Okazaki fragment processing during eukaryotic DNA replication

Lynne S. Cox

University Lecturer

Department of Biochemistry

University of Oxford

Oxford OX1 3QU

UK

e-mail: lscox@bioch.ox.ac.uk

Running title: Okazaki fragment processing

Key Words: Okazaki fragments; eukaryotic DNA replication; DNA polymerase; RNaseH1; FEN-1; PCNA; DNA ligase 1

Summary

Okazaki fragments containing an RNA primer are synthesised on the lagging strand of the eukaryotic replication fork. Major advances have recently been made in our understanding of the complex process of Okazaki fragment processing to form high molecular weight daughter DNA, identifying distinct stages. Firstly, primase synthesises an RNA primer which is rapidly extended into DNA by bound DNA polymerase α. DNA pol ε is then thought to synthesise the remainder of the fragment (50-200 nucleotides). The RNA primer is removed almost intact by RNaseH1, leaving a single ribonucleotide that is cleaved by the FEN-1 nuclease. DNA originally synthesised by pol α is replaced by DNA pol ε in co-operation with FEN-1. Finally, DNA ligase 1 seals the resulting nick by forming a phosphodiester bond between adjacent DNA nucleotides. A model whereby

these complex processes are modulated by DNA context and regulatory protein-protein interactions is presented.

Why is DNA synthesis semi-discontinuous ?

Eukaryotic DNA replication on duplex DNA of anti-parallel strands occurs in a bi-directional, and semiconservative manner (reviewed in [1]). Co-ordinate replication of both strands at each fork of a replication bubble must therefore result in one nascent strand being synthesised in a 5'->3' direction, while the complementary strand is synthesised in an overall 3'->5' direction. However, all known replicative DNA polymerases can synthesise DNA only in the 5'->3' direction. Pioneering studies in *E. coli* with or without phage infection showed that newly replicated DNA was present as short fragments of 1000-2000 nucleotides, sedimenting at 9-11S in alkaline sucrose. With time, these fragments matured into higher molecular weight DNA [2] (see Figure 1). These short nascent regions of DNA are now commonly known as Okazaki fragments. From these early experiments, it was concluded that DNA was replicated in a discontinuous fashion, on one or both strands [3]. Subsequent work, particularly using the eukaryotic virus SV40 [4], has demonstrated that one strand is continuously made in the 5'->3' direction - this is the leading strand. Its complementary strand is copied as short fragments, also synthesised in the 5'->3' direction, allowing an overall direction of synthesis of 3'->5' on this, the lagging strand, without contravening the DNA polymerase requirement to act in the 5'->3' direction (see Figure 2). Eukaryotic Okazaki fragments are approximately 10-fold shorter than those of prokaryotes, around 100-200 nucleotides, and sediment around 4S. This size difference may reflect the 10-fold slower fork rate in eukaryotes of approx. 50-10 nt/second, compared with prokaryotic rates of 500-1000 nt/second (reviewed in [1]). In order to understand the problems facing the cell when processing Okazaki fragments during DNA maturation, it is important to consider the structure of these fragments, as established by their specific mode of synthesis.

Each Okazaki fragment begins with an RNA primer

To ensure fidelity of DNA replication and prevent random DNA synthesis, DNA polymerases do not initiate DNA replication *de novo*, but catalyse phosphodiester bond formation between the free 3'OH group of a pre-exiting oligoribonucleotide primer base-paired to a DNA template strand, and the 5' phosphate group of an incoming deoxyribonucleotide. By contrast with DNA polymerases, RNA polymerases are able to polymerise nucleotides *de novo*, and it is an RNA polymerase-type enzyme known as primase that is responsible for producing the initial oligoribonucleotides. Following opening of the double helix at the replication origin, the primase, in tight complex with DNA polymerase α, binds to unwound single strand regions of template DNA and slides along until it encounters a preferred initiation site. Eukaryotic primase demonstrates some degree of template sequence preference, particularly selecting pyrimidine-rich template regions giving RNA primer sequences such as GGAAGAAAGC; not all pyrimidine-rich sequences support primer formation. Interaction of primase with the single strand DNA binding protein, RP-A, and the concentration of NTPs may also influence the choice of priming initiation site.

Primase consists of two subunits: p49 possesses RNA polymerase activity, while p58 is necessary for stabilising the binding of p49 to template DNA, and for mediating primase interaction with the p180 large subunit of DNA polymerase α [(5)]. Both primase subunits are able to bind to nucleotide triphosphates, allowing the enzyme to bind both to the newly added rNTP and the NTP that is to be added next to the RNA primer. The very first step in primer formation probably involves an intramolecular translocation event, whereby the 5' terminal nucleotide bound to the primase active site in p58 is transferred to the polymerase active site in p49 [(5)]. Primase, like eukaryotic RNA polymerase II, has a much higher K_m for initiation than for elongation [(5)], perhaps ensuring greater control of priming.

Extension of DNA primer

Once RNA primers of more than 7 ribonucleotides has been synthesised, DNA polymerase α that is tightly associated with primase takes over phosphodiester bond formation, incorporating deoxribonucleotides to extend the RNA primer with DNA. Analysis of replication using purified eukaryotic proteins demonstrated that DNA pol α, when in association with its accessory factor (AAF), could cover a single stranded circular DNA template with Okazaki-like fragments [(6)]. However, this monopolymerase system is unlikely to reflect the situation *in vivo*, as it lacks many of the factors known to be required for replication in whole cells. For example, DNA pol α activity (but not primase) is found to be stimulated by the single stranded binding protein RP-A. On reconstituting SV40 DNA replication *in vitro* using purified human proteins, DNA pol α was unequivocally shown to be capable of lagging strand DNA synthesis [(7)].

DNA polymerase switching

As DNA pol α/primase possesses the only known priming activity, it is thought that this enzyme initiates DNA synthesis on both leading and lagging strands. Replication of SV40 DNA *in vitro* in the absence of PCNA, a DNA polymerase δ auxiliary protein, resulted in production only of lagging strand products, suggesting that the switch to PCNA-dependent elongation involves a second DNA polymerase. The PCNA-dependent, highly processive, high-fidelity proof-reading DNA pol δ is the obvious candidate for synthesis of the leading strand. A mechanism by which leading strand synthesis is initiated by pol α /primase and then polymerases are exchanged, has been suggested based on further studies in the SV40 system. This polymerase switch is thought to involve the formation of a tight complex of replication factor C (RF-C) and PCNA at the 3'OH of nascent DNA at the primer-template junction. This complex prevents further synthesis by DNA pol α and may result in its dissociation from the template, while favouring association of DNA pol δ. RP-A may also prevent spurious binding by DNA pol α. The dissociated pol α/primase is then free to initiate synthesis on the lagging strand.

Such “dipolymerase” systems based on replication of SV40-origin-containing DNA *in vitro* have proven extremely useful in identifying roles for individual replication factors during the complex replication process. However, they too may not fully reflect the situation *in vivo*, since a further DNA polymerase, DNA pol ε, is essential in yeast. This DNA pol ε has also been implicated in mammalian lagging strand DNA replication, in studies using simple DNA template with purified mammalian proteins [(8))]. DNA pol ε was subsequently shown to complete the synthesis of artificial Okazaki fragments *in vitro* [(9)]. In conjunction with the essential nature of the DNA pol ε gene in yeast and its ability to carry out Okazaki fragment completion *in vitro*, it is highly possible that *in vivo*, a polymerase switch is made on the lagging as well as the leading strand, with DNA pol ε taking over from DNA pol α [(8)]. This idea of a switch from DNA pol α to DNA pol ε is supported by the finding that leading and lagging strand synthesis have the same error rate, while DNA pol α lacks 3’->5’exonuclease proof-reading activity and consequently has lower fidelity than either DNA pol δ or DNA pol ε, both of which possess proof-reading activity [(1)]. The validity of this model has not been conclusively demonstrated, and it should be borne in mind that the ability of DNA polymerases α or ε to carry out lagging strand synthesis *in vitro* is not necessarily indicative that they do so *in vivo*. Although the fidelity argument could also support a role for DNA pol δ rather than DNA pol ε in lagging strand synthesis, this has not been observed *in vitro* [(7)] and cannot account for the essential nature of the DNA pol ε gene in yeast. It is therefore assumed throughout the remainder of this review that polymerase switches occur on both leading and lagging strands, with DNA pol δ replicating the leading strand and DNA pol ε completing lagging strand synthesis (Figure 2).

RNaseH1 removes the RNA primer

As consequence of RNA-primed discontinuous DNA synthesis on the lagging strand, eukaryotic Okazaki fragments contain an RNA primer of 8-10 ribonucleotides. Since mature genomic DNA lacks any ribonucleotide regions, the RNA primer component must be removed and replaced by

DNA, and gaps in the sugar-phosphate backbone sealed in a critical step of maturation of replication products into full length genomic DNA[6,10,11]. Removal of the RNA component is carried out by an RNase enzyme. In eukaryotes, RNaseH levels and activity correlate with proliferative status, supporting a role for it in DNA synthesis. It is an ubiquitous enzyme, present at approximately 20,000 copies per cell nucleus [12]. Biochemical purification of RNaseH from several sources has helped to elucidate the properties and activity of this enzyme. Eukaryotic RNaseH1 from various species varies in mass between 68-90 kDa, has an isoelectric point of approximately 6.4 and uses either magnesium or manganese ions as cofactors. RNaseH is thought to interact with and stimulate DNA pol α/primase, by removing the initiating ribonucleotide primers [10].

During replication of single-stranded DNA in a mouse *in vitro* system, RNaseH1 was required for removal of all but the final ribonucleotide of the RNA primer [6]. Elegant experiments by Bambara and colleagues using synthetic Okazaki fragments *in vitro* have confirmed that a single monoribonucleotide is left at the 5' terminus of the RNA-DNA junction following RNaseH1 action [9]. RNaseH1 cleavage occurs independently of the RNA primer sequence or primer length. The RNA fragment is removed as an intact product, suggesting that RNaseH1 makes a single endonucleolytic cleavage at the phosphodiester bond one nucleotide 5' of the RNA-DNA junction [9,13] (Figure 3). For example, an RNA primer consisting of 21 nucleotides is released by RNaseH1 as an intact 20 nucleotide product, with the remaining ribonucleotide still attached to the DNA of the Okazaki fragment [9]. Cleavage requires Mg^{2+} and is less specific in the presence of Mn^{2+}. The specificity of cleavage is also lost if the RNA primer is not extended with DNA, or if a nick exists at the RNA-DNA junction. It is possible that the RNaseH1 enzyme recognises the transitional structure at the RNA-DNA junction, though mismatches at the junction, which would presumably distort the structure, do not prevent recognition or cleavage by RNaseH1 (reviewed in [13]).

Nucleases in Okazaki fragment processing - identification of FEN-1

Although RNaseH1 can remove a significant portion of the primer, it is apparent from several lines of evidence that an additional nuclease is also required for correct processing of Okazaki fragments. For example, yeast mutants null for RNaseH1 are still able to process Okazaki fragments, presumably using a "back-up" mechanism [12]. SV40 DNA replication *in vitro* in the presence of replication factors RP-A, RF-C, PCNA, topoisomerase I and II, DNA pol α and DNA pol δ, resulted in nascent DNA where Okazaki fragments were synthesised but were not processed to mature DNA [7]. Additional proteins required for processing of SV40 Okazaki fragments *in vitro* included a nuclease termed MF1 and DNA ligase 1, though no RNaseH1 was added [11]. Similarly in the monopolymerase system, form I DNA was produced only on addition of a 5'->3' exonuclease activity, together with RNaseH and DNA ligase [10]. While mouse RNaseH was shown to be able to remove 80% of RNA primers synthesised by pol α/primase *in vitro* on an M13 template, complete primer removal only took place on addition of a 5'->3' exonuclease [6]. The same 5'->3' exonuclease activity has been identified independently by various groups using a range of different biological systems (reviewed by [14]). Cloning and sequencing of the genes from various eukaryotic species shows several regions of strong conservation representing nuclease domains [14], together with a short C terminal region involved in contacting PCNA (reviewed in [15]). There is also considerable structural and functional homology with related viral and prokaryotic enzymes (see [14]). Throughout this review, I shall use the term FEN-1 [16] for this eukaryotic nuclease, with species designated as y for yeast (*S. cerevisiae*), m for mouse, c for calf and h for human. FEN-1 is a member of a larger protein family including the known and putative yeast DNA repair proteins RAD2, rad13, rad2 and vertebrate XP-G (reviewed by [14]).

In vitro replication studies therefore suggest that FEN-1 is essential for maturation of nascent DNA [10,11,]. Such *in vitro* evidence is further supported by genetic evidence from yeast lacking FEN-1 (*rth1*Δ), which arrest with a dumbbell phenotype at the restrictive temperature, characteristic of a replication defect [17]. At the permissive temperature, yFEN-1 mutants show a hyper-recombination

mutator phenotype, suggesting the persistence of single stranded, nicked or gapped DNA; this is consistent with a failure in Okazaki fragment processing [17,18]. Such failure to remove RNA primers has been correlated with an increased invasion frequency of ssDNAs and an elevated recombination rate, while the DNA rearrangements observed in yFEN-1 mutants may also be present in some human diseases e.g. familial hypertrophic myopathy, lethal junctional epidermolysis bullosa and somatic mutations in the p53 and APC tumour suppressor genes, perhaps predisposing to human cancer [18]. Understanding of Okazaki fragment processing at the molecular level is therefore highly relevant to the understanding, and eventually rational treatment, of human disease. To assist in understanding these processing steps, FEN-1 has been purified from various sources allowing a full molecular dissection of its substrate preferences and enzymatic activities.

Co-operation between RNaseH and FEN-1 in primer processing

FEN-1 is thought to act immediately after RNaseH cleavage, to remove the final ribonucleotide from the initiating RNA primer [9] by acting either as a 5' (five')->3' exonuclease, or a structure specific flap endonuclease [16]. Pre-digestion of the RNA primer of an Okazaki fragment by RNaseH1 stimulates cFEN-1 to remove the junctional ribonucleotide, though the two enzymes may compete with each other if present simultaneously [13]. At least *in vitro*, the optimal molar ratio of RNaseH1: FEN-1 is 1:1000 [13]. Since RNaseH1 is present as a low level contaminant of purified FEN-1, it is highly possible that RNaseH1 activity was present in the reconstituted dipolymerase system discussed above [11]. Hence, even though purified RNaseH1 had not been added [11], there may have been sufficient activity in complex with hFEN-1 to allow processing of RNA primers by RNaseH1, and therefore to produce optimal substrates for hFEN-1 cleavage. Again in the absence of added RNaseH1, cFEN-1 has been shown to remove a displaced RNA primer flap [12]. However, such experiments are also susceptible to the same criticism, namely that low levels of RNaseH1 contaminating the FEN-1 preparation are responsible for at least some of the observed primer removal.

FEN-1 endonuclease activity - replication or repair?

Substrate specificity of FEN-1 has been determined using synthetic oligonucleotides and primed M13 templates with or without mismatches to create flap structures (Figure 4A). Current evidence points towards FEN-1 binding to the terminal 5' ribo- or deoxyribo-nucleotide of a 5' flap structure, and sliding down until it reaches the junction with duplex DNA. FEN-1 bound at this junction region has been shown by footprinting analysis to protect 25 nucleotides [17]. FEN-1 can then cleave one nucleotide downstream of any 5' flap structure made of either DNA or RNA, acting irrespective of flap length or sequence [16]. It can also cleave pseudo-Y structures (see Figure 4A), although optimal cleavage has been reported to require the adjacent DNA strand (F_{adj}, Figure 4A). Subsequent studies showing that in some circumstances, an adjacent DNA strand is neutral or even inhibitory for FEN-1 cleavage may actually reflect the two modes (endo- or exo- nuclease) of FEN-1 activity being assayed.

After binding to the end of a 5' flap structure, FEN-1 is thought to slide to the point of cleavage, then make a single endonucleolytic cut to remove the single stranded flap region. Since cFEN-1 already loaded onto a flap structure can be trapped by subsequent coating of the single stranded region with SSB or by covalently binding streptavidin to a biotin-modified base at the 5' terminus of the flap, FEN-1 is probably capable of bi-directional sliding. Although FEN-1 cannot move past SSB or streptavidin, it can cleave flap structures with a terminal triphosphate, 7-methyl3'G5'ppp5'G cap [12] or a flap that contains biotin, 3^{o} butylsilyl adducts, or cisplatin adducts (CDDP) internally or at the 5' end [19]. However, no cleavage can occur if the CDDP adducts are present at the expected cleavage point [19]. This result has important implications for the replication of damaged DNA, and for the repair of bulky lesions. Hence, cleavage 5' to a lesion by a repair endonuclease followed by gap filling DNA synthesis would result in displacement of the strand containing the lesion, to form a 5' flap structure onto which FEN-1 could bind, slide past the lesion,

and remove the flap containing the adduct (Figure 4B). Displacement DNA synthesis may therefore be vital in producing the preferred substrate for FEN-1's endonucleolytic activities (see below).

Such activities suggest a role for FEN-1 in DNA repair as well as in DNA replication. To date, FEN-1 has not been directly shown to be required for nucleotide excision repair, though it is essential for the long patch component of base excision repair (BER) [20]. Consistent with its role in BER, FEN1 can also remove monoribonucleotides that have accidentally become ligated into DNA. Removal requires cleavage at the 5' side of the ribonucleotide by RNaseH, producing an 5' nicked substrate which FEN-1 can cleave by virtue of its exonuclease activity. The gap is then presumably filled by the action of DNA pol ε (or DNA pol α). Similarly, cleavage of 5' abasic sites through FEN-1 endonuclease activity can be demonstrated *in vitro* , requiring displacement of the abasic strand by adjacent DNA, perhaps mimicking displacement by an incoming DNA polymerase. Hence, as well as being crucial to genomic integrity by correctly processing Okazaki fragments, FEN-1 is probably intimately involved in several types of DNA repair. It is highly possible that it plays some part in recombinatorial repair, and perhaps also in physiological recombination events such as V(D)J joining during immunoglobulin gene rearrangements. The hyper-recombination, mutator phenotype of yeast with mutant FEN-1 [17,18] strongly supports this supposition, and highlights a need to look for FEN-1 mutation or dysregulation in human diseases where high recombination rates are diagnostic.

FEN-1 exonuclease activity and nick translation

The exonuclease activity of FEN-1 has been detected in many *in vitro* replication systems. For example, Goulian *et al*., [6] report that a relatively stable and abundant 5'->3' exonuclease attacks 5' phosphate termini of Okazaki fragments to yield 80-95% mononucleotide and 5-20% oligonucleotide products. Calf FEN-1 exonuclease activity is slow on pseudoY structures or if the downstream primer to be excised is separated from the incoming upstream fragment by a gap rather

than a nick. It is probable that under these circumstances, DNA pol ε fills the gap, creating a 5' nicked substrate that is then optimal for FEN-1 exonuclease activity. Upon cleavage of the single terminal 5' base, a gap is again generated, which DNA pol ε then fills. In this way, cycles of gap filling and FEN-1 cleavage can propagate down the DNA in a process known as "nick translation". It is possible that the exonuclease activity is more sensitive than the endonuclease activity to DNA substrate sequence, and it has been suggested that GC-rich regions may represent pause sites for FEN1-DNA polymerase translocation along duplex DNA during nick translation.

Although currently hypothetical, it is possible that the 5' flap endonuclease activity of FEN-1 is more important for RNA primer removal (particularly in the absence of RNaseH1) and for cleavage of DNA bearing bulky lesions, while nick translation may be the more usual mode of FEN-1 action during DNA primer removal. This is probably very important in allowing high fidelity DNA pol ε to replace all DNA made by lower fidelity DNA pol α.

For exonuclease activity, FEN-1 may therefore need to interact with an incoming DNA pol ε, and this interaction may be mediated via PCNA. In support of this notion, exonuclease activity of mFEN-1 is optimal at pH8 with no monovalent salts and 0.1 mM $MgCl_2$ [16], while in the presence of trimeric, DNA-bound PCNA, RF-C and ATP (presumably required to load the PCNA onto DNA), yFEN-1 can tolerate increasing salt concentrations. PCNA is thought to stimulate FEN-1 via direct physical interaction (reviewed in [14,15]). The PCNA probably approaches FEN-1 from the 5' side of the nascent strand in association with DNA pol ε, which synthesises DNA from an upstream primer towards the downstream primer. There has been much recent debate concerning the directionality of PCNA sliding on DNA, but it is now almost certain that the so-called C-terminal face of PCNA is the "forward" face that is presented for interaction with, and stimulation of, FEN-1 (see Figure 4C).

FEN-1 structure can account for its nuclease activity

Comparisons with prokaryotic FEN-1 homologues for which X-ray crystallography data is available suggest that FEN-1 may be a globular protein with a concave active site harbouring two metal ions (presumably Mg^{2+}), with an arch structure above supported by α-helices, and an aperture large enough for single stranded, but not duplex, DNA or RNA to pass through [21]. It is therefore probable that during endonucleolytic cleavage, the 5' terminus of the single strand DNA or RNA flap structure passes through the "arch", and that by sliding down the single stranded region, FEN-1 can then locate the junctional nucleotide in its active site within the concave region below the arch. It can further be predicted that one or both of the two metal ions in the active site will interact with the scissile phosphate of the DNA, stabilising a pentacovalent intermediate and generating an attacking hydroxyl ion that cleaves the phosphodiester bond. The metal ions may also assist in loss of the 3' oxyanion after cleavage. Mutagenesis studies have been carried out on hFEN-1, alanine-substituting those acidic and hydrophobic residues thought to be involved in metal ion binding. Although hFEN-1 mutants D34A, D86A, E160A and D181A could still bind DNA substrate to varying degrees, none was able to carry out cleavage. Alanine mutations in residues D158, D179 and D233 resulted in total loss of substrate binding. These results could suggest that one of the metal ions is involved in nucleophilic attack on the phosphodiester bond, while the other may help to stabilise the structure to allow optimal DNA substrate binding [22].

Role of DNA context in primer removal

As discussed above, FEN-1 can act as a structure-specific endonuclease, cleaving 5' flap structures, or a 5'->3' exonuclease, acting during nick translation. There is an absolute requirement for a free 5' end (either DNA or RNA) for FEN-1 to act either as an endo- or an exo-nuclease. This may have evolved as a mechanism to protect the lagging strand template from unintentional cleavage, as at the replication fork, single stranded template DNA mimics a "flap" structure but without free ends (see Figure 4D). DNA context is therefore probably regulatory for FEN-1 activity, and this context

is determined to no small degree by the action of replicative DNA polymerases. In the case of its endonuclease action, DNA pol ε activity is presumably required *in vivo*, to generate flap structures by displacement synthesis from an upstream primer. Similarly, in exonuclease mode, DNA pol ε must fill the gap created by a single FEN-1 cleavage event in order to regenerate the preferred nicked substrate. Therefore, association of Fen1 with DNA polymerases would be predicted from the known substrate specificities and nuclease activities of FEN1. In agreement with this prediction, calf thymus FEN-1 has been shown to co-purify with DNA pol ε and promote cleavage of an oligonucleotide hybridised to ssM13 DNA [(8)].

In order for DNA pol ε (or any other replicative polymerase) to carry out displacement synthesis, hydrogen bonds must be broken between the template strand nucleotides and those on the strand to be displaced. Duplex melting requires energy that is usually provided from ATP (or dATP) hydrolysis by ATP-dependent DNA helicases. The direction of helicase translocation is described with respect to the template strand to which it is bound, hence a 3'->5' helicase can accompany a DNA polymerase synthesising nascent DNA in a 5'->3' direction. Several 3'->5' helicases have been identified from eukaryotic sources that may therefore be involved in DNA replication, and it is of note that the inherited genetic lesion in cancer-prone Bloom's syndrome patients, and in patients with premature ageing Werner's syndrome, is a defect in a RecQ-type helicase [(23)]. Thus, helicases may soon become as important in the cancer clinic as in the replication laboratory.

Consistent with a role in strand displacement synthesis on the lagging strand, DNA pol ε associates with 3'->5' helicases DNA helicase E, thereby allowing displacement synthesis and the creation of 5' flap structures as FEN-1 substrate. Perhaps most excitingly, yFEN-1 has recently been shown to interact physically and genetically with the replicative helicase encoded by the *Dna2* gene of *S. cerevisiae* [(24)]. Overexpression of yFEN-1 suppresses the DNA replication phenotype of *Dna2-1* temperature sensitive mutants, and conversely, *Dna2* overexpression suppresses *rad27*Δ (yFEN-1

deletion) mutant phenotype. The inviability and arrest phenotype of double mutant of *Dna2-1* and *rad27*Δ support the suggestion that the proteins encoded by these two genes are required in Okazaki fragment processing.

Ligation of mature Okazaki fragments

Following removal of the RNA primer and probably also the DNA pol α-synthesised DNA primer, processed Okazaki fragments must be ligated together to form continuous stretches of DNA without any nicks. Of the described eukaryotic ligases, DNA ligase 1 is the best candidate for Okazaki fragment joining. It is the predominant ligase in proliferating cells and its activity increases on transition from the resting G_0 stage to the cell division cycle. There is an absolute requirement for DNA ligase 1 in formation of covalently closed circular DNA *in vitro* [6], and it is known to join Okazaki fragments efficiently *in vitro* [11]. The patient-derived cell line 46BR with missense mutations in both alleles of DNA ligase 1 shows abnormal joining of Okazaki fragments, and this deficiency in cell extracts can be complemented by addition of DNA ligase 1. Homozygous knockout of the DNA ligase 1 gene in mouse ES cells shows that it is essential for viability, while the yeast homologues, *CDC9* in *S. cerevisiae* and *cdc17* in *S. pombe* are necessary for replication, repair and recombination ([25], and references therein). These findings together strongly suggest that DNA ligase 1 is the ligase responsible for Okazaki fragment joining *in vivo*. Bond formation mediated by DNA ligase 1 is a three-step process: (i) ligase reacts with ATP to form an adenylate complex, releasing pyrophosphate; (ii) this adenylyl group is transferred to the 5' phosphate at the DNA nick; (iii) nucleophilic attack of the 3'OH at the nick onto the activated 5' phosphate results in phosphodiester bind formation with the release of AMP (reviewed by [1]).

The puzzle of how DNA ligase is precisely targeted to the nicks on the lagging strand of DNA has finally started to be unravelled. A direct physical interaction between DNA ligase 1 and PCNA has been reported [25], and although this interaction does not appear to stimulate ligase activity, it may

serve to localise the ligase to the appropriate site on the lagging strand. In support of this, the amino terminal 118 amino acids of DNA ligase 1 are found to be dispensable for catalytic activity but necessary for *in vivo* function, and it is this region that is required for localisation of DNA ligase 1 to sites of DNA replication in the cell nucleus (reviewed in [25]). DNA ligase 1 also shares very limited sequence homology with other PCNA-binding proteins such as FEN-1 and DNA pol ε, within the region required for replication targeting (LSC, manuscript in preparation).

Regulation of Okazaki processing through PCNA?

Taking into account these and other findings, it is possible to envisage a major regulatory role for PCNA in Okazaki fragment processing. Initially PCNA may be present at the Okazaki fragments in conjunction with DNA pol ε and/or FEN-1, but on stalling of displacement synthesis or nick translation, DNA pol ε may dissociate, possibly also resulting in loss of FEN-1. PCNA would remain clamped around the DNA and would serve as a binding site for DNA ligase 1, targeting it to the precise sites where it is required for bond formation. A similar model involving DNA pol δ has been proposed [25]; it is possible that ligation of the much longer leading strands occurs through a similar mechanism to Okazaki fragment joining. Interestingly, on DNA damage, the tumour suppressor protein, p53, transcriptionally induces a cyclin-kinase inhibitor protein, known as $p21^{Cip1}$, that can also bind to PCNA. The region of PCNA bound by $p21^{Cip1}$ is the same as that mapped for FEN-1 binding to PCNA and thought to be important for DNA pol ε and DNA ligase 1 interaction ([15]; LSC, manuscript in preparation). It is therefore possible that a further level of replication control may be exerted through $p21^{Cip1}$ competitive interactions for PCNA binding. The complexity of Okazaki fragment synthesis and processing is such that multiple control points are probable, and it may be predicted that many of these will centre around the PCNA clamp, which remains bound to the DNA throughout processing steps. PCNA may therefore serve to target proteins to their sites of action in a temporally ordered manner.

Model for Okazaki fragment synthesis and processing

A model of the individual but interdependent steps in the synthesis and processing of eukaryotic Okazaki fragments can now be proposed (Figure 5), based on the experimental evidence described above. Firstly, an RNA primer is synthesised by primase p49/p58, then DNA pol α polymerises a DNA oligonucleotide of approximately 5-20 nucleotides onto the RNA primer, creating an RNA-DNA primer. Through an as yet unknown mechanism that may require PCNA and RF-C, it is speculated that the higher fidelity DNA pol ε then substitutes for DNA pol α and synthesises the remainder of the Okazaki fragment up to the adjacent downstream primer. How the downstream RNA primer is removed may depend very much on the DNA context in which it exists, and this in turn is dependent upon the extent of synthesis from upstream primers. The RNA primer may be removed from a downstream Okazaki fragment immediately by RNaseH1 even before synthesis has started from an upstream site. This would result in duplex DNA with a 5' free end adjacent to a single stranded region of DNA. Although FEN-1 exonuclease can cleave this type of substrate, efficiency is low, and further cleavage of the junctional ribonucleotide may need to wait until DNA synthesis from an upstream primer generates a nicked substrate suitable for FEN-1 exonuclease action (Figure 5A). Alternatively, RNaseH1 may not act until the nascent upstream primer has been elongated right up to the downstream primer. Here, RNaseH1 cleavage followed by dissociation of the RNA oligoribonucleotide from the template DNA would leave a gapped molecule, which is presumably filled by DNA pol ε to create a nicked substrate optimal for FEN-1 exonuclease action (Figure 5B). If, however, the RNA primer reanneals with the template DNA after cleavage, a nicked substrate would be generated upon which FEN-1 may be expected to act immediately (Figure 5C). In this case, a single cleavage event would destroy the nicked substrate and create instead a gap, so no further cleavage could take place until DNA pol ε had firstly displaced the RNA fragment then filled the gap. In this way, Fen1 exonuclease activity would never be able to excise larger stretches of DNA without there also being a DNA polymerase adjacent to fill any gaps immediately. Alternatively, DNA pol ε may continue synthesis past the RNA primer,

displacing it by virtue of association with a DNA helicase (helicase E, Ku helicase or Dna2 helicase) and using the parental strand as template for further synthesis. This would result in formation of a 5' flap structure. RNaseH1 may cleave off even single stranded RNA from this flap, leaving a single ribonucleotide and possibly also DNA in the flap structure. This may be removed by FEN-1 endonucleolytic cleavage, with FEN-1 binding first to the terminal 5' nucleotide and then sliding down to the junction where cleavage occurs (Figure 5D). Such nick translation and/or strand displacement mechanisms may permit replacement not only of the RNA component of the Okazaki fragment, but also the DNA component originally synthesised by DNA pol α, thereby ensuring that lagging strand DNA is synthesised with as high fidelity as leading strand DNA.

The choice of mechanism must depend on RNA and DNA context in which the various enzymes operate, but it is becoming increasingly apparent that regulation of such complex processes may be occur by protein-protein interactions and intermolecular competition for such interactions. For example, the leading strand polymerase switch is probably mediated through RF-C/PCNA complex formation at the primer-template junction followed by preferential association of DNA pol δ and hence dissociation of DNA pol α /primase. A similar polymerase switch on the lagging strand may also occur through differential binding of DNA pol ε to PCNA. Indeed, PCNA is thought to be responsible for co-ordinating synthesis of leading and lagging strand DNA. Once at the primer junction, PCNA can associate with either DNA pol ε or FEN-1, perhaps through binding to the same interdomain connector on PCNA; it is possible that competition for PCNA regulates the extent of displacement synthesis and/or nick translation. Additionally, FEN-1 may interact physically with the DNA helicase responsible for permitting displacement synthesis. Finally, DNA ligase can also bind to PCNA, possibly displacing FEN-1. This would prevent further endo- or exo-nucleolytic cleavage, and thereby promote phosphodiester bond formation by DNA ligase 1. Recent data therefore suggest that PCNA may play a vital regulatory role in the synthesis and processing of Okazaki fragments.

The great advances made over the 30 years since the discovery of Okazaki fragments provide cause for optimism that soon we shall be able to extend our understanding of this complex process from the lab into the clinic, to intervene in diseases where human DNA replication control is aberrant.

References

1. Kornberg, A. and Baker, T. (1992) *DNA Replication*. W H Freeman and company.

2. Okazaki, R., Okazaki, T., Sakabe, K. and Sugimoto, K. (1967) *Jpn-J-Med-Sci-Biol*, **20**, 255-260.

3. Ogawa, T. and Okazaki, T. (1980) *Annu Rev Biochem*, **49**, 421-457.

4. Perlman, D. and Huberman, J. A. (1977) *Cell*, **12**, 1029-1043.

5. Copeland, W. C. and Wang, T. S.-F. (1993) *J. Biol. Chem.*, **268**, 26179-26189.

6. Goulian, M., Richards, S. H., Heard, C. J. and Bigsby, B. M. (1990) *J-Biol-Chem*, **265**, 18461-18471.

7. Tsurimoto, T. and Stillman, B. (1990) *Nature*, **346**, 534-539.

8. Siegal, G., Turchi, J. J., Myers, T. W. and Bambara, R. A. (1992) *Proc-Natl-Acad-Sci-U-S-A*, **89**, 9377-9381.

9. Turchi, J. J., Huang, L., Murante, R. S., Kim, Y. and Bambara, R. A. (1994) *Proc-Natl-Acad-Sci-U-S-A*, **91**, 9803-9807.

10. Ishimi, Y., Claude, A., Bullock, P. and Hurwitz, J. (1988) *J-Biol-Chem*, **263**, 19723-33.

11. Waga, S., Bauer, G. and Stillman, B. (1994) *J. Biol. Chem.*, **269**, 10923-10934.

12. Murante, R. S., Rumbaugh, J. A., Barnes, C. J., Norton, J. R. and Bambara, R. A. (1996) *J-Biol-Chem*, **271**, 25888-25897.

13. Huang, L., Rumbaugh, J. A., Murante, R. S., Lin, R. J., Rust, L. and Bambara, R. A. (1996) *Biochemistry*, **35**, 9266-9277.

14. Lieber, M. R. (1997) *Bioessays*, **19**, 233-240.

15. Cox, L. S. (1997) *Trends in Cell Biol.*, **7**, 493-497

16. Harrington, J. J. and Lieber, M. R. (1994) *EMBO J.*, **13**, 1235-1246.

17. Sommers, C. H., Miller, E. J., Dujon, B., Prakash, S. and Prakash, L. (1995) *J. Biol. Chem.*, **9**, 4193-4196.

18. Tishkoff, D. X., Filosi, N., Gaida, G. M. and Kolodner, R. D. (1997) *Cell*, **88**, 253-263.

19. Barnes, C. J., Wahl, A. F., Shen, B., Park, M. S. and Bambara, R. A. (1996) *J-Biol-Chem*, **271**, 29624-31.

20. Klungland, A. and Lindahl, T. (1997) *EMBO-J*, **16**, 3341-3348.

21. Ceska, T. A., Sayers, J. R., Stier, G. and Suck, D. (1996) *Nature*, **382**, 90-3.

22. Shen, B., Nolan, J. P., Sklar, L. A. and Park, M. S. (1997) *Nucleic-Acids-Res*, **25**, 3332-8.

23. Ellis, N. A., Groden, J., Ye, T.-Z., Strauhen, J., Lennon, D. J., Ciocci, S., Proytcheva, M. and German, J. (1995) *Cell*, **83**, 655-666.

24. Budd, M. E. and Campbell, J. L. (1997) *Mol-Cell-Biol*, **17**, 2136-42.

25. Levin, D. S., Bai, W., Yao, N. O. D., M. and Tomkinson, A. E. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 12863-12868.

Figure legends

Figure 1. Okazaki fragments represent short stretches of DNA made very early in the DNA replication process.

Okazaki fragments were first discovered by analysing DNA products obtained from replicating cells pulse labelled with radioactive deoxyribonucleotides on alkaline sucrose gradients[(2)]. This figure shows a typical result from such experiments (see references [2,3,4]). After a very short pulse (5 or 10 seconds), the majority of the label is found in short fragments of DNA that sediment at approximately 4S in eukaryotes, and 9-11S in prokaryotes. With time, the radiolabel appears in a higher molecular weight form of DNA, suggesting that the short fragments have been ligated. On the basis of these types of studies, discontinuous DNA synthesis was proposed [(2,3,4)]; these experiments have led directly to our understanding of semi-discontinuous DNA replication [(3)].

Figure 2: structure of the eukaryotic replication fork

(A) Asymmetry of the replication fork. The leading strand of DNA is synthesised in a 5'->3' direction, by DNA pol δ in association with its accessory factor PCNA. By contract the lagging strand must be synthesised in short, discontinuous fragments, primed with RNA (jagged line) by the action of primase, and then extended into DNA by the primase-associated DNA pol α. DNA pol ε is then thought to substitute for DNA pol α to synthesise the remainder of the fragment. Note that polymerase action is always in the 5'>3' direction, even though the overall direction of nascent DNA synthesis on the lagging strand 3'->5'.

(b) Detailed structure of an Okazaki fragment. The RNA synthesised by primase is shown as a jagged line, while primer DNA synthesised by DNA pol α is shown as a dashed line, and DNA pol ε-produced DNA as a continuous line. (These drawing conventions are used throughout Figures 2-5). "Upstream" is taken to mean 5' of the region of interest on the nascent strand of DNA.

Figure 3. Action of RNaseH1

RNA primers of Okazaki fragments are cleaved endonucleolytically by the enzyme RNaseH1. This enzyme is though to be structure specific, cleaving one nucleotide 5' of the RNA-DNA junction; it acts independently of RNA sequence or length. The majority of the RNA primer is therefore cleaved as an intact fragment, while one ribonucleotide is left behind on the DNA. It has recently been suggested that RNaseH1 may also cut at an RNA-DNA junction even when the RNA is single stranded i.e. not hybridised to template DNA. In this diagram, template DNA is indicated by the darker line.

Figure 4. FEN-1 substrates

A. The 5' flap endonuclease/ 5'(five)->3' exonuclease FEN-1, can cleave 5' DNA flap structures and pseudoY structures.

B. FEN-1 may facilitate removal of damaged DNA during DNA replication or DNA repair. Following a 5' incision by a repair endonuclease, and displacement DNA synthesis by DNA pol ε, a 5' flap structure is created. FEN-1 can bind onto the terminal 5' nucleotide, slide down the flap over the lesion, and cleave at the base of the flap, thereby releasing the damaged DNA fragment. DNA pol ε then fills the gap and ligase will seal the resulting nick.

(C) FEN-1 bound at the base of a 5' flap may interact with the face of PCNA on which the carboxy terminus is located (see ref. [15]), known as the "C face" of PCNA. PCNA may move from an upstream primer region to the flap region in conjunction with DNA pol ε and possibly DNA helicase E[8] or Dna2 helicase [24].

(D) FEN-1 does not cleave lagging strand template DNA. The lagging strand of the replication fork closely resembles a 5' flap or pseudoY structure (see A), according to the extent of unwinding in front of the leading strand synthesis. Although these substrates potentially resemble optimal FEN-1

cleavage substrates, inappropriate cutting of the lagging strand template does not occur, probably because FEN-1 has an absolute requirement to load onto a free 5' end of DNA and slide to the point of cleavage; there are no free ends on the template strand.

Figure 5. Model for Okazaki fragment processing according to DNA context.

The Okazaki fragments are as shown in Figure 2B. Various pathways may be followed to process the RNA and DNA primer prior to ligation by DNA ligase 1.

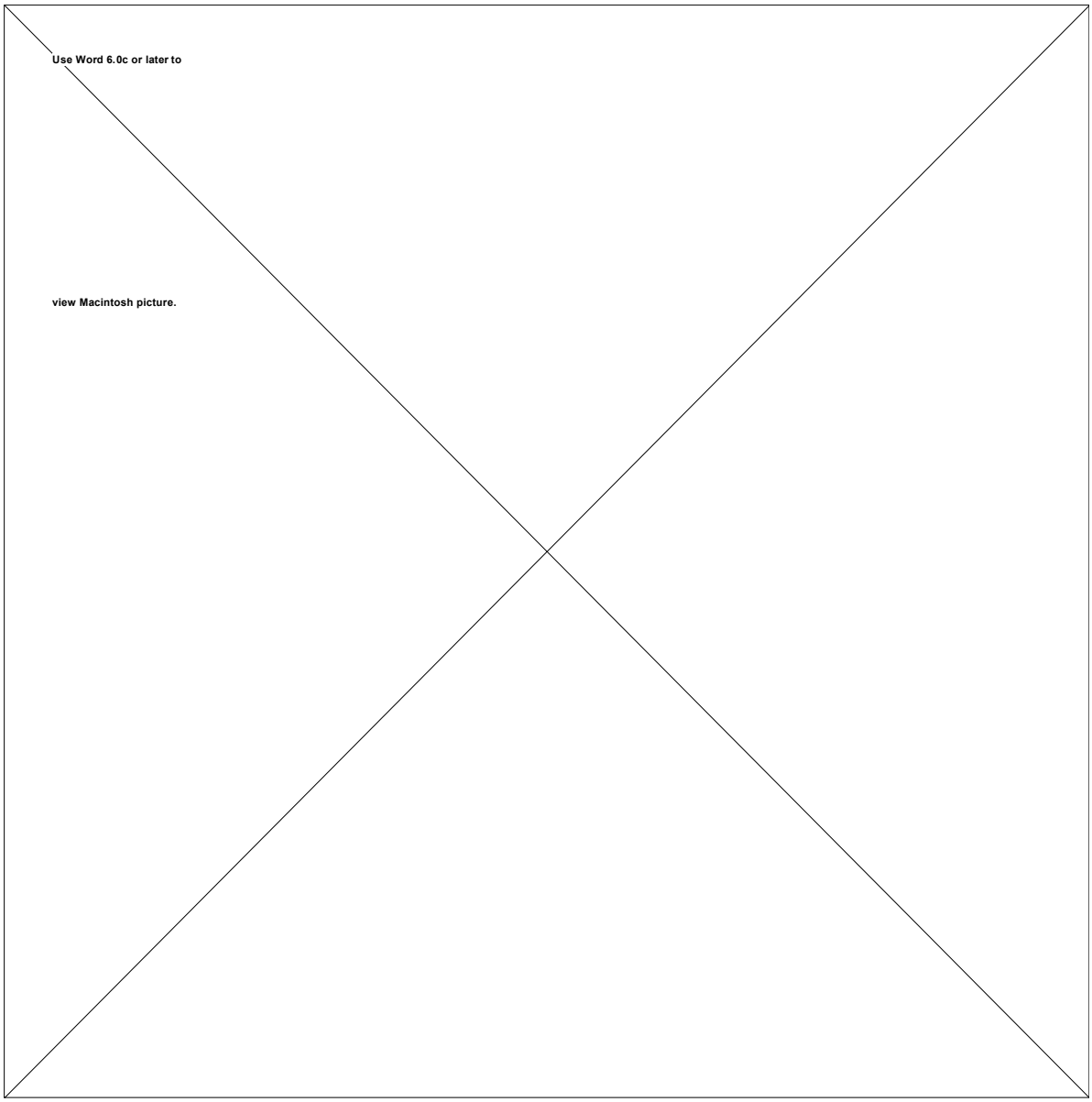
(A) RNaseH1 cleaves the downstream RNA primer before synthesis from the upstream primer has reached this site, leaving a gapped structure. DNA pol ε fills the gap, leaving a 5' nick that FEN-1 acts on exonucleolytically. The nick is propagated or "translated" by successive rounds of FEN-1 cleavage and DNA pol ε gap-filling.

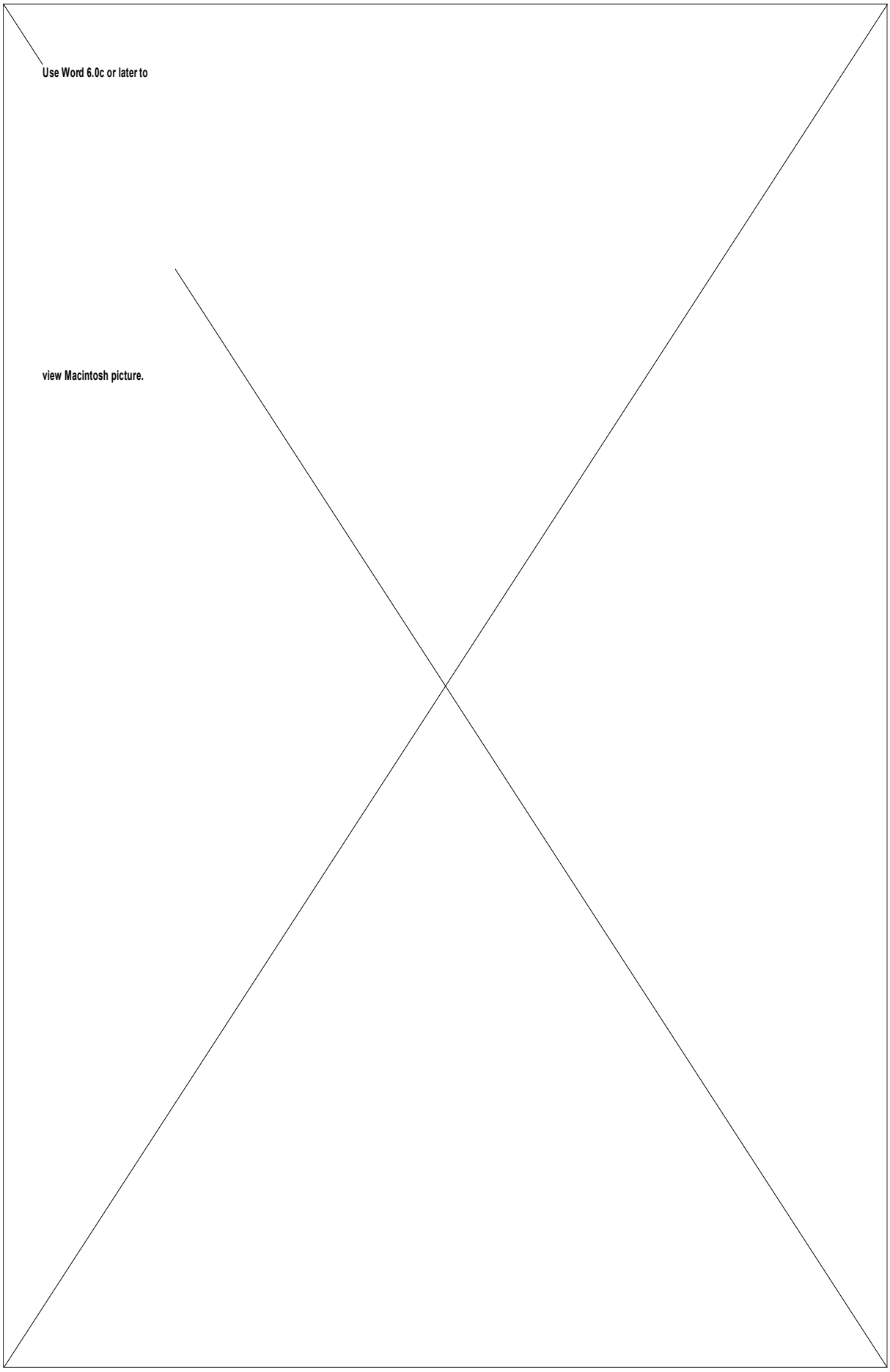
(B) DNA synthesis from an upstream primer displaces the downstream RNA primer, which is cleaved endonucleolytically by RNaseH1. DNA pol ε and FEN-1 then remove the primer DNA by nick translation.

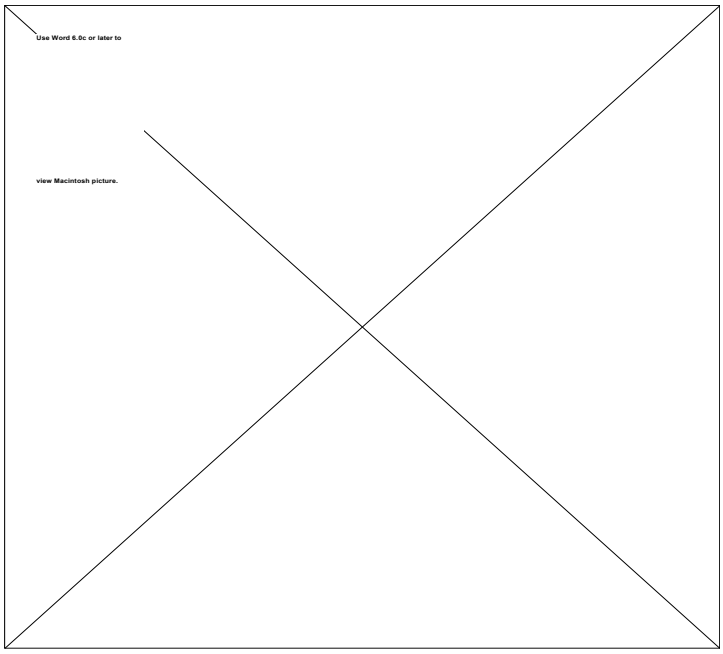
(C) RNA primer is removed prior to extension of DNA from an upstream primer, but the intact RNA fragment reanneals with its template DNA. This RNA must be displaced by incoming DNA pol ε before a suitable nicked substrate for FEN-1 exonuclease activity is generated.

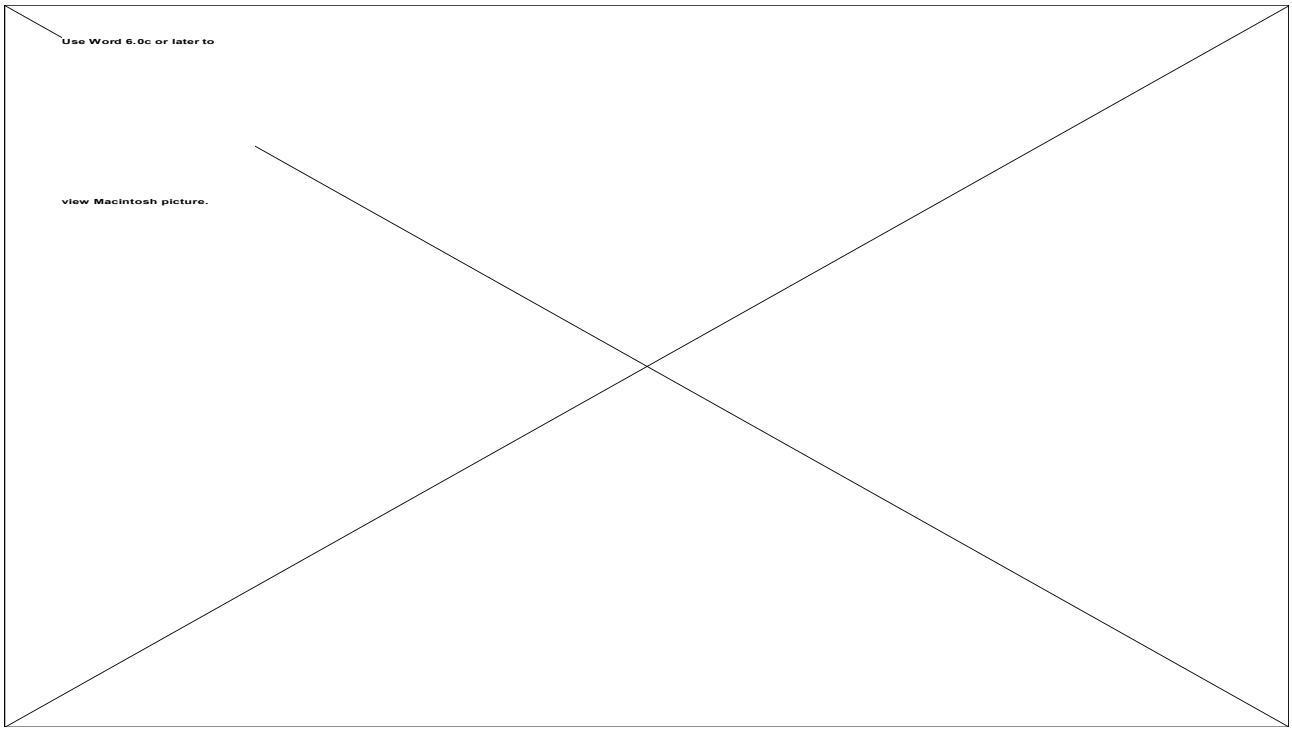
(D) Synthesis from the upstream primer displaces not only the primer RNA, but also some or all of the primer DNA. Even though the RNA component is single stranded, it can be cleaved by RNaseH1, leaving a 5' flap structure upon which FEN--1 acts endonucleolytically. DNA pol ε fills the gap.

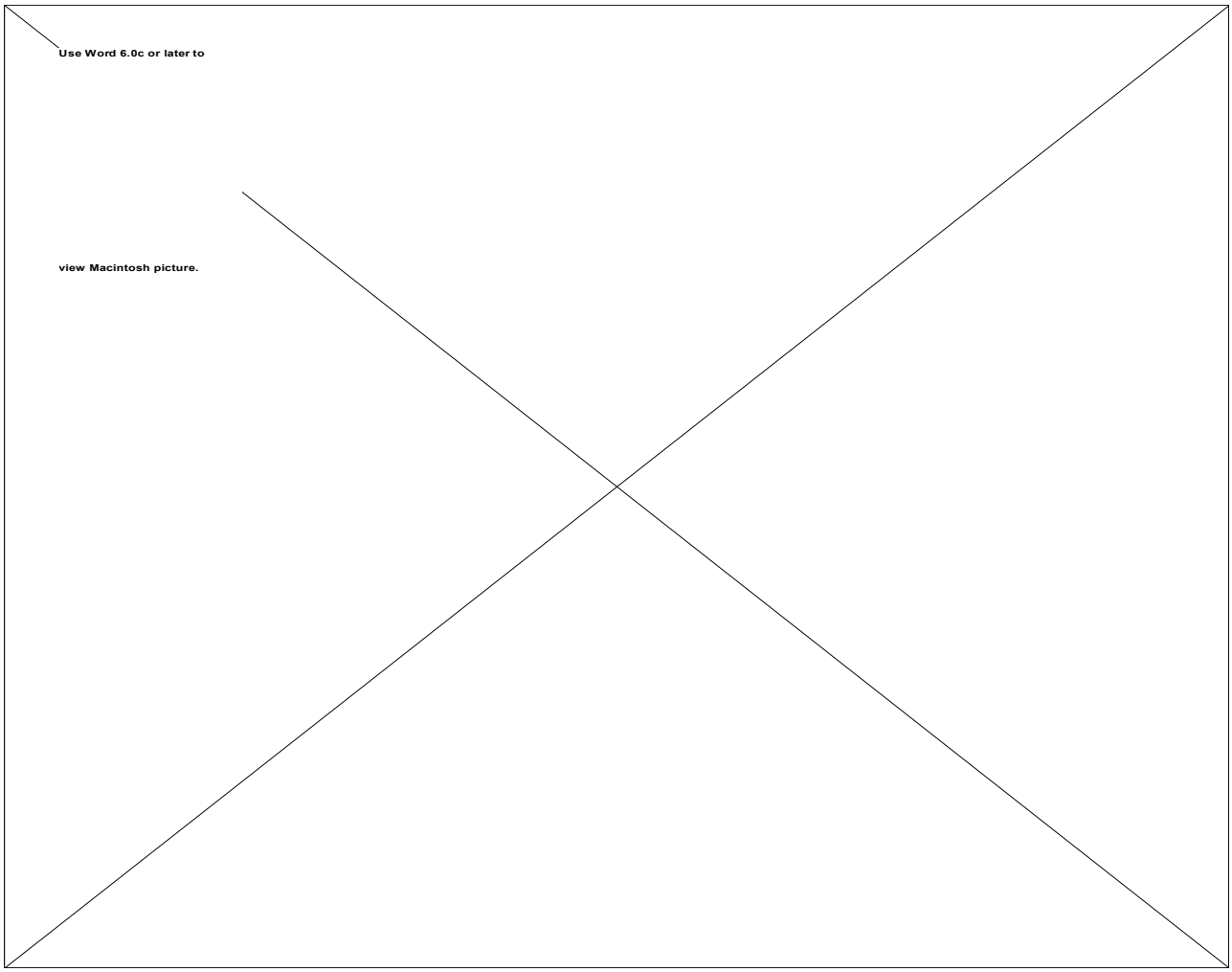
In all cases shown (A-D), DNA ligase 1 completes the Okazaki fragment processing by sealing the nick in the helicase backbone. Not that these model pathways may only represent a proportion of available processing options *in vivo*.


Use Word 6.0c or later to
view Macintosh picture.


Use Word 6.0c or later to
view Macintosh picture.


Use Word 6.0c or later to
view Macintosh picture.


Use Word 6.0c or later to
view Macintosh picture.


Use Word 6.0c or later to
view Macintosh picture.