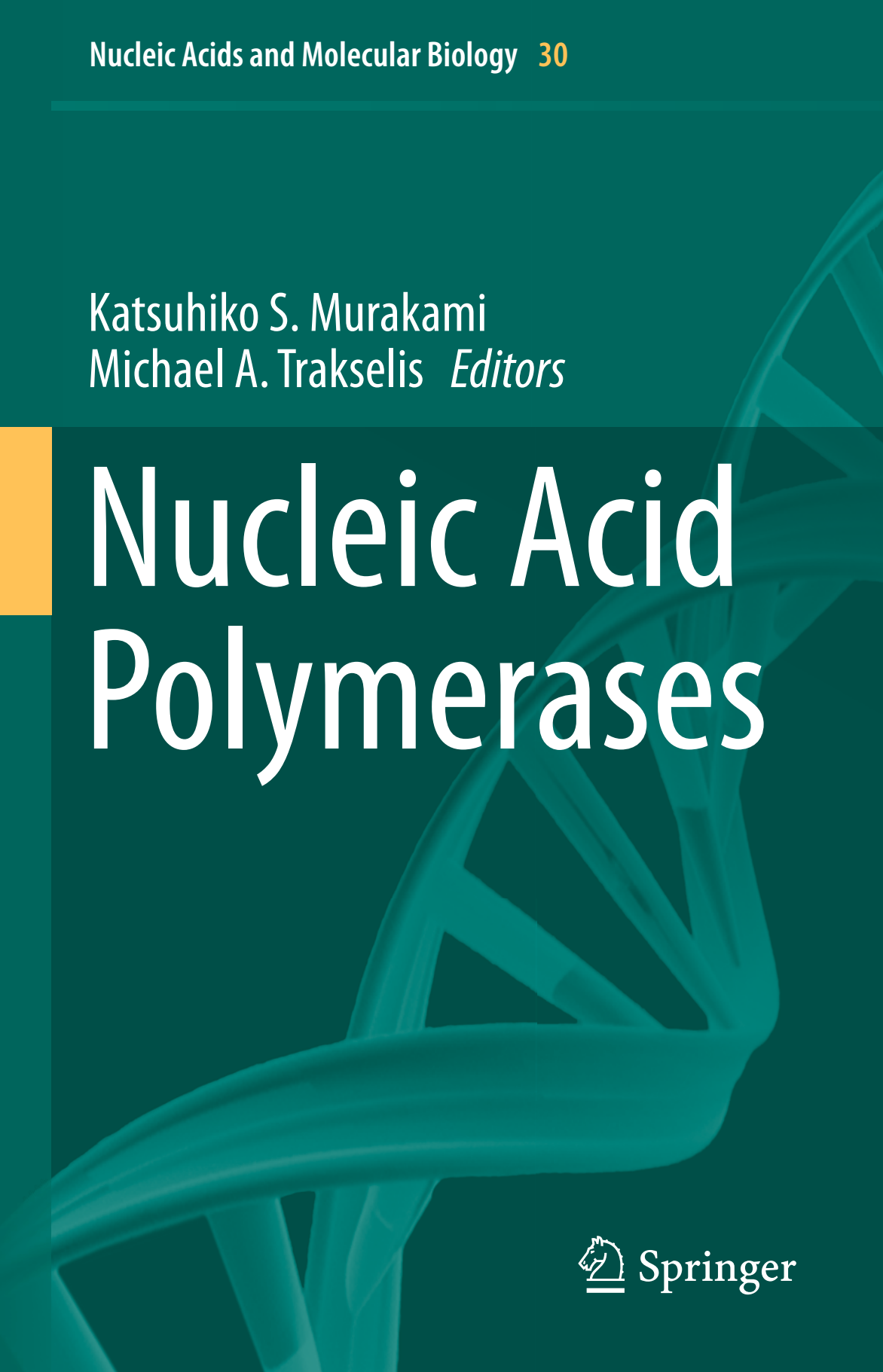
Nucleic Acids and Molecular Biology 30

Katsuhiko S. Murakami
Michael A. Trakselis *Editors*

Nucleic Acid Polymerases



Springer

Nucleic Acids and Molecular Biology

Volume 30

Series Editor
Janusz M. Bujnicki
International Institute of Molecular
and Cell Biology
Laboratory of Bioinformatics and
Protein Engineering
Trojdena 4
02-109 Warsaw
Poland

For further volumes:
http://www.springer.com/series/881

Katsuhiko S. Murakami • Michael A. Trakselis
Editors

Nucleic Acid Polymerases

Springer

Editors
Katsuhiko S. Murakami
Dept. of Biochem. and Mol. Biology
The Pennsylvania State University
University Park
Pennsylvania
USA

Michael A. Trakselis
Department of Chemistry
University of Pittsburgh
Pittsburgh
Pennsylvania
USA

ISSN 0933-1891
ISSN 1869-2486 (electronic)
ISBN 978-3-642-39795-0
ISBN 978-3-642-39796-7 (eBook)
DOI 10.1007/978-3-642-39796-7
Springer Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

More than any other class of enzymes, nucleic acid polymerases are directly responsible for an overabundance of enzymatic, regulatory, and maintenance activities in the cell. DNA polymerases accurately replicate copies of genomes in all forms of life as well as have specialized roles in DNA repair and immune response. RNA polymerases are most noted for their active roles in controlling gene expression during transcription but can also be utilized in self-replicating ribozymes and viral replication. Although the general sequence homology, structural architecture, and mechanism are conserved, they have evolved to incorporate deoxynucleotides (dNTPs) or ribonucleotides (rNTPs) explicitly. Various nucleic acid polymerases have specificities for RNA or DNA templates, incorporate dNTPs or rNTPs, and can be template dependent or independent. Here, we provide examples on the latest understanding of each class of nucleic acid polymerase, their structural and kinetic mechanisms, and their respective roles in the central dogma of life.

This book provides a catalog and description of the multitude of polymerases (both DNA and RNA) that contribute to genomic replication, maintenance, and gene expression. Evolution has resulted in tremendously efficient enzymes capable of repeated extremely rapid syntheses that have captivated researchers' interests for decades. We are inspired by work that started over 60 years ago and is actively pursued today for a fundamental understanding of life, contributions to human health and disease, and current and future biotechnology applications. Nucleic acid polymerases are fascinating on a number of levels, yet still continue to surprise us with novel modes of action revealed through ongoing and future studies described within this volume.

We wish to thank all the authors for their specific expertise and willingness to participate in this comprehensive review of nucleic acid polymerases. We are also grateful to the many investigators before us (including our research mentors: Stephen Benkovic and Akira Ishihama) who began and continue this important

line of research. We believe this book will be useful for a wide range of researchers in both the early and later stages of their careers. We would be thrilled if this volume becomes the go-to resource for nucleic acid polymerase structure, function, and mechanism for years to come.

Pittsburgh, PA — Michael A. Trakselis
University Park, PA — Katsuhiko S. Murakami

Contents

Chapter 1
Introduction to Nucleic Acid Polymerases: Families, Themes, and Mechanisms

Michael A. Trakselis and Katsuhiko S. Murakami

Keywords Polymerase • Mechanism • Structure • Function • Catalysis

Abbreviations

CPD	Cyclobutane pyrimidine dimers
E. coli	*Escherichia coli*
FDX	Fidaxomicin
FILS	Facial dysmorphism, immunodeficiency, livedo, and short statures
kDa	Kilodaltons
pol	Polymerase
Pol I	*E. coli* DNA polymerase I
RdRp	RNA-dependent RNA polymerase
Rif	Rifampicin
rRNA	Ribosomal RNA
TLS	Translesion synthesis
UV	Ultraviolet light
XPD	*Xeroderma pigmentosum*

M.A. Trakselis (✉)
Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260, USA
e-mail: mtraksel@pitt.edu

K.S. Murakami (✉)
Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA 16802, USA

The Center for RNA Molecular Biology, Pennsylvania State University, University Park, PA 16802, USA
e-mail: kum14@psu.edu

K.S. Murakami and M.A. Trakselis (eds.), *Nucleic Acid Polymerases*, Nucleic Acids and Molecular Biology 30, DOI 10.1007/978-3-642-39796-7_1,

1.1 Introduction/Discovery/Classification

Template-dependent and template-independent nucleotidyl transfer reactions are fundamentally important in the maintenance of the genome as well as for gene expression in all organisms and viruses. These reactions are conserved and involve the condensation of an incoming nucleotide triphosphate at the 3′ hydroxyl of the growing oligonucleotide chain with concomitant release of pyrophosphate. DNA polymerase I (Pol I) isolated from *E. coli* extracts was initially characterized in in vitro reactions well over 50 years ago by the seminal work of Arthur Kornberg's laboratory (Kornberg 1957; Lehman et al. 1958; Bessman et al. 1958). Inspired by this work, the discovery of a DNA-dependent RNA polymerase quickly followed in 1960 from a variety of researchers including Samuel Weiss (Weiss and Gladstone 1959), Jerald Hurwitz (Hurwitz et al. 1960), Audrey Stevens (Stevens 1960), and James Bonner (Huang et al. 1960). These early enzymatic characterizations of DNA-dependent deoxyribonucleotides and ribonucleotide incorporations gave credibility both to Watson and Crick's DNA double helix model (Watson and Crick 1953) and the transcription operon model proposed by François Jacob and Jacques Monod (Jacob and Monod 1961).

Prior to 1990, few DNA polymerase members were known. Pol I, Pol II, and Pol III from bacteria defined the initial A, B, and C families of polymerases (Braithwaite and Ito 1993), respectively. Eukaryotic polymerases adopted a Greek letter nomenclature (Weissbach et al. 1975) and included cellular B-family polymerases pol α, pol β, pol δ, pol ε, and pol ζ, and the mitochondrial A-family polymerase pol γ. Rapid progress in genome sequencing, search algorithms, and further biochemical analysis identified other putative DNA polymerases both in bacteria and eukaryotes prompting the expansion of the Greek letter nomenclature. Families D, X, and Y were created to classify unique polymerase in archaea as well as those with specialized functions in DNA repair (Burgers et al. 2001; Ishino et al. 1998; Ohmori et al. 2001). After inclusion of reverse transcriptase enzymes that are RNA-dependent DNA polymerases including telomerase, the DNA polymerase families now number seven (Table 1.1). Although the number of human DNA polymerases stands at 16 members, a recently characterized human archaeo-eukaryotic (AEP) DNA primase (Prim-Pol) has both RNA and DNA synthesis abilities (L. Blanco, personal communication) suggesting that other uncharacterized enzymes may have additional unidentified roles in DNA synthesis. This chapter introduces and highlights chapters within this series and puts the DNA and RNA polymerase families, structures, and mechanisms in context.

1.1.1 DNA Polymerase Families and Function

Most DNA-dependent DNA polymerases have a single catalytic subunit (Fig. 1.1). These single subunits are generally active on their own but are regulated with

Table 1.1 Model DNA polymerase family members

Family	Viral	Bacterial	Archaeal	Eukaryotic
A	T7 gp5	Pol I		Pol γ (mito)
		Klenow[a]		Pol θ
		Taq Pol		Pol ν
B	T4/RB69 gp43	Pol II	PolB1 (B2 and B3)[b]	Pol α
	phi29 Pol			Pol δ
				Pol ε
				Pol ζ
C		Pol III		
D			Pol D[c]	
X				Pol β
				Pol λ
				Pol μ
				TdT
Y		Pol IV	Pol Y	Pol η
		Pol V		Pol ι
				Pol κ
				Rev1
RT[d]	RT			Telomerase

[a]Klenow is the C-terminal truncation of *E. coli* Pol I
[b]Crenarchaea generally have thee Pol B enzymes, while euryarchaea have one
[c]Pol D is only found in euryarchaea phyla of archaea
[d]Reverse transcriptase (RT) is RNA-dependent DNA polymerase

regard to function through various accessory proteins that direct and restrain catalysis to specific DNA substrates. For the most part, Y-family DNA repair polymerases adopt a slightly more open active site to accommodate base damage and are devoid of any proofreading exonuclease domains (Chap. 4). These structural features are required for replication past a variety of lesions in the template strand during DNA replication to maintain the integrity of the fork. DNA repair polymerases (X and Y family) are actively involved in maintaining our genome under intense DNA-damaging stressors. The ability to prevent mutagenesis is their main cellular role, but changes in expression levels and disruptions of DNA repair pathways are common in promoting cancer and tumorigenesis (Chap. 3). Viral, bacterial, and some archaeal DNA polymerases (A and B families) are primarily single-subunit enzymes. They are held at the replication fork through dynamic interactions with accessory proteins to maintain high local concentrations during active replication (Chap. 6). The DNA replication polymerase enzymatic accuracy (fidelity) is unprecedented and is primarily responsible for maintaining stable genomes of all organisms. In bacteria and eukaryotes, the DNA replication polymerases (B and C families) have evolved to contain additional subunits that are almost always associated with the catalytic subunit as holoenzyme complexes (Chap. 2). The recently discovered D-family replication polymerases from certain archaea are also multisubunit enzymes and are presumably ancestral precursors to their eukaryotic homologs (Chap. 6). The polymerase accessory subunits have a variety of roles that are only just being identified including maintaining structural

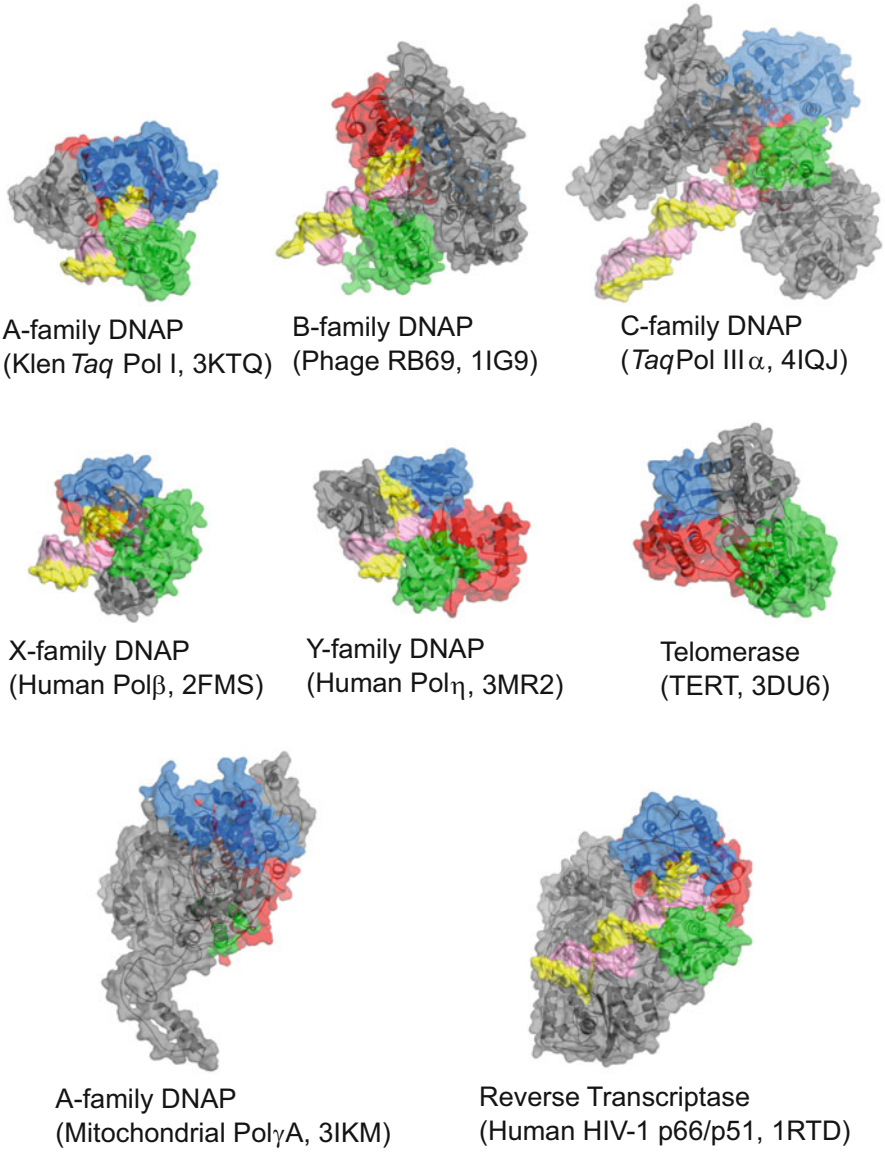


Fig. 1.1 Gallery of DNA polymerases. All polymerases and DNA are shown as cartoon models with partially transparent molecular surfaces. Their names of family, sources, and PDB accession codes are indicated. Protein structures are colored *gray* and key subdomains are colored (thumb, *green*; palm, *red*; fingers, *blue*). Nucleic acids are colored *yellow* for the template DNA and *pink* for the primer DNA. All polymerases are depicted using the same scale in this figure and also in Fig. 1.2 for direct comparison of their sizes

geometries, additional enzymatic activities, and interactions with other polymerase accessory proteins.

Although DNA polymerases are generally template-dependent enzymes and follow Watson and Crick base-pairing rules, there is a subclass of polymerases (primarily X family) that are template independent (Chap. 5). These polymerases are involved in aspects of DNA repair where DNA strands have lost connectivity and require additional nucleotide additions at the ends to facilitate repair. They also have a unique biological role contributing to random incorporations and corresponding diversity of antibodies required for immunological responses. A similar type of DNA extension is also required at the ends of chromosomes to maintain their length during DNA replication, but instead of random nucleotide additions, the enzyme telomerase uses an RNA template strand as a cofactor for sequence-specific DNA repeat additions called telomeres (Chap. 9). This RNA-dependent DNA polymerase, telomerase, is unique to organisms with linear genomes and is also implicated in a variety of human diseases and aging.

1.1.2 RNA Polymerase Families and Function

All cellular organisms including bacteria, archaea, and eukaryotes use multi-subunit DNA-dependent RNA polymerase for transcribing most of the RNAs in cells (Fig. 1.2) (Werner and Grohmann 2011). Bacterial RNA polymerase is the simplest form of this family (composed of the minimum five subunits), whereas archaeal and eukaryotic RNA polymerases possess additional polypeptides to form ~11–17 subunit complexes. Bacteria and archaea use a single type of RNA polymerase for transcribing all genes, whereas eukaryotes have three different enzymes, Pol I, Pol II, and Pol III, and synthesize the large ribosomal RNA (rRNA) precursor, messenger RNA (mRNA), and short untranslated RNAs including 5S rRNA and transfer RNA (tRNA), respectively (Chap. 12). In plant, there are two additional 12-subunit RNA polymerases, Pol IV and Pol V, that play important roles in RNA-mediated gene-silencing pathways (Chap. 13). The archaeal transcription system has been characterized as a hybrid of eukaryotic and bacterial transcription systems; the archaeal basal transcription apparatus is very similar to that of eukaryote, but its transcriptional regulatory factors are similar to those of bacteria (Hirata and Murakami 2009; Jun et al. 2011).

Bacteriophage encodes single-unit RNA polymerase of ~100 kDa molecular weight, which expresses bacteriophage genes on host bacterial cells for generating progeny phage particles (Chap. 10). Although bacteriophage RNA polymerase is about four times smaller than the cellular RNA polymerases (Fig. 1.2), it is able to carry out almost all functions in transcription cycle observed in cellular RNA polymerases. Its primary and three-dimensional structures are similar to A-family polymerase, which also includes mitochondrial RNA polymerase expressing genes from mitochondrial DNA (Fig. 1.2) (Chap. 11).

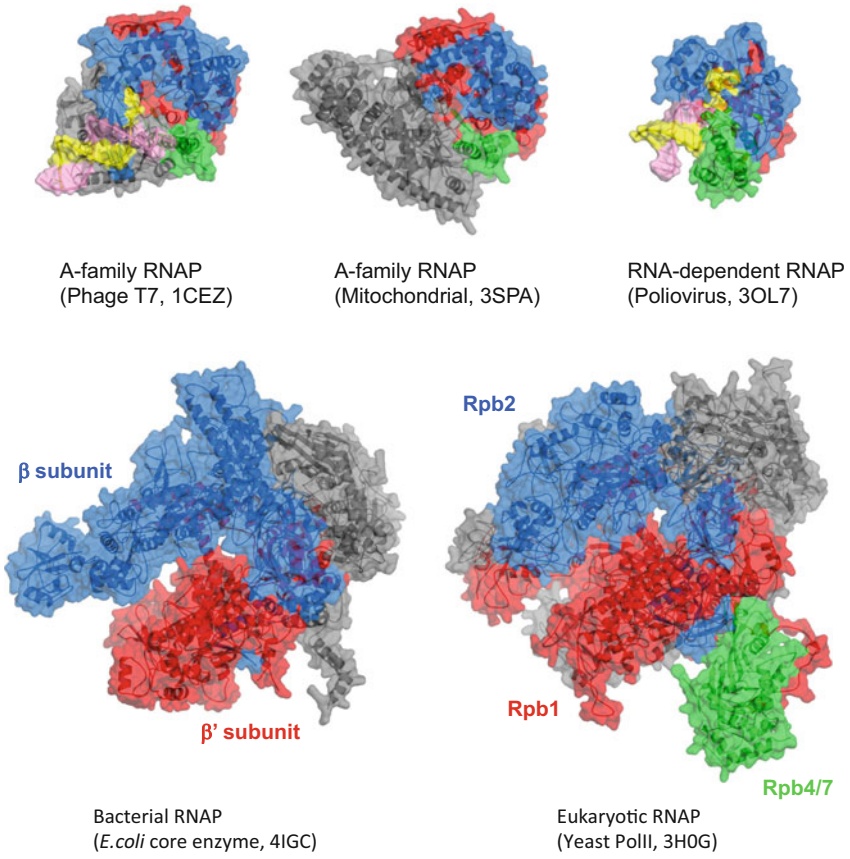


Fig. 1.2 Gallery of RNA polymerases. All polymerases and DNA are shown as cartoon models with partially transparent molecular surfaces. Their names of family, sources, and PDB accession codes are indicated. Protein structures are colored *gray* and key subdomains are colored (thumb, *green*; palm, *red*; fingers, *blue* for the A-family bacteriophage-type RNA polymerase and RdRp; largest subunit, *red*; second largest subunit, *blue*; protruding stalk, *green* for the cellular RNA polymerases). Nucleic acids are colored *yellow* for the template DNA and *pink* for the non-template DNA. All polymerases are depicted using the same scale in this figure and as in Fig. 1.1 for direct comparison of their sizes

RNA viruses including influenza, rhinovirus, hepatitis C, and poliovirus have RNA-dependent RNA polymerases (RdRps) that are responsible for replicating their RNA genomes and expressing their genes (Fig. 1.2). The RdRps are targets for antiviral therapies, but their higher mutation rates due to lack of proofreading endonuclease activity generate resistant variants to compromise antiviral therapies (Chap. 14).

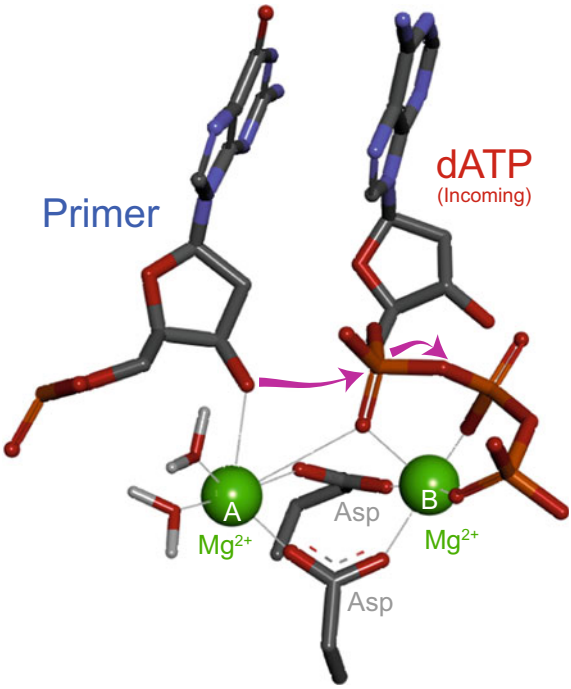


Fig. 1.3 Highlights the two-metal-ion mode of catalysis for DNA and RNA polymerases. Two conserved aspartates coordinate metals A and B in the active site. Metal A activates the 3′-OH for attack on the 5′ α-phosphate of the incoming nucleotide (either dATP or ATP) with release of β-γ pyrophosphate. Metal B neutralizes the negative charge on the phosphates as well as buildup in the transition state

1.2 Conserved Polymerase Structures

The original structure of the C-terminal fragment of *E. coli* Pol I (Klenow fragment) identified the general architecture of DNA polymerases to resemble a right hand with subdomains similar to fingers, thumb, and palm regions (Fig. 1.1) (Ollis et al. 1985). Although sequence homology from different DNA polymerase families has diverged quite significantly, the general organization of all polymerase structures is very similar (Figs. 1.1 and 1.2), suggesting that they may have evolved from a common ancestor. In fact, both the DNA and RNA polymerases catalyze essentially the same chemical reaction with subtle differences ensuring accurate incorporation of their respective nucleotides (Fig. 1.3) (Steitz 1993).

1.2.1 DNA Polymerase Structural and Kinetic Mechanisms

The two most important and conserved residues are aspartates contained within the palm domain that act to coordinate two metal ions (Mg^{2+}) for catalysis (Fig. 1.3). Metal A lowers the bonding potential of the hydrogen at the 3′-OH, activating the 3′-O^- for attack at the α-phosphate of the incoming nucleotide. Metal B aids pyrophosphate leaving and stabilizes structures of the pentacovalent transition

state. The mechanism was originally proposed based on the 3′–5′ removal of nucleotides in the exonuclease site of DNA Pol I (Beese et al. 1993). This two-metal-ion mechanism for phosphoryl transfer is identical for DNA and RNA polymerases and extremely similar to analogous reactions involving RNA-catalyzed reactions including splicing (Steitz and Steitz 1993). It is hypothesized that this mechanism was the basis of catalysis in the RNA world and has maintained its core features with all modern polynucleotide polymerases. Interesting, this basic two-metal-ion mechanism has recently been challenged by the observation of a third metal ion in the active site of pol η that acts to neutralize the negative charge buildup in the transition state and protonates the leaving group pyrophosphate (Nakamura et al. 2012). It will be interesting to see if this transient third metal ion also exists in other polymerases suggesting a common theme and expansion of the traditional two-metal-ion phosphoryl transfer mechanisms. The two other domains (fingers and thumb) have diverged significantly throughout the polymerase families but contain functionally analogous elements. The fingers domain acts to correctly position the incoming nucleotide with the template, while the thumb domain aids in DNA binding and successive nucleotide additions (processivity).

To increase the fidelity (accuracy) of continuous nucleotide incorporation, some DNA polymerases from the A, B, and C families have a separate exonuclease (3′–5′) domain which verifies correct incorporation and removes an incorrectly incorporated base. For bacterial and archaeal family A and B polymerases, the exonuclease activity is included in a separate domain within the contiguous polypeptide sequence. In the *E. coli* Pol III holoenzyme as well as eukaryotic B-family polymerase, the exonuclease activity is contained within a separate polypeptide. The first structure of DNA bound in the exonuclease domain was with the Klenow fragment and suggested a common two-metal-ion catalysis mechanism for removal of nucleotides as well (Beese and Steitz 1991). The exonuclease site was defined as having three conserved carboxylate residues coordinating both metal ions, binding to the DNA, and activating catalysis and removal of an incorrectly incorporate base.

In addition to exonuclease activity, high-fidelity DNA polymerases also maintain accuracy through kinetic checkpoints ensuring accurate base pairing (Fig. 1.4). The general consensus is that the polymerase domain alone accounts for fidelity values of 10^{-5} to 10^{-6} and inclusion of the exonuclease proofreading domain contributes another 10^{-2} for total fidelity values of 10^{-7} to 10^{-8} (1 error in every 100 million or 99.999999 % accurate) (Kunkel 2004). DNA polymerases from other families including X and Y have significantly lower fidelity values (10^{-2} to 10^{-5}) accounted by the more frequent error rates, lesion bypass abilities, and absent exonuclease domains (Chaps. 3–5).

For the majority of A-, B-, as well as some Y-family polymerases, a slow step prior to chemistry (step 3) ensures correct base pairing before phosphodiester bond formation. Based on the fusion of structural and kinetic data, it was originally postulated that an open-to-closed transition in the fingers domain was the slow step in the mechanism. More recently, the open-to-closed transition was measured directly using fluorescence and found to be fast relative to step 3, prompting the

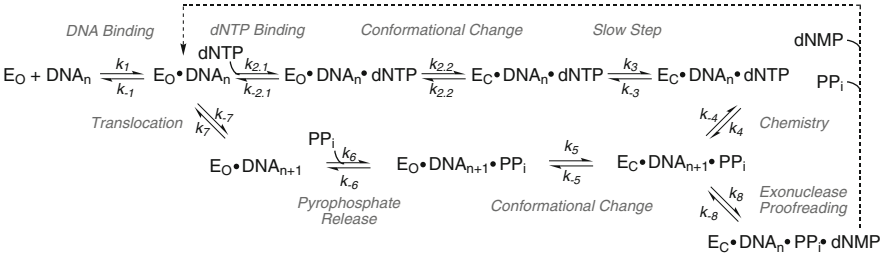


Fig. 1.4 General kinetic mechanism for high-fidelity DNA polymerases. The enzyme undergoes a fast open (E_O)-to-closed (E_C) transition (step 2.2) after binding DNA and nucleotide. Kinetic checkpoints include the slow step prior to chemistry (step 3) as well steps 5 and 8 after chemistry to activate the proofreading function if necessary

inclusion of steps 2a and 2b into the mechanism (Joyce et al. 2008; Johnson 2010). The identity of the slow step 3 is still unknown and may instead be associated with a change in metal ion coordination of either metal B or an incoming metal C in preparation of moving forward through the transition state towards chemistry (Nakamura et al. 2012).

The kinetic checkpoints themselves ensure that correct nucleotides are optimally positioned in the active site over incorrect ones to promote catalysis. Prevention of rNTP binding in DNA polymerase active sites is restricted by a steric gate towards the 2′-OH (step 2.1) (Delucia et al. 2006) as well as reduced rate of fingers closing (step 2.1) limiting their incorporation (Joyce et al. 2008). For Klenow and T7 pol, incorrect dNTP incorporation is prevented by a slower chemistry step 4 than for correct dNTPs defining polymerase fidelity (Dahlberg and Benkovic 1991). Step 5 following phosphoryl transfer is also considered a kinetic slow step and is important for increasing the probability for proofreading (step 8) in the case of a misincorporated base (Kuchta et al. 1988). Although this is not an absolute kinetic mechanism for nucleotide selection in all DNA polymerases (Fig. 1.4), the basic principles explain a number of the mechanistic facets required for maintaining high nucleotide fidelity. Whether or not this complex scheme holds as a general mechanism for all DNA polymerases remains to be determined, but it is likely to be accurate for high-fidelity DNA polymerases in particular.

Fast, successive, and accurate nucleotide additions require that the polymerase remains associated with the template after a translocation step (step 7) for multiple rounds of catalysis or processivity. DNA polymerases by themselves are not highly processive and are not able to incorporate more than 20–50 successive nucleotides in a single binding event. The exception seems to be the B-family DNA polymerase from phi29 which has extremely robust strand displacement activity and processivity of replication of several thousand bases (Blanco et al. 1989; Kamtekar et al. 2006). Phi29 Pol has a specific insertion called the terminal protein region 2 (TPR2) that acts with the palm and thumb subdomains to encircle and close around the DNA template limiting dissociation. Increased processivity has also been seen after oligomerization of some archaeal DNA polymerases effectively

encircling the template (Chap. 6). In both of these examples, the polymerases use a topological linkage of the protein to DNA to remain bound to the template for efficient and successive incorporations. More commonly, interactions of DNA polymerases with toroidal accessory factors (clamp proteins) achieve the same result of increased processivity by coupling the DNA polymerase with the template, limiting dissociation (Trakselis and Benkovic 2001; Bloom 2009). These clamp proteins (PCNA, in particular) have specific interaction domains that bind consensus sequences in DNA polymerases and other genomic maintenance proteins that act to recruit and retain enzymes at the replication fork (Moldovan et al. 2007).

1.2.2 RNA Polymerase Structural Mechanism

For the nucleotidyl transfer reaction by RNA polymerase, a two-metal-ion catalytic mechanism has been proposed, which is common in the DNA polymerase, as the enzyme possesses two divalent catalytic and nucleotide-binding metal cations (Mg^{2+}) chelated by two or three Asp residues at the enzyme active site (Fig. 1.3). Both metal ions are proposed to have octahedral coordination at physiological Mg^{2+} concentrations.

For transcribing RNA using DNA template, DNA-dependent RNA polymerase including cellular RNA polymerases (Chaps. 12 and 13), bacteriophage RNA polymerase (Chap. 10), and mitochondrial RNA polymerase (Chap. 11) unwinds a small region of double-stranded DNA to the single-stranded form and synthesizes RNA as a complementary sequence of the template. The unwound DNA region is called the transcription bubble that contains a DNA-RNA hybrid of ~8 base pairs. For synthesis of RNA, nucleotide substrate and catalytically essential divalent metals in addition to the single-stranded template DNA must be accommodated at the active site. One of four ribonucleotide triphosphates—ATP, GTP, CTP, and UTP—forms a Watson–Crick base pair with a DNA template base, and its α-phosphate group is attached by a 3′-hydroxyl of the growing end of the RNA. As a result, a linear RNA polymer is built in the 5′ to 3′ direction.

The overall shape of cellular RNA polymerases including bacterial, archaeal, and all types of eukaryotic enzymes is crab claw-like with a wide internal channel for double-stranded DNA binding (Fig. 1.2, Chaps. 12 and 13). The enzyme active site is located on the back wall of the channel, where an essential Mg^{2+} ion is chelated by three Asp of the absolutely conserved NADFDGD motif in the largest subunit. Compared to the bacterial RNA polymerase, archaeal and all eukaryotic RNA polymerases possess a characteristic protruding stalk that is formed by a heterodimer, and their relative positioning of the main body and stalk is also highly conserved.

The structure of bacteriophage-type RNA polymerases including mitochondrial enzyme resembles cupped right hand with palm, fingers, and thumb subdomains and a cleft that can accommodate double-stranded DNA (Fig. 1.2). Not only the overall structure of polymerases but also the secondary structures of subdomains in

the bacteriophage-type RNA polymerase are highly conserved in the A-family DNA polymerase. The enzyme active site is located on the palm subdomain, where an essential Mg^{2+} ion is chelated by two Asp of the conserved motifs A and C. The conserved motif B is in the mobile finger subdomain, which changes its position during the nucleotide addition cycle and plays an important role in the nucleotide selection (Chaps. 10 and 11).

The overall shape of RdRps is similar to other nucleic acid polymerases, having "cupped right hand" structure and fingers, thumb, and palm subdomains (Fig. 1.2, Chap. 14). Because of its extension of the fingers, RdRp has more fully enclosed the active site, which may enhance their protein stability and enzyme processivity for their genomic RNA replication function.

1.3 Implications in Disease/Therapy

Although accurate DNA synthesis is a hallmark of high-fidelity DNA replication polymerases, a number of other polymerases have been implicated in various diseases and aging underlying their importance for further study. The best known example involves telomerase. When normal somatic cells replicate in the absence of telomerase, they undergo successive shortening of their telomeric ends, termed the end replication problem (Allsopp and Harley 1995). The shortening of telomeres acts as a clock determining the life of a cell ultimately causing senescence and cell death. However, in cancer cells, telomerase is upregulated preventing telomeric shortening and increasing cellular survival giving rise to immortal cells found in tumors. Although telomerase deficiency is most notable in the genetic disorder, *dyskeratosis congenita*, mutations in telomerase are also associated with anemia, other bone marrow-related diseases, and lung fibrosis. The unifying diagnostic indicator in all cases is short telomeres (Armanios 2009). The mechanism of RNA-mediated DNA telomeric synthesis by telomerase will be discussed in great detail in Chap. 9.

Translesion DNA polymerases are specialized low-fidelity DNA polymerases that can insert bases opposite a lesion, bypassing the damage, while potentially inducing point mutations. It is hypothesized that potential mutagenesis is favored over complete replication arrest and fork collapse. Translesion synthesis (TLS) is regarded as being responsible for the large increase in point mutations found in cancer genomes (Bielas et al. 2006). These Y-family DNA polymerases generally have much less fidelity and more open active sites accommodating a variety of DNA template lesions including oxidations, deaminations, abasic sites, methylations, and a host of environmental mutagens and are described in detail in Chap. 4.

Mutations in the Y-family Pol η account for the inheritable genetic disease, *xeroderma pigmentosum* (XPD) (Masutani et al. 1999). This disease sensitizes cells to UV light, significantly increasing the risk of skin carcinomas. Pol η is known to bypass thymine cyclobutane pyrimidine dimers (CPD) caused by UV cross-linking

of adjacent residues (Johnson et al. 2000). Genetic mutations in Pol η associated with XPD disrupt the contacts with the DNA limiting its activity. Therefore, this translesion DNA polymerase has evolved a specific role in replication over UV-induced damage, and mutations in Pol η are responsible for replication fork collapse, double-strand breaks, and chromosomal breaks. The only other DNA polymerase found to be associated with an inheritable genetic disease is Pol ε where mutations in the large subunit give rise to splicing changes that cause decreased expression and predicted truncated protein products in FILS (facial dysmorphism, immunodeficiency, livedo, and short statures) syndrome patients (Pachlopnik Schmid et al. 2012).

The X-family base excision repair DNA polymerase (Pol β) (Chap. 3) has been found to have sporadic mutations in human tumors (Starcevic et al. 2004). Increased expression of Pol β has also been measured in a number of cancers interfering with normal DNA replication and causing mutations (Albertella et al. 2005; Tan et al. 2005). Several small-molecule inhibitors have been found to increase sensitivity to chemotherapeutic agents by blocking action of Pol β and seem to be a viable avenue for cancer therapy (Goellner et al. 2012). Other X-family DNA polymerases including terminal deoxytransferase (TdT) (Chap. 5) have been implicated in leukemia and carcinomas through altered expression levels. New nucleoside analogs have been shown to be specific towards TdT controlling expression levels and sensitizing cancer cells to conventional treatment.

Therefore, it seems there is an opportunity for targeted X- and Y-family DNA polymerase inhibition by either controlling expression levels or using as adjuvants with DNA-damaging radiation or chemotherapy (Lange et al. 2011). The challenges will be to avoid toxicity issues common with previous inhibitors, selectively target cancer cells, and act specifically on one of the 16 human DNA polymerases. Not an easy task, but with preliminary successes for Pol β and Tdt, the opportunity also exists for other DNA polymerases. The more we can emphasize structural differences in the active sites, identify allosteric regions, or detect novel mechanistic features, the better a position we are in to develop novel therapeutic agents. Success will require understanding the balance of DNA polymerase actions in a variety of cell types and developing screening methods to simultaneously measure effects on multiple DNA polymerases.

RNA polymerase is an essential enzyme in bacteria and virus and, as such, is a proven target for antibiotics and antiviral drugs (Chaps. 11 and 14). Fidaxomicin (FDX) is an inhibitor of bacterial RNA polymerase and is one of the newest antibiotics approved by the US Food and Drug Administration (FDA) for treatment of *Clostridium difficile*-associated diarrhea. The best characterized antibiotic against bacterial RNA polymerase is Rifampicin (Rif), which has been used as the first-line drug for infectious bacteria treatment, including tuberculosis, over four decades. However, a high incidence of Rif-resistant bacterial strains with RNA polymerase mutations is one of our public health challenges. Although many Rif-resistant *Mycobacterium tuberculosis* with RNA polymerase mutants have been derived in laboratory, only three residues with specific amino acid substitution account for ~85 % of *M. tuberculosis* Rif-resistant strains found in clinical isolates.

Therefore, structures of these three Rif-resistant RNAPs can move one step forward the structure-based discovery of improved Rif for tuberculosis treatment.

1.4 Remaining Questions and Future Directions

DNA and RNA polymerases have evolved naturally over millions of years to be highly accurate enzymes for incorporating the four available deoxyribonucleotides for DNA synthesis and ribonucleotides for RNA synthesis to faithfully maintain genomes and to express genes. Recent research efforts have focused on evolving these high-fidelity enzymes to have altered enzymatic properties or nucleotide specificities required for a variety of biotechnology applications (Chap. 7). Some of the goals are to amplify ancient genomes, incorporate alternative genetic alphabets, and replicate chemically and environmentally modified templates. In addition to traditional biotechnology development, these engineered polymerases have the potential to revolutionize synthetic biology by creating safe artificial living systems that incorporate unnatural DNA analogs for the creation of anything from drugs/metabolites to energy.

Although the basic mechanisms of incorporation, proofreading, and fidelity are well characterized for a number of DNA polymerases, there are many remaining questions on how polymerases function within the context of the replisome during normal DNA replication or repair. For example, do the kinetics or fidelities change drastically when accessory proteins are interacting with the polymerase? How does the unwinding rate of the DNA helicase control DNA polymerase kinetics? How is high-fidelity synthesis coordinated with error-prone lesion bypass when multiple polymerases are available? Answers to these questions will require the ability to both assemble and test more components of the replisome simultaneously in vitro and probe the kinetics within the context of an actively replicating cell.

The expression levels of DNA polymerases in various cancer cell types and stem cells are also an exciting avenue for study. Stem cells in particular need to maintain highly stable genomes. In these cells, the distribution of polymerases should favor high-fidelity enzymes and may even include suppressors against X- and Y-family polymerases. On the other hand, cancer cells are active mutators, and it would not be surprising to find inactivating mutations, loss, or rearrangements of DNA polymerases as more individual cancer cell sequencing results are available. In addition, DNA polymerases may be inactivated through alterations in DNA methylation patterns or RNAi changes. Pol θ in particular has been shown to have a significant difference in expression levels in breast tumors over non-tumor cells (Lemee et al. 2010). Therefore, it will also be important to assess any expression level deviations for DNA polymerases in individual cells to understand equilibrium changes that may be occurring at the replication fork and their resulting consequences on genomic stability. If DNA polymerase distributions can be determined first, it is conceivable that targeted DNA polymerase therapies will better sensitize cells to radiation or chemotherapy.

References

Albertella MR, Lau A, O'Connor MJ (2005) The overexpression of specialized DNA polymerases in cancer. DNA Repair (Amst) 4(5):583–593

Allsopp RC, Harley CB (1995) Evidence for a critical telomere length in senescent human fibroblasts. Exp Cell Res 219(1):130–136

Armanios M (2009) Syndromes of telomere shortening. Annu Rev Genomics Hum Genet 10:45–61

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

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

Bessman MJ, Lehman IR, Simms ES, Kornberg A (1958) Enzymatic synthesis of deoxyribonucleic acid. II. General properties of the reaction. J Biol Chem 233(1):171–177

Bielas JH, Loeb KR, Rubin BP, True LD, Loeb LA (2006) Human cancers express a mutator phenotype. Proc Natl Acad Sci USA 103(48):18238–18242

Blanco L, Bernad A, Lazaro JM, Martin G, Garmendia C, Salas M (1989) Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. J Biol Chem 264(15):8935–8940

Bloom LB (2009) Loading clamps for DNA replication and repair. DNA Repair (Amst) 8(5):570–578

Braithwaite DK, Ito J (1993) Compilation, alignment, and phylogenetic relationships of DNA polymerases. Nucleic Acids Res 21(4):787–802

Burgers PM, Koonin EV, Bruford E, Blanco L, Burtis KC, Christman MF, Copeland WC, Friedberg EC, Hanaoka F, Hinkle DC, Lawrence CW, Nakanishi M, Ohmori H, Prakash L, Prakash S, Reynaud CA, Sugino A, Todo T, Wang Z, Weill JC, Woodgate R (2001) Eukaryotic DNA polymerases: proposal for a revised nomenclature. J Biol Chem 276(47):43487–43490

Dahlberg ME, Benkovic SJ (1991) Kinetic mechanism of DNA polymerase I (Klenow fragment): identification of a second conformational change and evaluation of the internal equilibrium constant. Biochemistry 30(20):4835–4843

Delucia AM, Chaudhuri S, Potapova O, Grindley ND, Joyce CM (2006) The properties of steric gate mutants reveal different constraints within the active sites of Y-family and A-family DNA polymerases. J Biol Chem 281(37):27286–27291

Goellner EM, Svilar D, Almeida KH, Sobol RW (2012) Targeting DNA polymerase as for therapeutic intervention. Curr Mol Pharmacol 5(1):68–87

Hirata A, Murakami KS (2009) Archaeal RNA polymerase. Curr Opin Struct Biol 19(6):724–731

Huang RC, Maheshwari N, Bonner J (1960) Enzymatic synthesis of RNA. Biochem Biophys Res Com 3(6):689–694

Hurwitz J, Bresler A, Diringer R (1960) The enzymic incorporation of ribonucleotides into polyribonucleotides and the effect of DNA. Biochem Biophys Res Com 3(1):15–19

Ishino Y, Komori K, Cann IKO, Koga Y (1998) A novel DNA polymerase family found in Archaea. J Bacteriol 180(8):2232–2236

Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol 3:318–356

Johnson KA (2010) The kinetic and chemical mechanism of high-fidelity DNA polymerases. Biochim Biophys Acta 1804(5):1041–1048

Johnson RE, Washington MT, Prakash S, Prakash L (2000) Fidelity of human DNA polymerase eta. J Biol Chem 275(11):7447–7450

Joyce CM, Potapova O, Delucia AM, Huang X, Basu VP, Grindley ND (2008) Fingers-closing and other rapid conformational changes in DNA polymerase I (Klenow fragment) and their role in nucleotide selectivity. Biochemistry 47(23):6103–6116

Jun SH, Reichlen MJ, Tajiri M, Murakami KS (2011) Archaeal RNA polymerase and transcription regulation. Crit Rev Biochem Mol Biol 46(1):27–40

Kamtekar S, Berman AJ, Wang J, Lazaro JM, de Vega M, Blanco L, Salas M, Steitz TA (2006) The phi29 DNA polymerase: protein-primer structure suggests a model for the initiation to elongation transition. EMBO J 25(6):1335–1343

Kornberg A (1957) Enzymatic synthesis of deoxyribonucleic acid. Harvey Lect 53:83–112

Kuchta RD, Benkovic P, Benkovic SJ (1988) Kinetic mechanism whereby DNA polymerase I (Klenow) replicates DNA with high fidelity. Biochemistry 27(18):6716–6725

Kunkel TA (2004) DNA replication fidelity. J Biol Chem 279(17):16895–16898

Lange SS, Takata K, Wood RD (2011) DNA polymerases and cancer. Nat Rev Cancer 11(2):96–110

Lehman IR, Bessman MJ, Simms ES, Kornberg A (1958) Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from Escherichia coli. J Biol Chem 233(1):163–170

Lemee F, Bergoglio V, Fernandez-Vidal A, Machado-Silva A, Pillaire MJ, Bieth A, Gentil C, Baker L, Martin AL, Leduc C, Lam E, Magdeleine E, Filleron T, Oumouhou N, Kaina B, Seki M, Grimal F, Lacroix-Triki M, Thompson A, Roche H, Bourdon JC, Wood RD, Hoffmann JS, Cazaux C (2010) DNA polymerase theta up-regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability. Proc Natl Acad Sci USA 107(30):13390–13395

Masutani C, Araki M, Yamada A, Kusumoto R, Nogimori T, Maekawa T, Iwai S, Hanaoka F (1999) *Xeroderma pigmentosum* variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. EMBO J 18(12):3491–3501

Moldovan GL, Pfander B, Jentsch S (2007) PCNA, the maestro of the replication fork. Cell 129 (4):665–679

Nakamura T, Zhao Y, Yamagata Y, Hua YJ, Yang W (2012) Watching DNA polymerase eta make a phosphodiester bond. Nature 487(7406):196–201

Ohmori H, Friedberg EC, Fuchs RP, Goodman MF, Hanaoka F, Hinkle D, Kunkel TA, Lawrence CW, Livneh Z, Nohmi T, Prakash L, Prakash S, Todo T, Walker GC, Wang Z, Woodgate R (2001) The Y-family of DNA polymerases. Mol Cell 8(1):7–8

Ollis DL, Brick P, Hamlin R, Xuong NG, Steitz TA (1985) Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. Nature 313(6005):762–766

Pachlopnik Schmid J, Lemoine R, Nehme N, Cormier-Daire V, Revy P, Debeurme F, Debre M, Nitschke P, Bole-Feysot C, Legeai-Mallet L, Lim A, de Villartay JP, Picard C, Durandy A, Fischer A, de Saint BG (2012) Polymerase epsilon1 mutation in a human syndrome with facial dysmorphism, immunodeficiency, livedo, and short stature ("FILS syndrome"). J Exp Med 209(13):2323–2330

Starcevic D, Dalal S, Sweasy JB (2004) Is there a link between DNA polymerase beta and cancer? Cell Cycle 3(8):998–1001

Steitz TA (1993) DNA-dependent and RNA-dependent DNA-polymerases. Curr Opin Struct Biol 3(1):31–38

Steitz TA, Steitz JA (1993) A general two-metal-ion mechanism for catalytic RNA. Proc Natl Acad Sci USA 90(14):6498–6502

Stevens A (1960) Incorporation of the adenine ribonucleotide into RNA by cell fractions from, E. coli B. Biochem Biophys Res Com 3(1):92–96

Tan XH, Zhao M, Pan KF, Dong Y, Dong B, Feng GJ, Jia G, Lu YY (2005) Frequent mutation related with overexpression of DNA polymerase beta in primary tumors and precancerous lesions of human stomach. Cancer Lett 220(1):101–114

Trakselis MA, Benkovic SJ (2001) Intricacies in ATP-dependent clamp loading: variations across replication systems. Structure 9(11):999–1004

Watson JD, Crick FH (1953) Molecular structure of nucleic acids. Nature 171:737–738

Weiss SB, Gladstone L (1959) A mammalian system for the incorporation of cytidine triphosphate into ribonucleic acid. J Am Chem Soc 81(15):4118–4119

Weissbach A, Baltimore D, Bollum F, Gallo R, Korn D (1975) Nomenclature of eukaryotic DNA polymerases. Science 190(4212):401–402

Werner F, Grohmann D (2011) Evolution of multisubunit RNA polymerases in the three domains of life. Nat Rev Microbiol 9(2):85–98

Chapter 2
Eukaryotic Replicative DNA Polymerases

Erin Walsh and Kristin A. Eckert

Abstract DNA replication is a dynamic process that requires the precise coordination of numerous cellular proteins. At the core of replication in eukaryotic cells are three DNA polymerases, Pol α, Pol δ, and Pol ε, which function cooperatively to ensure efficient and high-fidelity genome replication. These enzymes are members of the B family of DNA polymerases, characterized by conserved amino acid motifs within the polymerase active sites. Pol α is a DNA polymerase of moderate fidelity that lacks 3′→5′ exonuclease activity, while Pols δ and ε are processive, high-fidelity polymerases with functional 3′→5′ exonuclease activities. Each polymerase exists as a holoenzyme complex of a large polymerase catalytic subunit and several smaller subunits. The Pol α holoenzyme possesses primase activity, which is required for de novo synthesis of RNA–DNA primers at replication origins and at each new Okazaki fragment. In one model of eukaryotic DNA replication, Pol ε functions in leading strand DNA synthesis, while Pol δ functions primarily in lagging strand synthesis. This chapter discusses the biochemical properties of eukaryotic replicative polymerases and how biochemical properties shape their functional roles in replication initiation, replication fork elongation, and the checkpoint responses.

Keywords DNA replication fork • S phase checkpoint • DNA polymerase fidelity • primase • proofreading exonuclease • replisome • genome stability

E. Walsh • K.A. Eckert (✉)
Department of Pathology, Jake Gittlen Cancer Research Foundation, Pennsylvania State University College of Medicine, Hershey, PA 17033, USA
e-mail: kae4@psu.edu

K.S. Murakami and M.A. Trakselis (eds.), *Nucleic Acid Polymerases*, Nucleic Acids and Molecular Biology 30, DOI 10.1007/978-3-642-39796-7_2,

2.1 Biochemical Properties of Replicative Polymerases

2.1.1 Polymerase α-Primase

2.1.1.1 Overview

A formal nomenclature for eukaryotic DNA polymerases has been adopted, in which cellular DNA polymerases are given Greek letters in order of their historical discovery (Weissbach et al. 1975; Burgers et al. 1990). Accordingly, Pol α-primase was the first mammalian polymerase to be purified and characterized (Yoneda and Bollum 1965). Pol α-primase holoenzyme is a heterotetrameric protein complex in all eukaryotes studied (Table 2.1) (see Muzi-Falconi et al. 2003 for review). The p180 large subunit contains the DNA polymerase active site (Plevani et al. 1985; Wong et al. 1986). The B subunit has no known enzymatic activity, but performs a regulatory function, possibly linking the pol α holoenzyme to components of the replication fork (Collins et al. 1993). DNA polymerases lack the ability to carry out de novo DNA synthesis and, in eukaryotes, require a 3′-OH provided by an RNA primer in order to initiate DNA synthesis. Primase activity is contained within the tightly associated p49/p58 complex (Plevani et al. 1985; Nasheuer and Grosse 1988). The architecture and subunit arrangement of the *Saccharomyces cerevisiae* (*S. c.*) Pol α-primase holoenzyme have been examined by X-ray crystallography and electron microscopy. The B subunit is tethered to the p180 subunit through a structured, flexible linker (Klinge et al. 2009). The four-subunit holoenzyme exists as a dumbbell-shaped particle, with the catalytic primase and polymerase active sites present in distinct lobes of the complex, separated by ~100 Å (Nunez-Ramirez et al. 2011).

2.1.1.2 An Essential Polymerase and Primase

Pol α polymerase activity is essential for chromosomal replication. Genetic analyses of mutants in *S. c.* have demonstrated that both the catalytic and regulatory B subunits are required in vivo for viability, and mutants in either gene exhibit defects in DNA replication and progression through S-phase (Johnson et al. 1985; Foiani et al. 1994; Budd and Campbell 1987). In mammalian cells, Pol α-primase neutralizing antibodies inhibit DNA synthesis (Miller et al. 1985; Kaczmarek et al. 1986). Thermosensitive yeast *POL1* mutants display elevated genetic instability due to defects in replication (Gutiérrez and Wang 2003; Liu et al. 1999). The primase activity is also essential in eukaryotic cells. Disruption of either the *PRI1* or *PRI2* genes is lethal in *S. c.* due to replication defects (Foiani et al. 1989; Lucchini et al. 1987). Characterization of conditional *PRI1* and *PRI2* mutants demonstrates that loss of primase activity causes increased mitotic recombination and spontaneous mutation rates, possibly due to defects in replication and impaired meiosis (Longhese et al. 1993).

Table 2.1 Composition of the major replicative DNA polymerases in eukaryotes

		Designation (gene/protein)		
Subunit	Function	Human	*S. cerevisiae*	*S. pombe*
Pol α holoenzyme				
A	Polymerase	*POLA1*/p180	*POL1*	*pol1*
B	Regulatory	*POLA2*/p70	*POL12*	*pol12*
	Primase regulatory	*PRIM2*/p58	*PRI2*	*pri2*
	Primase catalytic	*PRIM1*/p49	*PRI1*	*pri1*
Pol δ holoenzyme				
A	Polymerase, 3′→5′ exonuclease	*POLD1*/p125	*POL3*	*pol3*
B	Regulatory	*POLD2*/p50	*POL31*	*cdc1*
C	Regulatory	*POLD3*/p68[a]	*POL32*	*cdc27*
D	Regulatory (DNA damage)	*POLD4*/p12	–	*cdm1*
Pol ε holoenzyme				
A	Polymerase, 3′→ 5′ exo	*POLE1*/p261	*POL2*	*cdc20*
B	Regulatory	*POLE2*/p59	*DPB2*	*dpb2*
C	Double-stranded DNA binding	*POLE3*/p17	*DPB3*	*dpb3*
D	Double-stranded DNA binding	*POLE4*/p12	*DPB4*	*dpb4*

[a]Also referred to as the p66 subunit

2.1.1.3 Coordinated Primase and Polymerase Activities

In eukaryotic cells, in vivo studies have demonstrated that RNA primers of Okazaki fragments are attached to short DNA chains (Kitani et al. 1984). The Pol α holoenzyme is a unique replicative protein complex, possessing two coupled catalytic activities (Hu et al. 1984). Primase initiates synthesis of RNA at polypyrimidine tracts within the DNA template, preferentially within a T-rich region, and this occurs at many sites along the DNA template (Yamaguchi et al. 1985; Kitani et al. 1984). The primase catalytic site maps to the p49 subunit (Nasheuer and Grosse 1988), in which a carboxylic triad is essential for its function in generating RNA primers (Copeland and Tan 1995). The *S. c.* Pol α/p49 subunit is sufficient for synthesis of RNA, but highly inefficient without the p58 subunit (Santocanale et al. 1993). Biochemically, p58 is necessary for synthesis of the first dinucleotide of the RNA primer and modulates the rate and stability of subsequent extension steps (Copeland and Wang 1993). GTP is preferentially incorporated into the first dinucleotide, possibly to stabilize the short primer (Sheaff and Kuchta 1993). Although RNA primers generated by purified primase can vary in length from ~2 to 10 nucleotides in the presence of dNTPs (Nasheuer and Grosse 1988), only those $\geq$7 nucleotides are extended by Pol α, and formation of an RNA primer of this length terminates primase activity (Kuchta et al. 1990). The p58 subunit regulates primase processivity and ensures the correct RNA primer length is synthesized (Zerbe and Kuchta 2002). The DNA polymerase activity of Pol α extends RNA primers with ~20–30 dNTPs to generate an RNA–DNA hybrid molecule (Thompson et al. 1995).

Mechanistically, the primase and polymerase activities of the Pol α holoenzyme are tightly coupled to ensure that new primers are not synthesized until the previous primer has been extended by polymerase (Sheaff et al. 1994). Functional coordination of the two activities is dependent upon the presence of dNTPs (Hu et al. 1984). An intramolecular transfer of the RNA primer from the primase active site to the polymerase active site occurs rapidly when dNTPs are present (Sheaff et al. 1994). Structurally, flexible tethering of Pol α holoenzyme lobes containing the primase and polymerase centers increases the efficiency of primer transfer between the two active sites (Nunez-Ramirez et al. 2011). Structural and biochemical studies suggest that the p58 subunit is involved in the switch from primase to polymerase through conformational changes (Arezi et al. 1999; Agarkar et al. 2011). Consistent with its role in synthesizing short RNA–DNA primers, the Pol α holoenzyme lacks the high processivity possessed by the other major replicative holoenzymes (see below), although it does possess the capacity to carry out robust DNA synthesis in vitro. The rate of DNA synthesis by calf thymus Pol α-primase is similar to that of Pols ε and δ (Weiser et al. 1991), and Pol α-primase displays an in vitro processivity of ~20–100 nucleotides (Hohn and Grosse 1987). Pol α-primase interacts with the p70 subunit of replication protein A (RPA), which stimulates both the polymerase activity and processivity of the holoenzyme (Braun et al. 1997), possibly by increasing the affinity for primer termini (Maga et al. 2001).

2.1.1.4 A Moderately Accurate Polymerase

The in vitro fidelity of the mammalian Pol α holoenzyme purified from multiple sources has been determined using several genetic assays (Kunkel and Loeb 1981; Kunkel 1985; Eckert et al. 1997). The majority of errors created by Pol α are single-base substitutions, followed by one-base deletion errors (Kunkel 1985; Eckert et al. 1997). The purified *S. c.* Pol α/p180 catalytic subunit and the *S. c.* Pol α holoenzyme have similar overall fidelities but display statistically significant differences in error rates within specific sequence contexts (Kunkel et al. 1989). The *S. c.* Pol α/p180 subunit exhibits an error rate of 1/9,900 and 1/12,000 per nucleotide synthesized, for single-base substitutions and deletions, respectively (Kunkel et al. 1991). The human Pol α/p180 subunit base substitution error rate can be as low as 1/42,000 nucleotides in low-pH buffer conditions (Eckert and Kunkel 1993). Direct comparison of purified mammalian Pol α-primase, Pol δ, and Pol ε preparations demonstrated that Pol α possesses the lowest fidelity of the three major eukaryotic replicative polymerases (Thomas et al. 1991).

One mechanism by which Pol α maintains its moderate fidelity is the low rate of extending mismatched DNA primers (Perrino and Loeb 1989). Some calf thymus Pol α holoenzyme pause sites correspond to sites of increased mismatched 3′-primer termini (Fry and Loeb 1992), consistent with a low rate of mispair extension. Pol α holoenzyme pausing within microsatellite sequences also is correlated with the rate of misalignment-based errors (Hile and Eckert 2004).

Replication auxiliary factors enhance the DNA fidelity of Pol α-primase. RPA decreases dNTP misincorporation rate by decreasing the affinity of the Pol α holoenzyme to mismatches while increasing the affinity to matched DNA primer termini, suggesting that it may function as a "fidelity clamp" that modulates the Pol α holoenzyme for faithful DNA synthesis (Maga et al. 2001). Terminal misincorporation events at Pol α pause sites are also avoided in the presence of RPA (Suzuki et al. 1994). Pol α-primase lacks 3′→5′ exonuclease activity due to amino acid changes at catalytic residues within the domain, which limits its intrinsic fidelity (Pavlov et al. 2006b). However, *S. c.* Pol δ can perform intermolecular proofreading of errors made by Pol α-primase (Pavlov et al. 2006a). Based on an estimation that ~4–8 × 10^4 RNA–DNA primers are synthesized by Pol α-primase during replication in human cells (Muzi-Falconi et al. 2003), such mechanisms of enhanced fidelity may be essential in maintaining genome stability.

In contrast to its DNA polymerase fidelity, the priming activity exhibits very low fidelity during RNA synthesis. Primase readily misincorporates NTPs during in vitro RNA synthesis, often polymerizing primers containing consecutive mismatches (Sheaff and Kuchta 1994). Importantly, after intramolecular transfer, such mismatched RNA primers are readily extended by the DNA polymerase activity in the presence of dNTP substrates.

2.1.1.5 Posttranslational Regulation

The Pol α holoenzyme exists as an assembled complex throughout the cell cycle (Ferrari et al. 1996), and its activity is regulated by cyclin-dependent kinases in a cell cycle-dependent manner (Voitenleitner et al. 1999). The p180 subunit is a phosphoprotein that becomes hyperphosphorylated in G_2/M phase, while the B subunit is phosphorylated only in G_2/M (Nasheuer et al. 1991). Pol α phosphorylation results in lowered single-stranded DNA binding affinity, lowered DNA synthesis activity, and an inhibition of DNA replication (Nasheuer et al. 1991; Voitenleitner et al. 1999).

2.1.2 DNA Polymerase δ

2.1.2.1 Overview

Pol δ was originally identified as a novel DNA polymerase purified from rabbit bone marrow that possessed a very active 3′→5′ exonuclease activity (Byrnes et al. 1976). Subsequently, *S. c.* POL3 was identified as the yeast homolog of Pol δ (Boulet et al. 1989). In all eukaryotes, Pol δ is a multimeric complex, wherein the largest subunit harbors the DNA polymerase and 3′→5′ exonuclease catalytic domains, and is tightly associated with a regulatory B subunit (Table 2.1). Mammalian Pol δ holoenzyme is a heterotetrameric protein consisting of two additional

C and D subunits, p68 and p12. The fission yeast *Schizosaccharomyces pombe* (*S. p.*) also encodes a heterotetrameric Pol δ, possessing the Cdm1 subunit which shows some sequence similarity to the mammalian p12 subunit. The budding yeast *S. c.* Pol δ is a heterotrimer of the A and B subunits together with a C subunit ortholog, Pol32.

The structure of a truncated form of the *S. c.* Pol δ/p125 subunit in ternary complex with DNA and dNTP substrates has been solved at 2 Å resolution (Swan et al. 2009). The polymerase domain has the characteristic right-hand structure of palm, fingers, and thumb subdomains. The 3′→5′ exonuclease domain is separated from the polymerase domain by 45 Å. Pol δ also possesses a novel N-terminal domain that interacts with 10–20 nucleotides of the ssDNA template, upstream of the polymerase active site. The structure of a truncated *S. c.* Pol δ holoenzyme as determined by small-angle X-ray scattering analysis consists of a globular catalytic core (A subunit) flexibly linked to an elongated tail comprised of the B and C subunits (Jain et al. 2009). The interaction of Pol3 (A subunit) with the Pol31 and Pol32 (B and C subunits) is stabilized by binding of a 4Fe-4S metal cofactor cluster to four conserved cysteine residues in the CysB motif of the Pol3 C-terminal domain (Netz et al. 2012).

2.1.2.2 An Essential Polymerase

Evidence that Pol δ is an essential component of the eukaryotic replication machinery has been derived from several models. Disruption of Pol δ is lethal in *S. p.*, and thermosensitive mutants of pol3 arrest in S-phase of the cell cycle (Francesconi et al. 1993). Deletion of the *S. p.* cdc27 gene (Pol δ/C subunit) results in cell cycle arrest (MacNeill et al. 1996). Although the orthologous *S. c POL*32 gene is not essential, deletion mutants do display abnormal phenotypes (Gerik et al. 1998). *S. c. POL*3 mutants harboring different substitutions at a catalytically essential residue, Leu^{612}, exhibit a wide range of phenotypic deficiencies (Venkatesan et al. 2006). Only 8 of 19 mutants are viable and display varying degrees of genotoxic sensitivity, cell cycle defects, and morphological abnormalities (Venkatesan et al. 2006). Replication studies in *Xenopus* egg extracts demonstrated that immunodepletion of Pol δ leads to a significant decrease in DNA synthesis and the accumulation of unreplicated, single-stranded DNA gaps (Fukui et al. 2004). Knockdown of Pol δ/p125 in HeLa cells causes an accumulation of cells in early, middle, and late S-phase, and cells do not enter mitosis (Bermudez et al. 2011). Homozygous disruption of *Pold1* in mice was shown to cause embryonic lethality between E4.5 and E7.5, due to significant defects in DNA synthesis (Uchimura et al. 2009). Profoundly, homozygous mutations of highly conserved amino acids within motif A of the Pol δ active site (L604G and L604K) are embryonic lethal in mice, and mice heterozygote for the L604K mutant exhibit reduced life spans and accelerated tumorigenesis (Venkatesan et al. 2007). Mouse embryonic fibroblasts heterozygous for these mutant proteins display elevated mutation rates and chromosomal instability relative to wild-type cells (Venkatesan et al. 2007). The severe

replication defects associated with loss of Pol δ in each of these models were vital in uncovering the polymerase's role as a major component of the eukaryotic replication fork.

2.1.2.3 Efficient DNA Synthesis in the Presence of Replication Accessory Factors

Pol δ requires replication accessory proteins to achieve its greatest efficiency. Shortly after the discovery of Pol δ, an auxiliary protein was identified, which co-eluted with Pol δ during purification and significantly enhanced its in vitro DNA synthesis on several DNA templates (Tan et al. 1986). This protein was later identified as proliferating cell nuclear antigen (PCNA) and was shown to greatly enhance calf thymus Pol δ processivity in vitro (Prelich et al. 1987; Bravo et al. 1987). PCNA is now known as the eukaryotic sliding clamp protein, a homotrimer ring-shaped protein that encircles DNA and tethers replication proteins, allowing movement along the DNA template (Moldovan et al. 2007). Replication Factor C (RFC), a heteropentameric complex, is responsible for loading PCNA onto DNA through an ATP-dependent mechanism. In the presence of PCNA, RFC, and RPA, *S. c.* Pol δ processivity is enhanced from 6 nucleotides to >600 nucleotides (Chilkova et al. 2007). Human Pol δ holoenzyme activity increases >50-fold and processivity is stimulated in the presence of PCNA (Xie et al. 2002). However, the human Pol δ holoenzyme differs biochemically from *S. c.* Pol δ holoenzyme, as human Pol δ dissociates more readily from DNA templates, even in the presence of accessory factors (Bermudez et al. 2011).

Recombinant human Pol δ can be purified in several subassemblies, all of which retain DNA synthesis activity (Zhou et al. 2012; Podust et al. 2002). The Pol δ p125/p50 heterodimer (also referred to as the core dimer) is a tightly associated complex with low specific activity. The three subunit complex, Pol δ3 (core + p68 subunit), displays high DNA synthesis activity but is unstable in vitro. The p68 subunit is essential for synthesis of long DNA products by Pol δ. The p12 subunit increases stability of the holoenzyme and increases DNA synthesis activity. All four subunits of human Pol δ individually interact with PCNA, which may allow for flexibility during DNA replication, as many proteins functionally interact with PCNA during Okazaki fragment maturation (Zhang et al. 1999; Wang et al. 2011; Li et al. 2006b). PCNA interacting motifs also were identified within the C-termini of all three subunits of *S. c.* Pol δ, which are needed for efficient DNA synthesis (Acharya et al. 2011). PCNA stimulation of DNA synthesis activity differs quantitatively among the various human Pol δ subassemblies, in the order Pol δ4 > Pol δ3 (core + p68) > Pol δ2 core (Zhou et al. 2012). Kinetically, PCNA reduces the K_m for DNA template binding and increases the V_{max} of the calf thymus Pol δ catalytic core, suggesting that PCNA might stimulate Pol δ processivity by increasing its residence time on the DNA template and the rate of nucleotide incorporation (Ng et al. 1991).

2.1.2.4 An Accurate Polymerase

Replication of the genome requires accurate DNA synthesis in order to avoid the accumulation of deleterious mutations. As demonstrated using in vitro mutation assays, purified Pol δ generally possesses a high DNA synthesis fidelity (see Prindle and Loeb 2012 for review). The *S. c.* Pol δ holoenzyme incorporates less than one-base substitution error out of 80,000 nucleotides synthesized (Fortune et al. 2005). For human Pol δ, less than one-base misinsertion error occurred per 220,000 nucleotides (Schmitt et al. 2009). However, *S. c.* and human Pol δ holoenzymes do not exhibit high fidelity during in vitro synthesis of repetitive microsatellite DNA sequences (Hile et al. 2012; Abdulovic et al. 2011). At the heart of Pol δ's high fidelity is its intrinsic 3′→5′ exonuclease activity (Simon et al. 1991), which enables proofreading upon incorporation of incorrect nucleotides during DNA synthesis (Kunkel et al. 1987). Both human and *S. c.* exonuclease-deficient Pol δ forms exhibit approximately 10-fold higher base substitution error rates than wild-type Pol δ in vitro (Fortune et al. 2005; Schmitt et al. 2009). In vivo, yeast strains carrying inactivating mutations within the *POL3* exonuclease domain exhibit a 100-fold increased spontaneous mutation rate (Morrison and Sugino 1994). Furthermore, mice with homozygous substitutions at highly conserved residues within the proofreading domain of Pol δ exhibit a high incidence of cancer and decreased survival (Goldsby et al. 2002; Albertson et al. 2009).

Intrinsic kinetic properties of Pol δ also are key determinants of its high fidelity. *S. c.* Pol δ strongly favors incorporation of correct versus incorrect nucleotides during synthesis (Dieckman et al. 2010). A highly conserved leucine residue within the Pol δ active site has been extensively studied in yeast, mice, and humans (Leu-612, 604 and 606, respectively) for its role in ensuring high-fidelity DNA synthesis. Amino acid substitutions at this site cause allele-specific phenotypic effects in *S. c.* and mice (Venkatesan et al. 2006, 2007). In vitro, the human Pol δ holoenzyme L606G mutant is highly error prone (Schmitt et al. 2010). Interestingly, the L606K mutant exhibits higher fidelity than wild-type Pol δ but decreased bypass of DNA adducts (Schmitt et al. 2010), suggesting that the high fidelity conferred by this active site residue is balanced by the need for the wild-type enzyme to perform other activities, including translesion synthesis at the replication fork.

Efficient proofreading requires partitioning of DNA substrates containing 3′ terminal mispairs from the polymerase to the exonuclease active sites (Khare and Eckert 2002). An *S. c. POL3* active site mutation that impairs this partitioning results in decreased DNA synthesis fidelity and increased spontaneous mutation rate (McElhinny et al. 2007). Purified *S. c.* Pol δ interacts with base pairs distant from the templating base, which may allow the polymerase to "sense" base mismatches (Swan et al. 2009).

Although PCNA stimulates Pol δ activity, it may decrease its fidelity. In vitro, PCNA increases the rate of nucleotide misincorporation by *S. c.* Pol δ, resulting in a

significant reduction in fidelity (Hashimoto et al. 2003). Similarly, nucleotide misincorporations by calf thymus Pol δ increase ~27-fold in the presence of PCNA, and PCNA may enable Pol δ to extend mismatched base pairs by stabilizing the Pol δ-template-primer complex (Mozzherin et al. 1996). However, addition of both PCNA and RPA had no effect on *S. c.* Pol δ base substitution error rates, and the addition of PCNA and RPA to Pol δ DNA synthesis reactions contributed to a decreased rate of large deletion errors within directly repeated sequences (Fortune et al. 2006). Further studies are needed to fully understand the effect of accessory factors on Pol δ fidelity. Interestingly, the Werner syndrome protein (WRN), a RecQ family helicase and 3′→5′ exonuclease, can enable high-fidelity DNA synthesis by Pol δ through excision of primer-template mismatches prior to polymerase extension (Kamath-Loeb et al. 2012).

2.1.2.5 Posttranslational Regulation

Phosphorylation may represent an important mechanism of Pol δ regulation (see Lee et al. 2012 for review). Mammalian Pol δ protein levels peak at the G1/S border, and the p125 subunit is most actively phosphorylated during S-phase (Zeng et al. 1994). The B subunit (p50) is also phosphorylated in vivo and is an in vitro substrate of the cyclin A-Cdk2 cell cycle-dependent kinase (Li et al. 2006a). The C subunit (p68) can be phosphorylated by cyclin-dependent kinases (CycE-Cdk2, CycACdk1, or CycA-Cdk2) in vitro, and PCNA interferes with this phosphorylation (Ducoux et al. 2001). In addition, mammalian Pol δ/p125, p68, and p12 subunits can be phosphorylated by casein kinase 2 in vitro and subsequently dephosphorylated by protein phosphatase-1 (Gao et al. 2008), suggesting an additional regulatory circuit for regulation. Phosphorylation of Pol δ/p68 coincides with Pol δ association with chromatin at the start of S-phase (Lemmens et al. 2008). The Pol δ/p68 subunit also contains a phosphorylation site for protein kinase A, and phosphomimetic mutation of this residue decreases Pol δ affinity for PCNA and processivity (Rahmeh et al. 2011). Thus, phosphorylation may serve to regulate Pol δ activity by controlling its interaction with DNA and/or auxiliary proteins during replication.

2.1.3 DNA Polymerase ε

2.1.3.1 Overview

Purified Pol ε was first characterized as a larger form of Pol δ that was highly processive, but not stimulated by PCNA in vitro (Syvaoja and Linn 1989). The Pol ε holoenzyme exists as a heterotetrameric protein in all eukaryotes studied (Table 2.1) (Pursell et al. 2008). The large catalytic subunit contains the polymerase and 3′→5′ exonuclease active sites within the N-terminus, and binding domains for smaller

subunits and PCNA within the C-terminus. The C and D subunits form a subcomplex that is important for double-stranded DNA binding (Tsubota et al. 2003). The *S. c.* Pol ε holoenzyme structure has been solved at 20 Å resolution using cryo-electron microscopy and single-particle analyses (Asturias et al. 2006). The three small subunits (B–D) form a discrete extended tail structure, separated from the large catalytic subunit by a flexible hinge. Similar to the *S. c.* Pol δ, a 4Fe-4S cluster is bound to four conserved cysteine residues (the CysB motif) within the C-terminal domain of Pol2 and may be essential for stabilizing the *S. c.* Pol ε holoenzyme (Netz et al. 2012).

2.1.3.2 An Essential Component of the Eukaryotic Replication Machinery

The Pol ε holoenzyme is essential for chromosomal replication. In *S. c.*, *POL2* reading frame disruptions are nonviable due to arrest of DNA replication; however, truncation mutations that maintain catalytic activity are viable with a slow growth phenotype (Morrison et al. 1990). Deletion of the *POL2* N-terminus, which encodes the polymerase catalytic domain, also is viable, although mutants display severe replication defects, including slow fork movement, prolonged S-phase, and shortened telomeres (Kesti et al. 1999; Feng and D'Urso 2001; Dua et al. 1999). Thermosensitive yeast mutants also demonstrate that nonfunctional Pol ε leads to defective chromosomal replication and the accumulation of short DNA fragments (Araki et al. 1992; Budd and Campbell 1993). Immunodepletion of Pol ε from *Xenopus* egg extracts significantly impedes elongation of nascent DNA strands and causes the accumulation of short replication intermediates (Waga et al. 2001). In HeLa cells, transient knockdown of Pol ε causes an accumulation of cells in early S-phase and a decreased rate of replication fork movement (Bermudez et al. 2011). Disruption of Pol ε in mice results in embryonic lethality (Menezes and Sweasy 2012).

2.1.3.3 A Highly Efficient and Processive Polymerase

Pol ε and Pol δ differ in their biochemical interactions with PCNA. PCNA binds the Pol ε p261 (N-terminus), p59, and p12 subunits and increases the rate of nucleotide incorporation by the holoenzyme (Bermudez et al. 2011). In the presence of PCNA, RFC, and RPA, *S. c.* Pol ε processivity is stimulated ~6-fold, less than the ~100-fold stimulation observed for Pol δ (Chilkova et al. 2007). *S. c.* Pol ε has a high affinity for DNA and low affinity for PCNA, while *S. c.* Pol δ displays the opposing affinities for each, suggesting that Pol ε might load onto DNA independently of PCNA, while Pol δ requires preloading of PCNA (Chilkova et al. 2007). However, *S. c.* and human Pol ε are more processive than Pol δ, even in the absence of accessory factors (Dua et al. 2002) (Bermudez et al. 2011). High *S. c.* Pol ε processivity requires a minimal primer duplex stem length of 40 base pairs, and

structural studies suggest that the tail domain formed by the B–D subunits contributes directly to Pol ε processivity by binding double-stranded DNA, precluding the need for PCNA to enhance DNA affinity (Asturias et al. 2006). The processivities of both the *S. c.* Pol ε polymerase and exonuclease activities are reduced in subassemblies lacking the C and D subunits, relative to the holoenzyme form (Aksenova et al. 2010).

2.1.3.4 A Highly Accurate Polymerase

Pol ε is perhaps the most accurate eukaryotic DNA polymerase. Purified calf thymus Pol ε base substitution error rates determined in vitro ranged from 1/30,000 to 1/400,000 mutations per nucleotide synthesized, lower than either Pol δ or Pol α (Thomas et al. 1991). The *S. c.* Pol ε holoenzyme exhibits very low base substitution and single-base deletion error rates, $\leq 2 \times 10^{-5}$ and $\leq 5 \times 10^{-7}$ (Shcherbakova et al. 2003), respectively. Error rates of human Pol ε have not yet been determined because they are close to background rates for in vitro mutation assays (Korona et al. 2011). Although the in vitro fidelity of purified *S. c.* Pol ε lacking the C and D subunits is the same as the Pol ε holoenzyme form, deletion of *DPB3* and *DPB4* results in a slightly increased spontaneous mutation rate in vivo (Aksenova et al. 2010).

The high fidelity of Pol ε is due, in part, to its intrinsic 3′→5′ exonuclease activity. Amino acid substitutions of a conserved motif within the exonuclease active site result in a ~20-fold increase in the *S. c.* spontaneous mutation rate (Morrison et al. 1991). Purified exonuclease-deficient *S. c.* Pol ε exhibits single-base substitution and deletion errors rates that are ~10- and 100-fold higher, respectively, than wild-type Pol ε (Shcherbakova et al. 2003). Kinetically, *S. c.* Pol ε mutants harboring a C1089Y substitution within the polymerase active site exhibit an increased rate of base misincorporation that may result from the inability to perform proper DNA shuffling between the polymerase and exonuclease domains (Shimizu et al. 2002). Human exonuclease-deficient Pol ε is ~5-fold more accurate than wild-type *S. c.* Pol ε for both base substitutions and single-base deletions (Korona et al. 2011). Homozygous loss of Pol ε exonuclease activity in mice results in an elevated spontaneous mutation rate and tumor incidence (Albertson et al. 2009). Specific polymerase active site residues also play a role in determining Pol ε fidelity. For instance, an M644F substitution within the *S. c.* Pol ε active site reduces fidelity due to an increased base misincorporation rate (Pursell et al. 2007a). The Pol ε M644 active site residue may modulate fidelity by maintaining proper geometry of the substrate binding pocket (Pursell et al. 2007a).

2.1.3.5 Posttranslational Modification

The posttranslational regulation of Pol ε has not been extensively studied. However, the *S. c.* B subunit, required for chromosomal replication (Araki et al. 1991), is