Polymerase Dynamics at the Eukaryotic DNA Replication Fork*

Published, JBC Papers in Press, October 3, 2008, DOI 10.1074/jbc.R800062200
Peter M. J. Burgers†
From the Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110

This review discusses recent insights in the roles of DNA polymerases (Pol) δ and ε in eukaryotic DNA replication. A growing body of evidence specifies Pol ε as the leading strand DNA polymerase and Pol δ as the lagging strand polymerase during undisturbed DNA replication. New evidence supporting this model comes from the use of polymerase mutants that show an asymmetric mutator phenotype for certain mispairs, allowing an unambiguous strand assignment for these enzymes. On the lagging strand, Pol δ corrects errors made by Pol α during Okazaki fragment initiation. During Okazaki fragment maturation, the extent of strand displacement synthesis by Pol δ determines whether maturation proceeds by the short or long flap processing pathway. In the more common short flap pathway, Pol δ coordinates with the flap endonuclease FEN1 to degrade initiator RNA, whereas in the long flap pathway, RNA removal is initiated by the Dna2 nuclease/helicase.

Lessons Learned from SV40 Replication Studies

How are three DNA polymerases distributed over two strands at one single replication fork? For several decades, researchers have been faced with the enigmatic problem of assigning functions to the three major replicative DNA polymerases in the nucleus: Pol α, Pol δ, and Pol ε. It all started out much simpler. When the in vitro SV40 DNA replication system was developed in 1984 (1), replication studies of this small viral double-stranded DNA molecule promised to provide important insights in the cellular elongation machinery because only the replication initiator and the helicase functions are encoded by the viral large tumor antigen. For all other elongation factors including DNA polymerase(s), the virus depends on the host cell. Indeed, this system quickly led to the discovery of the single-stranded binding protein RPA and also implicated Pol δ and the replication clamp PCNA in SV40 DNA replication. Soon, a model emerged in which the primase activity of the Pol α-prime complex initiates DNA synthesis with an ~10-nt RNA primer that is elongated by its polymerase activity for another ~20 nt (Table 1). The loading of PCNA effects a switch to Pol δ for further elongation of an Okazaki fragment. On the leading strand, Pol δ continues elongation until all viral DNA is replicated, whereas on the lagging strand, an iterative switch between Pol α and Pol δ ensures initiation and elongation of Okazaki fragments, respectively (reviewed in Ref. 2). In this compact and efficient replication system, there was no necessity for a third DNA polymerase. The factors uncovered in the SV40 system were shown to be essential replication genes in yeast, and biochemical studies showed a strong functional similarity between the yeast and mammalian enzymes such that a unified system for the replication fork machinery could be easily developed (reviewed in Refs. 3 and 4).

However, important problems in replication fork structure and dynamics remained unsolved. In 1990, Pol ε was identified in yeast as an essential replication protein (5). As a second proofreading DNA polymerase together with Pol δ, Pol ε laid a legitimate claim to being a replicative polymerase as well (Table 1). However, it did not fit in the next SV40 package, and in vivo chromatin cross-linking studies in human cells showed that although Pol ε did cross-link to replicating chromosomal DNA, it did not cross-link to SV40 DNA (6). In addition, the completion process for DNA replication remained relatively unexplored, and no consistent view seemed to emerge with regard to the factors responsible for degradation of the initiator RNA, perhaps because of a redundancy of nuclease in cell extracts (reviewed in Ref. 7).

Division of Labor at the Fork

Because the physical placement of a DNA polymerase on a given strand has not yet been accomplished, most approaches to strand placement have been indirect, relying on genetic analyses in the yeast Saccharomyces cerevisiae. For the lagging strand machinery, strand placement of Pol δ has been inferred from genetic interactions between polymerase mutants and other lagging strand replication genes (see below). However, this approach has not worked for the leading strand because of a lack of firm knowledge about proteins that specifically occupy this strand. A second approach has relied on the analysis of mutation spectra produced by the use of mutator DNA polymerases. The obvious problem in a mutation spectrum analysis of this type is clear. If, for instance, a mutator DNA polymerase gave an increase in A-T → G-C transitions, these mutations could have originated either from misincorporation of dGMP across a template dT residue on the one strand or from misincorporation of dCMP across the template dA residue on the opposite strand. The design of a new class of polymerase mutants has overcome this problem. These active-site mutants have near normal polymerase activity, carry out proofreading, and show no replication defects in vivo. They are very modest mutators and, most important, show a strong asymmetry for reciprocal mismatches in vitro. Thus, the pol2-M644G mutant Pol ε was shown in a fidelity analysis in which single-stranded DNA templates were replicated in vitro and scored in Escherichia coli to have an increased A-T → T-A mutation rate that resulted entirely from an increased rate of template dT-dTMP misincorporation and not from the reciprocal template dA-dAMP misincorporation (8). Now, all that needed to be done...
MINIREVIEW: Polymerase Dynamics at the Replication Fork

TABLE 1
Replication fork DNA polymerases
The large subunit of each complex contains the polymerase activity and the 3’-exonuclease activity (except for Pol α); the Pri1 subunit of the Pol α-primase complex has the primase activity. Proposed replication functions are shown. For details, see Ref. 3.

<table>
<thead>
<tr>
<th>Genes and subunit sizes</th>
<th>S. cerevisiae</th>
<th>S. pombe</th>
<th>Human</th>
<th>Activity</th>
<th>Fidelity</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pol α-primase</td>
<td>POL1-p167</td>
<td>pol1-p159</td>
<td>POL1A-p166</td>
<td>Polymerase</td>
<td>$10^{-4}$–$10^{-7}$</td>
<td>Initiation of replication</td>
</tr>
<tr>
<td></td>
<td>POL12-p79</td>
<td>pol12-p66</td>
<td>POL2-p68</td>
<td>Primase</td>
<td></td>
<td>Initiation of Okazaki fragments</td>
</tr>
<tr>
<td></td>
<td>PR11-p48</td>
<td>pr11-p52</td>
<td>PRIM1-p48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PR12-p62</td>
<td>spp2-p53</td>
<td>PRIM2A-p58</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Pol δ              | POL3-p125 | pol3-p124 | POLD1-p124 | Polymerase | $10^{-6}$–$10^{-7}$ | Elongation and maturation of Okazaki fragments |
|                    | POL31-p55 | cdc1-p51 | POLD2-p51 | 3’-Exonuclease |          |          |
|                    | POL32-p40 | cdc27-p42 | POLD3-p66 |          |          |          |
|                    | cdm1-p19 |          | POLD4-p12 |          |          |          |

| Pol ε              | POL2-p256 | pol2-p253 | POLE-p261 | Polymerase | $10^{-4}$–$10^{-7}$ | Replication fork assembly |
|                    | DPB2-p78 | dpb2-p67 | POLE2-p59 | 3’-Exonuclease |          |          |
|                    | DPB3-p23 | dpb3-p22 | POLE3-p17 | DNA binding |          |          |
|                    | DPB4-p22 | dpb4-p24 | POLE4-p12 |          |          |          |

was to repeat this analysis in vivo in the pol2-M644G strain. The URA3 gene, a selectable mutational target, was placed in either the forward or opposite direction close to the ARS306 origin. Because this strong origin fires in nearly every cell cycle, the fork direction through the URA3 gene is unambiguously known. Finally, mutation spectra were determined, and A–T→T–A mutations were interpreted as having resulted from template dT-dTMP misincorporations based on the previous in vitro analysis. This interpretation was consistent with a model in which Pol ε carried out leading strand DNA replication much more frequently than lagging strand DNA replication (8).

Although one interpretation of this study substantiates Pol ε as the leading strand enzyme, it is also consistent with a model in which Pol ε carries out much more limited leading strand synthesis and still no lagging strand synthesis. To distinguish between these two interpretations, it is necessary to carry out a similar study with Pol δ and show that it does not carry out leading strand synthesis. Therefore, this type of analysis was carried out with a pol3-L612M asymmetric mutator mutant that, among other asymmetries, specifically increased the rate of the dT-dGMP mismatch but not the complementary dA-dCMP mismatch in an in vitro fidelity analysis. The subsequent in vivo analysis was entirely consistent with a model in which Pol δ carried out lagging and no leading strand replication (9). These results suggest the simple consensus fork model shown in Fig. 1.

These data are in agreement with previous studies of proofreading-defective mutants of Pol δ and Pol ε that had already suggested that these two enzymes proofread opposite strands of the replication fork and therefore likely replicate opposite strands (10). Together with the knowledge that Pol δ is the lagging strand enzyme (discussed below), this analysis also places Pol ε on the leading strand. However, these results are in apparent disagreement with mutational studies showing that the catalytic domain of Pol ε is dispensable for cell growth in both yeasts, S. cerevisiae and S. pombe (11–13). How can these disparate results be reconciled? To conclude simply from the domain deletion data that Pol ε does not normally replicate DNA would be to ignore the remarkable ability of the cell to adapt. In fact, the domain mutant shows severe phenotypic defects in the progression of DNA replication (14). Furthermore, chromatin immuno-precipitation studies in yeast show that Pol ε travels with the replication fork during DNA replication (15). The observation that the entire replication fork may have been rearranged under pressure of the Pol ε domain mutation suggests the notion that, even in wild-type cells, the fork may range into the simple SV40-type fork under some conditions (Fig. 1). Whether this is in response to replication stress or during the replication of late regions or heterochromatic regions should be a fascinating future area for study, especially in mammalian cells, where chromosomal regions show more variation in structure and properties than in the simple yeasts.

**Proofreading of Pol α Errors by Pol δ**

Lagging strand DNA replication is thought to proceed in several discrete stages: initiation by DNA primase; limited elongation of the RNA primer by Pol α; switching of the primer terminus from Pol α to Pol δ; elongation by Pol δ; and maturation.
MINIREVIEW: Polymerase Dynamics at the Replication Fork

of the elongated Okazaki fragment by Pol δ, FEN1, and DNA ligase. Each transition is believed to be mediated by a specific protein or protein complex and has to occur with very high efficiency. In a mammalian cell, this process occurs 20–50 million times during every cell cycle, and even in a yeast cell with its very compact genome, ~100,000 Okazaki fragments need to be initiated, elongated, and matured in a single S phase.

As the faithful duplication of the genome is of the utmost importance, it has always been a puzzling question what happens with the short segment of initiator DNA of ~20 nt that has been synthesized by Pol α, prior to entry of Pol δ. Given the lower fidelity of Pol α because of lack of exonucleolytic proofreading (Table 1), these initiation zones should turn into hotspots for mutations if they were not subject to some form of error correction. There is strong evidence that Pol δ can proofread errors made by Pol α. This is based on a recent study of Pol α mutant mutants that establishes an epistatic relationship between Pol α and Pol δ for fidelity (16). A pol1-L868M active-site mutant that showed a 6-fold increase in error frequency in vitro showed a marginal mutator phenotype in vivo. The single proofreading-defective pol3-exomutant had an ~7-fold increase in mutation rate. However, the pol1-L868M pol3-exo double mutant showed strong synergism, a 70-fold increase in mutation rates. In contrast, no synergism was observed between pol1-L868M and the analogous pol2-exo mutant defective for proofreading by Pol α. The simplest explanation for these results is that Pol δ, but not Pol ε, proofreads errors made by Pol α, and this is consistent with our replication fork model, which places Pol δ on the lagging strand and Pol ε on the leading strand.

Dynamic Interaction between FEN1 and Pol δ in Okazaki Fragment Maturation

Every 100–200 nt on the lagging strand, the replicating Pol δ runs into the RNA primer of the previous Okazaki fragment. An successful (18). As Pol δ reaches the 5’-end of the downstream Okazaki fragment, it continues an additional 1–2 nt of synthesis in a strand displacement mode (Fig. 2). FEN1 removes the small 5’-flap generated by strand displacement synthesis; this flap is most often only a mononucleotide in size. If additional initiator RNA is present, Pol δ and FEN1 go through iterative cycles of strand displacement and small flap cutting, a process called nick translation, until all RNA has been degraded. Once this process has generated a proper DNA-DNA nick, DNA ligase I acts to complete double-stranded DNA. PCNA complexes these three enzymes at the site of action to ensure processive maturation. Interestingly, crenarchaeal PCNA is a heterotrimer, with each individual PCNA subunit displaying preferential binding for a specific Okazaki fragment maturation enzyme, one PCNA subunit for the DNA polymerase, the second for FEN1, and the third for DNA ligase, thereby suggesting a precise architecture of the lagging strand replication complex (19).

The forward movement by Pol δ can be counteracted by its 3’-exonuclease activity. The nuclease activity of Pol δ generally proofreads polymerase insertion error to assure high fidelity DNA replication (Table 1). However, it also plays a crucial role in Okazaki fragment maturation (20). Exonuclease-mediated 3’-degradation can function to generate ligatable nicks from small 5’-flaps and to maintain those ligatable nicks by idling. Idling is the iterative process of limited strand displacement synthesis by the polymerase at a nick, followed by switching to the exonuclease domain and degradation of the displacing strand until the nick position has been reached again (Fig. 2). It is in idling that critical mechanistic differences are expressed between Pol δ and Pol ε. During idling, Pol δ maintains a dynamic relationship with the nick position, producing alternate substrates for FEN1 action and for DNA ligase I action; the leading strand Pol ε does not (18).
How long are these long flaps, and how are they generated? A yeast rad27Δ FEN1 accumulates duplications up to ~100 nt that likely result from the generation of up to ~100-nt flaps in the mutant strain due to defects in short flap processing (21, 30). However, any flap long enough to fold up or bind proteins and thereby inhibit FEN1 action is considered to be a long flap and needs processing by alternative pathways. In biochemical studies, flaps of ~30 nt in length bind RPA, inhibit FEN1 action, and activate Dna2 action (29, 31, 32). The generation of long flaps is normally prevented by the 3'-exonuclease activity of Pol δ. However, although idling can maintain Pol δ at a nick for some time, eventually the enzyme will shift to an irreversible strand displacement synthesis mode, during which extended regions of DNA are unwound (32, 33). Thus, long flaps likely result from a failure of the short flap pathway either because of FEN1 dysfunction at an unusual DNA sequence or structure or because of unusually efficient and extensive strand displacement synthesis by Pol δ. Genetic studies support the proposed backup mechanism for Dna2 when flaps have grown too long. When either the exonuclease activity of Pol δ or FEN1 activity is compromised, the tight control of the machinery to maintain a nick position is diminished, and it is lost in pol32-exo− rad27Δ double mutants, causing lethality of the double mutant. However, overexpression of Dna2 rescues the double mutant, suggesting that the accumulation of long flaps can be handled by increasing Dna2 protein levels (32).

The efficiency of strand displacement synthesis at the border of two Okazaki fragments is likely a prime determinant for the ratio of long to short flaps. In mutants that restrain strand displacement synthesis, the necessity for long flap processing by Dna2 should lessen in importance. Studies of the Pif1 helicase and the Pol32 subunit of Pol δ support this model. Pif1 is a 5′-3′-helicase that functions in mitochondrial DNA maintenance, telomere homeostasis, and lagging strand DNA replication (34, 35). The presence of Pif1 during gap filling by Pol δ favors the generation of longer flaps in vitro (36). Pol32 is the third, nonessential subunit of yeast Pol δ (Table 1). Compared with the three-subunit form of Pol δ, the two-subunit form lacking Pol32 has a decreased processivity of DNA synthesis (37) and shows reduced strand displacement synthesis (38). On the basis of their biochemical properties, one would expect that deletion of PIF1 and POL32 might suppress the lethality of either mutant strains that produce an excess of long flaps or of mutant strains that are defective for processing long flaps. Indeed, the lethality of rad27Δ pol3-exo− mutants that are predicted to produce long flaps is rescued by the additional deletion of PIF1 or POL32 (38). On the other hand, the lethality of dna2Δ
mutants that are defective in cutting long flaps is also suppressed by PIFI deletion and more robustly by the PIFI POL32 double deletion (35, 38).

By dividing displaced RNA strands over two compensatory pathways, the cell has developed a remarkable flexibility in dealing with different flap sizes and structures generated during Okazaki fragment synthesis. In addition, other proteins may perform auxiliary functions. Among these are RNase H2 and the RecQ-like helicase Sgs1 (7). Where and how these proteins function during maturation require further study, particularly because defects in them have been associated with human disease states.

Acknowledgment—I thank John Majors for critical discussions during the course of this work.

REFERENCES