Published in final edited form as:

DNA Repair (Amst). 2007 December 1; 6(12): 1709–1725.

The X Family Portrait: Structural Insights into Biological Functions of X Family Polymerases

Andrea F. Moon¹, Miguel Garcia-Diaz^{1,2}, Vinod K. Batra¹, William A. Beard¹, Katarzyna Bebenek^{1,2}, Thomas A. Kunkel^{1,2}, Samuel H. Wilson¹, and Lars C. Pedersen^{1,*}

¹ Laboratory of Structural Biology, National Institute of Environmental Health Sciences. 111 T.W. Alexander Drive, MD F3-09, Research Triangle Park, North Carolina 27709

² Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences. 111 T.W. Alexander Drive, MD F3-09, Research Triangle Park, North Carolina 27709

Abstract

The mammalian Family X DNA polymerases (DNA polymerases β , λ , μ , and TdT) contribute to base excision repair and double-strand break repair by virtue of their ability to fill short gaps in DNA. Structural information now exists for all four of these enzymes, making this the first mammalian polymerase family whose structural portrait is complete. Here we consider how distinctive structural features of these enzymes contribute to their biological functions *in vivo*.

Keywords

Family X Polymerase; DNA polymerase β ; DNA polymerase λ ; TdT; DNA polymerase μ ; gap-filling; misinsertion; misalignment; frameshift; lesion bypass; template-dependence; template-independence; substrate specificity

Introduction

The mammalian genome is large and complex, requiring multiple processes to faithfully replicate and maintain its integrity. Most of these processes require DNA synthesis by a DNA polymerase. The human genome encodes at least 15 different DNA polymerases (reviewed in [1]). They are grouped by primary sequence homology into Family A (Pols γ , ν and θ), Family B (Pols α , δ , ϵ and ζ), Family Y (Pols η , ι , κ , and Rev1) and Family X (Pols β , λ , μ , and terminal deoxyribonucleotidyltransferase). In general, most polymerases in Families A and B are high fidelity enzymes primarily involved in faithful DNA replication, and in repair of replication mistakes. Family Y polymerases are less accurate and typically participate in lesion bypass, allowing cells to continue DNA replication where it might otherwise become stalled. Mammalian Family X DNA polymerases are small, relatively inaccurate enzymes involved in a number of DNA repair processes, including base excision repair and repair of double-strand breaks.

What do all these polymerases do in a cell? Why are multiple DNA polymerases required, especially since they all catalyze essentially the same reaction and often share common features? Among several approaches to investigate DNA polymerase functions, one that has

*Corresponding author: Phone: 919-541-0444; Fax: 919-541-7880; Email: pederse2@niehs.nih.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

been particularly successful is structural biology. X-ray crystal structures have been solved for DNA polymerases in each family. While these highly informative structures are often from bacteriophage, viral, eubacterial, or archeal sources, Family X is the first for which structures are now available for all four mammalian members.

Family X enzymes are relatively small DNA polymerases whose primary role is to fill gaps of one to a few nucleotides during DNA repair. Pol β and Pol λ have DNA polymerase and 5'-deoxyribose 5'-phosphate (dRP) lyase activity [2,3], both of which are needed for single-nucleotide base excision repair (BER) [4,5]. BER is a repair process involving recognition and removal of certain DNA lesions (damaged bases and single-stranded nicks), often leaving an apurinic/apyrimidinic (AP) site intermediate. Processing of the AP site generates a single nucleotide gap with 3'-OH and 5'-deoxyribose phosphate termini (reviewed in [6]).

Pol λ , Pol μ , and terminal deoxyribonucleotidyle transferase (TdT) have an amino-terminal BRCT domain (similar to the BRCA-1 C-terminal protein-protein interaction domain [7]), that is associated with processing of DNA ends during nonhomologous end-joining (NHEJ) of double-stranded DNA breaks (DSBs) resulting from DNA damage, and/or during V(D)J recombination [8–11]. V(D)J recombination is a process by which the V, D, and J regions of germline immunoglobulin genes are broken, rearranged, and rejoined to create an array of immunoglobulin protein chains with variable specificities [12].

TdT is unique in that it is a template-independent polymerase [13], generating stretches of random sequence (N-regions) at immunoglobulin gene junctions during V(D)J recombination [14], resulting in increased immunological heterogeneity. Pol λ plays a role in V(D)J recombination at immunoglobulin heavy-chain loci, in an earlier stage than N-addition by TdT [15]. Conversely, Pol μ is involved in VJ recombination at immunoglobulin κ light-chain loci, after synthesis of the N-regions [16]. The source of these disparate biological functions is largely unknown. Yet, clues to understanding differences in activity may lie in the substrate preferences for each enzyme. It has been recently suggested that, among members of the Family X polymerases, there may be a decreasing gradient of template-dependence, correlating to the prescribed physiological functions of these enzymes [11].

In recent years, structural genomics has given rise to a vast array of knowledge for many related proteins—orthologous proteins from different species, and for proteins within a species that are related by primary or tertiary structure. This approach has been used to better understand the unique properties of each of the Family X DNA polymerases. Though these enzymes do not exhibit a high extent of sequence conservation (Figure 1A), the overall structures of their enzymatic domains are remarkably similar [17–20]. Conversely, though their structural folds are nearly identical, their biological activities and substrate preferences markedly differ. A multitude of NMR and X-ray crystal structures is currently available for these polymerases (Table 1), in complex with substrates chosen to coincide with the biological roles of these enzymes. Thorough analysis of these structures provides a deeper understanding of the unique activities and specificities attributed to each polymerase.

Structural Comparison of the Family X Polymerases

Primary sequence alignments (Figure 1A) indicate that Pol β and Pol λ are more similar to each other than to either TdT or Pol μ . Likewise, TdT and Pol μ are more similar to each other than they are to either Pol β or Pol λ . The first x-ray crystal structure of a ternary complex (polymerase with DNA substrate and an incoming nucleotide bound in the active site) of a Family X DNA polymerase was solved for rat Pol β [21] and demonstrated that, like polymerases in other families, the polymerase domain is composed of three subdomains that have been likened to a hand [22]. Also present is an 8 kDa lyase domain that plays an essential role in BER [4,5,23]. Structures of ternary substrate complexes of Pol λ [24] and Pol μ [19],

as well as a binary complex (polymerase with bound oligonucleotide substrate, in the absence of an incoming nucleotide) of TdT [20], indicate that the overall fold and subdomain organization of Family X polymerases are very similar (Figure 1B).

An N-terminal BRCT domain present in Pol λ , Pol μ , and TdT allows them to interact with proteins occupying the ends of DSBs, such as Ku, XRCC4, and Ligase IV (reviewed in [25–27]). Deletion of the BRCT domain abolishes the ability of these polymerases to participate in NHEJ [9,11,28]. The BRCT domains of Pol μ (PDB code 2DUN) and TdT (PDB code 2COE) have been recently solved by NMR spectroscopy (Figure 1C). These domains exhibit an α/β motif that is structurally similar to that of XRCC1 (PDB code 1CDZ [29], r.m.s.d of 2.8Å), another DNA repair protein. In this α/β motif, the core of the structure is comprised of a three-stranded parallel β -sheet, and flanked on either side by three α -helices. There is a high degree of sequence conservation between the BRCT domains of Pol μ and TdT, which translates to their tertiary structure (r.m.s.d of 1.4Å over 80 C α atoms). Interestingly, sequence comparison with the BRCT domain of Pol λ shows no significant similarity. Variations in BRCT domain structure may influence interactions with protein binding factors. For example, Pol μ interacts with the Ku70/Ku80 heterodimer, but this interaction is not sufficient to recruit Pol μ for binding to a DSB. A complex comprised of Pol μ , Ku, and XRCC4/Ligase IV is required for stable recruitment to DSB ends, and for subsequent gap-filling [8]. TdT forms a similar complex with these end-joining factors, and stably binds DNA ends in the presence of Ku and XRCC4/Ligase IV. In contrast, Pol λ can bind and repair a DSB break in complex with XRCC4/Ligase IV alone, in the absence of Ku [9]. Notably, Pol β lacks a BRCT domain and has not been associated with NHEJ.

Structural interpretation of the catalytic mechanism

The classical two metal ion mechanism for catalysis by polymerases was first proposed in 1993 [30]. Subsequently, a ternary substrate complex structure of Pol β was trapped by employing a dideoxy-terminated primer that prevented insertion of the bound nucleoside triphosphate [17]. Three conserved aspartates (D190, D192, and D256) in the catalytic palm subdomain position two Mg⁺² ions necessary for catalysis [31,32]. Although the topology of the palm subdomains of Family X polymerases is not homologous to Family A, B, Y and RT polymerases, the catalytic participants (metals, dNTP, and DNA) can be functionally aligned [33]. Indeed, Family X polymerases share many of the general structural and mechanistic features with polymerases from other families.

A more recent structure of Pol β using a nonhydrolyzable analog at the incoming position has provided a comprehensive view of the pre-catalytic complex just prior to catalysis, where all key atoms are present [34]. Crystal structures of pre- and post-catalytic ternary complexes of Pol λ have also been obtained, providing detail into the subtle but critical changes that occur at the point of bond breakage and formation [24](Garcia-Diaz, in press).

Proper geometry for catalysis is obtained upon binding of the correct incoming nucleotide and two divalent metal ions. A nucleotide-binding metal (site B, Figure 2), is coordinated by oxygen atoms from all three phosphates of the incoming nucleotide, and by two aspartate residues. The catalytic metal (site A, Figure 2) interacts with the α -phosphate, three aspartate residues, and the nucleophile for the reaction—the 3'-OH of the terminal deoxyribose sugar of the primer strand. The catalytic metal positions the O3' with respect to the α -phosphate of the incoming nucleotide, and may also lower its pK_a. Quantum mechanics/molecular mechanics simulations of the catalytic mechanism suggest that the reaction is associative in nature and proceeds via an in-line displacement mechanism whereby the proton from the 3'-OH is transferred to a nearby aspartate (D256) [35]. The nucleophile can then attack the α -phosphate, leading to a trigonal-bipyramidal pentacoordinated transition state. Creation of the transition state species

leads to inversion of the α -phosphate stereochemistry [36,37], and release of the pyrophosphate leaving group with its associated nucleotide-binding metal [17,35,38]. The negative charge build-up on the α -phosphate and the leaving group during the transition state are believed to be stabilized by the presence of the two metal ions [39]. The active site then returns to the pre-catalytic state for another round of catalysis.

Though all Family X enzymes likely utilize an identical chemical mechanism, their methods for active site assembly appear to vary. In Pol β , active site assembly requires large-scale subdomain motions (Figure 3A, right). Pol β exists in an 'open' conformation until binding of the incoming nucleotide, which triggers movement of the carboxyl-terminal thumb subdomain, thereby 'closing' the active site. Specifically, α -helix N of the thumb subdomain is repositioned to directly interact with the nascent base pair (i.e. templating and incoming nucleotides), effectively sandwiching the nascent base pair between α -helix N and the template-primer terminus (Figure 3A, middle) [40]. This provides the polymerase an opportunity to monitor geometric positioning of the incoming nucleotide and the primer terminus.

In addition to large-scale subdomain movement, several side chains rearrange to assemble the active site. Comparing the binary and ternary complex structures (with correct incoming dNTP) reveals that two of the catalytic aspartate residues (Pol β D190 and D256) occupy very similar conformations before and after binding of the incoming nucleotide. However, Y271 and F272 (YF motif) are shifted to lie in the minor groove of the DNA substrate. The third catalytic aspartate, D192, adopts two different conformations, depending on the presence or absence of the incoming nucleotide (Figure 3A, left). In the open conformation, D192 hydrogen bonds with R258, restraining D192 in a conformation that prevents interaction with the metals. Upon binding of the correct nucleotide, movement of the aromatic ring of F272 appears to disrupt the interaction between D192 and R258, allowing a change in the conformation of D192, such that it can coordinate the active site metals. The side chain of R258 also rotates away from the active site, initiating new hydrogen bonding interactions with E295 and/or Y296 [41,42]. It has been speculated that regulation of the two different conformations of D192 and R258 provides a significant barrier to catalysis, which does not occur until the enzyme enters the closed conformation [43].

Additional Pol β side chains undergo significant movement upon dNTP binding. The side chain of R283 from α -helix N alters its conformation to interact with the DNA minor groove [44]. Such minor groove interactions are necessary to maintain the fidelity of polymerization by these enzymes [45–47]. Many side chain motions are associated with closing of the carboxyl-terminal thumb subdomain. However, modeling of thumb motions suggests that these motions are not concerted, but occur in a systematic stepwise fashion [48].

In contrast to Pol β , large-scale domain motions are not observed when Pol λ binds the correct dNTP, suggesting that subdomain motions are not required for catalysis [24]. Pol λ relies primarily on correctly positioning the DNA template strand (Figure 3B). In a binary complex where the incoming nucleotide has not yet bound, the template strand is moved approximately 5 Å outward from the DNA binding cleft. However, protein conformation more closely resembles that of the Pol β ternary complex, where α -helix N is close to the active site. Upon binding of the nucleotide to form the ternary complex, the template strand is moved into the correct position. Like the subtle side chain motions occurring within the active site of Pol β , Pol λ employs a similar pattern of key side chain repositioning (Figure 3B). When comparing the binary and ternary complexes of Pol λ , the catalytic aspartate residues (Pol λ D427, D429, and D490) clearly occupy very similar conformations, regardless of the presence or absence of the incoming nucleotide. The key motions involve the Y505 and F506 side chains (YF motif), which shift, forming minor groove contacts with the correctly positioned DNA duplex [49]. In addition, R517, a structurally equivalent residue to Pol β R283, occupies the templating

position in the binary complex [18], and rearranges its conformation to hydrogen bond within the minor groove [24,50]. These subtle protein and DNA movements are the only significant changes observed for Pol λ . A comparison of pre- and post-catalytic structures of Pol λ shows that active site geometry remains the same throughout the catalytic cycle, and that only the conformation of the α -phosphate changes due to inversion of the stereochemistry as a result of bond formation with the O3' on the primer terminus (Figure 3C) [24].

Superpositions of active site residues from Pol β and Pol λ in liganded complexes are nearly identical (Figure 4A). In contrast, superpositions of active site residues from Pol β and Pol μ display very similar secondary structural architecture, but different side chain identities and positions (Figure 4B). Note also that the active sites of Pol μ and TdT lack the YF motif characteristic of Pol β and Pol λ (Figure 4C). These two residues are replaced by glycine and tryptophan (Pol μ G435/W436, TdT G449/W450). Although a DNA binary complex structure is not available for Pol μ , it seems likely that the GW motif would not be capable of the same concerted movement that occurs upon binding of the nucleotide in Pol β and Pol λ . This hypothesis is supported by crystal structures of unliganded TdT, and with bound single-stranded oligonucleotide or incoming nucleotide, all of which have the GW motif in the same conformation [20]. In Pol μ and TdT, this motif is implicated in decreased discrimination between deoxyribo- versus ribonucleotides [51] [52], as compared to Pol β and Pol λ , which are much more specific for deoxyribonucleotides [40,49,51,53].

Superpositions of active site residues from Pol μ and TdT (Figure 4C) show a level of conservation similar to that observed between Pol β and Pol λ [19,20]. TdT is present in a 'closed' conformation in three different crystal structures, even in the absence of bound substrates [20]. Interestingly, differences in active site configuration become more apparent among the four X family member as distance from the catalytic center increases. This suggests a 'gradient' of protein movements required for catalysis, with Pol β and TdT at opposite ends of the spectrum. Such differences may be relevant to primer-template preferences and biological activity, as discussed below.

Structural insights into enzyme functionality

The 8 kDa Domain—DNA Binding and Gap-filling

In accordance with the role of DNA as the genetic material for all living cells, maintaining the integrity of the DNA sequence is essential. Therefore, damage to the DNA must be detected and repaired quickly in order to prevent loss or misinterpretation of crucial genetic information. Such DNA damage can occur as a byproduct of misincorporation or polymerase slippage during the natural course of DNA synthesis, and as a result of exposure to DNA damaging agents (reviewed in [6]). DNA repair processes seek out DNA lesions, removing them from the DNA strands, and repairing the genetic sequence at the site of the damaged bases. As a byproduct of these DNA repair processes, single- and/or double-stranded gaps are created at certain points along the DNA (Figure 5). Therefore, specific enzymes must exist that are equipped to accommodate these nonstandard substrates, and resolve the gaps—the Family X polymerases have evolved to serve in this capacity. The unique structural feature that allows this family of enzymes to bind single- and/or double-strand gaps is the presence of an N-terminal 8 kDa domain upstream of the polymerization domain (Figure 1). The key role of this 8 kDa domain appears to be DNA binding, and global positioning of the enzyme on gapped or nicked substrates [21,41,54]—consistent with the roles of these enzymes in BER, and in repair of DSBs.

The 8 kDa domains of Pol β and Pol λ harbor an intrinsic dRP lyase activity required to remove the 5'-dRP group generated as an intermediate in single-nucleotide BER. The dRP lyase reaction mechanism proceeds through a β -elimination mechanism via a Schiff base

intermediate, ultimately releasing a 5'-terminal dRP group [2,55]. Recent publications have explored the structural aspects of lyase chemistry [44,56,57]. Though all four Family X enzymes contain the 8 kDa domain, only Pol β and Pol λ have detectable lyase activity [2,3]. The amino acid thought to serve as the catalytic nucleophile, Pol β K72 [58] and Pol λ K312 [3], is not conserved in Pol μ and TdT (V212 in hPol μ and V224 in hTdT), likely accounting for the absence of lyase activity.

To repair damaged or broken DNA, Family X DNA polymerases perform DNA synthesis 'between' two DNA duplex regions separated by one to a few nucleotides of single-stranded DNA. To accomplish this gap filling, these enzymes simultaneously bind to the primer-template junction and to the 'downstream' duplex, an interaction mediated by the 8 kDa domain. One of the mechanisms through which the 8 kDa domain aids in DNA binding is likely by direct interaction with the 5'-phosphate moiety on the downstream end of gapped DNA. In ternary complex structures of Pol β and Pol λ , the 5'-phosphate is bound in a positively charged pocket in the 8 kDa domain (Figure 6A and 6B). Binding is mediated by multiple hydrogen bonding interactions with basic side chains within the pocket. For Pol μ , the concentration of positively charged residues in this pocket is decreased relative to those in Pol β and Pol λ (Figure 6C), and there are concomitantly fewer hydrogen bonding interactions to hold this moiety in position [19]. As of yet, there are no structures of TdT with a DNA substrate containing a downstream duplex. TdT has a structurally similar pocket on the surface of the 8 kDa domain, but this pocket has the lowest concentration of positively charged residues (Figure 6D) [20]. The binding affinity of the different Family X polymerases for gapped DNA substrates likely correlates to the strength of their interactions with the 5'-phosphate. This hypothesis is supported by the observation that the polymerase activity of Pol β and Pol λ is more strongly stimulated by the presence of the 5'-phosphate on the downstream end of the gap [59,60] than is the activity of Pol μ [11].

Family X polymerases contain another structural motif that allows simultaneous binding to both ends of a gapped DNA substrate—the helix-hairpin-helix (HhH) motif. HhH motifs have been described in many proteins that bind either single- or double-stranded DNA in a non-sequence-dependent manner, with the aid of a coordinated metal cation [61,62]. In the Pol β , Pol λ , and Pol μ ternary complex structures, one HhH within the 8 kDa domain (α -helices C and D) interacts with the downstream end of the gap, while a second HhH within the fingers subdomain (α -helices F and G) interacts with the upstream duplex (Figure 1). These structures show that the trajectory of the downstream duplex is altered 90° with respect to the upstream duplex. The positions of the 3'-hydroxyl and 5'-phosphate in the structure of a nicked DNA product complex with Pol β indicate that these atoms are over 27 Å apart. These ends must be ligated in the final step of BER. These structures suggest that the function of the HhH motif is stabilization of the bent DNA, thereby facilitating proper positioning of the two free DNA ends. Although a structure of TdT with relevant repair intermediates has not been determined, similar HhH motifs have been identified [20].

Interestingly, slight differences exist in the structure of the 8 kDa domain HhH motif among the Family X enzymes (Figure 7). In Pol β and Pol λ , this HhH is structurally similar to those found in other DNA repair enzymes [61]. Protein residues in the hairpin loop are conserved, with the motif of G ϕ G (where ϕ is a hydrophobic residue) yielding a specific structural motif based on relative spatial orientation of the two helices (Figure 7A). However, in Pol μ and TdT, α -helix C is distorted, displaying non-canonical geometry (Figure 7B) [19,20]. This distortion is likely due to a lack of primary sequence conservation of hairpin loop residues. In human and mouse TdT, the sequence of residues in the hairpin is C ϕ G, where the hydrophobic residue is a leucine. For Pol μ , the hairpin sequence is HFG in the human enzyme, and YFG in the mouse enzyme. In the latter cases, the hydrophobic residue is larger, which could create steric hindrance with α -helix C and alter its geometry. For both Pol μ and TdT, the presence of a

larger residue (cysteine in TdT and histidine or tyrosine in Pol μ) in place of the first glycine could alter the interaction of this HhH motif with the downstream end of the DNA duplex (Figure 7C). Distortion of the 8 kDa domain HhH motif, combined with decreased interactions with the 5'-phosphate suggests that Pol μ and TdT may have decreased affinity for the downstream DNA, relative to that of Pol β and Pol λ . Consistent with these considerations, Pol β , Pol λ , and Pol μ all exhibit higher activity on gapped DNA duplexes than on template-primer substrates lacking a downstream primer. Pol β [63] and Pol λ [60] are processive in filling of small gaps, when there is a 5'-phosphate on the downstream end of the gap [59,60]. Pol μ activity is also stimulated by a downstream 5'-phosphate, but gap-filling is less processive [52,64].

Template-dependence versus template-independence

One of the key differences between Family X polymerases is the issue of template-dependent versus template-independent polymerization. Pol β and Pol λ are primarily template-dependent enzymes [60,65,66]. Conversely, TdT is a template-independent polymerase [13]. Pol μ exhibits both template-dependent [8] and template-independent [67] activities. It has been suggested that the capacity to perform template-independent synthesis may partly reflect the presence of a loop region between β -strands 3 and 4, referred to as Loop I [11,68]. Loop I is also present in TdT, occupies the same position in all three TdT structures [20], and is located in a region of the DNA binding cleft that would normally be occupied by the template strand (Figure 8A). Therefore, this loop could occlude binding of any DNA substrate possessing a template strand. Such a conformation for Loop I is consistent with the apparent affinity of TdT for single-stranded DNA and double-stranded substrates with a 3'-primer terminal overhang [19]. In Pol μ , Loop I is of a similar length (Figure 8B) to that of TdT. This loop is disordered in the crystal structure of a ternary complex of Pol μ bound to gapped DNA, suggesting conformational flexibility. However, the DNA duplex—including the template strand—is bound in the usual fashion within the DNA binding cleft. Clearly, Loop I of Pol μ cannot occupy the same position as that of TdT in this configuration. A comparison of the ends of the β -strands forming the loop shows that TdT's loop extrudes upward toward the DNA binding cleft, while that of Pol μ appears to turn downward, away from the cleft (Figure 8B) [19]. There are currently no structures of Pol μ bound only to a substrate used for template-independent synthesis, where this loop could occupy a position similar to that adopted by TdT. Interestingly the equivalent regions in Pol β and Pol λ would be less likely to interfere with binding of the template strand because they are much shorter (Figure 8C). Consistent with this idea, when Loop I in Pol μ is shortened to a length similar to that of Pol β , the altered polymerase has higher catalytic efficiency on gapped substrates, but is incapable of template-independent synthesis [11,68].

A second contribution to template-independent activity comes from a histidine residue within the active sites of Pol μ and TdT [19]. This side chain is conserved between Pol μ (H329) and TdT (H342) (Figure 1A), but is absent in Pol β (G189) or Pol λ (G426). In the Pol μ ternary structure, the H329 side chain adopts a conformation that would allow for hydrogen bonding between the histidine nitrogens and the phosphate of the primer terminal residue and/or the γ -phosphate of the incoming nucleotide (Figure 4B). These interactions appear to be crucial for proper positioning of the primer terminus and the incoming nucleotide during template-independent polymerization [19]. The conformation adopted by Pol μ 's H329 is not observed for H342 in any of the TdT structures [20], but could be easily adopted by a simple rotation of the side chain. Just as for Pol μ , the H342A mutant of TdT has substantially reduced template-independent activity [19]. Thus, one could imagine that Loop I and the histidine could cooperatively stabilize the primer terminal nucleotide in a catalytically competent conformation, even when no complementary template strand base is present. This is not only relevant to template-independent synthesis, but also to NHEJ of substrates with short 3'-

overhangs. These substrates require template-dependent extension of a primer terminus lacking its complementary template strand partners (Figures 5B and 5C). This type of polymerization reaction is unique to Pol μ in comparison to other polymerases examined to date, and is abrogated by deletion of Loop I or H329A substitution.

Thus, Pol β is most active for template-dependent synthesis with small gaps of fewer than six nucleotides (Figure 5A) [63]. Pol λ is also active on such gapped DNA substrates, but it also associates with Ku and XRCC4/Ligase IV to repair DSBs with as little as a single complementary nucleotide at the site of the break (Figure 5B) [9,11]. However, neither Pol β nor Pol λ have the active site histidine that allows for stabilization of the primer terminus in the absence of a template strand nucleotide opposite that position. Pol μ is capable of utilizing DSB substrates with microhomology (Figure 5B), but can also polymerize from a DSB with no homology (Figure 5C) [11,69], a property which requires Loop I and H329 [11,19]. In a similar fashion, template-independent synthesis by TdT appears to require both Loop I and H342 (Figure 5D) [19,68]. Overall, structural and biochemical analyses of these four Family X enzymes reveal a gradient of template-dependence [11].

Fidelity of Synthesis

Misinsertion and mispair extension

A gradient of template-dependence is further suggested by considering the fidelity of DNA synthesis conducted by these four Family X polymerases. Pol β has the highest fidelity of Family X polymerases. Pol λ is slightly less accurate for base substitutions and much less accurate for single-base deletions [70,71]. Pol μ is highly error-prone for frameshifts [72] and for substitutions via transient misalignment [73]. TdT, being template-independent, has the lowest biosynthetic 'fidelity'. Structures of Family X polymerases bound to nonstandard substrates provide some insights into their fidelity. Unlike major replicative DNA polymerases like Pol δ and Pol ϵ , Family X members do not possess a 3' to 5' exonuclease domain (Figure 1) and lack exonuclease activity to proofread errors. Similarly, the major replicative polymerases that are more accurate interact extensively with the DNA minor groove upstream of the active site (reviewed in [74]). In contrast, Pol β , Pol λ , and Pol μ have fewer such interactions and may therefore be somewhat more tolerant of DNA distortions upstream of the active site.

DNA polymerases stall or slow at sites of mispaired or adducted nucleotides. Being the most accurate and the most structurally well-characterized of the Family X polymerases, studies of Pol β have been extremely informative and have provided molecular insights into DNA polymerase stalling and/or termination. Structures of Pol β in complex with lesion-containing or mismatched DNA substrates have been determined [75–78]. These structures provide insight into the dynamic interplay between the polymerase and its DNA substrate, and how aberrant base pairs decrease nucleotide insertion efficiency. For example, the binary and ternary complex structures of Pol β with a gapped DNA substrate containing an A-A mismatch at the primer terminus illustrate the influence of DNA sequence on polymerase conformation and provide a structural view for accommodation of this mismatch and reduction of nucleotide misinsertion [77]. In the binary open complex, the mismatched adenines are not planar, but stack with one another, with the template adenine stacking over the primer terminal adenine. The templating (coding) cytidine is flipped extrahelical, permitting the adenines to stack. Pol β binds an incoming nucleotide with high affinity with homopurine mismatches at the primer terminus [79]. Upon addition of the correct nucleotide, dGTP, the structure of the ternary complex indicates that the enzyme has 'closed,' and there are no overt distortions of the DNA backbone (Figure 9A). In contrast to the binary complex, the mismatched adenosines are now planar, with the template strand A is present in the canonical *anti* conformation and the primer terminal A adopting a *syn* conformation to accommodate this mismatched base pair in the DNA

duplex and to conserve the Watson-Crick hydrogen bonding of the adjacent base pairs. In this conformation, the sugar-phosphate of the primer terminal A is slightly distorted, displacing the 3'-hydroxyl away ($>6.5\text{\AA}$ from the α -phosphate of the incoming dGTP) from the optimum catalytic position. The poor position of the primer 3'-hydroxyl likely decreases the frequency of extension of this mispair, permitting an extrinsic proofreading exonuclease to correct this base substitution error. Alternatively, the possibility exists that catalysis may occur from other less populated, but more catalytically efficient, conformations.

A common oxidative DNA lesion is 8-oxo-7,8-dihydro-2'-deoxyguanosine (8oxodG). Unmodified deoxyguanine prefers an *anti* glycosidic conformation, whereas 8oxodG favors a *syn* conformation that can form a Hoogsteen base pair with adenine. Due to ambivalent base pairing properties, DNA polymerases exhibit low fidelity when encountering 8oxodG, inserting dCTP or dATP. The structure of the ternary complex of Pol β with dCTP and gapped DNA, where the templating base is 8oxodG provides a simple explanation (Figure 9B) [75]. Pol β prefers to insert dCTP opposite 8oxodG, although dATP is also incorporated in some instances [80]. To accommodate the keto-group of C8 with 8oxodG in an *anti* conformation, the 5'-phosphate of the templating nucleotide is repositioned by $3.6\text{\AA}$, thereby permitting canonical Watson-Crick base pairing with dCTP. Thus, template backbone flexibility is one parameter that can modulate the *anti/syn* equilibrium of the templating nucleotide. Although dATP is inserted with a much greater efficiency opposite 8oxodG than unmodified dG, the structure of the ternary complex indicates that the incoming dATP had been hydrolyzed to dAMP. In this case, both 8oxodG and dAMP are in *anti* conformations and are observed stacking with one another.

More recently, the structure of Pol β bound to a benzo[c]phenanthrene diol epoxide (B[c]Ph DE)-adducted templating guanine has been reported [78]. The observed *syn* conformation of this mutagenic adducted guanine provides a structural explanation for the observed preference for insertion of incorrect purines over correct dCMP. These structures illustrate the dynamic nature of protein/DNA interactions influencing the coding properties of the templating base, as well as the binding and insertion properties of the incoming nucleotide.

Accommodation of misaligned intermediates

Pol β , Pol λ and Pol μ all generate single-base deletions during synthesis [70,72,81]. This type of error involves misaligned DNA strands with an extra nucleotide in the template strand. Such an intermediate may be generated by several possible mechanisms (reviewed in [82]), one of which is slippage of the template strand relative to the primer strand, as first proposed by Streisinger [83]. This model predicts that insertions and deletions should occur more frequently in repetitive DNA sequences where the misaligned intermediate can be stabilized by correct base pairing. In accordance with this prediction, the single-base deletion error rate of Pol β increases with increasing homopolymeric run length (reviewed in [82]). Interestingly, Pol λ generates single-base deletions at much higher rates than does Pol β [70], a feature that may correlate with the fact that Pol λ has fewer contacts with the duplex DNA upstream of the polymerase active site than does Pol β [18]. Differences in deletion error rates and upstream DNA contacts are consistent with the higher fidelity of Pol β as it fulfills a role in BER, and with the ability of Pol λ to accommodate primer-templates with less than perfect homology (e.g., containing unpaired or mismatched bases), contributing to its role in NHEJ.

The deletion error rate of Pol λ is so high that it was possible to obtain crystal structures of intermediates during formation of single-base deletions [84]. These structures reveal an unpaired nucleotide in the template strand, in an extrahelical position, located one base pair upstream of the polymerase active site (Figure 9C). This base is accommodated with limited structural perturbations of the DNA backbone and may be stabilized by interactions with K544 in β -strand 8. Superposition of the DNA duplexes from the Pol λ misaligned ternary and

canonical ternary structures shows that the substrates are nearly identical except for a slight distortion of the backbone precisely at the location of the 90° bend in the DNA [84]. Thus, no perturbations in protein structure are required in order to accommodate the deletion intermediate, and the positions of all key active site atoms are identical, allowing catalysis to occur. This accommodation, combined with stabilization of the extrahelical intermediate, may account for Pol λ 's high deletion rate. These features are ideal for a polymerase whose relevant substrates may have imperfections in the duplex upstream of the polymerase active site. Pol μ is also capable of deletion synthesis [72,73], and the resulting intermediates may be stabilized in a similar fashion. The stabilizing lysine from Pol λ , K544, is not conserved in Pol μ , but an alternate residue, W457, could perform a similar role [19].

Concluding Remarks

Despite the fact that mammalian Family X polymerase share considerable homology, a similar organization of polymerase domains, and the same catalytic mechanism, they differ in a number of key structural elements and enzymatic properties (as summarized in Table 2). In comparison to their global structural similarity, some of these differences are perhaps deceptively small (e.g., a single amino acid), yet they are critical in defining different biological functions. One theme that develops is that, within the X family, there appear to exist gradients of protein domain, protein side chain, and DNA strand movements. These gradients appear to correlate with increasing structural diversity as the distance increases from the active site where chemistry occurs, which correlates in turn with fidelity, template dependency, and biological function. Pol β and Pol λ participate in BER to maintain genome stability; Pol λ and Pol μ participate in NHEJ of damage-induced double-stranded DNA breaks to avoid lethality, sometimes at the expense of mutagenesis; and Pol λ and Pol μ , and TdT participate in V(D)J recombination to increase antibody diversity.

Pol β has the strictest guidelines for substrate utilization, with an absolute requirement for an unbroken template strand. This affinity for a template strand is consistent with a high concentration of positively charged residues in the DNA binding cleft [40], and extensive hydrogen bonding interactions between the protein and the DNA phosphate backbone. In addition, Pol β , like more accurate polymerases in Families A and B, undergoes large-scale subdomain transitions in order to assemble the active site for catalysis [85]. Catalytic activation through subdomain motion, rather than template repositioning as observed for Pol λ , may account for the relative high fidelity observed for Pol β compared to other X Family members.

Pol λ is slightly more flexible in its requirements for an undistorted template strand. This flexibility likely stems from decreased hydrogen bonding interactions between residues in the binding cleft and the phosphate backbone of the DNA. As a result, Pol λ exhibits a lower fidelity of synthesis than does Pol β and is capable of accommodating frameshift intermediates where the backbone of the template strand is distorted. Frameshift synthesis is accomplished in part by stabilization of the extrahelical nucleotide resulting from misalignment. In addition, Pol λ can fill short gaps so long as there is at least a single complementary base pair upstream of the gap to be filled. Catalysis by Pol λ does not require large-scale protein domain movements. Rather, it accomplishes active site assembly through movement of the DNA template strand, and by subtle alterations in side chain conformations.

Pol μ has even more permissive substrate requirements, fewer hydrogen bonding interactions with the DNA template strand, and decreased interaction with the downstream 5'-phosphate. Template-dependent synthesis by Pol μ is more distributive and does not require that the primer for extension be paired with a complementary template strand base. In addition, Pol μ can also conduct template-independent synthesis. For these types of reactions, Loop I and the active site histidine (H329) appear to be critical to stabilize position of the primer terminus.

Although there are as yet no structural data for the apoprotein or binary complexes of Pol μ , it seems likely that neither Pol λ , Pol μ , nor TdT require major protein domain movements for their catalytic cycles. The fact that these are all relatively inaccurate polymerase correlates with the generally low fidelity of Y Family DNA polymerases, which are also suggested to perform catalysis without major subdomain movements [86]. The four mammalian Family Y polymerase are strongly implicated in translesion synthesis, and interestingly, several studies report that Pol λ and Pol μ can also perform translesion synthesis [87,88].

The human genome encodes three Family A polymerases and four Family B polymerase whose functions are only partly understood. As yet, structures do not exist for any of the mammalian Family A or Family B enzymes. As illustrated by the structural portrait of mammalian Family X polymerases discussed here, one can anticipate that having such structures would greatly enhance our understanding of their biological functions.

Acknowledgements

We thank M. Miller and L. Pedersen for critical reading and thoughtful comments on the manuscript. This research was funded by the Division of Intramural Research of the National Institute of Environmental Health Sciences, US National Institutes of Health.

References

1. Bebenek K, Kunkel TA. Functions of DNA Polymerases. *Adv Protein Chem* 2004;69:137–165. [PubMed: 15588842]
2. Matsumoto Y, Kim K. Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair. *Science* 1995;269:699–702. [PubMed: 7624801]
3. Garcia-Diaz M, Bebenek K, Kunkel TA, Blanco L. Identification of an intrinsic 5'-deoxyribose-5-phosphate lyase activity in human DNA polymerase lambda: a possible role in base excision repair. *J Biol Chem* 2001;276:34659–34663. [PubMed: 11457865]
4. Srivastava DK, Berg BJ, Prasad R, Molina JT, Beard WA, Tomkinson AE, Wilson SH. Mammalian abasic site base excision repair. Identification of the reaction sequence and rate-determining steps. *J Biol Chem* 1998;273:21203–21209. [PubMed: 9694877]
5. Braithwaite EK, Prasad R, Shock DD, Hou EW, Beard WA, Wilson SH. DNA polymerase lambda mediates a back-up base excision repair activity in extracts of mouse embryonic fibroblasts. *J Biol Chem* 2005;280:18469–18475. [PubMed: 15749700]
6. Barnes DE, Lindahl T. Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu Rev Genet* 2004;38:445–476. [PubMed: 15568983]
7. Bork P, Hofmann K, Bucher P, Neuwald AF, Altschul SF, Koonin EV. A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins. *Faseb J* 1997;11:68–76. [PubMed: 9034168]
8. Mahajan KN, Nick McElhinny SA, Mitchell BS, Ramsden DA. Association of DNA polymerase mu (pol mu) with Ku and ligase IV: role for pol mu in end-joining double-strand break repair. *Mol Cell Biol* 2002;22:5194–5202. [PubMed: 12077346]
9. Fan W, Wu X. DNA polymerase lambda can elongate on DNA substrates mimicking non-homologous end joining and interact with XRCC4-ligase IV complex. *Biochem Biophys Res Commun* 2004;323:1328–1333. [PubMed: 15451442]
10. Lee JW, Blanco L, Zhou T, Garcia-Diaz M, Bebenek K, Kunkel TA, Wang Z, Povirk LF. Implication of DNA polymerase lambda in alignment-based gap filling for nonhomologous DNA end joining in human nuclear extracts. *J Biol Chem* 2004;279:805–811. [PubMed: 14561766]
11. Nick McElhinny SA, Havener JM, Garcia-Diaz M, Juarez R, Bebenek K, Kee BL, Blanco L, Kunkel TA, Ramsden DA. A gradient of template dependence defines distinct biological roles for family X polymerases in nonhomologous end joining. *Mol Cell* 2005;19:357–366. [PubMed: 16061182]
12. Schatz DG. V(D)J recombination. *Immunol Rev* 2004;200:5–11. [PubMed: 15242391]
13. BOLLUM F. Terminal deoxynucleotidyl transferase. *The Enzymes* 1974;10:145–171.

14. Gilfillan S, Benoist C, Mathis D. Mice lacking terminal deoxynucleotidyltransferase: adult mice with a fetal antigen receptor repertoire. *Immunol Rev* 1995;148:201–219. [PubMed: 8825288]
15. Bertocci B, De Smet A, Weill JC, Reynaud CA. Nonoverlapping functions of DNA polymerases mu, lambda, and terminal deoxynucleotidyltransferase during immunoglobulin V(D)J recombination in vivo. *Immunity* 2006;25:31–41. [PubMed: 16860755]
16. Bertocci B, De Smet A, Berek C, Weill JC, Reynaud CA. Immunoglobulin kappa light chain gene rearrangement is impaired in mice deficient for DNA polymerase mu. *Immunity* 2003;19:203–211. [PubMed: 12932354]
17. Pelletier H. Polymerase structures and mechanism. *Science* 1994;266:2025–2026. [PubMed: 7801132]
18. Garcia-Diaz M, Bebenek K, Krahn JM, Blanco L, Kunkel TA, Pedersen LC. A structural solution for the DNA polymerase lambda-dependent repair of DNA gaps with minimal homology. *Mol Cell* 2004;13:561–572. [PubMed: 14992725]
19. Moon AF, Garcia-Diaz M, Bebenek K, Davis BJ, Zhong X, Ramsden DA, Kunkel TA, Pedersen LC. Structural insight into the substrate specificity of DNA polymerase mu. *Nat Struct Mol Biol* 2007;14:45–53. [PubMed: 17159995]
20. Delarue M, Boule JB, Lescar J, Expert-Bezancon N, Jourdan N, Sukumar N, Rougeon F, Papanicolaou C. Crystal structures of a template-independent DNA polymerase: murine terminal deoxynucleotidyltransferase. *Embo J* 2002;21:427–439. [PubMed: 11823435]
21. Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J. Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. *Science* 1994;264:1891–1903. [PubMed: 7516580]
22. Beard WA, Wilson SH. Structural insights into the origins of DNA polymerase fidelity. *Structure* 2003;11:489–496. [PubMed: 12737815]
23. Sobol RW, Prasad R, Evenski A, Baker A, Yang XP, Horton JK, Wilson SH. The lyase activity of the DNA repair protein beta-polymerase protects from DNA-damage-induced cytotoxicity. *Nature* 2000;405:807–810. [PubMed: 10866204]
24. Garcia-Diaz M, Bebenek K, Krahn JM, Kunkel TA, Pedersen LC. A closed conformation for the Pol lambda catalytic cycle. *Nat Struct Mol Biol* 2005;12:97–98. [PubMed: 15608652]
25. Nick McElhinny SA, Ramsden DA. Sibling rivalry: competition between Pol X family members in V(D)J recombination and general double strand break repair. *Immunol Rev* 2004;200:156–164. [PubMed: 15242403]
26. Rooney S, Chaudhuri J, Alt FW. The role of the non-homologous end-joining pathway in lymphocyte development. *Immunol Rev* 2004;200:115–131. [PubMed: 15242400]
27. Ma Y, Lu H, Schwarz K, Lieber MR. Repair of double-strand DNA breaks by the human nonhomologous DNA end joining pathway: the iterative processing model. *Cell Cycle* 2005;4:1193–1200. [PubMed: 16082219]
28. Ma Y, Lu H, Tiffin B, Goodman MF, Shimazaki N, Koiwai O, Hsieh C, Schwarz K, Lieber MR. A biochemically defined system for mammalian nonhomologous DNA end-joining. *Mol Cell* 2004;16:701–713. [PubMed: 15574326]
29. Zhang X, Morera S, Bates PA, Whitehead PC, Coffey AI, Hainbuecher K, Nash RA, Sternberg MJ, Lindahl T, Freemont PS. Structure of an XRCC1 BRCT domain: a new protein-protein interaction module. *Embo J* 1998;17:6404–6411. [PubMed: 9799248]
30. Steitz TA. DNA- and RNA-dependent DNA polymerases. *Curr Opin Struct Biol* 1993;3:31–38.
31. Sawaya MR, Pelletier H, Kumar A, Wilson SH, Kraut J. Crystal structure of rat DNA polymerase beta: evidence for a common polymerase mechanism. *Science* 1994;264:1930–1935. [PubMed: 7516581]
32. Menge KL, Hostomsky Z, Nodes BR, Hudson GO, Rahmati S, Moomaw EW, Almasy RJ, Hostomska Z. Structure-function analysis of the mammalian DNA polymerase beta active site: role of aspartic acid 256, arginine 254, and arginine 258 in nucleotidyl transfer. *Biochemistry* 1995;34:15934–15942. [PubMed: 8519750]
33. Steitz TA, Smerdon SJ, Jager J, Joyce CM. A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science* 1994;266:2022–2025. [PubMed: 7528445]

34. Batra VK, Beard WA, Shock DD, Krahn JM, Pedersen LC, Wilson SH. Magnesium Induced Assembly of a Complete DNA Polymerase Catalytic Complex. *Structure* 2006;14:757–766. [PubMed: 16615916]
35. Lin P, Pedersen LC, Batra VK, Beard WA, Wilson SH, Pedersen LG. Energy analysis of chemistry for correct insertion by DNA polymerase beta. *Proc Natl Acad Sci U S A* 2006;103:13294–13299. [PubMed: 16938895]
36. Burgers PM, Eckstein F. A study of the mechanism of DNA polymerase I from *Escherichia coli* with diastereomeric phosphorothioate analogs of deoxyadenosine triphosphate. *J Biol Chem* 1979;154:6889–6893. [PubMed: 378995]
37. Gupta AP, Benkovic SJ. Stereochemical course of the 3'----5'-exonuclease activity of DNA polymerase I. *Biochemistry* 1984;23:5874–5881. [PubMed: 6098302]
38. Brautigam CA, Steitz TA. Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr Opin Struct Biol* 1998;8:54–63. [PubMed: 9519297]
39. Herschlag D, Jencks WP. The effect of divalent metal ions on the rate and transition-state structure of phosphoryl-transfer reactions. *J Am Chem Soc* 1987;109:4665–4674.
40. Beard WA, Wilson SH. Structural insights into DNA polymerase beta fidelity: hold tight if you want it right. *Chem Biol* 1998;5:R7–13. [PubMed: 9479474]
41. Sawaya MR, Prasad R, Wilson SH, Kraut J, Pelletier H. Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* 1997;36:11205–11215. [PubMed: 9287163]
42. Yang L, Beard WA, Wilson SH, Broyde S, Schlick T. Highly organized but pliant active site of DNA polymerase beta: compensatory mechanisms in mutant enzymes revealed by dynamics simulations and energy analyses. *Biophys J* 2004;86:3392–3408. [PubMed: 15189842]
43. Radhakrishnan R, Arora K, Wang Y, Beard WA, Wilson SH, Schlick T. Regulation of DNA Repair Fidelity by Molecular Checkpoints: "Gates" in DNA Polymerase beta's Substrate Selection. *Biochemistry* 2006;45:15142–15156. [PubMed: 17176036]
44. Beard WA, Prasad R, Wilson SH. Activities and mechanism of DNA polymerase beta. *Methods Enzymol* 2006;408:91–107. [PubMed: 16793365]
45. Beard WA, Osheroff WP, Prasad R, Sawaya MR, Jaju M, Wood TG, Kraut J, Kunkel TA, Wilson SH. Enzyme-DNA interactions required for efficient nucleotide incorporation and discrimination in human DNA polymerase beta. *J Biol Chem* 1996;271:12141–12144. [PubMed: 8647805]
46. Osheroff WP, Beard WA, Wilson SH, Kunkel TA. Base substitution specificity of DNA polymerase beta depends on interactions in the DNA minor groove. *J Biol Chem* 1999;274:20749–20752. [PubMed: 10409611]
47. Osheroff WP, Beard WA, Yin S, Wilson SH, Kunkel TA. Minor groove interactions at the DNA polymerase beta active site modulate single-base deletion error rates. *J Biol Chem* 2000;275:28033–28038. [PubMed: 10851238]
48. Yang L, Beard WA, Wilson SH, Broyde S, Schlick T. Polymerase beta simulations suggest that Arg258 rotation is a slow step rather than large subdomain motions per se. *J Mol Biol* 2002;317:651–671. [PubMed: 11955015]
49. Shevelev I, Blanca G, Villani G, Ramadan K, Spadari S, Hubscher U, Maga G. Mutagenesis of human DNA polymerase lambda: essential roles of Tyr505 and Phe506 for both DNA polymerase and terminal transferase activities. *Nucleic Acids Res* 2003;31:6916–6925. [PubMed: 14627824]
50. Beard WA, Shock DD, Vande Berg BJ, Wilson SH. Efficiency of correct nucleotide insertion governs DNA polymerase fidelity. *J Biol Chem* 2002;277:47393–47398. [PubMed: 12370169]
51. Boule JB, Rougeon F, Papanicolaou C. Terminal deoxynucleotidyl transferase indiscriminately incorporates ribonucleotides and deoxyribonucleotides. *J Biol Chem* 2001;276:31388–31393. [PubMed: 11406636]
52. Nick McElhinny SA, Ramsden DA. Polymerase mu is a DNA-directed DNA/RNA polymerase. *Mol Cell Biol* 2003;23:2309–2315. [PubMed: 12640116]
53. Ruiz JF, Juarez R, Garcia-Diaz M, Terrados G, Picher AJ, Gonzalez-Barrera S, Fernandez de Henestrosa AR, Blanco L. Lack of sugar discrimination by human Pol mu requires a single glycine residue. *Nucleic Acids Res* 2003;31:4441–4449. [PubMed: 12888504]

54. Prasad R, Beard WA, Wilson SH. Studies of gapped DNA substrate binding by mammalian DNA polymerase beta. Dependence on 5'-phosphate group. *J Biol Chem* 1994;269:18096–18101. [PubMed: 8027071]
55. Piersen CE, Prasad R, Wilson SH, Lloyd RS. Evidence for an imino intermediate in the DNA polymerase beta deoxyribose phosphate excision reaction. *J Biol Chem* 1996;271:17811–17815. [PubMed: 8663612]
56. Prasad R, Batra VK, Yang XP, Krahn JM, Pedersen LC, Beard WA, Wilson SH. Structural insight into the DNA polymerase beta deoxyribose phosphate lyase mechanism. *DNA Repair (Amst)* 2005;4:1347–1357. [PubMed: 16172026]
57. Garcia-Diaz M, Bebenek K, Gao G, Pedersen LC, London RE, Kunkel TA. Structure-function studies of DNA polymerase lambda. *DNA Repair (Amst)* 2005;4:1358–1367. [PubMed: 16213194]
58. Prasad R, Beard WA, Chyan JY, Maciejewski MW, Mullen GP, Wilson SH. Functional analysis of the amino-terminal 8-kDa domain of DNA polymerase beta as revealed by site-directed mutagenesis. DNA binding and 5'-deoxyribose phosphate lyase activities. *J Biol Chem* 1998;273:11121–11126. [PubMed: 9556598]
59. Singhal RK, Prasad R, Wilson SH. DNA polymerase beta conducts the gap-filling step in uracil-initiated base excision repair in a bovine testis nuclear extract. *J Biol Chem* 1995;270:949–957. [PubMed: 7822335]
60. Garcia-Diaz M, Bebenek K, Sabariego R, Dominguez O, Rodriguez J, Kirchhoff T, Garcia-Palomero E, Picher AJ, Juarez R, Ruiz JF, Kunkel TA, Blanco L. DNA polymerase lambda, a novel DNA repair enzyme in human cells. *J Biol Chem* 2002;277:13184–13191. [PubMed: 11821417]
61. Doherty AJ, Serpell LC, Ponting CP. The helix-hairpin-helix DNA-binding motif: a structural basis for non-sequence-specific recognition of DNA. *Nucleic Acids Res* 1996;24:2488–2497. [PubMed: 8692686]
62. Mullen GP, Wilson SH. DNA polymerase beta in abasic site repair: a structurally conserved helix-hairpin-helix motif in lesion detection by base excision repair enzymes. *Biochemistry* 1997;36:4713–4717. [PubMed: 9125491]
63. Singhal RK, Wilson SH. Short gap-filling synthesis by DNA polymerase beta is processive. *J Biol Chem* 1993;268:15906–15911. [PubMed: 8340415]
64. Roettger M, Fiala K, Sompalli S, Dong Y, Suo Z. Pre-Steady-State Kinetic Studies of the Fidelity of Human DNA Polymerase Mu. *Biochemistry* 2004;43:13827–13838. [PubMed: 15504045]
65. Quintana-Hau JD, Uribe-Luna S, Espinosa-Lara M, Maldonado-Rodriguez R, Logsdon N, Beattie KL. Construction and expression of a chimeric gene encoding human terminal deoxynucleotidyltransferase and DNA polymerase beta. *Gene* 1995;163:289–294. [PubMed: 7590283]
66. Aoufouchi S, Flatter E, Dahan A, Faili A, Bertocci B, Storck S, Delbos F, Cocea L, Gupta N, Weill JC, Reynaud CA. Two novel human and mouse DNA polymerases of the polX family. *Nucleic Acids Res* 2000;28:3684–3693. [PubMed: 10982892]
67. Dominguez O, Ruiz JF, Lain de Lera T, Garcia-Diaz M, Gonzalez MA, Kirchhoff T, Martinez AC, Bernad A, Blanco L. DNA polymerase mu (Pol mu), homologous to TdT, could act as a DNA mutator in eukaryotic cells. *Embo J* 2000;19:1731–1742. [PubMed: 10747040]
68. Juarez R, Ruiz JF, McElhinny SA, Ramsden D, Blanco L. A specific loop in human DNA polymerase mu allows switching between creative and DNA-instructed synthesis. *Nucleic Acids Res* 2006;34:4572–4582. [PubMed: 16963491]
69. Lieber MR. The polymerases for V(D)J recombination. *Immunity* 2006;25:7–9. [PubMed: 16860749]
70. Bebenek K, Garcia-Diaz M, Blanco L, Kunkel TA. The frameshift infidelity of human DNA polymerase lambda. Implications for function. *J Biol Chem* 2003;278:34685–34690. [PubMed: 12829698]
71. Picher AJ, Garcia-Diaz M, Bebenek K, Pedersen LC, Kunkel TA, Blanco L. Promiscuous mismatch extension by human DNA polymerase lambda. *Nucleic Acids Res* 2006;34:3259–3266. [PubMed: 16807316]
72. Zhang Y, Wu X, Yuan F, Xie Z, Wang Z. Highly frequent frameshift DNA synthesis by human DNA polymerase mu. *Mol Cell Biol* 2001;21:7995–8006. [PubMed: 11689691]

73. Tiffin B, Kobayashi S, Bertram JG, Goodman MF. To slip or skip, visualizing frameshift mutation dynamics for error-prone DNA polymerases. *J Biol Chem* 2004;279:45360–45368. [PubMed: 15339923]
74. Bebenek K, Kunkel TA. Streisinger revisited: DNA synthesis errors mediated by substrate misalignments. *Cold Spring Harb Symp Quant Biol* 2000;65:81–91. [PubMed: 12760023]
75. Krahn JM, Beard WA, Miller H, Grollman AP, Wilson SH. Structure of DNA polymerase beta with the mutagenic DNA lesion 8-oxodeoxyguanine reveals structural insights into its coding potential. *Structure* 2003;11:121–127. [PubMed: 12517346]
76. Krahn JM, Beard WA, Wilson SH. Structural insights into DNA polymerase beta deterrents for misincorporation support an induced-fit mechanism for fidelity. *Structure* 2004;12:1823–1832. [PubMed: 15458631]
77. Batra VK, Beard WA, Shock DD, Pedersen LC, Wilson SH. Nucleotide-induced DNA polymerase active site motions accommodating a mutagenic DNA intermediate. *Structure* 2005;13:1225–1233. [PubMed: 16084394]
78. Batra VK, Shock DD, Prasad R, Beard WA, Hou EW, Pedersen LC, Sayer JM, Yagi H, Kumar S, Jerina DM, Wilson SH. Structure of DNA polymerase beta with a benzo[c]phenanthrene diol epoxide-adducted template exhibits mutagenic features. *Proc Natl Acad Sci U S A* 2006;103:17231–17236. [PubMed: 17079493]
79. Beard WA, Shock DD, Wilson SH. Influence of DNA structure on DNA polymerase beta active site function: extension of mutagenic DNA intermediates. *J Biol Chem* 2004;279:31921–31929. [PubMed: 15145936]
80. Miller H, Prasad R, Wilson SH, Johnson F, Grollman AP. 8-oxodGTP incorporation by DNA polymerase beta is modified by active-site residue Asn279. *Biochemistry* 2000;39:1029–1033. [PubMed: 10653647]
81. Kunkel TA. The mutational specificity of DNA polymerase-beta during in vitro DNA synthesis. Production of frameshift, base substitution, and deletion mutations. *J Biol Chem* 1985;260:5787–5796. [PubMed: 3988773]
82. Garcia-Diaz M, Kunkel TA. Mechanism of a genetic glissando: structural biology of indel mutations. *Trends Biochem Sci* 2006;31:206–214. [PubMed: 16545956]
83. Streisinger G, Okada Y, Emrich J, Newton J, Tsugita A, Terzaghi E, Inouye M. Frameshift mutations and the genetic code. This paper is dedicated to Professor Theodosius Dobzhansky on the occasion of his 66th birthday. *Cold Spring Harb Symp Quant Biol* 1966;31:77–84. [PubMed: 5237214]
84. Garcia-Diaz M, Bebenek K, Krahn JM, Pedersen LC, Kunkel TA. Structural analysis of strand misalignment during DNA synthesis by a human DNA polymerase. *Cell* 2006;124:331–342. [PubMed: 16439207]
85. Beard WA, Wilson SH. Structure and mechanism of DNA polymerase Beta. *Chem Rev* 2006;106:361–382. [PubMed: 16464010]
86. Yang W. Damage repair DNA polymerases Y. *Curr Opin Struct Biol* 2003;13:23–30. [PubMed: 12581656]
87. Maga G, Villani G, Ramadan K, Shevelev I, Tanguy Le Gac N, Blanco L, Blanca G, Spadari S, Hubscher U. Human DNA polymerase lambda functionally and physically interacts with proliferating cell nuclear antigen in normal and translesion DNA synthesis. *J Biol Chem* 2002;277:48434–48440. [PubMed: 12368291]
88. Zhang Y, Wu X, Guo D, Rechko O, Taylor JS, Geacintov NE, Wang Z. Lesion bypass activities of human DNA polymerase mu. *J Biol Chem* 2002;277:44582–44587. [PubMed: 12228225]
89. DeLano, WL. The PyMOL Molecular Graphic System User's Manual. DeLano Scientific, San Carlos, CA, USA, 2002: 2002.

Figure 1. Mammalian Family X polymerases. (A) Structure-based primary sequence alignment. Catalytic aspartates, boxed in black; residues involved in minor groove interactions, orange; residues involved in stacking interactions with nucleotides, yellow; dRP lyase nucleophile, green; active site histidines (Pol μ H329, TdT H342), cyan; residue putatively stabilizing frameshift intermediate, magenta. Secondary structural elements are blue (α -helices), green (β -strands), and yellow (3¹⁰ helix) boxes. (B) Domain organization. α -helices are depicted as cylinders, and labeled alphabetically. β -strands are shown as directional arrows, labeled numerically. The bound DNA duplex is shown in the binding cleft: template (T), orange; upstream primer (P), khaki; and downstream primer (D), brown. The position of the 3'-OH and downstream 5'-phosphate are marked with green and yellow stars, respectively. The incoming nucleotide is cyan. The ternary complex of Pol β (PDB code 2FMS) was used for this illustration. (C) NMR solution structure of the Pol μ (maroon) and TdT (purple) BRCT domains. α -helices are labeled alphabetically and β -strands numerically. Numbering only applies within the BRCT domains.

Figure 2. Catalytic mechanism for nucleotidyl transfer. The primer terminal residue (orange) is drawn in stick, with the O3' in red. The incoming nucleotide is cyan, with its phosphates in magenta. Protein residues from the ternary structure of Pol β (PDB code 2FMS) are khaki. Magnesium ions are purple and water molecules completing coordination of the metal ions are small red spheres.

Figure 3. Comparison of binary and ternary complexes for Pol β and Pol λ . (A) Superposition of binary (PDB code 1 BPX; protein, light green; DNA, khaki) and ternary (PDB code 2FMS; protein, dark green; DNA, orange) crystal structures of Pol β (right). Left: Amino acid side chain motions within the active site. Middle: Variations in thumb subdomain position between the 'open' (light green) and 'closed' (dark green) forms of the enzyme. (B) Superposition of binary (PDB code 1XSL; protein, light blue; DNA, khaki) and ternary (PDB code 1XSN; protein, dark blue; DNA, orange) crystal structures of Pol λ (left). Right: Highlights of amino acid side chain motions in the active site of Pol λ . Binary complex: protein, light blue; DNA khaki. Ternary complex: protein, dark blue; DNA, orange. The incoming nucleotide is drawn in stick (cyan). The magnesium ion in metal site B shown as a purple sphere. (C) Superposition of the pre-catalytic (PDB code 2PFO, orange) and the post-catalytic (PDB code 1XSP, khaki) ternary complexes of Pol λ . The dUMPNNPP from the pre-catalytic complex is cyan. The manganese ion in metal site A is shown as a yellow sphere, and the magnesium ion in metal site B is shown

as a purple sphere. The α -phosphate undergoing stereochemical inversion is circled in dashed light blue.

Figure 4.

Family X polymerase active site organization. (A) Superposition of ternary complexes of Pol β (PDB code 2FMS; protein, green; DNA, khaki) and Pol λ (PDB code 2PFO; protein, blue; DNA, orange). Protein side chains from Pol β are drawn in light green, while those for Pol λ are light blue. (B) Superposition of the ternary complexes of Pol β (PDB code 2FMS; protein, green; DNA, khaki) and Pol μ (PDB code 2IHM; protein, maroon; DNA, orange). Protein side chains from Pol β are drawn in light green, while those for Pol μ are pink. Putative hydrogen bonding interactions between Pol μ H329 and the primer terminal phosphate or the γ-phosphate of the incoming nucleotide are shown as black dashed lines. Magnesium and sodium ions are shown as purple and green spheres, respectively. (C) Superposition of ternary complexes of

Pol μ (PDB code 2IHM; protein, maroon; DNA, khaki) and TdT apoprotein (PDB code 1JMS; protein, purple). Protein side chains from Pol μ are drawn in pink, while those for TdT are light purple. The incoming nucleotide from the structure of Pol μ is drawn in khaki. Magnesium and sodium ions are shown as purple and green spheres, respectively.

Figure 5.

Schematic representation of substrate preferences of Family X polymerases. The primer terminal residue is red and the incoming nucleotide is black. (A) Pol β, Pol λ, and Pol μ can fill small single-strand gaps in a template-dependent manner. (B) Pol λ and Pol μ can fill double-strand gaps with at least one complementary base pair proximal to the gap to be filled. (C) Pol μ, aided by Loop I and H329, can fill a gap from a primer terminal residue with no complementary template strand nucleotide opposite that position. (D) TdT, with the aid of Loop I and H342, utilizes DNA substrates with a 3'-primer terminal overhang, and carries out polymerization in a template-independent manner. Pol μ is also capable of template-independent synthesis in this fashion, but to a lesser extent.

Figure 6.

Comparison of the 8 kDa domain 5'-phosphate binding pocket. Electrostatic surface potential plots were calculated using the Adaptive Poisson-Boltzmann Solver tool in PyMOL [89]. The downstream primer is drawn in orange. All DNA and protein residues are labeled in white. The 5'-phosphate binding pocket in the 8 kDa domain of: (A) Pol β (PDB code 2FMS). (B) Pol λ (PDB code 1XSN). (C) Pol μ (PDB code 2IHM). (D) TdT (PDB code 1JMS). The downstream primer from PDB code 2IHM was modeled into the 5'-phosphate binding pocket of TdT, and is drawn in light gray.

Figure 7.

Comparing the 8 kDa domain Helix-hairpin-Helix motifs. (A) HhH motifs from the 8 kDa domains of Pol β (green) and Pol λ (blue) are shown in a ribbon diagram. Side chains in the hairpins of Pol β and Pol λ are light green and light blue, respectively. (B) HhH motifs from the 8 kDa domains of Pol μ (maroon) and TdT (purple). Side chains in the hairpins of Pol μ and TdT are pink and light purple, respectively. (C) HhH motifs from the 8 kDa domains of Pol μ (maroon) and Pol β (green). Side chains from residues in the hairpins of Pol μ and Pol β are pink and light green, respectively. Distortions in α -helix C are highlighted by a dashed circle (orange).

Figure 8.

Comparing Loop I in Family X polymerases. Secondary structural elements of Pol β are shown in the background in light gray. The DNA is orange. β-strands 3 and 4 are colored ribbons and labeled in black. (A) Superposition showing the Loop I region of Pol β (PDB code 2FMS, green) and TdT (PDB code 1JMS, purple). (B) Superposition showing the Loop I region of TdT (PDB code 1JMS, purple) and Pol μ (PDB code 2IHM, maroon). Loop I is disordered in Pol μ. The dashed maroon line shows a hypothetical path of the loop in relation to the DNA binding cleft. (C) Superposition of ternary complexes showing the Loop I region of Pol β (PDB code 2FMS, green) and Pol λ (PDB code 1XSN, blue).

Figure 9.

Accommodating distorted DNA substrates. (A) Superposition of a gapped substrate with an A-A mismatch at the primer terminus (PDB code 1ZJN, orange) and a normal gapped DNA substrate (PDB code 2FMS, khaki). The identities of the primer terminal and nascent base pairs are indicated (normal ternary in khaki, and mismatched ternary in orange). Identities of the DNA bases in the mismatch structure are shown to the right of each base (orange). (B) Superposition of a normal gapped DNA substrate (PDB code 2FMS, khaki) with canonical G-C base pair and a gapped DNA substrate containing an 8oxodG at the templating position (PDB code 1MQ3, orange), paired with incoming dCTP. (C) Pol λ accommodating an extra base. Superposition of a normal ternary complex of Pol λ (PDB code 1XSN; DNA, orange) with the

frameshift intermediate (PDB code 2BCV; DNA khaki). Secondary structural elements for the frameshift structure are shown as blue ribbons. The extrahelical nucleotide is depicted in cyan. β -strand 8 harbors the K544 residue proximal to the extrahelical nucleotide, near the top of the panel.

Table 1

Selective list of NMR and X-ray crystal structures for Family X DNA polymerases

Polymerase	PDB Code	DNA	Incoming Nucleotide
hPol β	1BPX	1 nt gap	ddCTP
	1BPY	1 nt gap	
	1BPZ	nicked DNA	
	1TV9	mismatch A:C	
	1TVA	mismatch T:C	
	1ZJM	primer A:A mismatch	dGTP ddCTP dUMPNPP dUMPNPP dAMP dCTP ddCTP
	1ZJN	primer A:A mismatch	
	2FMP	1 nt gap ($\text{Na}^+/\text{Mg}^{+2}$)	
	2FMQ	1 nt gap ($\text{Na}^+/\text{Mg}^{+2}$)	
	2FMS	1 nt gap ($\text{Mg}^{+2}/\text{Mg}^{+2}$)	
	1MQ2	gapped DNA 8oxodG	
	1MQ3	gapped DNA 8oxodG	
	2I9G	B[c]Ph DE adduct	
hPol λ	1RZT	2nt gap	ddTTP dUMPNPP PP _i
	1XSL	1 nt gap	
	1XSN	1 nt gap	
	2PFO	1 nt gap ($\text{Mn}^{+2}/\text{Mg}^{+2}$)	
	1XSP	nicked DNA	
	2BCQ	frameshift (extrahelical dTMP)	ddTTP
	2BCR	frameshift (extrahelical dAMP)	
	2BCS	frameshift (extrahelical dCMP)	
	2BCV	frameshift (extrahelical dTMP)	
	2BCU	T:T mismatch (unpaired dAMP)	
mTdT	1JMS	brominated 4 nt ssDNA	ddATP
	1KEJ		
	1KDH		
mPol μ	2COE*	1nt gap	ddTTP
	2IHM		
	2DUN*		

Pol β and Pol λ X-ray crystal structures with 39 kDa polymerase domain. Pol μ and TdT X-ray crystal structures with 40kDa polymerase domain. 2COE and 2DUN (marked by *) are NMR spectroscopy structures of the hTdT and hPol μ BRCT domains, respectively.

Table 2
Comparison of key structural and behavioral aspects of Family X DNA polymerases

	Pol β	Pol λ	Pol μ	TdT
8 kDa domain				
Lyase activity	Yes	Yes	Yes	Yes
5'-phosphate pocket	Yes	Yes	No	No
HhH1	Yes	Yes	Yes, but distorted	Few positively charged residues Yes, but distorted
Catalysis (Nucleotidyl Transfer)				
Two-metal mechanism	Yes	Yes	Yes	Unknown, but likely
Active site assembly	Subdomain motions & side chain motions	DNA repositioning, side chain motions	Unknown, but likely few	Unknown, but likely
Processivity	Processive with 5'-phosphate	Processive with 5'-phosphate	Distributive	Distributive
YF motif	Yes	Yes	No, GW motif	No, GW motif
Nucleotide selectivity	dNTPs	dNTPs	dNTPs & rNTPs	dNTPs & rNTPs
DNA strand interactions	Many, mostly with template	Fewer, most with template opposite primer	Fewer, most with upstream primer	Unknown
Minor groove interactions	Y271, F272, N279, R283	Y505, F506, N513, R517	R447, possible	Unknown
HhH2	Yes	Yes	Yes	Yes
Substrate Specificity				
Gap-filling	Yes	Yes	Yes	No
Template-dependence	Yes	Yes	Yes	No
Template-independence	No	No	Yes	Yes
DNA Repair				
Participation in BER	Yes	Yes	No	No
Ser/Pro-rich domain	No	Yes	No	No
BRCT domain	No	Yes	Yes	Yes
NHEJ	No	Yes	Yes	Perhaps
DSB microhomology	N/A	Yes, needs at least one complementary base pair	No homology required	Template-independent
V(D)J recombination	No	Yes, immunoglobulin heavy chain	Yes, immunoglobulin light chain	Yes, N-addition
Active site histidine	No	No	Yes	Yes
Fidelity				
Loop I	No	Small	Larger, not seen occluding template strand	Larger, likely occludes template strand binding
Misalignment synthesis	Yes, high in homopolymeric runs	Yes, high error rate	Yes, error rate unknown	Template-independent
Translesion synthesis	Some	Some	Some, error rate unknown	Template-independent
Mismatch extension	Some	Some	Some, error rate unknown	Template-independent