Eukaryotic DNA Polymerases

Secondary article

Sue Cotterill, *St George's Hospital Medical School, London, UK*

Stephen Kearsey, *Department of Zoology, University of Oxford, Oxford, UK*

DNA is replicated and repaired by a family of enzymes called DNA polymerases. Eukaryotic cells have a diversity of these enzymes that, while sharing a common biochemical activity, are specialized for particular roles.

Article Contents

Introduction

The major function of DNA polymerases is to replicate the genome and thus to allow transmission of genetic information from one generation to the next. In eukaryotic cells, this takes place during a discrete period (S phase) of interphase, and replicated chromosomes are subsequently segregated to daughter cells during mitosis. As well as replicating DNA, polymerases help to maintain the integrity of the genome by participating in various modes of DNA repair.

The basic catalytic reaction of DNA polymerases is to effect semiconservative replication of DNA, using a single-stranded DNA chain as a template and four deoxynucleotides (TTP, dCTP, dGTP, dATP) as precursors for DNA synthesis (**Figure 1**). The enzyme assembles the precursor nucleotides on the template to form a complementary DNA strand, selecting the incoming nucleotide using the base pair rules A•T and G•C. To start synthesis on a single-stranded DNA molecule, DNA polymerases need a primer. This is a length of RNA or DNA that is annealed to the single-stranded template. The primer provides a 3′-OH that can be extended by the polymerase; this configuration of the primer is important because polymerases can only extend a new chain in the 5′ to 3′ direction. If the DNA template to be replicated is double-stranded, polymerase action still needs a primer but, in addition, requires other enzymes to unwind the double helix.

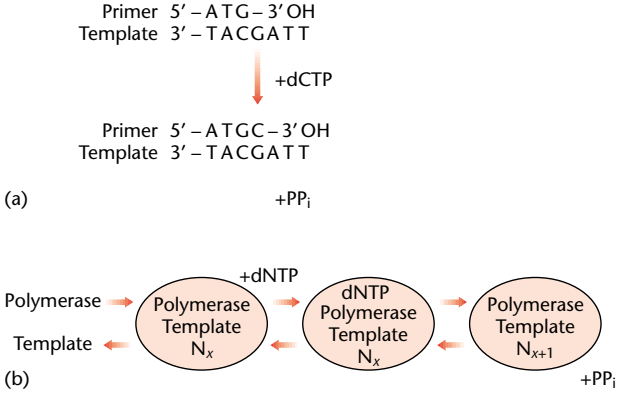

Figure 1 (a) Basic mechanism catalysed by DNA polymerases. Polymerization of nucleotides on the single-stranded template requires a primer (RNA or DNA) which provides a 3′-OH group to which the incoming nucleotide is joined and the direction of synthesis is thus 5′ to 3′. Only nucleotides that correctly base pair with the templates strand (according to A•T, G•C rules) are incorporated. One molecule of pyrophosphate (PP_i) is produced per nucleotide incorporated. (b) Steps in the polymerization reaction. The order of the reaction is binding to the template primer, followed by binding of the dNTP. dNTP is cleaved at the α/β bond to give dNMP that is added on to the chain, PP_i is released, and the polymerase translocates along to the next 3′ terminus or dissociates.

Structural Diversity of Polymerases

Although DNA polymerases all share the same basic catalytic mechanism, eukaryotic cells contain at least 15 distinct polymerases and more are likely to be discovered (Goodman and Tippin, 2000; Hubscher and Spadari, 2002; **Table 1**). The reason for this diversity seems to be that polymerases have become specialized by modifying the characteristics of the polymerization reaction or acquiring additional biochemical functions to tailor individual enzymes to particular tasks. At least three polymerases (α, δ and ε) are needed for chromosome replication, and even more polymerase diversity seems to be needed for DNA repair. Several modes of repair involve removal of damaged nucleotides followed by DNA synthesis. Different types of repair synthesis involve not only the replicative polymerases α, δ and ε but also a variety of specific repair polymerases.

Some polymerases consist of a single polypeptide chain, whereas others, such as those involved in chromosome replication, are composed of several different subunits. However, in each case one of the subunits is easily identifiable as containing the polymerase catalytic site and this allows classification of polymerases on the basis of amino acid sequence similarities. Eukaryotic DNA polymerases fall into five families, designated A (which also

Table 1 Summary of eukaryotic DNA polymerase functions

Polymerase	Family	Cellular function
α (alpha)	B	Chromosome replication (initiation, Okazaki fragment priming)
		Double-strand break repair
δ (delta)	B	Chromosome replication (elongation)
		Nucleotide-excision repair
		Base-excison repair
		Mismatch repair
		Double-strand break repair
ε (epsilon)	B	Chromosome replication (elongation)
		Nucleotide-excision repair
		Break-excision repair
		Mismatch repair
		Double-strand break repair
ζ (zeta)	B	Translesion synthesis
γ (gamma)	A	Mitochondrial DNA replication
θ (theta)	A	DNA repair
β (beta)	X	Base-excision repair
λ (lambda)	X	Base-excision repair
μ (mu)	X	Nonhomologous end joining
σ (sigma)	X	Sister chromatin cohesion
Telomerase	RT	Telomere maintenance
η (eta)	Y	Translesion synthesis
ι (iota)	Y	Translesion synthesis
κ (kappa)	Y	Translesion synthesis
Rev1	Y	Translesion synthesis

includes the archetypal *E. coli* DNA polymerase I); B (which includes all the eukaryotic polymerases involved in chromosome replication); and RT, X and Y, which include enzymes involved in DNA repair or specialized types of replication (**Table 1**). On the basis of sequence comparisons, polymerases in different families may have limited similarity, but polymerases in families A and B appear to have an identical mechanism for the polymerization reaction (see below).

Polymerases Involved in Chromosome Replication

Chromosomes are replicated during the S phase of the cell cycle by three multisubunit polymerases – DNA polymerases α, δ and ε (Hubscher *et al*., 2000) – which appear to be found in all eukaryotes. The process of DNA replication involves initiation, when DNA synthesis is activated at multiple replication origins in the chromosomes, and elongation, which involves extension of the short DNA chains generated during initiation. Only polymerase α has an associated primase activity capable of synthesizing short (~ 10 nucleotide) RNA primers, and this function is associated with two polymerase-associated subunits. DNA polymerase α is therefore the only enzyme that could be involved in the primer synthesis during initiation at origins of replication. Also, since the 5′ to 3′ directionality of DNA polymerases means that one strand of the DNA at any site has to be synthesized discontinuously, the α polymerase is also required during the elongation step for the priming of synthesis of Okazaki fragments on the lagging strand.

After DNA polymerase α has synthesized a short (30–40 nucleotide) stretch of DNA, a process called polymerase switching takes place in which polymerase α is displaced from the template and synthesis by polymerases δ and probably ε takes over. Polymerase δ is a multisubunit polymerase and probably functions at the leading and lagging strands of the replication fork. Polymerase ε consists of four subunits, and its precise role in chromosomal replication is unclear. It has been proposed that the function of DNA polymerase ε may be restricted to the lagging strand, perhaps only in the maturation of Okazaki fragments.

The polymerase switch is necessary for two reasons. Firstly because, in addition to their basic catalytic function, both δ and ε polymerases have a 3′–5′ exonuclease activity in their polymerase catalytic subunit that provides a proofreading function. This means that if an inappropriate nucleotide is incorporated (i.e. one that does not conform A•T, G•C pairing rules), the mispaired nucleotide is excised and the polymerase can then incorporate the correct nucleotide. Polymerase α does not have this activity and therefore this switch greatly enhances the fidelity of DNA replication. Second, both of these enzymes are able to associate with factors that increase their processivity. This is a measure of the number of nucleotides that a polymerase can incorporate on to a template without dissociating. Rapid replication of long stretches of DNA is needed during the synthesis of the complete genome. This process is therefore aided by the switch from a distributive polymerase (polymerase α) to a polymerase that is more processive.

Replication Accessory Factors

As can be seen from the discussion above, proper function of the replicative polymerases is dependent on a number of other proteins. DNA in a cell does not exist as a naked single-stranded structure on which polymerases can freely function, but rather as a double-stranded helix in a protein-covered chromosome. Polymerase duplication of chromosomal DNA can therefore only take place after a number of factors have acted to allow the polymerase access to the

DNA and create an environment conducive to polymerase activity. During initiation of replication, a whole network of proteins is involved in generating single-stranded DNA and loading polymerases on to the origin (Bell and Dutta, 2002). Initial synthesis of the primer by polymerase α takes place after the action of factors forming the pre-replicative complex, namely ORC (origin recognition complex), Cdc6, Cdt1 and the MCM (minichromosome maintenance) proteins. Polymerase α loading also needs the chromatin association of additional proteins (such as Cdc45 and replication protein A (RPA)) proteins, which occurs subsequent to pre-replicative complex assembly. Initiation and elongation require DNA unwinding to take place, and the MCM protein complex most likely provides this helicase activity. It may be critical to control the chromatin association of polymerase α, since it has intrinsic primase activity and can potentially initiate replication *de novo*.

Following priming of synthesis, additional factors are required to allow polymerase switching, in which polymerases δ and ε take over from polymerase α and elongate the nascent DNA chain. This process involves the combined action of the RPA, RFC (replication factor C) and PCNA (proliferating cell nuclear antigen) proteins (Waga and Stillman, 1998). When polymerase α encounters RPA on single-stranded DNA, the DNA polymerase α activity is inhibited. This allows binding of RFC, which recognizes the 3′ end of the DNA formed by polymerase α. RFC has been termed 'clamp loader' as it functions to assemble PCNA, which forms a trimeric clamp, around the DNA template. PCNA links polymerases δ and ε to the template and acts to increase the processivity of the polymerases by preventing premature dissociation from the template. On the leading strand, polymerase switching is only required once, but on the lagging strand it is needed every time DNA polymerase α initiates (100–200 nucleotides). Synthesis is terminated on the lagging strand when polymerase δ or ε collides with the 5′ end of an Okazaki fragment. At this point, other accessory factors such as the FEN exonuclease, Dna2 exonuclease, RNAseH and DNA ligase are needed to remove the RNA primer and to seal the nick between the DNA fragments, allowing completion of DNA synthesis.

As well as physically replicating the DNA in the chromosomes, polymerases have been proposed to have an additional cell cycle role by functioning in checkpoint controls. If DNA is damaged or replication is inhibited, checkpoint mechanisms can slow S phase or prevent entry into mitosis, allowing extra time for DNA repair to be carried out, and this helps to preserve the integrity of the genome. Mutations affecting DNA polymerase ε are defective in checkpoint control in budding yeast, perhaps because this polymerase has a role in sensing problems with DNA replication (Navas *et al*., 1995). DNA polymerase α/primase has also been implicated in the replication checkpoint (Michael *et al*., 2002).

Polymerases Involved in DNA Repair

Many types of DNA repair involve removal of damaged or mismatched nucleotides to form a single-stranded region, and ultimately any situation where DNA is removed will require the action of DNA polymerases (Wood and Shivji, 1997). A diverse set of proteins is needed for recognition of the various different types of damaged DNA and its processing to provide a suitable substrate for DNA polymerases, and the discussion here is restricted to polymerase function in the final synthetic stages of repair.

Polymerase β

In base-excision repair, removal of an altered base is followed by excision of a single abasic nucleotide. DNA polymerase β then serves to fill in the missing nucleotide; this involves an activity in the N-terminal domain of the protein that removes the 5′ phosphate remaining after nucleotide excision, followed by a polymerization reaction to fill in the missing nucleotide. This limited synthesis makes sense in terms of the biochemical properties of DNA polymerase β, which is distributive and has no proofreading activity. In a different mode of base-excision repair (termed 'long-patch'), a longer (2–10 nucleotide) stretch of DNA is removed, but it is unclear whether this is filled in by the β, or by the replicative δ and ε polymerases.

Polymerases α, δ, ε

Two other types of repair involve the replicative DNA polymerases δ or ε, in conjunction with the RFC and PCNA as cofactors. In nucleotide-excision repair, resynthesis by these polymerases follows removal of an oligonucleotide of 24–32 reduces that contains the damaged nucleotide. Mismatch repair corrects regions of DNA that contain mismatched nucleotides (i.e. where A is not paired with T and G is not paired with C). The mismatched segment is removed, and polymerases δ and ε may be required to resynthesize several hundred nucleotides around the repair.

A distinct mechanism is required for double-strand break repair, where cleavage of both strands of the double helix means that simple resynthesis using the undamaged strand as template cannot be used to effect repair. One mechanism of double-strand break repair involves homologous recombination between the damaged chromosome and a second undamaged copy (e.g. the homologous chromosome in a diploid cell). The undamaged chromosome then serves as a template to allow synthesis across the site of the break and therefore ultimately allows rejoining of the broken fragments. Genetic evidence from yeast indicates that all three replicative polymerases are needed for this mode of double-strand break repair, suggesting a model in which the broken DNA end invades the intact

chromosome, setting up a modified replication fork with leading-and lagging-strand synthesis (Kraus *et al*., 2001). Double-strand breaks are also formed in meiotic cells and are needed to promote homologous recombination. It is likely that DNA polymerases are required for this process, although the identity of the polymerases involved is not clear.

Repair by Error-Prone Polymerases

A number of polymerases have also been identified that appear to function by allowing a replication fork to pass the site of a damaged DNA strand. These polymerases are not halted by damage in the same way as the normal replicative polymerase are, but can synthesize across the site of damage, although often in a very error-prone fashion. They can also copy undamaged DNA, although with poor fidelity partly because they lack a proofreading activity but also because they may be able to extend primers that are mispaired with the template, or more readily incorporate nucleotides that do not form correct Watson–Crick base pairs with the template. The diversity of these enzymes may reflect the fact that they are specialized to cope with replicating across different types of lesion. This field is expanding rapidly and we describe only a few examples here; the reader is directed to comprehensive reviews for a more detailed account (Goodman, 2002; Hubscher and Spadari, 2002).

Polymerase η

The importance of polymerase η is demonstrated by the genetic disease xeroderma pigmentosum variant (XPV) in which affected individuals lack the enzyme and as a consequence suffer a high rate of sunlight-induced skin cancer. This defect seems to arise from the requirement for polymerase η in copying T-T photodimers, which are a product of UV irradiation of DNA. Normal replicative polymerases are impeded by these photodimers, but polymerase η correctly incorporates two As opposite each mutant T-T site, thus allowing error-free replication. Polymerase η may be able to bypass other types of bulky lesion but has a high misincorporation frequency (1 in 10^{-1}–10^{-3} compared to 10^{-6} for polymerase δ). Recent evidence suggests that pol η interacts with PCNA, as found with the replicative polymerases δ and ε.

The error-prone nature of polymerase η DNA synthesis may be exploited to generate mutations in antibody genes. This process, called somatic hypermutation, allows high-affinity antibodies to be generated by mutating the corresponding antibody genes at a rate that is 10^6 times higher than background levels (Goodman and Tippin, 2000). Polymerase η is one of several error-prone polymerases that have been proposed to play a role in this process. For instance, XPV patients lacking polymerase η have a reduced A•T mutation mutation rate in their antibody genes, although at least one other polymerase is required to provide mutation of G•C sites. Little is understood of how error-prone DNA replication is restricted to such a specific region of immunoglobulin genes.

Rev1 and polymerase ζ

Biochemical and *in vivo* data suggest that Rev1 and polymerase ζ are involved in dealing with a different kind of lesion, when a template base is missing. Bypass of an abasic site is promoted by Rev1p, which incorporates a C in a template-dependent (but not template-directed) manner. This incorporated C cannot be used as a primer by normal polymerases, as it is not correctly base-paired to the template, but synthesis may be resumed by polymerase ζ, which can use a mismatched primer with high efficiency, although subsequent synthesis is low fidelity.

Polymerases with Specialized Functions

Polymerase σ

Polymerases α, δ and ε replicate most of the chromosomal DNA, but specialized polymerases are required for replicating certain chromosomal regions and the extra-chromosomal DNAs found in organelles. Polymerase σ has recently been discovered to be necessary for ensuring that the two replicated chromosomes (sister chromatids) remain attached together after DNA replication. This continued association is important for ensuring the proper separation of the chromosomes in mitosis (Wang *et al*., 2000). Polymerase σ is another example of an error-prone polymerase, and it has been suggested that this polymerase could take over from polymerase δ or polymerase ε during replication of special parts of the chromosome in order to establish cohesion between the sister chromatids. However, a recent report has cast doubt on this proposed function of polymerase σ (Saitoh *et al*., 2002).

Telomerase

Telomerase is another polymerase required to complete chromosome synthesis, for replication of the telomeres at the ends of chromosomes. This enzyme is actually an RNA-directed DNA polymerase, and is unusual in that it is not template directed but uses an internal RNA molecule to direct synthesis of short repeated sequences that are added to the ends of chromosomes.

Polymerase γ

DNA polymerase γ is found in mitochondria and is required for replication and repair of the mitochondrial DNA.

Catalytic Function of Polymerases

Structural analyses of family A and family B polymerases has revealed similarity in their structures and catalytic mechanisms (Patel and Loeb, 2001; Steitz, 1999). The enzyme can be compared to a cupped right hand, with the palm providing the catalytic amino acids, the fingers binding the incoming nucleotide and the single-stranded template, and the thumb binding double-stranded DNA (**Figure 2**). The palm domain appears to be homologous for the A and B families and, while the finger and thumb domains are not, they show analogous features. Both A and B family polymerases show a conformational change on DNA binding that involves the DNA-binding thumb domain rotating towards the catalytic palm domain. On binding of the dNTP, the polymerase then adopts a 'closed conformation' in which the nucleotide-binding site closes up and the incoming dNTP becomes enclosed in a pocket with hydrophobic residues around the base and ribose and a hydrophilic environment around the triphosphate. Recent work has mapped in more detail the contact points of the nucleotide to two aspartic acid residues in the active site that are widely conserved in both DNA and RNA polymerases. These make contact with two magnesium ions that enter the binding cleft associated with phosphates of the incoming nucleotide. These metal ions are critical for catalysis of the polymerization reaction, and function by promoting the nucleophilic attack of the primer 3′-OH to the α-phosphate of the incoming dNTP (Steitz, 1998). This closed pocket conformation is only possible if the incoming nucleotide achieves correct Watson–Crick pairing with the template, thus helping to promote fidelity of the enzyme.

As mentioned above, a number of the replicative polymerases have, in addition to the polymerase activity, an exonuclease activity that is involved in proofreading – the removal of an inappropriate nucleotide that does not have correct Watson–Crick pairing with the template. Analysis of the crystal structure has also been able to shed light on this aspect of polymerase function. The exonuclease domain is not within the polymerization active site but is about 30–40 Å away. If the wrong nucleotide is incorporated on to the growing chain, it slows polymerization, probably because the 3′-OH of the primer to which the next nucleotide must be added is misorientated. This allows time for the primer terminus to move into the exonuclease active site. The DNA loses contact with the polymerization cleft but remains attached to the polymerase via the thumb domain, which rotates allowing movement of the DNA chain into the exonuclease domain. The fact that the DNA template does not lose contact with the polymerase at any time during this operation allows misincorporations to be removed efficiently and this

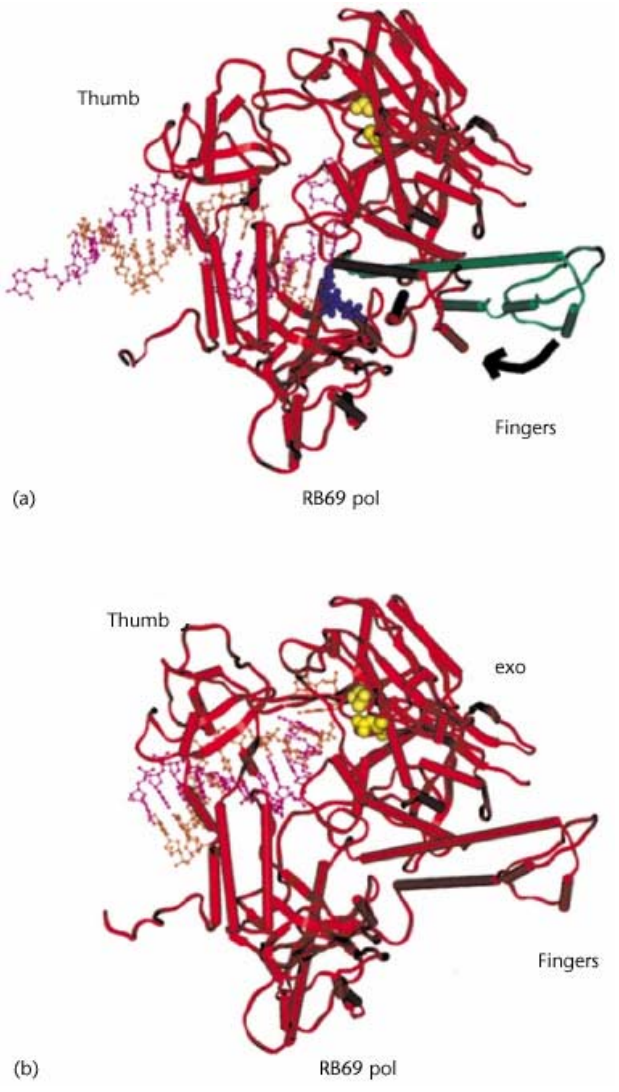

Figure 2 Structure of RB69 polymerase. RB69 is a prokaryotic family B polymerase, and is therefore likely to have a structure representative of eukaryotic replicative polymerases (α, δ, ε; Franklin *et al.*, 2001). (a) Polymerization mode. DNA template-primer (pink-orange) first binds to the polymerase thumb domain, followed by binding of a nucleoside triphosphate (blue), complementary to the template base immediately adjacent to the 3′-OH of the primer. On binding the dNTP, the finger domain (shown here in the open conformation) rotates (black arrow) towards the palm domain to form a closed ternary complex. This facilitates phosphodiester bond formation between primer 3′-OH and the α-phosphate of the incoming dNTP. (b) In the exonuclease (proofreading) mode, the thumb tip rotates to partition the DNA to the exonuclease site (shown in yellow), allowing removal of a noncomplementary base. Reproduced with permission from Patel and Loeb (2001).

mechanism therefore contributes to both the fidelity and processivity of the polymerase reaction.

Assays

The incorporation of nucleotides easily lends itself to an assay for the enzyme. All assays for DNA polymerases involve measurement of incorporation of labelled nucleotides into a primed DNA template (Wang *et al*., 1999). Various templates have been used. The earliest was activated calf thymus DNA, where single-stranded gaps are made in a double-stranded template by DNAse I. In this case the ability of the enzyme to fill in the gaps (on average 30–40 bp) is measured. More modern assays use single-stranded DNA molecules (such as bacteriophage M13) with a single primer, or use homopolymer type substrates such as poly(dA) sparsely primed with dT. These templates have an additional advantage in that the ability of a polymerase to synthesize a long stretch of DNA, or to pass a particular sequence or DNA structure, can also be ascertained.

Biochemical characterization of polymerase activity *in vitro* routinely measures three parameters: the rate of polymerization, fidelity and processivity. The fidelity for polymerases with a proofreading activity is generally around 1 error per 10^5 to 10^6 nucleotides incorporated. Processivity can be very variable depending on the polymerase and is often increased by accessory factors.

Current Research Topics and Unanswered Questions

The recent increase in the number of known eukaryotic polymerases makes it likely that others remain to be discovered, some of which could have crucial cellular roles. Given this similarity of catalytic mechanism between polymerases in families A and B, it will be interesting to see whether this extends to these other newly discovered polymerase families. If the basic mechanisms are the same, what are the modifications that occur to allow altered properties such as error-prone synthesis? Even for well-characterized polymerases there are outstanding basic questions concerning their function. For instance, at present we have only a vague idea how polymerases are recruited to replication origins. It is not clear how polymerases δ and ε cooperate in chromosomal replication and, given that potentially any primer 3′-OH can potentially be extended by any polymerase, how is polymerase selection effected to ensure that only the right polymerase is used in the right place? Finally, it will be important to establish the roles of DNA polymerases outside of chromosome replication and repair in processes such as recombination, checkpoint control and chromosome cohesion.

References

Bell SP and Dutta A (2002) DNA replication in eukaryotic cells. *Annual Review of Biochemistry* **71**: 333–374.

Franklin MC, Wang J and Steitz TA (2001) Structure of the replicating complex of a pol alpha family DNA polymerase. *Cell* **105**: 657–667.

Goodman MF (2002) Error prone repair DNA polymerases in prokaryotes and eukaryotes. *Annual Review of Biochemistry* **71**: 17–50.

Goodman MF and Tippin B (2000) The expanding polymerase universe. *Nature Reviews of Molecular Cell Biology* **1**: 101–109.

Hubscher U, Nasheuer HP and Syvaoja JE (2000) Eukaryotic DNA polymerases, a growing family. *Trends in Biochemical Sciences* **25**: 143–147.

Hubscher U and Spadari S (2002) Eukaryotic DNA polymerases. *Annual Review of Biochemistry* **71**: 133–163.

Kraus E, Leung W-Y and Haber JE (2001) Break-induced replication: a review and an example from budding yeast. *Proceedings of the National Academy of Sciences of the USA* **98**: 8255–8262.

Michael WM, Ott R, Fanning E and Newport J (2002) Activation of the DNA replication checkpoint through RNA synthesis by primase. *Science* **289**: 2133–2137.

Navas TA, Zhou Z and Elledge SJ (1995) DNA polymerase epsilon links the DNA replication machinery to the S phase checkpoint. *Cell* **80**: 29–39.

Patel PH and Loeb LA (2001) Getting a grip on how DNA polymerases function. *Nature Structural Biology* **8**: 656–659.

Saitoh S, Chabes A, McDonald WH *et al*. (2002) Cid13 is a cytoplasmic poly(A) polymerase that regulates ribonucleotide reductase mRNA. *Cell* **109**: 563–573.

Steitz TA (1998) A mechanism for all polymerases. *Nature* **391**: 231–232.

Steitz TA (1999) DNA polymerases: structural diversity and common mechanisms. *Journal of Biological Chemistry* **274**: 17395–17398.

Waga S and Stillman B (1998) The DNA replication fork in eukaryotic cells. *Annual Review of Biochemistry* 67: 721–751.

Wang TS, Conger KL, Copeland WC and Arroyo MP (1999) Eukaryotic DNA polymerases. In: Cotterill S (ed.) *Eukaryotic DNA Replication*, pp. 67–92. Oxford: Oxford University Press.

Wang Z, Castano IB, De Las Penas A, Adams C and Christman MF (2000) Pol kappa: a DNA polymerase required for sister chromatid cohesion. *Science* **289**: 774–779.

Wood RD and Shivji MKK (1997) Which DNA polymerases are used for DNA repair in eukaryotes? *Carcinogenesis* **18**: 605–610.

Further Reading

Baker TA and Bell SP (1998) Polymerases and the replisome: machines within machines. *Cell* **92**: 295–305.

Burgers P (1998) Eukaryotic DNA polymerases in DNA replication and DNA repair. *Chromosoma* **107**: 218–227.

Wang TF (1996) Cellular DNA polymerases. In: DePamphilis M (ed.) *DNA Replication in Eukaryotic Cells*, pp.461–494. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.