Structural and biochemical basis for the high fidelity and processivity of DNA polymerase ε

Rais Ahmad Ganai
Till min familj
# Table of Contents

ABSTRACT \hspace{1cm} ii  
LIST OF PAPERS \hspace{1cm} iii  
ABBREVIATIONS \hspace{1cm} iv  
INTRODUCTION \hspace{1cm} 1  
BACKGROUND \hspace{1cm} 1  
  DNA replication \hspace{1cm} 1  
  DNA polymerase \hspace{1cm} 4  
  Eukaryotic DNA polymerase \hspace{1cm} 5  
  Mechanism of DNA polymerization \hspace{1cm} 9  
  Nucleotide binding \hspace{1cm} 11  
  Stabilization of primer terminus in polymerase active site \hspace{1cm} 12  
  3′–5′ exonuclease activity and fidelity \hspace{1cm} 13  
  Partitioning of primer terminus between two active sites \hspace{1cm} 15  
  Structural insights of the catalytic core \hspace{1cm} 18  
  Role of Pol ε at the replication fork \hspace{1cm} 22  
  Strand displacement \hspace{1cm} 25  
AIM OF THE THESIS \hspace{1cm} 26  
RESULTS AND DISCUSSION \hspace{1cm} 27  
FUTURE PERSPECTIVES \hspace{1cm} 34  
ACKNOWLEDGEMENTS \hspace{1cm} 36  
REFERENCES \hspace{1cm} 38
Abstract

DNA polymerase epsilon (Pol ε) is a multi-subunit B-family DNA polymerase that is involved in leading strand DNA replication in eukaryotes. DNA Pol ε in yeast consists of four subunits, Pol2, Dpb2, Dpb3, and Dpb4. Pol2 is the catalytic subunit and Dpb2, Dpb3, and Dpb4 are the accessory subunits. Pol2 can be further divided into an N-terminal catalytic core (Pol2core) containing both the polymerase and exonuclease active sites and a C-terminus domain. We determined the X-ray crystal structure of Pol2core at 2.2 Å bound to DNA and with an incoming dATP. Pol ε has typical fingers, palm, thumb, exonuclease, and N-terminal domains in common with all other B-family DNA polymerases. However, we also identified a seemingly novel domain we named the P-domain that only appears to be present in Pol ε. This domain partially encircles the nascent duplex DNA as it leaves the active site and contributes to the high intrinsic processivity of Pol ε.

To ask if the crystal structure of Pol2core can serve as a model for catalysis by Pol ε, we investigated how the C-terminus of Pol2 and the accessory subunits of Pol ε influence the enzymatic mechanism by which Pol ε builds new DNA efficiently and with high fidelity. Pre-steady state kinetics revealed that the exonuclease and polymerization rates were comparable between Pol2core and Pol ε. However, a global fit of the data over five nucleotide-incorporation events revealed that Pol ε is slightly more processive than Pol2core. The largest differences were observed when measuring the time for loading the polymerase onto a 3’ primer-terminus and the subsequent incorporation of one nucleotide. We found that Pol ε needed less than a second to incorporate the first nucleotide, but it took several seconds for Pol2core to incorporate similar amounts of the first nucleotide.

B-family polymerases have evolved an extended β-hairpin loop that is important for switching the primer terminus between the polymerase and exonuclease active sites. The high-resolution structure of Pol2core revealed that Pol ε does not possess an extended β-hairpin loop. Here, we show that Pol ε can processively transfer a mismatched 3’ primer-terminus between the polymerase and exonuclease active sites despite the absence of a β-hairpin loop. Additionally we have characterized a series of amino acid substitutions in Pol ε that lead to altered partitioning of the 3’ primer-terminus between the two active sites.

In a final set of experiments, we investigated the ability of Pol ε to displace the downstream double-stranded DNA while carrying out DNA synthesis. Pol ε displaced only one base pair when encountering double-stranded DNA after filling a gap or a nick. However, exonuclease deficient Pol ε carries out robust strand displacement synthesis and can reach the end of the templates tested here. Similarly, an abasic site or a ribonucleotide on the 5’-end of the downstream primer was efficiently displaced but still only by one nucleotide. However, a flap on the 5’-end of the blocking primer resembling a D-loop inhibited Pol ε before it could reach the double-stranded junction. Our results are in agreement with the possible involvement of Pol ε in short-patch base excision repair and ribonucleotide excision repair but not in D-loop extension or long-patch base excision repair.
List of papers included in the thesis


II. Yeast DNA polymerase ε catalytic core and holoenzyme have comparable catalytic rates. Rais A. Ganai, Pia Osterman and Erik Johansson (JBC)

III. Switching between polymerase and exonuclease sites in DNA polymerase ε. Rais A. Ganai, Göran Bylund and Erik Johansson (Nucleic Acids Research)

IV. Modulation of strand displacement synthesis of DNA polymerase ε by processive 3’- 5’ exonuclease activity. Rais A. Ganai and Erik Johansson (Manuscript)

List of papers not included in the thesis


## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNA Pol</td>
<td>DNA polymerase</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxyribonucleotides</td>
</tr>
<tr>
<td>Pol site</td>
<td>Polymerase site</td>
</tr>
<tr>
<td>Exo site</td>
<td>Exonuclease site</td>
</tr>
<tr>
<td>Pol α</td>
<td>DNA polymerase alpha</td>
</tr>
<tr>
<td>Pol ε</td>
<td>DNA polymerase epsilon</td>
</tr>
<tr>
<td>Pol δ</td>
<td>DNA polymerase delta</td>
</tr>
<tr>
<td>E • DNA</td>
<td>Enzyme • DNA</td>
</tr>
<tr>
<td>φ29</td>
<td>Phi 29 DNA polymerase</td>
</tr>
<tr>
<td>Fe-S</td>
<td>Iron-sulfur cluster</td>
</tr>
<tr>
<td>NER</td>
<td>Nucleotide excision repair</td>
</tr>
<tr>
<td>BER</td>
<td>Base excision repair</td>
</tr>
<tr>
<td>PCNA</td>
<td>Proliferating cell nuclear antigen</td>
</tr>
<tr>
<td>RFC</td>
<td>Replication factor C</td>
</tr>
<tr>
<td>RPA</td>
<td>Replication protein A</td>
</tr>
<tr>
<td>Mcm2-7</td>
<td>Minichromosome maintenance 2-7</td>
</tr>
<tr>
<td>CDK</td>
<td>Cyclin dependent kinase</td>
</tr>
<tr>
<td>DDK</td>
<td>Dbf4-dependent Cdc7 kinase</td>
</tr>
</tbody>
</table>
Introduction

DNA encodes all of the information necessary for the growth and survival of an organism, and it exists as a double helix of two antiparallel strands. Before cell division, the duplication of each strand is carried out by special enzymes called DNA polymerases. Three major enzymes are involved in duplication of the genome: Pol α, Pol δ, and Pol ε. DNA polymerase ε (Pol ε) plays a central role in the duplication of DNA by performing synthesis on the leading strand. It consists of a catalytic subunit (Pol2) and the three non-catalytic subunits Dpb2, Dpb3, and Dpb4. The N-terminus of Pol2 contains both a 5'–3' polymerase activity and a 3'–5' proofreading activity. Pol ε is a high-fidelity polymerase due to efficient selectivity in the polymerase active site combined with proofreading in the exonuclease active site. The C-terminus of Pol2 is important for interactions with the accessory subunits. DNA Pol ε is a highly processive enzyme and synthesizes long stretches of DNA without the requirement of a PCNA clamp. The mechanism by which Pol ε synthesizes DNA is poorly understood, however, due to a lack of structural information. This thesis will focus on the determination of the high-resolution structure of Pol ε and the biochemical and kinetic properties of Pol ε. The work provides insights into how Pol ε synthesizes DNA with high processivity and high fidelity.

Background

DNA replication

DNA replication involves the accurate duplication of DNA into two identical copies. For the survival of an organism, DNA replication has to be carried out very accurately and efficiently. The mechanism by which DNA is duplicated into two strands is conserved in all three domains of life, archaea, prokaryotes, and eukaryotes (2-5). During DNA replication, one strand of DNA acts as a template against which DNA building blocks are added in a specified direction. Therefore, when a cell divides into two daughter cells, each daughter cell receives one old and one newly synthesized strand. This mode of replication is called semi-conservative replication. The mechanism of initiation, elongation, and termination of DNA replication in *Saccharomyces cerevisiae* is briefly described below.
Initiation and elongation of DNA replication

The process of DNA replication begins at special sites on the DNA called origins of replication. In budding yeast, these are referred to as autonomously replicating sequences (6). These are 100–200 base pair regions in the DNA characterized by an 11 base pair AT-rich consensus sequence that is essential for replication to occur. In eukaryotes, DNA replication starts at the same time on multiple origins to facilitate rapid replication of the DNA. During mitosis, the origin recognition complex (ORC) and Cdc6 bind to origins to begin replication. The binding of ORC and Cdc6 promote the recruitment of Cdt1, which in turn loads the Mcm2-7 helicase complex onto the DNA through hydrolysis of ATP (7,8). The sequential binding of these factors to DNA gives rise to the formation of a pre-replicative complex (preRC). The binding of Mcm2-7 licenses the origins for replication, and this is followed by the binding of several other proteins that transform the preRC into a pre-initiation complex (preIC) and lead to the activation of the Mcm2-7 helicase. All of this takes place when cells transit from G1 to S phase or when quiescent cells transit from G0 to the G1 phase of the cell cycle, and several different mechanisms ensure that replication occurs only once during the cell cycle. Two kinases, CDK and DDK, play very important roles in the formation of the preIC. The first step in preIC complex formation requires the binding of Mcm10, which leads to the dissociation of Cdc6 and Cdt1. This is followed by phosphorylation of several subunits of Mcm2-7 by DDK, in particular the phosphorylation of Mcm4, which activates the Mcm2-7 helicase for interaction with GINS and Cdc45 (9). At the same time, CDK phosphorylates Sld2 and Sld3 and activates them for binding to Dpb11 (10,11). Dpb11 contains two N-terminal BRCT domains and two C-terminal BRCT domains. The N-terminal BRCT domains bind to phosphorylated Sld3, and the C-terminal BRCT domains bind to phosphorylated Sld2 (12). Cdc45 and Sld3 form a complex and are recruited to origins of replication in a mutually dependent manner (13). Sld2 can form a complex with Pol ε and GINS, and this is called the preloading complex (12). The function of Dpb11 is to recruit both Sld3-Cdc45 and Sld2-GINS-Pol ε complexes to the origins of replication. Upon recruitment of Pol α to the replication origin, which completes the preIC complex, the replication fork becomes activated. Activation of the replication fork leads to the dissociation of Sld2, Sld3, and Dpb11 and the assembly of the replicative helicase – a complex of Cdc45, Mcm2-7, and GINS (CMG). The CMG complex unwinds the DNA and allows Pol α to prime both the leading and lagging strands. Finally, Pol δ is loaded onto the replication fork for DNA synthesis on the lagging strand and Pol ε on the leading strand. In eukaryotes, replication is bidirectional; therefore, two replication forks are assembled at each origin that move in opposite directions. In addition, because
polymerization moves only in the 5′–3′ direction due to the antiparallel nature of the two DNA strands, synthesis on one of the DNA strands is carried out in short stretches called Okazaki fragments (Figure 1). Pol δ is dependent on the PCNA clamp for processive DNA synthesis, whereas Pol ε interacts with GINS, which might enhance the processivity of Pol ε (14-16). This complex of several different proteins assembled on a replication fork was referred to as a *replisome progression complex* by Karim Labib et al. (17).

![Figure 1. Schematic representation of a eukaryotic replication fork.](image)

**Termination of DNA replication**

The CMG complex remains associated with the replication fork throughout replication (18). However, what happens to the CMG complex when two replication forks run into each other at the end of replication was not known until recently (19). SCF^{Dia2} E3-ubiquitin ligase mediates ubiquitinylation of the Mcm7 subunit of the CMG complex. F-box protein Dia2 acts as substrate targeting component of SCF^{Dia2} and is essential for SCF^{Dia2}–mediated ubiquitinylation. Ubiquitinylation of Mcm7 occurs only for a very short period of time at the end of S phase and only when two replication forks collide with each other. Ubiquitinylation of Mcm7 licenses Cdc48 segregase binding to the CMG complex. Cdc48 is an AAA+ ATPase, and binding of ATP to Cdc48 induces a conformational change in Cdc48 that triggers the disassembly of the CMG complex. Disassembly of the CMG complex from DNA starts by dissociation of GINS and Cdc45 followed by disintegration of the Mcm2-7 complex. However, how Mcm7 is activated for ubiquitination by SCF^{Dia2} remains unknown. In experiments where SCF^{Dia2} was limited, the CMG complex remained associated with chromatin even after S phase (19). The
topological constraints after the dissociation of the CMG complex at two converging forks are resolved by topoisomerase II (20,21). A similar mechanism of fork disassembly was observed in the higher eukaryote *Xenopus levis* indicating a universal mechanism of disassembling the CMG complex (22).

**DNA polymerases**

DNA polymerases are key enzymes involved in the replication of the genome and act as molecular motors that synthesize DNA in a template-directed fashion. *Escherichia coli* has five different DNA polymerases, yeast has eight, and mammals have seventeen different DNA polymerases. These enzymes are broadly classified into two categories, replicative polymerases that are involved in DNA replication and non-replicative polymerases that are involved in DNA repair. These replicative and non-replicative polymerase can be further classified into seven different DNA polymerase families (Table 1) (23) based on the homology of the primary amino acid sequence (24,25). Because this thesis focuses on Pol ε, the discussion below presents the biochemical properties and functional roles played by eukaryotic replicative polymerases.
Table 1. Classification of DNA polymerases into different families (modified from (26)).

<table>
<thead>
<tr>
<th>Family</th>
<th>Prokaryotic</th>
<th>Eukaryotic</th>
<th>Archaea</th>
<th>Viral</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Pol I</td>
<td>Pol γ, θ, v</td>
<td></td>
<td>T3, T7, T5</td>
</tr>
<tr>
<td>B</td>
<td>Pol II</td>
<td>Pol α, δ, ε, ζ</td>
<td>Pol B</td>
<td>RB69, T4, T6</td>
</tr>
<tr>
<td>C</td>
<td>Pol III</td>
<td>Pol D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td>Pol D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>Pol β, λ, μ, TdT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>Pol IV, V</td>
<td>Pol ι, κ, η, Rev1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT</td>
<td></td>
<td>Telomerase</td>
<td>Reverse transcriptase</td>
<td></td>
</tr>
</tbody>
</table>

Eukaryotic replicative DNA polymerases

At the core of the eukaryotic replication fork there are three different DNA polymerases present, Pol α, Pol δ, and Pol ε. Under normal physiological conditions, these three polymerases are sufficient for the duplication of the yeast genome. Pol α is responsible for initiating the de novo synthesis of primers, Pol δ replicates the lagging strand, and Pol ε replicates the leading strand (27-29). A specific amino acid substitution in the polymerase active site of Pol δ (L612M) and Pol ε (M644G) allowed for the strand specificity of these polymerases to be determined (28,29). These amino acid substitutions were chosen because they were close to the nucleotide binding pocket, the polymerase activity was not drastically affected, and the mutants displayed reduced fidelity and gave specific mutation spectra. The Pol ε M644G mutant gave 39-fold higher error rates for dTTP-dTTP than dATP-dATP mismatches. Similarly, Pol δ L612M was shown to have a 28-fold higher error rate for dTTP-dGTP than dATP-dCTP mismatches. To identify the polymerase acting on each strand, the URA3 gene was cloned on both sides of the replication
origin ARS306 in yeast strains carrying the amino acid substitutions in the respective polymerases. The patterns of the mutation spectra obtained by sequencing both strands of the URA3 gene demonstrated that Pol ε is the leading strand polymerase and Pol δ is the lagging strand polymerase. In addition, earlier biochemical studies demonstrated that Pol δ efficiently collaborates with PCNA, Fen1, and Lig1 during Okazaki fragment maturation, which supported the role of Pol δ in lagging strand replication (30). These proteins allow polymerase δ to perform limited strand displacement synthesis in order to remove downstream RNA primers at each Okazaki fragment. Very recent biochemical studies with the reconstituted replication fork showed that Pol ε was selected for the leading strand by its interaction with the CMG complex and that PCNA preferred Pol δ over Pol ε on the lagging strand (15). These results validated the previous genetic data that Pol ε is the leading strand polymerase and Pol δ is the lagging strand polymerase.

**DNA polymerase α**

In *S. cerevisiae*, Pol α is composed of four different subunits and it can be divided into two functional assemblies. Pri1 and Pri2 provide the DNA-dependent RNA primase activity, and Pol1 and Pol12 make up the DNA polymerase component (Table 2). Pri1 and Pol1 are catalytic subunits and Pri2 and Pol12 are accessory subunits (31). Pol α possesses 5′–3′ polymerase activity but lacks a 3′–5′ proofreading function. The possible function of the accessory subunits is to link Pol α with other proteins at the replication fork. The structure of the four-subunit holoenzyme resembles a dumbbell-shaped structure with two distinct parts, the polymerase and the primase, that are approximately 100 Å apart from each other (32).

The initiation of DNA replication begins with *de novo* synthesis of a short RNA primer of 10 nucleotides by DNA-dependent RNA polymerase Pri1 (33). The initial synthesis of the RNA primer leads to a switch from the primase to the polymerase Pol1, which extends the RNA primer by twenty more dNTPs before being displaced by the replicative polymerases (34). The functional interaction between the DNA polymerase and RNA polymerase components of Pol α is highly regulated to ensure efficient switching. Pol α can processively switch from RNA synthesis to DNA synthesis in the presence of dNTPs (35), and this switch is facilitated by the flexible nature of the two domains. Pol α is a non-processive enzyme, and this is consistent with its role in synthesis of short primers. The polymerase activity by Pol α is essential for chromosome duplication. Mutational analysis of individual subunits of *S. cerevisiae* Pol α reveals that they are all important for cell viability and cell cycle progression (36-39). Temperature sensitive mutations in the Pol1 protein exhibit high
genomic instability due to problems in DNA replication, and deletion of either PRI1 or PRI2 leads to unviable phenotypes (40,41).

**Polymerase δ**

Pol δ is a multi-subunit protein that is involved in replication of the lagging strand. In *S. cerevisiae*, Pol δ consists of three subunits, Pol3 (125 kDa), Pol31 (50 kDa), and Pol32 (40 kDa) (Table2) (42). Pol δ in humans and *Schizosaccharomyces pombe* contain the additional accessory subunits P12 and Cdm1, respectively. Pol δ is indispensable for DNA replication and it proofreads the replication errors made by Pol α (43). Pol δ by itself is a non-processive enzyme, and it requires accessory factors such as PCNA (clamp) for processive DNA synthesis. PCNA is a heterotrimeric sliding clamp that is loaded onto DNA by the RFC (clamp loader) in an ATP-driven process. Together, PCNA, RFC, and RPA (ssDNA binding protein) increase the processivity of Pol δ from 6 nucleotides to more than 600 nucleotides (16).

Apart from being involved in replication of the lagging strand, Pol δ is also involved in homologous recombination, DNA repair, and Okazaki fragment maturation. Small-angle X-ray scattering analysis of *S. cerevisiae* Pol δ reveals that it consists of a globular catalytic core of Pol3 and a flexible tail domain composed of the non-catalytic subunits Pol31 and Pol32 (44).

**Pol3**

Pol3 is the catalytic subunit of Pol δ and contains both a 5’–3’ polymerase active site and a 3’–5’ exonuclease active site (42,45). Pol3 is indispensable for replication of the lagging strand. The C-terminus of Pol3 contains a [4Fe-4S] cluster that is essential for interaction with Pol31 (46), and Pol3 also contains a PiP box for interaction with PCNA.

**Pol31 and Pol32**

Pol31 and Pol32 are the accessory subunits of Pol δ. Pol32 is an elongated protein that contributes to the overall elongated structure of Pol δ (47). Pol32 contains a PCNA motif for interaction with PCNA (48), and PCNA also interacts with Pol δ through other sites present on the different subunits of Pol δ (49). Pol32 is involved in mutagenic bypass pathways (50). Pol31 is an essential protein and acts a bridge between Pol3 and Pol32. It was recently shown that Pol31 and Pol32 form a complex with Rev3 and Rev4 to create the four-subunit Pol ζ (51).
DNA polymerase ε

Pol ε is a multi-subunit DNA polymerase involved in leading strand replication and is indispensable for chromosome duplication (29). Pol ε consists of four subunits, Pol2, Dpb2, Dpb3, and Dpb4 (Table 2). Pol ε is intrinsically a highly processive enzyme and is only stimulated about 6-fold by the presence of the accessory proteins PCNA, RFC, and RPA (16). Biacore experiments revealed that it has high affinity for DNA and low affinity for PCNA, which is the opposite of Pol δ. The authors of the Biacore study suggest that Pol ε and Pol δ are loaded onto DNA via different mechanisms (16). The cryo-EM structure of Pol ε was solved at 20 Å resolution, and in this structure Pol2 forms a globular head domain and the non-catalytic subunits form a flexible tail domain (52).

Pol2

Pol2 is the catalytic subunit of Pol ε. Pol2 can be divided into an N-terminal domain called Pol2core that contains both polymerase and exonuclease active sites and a C-terminal domain that is important for the interaction with the non-catalytic subunits. Pol2 is an intrinsically processive enzyme that functions in various cell cycle regulatory pathways such as cell cycle checkpoint, DNA repair, and recombination. Deletion of Pol2 in yeast leads to an unviable phenotype. Interestingly, deletion of only the N-terminal catalytic domain of Pol2 results in severe growth defects but point mutations in catalytic carboxylate residues of the polymerase active site leads to an unviable phenotype (53,54). Deletion of the C-terminus of Pol2 also leads to cell death, perhaps due to a defect in the initiation of replication (55). The N-terminus of Pol2 contains an [4Fe-4S] cluster that, when inactivated, abolishes the polymerase activity but not the exonuclease activity (56).

Dpb2

Dpb2 is an essential but non-catalytic subunit of Pol ε with several functions in the cell (57). The N-terminus of Dpb2 is important for interactions with the Psf1 subunit of GINS, which links Pol ε to the replicative helicase during DNA replication (14) and the C-terminus of Dpb2 has been shown to interact with Pol2 (58). Dpb2 is phosphorylated by the Cdc28 kinase. This phosphorylation is important for normal cell cycle progression, but substituting a non-phosphorylatable amino acid in Dpb2 is not lethal (59).
Dpb3 and Dpb4

Dpb3 and Dpb4 are non-catalytic and non-essential subunits in *S. cerevisiae*, but Dpb3 in *S. pombe* is essential for survival of the cell. Deletion of *S. cerevisiae* Dpb4 leads to slow progression of the S phase, and deletion of Dpb3 leads to an increase in spontaneous mutation rates (60,61). Both of these proteins interact with the C-terminus of Pol2 and confer increased processivity to Pol ε by enhancing the binding of Pol ε to double-stranded DNA (62). Dpb4 also associates with Dls1, and the function of this complex is to recruit the chromatin remodeling complex ISW2 to DNA (63).

<table>
<thead>
<tr>
<th>Polymerase subunit</th>
<th>Primase subunits</th>
<th>Accessory subunits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pol α</td>
<td>Pol1 Pri1</td>
<td>Pol12 and Pri2</td>
</tr>
<tr>
<td>Pol δ</td>
<td>Pol3</td>
<td>Pol31 and Pol32</td>
</tr>
<tr>
<td>Pol ε</td>
<td>Pol2</td>
<td>Dpb2, Dpb3, and Dpb4</td>
</tr>
</tbody>
</table>

Mechanism of DNA polymerization

All DNA polymerases share the same basic mechanism of nucleotide incorporation despite differences in amino acid sequence at the active site, and the minimal kinetic mechanism has been established for all polymerase families (Figure 2) (23). The first step of catalysis is the binding of DNA polymerase to DNA to form a binary complex. The structural basis of DNA binding is inferred from crystal structures of Klentaq 1 with and without DNA (64). These structures reveal that the initial binding of DNA polymerase to DNA leads to large structural changes in the tip of the thumb domain. These changes are followed by movement of the thumb in the opposite direction that allows the thumb to encircle the DNA. The next step in catalysis involves the
binding of a dNTP and two metal ions (usually Mg$^{2+}$) in the active site of the polymerase. The polymerase then changes to a closed conformation to position the incoming dNTP, the Mg$^{2+}$ ions, and the 3'-OH of the nascent DNA strand in a perfect geometric arrangement for bond formation. From a structural perspective, the change to a closed conformation primarily involves the rotation of the fingers domain inward towards the palm domain. The 3'-OH of the primer terminus engages in a nucleophilic attack on the α-phosphate of the incoming nucleotide resulting in phosphodiester bond formation. Catalysis is not possible without the involvement of two divalent metal ions (usually Mg$^{2+}$), which form transient bonds with the β- and γ-phosphates of the incoming nucleotide and the conserved catalytic carboxylates of the palm domain. The metal ion in the A-site forms a transient bond with the α-phosphate of the incoming nucleotide and the oxygen of the 3'-OH. The metal ion in B-site forms a bond with the β- and γ-phosphates of the incoming nucleotide. The function of the metal ions is to stabilize the net charge during catalysis. Phosphodiester bond formation leads to a second conformational change in the polymerase and the release of inorganic pyrophosphate. The polymerase then either dissociates from the DNA or translocates along the DNA to perform another round of DNA synthesis. There is significant debate about whether the rate-limiting step during polymerization is a conformational change or if it is a chemistry step. However, recent data suggest that structural arrangements at the active site of a polymerase are slower than the movement of the fingers closing in on the active site (65-67).
Figure 2. Kinetic mechanism of DNA polymerization. E indicates the enzyme, *(asterisk) denotes the conformational change and PPi indicates pyrophosphate.

Nucleotide binding

Binding of a nucleotide at the active site involves interactions between the nucleotide and the template base, the active site residues, the fingers domain, the metal ions, and the terminal nucleotide of the primer strand. These interactions not only allow the polymerase to align the dNTP against the template strand for proper Watson-Crick interactions, but they also allow the polymerase to differentiate between the correct and incorrect nucleotide thus imparting high selectivity to DNA polymerases. As soon as the matched nucleotide is selected in the nucleotide binding pocket, it stimulates the movement of the fingers toward the palm domain for bond formation. However, if an incorrect nucleotide is bound in the nucleotide binding pocket, the lack of proper geometry of the incoming nucleotide at the active site will inhibit closing of the fingers. The replicative polymerases differentiate between correct and incorrect nucleotides with high efficiency. T7, T4, and HIV Reverse Transcriptase polymerases bind correct nucleotides 250-fold to 350-fold more efficiently than incorrect nucleotides (68-71). The difference in binding a correct vs. an incorrect nucleotide in repair polymerases, however, is much lower. Klenow fragment shows only a 3.4-fold selectivity, yeast Pol η a 4-fold selectivity, and Pol β a 20-fold selectivity (72-76). The low
discrimination against incorrect nucleotides by repair polymerases corresponds to their function in error-prone synthesis.

**Stabilization of the primer terminus in the polymerase active site**

The active sites of DNA polymerases have evolved a configuration that allows them to make non-sequence-specific contacts with the DNA and the incoming nucleotide. DNA in the polymerase active site is stabilized by the interactions of the polymerase with the minor groove of the DNA. The thumb domain plays a major role in stabilizing the DNA in the active site of the polymerase, and the amino acid side chains of the tip of the thumb domain make hydrogen bonds with the primer strand in the polymerase active site. Movement of the thumb domain has been observed when DNA switches between the polymerase and exonuclease active sites (77,78). Pol δ stabilizes DNA in the polymerase active site by extensive interactions between the thumb domain and the newly built DNA as far as seven base pairs into the newly synthesized double-stranded DNA (79).

The KKRY motif in the thumb domain of B-family polymerases has been shown to be important for stabilizing DNA in the polymerase active site (80). K498 in φ29 DNA Pol is homologous to K967, K814, and K706 of Pol ε, Pol δ, and RB69 gp43, respectively. When K498 in φ29 DNA Pol was replaced with arginine (K498R) or tyrosine (K498Y), the mutant exhibited higher exonuclease activity compared to wild-type φ29 DNA Pol and required high dNTP concentrations to synthesize DNA (80). The strong preference for exonuclease activity over polymerase activity in these mutants reflects their inability to stabilize the primer in the polymerase active site. Similarly, the K706A amino acid substitution in RB69 gp43 displayed a large decrease in polymerase activity due to partitioning of the primer into the exonuclease active site (81). Thus, the KKRY domain appears to contribute to the stabilization of the primer-template in the polymerase active site.

Amino acid K555 in the thumb domain of φ29 DNA Pol is homologous to R988 of Pol ε, R839 of Pol δ, and K734 of RB69 gp43. The K555A amino acid substitution in φ29 DNA Pol resulted in a polymerase with reduced exonuclease activity than wild type and a tendency of the polymerase to retain the primer in the polymerase site (82). These results were unexpected because K555 interacts with the primer in the polymerase site. The authors of that study postulated that K555 was either playing a role in transferring the primer
between the two active sites or it was somehow stabilizing the primer in the exonuclease site without affecting single-stranded DNA binding.

The A-, B- and Y-family polymerases contain three conserved motifs called Motif A, B, and C that form part of the polymerase active site (Table 3). These motifs contain the conserved aspartates essential for catalysis, and these motifs are also important for stabilizing the DNA in the active site. In Motif A, Pol ε contains a methionine at position 644 whereas other members of the B-family polymerases contain leucine at this position. Different amino acid substitutions in this motif have been shown to increase or decrease stabilization of the primer terminus in the polymerase active site. Amino acid substitutions L412M in T4, L868F and L868M in Pol α, L612M in Pol δ, L415G and L415F in RB69 gp43, and M644G in Pol ε resulted in polymerases with increased ability to stabilize a mismatch in the polymerase site. The M644L substitution in Pol ε resulted in a polymerase with decreased ability to extend a mismatch (83).

<table>
<thead>
<tr>
<th>Motif A</th>
<th>Motif B</th>
<th>Motif C</th>
</tr>
</thead>
<tbody>
<tr>
<td>yPol2</td>
<td>640-DVASHPNIM</td>
<td>824-VILNSFYGYVM</td>
</tr>
<tr>
<td>hp261</td>
<td>626-DVGAMYPNII</td>
<td>809-KCILNSFYGYVM</td>
</tr>
<tr>
<td>yPol3</td>
<td>688-DFNSLYPSIM</td>
<td>781-KISANSVYGFTE</td>
</tr>
<tr>
<td>RB69 gp43</td>
<td>411-DLTSLYPSIS</td>
<td>568-KLINSLYGALG</td>
</tr>
<tr>
<td>Phi29 Pol</td>
<td>249-DVNSLYPAQM</td>
<td>383-KLMLNSLYGKFA</td>
</tr>
<tr>
<td>yPol1</td>
<td>864-DFNSLYPSI</td>
<td>944-KLTANSMYGCLG</td>
</tr>
</tbody>
</table>

Table 3. Sequence alignment of conserved polymerase motifs in B-family DNA polymerases. The highlighted D amino acids indicate the catalytic residues and highlighted M indicates the M644 amino acid. Y indicates the yeast whereas h denotes human.

3’–5’ exonuclease activity and fidelity

Some DNA polymerases have evolved an exonuclease active site to increase the overall incorporation efficiency by proofreading the misincorporated nucleotides. DNA sequence analysis of B-family DNA polymerases reveals three conserved motifs – Exo I, Exo II, and Exo III – at the N-terminal end of the protein (Table 4). Together, these motifs form the exonuclease active site and contribute catalytic aspartates and carboxylates that bind two Mg²⁺ ions together with the attacking water molecules to catalyze the cleavage of the terminal phosphodiester bond and the release of dNMP (78,84). Substitution of alanine for any of the catalytic aspartates in the exonuclease site reduces or abolishes exonuclease activity and increases the spontaneous mutation
frequency by several hundred-fold (45). Yeast replicative polymerases incorporate a wrong nucleotide once every $10^4$ to $10^5$ nucleotides replicated. This means that yeast replicative polymerases misincorporate one or two nucleotides every cell division. Most of the misincorporated nucleotides are excised by the associated 3′–5′ exonuclease activity, which acts as a first line of defense against misincorporations. Efficient proofreading by DNA polymerases is dependent on several factors, including the nature of the misincorporated nucleotide, the sequence context of the misincorporation, and the concentration of dNTPs at the replication fork. If the misincorporated nucleotides escape the proofreading, they are generally corrected by the mismatch repair system.

Yeast Pol α is a low-fidelity polymerase due to a lack of exonuclease activity. Based on amino acid sequence alignment with other polymerases, the catalytic carboxylate residues in the exonuclease domain Pol α have been replaced by non-catalytic residues. The majority of errors made by Pol α are single-base substitution errors, and Pol α makes one substitution error for every 9,900 nucleotide incorporations. The second most common errors observed are single-base deletions with one deletion error for every 12,000 nucleotides replicated (85-87). Despite the lack of exonuclease activity, Pol α has a low error rate compared to translesion polymerases due to its low propensity to extend mismatches. RPA (single-stranded DNA binding protein) modulates the fidelity of Pol α by decreasing its affinity for mismatched primer termini and by increasing its affinity for matched primer-termini by acting as a fidelity clamp (88,89). Mistakes made by Pol α are corrected by the proofreading activity of Pol δ (43).

Yeast Pol δ is a high-fidelity polymerase, and kinetic analysis of Pol δ reveals that it has a strong preference for incorporation of correct nucleotides over incorrect nucleotides (90). *S. cerevisiae* Pol δ exhibits a single-base substitution error rate of one in every 100,000 nucleotides replicated (91) and a single-base deletion error rate of one in every 100,000 bases replicated. Inactivation of the exonuclease activity of Pol δ leads to a 100-fold increase in spontaneous mutation rates in vivo (45), but only a 10-fold increase in mutation rates in vitro (91).

Yeast Pol ε is also a high-fidelity polymerase, and it is the most accurate of all the B-family DNA polymerases (92). *S. cerevisiae* Pol ε exhibits a low error rate of one single-base substitution for every 200,000 nucleotides replicated and one single-base deletion for every 2,000,000 nucleotides replicated (93). Exonuclease-deficient Pol ε exhibits 10-fold and 100-fold higher error rates for single-base substitutions and single-base deletions, respectively, in vitro (93).
Alleles that inactivate the exonuclease activity of Pol ε or Pol δ in mice result in a strong mutator phenotype that leads to increased incidence of cancers (94,95). The proofreading activities of Pol ε and Pol δ act as tumor suppressors in mice, and this emphasizes the importance of the exonuclease activity for genomic integrity.

<table>
<thead>
<tr>
<th></th>
<th>Exo I</th>
<th>Exo II</th>
<th>Exo III</th>
</tr>
</thead>
<tbody>
<tr>
<td>yPol2</td>
<td>289-FDIET</td>
<td>377-FNGDFFDwPFIHN</td>
<td>473-YSVSD</td>
</tr>
<tr>
<td>hp261</td>
<td>274-FDIET</td>
<td>362-YNGDFDFwPFVEAR</td>
<td>458-YSVSD</td>
</tr>
<tr>
<td>yPol13</td>
<td>320-FDIEC</td>
<td>401-YNTTNFDIPYLLNR</td>
<td>516-YCLKD</td>
</tr>
<tr>
<td>RB69 Pol</td>
<td>113-FDIEV</td>
<td>216-WNVESFDIPYVYNN</td>
<td>323-YNIID</td>
</tr>
<tr>
<td>Phi29 Pol</td>
<td>011-CDFET</td>
<td>61-HNLKFDGAfIIINWLN</td>
<td>165-YIKND</td>
</tr>
<tr>
<td>Pol1</td>
<td>520-DKPOQ</td>
<td>638-HRLQNVYLDVLHAR</td>
<td>810-FIVPD</td>
</tr>
</tbody>
</table>

**Table 4.** Sequence alignment of the exonuclease motifs in B-family polymerases. The highlighted amino acids represent the catalytic residues.

**Partitioning of the primer terminus between the two active sites**

During DNA polymerization, DNA synthesis is favored over switching of the primer into the exonuclease active site so as to prevent unnecessary degradation of correctly base-paired DNA (96). As described previously, replicative DNA polymerases incorporate incorrect nucleotides very rarely. However, if an incorrect nucleotide is incorporated, the polymerase active site is unable to close the fingers domain when trying to extend the misincorporated nucleotide. This causes the polymerase to stall, and as a result the DNA polymerase either switches the primer to the exonuclease site or dissociates from the primer-template to rebind in the exonuclease site. These are referred to intramolecular and intermolecular proofreading pathways, respectively. The decision to proofread is dependent on many factors, including dNTP concentrations, the DNA sequence, and the DNA polymerase involved (96-98). From a mechanistic standpoint, in the A-family *Bacillus stearothermophilus* BF polymerase (Bacillus fragment), the mismatch induces a distortion in the template strand at the pre-insertion site, and this affects the binding of the incoming dNTP in the active site and base pairing with the template base resulting in the stalling of the DNA polymerase (99). In the case of C-family polymerases, for example, Pol C (PolIIIC), the presence of a mismatch in the polymerase active site weakens the interactions...
between the thumb and the phosphodiester backbone of the primer-template resulting in destabilization of the DNA in the polymerase active site (100). Intramolecular and intermolecular proofreading pathways are described below.

**Intramolecular proofreading pathway**

The intramolecular pathway involves the transfer of DNA between the exonuclease and polymerase site without dissociation of the polymerase from the DNA. The first step in proofreading is the ability to differentiate between a matched and a mismatched primer in the polymerase site. The altered geometry of a mismatch at the polymerase site prevents the further addition of nucleotides and allows the polymerase to separate the nascent DNA strand from the template strand. The template strand is held within the polymerase site while the primer strand is transferred to the exonuclease site – which is located 30–40 Å from the polymerase active site – for trimming (Figure 3). The formation of the exonuclease complex involving strand separation is the rate-limiting step of this pathway. Strand separation is an energetically unfavorable step, and it involves a specific movement of the thumb domain that allows the primer to move into the exonuclease site. This mechanism is supported by the observations that primer termini consisting of G+C base pairs are more difficult to separate than primer termini consisting of A+T base pairs (98). Deletion of the β-hairpin loop in DNA polymerase from RB69 does not affect the exonuclease activity on single-stranded DNA but affects the exonuclease activity on double-stranded DNA, thus revealing a defect in strand separation of duplex DNA (101). Studies on the structure of Klenow fragment have revealed that four terminal base pairs of the primer-template are melted to allow the transfer of the primer to the exonuclease active site (102).

The trimmed primer in the exonuclease site must be returned back to the polymerase site for the resumption of DNA synthesis. Precise DNA-enzyme contacts are needed to position the primer in correct alignment to avoid frameshift mutations. DNA polymerases, such as Klenow fragment, polymerase from T4, and polymerase from RB69, are able to excise a mismatch in the exonuclease site and transfer the primer to polymerase site intramolecularly (101,103,104). There are clear structural differences between how A- and B-family polymerases shift primers between the two active sites. The biggest difference is that the exonuclease active sites of A- and B-family polymerases are positioned on opposite sides of the polymerase active site (105). In addition A-family polymerases do not contain the β-hairpin loop that extends from the exonuclease sites of B-family polymerases and has been shown to be important for switching the primer between the two active sites.
Antimutator amino acid substitutions increase the likelihood of exonuclease complex formation and decrease the ability of the enzymes to form a polymerase complex with the DNA. This indicates that some assistance is perhaps needed to shift the trimmed primer end from the exonuclease to the polymerase site. In the polymerase from bacteriophage T4, the I417V amino acid substitution results in a polymerase with increased tendency to shift DNA from the polymerase site to the exonuclease site, and this leads to increased proofreading and decreased misincorporations (106). The A737V substitution in the polymerase from bacteriophage T4 has a more processive exonuclease activity than the wild-type polymerase (107). In the Klenow fragment of DNA Pol I, residues in helix J (Q677 in particular) have been shown to participate in switching the primer terminus between the two active sites. The Q677A substitution was shown to decrease polymerase activity with a concomitant increase in 3′–5′ exonuclease activity because the mutant was defective in switching the primer from the exonuclease site to the polymerase site (108). Several antimutators have also been identified in Pol ε, including V522A in the N-terminal domain, G435C in the exonuclease domain, T850M and K966Q in the palm domain, and A1153D in the thumb domain. These mutations rescue the lethal phenotype of Pol2 exo− in combination with the deletion of mismatch repair genes (pol2−4 msh2Δ). The biochemical basis of the antimutator effect has not yet been determined.

**Intermolecular proofreading pathway**

The intermolecular proofreading pathway involves several dissociation and reassociation events between the DNA polymerase and the DNA. Incorporation of an incorrect nucleotide at the polymerase active site can also lead to dissociation of the polymerase from the DNA. The DNA polymerase can then rebinding the mismatched DNA directly in the exonuclease active site where it excises the mismatch and then dissociates again. This is followed by re-association with the DNA with the edited primer-template in the polymerase site and the resumption of DNA synthesis (Figure 3). This pathway is supported by experiments showing that T4 polymerases exchange positions on PCNA during processive polymerization (109). Several molecules of T4 are loaded onto the DNA at the same time by PCNA, and when one polymerase dissociates after making a mismatch, other polymerases associated with PCNA can bind the mismatched primer-terminus in the exonuclease site and excise the mismatch. These results are further supported by the high mutation frequency observed in strains carrying low concentrations of Pol δ (110). Low concentrations of Pol δ reduce the chances of tethering multiple Pol δ molecules to PCNA, and this might allow for error-prone polymerases to extend mismatches and thus reduce the fidelity of replication.
Structural insights of the catalytic core

The crystal structures of the replicative DNA polymerases Pol α and Pol δ have recently been solved (79,111). Their structures resemble a cupped right hand encircling the DNA (Figure 3 and 4). Pol α and Pol δ share the same basic structural domains as found in other polymerases. The catalytic cores of all B-family polymerases are composed of the palm, the thumb, the fingers, the exonuclease, and the N-terminal domains, and these domains perform different functions. The polymerase active site is formed by the palm, the thumb, and the fingers domain. The function of this site is to position the primer terminus and the incoming nucleotide for catalysis. The fingers domain binds the incoming nucleotide, the palm domain contains the catalytic carboxylates, and the thumb domain stabilizes the DNA in the active site. The
3’–5’ exonuclease domain performs the proofreading function. One of the major differences in the structures of B-family and A-family polymerases is that the exonuclease domain is on opposite sides of the palm domain.

**Structure of Pol α**

The crystal structure of Pol α was recently solved by L. Pelligrini et al. (Figure 3) (111). In order to obtain diffracting crystals, the polymerase was truncated to remove the unstructured N-terminal end and the extreme C-terminal end containing an [4Fe-4S] cluster. Amino acid residues 39 to 1258 were used for crystallization studies (111). The Pol α structure was solved as the enzyme alone, as a binary complex with an RNA/DNA hybrid (RNA primer and DNA template), and as a ternary complex with an RNA/DNA hybrid and an incoming dNTP. Alignment of all three structures revealed significant structural changes in the fingers, thumb, and palm domains in response to dNTP binding, and the RNA/DNA took on an A-form conformation but not a B-form conformation in the active site. The polymerase contacts the RNA up to 7 base pairs upstream of the 3’-end of the primer. The palm domain contacts the first 3 base pairs of RNA/DNA double helix, and the thumb contacts the RNA strand from the second nucleotide to the sixth nucleotide. The thumb domain grips the RNA by interacting with the ribose and phosphates of the RNA strand. The crystal structures provide mechanistic insights into how Pol α disengages itself from continuous DNA synthesis. Biochemical evidence indicates that Pol α preferred RNA primer over a DNA primer for extension. It processively synthesizes DNA equivalent to one turn of the helix, and most of the extension products terminate after synthesis of about 12 nucleotides of DNA. Pol α appears to prefer an RNA/DNA hybrid in the active site because the polymerase interacts with the ribose and phosphate of the nascent RNA primer. The incorporation of dNTPs after laying down a stretch of RNA into the nascent strand impairs the interaction between the DNA and the amino acid side chains of Pol α, and this causes Pol α to dissociate from the DNA. This process appears to have evolved to allow for efficient switching of polymerases after the initial RNA primer is long enough for extension by replicative polymerases. The crystal structures of Pol α also revealed that it lacks the extended β-hairpin loop emanating from the exonuclease domain of most other B-family DNA polymerases. The β-hairpin loop is instead replaced by a helix that contacts the second and third nucleotides downstream of the
unpaired template strand and holds the template strand in the active site for efficient Watson-Crick base pairing with the incoming nucleotide.

**Structure of Pol δ**

The crystal structure of Pol3, the catalytic core of Pol δ, was solved as a ternary complex containing an incoming dCTP and Ca$^{2+}$ by Aggarwal et al. in 2009 (Figure 4) (79). The structure resembles a typical B-family polymerase containing palm, thumb, fingers, exonuclease, and N-terminal domains, and the polymerase and exonuclease domains are separated by a distance of approximately 45 Å. The DNA in the active site is in a B-form conformation. The structure provides clues as to how Pol3 catalyzes DNA polymerization with high fidelity. The first two base pairs of the nascent strand interact with the amino acid side chains by hydrogen bonding, and these contacts were postulated to be important for sensing the correctly incorporated nucleotides. Incorporation of an incorrect nucleotide into the growing strand will destabilize hydrogen bonding between the nascent strand and the amino acid side chains at the polymerase active site. Altered geometry at the polymerase active site will shift the primer towards the exonuclease site for excision. In addition, amino acid side chains of the thumb domain were found to contact the nascent DNA through hydrogen binding up to 5 base pairs upstream from the primer terminus. This led to the hypothesis that if a mismatch evades
recognition at the primer terminus junction the distorted non-Watson–Crick base paired DNA would be sensed as far as four or five nucleotides upstream and shifted to the exonuclease active site for editing. The crystal structure of Pol δ also revealed that it has a β-hairpin loop extending from the exonuclease domain towards the polymerase active site in close proximity to the primer terminus. The β-hairpin loop wedges between the primer and template strands. The β-hairpin loop in T4 and RB69 gp43 has been shown to be important for the transfer of DNA between the polymerase and exonuclease sites as discussed above.

Three metal ions (Ca^{2+}) were found in the active site of Pol δ, and two of the metal ions were in position for coordinating the nucleophilic attack by the 3'-OH of the primer onto the α-phosphate of the incoming nucleotide. The third metal ion was coordinated by the γ-phosphate of the incoming nucleotide and two glutamine residues in the active site. Substituting these glutamines for alanine’s resulted in a polymerase with decreased polymerization rates for both correct and incorrect nucleotides suggesting a role for the third metal ion in catalysis (79).

Figure 4. Structure of a ternary complex of Pol δ. The structure was modified from (1) (with permission from the publisher)
Functions of DNA polymerase ε during the cell cycle

In addition to the involvement of DNA polymerase ε in leading strand synthesis, a growing body of evidence supports a role for Pol ε in DNA damage responses and DNA repair, DNA recombination, and chromatin remodeling (112-115). However, the role of Pol ε remains unclear due to high levels of redundancy in these pathways. Several other polymerases have also been associated with the same processes. I will briefly discuss the role of polymerase ε in these processes.

Checkpoint activation

Checkpoint activation is a surveillance mechanism that restricts cell cycle progression in response to problems with cell division or the detection of DNA damage. The checkpoint in S phase is activated in response to stalling of the replication fork, which results in uncoupling of the CMG helicase and DNA Pol ε (116). This uncoupling leads to long stretches of single-stranded DNA remaining on the leading strand. Binding of RPA to these single-stranded stretches recruits the Mec1/ATR kinase that activates a cascade of events leading to the stabilization of stalled replication forks and checkpoint activation. Pol ε is an important component of the S phase checkpoint response. The role of the C-terminus of Pol2 in checkpoint activation emerged from a substitution at Q2195 by a stop codon. The mutant cells were unviable in the presence of hydroxyurea. The cells were unable to activate the checkpoint kinase Dun1 or the transcription of RNR3, the regulatory subunit of RNR that is only transcribed in response to depletion of dNTP pools (115). Thus Pol ε acts as a sensor of DNA damage and transduces signals to other members of the checkpoint cascade.

Mrc1 interacts with both the N-terminus and C-terminus of Pol2 and Mcm2 (117). In response to DNA damage in the presence of hydroxyurea, Mrc1 is phosphorylated and loses its interaction with the N-terminus of Pol2 but not with the C-terminus of Pol2 (117). The loss of Mrc1 interaction with the N-terminus of Pol2 could allow the N-terminus of Mrc1 to interact with CMG components such as Cdc45 or Mcm2-7 and stabilize the replication fork. This loss of interaction could also free the N-terminus of Mrc1 to perhaps recruit Rad53 and thus activate the checkpoint cascade, or the impaired interaction with the N-terminus of Pol2 could induce a conformational change in the Pol2 C-terminus thus activating the S phase checkpoint. Further work needs to be done to establish the exact role of Pol ε in checkpoint activation.
More recently, it was shown that the Dpb4 subunit of Pol ε associates with Dpb11 and Sld2 in response to replication fork stalling (118). Together these proteins activate the Mec1 kinase independent of the 9-1-1 clamp, and this leads to activation of Rad53 kinase.

**Excision repair**

DNA is constantly exposed to exogenous and endogenous agents that modify or distort the structure of the DNA. The damages can either lead to a loss of a base to create an abasic site or to a modification of a base such as the formation of 8-oxo-G or T-T dimers. These lesions act as barriers to a moving replication fork and pose a threat to genomic integrity. Therefore, to prevent stalling of a replication fork, cells have evolved multiple mechanisms to repair these damages, including base excision repair and nucleotide excision repair (119).

**Base excision repair (BER)**

DNA damage that does not distort the geometry of DNA is generally repaired by base excision repair (BER). The BER repair pathway requires the concerted effort of DNA glycosylases, AP endonucleases, polymerases, and a ligase. DNA glycosylases recognize and remove the damaged bases by breaking the glycosydic bond between the damaged base and ribose sugar creating an abasic site. AP endonuclease binds to this abasic site in DNA and makes a single-stranded nick on the 5’-end of the abasic site, and this is followed by filling of the gap by a DNA polymerase and sealing of the nick by ligase. BER is divided into two subpathways: short-patch BER and long-patch BER. Short-patch BER involves the displacement of a single nucleotide by the DNA polymerase to generate a short downstream flap. Long-patch repair involves strand displacement synthesis of several kilobases of DNA before ligation. In mammals, Pol β is the primary polymerase involved in BER (120). Work with purified mouse proteins showed that both Pol δ and Pol ε are involved in long-patch repair in a PCNA-dependent manner (121). In humans, both Pol ε and Pol δ are considered to act as backup polymerases for BER, and purified yeast Pol ε has been shown to complement the BER pathway in cell extracts containing temperature-sensitive Pol ε mutants (122). These findings are further supported by the observed interactions of Pol2 with BER intermediates in cell-free extracts of *S. cerevisiae* (123).

**Nucleotide excision repair (NER)**

DNA damage that structurally distorts the DNA and affect base pairing between the two strands is repaired by nucleotide excision repair (NER). The lesions repaired by NER are generally products of environmental agents such
as chemically induced bulky adducts and UV light-induced cyclopyrimidine dimers and 6-4 photoproducts. In mammals, a defect in NER leads to a disorder called xeroderma pigmentosum. During NER, DNA is incised on both sides of the DNA damage resulting in the removal of approximately 30 nucleotides from the DNA. In an earlier reconstitution experiment, it was shown that purified Pol ε is involved in NER independent of PCNA; however, in the presence of RPA, Pol ε was entirely dependent on PCNA for NER (124). NER in human cells was later shown to be carried out by both Pol δ and Pol ε (125). The recruitment of the polymerase was dependent on PCNA being present at the DNA damage site. The alternative Ctf18-RFC clamp recruits Pol ε, whereas the standard RFC-PCNA clamp recruits Pol δ, and the repair is performed in concert with ubiquitylated PCNA and Pol κ (Polymerase kappa-a repair polymerase). Recent work with cell-free yeast extracts suggested that both Pol δ and Pol ε, but not Pol η, are involved in NER (126). Yeast extracts containing temperature-sensitive mutants of Pol2 and Pol3 were defective in NER, and this was rescued by the addition of purified Pol2 and Pol3, respectively.

**Ribonucleotide excision repair (RER)**

Ribonucleotides are the most common DNA lesions present inside the cell, and it was very recently shown that replicative polymerases incorporate ribonucleotides into the genome (127,128). Pol ε incorporates more ribonucleotides than Pol δ and proofreads only about one third of the incorporated ribonucleotides (129,130). Unrepaired ribonucleotides embedded within the DNA lead to short deletions in repeated genomic sequences involving topoisomerase I, and this increases genomic instability. In addition, their presence in DNA makes it prone to cleavage by nicking enzymes leading to single-strand breaks. They also alter the conformation of DNA from B-form to A-form thus affecting replication past ribonucleotides. Therefore, timely removal of ribonucleotides is important for the maintenance of genomic stability. There are several pathways that remove the ribonucleotides from the genome, such as the ones mediated by RNase H2, topoisomerase I, and post-replication repair mechanisms (127,131). Recently, one of these pathways mediated by RNase H2 was reconstituted *in vitro* with purified proteins (132). RNase H2 makes a strand incision on the 5'-end of the incorporated ribonucleotide, and this is followed by extension of the 3'-OH by Pol δ and subsequent displacement of the ribonucleotide in a PCNA-dependent manner. The flap is subsequently cleaved by Fen1, and nicks are sealed by ligase I. Exo1 and Pol ε were also shown to perform RER, although with less efficiency in comparison to Fen1 and Pol δ.
**Okazaki fragment maturation**

Synthesis of DNA in short stretches on the lagging strand requires the removal of the RNA primer and sealing of the gap in the DNA. Single-stranded gaps in the DNA can lead to formation of double-stranded breaks that are often lethal to the cell. Haploid yeast cells contain a genome of $12 \times 10^6$ bp where the lagging strand replication is carried out in short stretches of 150–200 nucleotides called Okazaki fragments. This means that about 100,000 of these fragments must be processed prior to cell division. Okazaki fragment maturation entails displacement of downstream RNA by a DNA polymerase into a flap followed by excision of the flap and finally sealing of the nick by DNA ligase I. The entire process is called nick translation. Numerous pathways have been proposed for Okazaki fragment maturation in eukaryotes, but two pathways appear to be the most prevalent. One involves only Fen1 and other involves both the Dna2 and Fen1 flap endonucleases (30). Pol δ is the primary polymerase involved in Okazaki fragment maturation because of its ability to displace the downstream double-stranded RNA/DNA hybrid and double-stranded DNA in limited stretches. This allows Fen1 to cleave the short flaps and thus limit excessive strand displacement, and this is followed by sealing the nick by ligase. The displacement of only 2 or 3 nucleotides is due to idling by Pol δ, which is regulated by the switch between polymerase and exonuclease sites. However, if the flap generated is too long, RPA binds to the flap and recruits Dna2 that cleaves the flap followed by fine-tuning of the flap by Fen1 and subsequent sealing of the nick (133,134). The role of Pol ε is limited to being a backup polymerase in Okazaki fragment maturation.

**Strand displacement**

The ability of DNA polymerases to open up the downstream double-stranded DNA is called strand displacement synthesis. Strand displacement synthesis is key to many DNA repair processes, including BER, NER, D-loop extension during homologous recombination, and Okazaki fragment maturation. DNA polymerases differ in their ability to perform strand displacement, and replicative polymerases are generally very inefficient in this process (135-138). Strand displacement is limited by the regression pressure exerted by the double-stranded DNA. The force shifts the primer terminus from the polymerase site to the exonuclease site thus limiting the strand displacement (139). DNA Pol III in bacteria and DNA polymerase from bacteriophage T4 cannot perform processive strand displacement synthesis, whereas the bacteriophage φ29 polymerase can perform processive strand displacement synthesis. The TPR2 subdomain found in φ 29 polymerase confers high processivity and strand displacement activity to the enzyme (140).
Aims of the thesis

The focus of this dissertation was to study the structural and functional properties of Pol ε in order to better understand its role in DNA replication and repair. My specific aims were:

1. To obtain a high-resolution structure of Pol2
2. To study the biochemical and kinetic properties of Pol ε
Results and discussion

**Paper I.** Structural basis for processive polymerization by yeast DNA polymerase ε

In this paper, we used X-ray crystallography to determine the ternary structure of Pol2\(_{\text{core}}\) (amino acids 1–1228) at 2.2 Å resolution. Pol2\(_{\text{core}}\) was crystallized in complex with DNA and an incoming dATP. To prevent the exonucleolytic degradation of the DNA, two catalytic residues in the exonuclease site (D290 and E292) were substituted with alanines. The overall structure of Pol2\(_{\text{core}}\) resembles a typical B-family polymerase such as Pol δ or RB69 gp43. The structure bears a resemblance to a right hand cupped around the DNA. The structure contains the typical thumb, palm, fingers, exonuclease, and N-terminal domains. The thumb, palm, and fingers form the polymerase catalytic center. Due to the unstructured nature of the N-terminus, the first 30 amino acids were not visible in the electron density maps.

We observed two distinct differences between Pol ε and other B-family polymerases. The first is the absence of an extended β-hairpin loop in the exonuclease domain, and the second is the presence of a novel domain we have called the P-domain. In B-family polymerases such as Pol δ and RB69 gp43, the β-hairpin consists of a pair of β-strands connected by a loop that extend away from the exonuclease domain into the polymerase active site and toward the single-stranded DNA of the template strand. The function of this β-hairpin loop appears to be to assist the DNA polymerase in switching DNA between the polymerase and exonuclease active sites. Pol ε contains a short β-hairpin loop that does not appear to contact the DNA or other polymerase domains. This seems like somewhat of a paradox because Pol ε has the highest fidelity among B-family polymerase for the incorporation of dNTPs.

The P-domain is an insertion of 100 amino acids emanating from the palm domain that does not appear to be present in any other B-family DNA polymerase. The domain is made of three beta strands, two alpha helices, and a beta hairpin loop. Because of its contacts with DNA, we postulated that it might affect polymerase activity by stabilizing the duplex DNA as it leaves the active site. In fact, deletion of this domain almost completely eliminates the polymerase activity most likely due to the inability of the polymerase to bind DNA. Substituting the three-amino-acid HRK motif with AAA in this domain did not affect the total polymerase activity but affected the processivity of the polymerase, thus suggesting that this domain contributes to the processivity of Pol ε. We attempted to rescue the processivity of these mutants by adding PCNA in a holoenzyme assay, but we observed no stimulation of processivity.
This supports our hypothesis that the P-domain contributes to the intrinsic processivity of Pol ε. The P domain encircles the newly synthesized duplex DNA as it leaves the active site, and we presume this might also be important for sensing replication errors escaping the active site as far as 30 Å to 40 Å from the active site. Furthermore, the base of the P-domain contains two cysteine residues that are tetrahedrally coordinated to a metal ion, most likely Zn^{2+}.

We observed only one metal ion in the polymerase active site, and this metal ion was present in the B-site. The absence of a metal ion in the A-site could be due to the fact that we used a ddC-terminated primer in the crystallization experiments and the missing hydroxyl group disrupts the coordination of the metal ions. The newly synthesized duplex DNA is contacted by conserved amino acid residues in the active site through hydrogen bonds, and these contacts have been shown in other polymerases to be important for replication fidelity. K967 in Pol ε contacts the n−1 and n−2 nucleotides of the primer strand and R988 contacts the n−4 and n−5 phosphate backbone of the newly built DNA. We hypothesize that these amino acid interactions are important for sensing mismatches that escape the detection at the active site.

In conclusion, we determined the structure of Pol2core and found that it has a right-handed cup-shaped structure encircling the DNA. The polymerase contains the usual palm, thumb, fingers, exonuclease, and N-terminal domains of all other B-family polymerases and is structurally similar to RB69 gp43 and Pol δ (77,79). We discovered an apparently unique domain we named the P-domain that is not present in other B-family polymerases. This domain contacts the minor groove of the DNA and imparts high intrinsic processivity to Pol ε. Y-family polymerases also contain an extra domain called the little finger domain, the polymerase-associated domain (PAD), or the wrist domain that encircles the leaving DNA except that this domain is located on the opposite side of the thumb compared to the P-domain and contacts the major groove of the DNA (141-143). The DNA polymerase from bacteriophage φ29 also contains an extra domain that encircles the single-stranded template as it enters the active site (140), and this domain confers high processivity and strand displacement synthesis capability to the DNA polymerase. In addition, Pol ε has a truncated β-hairpin loop extending from the exonuclease domain whereas Pol δ and RB69 have long β-hairpin loops extending into the polymerase active site. The structure of Pol2core will be useful for modeling cancer-associated mutations and for providing a structural perspective on the polymerase’s biochemical properties.
**Paper II.** Yeast DNA polymerase ε catalytic core and holoenzyme have comparable catalytic rates

The second paper deals with studying the kinetic mechanism of DNA polymerization by Pol ε. We compared the kinetic constants of Pol ε with those of Pol2core to understand how the C-terminus of Pol2 and the non-catalytic subunits influence the kinetic properties of Pol ε. We used rapid quench and single turnover kinetics to measure the kinetic constant of Pol ε and Pol2core.

To compare exonuclease rates of Pol2core and Pol ε, we performed exonuclease assays on four different DNA substrates, including single-stranded DNA, perfectly primed-template, single mismatched DNA, and doubly mismatched DNA. The exonuclease rates obtained were similar between the two enzymes except for the degradation of single mismatched DNA where the rate of exonucleolytic degradation by Pol2core was slightly higher than Pol ε. As expected, lower exonuclease rates were obtained for matched DNA than for mismatched DNA or single-stranded DNA.

To measure the maximum rate of polymerization, we performed polymerase assays at different dTTP concentrations, and the products formed were plotted against time in a biphasic exponential association equation to yield \(k_{\text{obs}}\). The \(k_{\text{obs}}\) obtained at each dTTP concentration was plotted against dTTP concentration and fit to a hyperbolic equation to give the maximum rate of polymerization \(k_{\text{pol}}\) and the dissociation constant \(K_{d\text{dTTP}}\). The rates obtained for Pol ε exo− were a \(k_{\text{pol}}\) of \(321 \, \text{s}^{-1}\) and a \(K_{d\text{dTTP}}\) of \(21 \, \mu\text{M}\), and the rates for Pol2core exo− were a \(k_{\text{pol}}\) of \(330 \, \text{s}^{-1}\) and a \(K_{d\text{dTTP}}\) of \(25 \, \mu\text{M}\). The similarity in rate constants suggests that the C-terminus of Pol2 and the catalytic subunits do not affect DNA synthesis.

We also measured the elemental effect of incorporation to study whether a conformational change or the chemistry of bond formation is the rate-limiting step. Using an analog of dTTP called sp-dTTP-\(\alpha\)S, we demonstrated that a conformational change preceding bond formation was rate limiting for both Pol2core exo− and Pol ε exo−. In addition, we measured the elongation rates of DNA polymerization because this not only measures the bond formation but also the translocation of the polymerase along the DNA. A processive polymerization assay was performed at physiological dNTP concentrations. The incorporation of the first six nucleotides was fit to a model based on the assumption that at each nucleotide incorporation event there is a choice between the next nucleotide incorporation, pyrophosphorolysis, and dissociation. The favored process is dependent on the individual rate constant of each reaction. The fitting of the data in the KinTek Explorer software
revealed an average elongation rate of $242 \text{ s}^{-1}$ for Pol $\varepsilon$ exo$^-$ and $176 \text{ s}^{-1}$ for Pol2core exo$^-$.

We next measured the formation of binary complexes and subsequent nucleotide incorporation by Pol $\varepsilon$ exo$^-$ and Pol2core exo$^-$. We found that Pol $\varepsilon$ exo$^-$ took around 1.5 seconds to load onto DNA and incorporate the first nucleotide, but it took significantly longer for Pol2core exo$^-$ to load onto DNA and incorporate the first nucleotide. We also found a correlation between the size of the DNA substrate and the time to incorporate the first nucleotide. The formation of the binary complex and subsequent incorporation of the nucleotide was much faster for Pol $\varepsilon$ exo$^-$ than Pol2 exo$^-$ on a 50/80-mer substrate. However, when we decreased the length of the substrate (12/18-mer), the formation of the binary complex and subsequent incorporation of the first nucleotide was comparable between Pol $\varepsilon$ exo$^-$ and Pol2 exo$^-$. To confirm that the differences on longer substrates were not due to inactive Pol2core exo$^-$, we measured the active site concentration of the polymerases and found them to be almost equal. We hypothesized that the loading of Pol $\varepsilon$ was enhanced by the C-terminus of Pol2 and the non-catalytic subunits of Pol $\varepsilon$.

In conclusion, we found that the exonuclease rates were similar between Pol $\varepsilon$ and Pol2core, which is consistent with similar mutation rates observed in vitro (93). The exonuclease rates obtained with yeast Pol2core are much higher than rates obtained for the human Pol2core for both matched and single mismatched DNA (130,144). This difference could very well be due to the temperature at which the reactions were performed and the DNA substrates used. The $k_{pol}$ of $330 \text{ s}^{-1}$ and $K_{d\text{dATP}}$ of 25 $\mu$M for yeast Pol2core exo$^-$ in this study were comparable to the $k_{pol}$ of $248 \text{ s}^{-1}$ and $K_{d\text{dATP}}$ of 31 $\mu$M previously obtained for human Pol2 exo$^-$ (145). The values obtained were also similar to other replicative polymerases from bacteriophage T7, T4, and RB69 (146-148). The elongation rates for both Pol2core and Pol $\varepsilon$ (100 nucleotides/second) were much faster than the rate of replication fork movement, and this suggests that the rate of replication fork movement is not limited by the polymerase but by the total movement of the replisome, which is perhaps regulated by the movement of the helicase (149,150). The largest differences between Pol2core and Pol $\varepsilon$ were observed during loading of these enzymes onto DNA and subsequent nucleotide incorporation. We propose that the P-domain, which makes the active site less accessible, might obstruct the loading of DNA into the active site. Conformational changes are perhaps needed to position the primer terminus into the active site. This defect in loading is alleviated by the accessory subunits and the C-terminus of Pol2. We suspect that the main factors contributing to the loading of Pol $\varepsilon$ onto DNA are Dpb3 and Dpb4.
because they have been previously shown to increase the affinity of Pol ε for DNA (62).

**Paper III. Switching between polymerase and exonuclease sites in DNA polymerase ε**

This paper deals with how Pol ε processively switches DNA between two active sites despite the absence of an extended β-hairpin loop. The crystal structure of Pol2_core revealed that the polymerase has a truncated β-hairpin loop extending from the exonuclease domain (151). In contrast, the extended β-hairpin loop in RB69 gp43 and Pol δ reaches toward the polymerases active site. In Pol δ, the β-hairpin loop appears to reside between the primer and the template region in the polymerase active site. In RB69 gp43, this loop has been shown to be important for processive switching of the DNA between the two active sites (101). Here, we show that Pol ε can processively switch DNA from the polymerase active site to the exonuclease active site for editing and that the edited primer can be switched back to polymerase active site for resumption of DNA synthesis.

Motif A in Pol ε contains a methionine in position 644 that lies beneath the incoming nucleotide in the active site. Amino acid substitutions at this position have been used to decipher the biological roles played by Pol ε, and genetic and biochemical analysis of M644G and M644L mutants show that these mutants have contrasting properties (127). M644G is a mutator polymerase whereas M644L is an antimutator polymerase. Here, we show that this is perhaps due to altered switching in these two mutants. Compared to the wild-type polymerase, we demonstrate that M644G is poor at switching DNA from the polymerase site to the exonuclease site and is mildly affected in switching DNA from the exonuclease site to the polymerase site. However, the M644L variant is better at switching DNA from the polymerase site to the exonuclease site and is very inefficient at switching DNA from the exonuclease site to the polymerase site. The differences in switching DNA between the two active sites demonstrate a disparity in how the two variants form polymerase and exonuclease complexes. This is supported by our dNTP titration experiment that showed that the M644L mutant is more sensitive to lack of dNTPs than the M644G variant and wild-type Pol ε.

B-family polymerases contain a conserved KKRY motif that is important for stabilizing DNA in the active site. K967 in Pol ε is part of this motif and is homologous to K706 in RB69 gp43, K498 in φ29 DNA Pol, and K814 in Pol δ (80). This amino acid has been shown to form hydrogen bonds with the base of the n–2 nucleotide of the nascent strand. Substitution of this amino acid residue by alanine (K967A) severely affects the stabilization of the primer in
the polymerase active site and leads to the elimination of polymerase activity and the induction of processive exonuclease activity. Similar effects were observed by substituting the analogous amino acid by alanine in RB69 gp43 (K706A) (81). The K498T substitution in Phi29 DNA Pol also affected the polymerase’s ability to synthesize DNA (80).

Amino acid R988 in Pol ε interacts with the n−4 and n−5 bases of the newly synthesized DNA. The homologous amino acids in Pol δ, RB69 gp43, and Phi29 DNA Pol are R839, K734, and K555, respectively. Replacement of R988 with alanine destabilized the primer terminus in the polymerase active site, and mutants were unable to synthesize DNA. The R988K substitution in Pol ε did not yield as severe a polymerization defect as R988A. The R988K mutant polymerase was very inefficient in transferring DNA from the exonuclease site to the polymerase site but was better in transferring DNA from the polymerase to exonuclease site compared to wild-type. Substitution of the homologous amino acid K555 to alanine (K555A) in Phi29 DNA Pol resulted in a polymerase with normal DNA polymerase activity but severely degraded exonuclease activity on double-stranded DNA. This revealed a defect in switching DNA from the polymerase site to the exonuclease site (82).

The mutants were severely affected in elongation, and under multiple-turnover conditions K967A, R988A, and R988K exhibited processive exonuclease activity. Under single-turnover conditions, K967A also exhibited exonuclease activity but a faint band of n+1 synthesis was also observed. Similarly, R988A also exhibited processive exonuclease activity and only produced tiny amounts of 1-nucleotide and 2-nucleotide products. Under single-turnover conditions, the R988K variant synthesized 4 or 5 nucleotides, but most of the products ended at 2 or 3 nucleotides.

This prompted us to insert these mutations into the yeast genome, and the results of this were in agreement with the elongation defects. The K867A cells were inviable, the R988A cells had severe growth defects, and the R988K cells grew as well as wild-type. This experiment also reinforced the importance of the polymerase activity of Pol ε for survival of the cell.

In conclusion, we show that Pol ε can processively transfer DNA between the polymerase and exonuclease active sites despite the absence of a β-hairpin loop. The mechanism of transfer appears to be different than other B-family polymerase due to the truncation of the β-hairpin loop. We demonstrate that the mutator phenotype observed with the M644G variant and the antimutator phenotype observed with the M644L variant are due to altered abilities to switch DNA between active sites. Furthermore, we show that K967 and R988 are important for stabilizing DNA in the polymerase site. The results from the
R988K variant suggest that this amino acid might be important for transferring DNA from the exonuclease site to the polymerase site.

**Paper IV.** Modulation of strand displacement synthesis of DNA polymerase ε by processive 3'-5' exonuclease activity

To examine if Pol ε can perform strand displacement synthesis, we designed a gapped DNA substrate containing an 80-mer template annealed to a 50-mer primer and a 22-mer blocking primer. The two primers were separated by a gap of 8 nucleotides. We found that Pol ε is unable to carry out processive strand displacement synthesis, and it only displaces one base pair when encountering double-stranded DNA after filling a gap. However, exonuclease deficient Pol ε (Pol ε exo−) carries out robust strand displacement synthesis and reaches to the end of the template. Strand displacement is a key event in many cellular processes such as BER, Okazaki fragment maturation, homologous recombination, etc. Therefore, we used biologically relevant substrates to study the possible involvement of Pol ε in these processes. Previous studies have indicated that Pol ε is not part of Okazaki fragment maturation process, and here we expanded on that work and found that it only manages to displace and extend one base pair into the RNA/DNA hybrid (30). Next we asked if Pol ε was able to recognize and extend a nick, and we designed a substrate mimicking a nicked substrate with a 5'-phosphate on the blocking primer. We observed that Pol ε opened up the double-stranded region by one nucleotide whereas Pol ε exo− displaced the downstream blocking primer all the way to the end of the template. To mimic a substrate containing an abasic site, we designed a substrate containing a tetrahydrofuran moiety at the 5'-end. We observed that strand displacement synthesis was much more efficient than on the nicked substrate, and Pol ε displaced most of the tetrahydrofuran moiety into a flap.

Ribonucleotides are the most common form of DNA lesions present in the cell, and it was recently shown that large numbers of ribonucleotides are incorporated into the genome by replicative polymerases (128). These can act as blocks to moving replication forks, and inefficient repair can lead to 2–5 base pair deletions. Therefore, their repair is important for genomic integrity. We designed a substrate containing a 5'-phosphorylated ribonucleotide and examined if it could be displaced by Pol ε. Indeed, Pol ε carried out strand displacement synthesis and displaced the erroneous ribonucleotide region.

To further investigate the effect of a 5'-flap on strand displacement synthesis by Pol ε, we added five more nucleotides onto the 5'-end of the 22mer blocking primer to form a flapped structure. The structure resembles a D-loop, which
is formed during repair of double-strand breaks by homologous recombination. In these experiments, Pol ε was unable to reach the double-stranded junction. We also carried out an analogous experiment with Pol δ and observed that the 5’-flap did not hinder its strand-displacement synthesis. Most Pol δ displaced the DNA strand by one nucleotide and a fraction displaced it by two nucleotides. Our results suggest that the inhibition of Pol ε activity by a flap makes it very unlikely that Pol ε carries out D-loop extension during homologous recombination and long-patch BER. Finally, we showed that limited strand displacement by Pol ε is due to efficient switching of the primer from the polymerase to the exonuclease site. By performing the experiment under single hit conditions in the presence of a heparin trap, we showed that Pol ε displaces only one nucleotide and that Pol ε exo− displaces two nucleotides. The single-nucleotide strand-displacement synthesis by Pol ε means that the switch occurred from the polymerase to the exonuclease site, and this modulated the strand displacement synthesis by removing the second nucleotide as soon it was incorporated.

In conclusion, we show that Pol ε performs strand displacement synthesis of only one nucleotide due to highly processive exonuclease activity when encountering a blocking oligonucleotide. Moreover, Pol ε is inhibited by the downstream flap that mimics a D-loop in front of the moving polymerase. Our results support a role for Pol ε in short-patch BER and RER but not in D-loop extension or long-patch BER.

**Concluding remarks and future perspectives**

In this thesis we examined the structural and functional basis for the high processivity and the high fidelity of DNA synthesis by Pol ε. We determined the high resolution structure of Pol2 core, and measurement of the kinetic rate constants for both the Pol ε holoenzyme and Pol2 core indicate that the structure of the catalytic core is a valid model for studying the biochemical properties of Pol ε. In addition, we demonstrated that Pol ε can processively switch DNA between the polymerase and exonuclease active sites despite the absence of a β-hairpin loop and that Pol ε has a unique P domain that increases the processivity of the polymerase. Future work will be focused on studying the influence of cancer-associated mutations on the biochemical properties of Pol ε.

Future structural work will be focused on obtaining the high-resolution structure of Pol ε with a mismatch at the primer terminus. This will enable us to understand the structural basis for the transfer of DNA between the polymerase and exonuclease active sites. In addition, biochemical experiments with amino acid substitutions in the active sites of Pol ε will also
be performed to study specific amino acids contacts required to transfer DNA between the two active sites.

We also obtained data pertaining to how amino acids K967 and R988 of the thumb domain of Pol ε stabilize DNA in the polymerase active site. In the future, our work will also be focused on studying how the polymerase stabilizes the primer terminus in the exonuclease active site. Based on the sequence alignment with φ29 and the biochemical characterization of φ29 mutants, T293 in the Exo I motif of Pol ε and N388 and F281 in the Exo II motif of Pol ε appear to be potential amino acid residues for interaction with the primer terminus in the exonuclease site.
Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Erik Johansson for your guidance, support and willingness to take time to discuss my research. For giving me the freedom to work on my own way and encouragement to think more independently about my experiments and results. Your guidance will have an everlasting influence on my scientific temperament. I am glad that I have worked under your supervision.

I would like to offer my special thanks to Umesh Varshney, IISc, Bangalore, for introducing me to the world of science. You guidance has been invaluable. I thank you for everything you have done for me.

I want thank all the members of EJ lab past and present. I have been blessed with a friendly and cheerful group of lab mates; I truly admire your helpfulness. I will MISS YOU. Ulf the perfectionist of our lab, I thank you for the amazing times in the lab and the office, lots of memories to cherish. Thank you for your help with the presentations and english pronunciations. Pia Undoubtedly one of the most intelligent persons I know, thank you for critically reading my manuscripts and keeping everything arranged in the lab. Goran, Learned a lot from you hardworking guy, thank you for the parties we had at your place and inviting us to your summer house, we had a blast there. Peter Thank you for always being amusing and your help with the MATLAB. Kristoffer for the good times in the lab, Dominika for the fun loving attitude in the lab and for the parties. Else-Britt for your kindness and keeping lab stocked with different supplies Matthew for introducing me to rapid quench and for proofreading my manuscripts and the thesis. Jarek you are so nice and kind. Pia, Ulf, Dominika thank you for the many hours spent on correcting my thesis (my disastrous typing skills). I really had amazing time with you guys and all the best for your future endeavors.

I thank my co-supervisor Andrei Chabes for your help and recommendations letters. I big thank you to the Chabes lab members.

Nasim thank you sharing you experience with us, you are our inspiring senior. I thank Anna Lena Chabes for enduring our noise in the lab and for the RE enzymes.

I would like to take this opportunity to thank all my friends in other labs for providing very friendly environment and unforgettable coffee breaks. In particular, Jeanette lovely, kind and a caring person I know, it is pleasure to be your friend. Monica for your kind friendship, encouragement and for listening to my non sense Verena for the fun conversations, you filled our days with laughter, Rob the only British in our department, for the humor and the discussions. Great to be your friend, will miss you. Lina Thank you for your friendship and sweet smiles. Mikkeal for the advice and encouragement Irina had a lot of fun chatting with you Ximena perky girl
for the nice times Marcus had lot fun with you both in department and in IKSU. Kristoffer thank you for sharing your enzyme kinetics knowledge with me. Meunender I miss you Munna bhai. Anna Shestova, thank you for taking care of radioactive room and never complaining about it.

I want to thank all the members of the “Friday after work group” for making sure we had fun in the department and built friendship with new recruits. I will always remember the good times.

Over the years I made lot of friends in the department and I thank you all for making my stay in the department so much fun. I would like to offer my special thanks to Dinesh, Miriam, Kamraj, Farah, Jagan, Venki and, Jani. My best wishes for your future.

My special thanks are extended to the staff of the department Clas and Ingrid, I am extremely grateful for your help. I don’t remember a day when your said come back later, you have always dropped your work and helped me. Elisabeth, I am grateful for the early morning autoclaving of my limitless media. A. Sjöström for the help with my applications.

Rahul and Jiyoti for helping me settle in Umeå and for everything else you have done for me.

Nighat and Ajaz Farooqi for making me feel home and always inviting me over for Eid.

My desi friends in Umeå, Khanzeb, Arif and Viqar and your families for the tasty food and the Saturday night parties. I had great time with you guys. Ateeq, Salah and Rashid for your friendship and help. I thank all the friends I made from the other side of the border, I will miss you all.

I am indebted to my pals back home who supported me through all these years, Imran, Dhuha, Intiyaz S, Intiyaz K, Younis, Rafiq, Shahzad and Misbah. I had some of the best moments of life with you guys. Thank you Imran for inquiring about my PhD every other month and staying in touch with me regularly over the years. Intiyaz U20 and Paris memories will last forever. Mushtaq, Javed and Sajad a big thank your help in Bangalore.

I sincerely thank the generous funding agencies that provided financial support during my graduate studies.

Finally, with all my gratitude, I thank my entire family for your endless love and support. I would not have been able to do any of this without you. Thank you for believing in me.
References


function at the initial stage of DNA replication. Molecular and cellular biology **14**, 923-933


DNA polymerase and a polymerase editing complex. *Cell* **99**, 155-166


DNA polymerase alpha by multiple mechanisms. Biochemistry **36**, 8443-8454


exonuclease activity. *The Journal of biological chemistry* 271, 28903-28911


as a 3'→5' exonuclease. *The Journal of biological chemistry* **269**, 438-446


the termination of each Okazaki fragment. *The Journal of biological chemistry* **269**, 24221-24228


