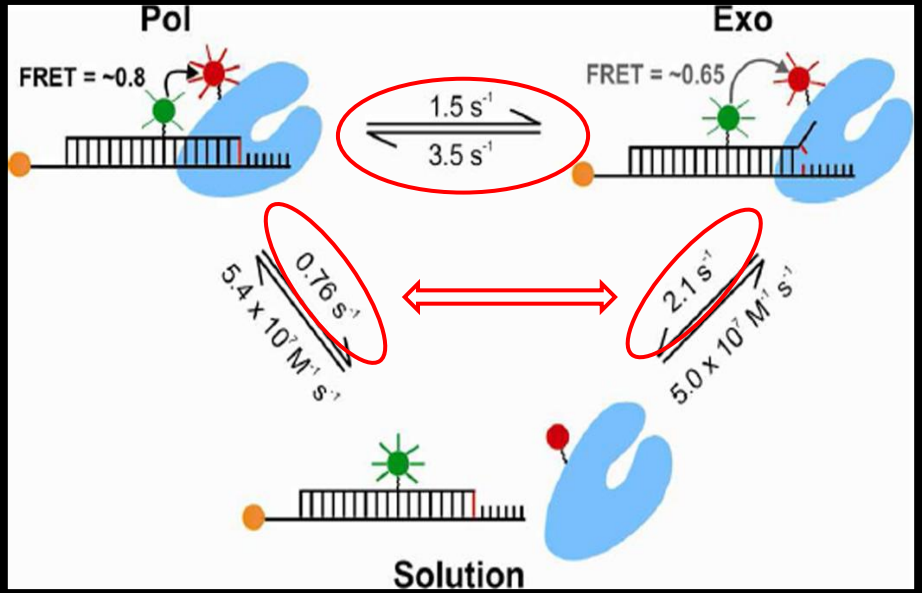
Fidelity and site switching



Source: Lamichhane *et al.* Dynamics of site switching in DNA polymerase. J Am Chem Soc. 2013; 135 (12): 4735-42.

3′ → 5′ exonuclease domain

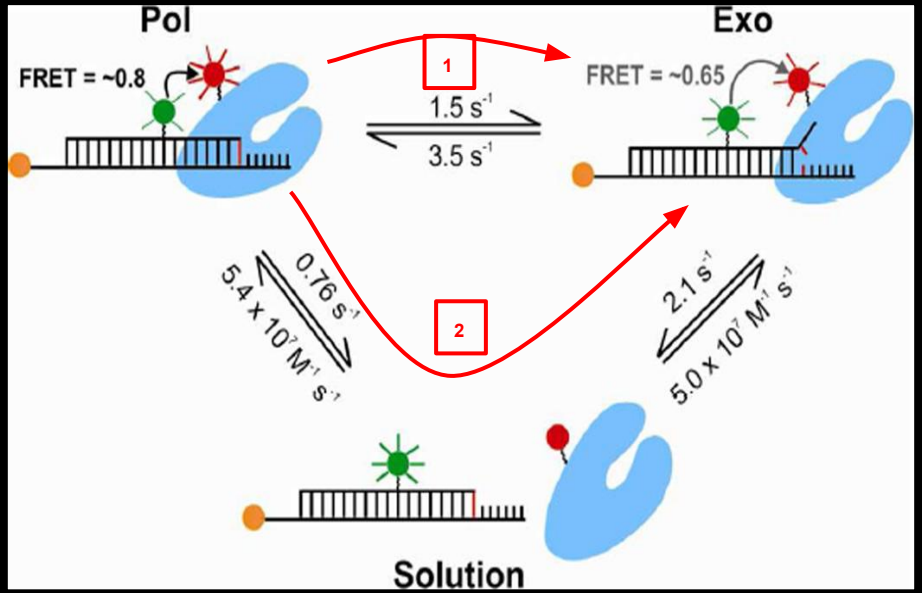
Fidelity and site switching



Source: Lamichhane *et al.* Dynamics of site switching in DNA polymerase. J Am Chem Soc. 2013; 135 (12): 4735-42.

3′ → 5′ exonuclease domain

Fidelity and site switching



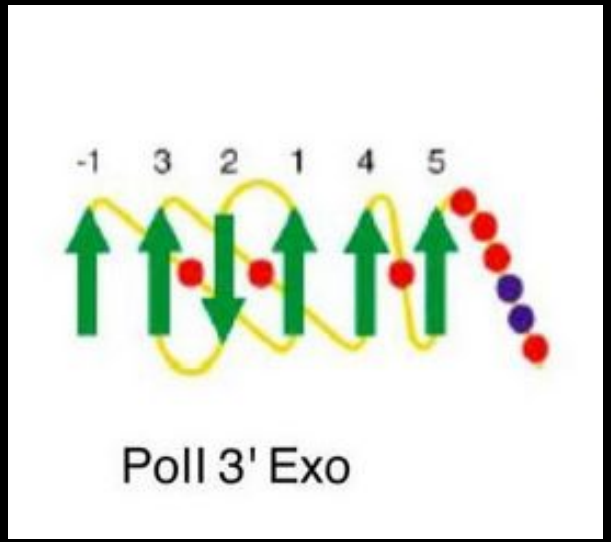
Rate of extension from a mispaired primer strand terminus:
0.027 s−1 for KF

Source: Lamichhane *et al.* Dynamics of site switching in DNA polymerase. J Am Chem Soc. 2013; 135 (12): 4735-42.

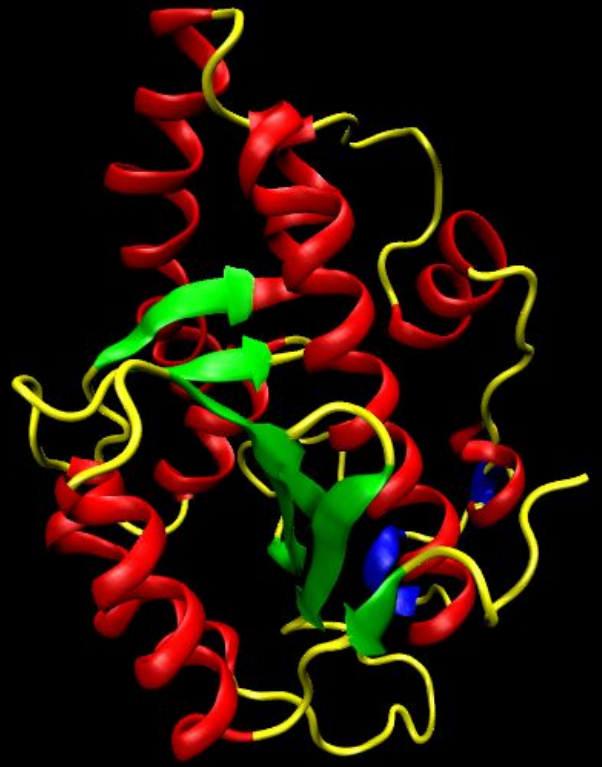
3′ → 5′ exonuclease domain

Structure: RNAse H-like fold

Topology scheme



3D Structure

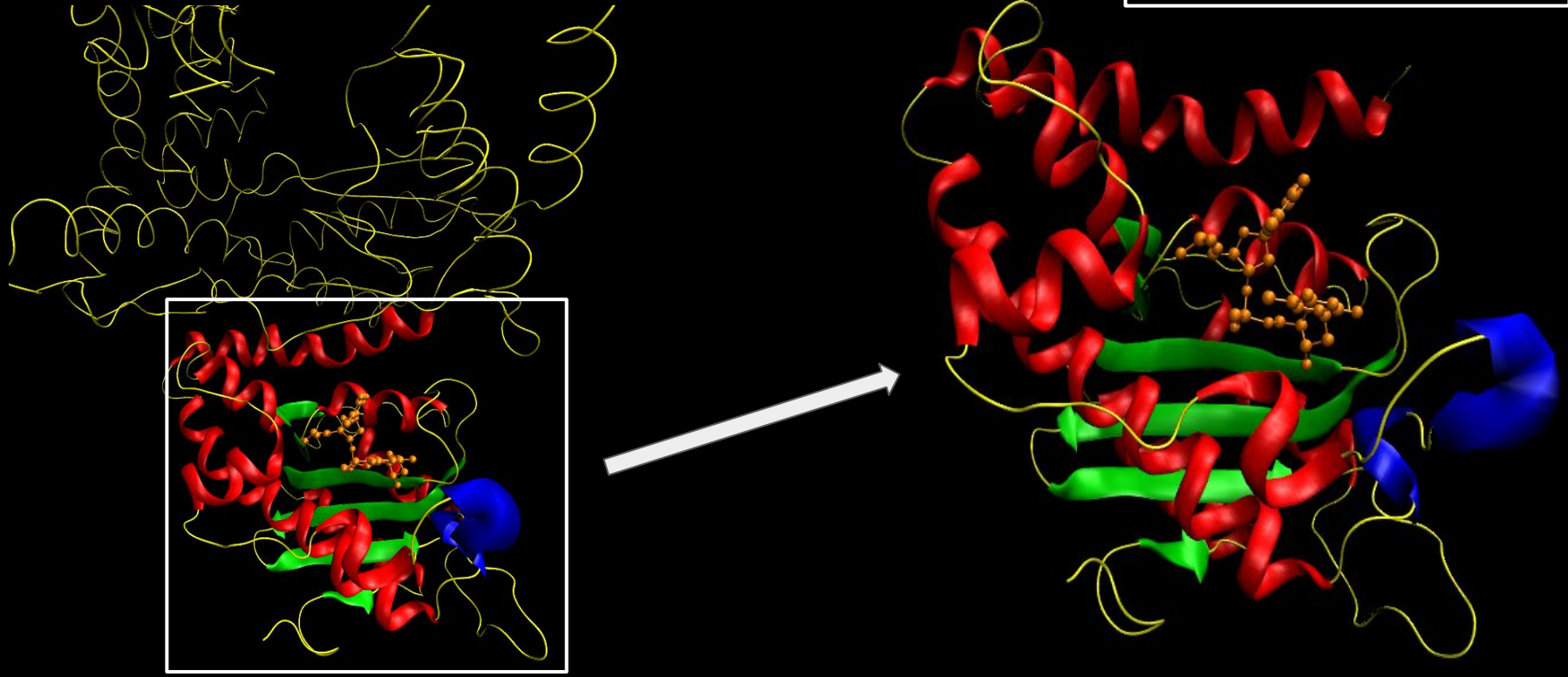


Source: Lovett ST. The DNA polymerases of *E. coli*. EcoSal Plus. 2011 Dec; 4(2)

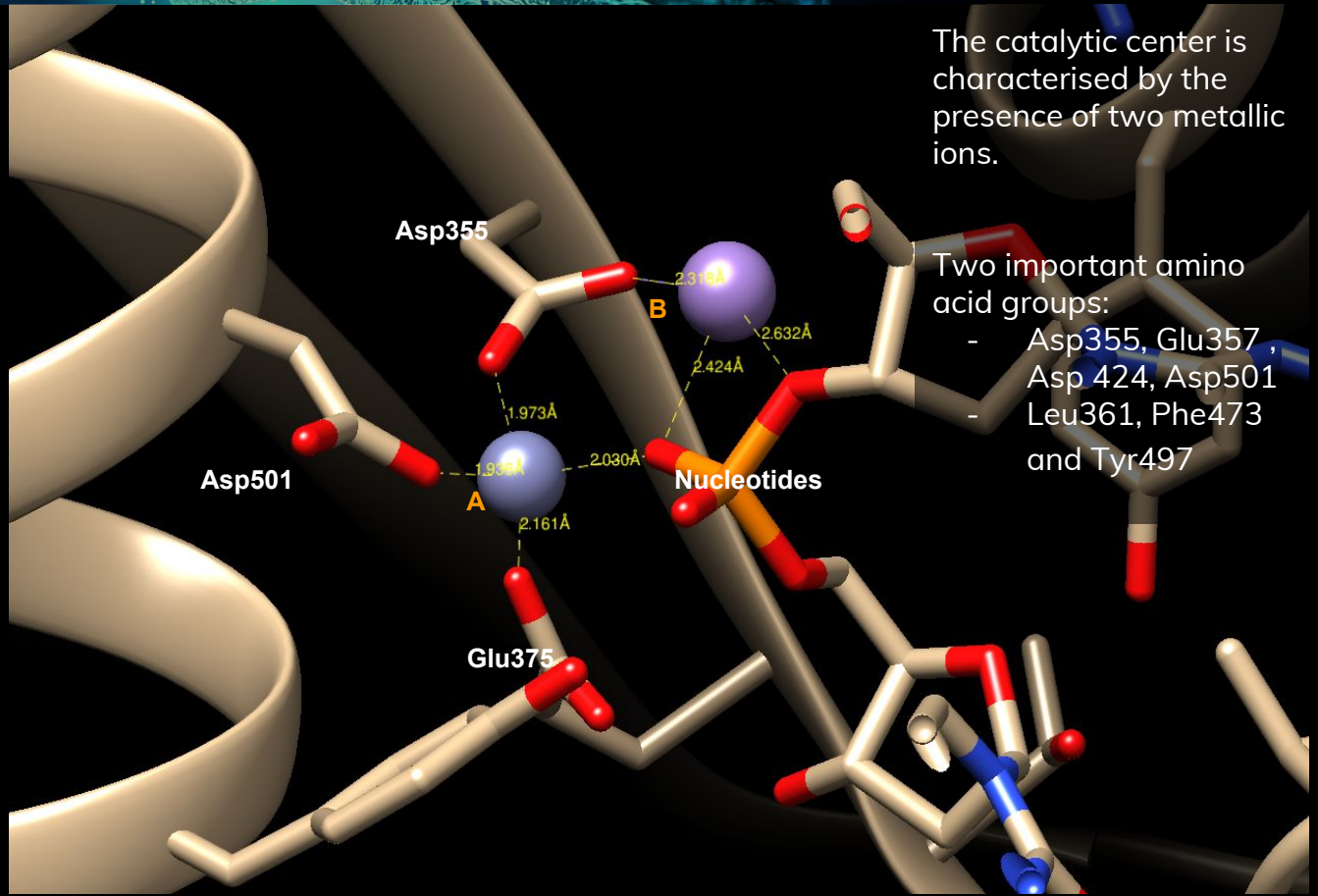
3′ → 5′ exonuclease domain

Active site structure

Alpha helix	3-10 helix
Beta strand	DNA



Zn^{2+} and Mn^{2+} binding



The catalytic center is characterised by the presence of two metallic ions.

Two important amino acid groups:

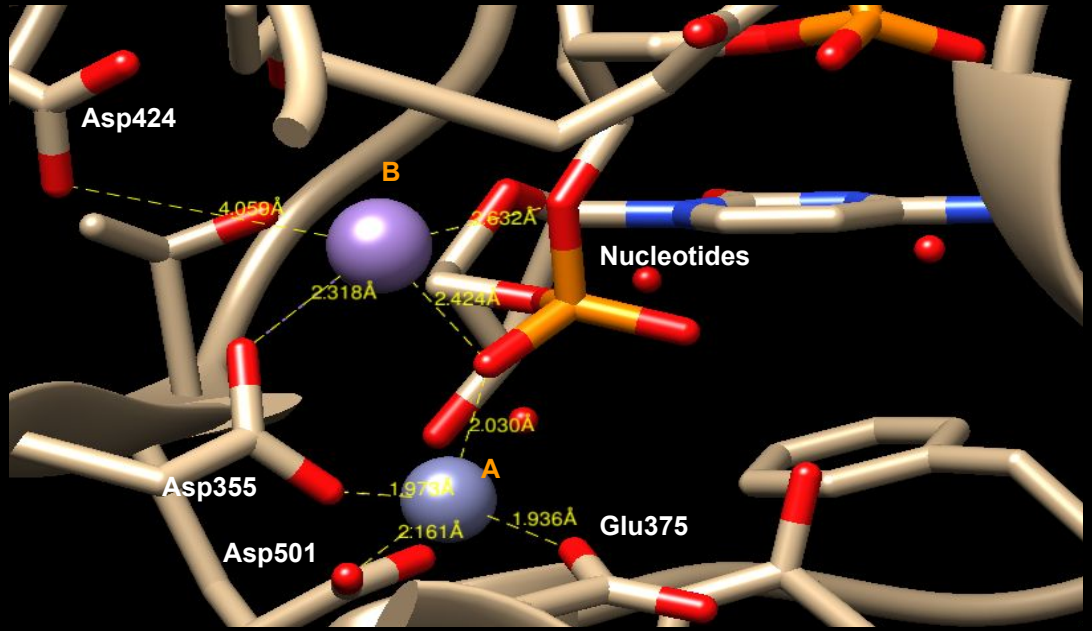
- Asp355, Glu357 , Asp 424, Asp501
- Leu361, Phe473 and Tyr497

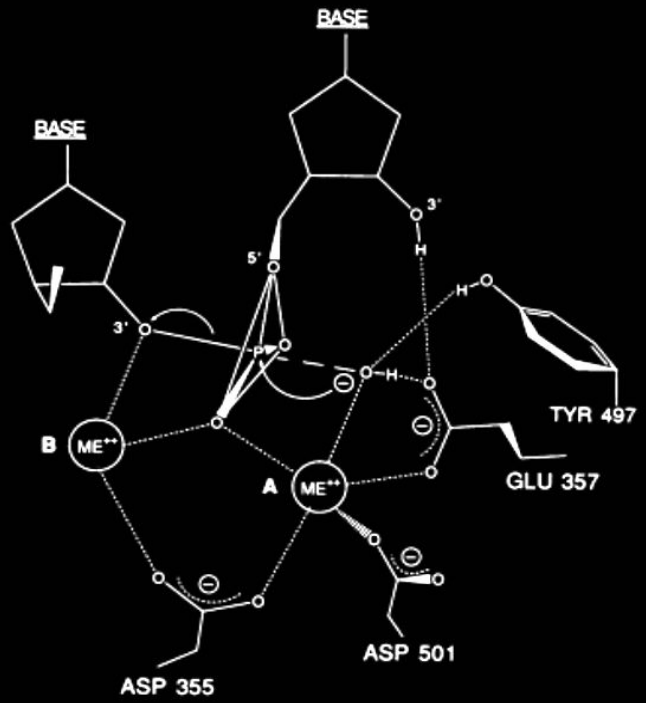
3′ → 5′ exonuclease domain

Active site structure

Excision reaction:

$$[(DNA)_{n\ residues} + H_2O \rightarrow (DNA)_{n-1\ residues} + dNMP]$$





Source: Beese LS, Steitz TA. Structural basis for the 3′-5′ exonuclease activity of Escherichia coli DNA polymerase I: a two metal ion mechanism. EMBO J. 1991; 10(1): 25-33.

Conclusions

- There are highly conserved regions in the sequence of DNA polymerase that play a key role in the development of its function
- Structure is also conserved within prokaryotic organisms
- Incorporation of dNTPs involves hydrophobic and hydrophilic interactions, two metallic cations, the 3'-primer OH and W-C base pairing
- Apart from open and close states, there is an intermediate state called ajar conformation that also appears when a mismatch occurs

Bibliography

1. Astatke M, Grindley ND, Joyce CM. How *E. coli* DNA polymerase I (Klenow fragment) distinguishes between deoxy- and dideoxynucleotides. *J Mol Bio.* 1998;278(1):147-146.

2. Beese LS, Derbyshire V, Steitz TA. Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science.* 1993;260(5106):352-355.

3. Beese LS, Steitz TA. Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *EMBO J.* 1991;10(1):25-33.

4. Bermek O, Grindley ND, Joyce CM. Prechemistry nucleotide selection checkpoints in the reaction pathway of DNA polymerase I and roles of glu710 and tyr766. *Biochemistry.* 2013;52(36):6258-6274.

5. Doublilé S, Ellenberger T. The mechanism of action of T7 DNA polymerase. *Curr Opin Struct Biol.* 1998;8(6):704-712.

6. Foti JJ, Walker GC. SnapShot: DNA polymerases I prokaryotes. *Cell.* 2010;141(1):192-192.e1.

7. Joyce CM, Steitz TA. Function and structure relationships in DNA polymerases. *Annu Rev Biochem.* 1994;63:777-822.

8. Kaushik N, Pandey VN, Modak MJ. Significance of the O-helix residues of *Escherichia coli* DNA polymerase I in DNA synthesis: dynamics of the dNTP binding pocket. *Biochemistry.* 1996;35(22):7256-7266.

9. Lovett ST. The DNA exonucleases of *Escherichia coli*. *EcoSal Plus.* 2011;4(2):10.1128/ecosalplus.4.4.7.

10. Meli M, Sustarsic M, Craggs TD, Kapanidis AN, Colombo G. DNA polymerase conformational dynamics and the role of fidelity-conferring residues: insights from computational simulations. *Front Mol Biosci.* 2016;3:20.

11. Miller BD, Beese LS, Parish CA, Wu EY. The closing mechanism of DNA polymerase I at atomic resolution. *Structure.* 2015;23(9):1609-1620.

Bibliography

12. Patel PH, Loeb LA. Getting a grip on how DNA polymerases function. *Nat Struct Biol*. 2001;8(8):656-659.

13 Patel PH, Suzuki M, Adman E, Shinkai A, Loeb LA. Prokaryotic DNA polymerase I: evolution, structure, and “base flipping” mechanism for nucleotide selection. *J Mol Bio*. 2001;308(5):823-837.

14. Singh K, Srivastava A, Patel SS, Modak MJ. Participation of the fingers subdomain of *Escherichia coli* DNA polymerase I in the strand displacement synthesis of DNA. *J Biol Chem*. 2007;282(14):10594-10604.

15. Wu EY, Beese L. The structure of a high fidelity DNA polymerase bound to a mismatched nucleotide reveals an "ajar" intermediate conformation in the nucleotide selection mechanism. *J Biol Chem*. 2011;286(22):19758-19767.

16. Yang W, Lee JY, Nowotny M. Making and breaking nucleic acids: two -Mg^{2+}-ion catalysis and substrate specificity. *Mol Cell*. 2006;22(1):5-13.

17. Berg JM, Tymoczko JL, Stryer L. Biochemistry, 5th edition. New York: W H Freeman; 2002.

18. Martina CE, Lapenta F, Silva AM, Hochkoeppler A. HoLaMa: A Klenow sub-fragment lacking the 3'–5' exonuclease domain. Arch Biochem Biophys. 2015; 575:46-53.

19. Yang W, Lee JY, Nowotny M. Making and Breaking Nucleic Acids: Two-Mg^{2+}-Ion Catalysis and Substrate Specificity. Molecullar Cell. 2006; 22 (1): 5–13.

20. Derbyshire V, Grindley ND, Joyce CM. The 3'-5' exonuclease of DNA polymerase I of Escherichia coli: contribution of each amino acid at the active site to the reaction. EMBO Journal. 1991; 10 (1): 17-24.

21. Hamdan S, Carr PD, BROWN se, Ollis DL, Dixon NE. Structural Basis for proofreading during replication of the *Escherichia coli* chromosome. Cell. 2002; 10 (4), 535-546.

22. Majorek KA, Dunin-Horkawicz S, Steczkiewicz K, Muszewska A, Nowotny M, Ginalski K, Bujnicki JM. The RNase H-like superfamily: new members, comparative structural analysis and evolutionary classification. Nucleic Acids Res. 2014; 42 (7): 4160-4179.

PEM questions

1. Which domains of the DNA polymerase I constitute the Klenow fragment?
    a. Only the 5’ → 3’ polymerase domain.
    b. 5’ → 3’ polymerase and 3’ → 5’ exonuclease domains.
    c. 5’ → 3’ polymerase and 5’ → 3’ exonuclease domains.
    d. 3’ → 5’ exonuclease and 5’ → 3’ exonuclease domains.
    e. 5’ → 3’ polymerase, 3’ → 5’ exonuclease, and 5’ → 3’ exonuclease domains.
2. Which is the principal function of DNA polymerase I?
    a. Synthesis of DNA in the leading strand.
    b. Untargeted mutagenesis.
    c. Repair of pyrimidine dimers.
    d. Synthesis of RNA primers.
    e. Removal RNA primers and replacement the strand with DNA
3. Regarding the deoxynucleotides incorporation by DNA polymerase I. Which affirmation is true?
    a. The advancement of the polymerase ternary complex to the closed state (Ec) is energetically favorable.
    b. The ajar conformation appears due to the presence of a mismatch.
    c. A and B are correct.
    d. The catalytic process requires the presence of two atoms of iron.
    e. All of them are correct
4. Regarding the structure of DNA polymerase I, which affirmation is false?
    a. The 5’ → 3’ polymerase domain presents a hand shape.
    b. The 3’ → 5’ exonuclease domain is between the 5’ → 3’ exonuclease and the 5’ → 3’ polymerase domain.
    c. Amino acids involved in deoxynucleotide binding are in the thumb region of the 5’ → 3’ polymerase.
    d. It is similar to the structure of human DNA polymerase nu.
    e. The two catalytic sites of the Klenow fragment are separated by approximately 35 Å.

PEM questions

5. Which ion is important for the nucleophilic attack in DNA polymerization?
   a. Mn
   b. Zn
   c. A and B are correct
   d. Mg
   e. All of them are correct

6. How many exonuclease domains does DNA pol I have?
   a. Only 5'-3' exonuclease domain.
   b. Only 3'-5' exonuclease domain.
   c. A and B are correct.
   d. DNA pol I does not have any exonuclease domain.
   e. DNA pol I is in fact only an exonucleolytic enzyme and has 3 exonuclease domains.

7. Which is the specific function of DNA pol I 3'-5' exonuclease domain?
   a. It is a non functional domain in DNA polymerase I.
   b. It recognises primers and eliminates them.
   c. It participates in protein folding.
   d. It eliminates mismatched nucleotides from the primer strand, also known as proofreading.
   e. None of the above.

8. Regarding DNA and primer binding, which affirmations are correct?

   a. The thumb subdomain contacts the minor groove of the DNA
   b. Region 2 and region 6 interact with the template strand.
   c. A and B are correct.
   d. There are 4 conserved amino acids involved in the primer binding (Asp705, Asp882, Glu883 and Lys635)
   e. All of them are correct

PEM questions

9. Which elements are important on nucleophilic attack?
   a. Metal ions
   b. 3'OH of primer
   c. A and B are correct
   d. Incoming triphosphate
   e. All of them are correct

10. Which bond is formed after the nucleophilic attack of polymerase?
   a. Phosphodiester bond
   b. Salt bridges
   c. A and B are correct.
   d. Hydrogen bond
   e. All of them are correct

DNA polymerase I

Pedro Fortes, Sergi Garcia, Nuria Güil, Alba Ponce

Structural Biology - Course 2016-2017

Annex 1. Conformational changes during the binding of polymerase to DNA

Thumb subdomain:

1. It rotates towards the palm subdomain
2. The conserved amino acids residues located at the tip (helices H1 and H2) rotate in the opposite direction relative to the rest of the thumb such that the tip is in proximity to the DNA. →

DNA binding site