DNA polymerase iota and related Rad30-like enzymes

John P. McDonald¹, Agnès Tissier¹, Ekaterina G. Frank¹, Shigenori Iwai²,
Fumio Hanaoka^{3,4} and Roger Woodgate^{1*}

¹Section on DNA Replication, Repair and Mutagenesis, National Institute of Child Health and Human Development,
9000 Rockville Pike, Bethesda, MD 20892-2725, USA

²Biomolecular Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565-0874, Japan

³Institute for Molecular and Cellular Biology, Osaka University, and Core Research for Evolutional Science and Technology,
Japan Science and Technology Corporation, 1-3 Yamada-oka, Suita, Osaka 565-0871, Japan

⁴The Institute of Physical and Chemical Research (RIKEN), Wako-shi, Saitama 351-0198, Japan

Until recently, the molecular mechanisms of translesion DNA synthesis (TLS), a process whereby a damaged base is used as a template for continued replication, was poorly understood. This area of scientific research has, however, been revolutionized by the finding that proteins long implicated in TLS are, in fact, DNA polymerases. Members of this so-called UmuC/DinB/RevI/Rad30 superfamily of polymerases have been identified in prokaryotes, eukaryotes and archaea. Biochemical studies with the highly purified polymerases reveal that some, but not all, can traverse blocking lesions in template DNA. All of them share a common feature, however, in that they exhibit low fidelity when replicating undamaged DNA. Of particular interest to us is the Rad30 subfamily of polymerases found exclusively in eukaryotes. Humans possess two Rad30 paralogs, Rad30A and Rad30B. The *RAD30A* gene encodes DNA polymerase η and defects in the protein lead to the xeroderma pigmentosum variant (XP-V) phenotype in humans. Very recently *RAD30B* has also been shown to encode a novel DNA polymerase, designated as Pol ι . Based upon *in vitro* studies, it appears that Pol ι has the lowest fidelity of any eukaryotic polymerase studied to date and we speculate as to the possible cellular functions of such a remarkably error-prone DNA polymerase.

Keywords: DNA polymerase η ; Rad30; Rad30B; DNA polymerase κ ; xeroderma pigmentosum variant

1. A PLETHORA OF POLYMERASES: THE PROKARYOTIC PARADIGM

The identification of genes in *Escherichia coli*, namely *umuDC*, that are required for DNA damage-induced mutagenesis led to the understanding that living organisms possess mechanisms that play an active role in the creation and fixation of mutations (Kato & Shinoura 1977; Steinborn 1978; Walker 1984). These so-called 'DNA-damage tolerance' mechanisms ensure cellular survival by enabling replication to be completed despite the persistence of a normally replication blocking lesion, such as a *cis-syn* cyclobutane pyrimidine dimer (CPD), a 6-4 ultraviolet (UV) photoproduct (6-4PP) or an abasic site (AP) (Koffel-Schwartz *et al.* 1996). The most extensively studied of these 'damage tolerance' mechanisms is translesion DNA synthesis (TLS), or lesion bypass, which in *E. coli* is mediated by the UmuD'₂C protein complex (reviewed in Woodgate & Levine 1996; Smith & Walker 1998). It was originally hypothesized that when DNA polymerase III holoenzyme (Pol III) complex stalls at a site of DNA damage, the Umu complex would function as

a specialized processivity factor to encourage Pol III to replicate past the DNA lesion (Bridges & Woodgate 1984, 1985; and reviewed in Friedberg *et al.* 1995). However, the recent purification of the native UmuD'₂C complex (Bruck *et al.* 1996) and reconstitution of UmuD'₂C-mediated TLS *in vitro* (Tang *et al.* 1998, 1999, 2000) led to the discovery that the UmuD'₂C complex is, in fact, a DNA polymerase iota in its own right. Studies with a maltose-binding protein fused to UmuC suggest that the catalytic activity of this enzyme resides in UmuC and that UmuD' is one of many factors that stimulates its activity (Reuven *et al.* 1999). However, as UmuD'₂C purify as a complex under native conditions (Bruck *et al.* 1996), the complex itself, rather than UmuC alone, has been designated as *E. coli* Pol V (Tang *et al.* 1999). *E. coli* Pol V is highly 'error-prone' when replicating both damaged and undamaged DNA (Tang *et al.* 1999; Maor-Shoshani *et al.* 2000) and is capable of performing translesion replication of an abasic site by preferentially inserting a dAMP (Tang *et al.* 1999; Reuven *et al.* 1999). In addition, Pol V can efficiently bypass a T-T CPD and a T-T 6-4PP (Tang *et al.* 2000). Pol V primarily inserts a dGMP opposite the 3'T of a 6-4PP in good agreement with the known *in vivo* mutational spectra in *E. coli* (Tang *et al.* 2000). These findings have led to a new model of

*Author for correspondence (woodgate@helix.nih.gov).

TLS in which Pol V, at a site of DNA damage, takes over for the stalled Pol III and bypasses the lesion (Woodgate 1999; Goodman 2000). Pol V-dependent replication past the lesion provides an undamaged 3' DNA primer terminus which Pol III can use to resume replication of the genome.

Interestingly, *E. coli* also contains another homologue of *umuC*, called *dinB* (previously called *dinP*; see Ohmori *et al.* 1995) which is involved in untargeted mutagenesis (UTM) of λ -phage and F' plasmids (Brotcorne-Lannoye & Maenhaut-Michel 1986; Kim *et al.* 1997). Overexpression of the *dinB* gene results in an increase in -1 frameshift mutations resulting in single base deletions primarily in homopolymeric runs of bases (Kim *et al.* 1997). Recently, *in vitro* replication assays using a purified histidine-tagged DinB protein revealed that like UmuC, DinB is also a DNA polymerase and has been designated as *E. coli* DNA Pol IV (Wagner *et al.* 1999). Pol IV is a distributive enzyme with a higher degree of fidelity than Pol V (Wagner *et al.* 1999; Tang *et al.* 2000). It is also capable of replicating misaligned primer templates consistent with its role in frameshift mutagenesis (Wagner *et al.* 1999). Unlike Pol V, however, Pol IV preferentially inserts dAMP opposite the 3'T of a 6-4PP, and is incapable of catalysing complete bypass of the CPD, 6-4PP or abasic site lesions (Tang *et al.* 2000).

2. IDENTIFICATION OF UmuC/DinB ORTHOLOGUES AND IN OTHER ORGANISMS

Studies from a number of laboratories revealed that orthologues of *umuDC* existed in the genomes of a variety of bacteria as well as on multidrug-resistance plasmids (Perry and Walker, 1982; Perry *et al.* 1985; Kitagawa *et al.* 1985; Mortelmans & Dousman 1986; Sedgwick *et al.* 1989; Lodwick *et al.* 1990; Smith *et al.* 1990; Thomas *et al.* 1990; Lodwick & Strike 1991; Sedgwick *et al.* 1991). Furthermore, it was known that mutagenesis induced by genotoxic agents occurs not only in eubacteria but also in archaea (Wood *et al.* 1997) and eukaryotes (reviewed in Friedberg *et al.* 1995), suggesting that the process of mutagenic translesion DNA synthesis is evolutionarily conserved. For example, it has been known for some time that the *Saccharomyces cerevisiae* Rev1 protein, which is required for damage-induced mutagenesis, shares a small region of limited similarity to UmuC (Larimer *et al.* 1989).

In our attempt to identify UmuDC orthologues from non-eubacterial origins, we designed degenerate polymerase chain reaction primers to conserved regions of the *umuC* gene and were successful in identifying an orthologue in the Archae *Sulfolobus solfataricus* that is, in fact, slightly more homologous to *dinB* than *umuC* (Kulaeva *et al.* 1996). Searches of Genbank at that time also revealed potential orthologues in the yeast *Saccharomyces cerevisiae* and the nematode *Caenorhabditis elegans* (Kulaeva *et al.* 1996). Based upon these observations, we proposed that there is a superfamily of UmuC/DinB-like proteins which comprises three subfamilies; the UmuC-subfamily, the DinB-subfamily and the Rev1-subfamily. We also noted that members of the UmuC-subfamily of proteins are only present in eubacteria, while the members of the Rev1-subfamily are only present in eukaryotes and

members of the DinB subfamily are present in Archaea, eubacteria and eukaryotes.

When an alignment of all known UmuC-like and DinB-like proteins was used to draw a phylogenetic tree, there was, however, one protein from *S. cerevisiae*, designated D9461 (YDR491w), which did not fit into any of these three subfamilies (Kulaeva *et al.* 1996). This gene was subsequently redesignated as *RAD30* upon the analysis of disruption mutant yeast strains that revealed a mild sensitivity to UV light, and no epistatic relationship with other genes involved in cellular resistance to UV light (McDonald *et al.* 1997; Roush *et al.* 1998). In addition, no defect in UV-induced mutagenesis was observed in *rad30* mutant strains (McDonald *et al.* 1997). These findings indicated that *RAD30* is involved in a novel 'error-free' post-replication repair mechanism of UV-induced DNA damage in yeast (McDonald *et al.* 1997). Based on the similarity of the Rad30 protein to both UmuC and DinB it was suggested that Rad30 might participate in a translesion replication mechanism similar to that found in *E. coli*, but that Rad30-dependent replication of lesions would be 'error-free' by exhibiting higher fidelity than UmuC. Biochemical analysis of purified yeast Rad30 protein revealed that it is a DNA polymerase and was therefore designated as yeast Pol η (Johnson *et al.* 1999b). Expression of a mutant Rad30 protein that lacks polymerase activity in *rad30* null mutant yeast demonstrated that the polymerase activity of Rad30 is necessary for its biological role in resistance to UV light (Johnson *et al.* 1999c). Although Pol η replicates undamaged DNA with a somewhat lower fidelity than other replicative polymerases, it still exhibits a preference for incorporation of the correct nucleotide by a factor of 40- to 2500-fold (Washington *et al.* 1999). Pol η is capable of efficient bypass of a T-T CPD dimer by preferentially inserting the correct dAMP nucleotide opposite both thymines, shedding light on the 'error-free' nature of Pol η 's role in UV resistance (Johnson *et al.* 1999b; Washington *et al.* 1999). In addition, Pol η accurately bypasses the miscoding 8-oxo-7,8-dihydro-deoxyguanine (8-oxoguanine) lesion by predominantly inserting dCMP (Yuan *et al.* 2000). However, Pol η is unable to completely bypass either an abasic site (AP) or an AAF-adducted guanine (AAF-G) but instead inserts a single dGMP or dAMP opposite the abasic site and a single dCMP opposite the AAF-G (Yuan *et al.* 2000).

3. HUMAN ORTHOLOGUES OF *S. CEREVISIAE* Rad30 (Pol η) AND *E. COLI* DinB (Pol IV)

The identification of the novel translesion polymerase, Pol η , in yeast prompted a number of groups to explore the possibility that higher eukaryotes might also harbour orthologues of Pol η . The idea of mammals possessing Pol η orthologues was supported by the fact that cells from humans who exhibit the xeroderma pigmentosum variant (XP-V) phenotype have no defect in nucleotide excision repair but nevertheless have an increased sensitivity to UV light that results in much higher levels of UV-induced mutagenesis (Cleaver & Carter 1973; Wang *et al.* 1993; Raha *et al.* 1996; McGregor *et al.* 1999). Moreover, cell extracts derived from XP-V cells have a severely diminished ability to promote the replication of lesion-

Figure 1. Unrooted phylogenetic tree of the eukaryotic members of the UmuC/DinB/Rev1/Rad30 superfamily of polymerases. A protein alignment of all of the proteins shown was performed using the Clustal X (v. 1.62b) program (Jeanmougin *et al.* 1998). This protein alignment was used to calculate the tree using the Draw N-J Tree algorithm. The unrooted tree was drawn from the calculated tree using the TreeView (v. 1.6) program (Page 1996). The subfamilies are indicated as follows: Rad30, green; DinB, red; Rev1, blue. Sc, *Saccharomyces cerevisiae*; Sp, *Schizosaccharomyces pombe*; Ce, *Caenorhabditis elegans*; Dm, *Drosophila melanogaster*; At, *Arabidopsis thaliana*; Mm, *Mus musculus*; Hs, *Homo sapiens*.

containing plasmids when compared to wild-type cells (Svoboda & Vos 1995; Cordeiro-Stone *et al.* 1997; Svoboda *et al.* 1998; Cordonnier *et al.* 1999; Masutani *et al.* 1999a). These findings led to the suggestion that the XP-V phenotype might arise from a defect in a human Pol η ortholog (Johnson *et al.* 1999b). Indeed, complementation of the XP-V defect in lesion bypass by HeLa cell extracts allowed the purification of a polymerase that restores replication past UV-induced thymine dimers (Masutani *et al.* 1999a). Protein sequencing of the purified translesion polymerase revealed that it was, in fact, an orthologue of yeast Pol η and was therefore designated as human Pol η (Masutani *et al.* 1999b). The official name for the gene is *POLH*, but on occasions, both *XPV* and *RAD30A* synonyms have also been used (Masutani *et al.* 1999b). Reverse transcriptase–polymerase chain reaction amplification and sequencing of the human *POLH* gene from several different XP-V cell lines revealed various inactivating mutations of the Pol η protein, confirming that defects in *POLH* are responsible for the XP-V phenotype (Masutani *et al.* 1999b; Johnson *et al.* 1999a). Like its yeast counterpart, lesion bypass by human Pol η can be characterized as ‘error free’ since it preferentially inserts two dAMP nucleotides opposite a T–T CPD (Masutani *et al.* 1999a, 2000; Johnson *et al.* 2000b).

Another human gene that is a member of this superfamily of translesion polymerases is human *DINB1*, which was identified from the dbEST cDNA database via homology searching for a polypeptide with homology to *E. coli* DinB (Gerlach *et al.* 1999; Ogi *et al.* 1999).

Translation of the full-length cDNA revealed that this cDNA encodes a protein that is indeed more related to *E. coli* DinB (Pol IV) than to yeast or human Pol η . Given its phylogenetic relationship to other polymerases, it is not surprising that *DINB1* also encodes for a DNA polymerase and has been designated as Pol κ by the Human Genome nomenclature committee (Ohashi *et al.* 2000; Friedberg *et al.* 2000), but as Pol θ by Johnson and colleagues (Johnson *et al.* 2000a). As homology would indicate, Pol κ ’s activity is more akin to *E. coli* Pol IV than to Pol η in that it can extend from misaligned primer templates and can bypass an AAF-G adduct located in runs of Gs (Ohashi *et al.* 2000). Human Pol κ can also bypass an abasic site preferentially inserting dAMP and generating –1 deletions (Ohashi *et al.* 2000). However, Pol κ does not bypass a T–T CPD, a T–T 6-4PP or a cisplatin adduct (Ohashi *et al.* 2000; Johnson *et al.* 2000a). These findings suggest that Pol κ , similar to *E. coli* Pol IV, plays a role in frameshift mutagenesis within certain sequence contexts and at certain DNA lesions.

4. IDENTIFICATION OF A SECOND Rad30 ORTHOLOGUE IN HUMANS

Shortly after human *POLH* (*RAD30A*) was shown to be defective in humans with XP-V, we reported that humans contained a second *RAD30* orthologue (McDonald *et al.* 1999). Homology searches of the dbEST cDNA database using *S. cerevisiae* *RAD30* uncovered many partial human expressed sequence tags from various tissue sources that

Figure 2. Ability of Pol ι to incorporate nucleotides at various template sites. Pol ι -dependent (mis)incorporation was measured at each template site in the absence of dNTPs (0) all four dNTPs (4) or each individual dNTP (100 μ M) (G, A, T, C). The sequence context of each template is given above each panel. In most cases, Pol ι inserts the correct nucleotide but the discrimination between the correct and incorrect nucleotide is poor. At template T, G is, in fact, inserted more frequently than the correct base, A. Reproduced with permission from Tissier *et al.* 2000b.

encode proteins homologous to yeast Pol η yet different from human Pol η . Isolation of a full-length cDNA (GenBank accession number AF140501) and sequencing revealed that this cDNA encodes a paralogue of human *RAD30A* (*POLH*) and was therefore designated as *RAD30B* (McDonald *et al.* 1999). Northern analysis of *RAD30B* showed that it is expressed at low levels in many different tissues but is highly expressed in the testis. *In situ* hybridization using the mouse *Rad30b* gene to mouse testis showed that the highest level of expression of *Rad30b* is in postmeiotic round spermatids suggesting a role for Rad30b in bypass of DNA lesions after meiotic DNA replication has been completed (McDonald *et al.* 1999). The *RAD30B* gene was localized to chromosome 18q21 by FISH analysis and the location was confirmed by the fact that a sequence tagged site, R37837, which had previously been mapped to human chromosome 18q21.1 overlapped the sequence of *RAD30B*. Interestingly, the 18q21.1 region of is known to be involved in the aetiology of several types of cancers, suggesting that perhaps *RAD30B* may play a role in the development and/or progression of human cancers (McDonald *et al.* 1999).

Expanding on the previously published phylogenetic tree of the UmuC/DinB/Rev1 superfamily which consisted of 16 proteins and three subfamilies (Kulaeva *et al.* 1996), we presented an updated phylogenetic tree consisting of 36 proteins, including several newly identified Rad30 orthologues, and proposed that Rad30 represents a bona fide fourth subfamily of proteins within this superfamily (McDonald *et al.* 1999). In the short time since we published the 36-protein UmuC/DinB/Rev1/Rad30 phylogenetic tree (McDonald *et al.* 1999) several new members of this family have been identified, most notably in the fruitfly *Drosophila melanogaster*, the genome of which has recently been completely sequenced (Adams *et al.* 2000). In figure 1, we present an updated phylogenetic tree of the presently known eukaryotic members of this superfamily. Note that this tree lacks UmuC-like proteins, which, as reported earlier, appear to be confined

to prokaryotic organisms, and the sole archeal *dinB* homologue that we identified from *S. solfataricus* (Kulaeva *et al.* 1996). Interestingly, whereas the yeast *S. cerevisiae* possesses only Rad30 (Pol η) and Rev1, the yeast *Schizosaccharomyces pombe* possesses Rad30, Rev1 and DinB (Pol IV) orthologues. Despite the fact that two lower eukaryotes, *S. pombe* and *C. elegans*, and higher eukaryotes such as mice and humans possess DinB orthologues, the fruitfly *D. melanogaster* does not. In contrast, all six of these eukaryotic organisms possess both Rad30 and Rev1 orthologues. These observation suggests that perhaps there is less selective pressure to maintain a DinB orthologue than a Rad30 or a Rev1 orthologue. It is also worth noting that lower eukaryotes such as *S. cerevisiae*, *S. pombe* and *C. elegans* contain only a single *RAD30*-like gene, while higher eukaryotes such as *D. melanogaster*, mice and humans possess two Rad30 orthologues, Rad30A and Rad30B, indicating that these paralogues were duplicated some time prior to the evolution of *Drosophila*.

5. HUMAN Pol ι (Rad30B): AN IMMORAL-AMORAL POLYMERASE

Given that the prototype of each of the four subfamilies, *E. coli* Pol V (UmuC), *E. coli* Pol IV (DinB), *S. cerevisiae* Rev1 and *S. cerevisiae* (Pol η) Rad30, are DNA polymerases, it was reasonable to hypothesize that human Rad30B would also possess polymerase activity (McDonald *et al.* 1999). In order to test this hypothesis, full-length Rad30B protein was purified as a GST fusion from baculovirus injected insect cells and added to primer extension replication assays (Tissier *et al.* 2000b). As expected, human Rad30B exhibited polymerase activity and has been renamed DNA polymerase iota (Pol ι) (Tissier *et al.* 2000b). Like other members of this superfamily of polymerases, Pol ι is a distributive enzyme with usually only one to two bases being inserted per enzyme binding extension event. In most cases, Pol ι replicates DNA in a template-directed manner, with the correct

Figure 3. Ability of Pol ι to bypass various DNA lesions. Each template was incubated with 0, 2.5 and 10 nM GST-Pol ι for 15 min at 37 °C. The DNA sequence context of each template is given above each panel. The left most panel depicts replication of undamaged DNA. The remaining tracks show replication of five common DNA lesions. CPD, a thymine–thymine cyclobutane pyrimidine dimer; 6-4, a thymine–thymine 6-4 photoproduct; AAF, an acetyl aminofluorene guanine adduct; CisPI, a 1,2 cisplatinated guanine adduct; Ab, a synthetic tetrahydrofuran abasic lesion.

‘Watson–Crick’ base pair being inserted opposite the template (figure 2). However, Pol ι has a very low discrimination between right and wrong nucleotides with frequent mis-incorporations occurring at most template sites (figure 2). At these template sites, we consider that Pol ι is an ‘immoral’ DNA polymerase, in that it knows which base should be correctly inserted, but does not appear to care about this fact and makes frequent errors. Like most rules, there are exceptions and in the case of Pol ι it is at template T, where it appears to insert G opposite T just as efficiently as the correct base A (figure 2). At this site, Pol ι is an ‘amoral’ polymerase, not knowing the difference between right and wrong nucleotides. Quantitative analysis of the propensity of Pol ι to incorporate nucleotides at undamaged template sites revealed striking differences in mis-incorporation frequencies that were template dependent. For example, Pol ι replicates template A with the greatest fidelity with mis-incorporation frequencies of only $1\text{--}2 \times 10^{-4}$. At template C or G, Pol ι mis-incorporation frequencies are nearly 100-fold higher, averaging 1×10^{-2} . At template T, mis-insertion frequencies were higher still, with the mis-incorporation of G opposite T in fact occurring roughly threefold better than that of the correct base, A, and mis-incorporation of T opposite T at a frequency of approximately 6.7×10^{-1} .

These mis-insertions are, however, only biologically relevant if they are extended and the mispaired base fixed as a mutation. Indeed, Pol ι appears able to extend some mispairs efficiently. For example, after efficient mis-incorporation of a G opposite template T, Pol ι is able to elongate this G:T mispair by inserting the correct next nucleotide of G opposite template C (figure 2).

Given the fact that Pol ι is related to Pol η , and Pol η has been shown to bypass certain types of DNA damage (Masutani *et al.* 1999b, 2000; Johnson *et al.* 2000b; Vaisman *et al.* 2000), we were obviously interested in determining the ability of Pol ι to facilitate translesion DNA synthesis (figure 3). We find that when given all four dNTPs, Pol ι can, in fact, perform limited bypass of a

T–T CPD. In contrast to the CPD, however, Pol ι readily mis-inserts two nucleotides opposite the T–T 6-4PP, but lesion bypass occurs only a fraction of the time of that seen with the CPD. Similarly, Pol ι mis-inserts nucleotides opposite an abasic site and a small amount of bypass is observed. In contrast, Pol ι appears to insert a base opposite the AAF-G adduct, but does not bypass it and very little Pol ι -dependent mis-incorporation opposite adjacent cisplatinated Gs is observed. These findings suggest that Pol ι can mis-insert bases opposite certain DNA lesions and, in some cases, perform complete unassisted translesion DNA synthesis. A more detailed analysis of Pol ι -dependent lesion bypass is presented elsewhere (Tissier *et al.* 2000a).

6. POSSIBLE CELLULAR ROLES OF HUMAN Pol ι

From our biochemical characterization of human Pol ι , we know that Pol ι is a highly ‘error-prone’ polymerase which mis-incorporates bases in a template-dependent manner in undamaged DNA and then extends these mispaired bases (Tissier *et al.* 2000b). Furthermore, we know that Pol ι can bypass certain DNA lesions and, based on its low fidelity on undamaged DNA, it is likely that translesion replication mediated by Pol ι is also ‘error prone’ (Tissier *et al.* 2000a). It would, therefore, seem rather dangerous for a cell to possess such an extraordinarily mutagenic polymerase unless it were tightly regulated or might perform a specialized function with limited scope. We know that *POL*I orthologues are evolutionarily conserved in higher eukaryotes from *Drosophila* to humans (figure 1), implying that Pol ι presumably provides some selective advantage. The cellular function of Pol ι remains, however, to be determined. Perhaps we can conjecture as to its function based upon our preliminary characterization of its properties. We already know of two levels of regulation that the cell uses to restrict the amount of Pol ι . First, Northern analysis reveals that the *POL*I gene, although expressed in many cell types, is

Figure 4. Asymmetry of Pol ι -dependent spontaneous mutagenesis. Quantitative analysis of Pol ι misincorporation reveals that there is a striking asymmetry in mutagenesis that is template dependent (Tissier *et al.* 2000b). At template A, misincorporations occur relatively infrequently and as a consequence there are few mutations from T. In contrast, at template T there are frequent misincorporations of G or T that result in a large number of mutations from A. Interestingly, the asymmetry is strikingly similar to that observed during somatic hypermutation of IgV genes.

expressed at low levels in most tissues, with the notable exception being the testis, suggesting that the intracellular levels of Pol ι are likely to be limited (McDonald *et al.* 1999). Second, our replication assays reveal that Pol ι is a distributive enzyme, so that its error-prone mode of replication will be restricted to short patches of DNA.

One possible role for Pol ι is that it may function as a backup to Pol η in translesion synthesis of UV-induced CPDs. Pol η has clearly evolved to replicate this type of DNA damage accurately and efficiently (Johnson *et al.* 2000b; Masutani *et al.* 2000). However, even in XP-V cells devoid of Pol η , there is a low level of residual translesion DNA synthesis (Cordonnier *et al.* 1999). Furthermore, this replication appears highly inaccurate as XP-V cells exhibit increased UV-induced mutagenesis and an altered spectra of mutation (Wang *et al.* 1993; Raha *et al.* 1996; McGregor *et al.* 1999). Interestingly, our preliminary characterization of Pol ι -dependent mis-insertion opposite a T–T CPD suggests that it is remarkably similar to the spectra of mutations recovered in XP-V cells (Tissier *et al.* 2000a). Thus, mutagenic bypass replication promoted by Pol ι is likely to induce the mutations which lead to the skin cancers associated with the XP-V defect. Moreover, inappropriate expression of this ‘error-prone’ polymerase will also probably result in higher levels of spontaneous mutagenesis, that in many cell types will lead to increased carcinogenesis.

The possibility that Pol ι functions as a back-up to Pol η in XP-V cells, and is therefore responsible for some UV-induced skin cancers and perhaps other types cancers, can hardly be interpreted as providing a selective advantage. So, is there a crucial biological system in which a low-fidelity polymerase, like Pol ι , would be an integral component? One such mechanism is that of the

so-called ‘somatic hypermutability’ of the variable region of the human immunoglobulin (IgV) genes (reviewed in Harris *et al.* 1999; Spencer & Dunn-Walters 1999). This ‘hypermutation’ is thought to provide the necessary diversity in antibody genes so as to enable the body to produce an efficient immune response to virtually an unlimited array of antigens (see other manuscripts in this issue for more details). Indeed, the ability to produce such a broad immune response in order to ward off potential pathogens can certainly be thought of as providing a selective advantage. Although the molecular mechanisms that lead to somatic hypermutation remain to be resolved, it is hypothesized to result from an error-prone DNA polymerase that increases the mutation rate in IgV genes nearly one millionfold over the rest of the genome (Diaz & Flajnik 1998; Jacobs *et al.* 1998; Winter *et al.* 1998; Diaz *et al.* 1999; Kim *et al.* 1999). An error frequency of 10^{-2} to 10^{-3} (Yelamos *et al.* 1995; Smith *et al.* 1996) is certainly well within the range of Pol ι (Tissier *et al.* 2000b). Perhaps the most telling link between Pol ι and somatic hypermutation is that mis-incorporation by Pol ι is similar to the mutation spectra observed during somatic hypermutation. For example, it is known that somatic hypermutability results in an asymmetry in mutations. In particular mutations from T occur much less frequently than the complementary A (Yelamos *et al.* 1995; Smith *et al.* 1996). Recall that Pol ι has a preference for insertion of dTMP and therefore replicates template A very accurately. This would result in relatively few mutations from T (figure 4). On the other hand, Pol ι has significantly lower fidelity when replicating template T, which results in many more mutations from A (figure 4). Thus, the preliminary *in vitro* characterization of Pol ι suggest that it might account for many of the features of ‘somatic hypermutation’. While these observations alone do not demonstrate that Pol ι is the fabled error-prone DNA polymerase that carries out somatic hypermutation, it certainly warrants further study and experiments are currently in progress to test this intriguing hypothesis.

This work was supported in part by the National Institutes of Health intramural research programme and by a Core Research for Evolutional Science and Technology grant to F.H. We would like to thank all of the members of the Section on DNA Replication, Repair and Mutagenesis for their helpful comments and suggestions during the course of this work.

REFERENCES

- Adams, M. D. (and 194 others) 2000 The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195.
- Bridges, B. A. & Woodgate, R. 1984 Mutagenic repair in *Escherichia coli*. X. The *umuC* gene product may be required for replication past pyrimidine dimers but not for the coding error in UV mutagenesis. *Mol. Gen. Genet.* **196**, 364–366.
- Bridges, B. A. & Woodgate, R. 1985 Mutagenic repair in *Escherichia coli*: products of the *recA* gene and of the *umuD* and *umuC* genes act at different steps in UV-induced mutagenesis. *Proc. Natl Acad. Sci. USA* **82**, 4193–4197.
- Brotcorne-Lannoye, A. & Maenhaut-Michel, G. 1986 Role of RecA protein in untargeted UV mutagenesis of bacteriophage λ : evidence for the requirement for the *dinB* gene. *Proc. Natl Acad. Sci. USA* **83**, 3904–3908.

- Bruck, I., Woodgate, R., McEntee, K. & Goodman, M. F. 1996 Purification of a soluble UmuD'C complex from *Escherichia coli*: cooperative binding of UmuD'C to single-stranded DNA. *J. Biol. Chem.* **271**, 10 767–10 774.
- Cleaver, J. E. & Carter, D. M. 1973 Xeroderma pigmentosum variants: influence of temperature on DNA repair. *J. Invest. Dermatol.* **60**, 29–32.
- Cordeiro-Stone, M., Zaritskaya, L. S., Price, L. K. & Kaufmann, W. K. 1997 Replication fork bypass of a pyrimidine dimer blocking leading strand DNA synthesis. *J. Biol. Chem.* **272**, 13 945–13 954.
- Cordonnier, A. M., Lehmann, A. R. & Fuchs, R. P. P. 1999 Impaired translesion synthesis in xeroderma pigmentosum variant extracts. *Mol. Cell. Biol.* **19**, 2206–2211.
- Diaz, M. & Flajnik, M. F. 1998 Evolution of somatic hypermutation and gene conversion in adaptive immunity. *Immunol. Rev.* **162**, 13–24.
- Diaz, M., Velez, J., Singh, M., Cerny, J. & Flajnik, M. F. 1999 Mutational pattern of the nurse shark antigen receptor gene (NAR) is similar to that of mammalian Ig genes and to spontaneous mutations in evolution: the translesion synthesis model of somatic hypermutation. *Int. Immunol.* **11**, 825–833.
- Friedberg, E. C., Walker, G. C. & Siede, W. 1995 *DNA repair and mutagenesis*. Washington, DC: American Society for Microbiology.
- Friedberg, E. C., Feaver, W. J. & Gerlach, V. L. 2000 The many faces of DNA polymerases: strategies for mutagenesis and for mutational avoidance. *Proc. Natl Acad. Sci. USA* **97**, 5681–5683.
- Gerlach, V. L., Aravind, L., Gotway, G., Schultz, R. A., Koonin, E. V. & Friedberg, E. C. 1999 Human and mouse homologs of *Escherichia coli* DinB (DNA polymerase IV), members of the UmuC/DinB superfamily. *Proc. Natl Acad. Sci. USA* **96**, 11 922–11 927.
- Goodman, M. F. 2000 Coping with replication 'train wrecks' in *Escherichia coli* using Pol V, Pol II and RecA proteins. *Trends Biochem. Sci.* **25**, 189–195.
- Harris, R. S., Kong, Q. & Maizels, N. 1999 Somatic hypermutation and the three Rs: repair, replication and recombination. *Mutat. Res.* **436**, 157–178.
- Jacobs, H. (and 13 others) 1998 Hypermutation of immunoglobulin genes in memory B cells of DNA repair-deficient mice. *J. Exp. Med.* **187**, 1735–1743.
- Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. & Gibson, T. J. 1998 Multiple sequence alignment with Clustal X. *Trends Biochem. Sci.* **23**, 403–405.
- Johnson, R. E., Kondratik, C. M., Prakash, S. & Prakash, L. 1999a *hRAD30* mutations in the variant form of xeroderma pigmentosum. *Science* **285**, 263–265.
- Johnson, R. E., Prakash, S. & Prakash, L. 1999b Efficient bypass of a thymine–thymine dimer by yeast DNA polymerase, pol η . *Science* **283**, 1001–1004.
- Johnson, R. E., Prakash, S. & Prakash, L. 1999c Requirement of DNA polymerase activity of yeast Rad30 protein for its biological function. *J. Biol. Chem.* **274**, 15 975–15 977.
- Johnson, R. E., Prakash, S. & Prakash, L. 2000a The human *DINB1* gene encodes the DNA polymerase Pol θ . *Proc. Natl Acad. Sci. USA* **97**, 3838–3843.
- Johnson, R. E., Washington, M. T., Prakash, S. & Prakash, L. 2000b Fidelity of human DNA polymerase η . *J. Biol. Chem.* **275**, 7447–7450.
- Kato, T. & Shinoura, Y. 1977 Isolation and characterization of mutants of *Escherichia coli* deficient in induction of mutations by ultraviolet light. *Mol. Gen. Genet.* **156**, 121–131.
- Kim, N., Bozek, G., Lo, J. C. & Storb, U. 1999 Different mismatch repair deficiencies all have the same effects on somatic hypermutation: intact primary mechanism accompanied by secondary modifications. *J. Exp. Med.* **190**, 21–30.
- Kim, S. R., Maenhaut-Michel, G., Yamada, M., Yamamoto, Y., Matsui, K., Sofuni, T., Nohmi, T. & Ohmori, H. 1997 Multiple pathways for SOS-induced mutagenesis in *Escherichia coli*: an overexpression of *dinB/dinP* results in strongly enhancing mutagenesis in the absence of any exogenous treatment to damage DNA. *Proc. Natl Acad. Sci. USA* **94**, 13 792–13 797.
- Kitagawa, Y., Akaboshi, E., Shinagawa, H., Horii, T., Ogawa, H. & Kato, T. 1985 Structural analysis of the *umu* operon required for inducible mutagenesis in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **82**, 4336–4340.
- Koffel-Schwartz, N., Coin, F., Veaute, X. & Fuchs, R. P. P. 1996 Cellular strategies for accommodating replication-hindering adducts in DNA: control by the SOS response in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **93**, 7805–7810.
- Kulaeva, O. I., Koonin, E. V., McDonald, J. P., Randall, S. K., Rabinovich, N., Connaughton, J. F., Levine, A. S. & Woodgate, R. 1996 Identification of a DinB/UmuC homolog in the archeon *Sulfolobus solfataricus*. *Mutat. Res.* **357**, 245–253.
- Larimer, F. W., Perry, J. R. & Hardigree, A. A. 1989 The *REVI* gene of *Saccharomyces cerevisiae*: isolation, sequence and functional analysis. *J. Bacteriol.* **171**, 230–237.
- Lodwick, D., Owen, D. & Strike, P. 1990 DNA sequence analysis of the *imp* UV protection and mutation operon of the plasmid TP110: identification of a third gene. *Nucleic Acids Res.* **18**, 5045–5050.
- Lodwick, D. & Strike, P. 1991 Distribution of sequences homologous to the *impCAB* operon of TP110 among bacterial plasmids of different incompatibility groups. *Mol. Gen. Genet.* **229**, 27–30.
- McDonald, J. P., Levine, A. S. & Woodgate, R. 1997 The *Saccharomyces cerevisiae* *RAD30* gene, a homologue of *Escherichia coli* *dinB* and *umuC*, is DNA damage inducible and functions in a novel error-free postreplication repair mechanism. *Genetics* **147**, 1557–1568.
- McDonald, J. P., Rapic-Otrin, V., Epstein, J. A., Broughton, B. C., Wang, X., Lehmann, A. R., Wolgemuth, D. J. & Woodgate, R. 1999 Novel human and mouse homologs of *Saccharomyces cerevisiae* DNA polymerase η . *Genomics* **60**, 20–30.
- McGregor, W. G., Wei, D., Maher, V. M. & McCormick, J. J. 1999 Abnormal, error-prone bypass of photoproducts by xeroderma pigmentosum variant cell extracts results in extreme strand bias for the kinds of mutations induced by UV light. *Mol. Cell Biol.* **19**, 147–154.
- Maor-Shoshani, A., Reuven, N. B., Tomer, G. & Livneh, Z. 2000 Highly mutagenic replication by DNA polymerase V (UmuC) provides a mechanistic basis for SOS untargeted mutagenesis. *Proc. Natl Acad. Sci. USA* **97**, 565–570.
- Masutani, C., Araki, M., Yamada, A., Kusumoto, R., Nogimori, T., Maekawa, T., Iwai, S. & Hanaoka, F. 1999a Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. *EMBO J.* **18**, 3491–3501.
- Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M., Araki, M., Iwai, S., Takio, K. & Hanaoka, F. 1999b The *XPV* (xeroderma pigmentosum variant) gene encodes human DNA polymerase η . *Nature* **399**, 700–704.
- Masutani, C., Kusumoto, R., Iwai, S. & Hanaoka, F. 2000 Mechanisms of accurate translesion synthesis by human DNA polymerase η . *EMBO J.* **19**, 3100–3109.
- Mortelmans, K. E. & Dousman, L. 1986 Mutagenesis and plasmids In *Chemical mutagens: principles and methods for their detection*, vol. 10 (ed. F. de Serres), pp. 469–507. New York: Plenum Press.
- Ogi, T., Kato, T. & Ohmori, H. 1999 Mutation enhancement by DINB1, a mammalian homologue of the *Escherichia coli* mutagenesis protein DinB. *Genes Cells* **4**, 607–618.

- Ohashi, E., Ogi, T., Kusumoto, R., Iwai, S., Masutani, C., Hanaoka, F. & Ohmori, H. 2000 Error-prone bypass of certain DNA lesions by the human DNA polymerase κ . *Genes Dev.* **14**, 1589–1594.
- Ohmori, H., Hatada, E., Qiao, Y., Tsuji, M. & Fukuda, R. 1995 *dinP*, a new gene in *Escherichia coli*, whose product shows similarities to UmuC and its homologues. *Mutat. Res.* **347**, 1–7.
- Page, R. D. 1996 TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12**, 357–358.
- Perry, K. L. & Walker, G. C. 1982 Identification of plasmid (pKM101) coded proteins involved in mutagenesis and UV resistance. *Nature* **300**, 278–281.
- Perry, K. L., Elledge, S. J., Mitchell, B., Marsh, L. & Walker, G. C. 1985 *umuDC* and *mucAB* operons whose products are required for UV light and chemical-induced mutagenesis: UmuD, MucA, and LexA products share homology. *Proc. Natl Acad. Sci. USA* **82**, 4331–4335.
- Raha, M., Wang, G., Seidman, M. M. & Glazer, P. M. 1996 Mutagenesis by third-strand-directed psoralen adducts in repair-deficient human cells: high frequency and altered spectrum in a xeroderma pigmentosum variant. *Proc. Natl Acad. Sci. USA* **93**, 2941–2946.
- Reuven, N. B., Arad, G., Maor-Shoshani, A. & Livneh, Z. 1999 The mutagenesis protein UmuC is a DNA polymerase activated by UmuD', RecA, and SSB and is specialized for translesion replication. *J. Biol. Chem.* **274**, 31763–31766.
- Roush, A. A., Suarez, M., Friedberg, E. C., Radman, M. & Siede, W. 1998 Deletion of the *Saccharomyces cerevisiae* gene *RAD30* encoding an *Escherichia coli* DinB homolog confers UV radiation sensitivity and altered mutability. *Mol. Gen. Genet.* **257**, 686–692.
- Sedgwick, S. G., Thomas, S. M., Hughes, V., Lodwick, D. & Strike, P. 1989 Mutagenic DNA repair genes on plasmids from the 'pre-antibiotic era'. *Mol. Gen. Genet.* **218**, 323–329.
- Sedgwick, S. G., Ho, C. & Woodgate, R. 1991 Mutagenic DNA repair in Enterobacteria. *J. Bacteriol.* **173**, 5604–5611.
- Smith, B. T. & Walker, G. C. 1998 Mutagenesis and more: *umuDC* and the *Escherichia coli* SOS response. *Genetics* **148**, 1599–1610.
- Smith, C. M., Koch, W. H., Franklin, S. B., Foster, P. L., Cebula, T. A. & Eisenstadt, E. 1990 Sequence analysis of the *Salmonella typhimurium* LT2 *umuDC* operon. *J. Bacteriol.* **172**, 4964–4978.
- Smith, D. S., Creadon, G., Jena, P. K., Portanova, J. P., Kotzin, B. L. & Wysocki, L. J. 1996 Di- and trinucleotide target preferences of somatic mutagenesis in normal and auto-reactive B cells. *J. Immunol.* **156**, 2642–2652.
- Spencer, J. O. & Dunn-Walters, D. K. 1999 Somatic hypermutation and B-cell malignancies. *J. Pathol.* **187**, 158–163.
- Steinborn, G. 1978 *Uvm* mutants of *Escherichia coli* K12 deficient in UV mutagenesis. I. Isolation of *uvm* mutants and their phenotypical characterization in DNA repair and mutagenesis. *Mol. Gen. Genet.* **165**, 87–93.
- Svoboda, D. L. & Vos, J. M. 1995 Differential replication of a single, UV-induced lesion in the leading or lagging strand by a human cell extract: fork uncoupling or gap formation. *Proc. Natl Acad. Sci. USA* **92**, 11975–11979.
- Svoboda, D. L., Briley, L. P. & Vos, J. M. 1998 Defective bypass replication of a leading strand cyclobutane thymine dimer in xeroderma pigmentosum variant cell extracts. *Cancer Res.* **58**, 2445–2448.
- Tang, M., Bruck, I., Eritja, R., Turner, J., Frank, E. G., Woodgate, R., O'Donnell, M. & Goodman, M. F. 1998 Biochemical basis of SOS-induced mutagenesis in *Escherichia coli*: reconstitution of *in vitro* lesion bypass dependent on the UmuD'₂C mutagenic complex and RecA. *Proc. Natl Acad. Sci. USA* **95**, 9755–9760.
- Tang, M., Shen, X., Frank, E. G., O'Donnell, M., Woodgate, R. & Goodman, M. F. 1999 UmuD'₂C is an error-prone DNA polymerase, *Escherichia coli*, DNA pol V. *Proc. Natl Acad. Sci. USA* **96**, 8919–8924.
- Tang, M., Pham, P., Shen, X., Taylor, J.-S., O'Donnell, M., Woodgate, R. & Goodman, M. F. 2000 Roles of *E. coli* DNA polymerases IV and V in lesion-targeted and untargeted SOS mutagenesis. *Nature* **404**, 1014–1018.
- Thomas, S. M., Crowne, H. M., Pidsley, S. C. & Sedgwick, S. G. 1990 Structural characterization of the *Salmonella typhimurium* LT2 *umu* operon. *J. Bacteriol.* **172**, 4979–4987.
- Tissier, A., Frank, E. G., McDonald, J. P., Iwai, S., Hanaoka, F. & Woodgate, R. 2000a Misinsertion and bypass of thymine–thymine dimers by human DNA polymerase ι . *EMBO J.* **19**, 5259–5266.
- Tissier, A., McDonald, J. P., Frank, E. G. & Woodgate, R. 2000b polt, a remarkably error-prone human DNA polymerase. *Genes Dev.* **14**, 1642–1650.
- Vaisman, A., Masutani, C., Hanaoka, F. & Chaney, S. G. 2000 Efficient translesion replication past oxaliplatin and cisplatin GpG adducts by human DNA polymerase η . *Biochemistry* **39**, 4575–4580.
- Wagner, J., Gruz, P., Kim, S. R., Yamada, M., Matsui, K., Fuchs, R. P. P. & Nohmi, T. 1999 The *dinB* gene encodes a novel *Escherichia coli* DNA polymerase (DNA pol IV) involved in mutagenesis. *Mol. Cell.* **4**, 281–286.
- Walker, G. C. 1984 Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. *Microbiol. Rev.* **48**, 60–93.
- Wang, Y. C., Maher, V. M., Mitchell, D. L. & McCormick, J. J. 1993 Evidence from mutation spectra that the UV hypermutability of xeroderma pigmentosum variant cells reflects abnormal, error-prone replication on a template containing photoproducts. *Mol. Cell. Biol.* **13**, 4276–4283.
- Washington, M. T., Johnson, R. E., Prakash, S. & Prakash, L. 1999 Fidelity and processivity of *Saccharomyces cerevisiae* DNA polymerase η . *J. Biol. Chem.* **274**, 36 835–36 838.
- Winter, D. B., Phung, Q. H., Umar, A., Baker, S. M., Tarone, R. E., Tanaka, K., Liskay, R. M., Kunkel, T. A., Bohr, V. A. & Gearhart, P. J. 1998 Altered spectra of hypermutation in antibodies from mice deficient for the DNA mismatch repair protein PMS2. *Proc. Natl Acad. Sci. USA* **95**, 6953–6958.
- Wood, E. R., Ghane, F. & Grogan, D. W. 1997 Genetic responses of the thermophilic archaeon *Sulfolobus acidocaldarius* to short-wavelength UV light. *J. Bacteriol.* **179**, 5693–5698.
- Woodgate, R. 1999 A plethora of lesion-replicating DNA polymerases. *Genes Dev.* **13**, 2191–2195.
- Woodgate, R. & Levine, A. S. 1996 Damage inducible mutagenesis: recent insights into the activities of the Umu family of mutagenesis proteins In *Cancer surveys: genetic instability in cancer* (ed. T. Lindahl), pp. 117–140. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Yelamos, J., Klix, N., Goyenechea, B., Lozano, F., Chui, Y. L., Gonzalez Fernandez, A., Pannell, R., Neuberger, M. S. & Milstein, C. 1995 Targeting of non-Ig sequences in place of the V segment by somatic hypermutation. *Nature* **376**, 225–229.
- Yuan, F., Zhang, Y., Rajpal, D. K., Wu, X., Guo, D., Wang, M., Taylor, J. S. & Wang, Z. 2000 Specificity of DNA lesion bypass by the yeast DNA polymerase η . *J. Biol. Chem.* **275**, 8233–8239.