Structural basis of error-prone replication and stalling at a thymine base by human DNA polymerase ι

Kevin N Kirouac and Hong Ling*

Department of Biochemistry, University of Western Ontario, London, Ontario, Canada

Human DNA polymerase ι (pol ι) is a unique member of Y-family polymerases, which preferentially misincorporates nucleotides opposite thymines (T) and halts replication at T bases. The structural basis of the high error rates remains elusive. We present three crystal structures of pol ι complexed with DNA containing a thymine base, paired with correct or incorrect incoming nucleotides. A narrowed active site supports a pyrimidine to pyrimidine mismatch and excludes Watson–Crick base pairing by pol ι . The template thymine remains in an *anti* conformation irrespective of incoming nucleotides. Incoming ddATP adopts a *syn* conformation with reduced base stacking, whereas incorrect dGTP and dTTP maintain *anti* conformations with normal base stacking. Further stabilization of dGTP by H-bonding with Gln59 of the finger domain explains the preferential T to G mismatch. A template ‘U-turn’ is stabilized by pol ι and the methyl group of the thymine template, revealing the structural basis of T stalling. Our structural and domain-swapping experiments indicate that the finger domain is responsible for pol ι ’s high error rates on pyrimidines and determines the incorporation specificity.

The EMBO Journal (2009) 28, 1644–1654. doi:10.1038/emboj.2009.122; Published online 14 May 2009

Subject Categories: genome stability & dynamics; structural biology

Keywords: incorporation specificity; mutagenesis; pol ι ; translesion synthesis; Y family DNA polymerase

Introduction

A strict adherence to Watson–Crick base pairing is a key feature in determining the high fidelity of replicative DNA polymerases. An induced fit mechanism is employed when a correct incoming nucleotide is optimally paired with the template base (Doublié *et al*, 1999). A highly restrictive active site with limited solvent accessibility enables this efficient base pair selection (Doublié *et al*, 1998). In contrast, Y-family DNA polymerases, which specialize in traversing DNA lesions, have evolved an open and solvent accessible active site allowing for permissive base pairing (Ling *et al*, 2001).

The Y-family polymerases contain a similar catalytic core, consisting of ‘palm’, ‘finger’ and ‘thumb’ domains as in high-fidelity DNA polymerases. The smaller finger and thumb domains in Y-family polymerases generate a solvent-accessible and spacious active site. The finger domain contacts the replicating base pair and is therefore the substrate recognition site. In addition, the Y-family polymerases possess a unique C-terminal domain, called the ‘little finger’ or polymerase-associated domain. The little finger (LF) holds the DNA substrate along with the thumb domain. Around the active site, the major groove of the DNA duplex is fully solvent exposed. Thus, bulky DNA adducts can be accommodated in the active site and multiple nucleotide conformations can be adopted for the enzyme to replicate through the lesion. The caveat to performing translesion DNA synthesis is a high error rate of replication on undamaged DNA (Johnson *et al*, 2000b; Zhang *et al*, 2000; Boudsocq *et al*, 2001).

Human Y-family DNA polymerase ι (pol ι) is a specialized polymerase that does not utilize Watson–Crick base pairing for DNA replication. Instead, this enzyme functions by inducing a *syn* conformation on template purines, which results in Hoogsteen base pairing with the correctly matched incoming nucleotide (Johnson *et al*, 2005; Nair *et al*, 2005a, 2006b). The ability to induce a nucleotide *syn* conformation by pol ι seems to serve as the mechanism for replication opposite damaged template purines. Structural evidence shows that the 1, *N*⁶ ethenodeoxyadenosine and N2 ethyl guanine lesions are presented in the *syn* conformation protruding into the solvent-accessible major groove of the DNA helix (Nair *et al*, 2006a; Pence *et al*, 2008), which allows base pairing with the correct incoming nucleotide.

DNA replication by pol ι on template pyrimidines shows extremely high error rates, whereas incorporation opposite template purines is more accurate (Tissier *et al*, 2000, 2001; Zhang *et al*, 2000; Kunkel *et al*, 2003). Opposite a template thymine (T), pol ι prefers to incorporate a guanine (G) up to 2.5-fold over the correctly paired adenine (A) in a metal-dependent manner (Frank and Woodgate, 2007). In addition, pol ι has inefficient replication past a template T base causing a signature T template stall (Zhang *et al*, 2000). A similar pattern of misincorporation and replication stalling by pol ι is observed opposite template uracil (U) (Vaisman and Woodgate, 2001). It has been proposed that G misincorporation opposite template U could restore the genomic sequence of cytosines (C) that have undergone deamination. The biological role of this preferred misinsertion of G opposite template T or U within cells remains unknown. However, such an unusual and highly specific property could serve a unique function in DNA maintenance.

The error-prone replication on template T by pol ι has been implicated in the high rates of DNA mutagenesis presented in patients with the UV-sensitive disorder, Xeroderma Pigmentosum Variant (XP-V) syndrome (Wang *et al*, 2007). When the human Y-family DNA polymerase η (pol η) is

*Corresponding author. Department of Biochemistry, University of Western Ontario, London, Ontario, Canada N6A 5C1.
Tel.: +1 519 661 3557; Fax: +1 519 661 3175; E-mail: hling4@uwo.ca

Received: 6 January 2009; accepted: 3 April 2009; published online: 14 May 2009

inactivated by mutations, pol ϵ takes over its specialized role of bypassing UV-induced thymine–thymine (T–T) dimers. However, the preference of misinsertion opposite template T by pol ϵ results in an increase in mutagenesis and the presentation of the disease. Although pol ϵ is responsible for increased DNA mutagenesis when functioning out of context, this enzyme does seem to play a role in tumour suppression. Mice deficient in both pol η and pol ϵ have an earlier onset on UV-light-induced tumours than those with pol η deficiency alone (Dumstorf *et al*, 2006; Ohkumo *et al*, 2006), indicating a role for pol ϵ in UV-induced lesion bypass. In addition, it has recently been observed that pol ϵ plays a significant role in cellular protection from oxidative damage (Petta *et al*, 2008). Although pol ϵ probably facilitates the repair of oxidative DNA lesions, the specificity and mechanism of this repair are unknown. The unique T template misincorporation by pol ϵ has been extensively reported, but has remained mechanistically unexplained.

Here, we report three crystal structures of pol ϵ in complex with DNA containing a template T base at the active site, which is paired with either correct (A) or incorrect (T or G) incoming nucleotides. Our results show, for the first time, the structural basis of preferred G misincorporation and stalling on a template T base by pol ϵ .

Results

Pok–DNA–dNTP complexes with template bending back at the T base

In order to position the thymine base at the pol ϵ active site, DNA substrates for crystallization containing a thymine 5'

to the template–primer junction were designed. The first (substrate 1) is a 15/9-nt duplex DNA with two thymines 5' to the template–primer junction (see Materials and methods). The second (substrate 2) is an 18-nt self-complementary duplex (14 base pairs) containing a 2,3-dideoxy 3' primer end for trapping ternary complexes (see Materials and methods). Incoming 2,3-dideoxy ATP (ddATP) was incubated with DNA substrate 1 and co-crystallized with pol ϵ , whereas dGTP and dTTP were incubated with DNA substrate 2 and co-crystallized with pol ϵ . Interestingly, the pol ϵ –DNA–ddATP complex was trapped at the first T from the template–primer junction without the expected one incorporation, probably because of the replication stalling at T. The resulting three crystal structures are denoted as T:ddADP, T:dGTP and T:dTTP according to the replicating base pair in the active site. The ternary complex crystals are in two different space groups (C2 for T:ddADP; P6 $_5$ 22 for T:dGTP and T:dTTP) and diffract to 2.0, 2.0 and 2.2 Å resolutions, respectively (Table I). In T:ddADP, the incoming ddATP was hydrolysed to ddADP. Hydrolysis in this manner has been observed in Dpo4 from *Sulfolobus solfataricus*, the model enzyme of the Y-family DNA polymerase, because of a weak phosphatase activity of the polymerase (Ling *et al*, 2001). In the T:ddADP structure, the active site metal ions have been refined as Ca $^{2+}$, due to the presence of 150 mM CaCl $_2$ in the crystallization buffer and the high electron density. In addition, anomalous signal peaks were observed at the metal ion sites, which are distinct from surrounding non-metal atoms, though weak at the A site (Supplementary Figure 4S). This is analogous to Dpo4 structures crystallized with 100 mM

Table I Summary of crystallographic data

Crystal	T:ddADP	T:dTTP	T:dGTP
Space group	C2	P6 $_5$ 22	P6 $_5$ 22
Complexes per AU ^a	2	1	1
<i>Unit cell</i>			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	140.2, 71.8, 127.4	97.2, 97.2, 201.9	98.1, 98.1, 203.7
β (deg)	112.5		
Resolution range (Å) ^b	52.0–2.0 (2.04–2.00)	27.0–2.2 (2.26–2.20)	24.0–2.0 (2.07–2.00)
<i>R</i> _{merge} ^b	7.2 (61.7)	8.44 (83.8)	12.1 (57.7)
<i>I</i> / σ <i>I</i>	24.6 (2.1)	36.8 (2.3)	57.9 (3.0)
Completeness (%) ^b	99.0 (96.3)	99.9 (100)	99.7 (100)
Redundancy ^b	3.6 (3.1)	11.6 (11.7)	13.3 (8.9)
No. reflections (test)	73972 (2%, 1551)	28762 (2%, 604)	38810 (2%, 826)
<i>R</i> _{work} / <i>R</i> _{free}	20.6/25.3	20.9/25.8	21.8/ 24.9
<i>No. atoms</i>			
Protein	6008	2978	2903
DNA	892	323	323
DNTP	50	28	30
Mg $^{2+}$ ions	—	2	2
Ca $^{2+}$ ions ^c	8	—	—
Waters	757	265	238
<i>B-factors</i>			
Protein	62.5	49.6	43.5
DNA	52.7	46.6	44.5
DNTP	29.0	44.8	48.8
Metal ions	35.0	46.1	47.7
Water	39.2	59.9	46.8
R.m.s.d. bond lengths (Å)	0.016	0.013	0.019
R.m.s.d. bond angles (deg)	1.96	1.56	1.82

^aAU means asymmetric unit.

^bData in the highest resolution shell are in parentheses.

^cThere are five non-catalytic Ca $^{2+}$ in the structure.

$\text{Ca}(\text{AC})_2$ (Ling *et al.*, 2001; Wong *et al.*, 2008). Such anomalous peaks were not observed for the T:dTTP or T:dGTP structures, which were crystallized in the absence of Ca^{2+} ions (Supplementary Figure S5). Thus, the T:dTTP and T:dGTP structures were refined with two active site Mg^{2+} ions. Primer extension assays have been carried out on polI in the presence of 150 mM CaCl_2 , which shows that Ca^{2+} ions do not change the nucleotide incorporation specificity of polI (Supplementary Figure S1). The divalent cation in the B site is positioned identically within all three T template structures, as well as with earlier polI structures containing template purines (Nair *et al.*, 2006b) and Dpo4 ternary structure (Vaisman *et al.*, 2005) (Supplementary Figure S2). The divalent cation in the A site, however, is mobile, with variable positions in all three structures (Supplementary Figure S2). Divalent ion mobility within the A site has also been reported earlier for polI (Nair *et al.*, 2005a) and Dpo4 (Vaisman *et al.*, 2005).

PolI in all three ternary structures is essentially identical to that of the purine-template polI structures solved earlier (Nair *et al.*, 2004, 2005a, 2006a,b) (Figure 1A). The pair-wise comparisons on all C α atoms produced root-mean-square

deviations (r.m.s.d.) within 0.7 Å among our three complex structures. In addition, the C α r.m.s.d. is ~0.8 Å between T:ddADP and a polI ternary complex solved earlier (PDB:2ALZ) containing a purine base at the template position. The close agreement between all of these polI structures indicates that polI, like other Y-family polymerases, does not undergo significant conformational change when replicating through different DNA substrates (Ling *et al.*, 2001; Nair *et al.*, 2006a,b; Bauer *et al.*, 2007).

We use the first complex structure (T:ddADP, Figure 1C) to describe the general features of the three ternary complexes, as these three structures are identical, except for the replicating base pair. All residues of polI have the same side-chain conformations in the current three structures, which are identical to those of the purine-template polI structures reported earlier, except for Tyr 61. Tyr 61 flips its side-chain conformation 100° from that seen in the purine-template polI structures, and moves its aromatic ring 9 Å closer to the template DNA (Figure 1A). The unique Tyr61 orientation is observed in all three of our template thymine structures reported here, which are in two different crystal forms. Thus, this conformation is independent of crystal packing

Figure 1 Comparison of polI–DNA–nucleotide ternary structures. The colour schemes are shown either as the colour of the appropriate labels or as the colour bars in the panels. (A) Superposition of T:ddADP (yellow), T:dTTP (cyan), T:dGTP (magenta), and a polI ternary complex solved earlier (PDB: 2ALZ, grey). Proteins are shown as C α traces and Tyr 61 is shown as sticks to highlight its conformational change as a result of the DNA U-turn. Incoming nucleotides were omitted for clarity. (B) DNA superposition of T:ddADP (yellow), Dpo4 ternary structure (PDB: 2AGQ, blue) and a polI ternary complex solved earlier (PDB: 2ALZ, grey). The top view is also shown with the template T base in orange and the incoming ddADP in red. (C, D) PolI (T:ddADP) and Dpo4 (PDB: 2ALZ) ternary complexes. DNA template strands are shown in yellow, T bases in orange and primer strands in grey. The U-turn DNA and position of the 5' template end are indicated by the appropriate arrows. LF represents the little finger domain.

and is likely induced by the DNA substrate that contains a template T base (see details below).

A striking structural change is observed in the single-stranded DNA template when comparing our T template base structures with earlier pol ι structures containing template purines and with Y-family polymerase Dpo4 (Figure 1). The earlier pol ι and Dpo4 structures project the single-stranded template DNA away from the active site in extended conformations (Figure 1B–D). In our T template base structures, the single-stranded DNA is flipped back upon itself in a ‘U’-shaped conformation, enclosing the replicating base pair from the DNA’s major groove and approaching the polymerase thumb domain across the major groove (Figure 1B and C). The DNA backbone is bent $\sim 90^\circ$ after the template T base towards the major groove, and the +1 nucleotide (5' to T) is oriented 90° to the template T base. This latter difference is completely different from the template/+1 base relationships in the extension template strands in the other Y-family polymerase structures (Figure 1B). In our structures, the +1 base lies perpendicular to the template T base (position 0) due to DNA strand bending. All three of our pol ι structures show this unique ‘U-turn’ DNA conformation after the T base, irrespective of the incoming nucleotide or the identity of the bases flanking the template T.

Contributions of pol ι domains and T template base to template stabilization

The unique ‘U-turn’ DNA conformation in our pol ι structures is stabilized by both pol ι and the unique T template base within the active site, which likely induces replication stalling. The single-stranded template DNA downstream of the T base is held in position by the finger domain, the LF domain and the thumb domain of pol ι (Figure 2A). Tyr 61, Leu 62 and Leu 78 from the finger domain contact the +1 nucleotide and provide strong hydrophobic interactions to the backbone sugar through the aromatic side chain of Tyr 61 and to the +1 nucleotide base from interactions with the two leucine residues (Figure 2B). The positively charged Arg 347 and polar Ser 307 from the LF domain contact the phosphate backbone of the template at position 0 and +1 nucleotides (Figure 2B). The 5' end of the template is in contact with the thumb domain at Tyr 244 and Lys 237 (Figure 2C). Tyr 244 stacks with the sugar of the +3 nucleotide, and the positively charged Lys 237 interacts with the negatively charged phosphate to fix the free 5' end of the template strand in front of the replicating base pair (Figure 2C). Furthermore, the $\sim 90^\circ$ bend at the T base is stabilized by the interactions between the methyl group of the template T base and the bent single-stranded template DNA (Figure 2D). Although the extensive

Figure 2 Structure of T:ddADP showing template DNA ‘U-turn’ stabilization. Numbers indicate template nucleotide positions relative to T at position 0. Hydrogen bonding is shown as blue dashed lines. (A) Overall T:ddADP structure is shown with DNA template strand in yellow, T base in orange and the primer strand in grey. The finger domain is shown in blue, little finger (LF) domain in purple, thumb domain in green and palm domain in grey. (B) Zoom-in view of the ‘U-turn’ stabilization by the pol ι finger (blue) and LF (magenta) domains. (C) Zoom-in view of the ‘U-turn’ stabilization by the pol ι thumb domain (green). (D) Zoom-in view of the ‘U-turn’ stabilization by the template T base (orange). Hydrophobic interactions are shown as black dashed lines. The view is as seen from underneath and looking up through the red square in panel A; the vertical arrow in panel A points to the side that is at the top of panel D.

contacts between this unique methyl group and the +1 nucleotide reinforce the unusual bending template, the methyl group may not be an absolute requirement, because of a similar stalling effect opposite template U (Vaisman and Woodgate, 2001). The unique 'U-turn' conformation is not observed in the presence of template purine bases (Nair *et al.*, 2005a, 2006a,b). Two of our template thymine structures (T:ddTTP and T:ddGTP) have the same-sized DNA substrate, single-stranded DNA overhang and crystal form as the earlier purine-template structures. However, these template thymine structures adopted the 'U-turn' conformation, similar to the T:ddADP structure, which has a different DNA substrate and crystal form (Figure 1A). This excludes any structural variation caused by differences in the single-stranded DNA and packing environments of the complexes. It seems that the purine bases A and G are too large to be accommodated in the U-turn conformation observed in our structures, which would disrupt bending. However, there are probably other unidentified factors also involved in preventing this conformation. Although template C has a size similar to T and U, which may

lend itself to a 'U-turn' structure, such a conformation may not be stable when template C is in the active site, as no significant replication stalling has been observed with this template base. The U-turn interactions are not involved with the template bases in the double strand DNA except the template T, but are mainly involved with backbone atoms on the downstream single-strand DNA. This is consistent with the observations that the stalling only depends on the T base and not on the bases flanking it (Zhang *et al.*, 2000). The current T template structures clearly show the structural basis for the signature T template stalling by pol ϵ .

Role of finger domain at pol ϵ active site

Structural comparison of pol ϵ and Dpo4 ternary complexes indicate that the Y-family polymerases are highly conserved structurally in the core area that forms the DNA-binding cleft, except for the finger domains (Figure 3B). The finger domains are structurally conserved, with most secondary structural elements aligned between pol ϵ and Dpo4 (Figure 3A and B). However, there are two striking differences in these finger

Figure 3 Polt and Dpo4 active site comparison. (A) Structure-based sequence alignment of amino acids for the finger domains of pol ϵ (cyan) and Dpo4 (grey). Numbers 2, 3 and 4 indicate the second, third and fourth β -sheets. Secondary structure is indicated as rectangles for α -helices and arrows for β -sheets. Residues interacting with the replicating base pair are highlighted in magenta. (B) Superposition of T:ddADP (cyan, purple) and ternary Dpo4 (type I)-DNA-nucleotide (1JX4, grey). The incoming nucleotides are shown as sticks for Dpo4 (grey) and T:ddADP (yellow). LF represents the little finger domain. (C, D) Close-up views of active sites showing finger-LF domain interactions in pol ϵ and Dpo4. Finger domains are cyan, LF domains are purple and DNA is yellow. (E) Active site superposition from (B) of T:ddADP (finger: cyan; little finger: purple; T base: orange; ddADP: yellow) with Dpo4 (grey). Positioning of the replicating base pair by pol ϵ side chains is indicated with black arrows.

domains, which affect the shape of the active sites and their interactions with DNA substrate. These differences are concentrated in the fragment between the β -strands $\beta 2$ and $\beta 3$, which is in the non-conserved substrate recognition site that contacts the template DNA and the replicating base pair in the active site (Yang, 2003). First, polI has a much shorter loop between $\beta 2$ and $\beta 3$ than that of Dpo4. This loop (L23) forms a structural interface for the finger domains to interact with the LF domains, as well as the single-stranded template DNA in Dpo4 (Figure 3B–E). The shorter loop of polI causes the LF domain to rotate 12° inward to the finger domain and the $\beta 9$ strand moves $\sim 2 \text{ \AA}$ towards the template strand, creating a narrowed active site in polI (Figure 3B and C). The narrowed active site limits the C1'–C1' distance of the replicating base pair to within $9 \text{ \AA}$ in polI. The C1'–C1' distances are 8.3, 8.6 and $8.9 \text{ \AA}$ in the structures of T:ddADP, T:dTTP and T:dGTP, respectively. By contrast, the replicating base pair in the active site of Dpo4 has a C1'–C1' distance of $\sim 10.6 \text{ \AA}$ (Figure 3D), which is a common strand width for B-form DNA in all other Y-family polymerase structures (Nair *et al*, 2005b; Alt *et al*, 2007; Lone *et al*, 2007).

The second difference between polI and Dpo4 finger domains lies in the fragment that contacts the replicating base pair within the active site. PolI has relatively large amino-acid side chains (Gln 59, Lys 60, Leu 62, Val 64 and Leu 78) contacting the replicating base pair at the active site (Figure 3E) when compared with Dpo4, which has relatively small amino acids (Val 32, Gly 41, Ala 42, Ala 44 and Gly 58) for the contacts. Although the finger domain is rolled out by $\sim 15^\circ$ relative to Dpo4, the larger side chains of Gln 59, Leu 62, Val 64 and Leu 78 in polI still push the replicating base pair towards the major groove, effectively tilting it off-plane relative to that of Dpo4 (Figure 3E). Lys 60 and the residues from strand $\beta 9$ of the polI LF domain squeeze the template base towards the incoming nucleotide and make the C1'–C1' distance shorter than $9 \text{ \AA}$ (Figure 3E). These structures show that the finger domain is not only important for contacting the replicating base pair but is also an essential factor for restricting the C1'–C1' distance in polI. Therefore, the polI finger domain is most likely the functional domain responsible for the nucleotide specificity during replication.

Conformation and stability of replicating base pairs in polI active site

The unique polI active site and the T template base make the replicating base pairs in the ternary complexes different in conformation from those in other Y-family polymerase ternary structures. All three of our polI structures have the template T base in a normal *anti* conformation when it is paired with an incoming nucleotide in the active site (Figure 4). Nucleotide binding does not induce a conformational change in the T template base as observed in the purine-template structures (Nair *et al*, 2005a, 2006a,b). Instead, the incoming dNTPs of the replicating base pairs in our polI structures adopt different conformations, depending on their fit in the enzyme active site. Owing to large residues from finger domain, the template T base is pushed out of the stacking area of the underneath bases and tilted off-plane by the finger domain in all three structures (Figure 4). The tilt (τ) and roll (ρ) angles of the off-plane T from the underlying base are around 6° and 16° , respectively.

In the T:ddADP structure, the incoming ddADP adopts a *syn* conformation and forms a Hoogsteen base pair with the template T (Figure 4A). The Hoogsteen base pair in the T:ddADP structure fits the narrowed active site with a C1'–C1' distance of $8.4 \text{ \AA}$, which is similar to other reported polI structures (Nair *et al*, 2005a, 2006a,b) and is smaller than the required C1'–C1' distance of $\sim 10.6 \text{ \AA}$ for proper Watson–Crick base pairing (Ling *et al*, 2001, 2003, 2004a,b; Wong *et al*, 2008). In addition, the incoming ddADP is flipped out of the stacking area of the underlying base pair towards the major groove because of its *syn* conformation. The ddADP is tilted $\sim 20^\circ$ off-plane with the underlying base pair and has an elongated stacking distance of about $4 \text{ \AA}$, which weakens the stability of the replicating base pair further (Figure 4A).

In the T:dTTP structure, the mismatched incoming dTTP is in an *anti* conformation (Figure 4B). The narrowed active site holds the pyrimidine–pyrimidine base pair well, due to the pair being smaller than the common pyrimidine–purine base pair in contacting distance. The narrowed active site thus stabilizes the small pyrimidine–pyrimidine mismatched base pair. The C1'–C1' distance of the T to T base pair is $8.5 \text{ \AA}$, which would not be stable in an active site that accommodates a standard Watson–Crick base pair with a C1'–C1' distance of $\sim 10.6 \text{ \AA}$. Interestingly, the incoming dGTP is also in the *anti* conformation, which has not been observed in other polI structures that contain template purine bases in the active site (Nair *et al*, 2005a, 2006a,b) (Figure 4C). The C1'–C1' distance is restricted to $8.9 \text{ \AA}$, which causes the template T to tilt an extra 15° off-plane in order to accommodate the *anti* conformation of the dGTP nucleotide. Our structural observation shows that a *syn* conformation of purine nucleotides can occur in the template or incoming nucleotide position and is the result of a narrowed active site that constrains the C1'–C1' distance of the replicating base pair. In contrast to ddADP, the bases of dTTP and dGTP in *anti* conformations remain within the active site, parallel to the underlying base pair (Figure 4B, C), with stacking distances in the normal range of $3.2\text{--}3.6 \text{ \AA}$. Compared with incoming A and T bases, the G base of dGTP has the largest stacking surface because of its purine base and *anti* conformation. Base stacking between the incoming nucleotide and the underlying base pair is crucial for the stability and preference of nucleotide incorporation (Yang, 2006). Therefore, G is most favourable and A is the least favourable opposite template T in terms of their base-stacking properties.

In addition to base stacking, incoming bases are also stabilized by hydrogen bonding with template bases. There are two hydrogen bonds between incoming A (N^6 and $N7$) and the template T (O^4 and $N3$) in the Hoogsteen base pair in T:ddADP (Figure 4A). In the T:dTTP structure, there are also two hydrogen bonds formed between the template T (O^2 and $N3$) and the incoming dTTP ($N3$ and O^4) (Figure 4B). Accordingly, the hydrogen-bonding forces of the replicating base pairs are comparable in these two complexes. Thus, the loss of base stacking on ddADP makes the mismatched dTTP more favourable for incorporation than the A base. Incoming dGTP also forms two hydrogen bonds with the template T, as well as a unique third hydrogen bond between its N^2 atom and OE1 of Gln59 from the finger domain (Figure 4C). This special hydrogen bonding with Gln59 is the first to be identified in a polI structure and reveals a unique stabilizing force that favours, over other

Figure 4 Base stacking and hydrogen bonding of replicating base pairs with $2F_o - F_c$ electron density maps at 1σ contour level. (A) T:ddADP structure; (B) T:dTTP structure; (C) T:dGTP structure. The template T base is shown in brown, the incoming nucleotide is shown in yellow and the underlying base pair is shown in grey. Hydrogen bonds represented as blue dashed lines are shown in the top views on the left side. Protein side chains involved in hydrogen bonding are shown in cyan. Green spheres represent divalent cations.

bases, the mis-incorporation of dGTP opposite template T by pol ϵ . The Gln 59 is conserved in pol ϵ homologues (Supplementary Figure S3), signifying its functional importance. The structural observation that incoming dGTP is the most stable incoming nucleotide is supported by the observation that the dGTP binding affinity opposite template T by pol ϵ is greater than dTTP or dATP binding ability and is the same as the dTTP binding affinity opposite template A (Washington *et al.*, 2004).

Role of pol ϵ finger domain in base incorporation specificity and replication stalling

The finger domain of pol ϵ contacts the replicating base pair and pulls the LF domain towards the active site, which contributes to the replication specificity. In order to confirm that the pol ϵ finger domain determines the replication specificity, we generated two Dpo4- ϵ chimeric proteins with functional domains switched between Dpo4 and pol ϵ . The LF domain has been implicated in contributing to replication specificity (Boudsocq *et al.*, 2004); thus, the finger or LF domains of Dpo4 were replaced with the corresponding counterparts of pol ϵ in the chimeras. Four DNA substrates containing T, A, C or G at the first replicating position and a T base at the eighth replicating position were used for the functional assays (Figure 5). The chimeric proteins were

tested by primer extension assays with wild-type Dpo4 and pol ϵ as controls.

Opposite the T template base, pol ϵ has a high misincorporation rate of G and T (vertical arrows in Figure 5A) opposite template T, as the primers (bottom bands) are almost fully reacted for dGTP (lane G) and dTTP (lane T). Multiple bands are observed because of the low processivity of these enzymes. In contrast, Dpo4 incorporates the correct A nucleotide preferentially (lane A), with dramatically reduced reactions with dTTP (lane T), dGTP (lane G) and dCTP (lane C) compared with dATP (Figure 5A). Opposite the A template (Figure 5B), both Dpo4 and pol ϵ have quite accurate incorporation, with preference for the correct incoming nucleotide dTTP (lanes T in Figure 5B). The polymerases against the G template in Figure 5C show similar patterns, with C being inserted preferentially. Interestingly, Dpo4 replicates the C template (Figure 5D) accurately, whereas pol ϵ preferentially inserts G with significant misincorporations of T (vertical arrows in Figure 5D). Remarkably, Dpo4- ϵ finger domain chimera (Dpo4- ϵ) adopted a high misincorporation rate of G and T opposite template T similar to pol ϵ , as the primer bands in the lanes T and G are almost fully reacted (vertical arrows in Figure 5A), which is similar to pol ϵ and different from Dpo4 (Figure 5A). Accordingly, opposite template A, G and C, the Dpo4- ϵ resembles pol ϵ and differs from

Figure 5 The role of the finger domain in incorporation specificity. Primer extension analysis was used to examine nucleotide incorporation opposite (A) template T base, (B) template A base, (C) template G base and (D) template C base by polI, Dpo4-1 (finger domain) chimera, and Dpo4. The first replicating template base and the T template base at the eighth position are bolded. Horizontal arrows indicate replication stalling, whereas vertical arrows indicate misincorporation. Enzymes were incubated with DNA and no nucleotides (0), all four dNTPs (4) or individual dNTPs (A, T, C or G).

wild-type Dpo4 (Figure 5B,C). The primer extension assays indicate that replication specificity is dominated by the finger domain of the Y-family polymerases as the finger swapping converts Dpo4 into a polI-like protein in terms of nucleotide incorporation. Interestingly, the other Dpo4-1 chimera with the LF domain swapped into Dpo4 (Dpo4-1-LF) does not show any base incorporation pattern changes from Dpo4 to polI (data not shown), in contrast to what we observed in the Dpo4-1 mutant. This LF-replacement chimera is very different from the Dpo4-Dbh chimeric proteins, in which the enzymatic properties of the mutants are mainly influenced by their LF domains (Boudsocq *et al.*, 2004). In the latter case, the LF domain is swapped between two very similar Y-family polymerases, Dpo4 and Dbh, which have almost identical substrate recognition sites. The Dpo4-Dbh chimeras show functional differences of the LF domains between homologues sharing very similar finger domains (Boudsocq *et al.*, 2004). In our case, the dramatic difference between the finger domains of Dpo4 and polI masks the influence of the LF domains. Overall, the mutagenesis data clearly support the structural observations that the finger domain plays an important role in nucleotide incorporation specificity, particularly for G and T misinsertion opposite template T and in determining the replication specificity.

Replication stalling is observed for polI at the T template bases (horizontal arrows in Figure 5), but not for Dpo4 and Dpo4-1. When all four nucleotides (lanes 4) are present in the assay, polI has poor extension beyond the first T template base (Figure 5A) and stops at the downstream eighth T base (labelled with horizontal arrows in Figure 5), whereas Dpo4 extends the primer to the end of the template DNA (top bands in Figure 5) with better processivity than polI. The finger domain alone does not seem to control the stalling property

of the enzymes, as the chimeric protein Dpo4-1 extends replication beyond the eighth T base. This finding is consistent with our structural observation that three domains, instead of the finger domain alone, contribute to the stabilization of the 'U-turn' DNA, which leads to replication stalling at the T base.

Discussion

The polI finger domain creates a unique active site that induces low fidelity opposite pyrimidines

PolI shows a wide diversity of its fidelity between template purines and template pyrimidines (Tissier *et al.*, 2001; Kunkel *et al.*, 2003). Although this enzyme has low error rates opposite template purines, it has the highest error rate of any known polymerase opposite template thymines (Johnson *et al.*, 2000a,b; Tissier *et al.*, 2000; Zhang *et al.*, 2000). The diversified substrate-recognition site in the finger domain changes in size and residue identity across the Y-family members and is expected to be responsible for the specificity of nucleotide incorporation (Ling *et al.*, 2001; Yang, 2003). The β 2-turn- β 3 loop (L23) is the only part of the polymerase core that contacts the LF domain in the Y-family ternary complex structures. Our structural analyses indicate that the shorter polI L23 of the finger domain induces a movement of the LF domain towards the template strand, which results in a narrowed active site. It is conceivable that the finger domain causes the LF shifting towards the template in the active site, as the LF is the most flexible domain in the Y-family structures (Wong *et al.*, 2008).

PolI promotes a T to G mismatch by its unique narrowed active site and specific interactions with the replicating base pair. As a good structural fit, the narrowed active site sup-

ports pyrimidine to pyrimidine mismatches when the template base is a pyrimidine. Higher misincorporations of T to T, T to C, C to T and C to C were observed in our primer extension assays for both pol ι and Dpo4- ι relative to misincorporations opposite template purines. When the template base is a purine, the smaller, incoming pyrimidine maintains its base stacking in *anti* conformations (Nair *et al*, 2006a, b), whereas a larger, incoming purine nucleotide would be difficult to fit in the narrowed active site for a purine to purine mismatch. This mechanism prevents misincorporation against purine-template bases and allows for accurate replication. Our structural and biochemical analyses are consistent with the well-documented, high error rates against template pyrimidine bases and relatively high fidelity of pol ι for purine bases (Johnson *et al*, 2000a, b; Tissier *et al*, 2000; Zhang *et al*, 2000).

Interestingly, in the presence of Mn²⁺, pol ι has increased fidelity opposite template thymine, with a preference of incorporating the correct A nucleotide instead of incorrect G (Frank and Woodgate, 2007). The Mn²⁺ ion has a more relaxed and mobile coordination than Mg²⁺ within the active site of polymerases. This effect likely allows the incoming nucleotide to adopt a variety of conformations that would not be possible with Mg²⁺. In this manner, Mn²⁺ ion coordination by pol ι may render a favourable interaction that selects A over G opposite template T.

Replication stalling is stabilized by conserved residues over three domains

Another unique feature of pol ι is a pronounced stalling of replication in extending a primer strand opposite a T base (Zhang *et al*, 2000). Our structures of pol ι show a unique template DNA 'U-turn' conformation at the DNA's single-stranded side that may effectively stall replication. The back-bending 'U-turn' conformation is stabilized by specific interactions from the unique methyl group of the template T base and a collection of interactions from three domains of pol ι . Three domains are involved in interacting with the 'U-turn': the finger domain (Tyr 61, Leu 62 and Leu 78), little finger domain (Ser 307 and Arg 347) and thumb domain (Lys 237 and Tyr 244). These combined interactions stabilize the bent single-stranded DNA and seem to hinder its translation into the active site for primer elongation. In addition, the highly bent DNA may also reduce the catalytic efficiency of pol ι , because of the observation that dNTP incorporation opposite template T is much slower than that opposite template A (Washington *et al*, 2004). Interestingly, most of the residues contacting the single-stranded template are unique to and conserved in pol ι from difference species (Supplementary Figure S3), suggesting conservation of specific functions. Replication stalling by pol ι may be involved in recruiting another polymerase for primer extension after insertion opposite template T or U.

Conclusions

Pol ι uniquely replicates DNA with a constrained active site, creating shorter C1'-C1' strand distances, with the finger domain projecting the template base out towards the solvent-accessible major groove and stabilizing a mismatched G base through H-bonding. The finger domain of pol ι is responsible for the unique active site, and in turn, its replication specificity. This feature allows pol ι to maintain a

relatively high fidelity on template purines, yet induce high rates of misincorporation on template pyrimidines. The high fidelity on template purines by pol ι seems to play a role in translesion synthesis by allowing accurate replication through adducted purine bases. The biological role of pol ι 's low fidelity is still unclear; however, it is apparent that when functioning out of context, this unique replication specificity induces high rates of DNA mutagenesis.

Materials and methods

Dpo4- ι chimeric proteins

To construct the Dpo4- ι finger domain chimera (Dpo4- ι) and the Dpo4- ι little finger domain chimera (Dpo4- ι -LF), plasmid vector pET-22b containing the Dpo4 gene and plasmid vector pHis-parallel1 containing the pol ι gene were used as templates for PCR. For Dpo4- ι , the N-terminus of Dpo4 was cloned up to the beginning of the finger domain using primers A (5'-CGTTACTGCC ATGGTTGTTCTTTTCGTTG-3') and B (5'-TTCTACTTGTGCATAAA AGCAGTCAAAATCAACGAAAAGAACAATC-3'). The result was an N-terminus Dpo4 PCR product containing an *NcoI* cutting site at the N terminus and a C-terminal overhang, which was complementary to the beginning of the pol ι finger domain. The C-terminus of Dpo4 was cloned past the end of the finger domain using primers C (5'-GTTGGTATTAGTTAATGGAGAAGACAAGGAAGTATATCAGCAAG TTTC-3') and D (5'-GCTAGTTATTGCTCAGC-3'). The result was a C-terminal Dpo4 PCR product with an N-terminal overhang, complementary to the end of the pol ι finger domain. The finger domain of pol ι was cloned using primers E (5'-TGCTTTTATGCACA AGTAGAAATG-3') and F (5'-GTCTTCTCCATTAATAACCAAC-3'). The N-terminal Dpo4 product was joined with the pol ι finger domain product using primers A and F. The resulting N-terminal Dpo4-pol ι finger domain product was joined with the C-terminal Dpo4 product using primers A and D to produce the final product of a Dpo4 gene containing the finger domain of pol ι (Dpo4- ι).

For Dpo4- ι -LF, the N-terminus of Dpo4 was cloned up to the beginning of the little finger domain using primers A (5'-CGTTACTG CCATGGTTGTTCTTTTCGTTGATTGTTGACTACTTTTACGCTC-3') and B (5'-AGTTCTTATAGGCTCGTTATACTCGTCTCTAGCTAGA-3'). The result was an N-terminus Dpo4 PCR product containing an *NcoI* cutting site at the N terminus. The little finger domain of pol ι was cloned using primers C (5'-GCCGTTACTGCCATGGTTGTTCTTTTCG TTGATTTTGACTACTTTTACGCTC-3') and D (5'-CATCCTCGAGACCT ACTTAGCAGTATTTAGTGCTTAAAGGTGCAGAAC-3'). The result was a pol ι little finger domain with an N-terminal overhang, complementary to the end of the Dpo4 product. The N-terminal Dpo4 product was joined with the pol ι little-finger domain product using primers A and D to produce the final product of a Dpo4 gene containing the little finger domain of pol ι (Dpo4- ι -LF). Both Dpo4- ι and Dpo4- ι -LF genes were cloned into the pHis-parallel1 vector and confirmed by sequencing.

Primer extension assays

DNA substrate (10 nM) was incubated with Dpo4, Dpo4- ι , Dpo4- ι -LF or hpol ι (10 nM) and 100 μ M of either all four dNTPs or individual dNTPs at 37°C for 2 min in a reaction buffer containing 40 mM Tris (pH 8.0), 5 mM MgCl₂, 250 μ g/ml BSA, 10 mM DTT and 2.5% glycerol. Reactions carried out in the presence of 150 mM CaCl₂ were incubated at 37°C for 60 min. Reactions were terminated with loading buffer (95% formamide, 20 mM EDTA, 0.025% xylene and 0.025% bromophenol blue) and resolved on a 20% polyacrylamide gel containing 7 M urea. Gels were visualized using a PhosphorImager.

Protein preparation

Human DNA pol ι (amino acid 1-420) was cloned into pGST-parallel1 vector and the subsequent glutathione S-transferase-tagged pol ι was overexpressed in *Escherichia coli* strain DE3. The pol ι -GST fusion protein was purified using affinity chromatography and cleaved using a histidine-tagged tobacco etch virus (TEV) protease, which was subsequently removed using nickel affinity chromatography. The cleaved pol ι containing two extra N-terminal residues was further purified using an SP column. Dpo4 used for

functional assays was purified as described earlier (Ling *et al*, 2001). The His-tagged Dpo4- κ chimeric proteins used for functional assays were overexpressed in *E. coli* strain DE3 and purified using nickel affinity chromatography, followed by further purification on an SP column.

DNA preparation

Oligonucleotides for crystallization were purchased from Keck Oligo Inc. and were gel purified. The 9-nt primer (5'-GTGGATGAG-3') was annealed to a 15-nt template (5'-CTCATTCATCCAC-3'), and the self-annealing 18-nt oligonucleotide (5'-TCATGGGTCCTAGGAC CC^{dd}-3') was annealed with itself to give a DNA substrate with two replicative ends. Oligonucleotides used for primer extension assays were purchased from Sigma-Aldrich and were gel purified. A 30-nt template (5'-GAGCAGTCGCACATGTAGTATCTCTGTGAC-3') was annealed to a 16-nt primer (5'-GTCACAGAGATACTAC-3') resulting in a template T base at the first and eighth position beyond the primer-template junction. The primer was 5'-end labelled using [γ -32P]ATP and T4 polynucleotide kinase. The 5'-labelled primer was mixed with template DNA at a 1.5:1 molar ratio and heated at 95 °C, followed by slow cooling to form the annealed DNA substrate.

Crystallization and structure determination

Ternary complexes were formed for T:ddADP, T:dGTP and T:dTTP by incubating protein (0.2 mM) and DNA in 1:1.2 ratio with either ddNTP or dNTP (1 mM) and MgCl₂ (5 mM). Crystals of the T-dGTP and T-dTTP complex were obtained in 12% PEG 5000 MME + 0.2 M NH₄SO₄ + 5% glycerol + 0.1 M MES, pH 6.5, whereas crystals of the T-ddADP complex were obtained in 12% PEG 3350 + 0.15 M CaCl₂ + 0.01 M DTT + 5% glycerol. All crystals were flash frozen in liquid nitrogen using paratone-N as a cryo protectant. X-ray diffraction data were collected on the beamline 24-ID-C at the Advanced Photon Source in Argonne National Laboratory. The data were processed and scaled using HKL (Otwinowski and Minor, 1997).

References

Alt A, Lammens K, Chiocchini C, Lammens A, Pieck JC, Kuch D, Hopfner KP, Carell T (2007) Bypass of DNA lesions generated during anticancer treatment with cisplatin by DNA polymerase ϵ . *Science* **318**: 967–970

Bauer J, Xing G, Yagi H, Sayer JM, Jerina DM, Ling H (2007) A structural gap in Dpo4 supports mutagenic bypass of a major benzo[a]pyrene dG adduct in DNA through template misalignment. *Proc Natl Acad Sci USA* **104**: 14905–14910

Boudsocq F, Iwai S, Hanaoka F, Woodgate R (2001) Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4): an archaeal DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic pol ϵ . *Nucleic Acids Res* **29**: 4607–4616

Boudsocq F, Kokoska RJ, Plosky BS, Vaisman A, Ling H, Kunkel TA, Yang W, Woodgate R (2004) Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication. *J Biol Chem* **279**: 32932–32940

Double S, Sawaya MR, Ellenberger T (1999) An open and closed case for all polymerases. *Structure* **7**: R31–R35

Double S, Tabor S, Long AM, Richardson CC, Ellenberger T (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* **391**: 251–258

Dumstorf CA, Clark AB, Lin Q, Kissling GE, Yuan T, Kucherlapati R, McGregor WG, Kunkel TA (2006) Participation of mouse DNA polymerase ι in strand-biased mutagenic bypass of UV photo-products and suppression of skin cancer. *Proc Natl Acad Sci USA* **103**: 18083–18088

Frank EG, Woodgate R (2007) Increased catalytic activity and altered fidelity of human DNA polymerase ι in the presence of manganese. *J Biol Chem* **282**: 24689–24696

Johnson RE, Prakash L, Prakash S (2005) Biochemical evidence for the requirement of Hoogsteen base pairing for replication by human DNA polymerase ι . *Proc Natl Acad Sci USA* **102**: 10466–10471

Johnson RE, Washington MT, Haracska L, Prakash S, Prakash L (2000a) Eukaryotic polymerases ι and ζ act sequentially to bypass DNA lesions. *Nature* **406**: 1015–1019

All three structures were solved using molecular replacement with a previously solved ternary complex (PDB:2ALZ) as a search model. Rigid body refinement was carried out using REFMAC (Murshudov *et al*, 1997), followed by restrained refinement and then TLS refinement. Electron density was well defined for all structures except for the first 27 residues of the N-terminus, loop regions 332–337, 350–356 and 371–378, and the last six residues of the C-terminus. Additionally, the +4 and +5 nucleotides within T:ddADP and the +3 and +2 nucleotides within T:dGTP and T:dTTP were disordered. All structures have good stereochemistry, with over 95% of the residues in the most favoured region of the Ramachandran plot.

Coordinates

The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.rcsb.org, with accession codes 3GV5, 3GV7 and 3GV8 for the structures T:ddADP, T:dGTP and T:dTTP, respectively.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

Acknowledgements

We thank G Xing for subcloning the N-terminal (1–420) polymerase domains from the full-length pol κ and K Rajashankar for beamline support at 24-ID of APS in Argonne National Laboratory. This research was funded by the Canadian Institutes of Health Research (HL).

Conflict of interest

The authors declare that they have no conflict of interest.

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

Kunkel TA, Pavlov YI, Bebenek K (2003) Functions of human DNA polymerases ϵ , κ and ι suggested by their properties, including fidelity with undamaged DNA templates. *DNA Repair (Amst)* **2**: 135–149

Ling H, Boudsocq F, Plosky BS, Woodgate R, Yang W (2003) Replication of a *cys*–*syn* thymine dimer at atomic resolution. *Nature* **424**: 1083–1087

Ling H, Boudsocq F, Woodgate R, Yang W (2001) Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. *Cell* **107**: 91–102

Ling H, Boudsocq F, Woodgate R, Yang W (2004a) Snapshots of replication through an abasic lesion; structural basis for base substitutions and frameshifts. *Mol Cell* **13**: 751–762

Ling H, Sayer JM, Plosky BS, Yagi H, Boudsocq F, Woodgate R, Jerina DM, Yang W (2004b) Crystal structure of a benzo[a]pyrene diol epoxide adduct in a ternary complex with a DNA polymerase. *Proc Natl Acad Sci USA* **101**: 2265–2269

Lone S, Townson SA, Uljon SN, Johnson RE, Brahma A, Nair DT, Prakash S, Prakash L, Aggarwal AK (2007) Human DNA polymerase κ encircles DNA: implications for mismatch extension and lesion bypass. *Mol Cell* **25**: 601–614

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**: 240–255

Nair DT, Johnson RE, Prakash L, Prakash S, Aggarwal AK (2005a) Human DNA polymerase ι incorporates dCTP opposite template G via a G.C + Hoogsteen base pair. *Structure* **13**: 1569–1577

Nair DT, Johnson RE, Prakash L, Prakash S, Aggarwal AK (2005b) Rev1 employs a novel mechanism of DNA synthesis using a protein template. *Science* **309**: 2219–2222

Nair DT, Johnson RE, Prakash L, Prakash S, Aggarwal AK (2006a) Hoogsteen base pair formation promotes synthesis opposite the 1,N⁶-ethenodeoxyadenosine lesion by human DNA polymerase ι . *Nat Struct Mol Biol* **13**: 619–625

- Nair DT, Johnson RE, Prakash L, Prakash S, Aggarwal AK (2006b) An incoming nucleotide imposes an anti to syn conformational change on the templating purine in the human DNA polymerase- ϵ active site. *Structure* **14**: 749–755
- Nair DT, Johnson RE, Prakash S, Prakash L, Aggarwal AK (2004) Replication by human DNA polymerase- ϵ occurs by Hoogsteen base-pairing. *Nature* **430**: 377–380
- Ohkumo T, Kondo Y, Yokoi M, Tsukamoto T, Yamada A, Sugimoto T, Kanao R, Higashi Y, Kondoh H, Tatematsu M, Masutani C, Hanaoka F (2006) UV-B radiation induces epithelial tumors in mice lacking DNA polymerase ϵ and mesenchymal tumors in mice deficient for DNA polymerase ϵ . *Mol Cell Biol* **26**: 7696–7706
- Otwinowski Z, Minor W (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol* **276**: 307–326
- Pence MG, Blans P, Zink CN, Hollis T, Fishbein JC, Perrino FW (2008) Lesion bypass of N2-ethylguanine by human DNA polymerase ϵ . *J Biol Chem* **284**: 1732–1740
- Petta TB, Nakajima S, Zlatanou A, Despras E, Couve-Privat S, Ishchenko A, Sarasin A, Yasui A, Kannouche P (2008) Human DNA polymerase ϵ protects cells against oxidative stress. *EMBO J* **27**: 2883–2895
- Tissier A, Frank EG, McDonald JP, Vaisman A, Fernandez de Henestrosa AR, Boudsocq F, McLenigan MP, Woodgate R (2001) Biochemical characterization of human DNA polymerase ϵ provides clues to its biological function. *Biochem Soc Trans* **29**: 183–187
- Tissier A, McDonald JP, Frank EG, Woodgate R (2000) poliota, a remarkably error-prone human DNA polymerase. *Genes Dev* **14**: 1642–1650
- Vaisman A, Ling H, Woodgate R, Yang W (2005) Fidelity of Dpo4: effect of metal ions, nucleotide selection and pyrophosphorolysis. *EMBO J* **24**: 2957–2967
- Vaisman A, Woodgate R (2001) Unique misinsertion specificity of poliota may decrease the mutagenic potential of deaminated cytosines. *EMBO J* **20**: 6520–6529
- Wang Y, Woodgate R, McManus TP, Mead S, McCormick JJ, Maher VM (2007) Evidence that in xeroderma pigmentosum variant cells, which lack DNA polymerase ϵ , DNA polymerase ϵ causes the very high frequency and unique spectrum of UV-induced mutations. *Cancer Res* **67**: 3018–3026
- Washington MT, Johnson RE, Prakash L, Prakash S (2004) Human DNA polymerase ϵ utilizes different nucleotide incorporation mechanisms dependent upon the template base. *Mol Cell Biol* **24**: 936–943
- Wong JH, Fiala KA, Suo Z, Ling H (2008) Snapshots of a Y-family DNA polymerase in replication: substrate-induced conformational transitions and implications for fidelity of Dpo4. *J Mol Biol* **379**: 317–330
- Yang W (2003) Damage repair DNA polymerases Y. *Curr Opin Struct Biol* **13**: 23–30
- Yang W (2006) Poor base stacking at DNA lesions may initiate recognition by many repair proteins. *DNA Repair (Amst)* **5**: 654–666
- Zhang Y, Yuan F, Wu X, Wang Z (2000) Preferential incorporation of G opposite template T by the low-fidelity human DNA polymerase ϵ . *Mol Cell Biol* **20**: 7099–7108