We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Evolving DNA Repair Polymerases: From Double—Strand Break Repair to Base Excision Repair and VDJ Recombination

Maria Jose Martin and Luis Blanco

Additional information is available at the end of the chapter

1. Introduction

Currently five polymerases have been identified in *Escherichia coli*, at least eight in *Saccharomyces cerevisiae*, nine in *Schizosaccharomyces pombe*, and fourteen in humans [1-4]. Based on the primary structure of the catalytic subunits, DNA polymerases have been classified into different families. Eukaryotic organisms have four families: A family (Polγ, Polθ and Polν), B family (Polα, Polδ, Polε and Polζ), X family (Polß, Polλ, Polμ and TdT) and Y family (Polη, Polι, Polκ and Rev1), whose members were discovered in the last decade [5], and are involved in replication through DNA lesions. Another significant development was the discovery of Polλ [6] and Polμ [7], which doubled the number of known enzymes of the X family of DNA polymerases, whose members are involved in DNA repair and generation of variability.

2. Evolution of the X family of DNA polymerases

The members of the X family are present in many organisms in all monophyletic taxa: Eukarya, Bacteria and Archaea, and even viruses with DNA genome [8]. The high degree of conservation at the structural and amino acid sequence levels between X family members suggests that they originate from a common ancestor.

Unlike viruses, prokaryotes and yeast, higher eukaryotes have more than one member of the X family. However, there are species in which no member of this family has been described, like the model organisms *Caenorhabditis elegans* and *Drosophila melanogaster* [2], so it becomes a matter of special interest to learn how they have solved the absence of these DNA polymer-
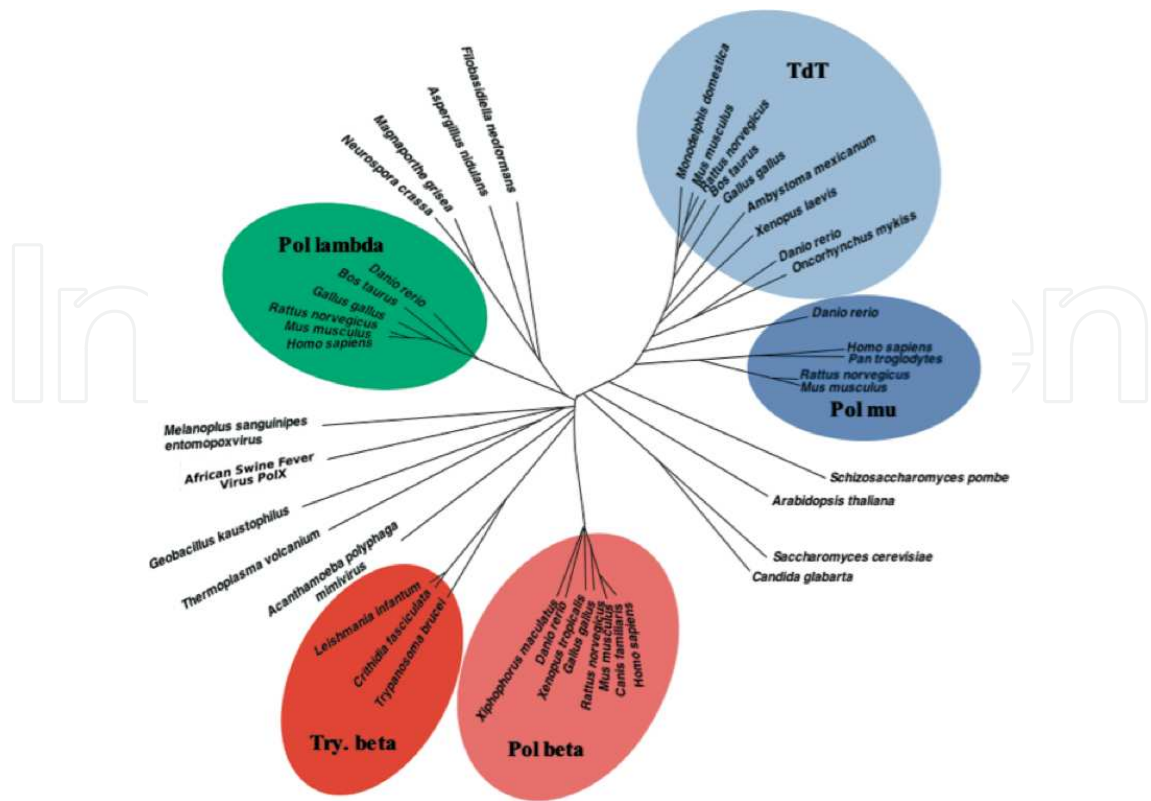
Figure 1. Evolutionary relationships between family X members. An unrooted phylogenetic tree built using a primary sequence alignment of a segment of the catalytic domain. Different enzymes are grouped (shaded areas) into the five main enzyme classes in the family: Polß, Pol λ, Polμ, TdT, and trypanosomatid (Try) Polß-like enzymes.

ases in DNA repair processes. Recent data indicate that recombination repair protein 1, the Drosophila homolog of human AP endonuclease 1 (APE1), interacts with DNA polymerase ζ [9]. It is possible that in protostomes (which include insects and nematodes), APE1-like genes are able to recruit a DNA polymerase other than an X family enzyme to AP sites on DNA. It has been proposed that protostomes evolved from organisms in the coelenterate phylum that lost a Polλ-like gene before other X family DNA polymerases were derived, since it is unlikely that multiple X family genes were lost as soon as coelenterates appeared [10].

Figure 1 shows the phylogenetic relationships between the known members of X family from different organisms. The phylogenetic tree was made using a short and highly conserved segment of the polymerization active site, in order to avoid the presence of accessory domains or small insertions or deletions that may interfere in the analysis. The results suggest that the several subfamilies that can be identified within the X family (Polß, Polλ, Polμ and TdT) have evolved from a common ancestor, perhaps to accommodate different functional requirements. The emergence of more complex organisms seems to promote the specialization of the X family members in order to increase the efficiency of the DNA synthesis processes in which they are involved. The distribution of X family DNA polymerases among different species suggests that the ancestor of the X family DNA polymerase was a Polλ-like gene, which diversified into Polß, Polμ and TdT during evolution. Polλ would have originally been involved in NHEJ to eliminate DNA damage. Subsequently, other X family DNA polymerases would have been

generated in some animals and fungi through gene duplications, acquiring novel roles in DNA metabolism such as in BER and V(D)J recombination. According to very recent results [11], these evolutionary forces driving creation of new polymerases are still taking place among primates: codon-based models of gene evolution yielded statistical support for the recurrent positive selection of Polλ, among other four NHEJ genes during primate evolution: XRCC4, NBS1, Artemis, and CtIP. Moreover, analysis of the mutations on the crystal structures available for XRCC4, Nbs1, and Polλ show that residues under positive selection fall exclusively on the surface of these proteins. Studies of positive local evolution on human populations show that, indeed, a single allele of Polλ has previously been reported to be under positive selection in both Asian and Sub-Sahara African populations [12]. Also, sliding-window analyses and pairwise comparisons of several strains of *Saccharomyces* indicated that several of the yeast NHEJ genes show evidence of positive selection, including POL4 [13]. A first hypothesis explaining the high level of positively selected mutations implies that as certain NHEJ components evolve, compensatory mutations may arise in other NHEJ components to re-optimize protein-protein interactions between the various partners. On the other hand, many viruses such as adenovirus, and retroviruses like HIV, interact with the proteins of the NHEJ pathway as part of their infectious life cycle [14-21]. The Corndog and Omega bacteriophages of mycobacteria have even incorporated the first gene of the bacterial NHEJ pathway, Ku, into their own genome [22]. This viral Ku now evolves under the selective pressures of the virus in order to recruit the bacterial NHEJ ligase, LigD, to circularize phage DNA. Therefore, a second hypothesis would explain the surprisingly rapid evolution of NHEJ genes as an ongoing evolutionary arms race between viruses and these critical genes.
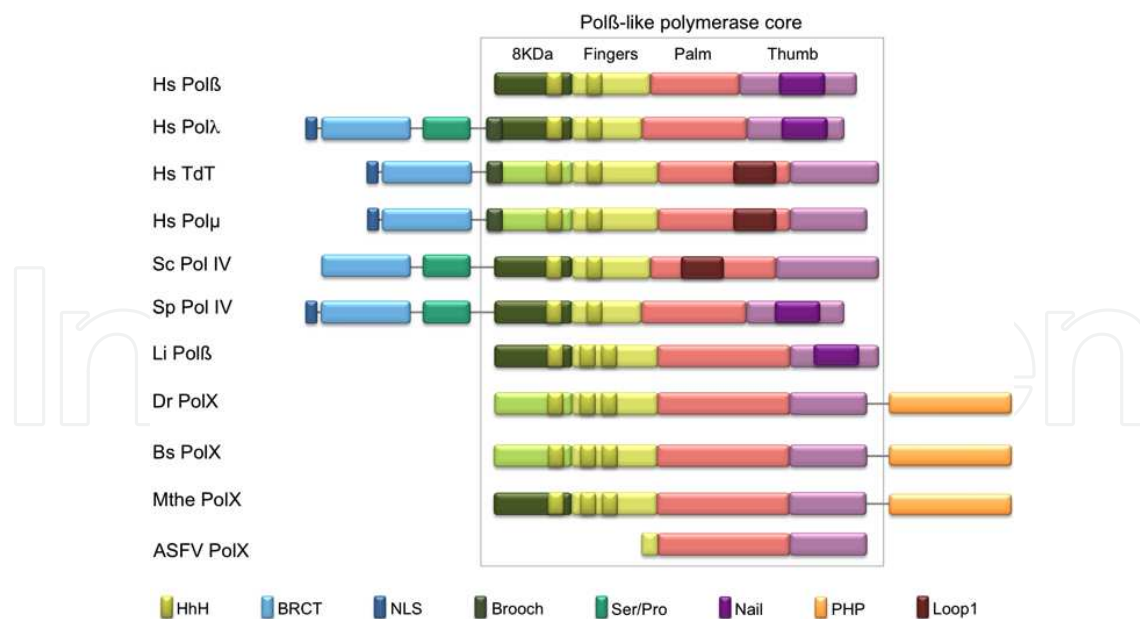
Figure 2. Modular organization of the family X polymerases. Schematic representation of the domains present in family X members from viruses to higher eukaryotes. Regarding the coloring of the 8 kDa domain, dark green represents dRP lyase-containing domains while bright green color indicates the lack of such activity. *Sc: Saccharomyces cerevisiae*; *Sp: Schizosaccharomyces pombe*; *Li: Leishmania infantum*; *ASFV: African swine fever virus*; *Bs: Bacillus subtilis*; *Mthe: Methanobacterium thermoautotrophicum*; *Dr: Deinococcus radiodurans.*

3. Comparative genomic organization of human DNA polymerases from family X

The modular organization of different members of the X family from viruses to eukaryotes indicates the existence of a conserved Polß-type core (Fig. 2), whose minimal version is the PolX from the African swine fever virus (ASFV), which contains only the palm and thumb subdomains of the polymerase domain [8]. The absence of the 8 kDa domain of both ASFV PolX and MSEV (*Melanoplus sanguinipes* entomopoxvirus) may reflect the existence of other proteins encoded by the viral genome to provide the catalytic (dRP lyase) and/or DNA binding properties residing in this domain in most of the DNA polymerases of the X family. Despite the small size of ASFV PolX, it has a second enzymatic activity: the AP-lyase, indicating a possible role in the viral BER pathway [23]. The evolutionary divergence of the members of the X family has occurred by acquisition of additional domains with regulatory properties and/or enzymatic activities. X family members from eubacteria (*Bacillus subtilis*) and Archaea (*Methanobacterium thermoautotrophicum*) have a phosphodiesterase domain (PHP, Fig. 2) fused to the Polß core domain, and thus possess polymerase and nuclease activities in the same polypeptide, a great functional benefit to carry out repair processes in the BER pathway. In eukaryotes there are members of this family from protozoa (*Leishmania infantum*) to mammals. However, there are major differences in the accessory domains that keep a very close relationship with their physiological function. The percentage of similarity at the amino acid sequence level of the Polß core between different members of this family varies from 91% between the Polß enzymes from *Crithidia* and *Leishmania* (LiPolß), and 42% between Polµ and TdT, to 19% identity between LiPolß and TdT [24]. LiPolß shows a 31% of amino-acid identity with mammalian Polß, close to the 32% between Polλ and Polß. Interestingly, both Polß enzymes from *Crithidia* and *Leishmania* present inserts within the core that allow protein-protein and protein-DNA interactions. Contrary to mammals, yeast cells have a single DNA polymerase from the X family, Pol4. Both Pol4 from *S. cerevisiae* and *S. pombe* possess two additional domains at their N-terminus: a BRCT domain followed by a regulatory Ser/Pro domain (Fig. 2). In addition, both Pol4 have a dRP-lyase activity associated with the 8 kDa domain suggesting a role in repair processes such as BER [25, 26]. Although both Pol4 enzymes share a common structural organization, they differ in terms of sequence similarity with their human counterparts. While *Sc*Pol4 is more similar in the composition of the basic Polß structure to Polλ, sharing a 25% of amino-acid identity [25], *Sp*Pol4 is closer to Polµ (27% amino-acid identity) than to Polλ (24% amino-acid identity). Based on sequence similarity one can speculate that, in yeast, *Sp*Pol4 is the orthologue of human Polµ while *Sc*Pol4 could be the orthologue of human Polλ.

The presence of BRCT domains in Pol4, Polλ, Polµ and TdT relates to the role that this domain plays in processes such as V(D)J recombination and NHEJ repair. The BRCT domain of Pol4 mediates the interaction of the polymerase with factors involved in the NHEJ pathway during repair of double-strand breaks in DNA [27, 28]. Similarly, the BRCT domains of Polλ, Polµ and TdT allow these proteins to participate in both NHEJ repair and V(D)J recombination in higher eukaryotes. It is possible that subtle differences in the amino acid sequence of the BRCT domain of each polymerase have great importance in regulating the access of each DNA polymerase to a specific substrate or protein of the route.

Finally, the eukaryotic Polß (initially thought to be exclusive of mammals) has lost some accessory domains during evolution, in a crucial step for its specialization as a housekeeping DNA repair polymerase that protects against the large amount of oxidative damage present as a result of aerobic metabolism. The conservation of the 8 kDa domain (Fig. 2), where the dRP-lyase activity resides, is central for participation in the BER pathway.

4. A BRCT domain as an ancient feature required for NHEJ

The members of the X family of polymerases are recruited to form a complex with the NHEJ core factors XRCC4/Ligase IV and Ku at the DNA break [27, 29, 30]. Recent evidence has shown that BRCT domains can be specifically involved in the interaction with phospho-serine or phospho-threonine containing motifs [31, 32], an ability that may be involved in granting access of regulated proteins to the break, even though no evidence has shown to date a phosphorylation-dependent, BRCT-mediated, interaction of NHEJ factors.

Interestingly, sequence comparisons show that the BRCT of Polµ is most similar to TdT, with 39% sequence identity that includes the residues important for NHEJ-complex formation [33]. That high level of sequence conservation is also observed at the 3D- structural level in the BRCT domains of Polµ (PDB ID: 2DUN) and TdT (PDB ID: 2COE), that in turn exhibit an a/ß motif that is similar to the BRCT found in XRCC1 (PDB ID: 1CDZ), a BER repair protein. The main differences include a shorter α -helix 2 in the TdT BRCT domain, as well as the positioning of the loop connecting α -helix 2 and ß-strand 4. The electrostatic surfaces of Polµ and TdT BRCT domains are also very similar, containing both a positively charged ridge on one face of the protein, and large negatively charged regions on the opposite faces. In the Polµ BRCT the positive ridge is formed by Arg^{44}, Arg^{52}, Arg^{85} and Arg^{86}. This positive patch has been proposed to be involved in the interaction with a phospho-modified protein [33], or most likely in the interaction with the downstream part of the DNA substrate [34]. Point mutations in several residues of the positive ridge as wells as the complete lack of the domain resulted in a diminished interaction with and activity on NHEJ substrates [34, 35]. By using the "brooch" motif (described below) to correctly orient and over-impose the crystals of the BRCT domain and the Polµ core, we found out that one of the positive patches in the BRCT domain perfectly accommodates the downstream part of the DNA substrate (Fig. 3; colored in dark blue). We then modeled the interaction of the BRCT domain of Polµ with the Ku70/Ku80 heterodimer by orienting the DNA substrate. Strikingly, the side of the BRCT domain facing the Ku heterodimer in the model was exactly the one containing the residues reported to be involved in this interaction (Fig. 3; colored in red). According to this model, the portion of the DNA substrate that would be contacted by the BRCT domain flawlessly correlates with the length of the BRCT-specific protection (6 bp) observed in our footprinting assays [34]

This DNA binding function of Polµ BRCT, independent of the core NHEJ factors, may enable a role for Polµ in the alternative NHEJ pathway, which occurs independently of Ku or Ligase IV. Polµ might bind the DNA break based on its own specificity for the 5′-P and then via the BRCT domain and using its terminal transferase activity, be in charge of the additions that
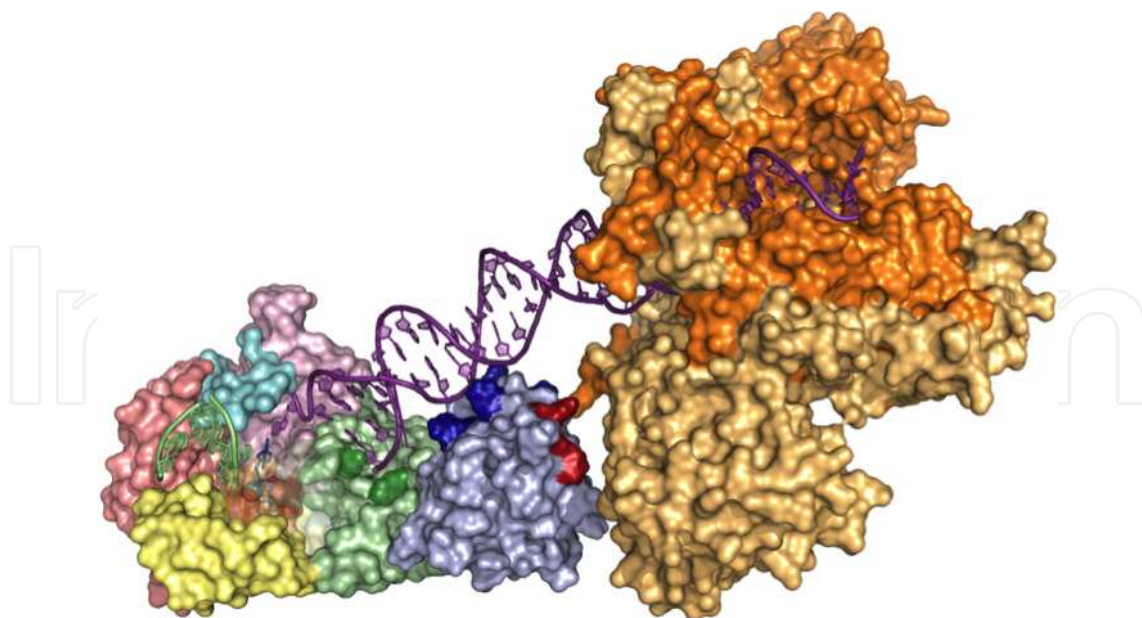
Figure 3. Model of the interaction of Polμ with the Ku heterodimer and the DNA substrate through the BRCT domain.

create the so-called polymerase-generated microhomology. In agreement with this proposed function, recent observations indicate that Polμ BRCT is atypical in the sense of not being involved in dimerization or multimerization. In fact, comparison of the structure of Polμ BRCT with other BRCT domains that effectively dimerize shows important differences, especially regarding R2 helix [36].

The sequence conservation among BRCT domains from family X polymerases is very low, with only 10 residues conserved and five of them (His^{82}, Val^{84}, Leu^{109}, Trp^{114}, and Leu^{115} in human Polλ) involved in the architecture of the domain. The other five (Gly^{54}, Arg^{57}, Gly^{69}, Thr^{81}, and Val^{125} in human Polλ) are exposed to the solvent in the surface of the protein. One of them, Arg^{43} in Polμ (Arg^{57} in Polλ), is implicated in interactions with other components of the NHEJ complex [33].

This low sequence similarity is reflected in structural variations of the family X polymerases' BRCT domains, which in turn influence the interactions established with other NHEJ factors, including an improved/preferential access of the polymerase to the DNA break. Deletion of the BRCT domain in the NHEJ-related polymerases [27, 29, 37], or point-mutagenesis of key-residues [33, 36], block the formation of complexes between the polymerase, Ku and XRCC4/LigaseIV at DNA ends.

The ability of X family polymerases to act during classical NHEJ thus relies on their interactions with other NHEJ factors through their BRCT domains, but PolXs have intrinsic capacities of gap-recognition and binding involving simultaneous recognition of both sides of the gap. As shown for Polß, the polymerase can bind both the template/primer part of the gap and also the template/downstream part, being the latter the strongest anchor point [38]. In the Polß co-crystal with a DNA gap this dual binding is clearly observable: contacts are established with the DNA backbone through a positively charged platform onto which the DNA is leaning.

Such a dual DNA binding is even more crucial for Polλ and Polμ, polymerases not as specialized as Polß in always confronting substrates with continuous template strands (i.e. gaps), but also in charge of bridging two separate DNA ends. The ability to independently bind and orient two DNA ends is thus closely related to their function during NHEJ, but is still found in the more recently evolved Polß as an appropriate solution for gap-filling. This tight binding to both sides of the templating base forces the formation of a sharp bend of 90° in the template strand, that has been proposed to increase nucleotide selectivity and sensitivity to mismatches, and in general is a mechanistic feature used by X polymerases to improve fidelity [39].

5. A small (8 kDa) DNA binding domain, critical for NHEJ

One of the structural features that allows polymerases from X family to bind gapped and NHEJ substrates is the 8 kDa domain (Fig. 4A), located either at the N-terminus (Polß from higher eukaryotes, bacteria and archaea), or at the N-terminal portion just after the flexible linker that contains the Ser-Pro domain (Polλ, Polμ, TdT and yeast Pol4). This 8 kDa domain is involved in contacting several parts of the DNA substrate through different motifs [40], but in some of the members of the X family bears a dRP-lyase activity, highly related to the BER pathway [41, 42].
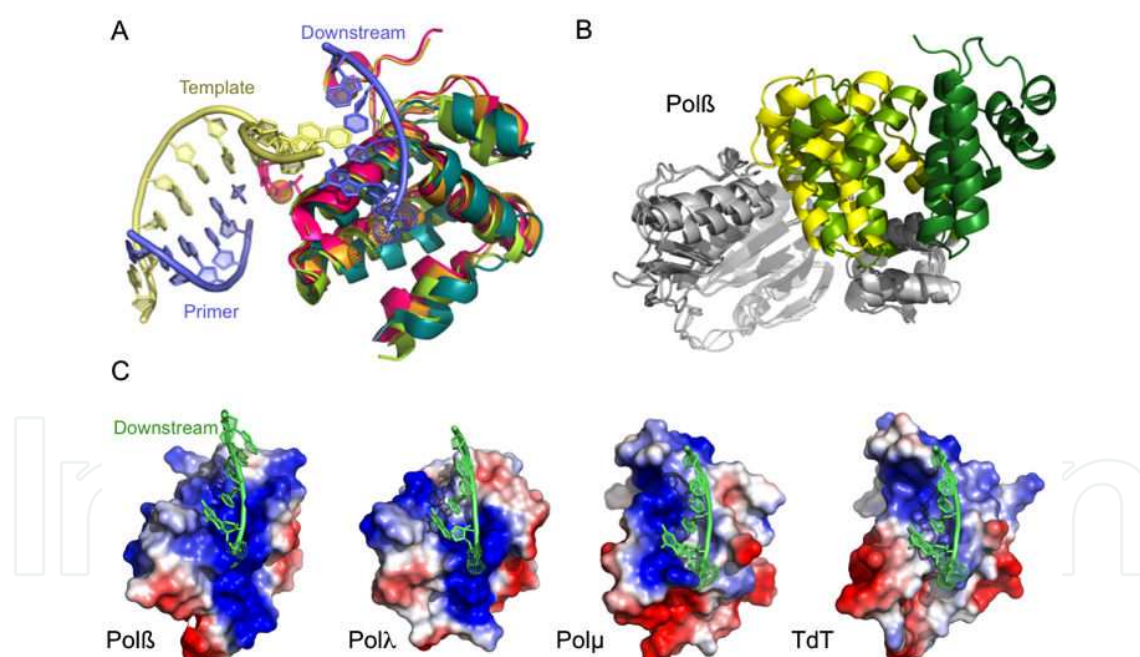
Figure 4. 8 kDa domain of the human X family members. A) Over-imposition of the 8 kDa domains of the X family members: Polß (dark pink), Polμ (orange) and Polλ (teal), shown in cartoon. B) Superimposition of the structures of Polß: apoenzyme (1BPX, dark green), binary (1BPY), light green and ternary (1BPZ, yellow) complexes. C) Electrostatic surface of the 8 kDa domain of Polß (1BPZ), TdT (1KDH), Polμ (2IHM) and Polλ (1XSN), with the downstream strand shown in green. Polμ DNA was over-imposed on the TdT structure.

With the resolution of the first crystal structures of rat and human Polß, the 8 kDa domain was found to be highly mobile (Fig. 4B), not freely, but displaying a small number of stable

positions: 1) in the absence of DNA and incoming nucleotide, the 8 kDa domain is located far away from the thumb subdomain, and the polymerase is in an open conformation; 2) in the presence of a DNA gap, the 8 kDa domain moves and comes closer to the thumb through binding of the 5'-phosphate group of the downstream strand; 3) after arrival of the nucleotide, there is a further movement of the 8 kDa domain, and Polß finally adopts the closed conformation. The model proposed originally [39] explains the formation of the 90° bend in the DNA substrate in two steps: first, binding of the 8 kDa domain to the downstream part of the gap stabilizes the initial positioning of the enzyme; secondly, upon folding the polymerase domain and binding of the primer part of the substrate, the bend of the DNA duplex is created. This bending causes the downstream part to rotate out, exposing the 3' end of the primer.

This two-step model is confirmed by the observations derived form the solved Polλ structures, the most indicative in this matter being the co-crystal with a 2 nt gap ([43], PDB ID: 1RZT). In this case, the 5'-P is located in its correct position and bound by the 8 kDa domain, but the place of the templating base is occupied by the second template nucleotide of the gap, i.e. the one adjacent to the downstream duplex. This causes the 3'-OH of the primer to be displaced to the -1 position relative to the catalytic position, adjacent to the NTP binding site observed in the 1 nt gap co-crystal (PDB ID: 1XSN). Therefore, the location of the polymerase domain in a gap (1-nt or longer) is dictated by the binding of the 8 kDa domain to the 5'-P, and not by interactions with the primer terminus.

This conclusion has implications of great interest for the binding of the polymerase to NHEJ substrates, since 8 kDa-mediated binding would occur irrespective of the conformation of the 3' end. The polymerase in charge for this has to be able to take advantage of micro-homologies for aligning the 3′ ends, and the 8 kDa domain provides an anchoring point for this complicated task.

5.1. Phosphate pocket

As already noted, the main function of the 8 kDa domain is the binding of the 5'-P group of the downstream strand of the DNA substrate. In fact, polymerization rates by template-instructed polymerases of the X family are greatly enhanced when the substrate contains this 5'-P group. In the case of Polß and Polλ, the processivity is also improved on long gaps (5 nt [44, 45]). In the ternary structures of Polß (PDB ID: 1BPY), Polλ (PDB ID: 1XSN) and Polμ (PDB ID: 2IHM) this 5'-P moiety is located at a positively charged pocket where binding is mediated by several hydrogen bonding interactions with basic side chains within the pocket (Fig. 4C). However, in Polμ there are fewer interactions than in Polß or Polλ, and the binding pocket is not as positively charged (Fig. 4C). There is no structure of TdT containing a downstream strand, but this enzyme still conserves the 8 kDa domain, that could be used to coordinate terminal addition of N-nucleotides with the joining of the two DNA ends generated during V(D)J recombination.

5.2. HhH domain

The 8 kDa domain contains another structural motif implicated in DNA binding, the helix-hairpin-helix (HhH) motif. These motifs bind single- or double-stranded DNA in a sequence independent manner, with the aid of a coordinated metal cation [46, 47]. In Polß, Polλ and
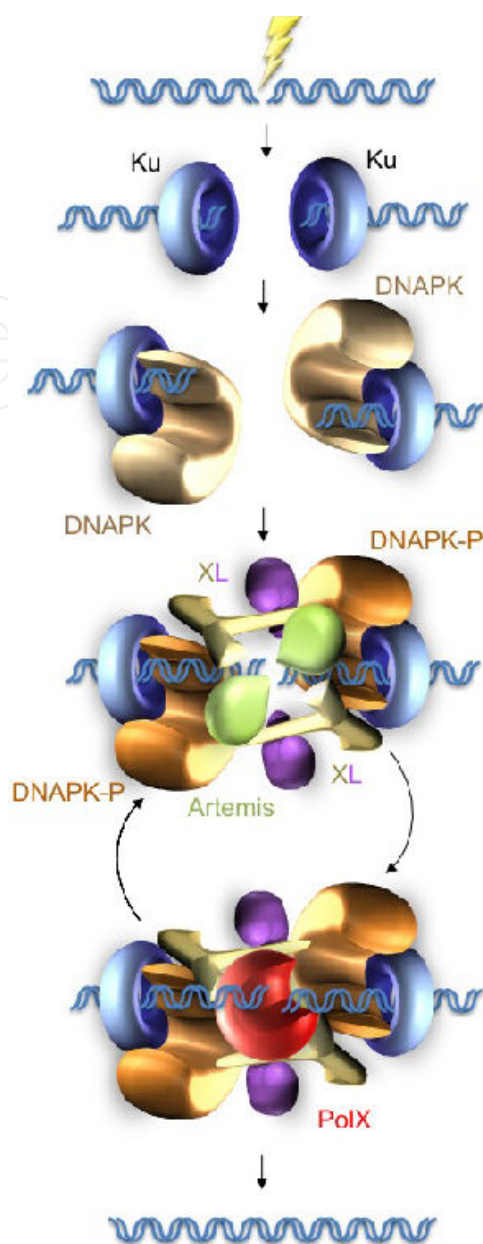
Figure 5. Non-homologous end joining pathway in eukaryotes. This pathway acts repairing damage-generated DSBs. The Ku70/80 heterodimer is the first protein factor to arrive at the site of the break and bind the DNA ends. The DNA PKcs is the recruited and forms a complex with Artemis. The phosphorylated Artemis acts as an endonuclease, generating ssDNA-protruding regions at the ends, and after this the complex dissociates from the DNA. The X family DNA polymerases are then in charge of searching for micro-homologies or generating them, and filling-in the gaps generated. Finally, the XRCC4/Ligase IV complex seals the break.

Polµ structures, this HhH interacts with the downstream part of the substrate, suggesting that its function is the stabilization of the bent DNA thereby facilitating the positioning of the two DNA ends in a NHEJ reaction.

The structures of the 8 kDa HhH motifs from the X family enzymes are not exactly the same: in Polß and Polλ this motif is similar to those found in other repair enzymes, with the GxG sequence of the hairpin and other protein residues being conserved. In Polµ and TdT, on the

other hand, one of the helices is distorted, probably as a consequence of the lack of primary sequence conservation in the hairpin (CLG in TdT, HFG and YLG in mouse and human Polμ, respectively).

5.3. Polλ dRP lyase allows repair of "dirty" DSBs

The 8 kDa domains of Polß and Polλ harbor an intrinsic dRP lyase activity that is required during single-nucleotide BER to remove the residual 5'-deoxyribose-phospate moiety left by the AP-endonuclease after elimination of the nitrogenous base. This reaction proceeds through a ß-elimination mechanism *via* an Schiff base intermediate, and has been shown to be the rate-limiting step in the elimination of several DNA lesions *in vivo* [41, 48]. The studies on the structural aspects of dRP-lyase chemistry [49-51] have led to the conclusion that the amino acids serving as catalytic nucleophiles are Lys^{72} in Polß [42] and Lys^{312} in Polλ [41]. This positively charged residue is not conserved in Polμ (Val^{212}) or TdT (Val^{224}), and thus the dRP-lyase activity is not present in these enzymes.

6. Polμ: A "Jekyll & Hide" DNA polymerase at the edge between genomic stability and variability

Polμ is a DNA polymerase belonging to the X family with a strong similarity to TdT, its closest counterpart in the X family. They share 42% identity at the amino acid sequence, and also have a very similar structural organization: their N-terminal portion contains a nuclear localization sequence, followed by a BRCT domain and then the Polß-core structure already mentioned.

Regarding Polμ biochemical properties, it displays a certain terminal transferase activity [7], although it is primarily a DNA-dependent DNA polymerase [7, 52] and its activity increases strongly in the presence of a template strand of DNA. It is also known that both types of polymerization are stimulated *in vitro* in the presence of Mn^{2+} ions, the preferred metal activator, and in the presence of this cofactor Polμ exhibits a strong mutator phenotype, with a very high probability of erroneous nucleotide incorporation, being one of the most error-prone polymerases known in higher eukaryotes [7]. This strong mutator ability is based on a dislocation mechanism [53, 54] through which Polμ is capable of repositioning the template strand so that incorporation is dictated by templating bases away from the end of the primer. The mutator capacity of Polμ is further enhanced by its low sugar discrimination, being able to incorporate not only dNTPs but also NTPs [55, 56]. This may have implications in cell cycle phases in which the levels of dNTPs are very low as NTPs reserves remain high throughout the cycle.

Although the *in vivo* role of Polμ has not been clarified yet, a number of functions for the polymerase have been proposed, including its participation in the non-homologous end-joining (NHEJ) pathway, in charge of repairing the highly harmful double strand breaks in DNA. The NHEJ system relies on little or no homology between sequences to achieve repair, since the proteins involved in the process recognize the ends of DNA based on their structure

rather than its sequence (reviewed in [57]). This pathway may lead to mutagenesis, contribu-ting to the variability of the genomes [58, 59], and is key to certain cellular processes such as antibody repertoire generation. NHEJ is the main mechanism to repair DSBs in higher eukaryotes, as it is operative throughout the cell cycle, unlike homologous recombination, a second DSB repair mechanism which is inhibited during the G0, G1 and S phases [57].

The first step of NHEJ is the binding of specific protein factors to the ends of the DNA break (Fig. 5). The Ku70/Ku80 heterodimer recognizes the ends of the break, and due to its toroidal shape accommodates the duplex DNA, preventing possible nucleolytic degradation [60]. Then, the DNA-PK kinase is recruited [61, 62], inducing a slight internalization of the Ku heterodimer [63], and allowing both sides of the break to approach through specific protein-protein interactions [64-66]. Once the ends are juxtaposed, generally cannot be directly linked, but require pre-processing. Analysis of the sequences repaired by NHEJ at the break points suggests that some of these events involve the alignment of the ends through micro-homolo-gies (complementary sequences from 1 to 4 nt) near the site of rupture [67-69]. When there is no direct microhomology the system must generate it by certain mechanisms that involve nucleases and/or DNA polymerases [70, 71], which would be needed to process distortions, flaps or gaps that may arise as a result of the alignment of the chains (reviewed in [72]). The Ku-DNAPK complex recruits the proteins needed for processing and subsequent ligation of the ends. Artemis, an ssDNA 3'-5' exonuclease, is activated through phosphorylation by DNA-PK [71]. Polynucleotide kinase (PNK), which has kinase and phosphatase activities [73], may also intervene in end-processing [74]. If the ends at this point were compatible, the last step of the mechanism would be the recruitment of the XRCC4/LigaseIV complex by Ku, which would carry out the ligation of the ends [75-78]. If, on the contrary, the ends were not compatible, a DNA polymerase would be needed, since its activity would be critical for filling the gaps generated during the alignment of the chains of DNA [70, 79]. Polµ could even perform template-independent polymerization to create the necessary complementary sequences [80, 81]. Finally, after processing the ends, the complex formed by DNA Ligase IV/XRCC4 would be responsible for sealing the joint between the ends of the break [64, 75]. Another factor similar to the protein XRCC4 has been recently identified in mammals. It has been called XLF (XRCC4-like factor) or Cernunnos, and interacts with the DNA LigaseIV/XRCC4 complex to promote end ligation [82, 83].

On the other hand, Polµ preferential expression in lymphoid tissues, especially in the germinal centers of secondary lymphoid organs, suggests a specific role of this polymerase in processes occurring in these regions. Its resemblance to TdT at the structural level, and its ability to conduct untemplated nucleotide additions, together with the fact that TdT is not expressed in secondary lymphoid organs, allowed to propose a function for Polµ in somatic hypermutation in the germinal centers [52], which occurs in these regions as an additional mechanism for diversification of the immune response [84]. Moreover, Polµ is present also in the thymus and bone marrow, and thus may be required during the normal process of V(D)J recombination as DNA-dependent polymerase to generate palindromic sequences (P sequences) at the ends of the coding fragments, or during gap-filling reactions required for coupling N additions to the DNA ends [52]. It was recently demonstrated an *in vivo* role of Polµ in the V(D)J recombination process of the light chain (kappa) of immunoglobulins, based on the observed deletions at the

junctions between these gene segments in the case of Polμ deficiency [85]. Also, recent data implicated Polμ in the DJ_H recombination in mice embryos, a stage in which TdT is still not expressed [86]. In this case, all the N-additions observed in wild type mice were completely attributable to Polμ, as shown by comparison with Polμ-KO mice. This evidence suggests a role for Polμ in the V(D)J mechanism.
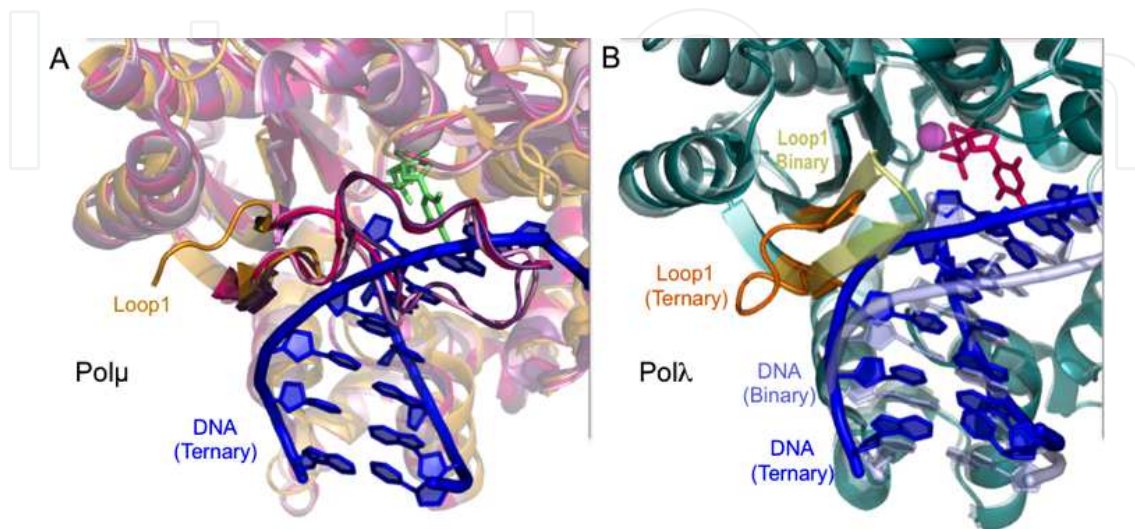
**Figure 6. Loop 1 in Polμ, TdT and Polλ: movements from the binary to the ternary complexes.** A) Over-imposition of the three available crystal structures of TdT (1JMS, 1KDH, 1KEJ, in light pink, dark pink and purple, respectively) and the ternary complex of Polμ (orange). B) Over-imposition of the binary (light teal) and ternary (dark teal) complexes of Polλ, shown in cartoon. Loop 1 from both structures are shown in yellow and orange, and the DNA substrates in light and dark blue, respectively.

7. A mobile loop in Polμ provides the ability to join non–homologous DNA ends

Template instruction is a general feature of most members of the X family, with the exception of TdT. TdT is the only known fully template-independent DNA polymerase, as it is able to add nucleotides to a primer DNA molecule in the absence of a template strand. This feature is crucial for its function in V(D)J recombination, where TdT adds nucleotides to the recombinational junctions of immunoglobulins and TCR receptor genes, generating variability as it creates new information [87, 88]. Interestingly, Polμ shows hybrid biochemical properties: it has an intrinsic terminal transferase activity, but it is strongly activated by a template DNA chain [7].

Understanding the structural and functional basis of the template-independence of TdT had to await the resolution of the crystal structure of the Polß-like core of TdT [89]. A loop region between ß-strands 3 and 4, referred to as Loop 1, has a similar position in all three TdT structures, and is located in a region of the DNA binding cleft that would normally be occupied by the template strand (Fig. 6A). Therefore, this loop would preclude binding of any DNA substrate possess-

ing a template strand, thus explaining its null activity on these substrates. On that basis, and by extrapolation of the TdT structural model to Polμ, it was predicted that Loop 1, specifically present in these two enzymes, could be directly responsible for their template-independent terminal transferase activity, but in Polμ Loop 1 must be flexible enough to also allow template-directed polymerization [80]. In agreement with this prediction, when the crystal structure of Polμ bound to a gapped DNA was solved [40], Loop 1 was disordered suggesting conformational flexibility (Fig. 6A). In this structure, the DNA duplex was bound in the usual fashion within the DNA binding cleft. It was then clear that Loop 1 of Polμ cannot occupy the same position as that of TdT when a template strand is present. A comparison of the ends of the ß-strands flanking the loop shows that TdT's Loop 1 extrudes upwards, toward the DNA binding cleft, while that of Polμ appears to turn downwards, away from the cleft [40]. Although no crystal structure is available of Polμ with a single stranded or 3'-protruding DNA substrate, it is likely that Loop 1 would then be found in the same conformation as in TdT, i.e. interacting with the primer strand, somehow mimicking a template strand. The structural evidence suggested that Loop 1 in Polμ may adopt different conformations depending on the nature of the substrate: the inherent flexibility of this loop in Polμ is distinct from TdT and suggests how Polμ can accommodate different substrates. Studies including the Loop 1 chimeras on Polμ [80] and TdT [90] confirmed this hypothesis: replacement of the TdT Loop 1 with that of Polμ is sufficient to allow template-dependent additions, while the reciprocal chimera (Polμ with the TdT Loop 1) is much less inclined to perform template-dependent additions.
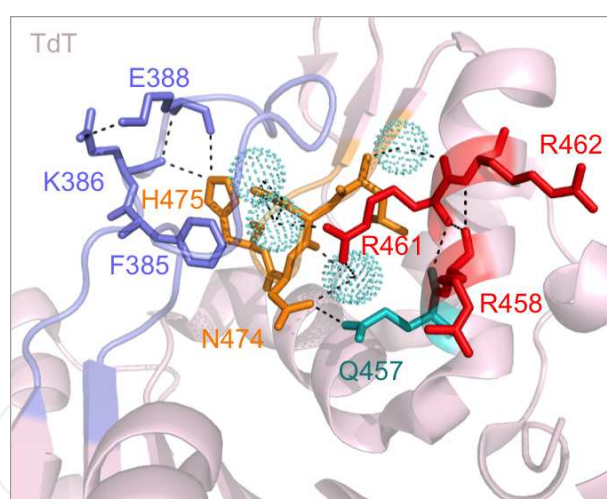
Figure 7. The Loop 1 network. Cartoon representation of the TdT apoenzyme (1JMS). Loop1 is shown in blue cartoon with selected residues involved in interactions shown in sticks, the thumb mini-loop is shown in orange with selected residues shown in sticks, arginines from the helix N are shown in red sticks, water molecules and other are shown in light teal.

The equivalent regions in Polß and Polλ would be less likely to interfere with binding of the template strand because they have a much shorter Loop 1: small enough in Polß to be described as a turn and of intermediate length in Polλ (Fig. 6B). Consistent with this idea, when Loop 1 in Polμ is shortened to a length similar to that of Polλ, the altered polymerase has higher catalytic efficiency on template-containing substrates, but is incapable of template-independ-

ent synthesis [29, 80]. Consistent with all this, Polλ has a strongly reduced ability to catalyze template-independent synthesis, but retains the ability to perform template-instructed additions. Polλ Loop 1 may be involved in a function somehow related to that in Polμ: modulation of fidelity by controlling dNTP-induced movements of the template strand and 3'-primer terminus in the transition from an inactive to an active conformation of the enzyme [91]. In fact, dNTP binding induces Polλ to transition from an inactive to an active conformation: ß-strands 3 and 4 partially unravel to form Loop 1, a nine-residue loop that repositions as the DNA template strand assumes its active conformation (Fig. 6B). Such a "fidelity checkpoint" would then be related to the energetic penalty of changing the structure of these ß-strands, that would only be overcome in the case of the formation of a correct match.

The role of Loop 1 during terminal transferase additions has been now established, but a more in depth study of how Polμ fixes and/or orients this mobile part of the protein in accordance with the substrate on which it is polymerizing is necessary. In the case of TdT, residue Phe^{401} (corresponding to Phe^{385} in Polμ), is involved in maintaining the fixed position of Loop 1 via a strong stacking interaction between its aromatic ring and His^{475} (His^{459} in Polμ), located in a mini-loop at the thumb subdomain (Fig. 7). Mutant F401A in TdT had a striking phenotype, turning a completely template independent enzyme into a DNA-instructed DNA polymerase [90]. This mutation clearly disrupted the network of interactions needed to maintain a fixed orientation of TdT Loop 1, that is now endowed with a greater degree of flexibility, as in Polμ, thus allowing TdT to accept a template strand. Phe^{389} is again conserved among Polμs and TdTs (Phe^{405}) of different species, and in both cases it seems to be involved in maintaining the shape and orientation of this motif. Mutation of this residue to alanine in TdT abolishes terminal transferase activity and allows templated insertion of only one nucleotide on a template/primer substrate [90]. We produced mutants in the implicated residues of Polμ and all of them lacked terminal transferase activity, indicating that the network of interactions maintaining the conformation of Loop 1 in TdT is conserved in Polμ [92]. Also, in TdT Loop 1 is interacting with another very small loop located in the thumb through His^{475} (Fig. 7), that is conserved in Polμ (His^{459}). This mini-loop is also present in the other members of the X family, but its function is different: residues from this loop directly interact with the template strand. In Polμ this mini-loop has both roles: depending on the substrate used and the desired conformation of Loop 1, the mini-loop may interact either with the template strand (through Asn^{457}) or with Loop 1 itself (through His^{459}). Accordingly, the asparagine is only needed during templated additions, and dispensable for terminal transferase activity of Polμ, while the histidine had the opposite effect [92]. We propose a regulatory function for the NSH motif in the thumb mini-loop, helping to accommodate either the template strand (as in Polß of Polλ) or Loop 1 (as in TdT) as suits best for each individual situation.

## 8. A single arginine in Polμ limits terminal transferase to favor fidelity during NHEJ

Having now a general idea of how these two polymerases, Polμ and TdT, are specially designed to perform this untemplated additions of nucleotide units, another question still

remains: why and how the terminal transferase activity of TdT is much higher than that of Polμ? Combined structural and functional evidences for both Polμ and TdT indicate that there is one residue modulating the terminal transferase activity of both enzymes. That residue (Arg[387] in Polμ and Lys[403] in TdT) tunes the catalytic efficiency of the terminal transferase reaction, by regulating the rate-limiting step. Judging by the structural data available, this residue could be establishing dual and alternative interactions during the catalytic cycle of both Polμ and TdT: when the primer is bound at the unproductive position (TdT crystal 1KDH), the residue is interacting with the primer strand, while in the Polμ crystal in which the primer strand is correctly positioned in a productive complex (2IHM), the arginine is interacting with the -3 position of the template strand (Fig. 8B). In the case of Polμ, and assuming an alternative interaction as that seen in TdT, Arg[387] acts as a brake for the necessary movement of the primer, to limit nucleotide additions before end bridging. In fact, the single change of this residue for the TdT counterpart (Polμ mutant R387K) showed an increase in untemplated additions that ranged from 10- to 100-fold, reaching levels comparable to those of TdT itself [93]. Interestingly, mutant R387K produced a very specific blockage at position +4 when continuous terminal transferase extension of a blunt end was tested [93]. This situation is such that, in a 3-protrusion of 4 nt, the second proposed protein-DNA interaction for this residue cannot occur, since the -3 position of the template strand is not available. In these substrates (ssDNA, 3' protrusions longer than 3 nt), this residue must be adopting a new partner for this second interaction, most surely a portion of the protein that is now located in place of the template strand: Loop 1. TdT Loop 1 contains a histidine (His[400]) that completely superimposes with the -3 position of the template strand, and this histidine is surely acting as a partner for Lys[403] when it is not interacting with the primer (catalytically active configuration; Fig. 8A. left panel). In agreement with this, our results measuring TdT activity on substrates ranging from blunt to 11 nt 3'-protruding indicate that polymerization was inhibited when the protrusion was shorter than 3 nt (these substrates would not allow correct positioning of Loop 1 and His[400]). A similar protein-protein interaction between Arg[387] and Loop 1 is surely occurring in Polμ when the -3 position of the template is not available (Fig. 8B, right panel), and it is distorted when the arginine is mutated to alanine, as indicated by the completely defective terminal transferase activity of mutant R387A [92].

Interestingly, the equivalent residue in human Polλ (Lys[472]) is also involved in regulating the catalytic cycle by means of inhibitory interactions with the primer strand [91]. Recent results suggest that Lys[472] may help to modulate template-dependent synthesis. In the wild type Polλ binary complex (1XSL), Lys[472] is within H-bonding distance of the 3'-O of the primer terminal nucleotide. Such hydrogen bond between Lys[472] and the primer terminus that could stabilize the inactive conformation must be disrupted in order for the 3'-O to assume its catalytically competent position. A weakened interaction between Lys[472] and the primer terminus would allow the 3'-O to more easily adopt a conformation that would support catalysis with an incorrect nucleotide bound, reducing the discrimination between correct and incorrect incorporation [94].

Thus, Arg[387] plays a key role in modulating template-independent synthesis by Polμ, having a dual role: it allows terminal transferase additions to occur, but also acts as a brake that limits these additions. Substituting the homologous lysine in TdT with arginine or alanine [90] also
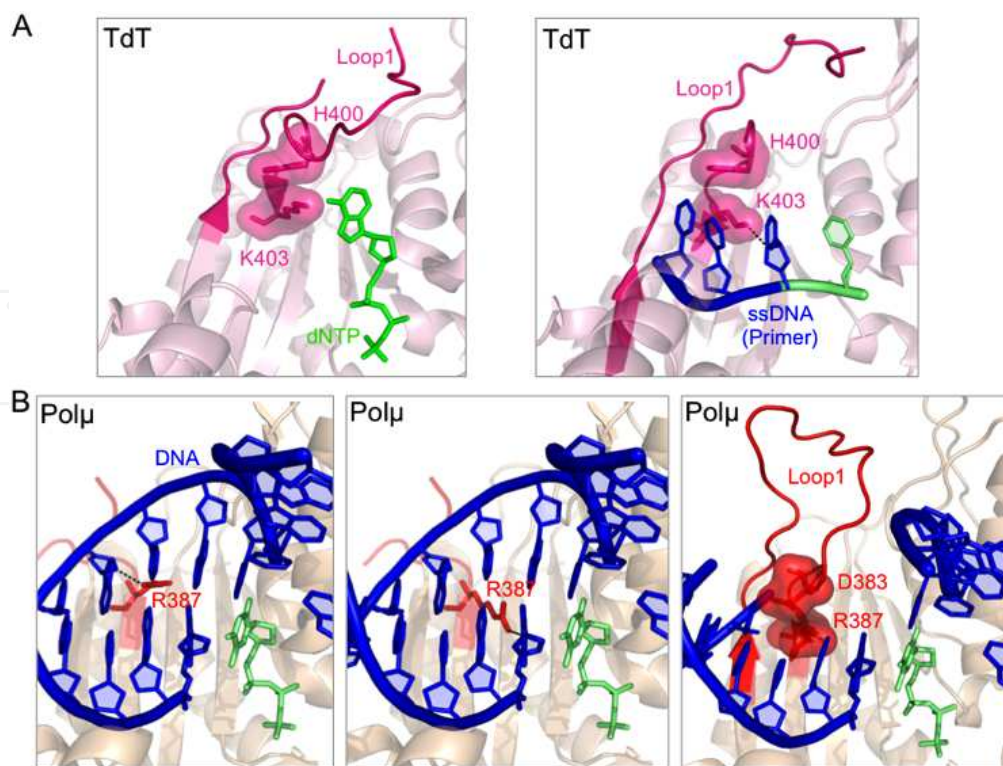
Figure 8. Arg³⁸⁷ triple interactions with the primer and template strands and with Loop 1. A) Cartoon representations of the binary complexes of TdT bound to dNTP (1KEJ) or ssDNA (1KDH). Loop 1 shown in dark pink and Lys[403] and His[400] show in sticks with semi-transparent surface. B) Cartoon representations of the ternary complex of Polμ: left panel, the original position of Arg[387] contacting the template strand; middle panel, the predicted interaction with the primer; right panel, over-imposition of Loop 1 from the TdT structure (1JMS) and proposed interaction of Arg[387] and Asp[383].

results in loss of template-independent activity, although the properties of the two TdT mutants are not identical. In the case of TdT, residue Lys[403] likely establishes a weaker interaction with the primer compared to its orthologue Arg[387] in Polμ. Thus, TdT has been optimized to efficiently overcome the rate-limiting step of the terminal transferase, to exclusively perform creative synthesis.

What is the reason for this limited terminal transferase activity in Polμ? Our results indicate that when a templating base is provided *in trans* during NHEJ, the rate-limiting step is relieved. A templating base provided *in trans* by the approaching end that could be located in a proper register will stabilize the incoming (and complementary) nucleotide, thus facilitating primer translocation. As a result of this, NHEJ of many incompatible ends can be efficient and accurate. During NHEJ of this fraction of incompatible ends, an excessive terminal transferase as that displayed by mutant R387K would be disadvantageous in terms of genomic stability. On the other hand, our findings also explain the need for a mild terminal transferase activity in Polμ, not only to create connectivity in those other DNA ends that cannot be efficiently joined on a templating basis, but perhaps contributing to gain a certain degree of genome variability. Additionally, it can be inferred that TdT evolved to maximize the efficiency of the translocation mechanism in the absence of template, at the cost/benefit of introducing untemplated nucleotides, thus being devoted to generate variability at V(D)J recombination intermediates.

Is this the physiological role of the terminal transferase activity of Polµ? NHEJ of short incompatible ends can be accurate in many cases, but imprecise in others depending on both the length and sequence of each protrusion. For the latter cases, when a templating base is not in a proper register, untemplated terminal transferase addition in a NHEJ context provides a valid, although mutagenic, solution that would be conceptually similar to translesion DNA synthesis. Besides, it cannot be ruled out that Polµ's terminal transferase can extend a single short 3'-protrusion to facilitate end joining of this fraction of non-complementary ends. There is also *in vivo* evidence of untemplated insertions made by Polµ. It has been shown that mice that are TdT-/- still contain 5% of V(D)J junctions with template-independent additions, which suggested a possible role of Polµ in these reactions [95]. In agreement with that, the terminal transferase activity of Polµ has been directly implicated in variability/repair processes occurring at embryo developmental stages in which TdT is still not expressed [86].

9. From Polµ to TdT: A new variability–generation mechanism for our immune system

Polµ and TdT are the most closely related of the four members of the human X family, with a 42% identity at the level of the aminoacid sequence. Although the branch of the phylogenetic tree of the X family that contains these two enzymes appeared much sooner than that of Polß, the strict template-independent activity of TdT appears to be a recent evolutionary event that coincides with the development of V(D)J recombination in mammals (Fig. 9). TdT shares the common Polß-like core with 8 kDa, fingers, palm and thumb and also possess the C-terminal BRCT domain that allows recruitment by the Ku proteins to the site of the break. But there are some differences: even though TdT still conserves a positively charged pocket to bind a downstream 5'-P, it contains the lowest amount of positive charges of all the members of the family, and, equal to what happens in Polµ, it has lost the residues essential for the dRP-lyase activity. This first modification, together with the tightly regulated expression of TdT confined to primary lymphoid tissues including thymus and bone marrow [96-98], already indicates that TdT, even though devoted to work at DSBs, is not able to deal with damaged nucleotides and the break points must be "clean", as they are in the case of programmed breaks such as those occurring during the development of the immune response. TdT has been in fact engineered through evolution to "misbehave" and break almost every rule that can apply to a conventional DNA polymerase: it incorporates nucleotides in a template independent manner, using only single stranded DNA [99, 100] or dsDNA with a 3'-overhang longer than four nucleotides [86]. This strict preference for the DNA substrate is dictated by its long Loop 1, of about the same length as the one present in Polµ, but immobilized by several interactions not present in Polµ, such as the ones established between Loop 1 and the small thumb loop [92]. The position of Loop 1 in the crystal structure completely over-imposes with the template strand from the Polµ ternary complex, thus explaining why the length of the single stranded primer needs to be of at least 4 nucleotides for an efficient reaction to take place. This protein piece helps locate the nucleotide in place, and probably is to be blamed for the different order of substrate binding displayed by TdT in contrast with other polymerases: efficient polymer-

ization for a template-dependent polymerase would be optimal through the strictly ordered binding of DNA substrate prior to dNTP, as the converse order of dNTP binding prior to DNA would be error-prone, being correct only once out of four times. Indeed, numerous steady-state and pre-steady state studies have validated that all template-dependent polymerases obey this mechanism [101]. The order by which TdT binds DNA and dNTP is indeed random as determined through a series of initial velocity studies [102]: TdT forms the catalytic competent ternary complex *via* binding of dNTP prior to DNA or vice versa. This scenario is similar to that observed for the *Mycobacterium* NHEJ polymerase, in which a pre-ternary complex can be formed with the nucleotide being present in the absence of a primer strand [103]. This situation could apply also to Polμ, as it would be beneficial for the efficiency of DSB repair, and could have been maintained in TdT since the ability to randomly bind substrates might play a physiological role in generating random nucleotide additions during recombination. Another feature that is present in Polμ and has been maintained in TdT during evolution is the ability to incorporate ribonucleotides. This loss of the "steric gate" probably appeared in Polμ as a collateral effect of the need for a spacious active site able to accommodate misalignments during the search for microhomology, and has been positively selected due to the optimal characteristics of the ribonucleotides as the most abundant substrates, but also due to the "length control" mechanism that the incorporation of ribonucleotides implies during un-templated addition of nucleotides: for both Polμ and TdT, further elongation of a ribonucleotide-containing primer occurs at a slower rate and the addition of more than two ribonucleotides does inhibit activity [55, 56, 104].
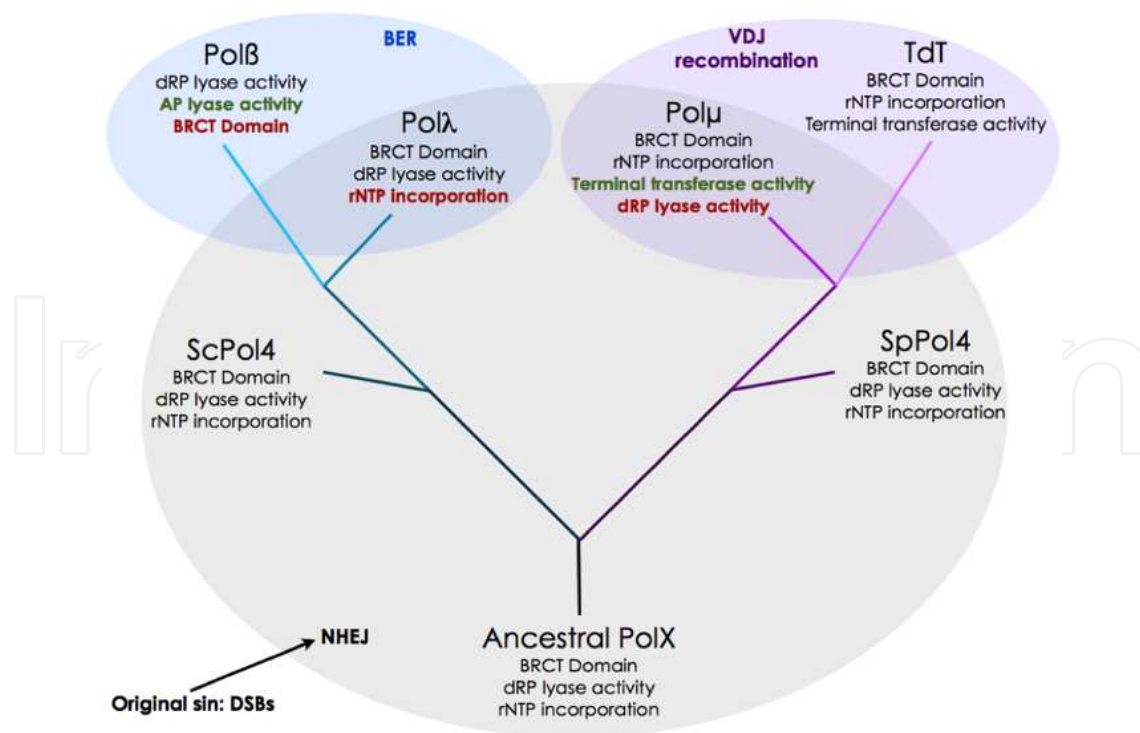
Figure 9. Evolution of family X polymerases. Red color indicates the loss of an activity or feature, green color indicates the gain of an activity or feature. See text for details.

Despite all the similarities between Polμ and TdT, such as the loss of the dRP-lyase activity, the ability to incorporate ribonucleotides and the presence of Loop 1, Polμ has remained preferentially a template-directed polymerase. In the first place, being a more ancient product of evolution than TdT means that its function had to be a more general one: Polμ is devoted mainly to its DNA repair function in the NHEJ pathway. The differential expression patterns of TdT and Polμ also speak in favor of this hypothesis: even though Polμ is strongly expressed in lymphoid tissues in humans, in contrast to TdT, a basal expression of Polμ is observed in a wide range of tissues, more specifically in the brain [7], that suffers from a high level of oxidative damage. Also, the structural features of Polμ support its role as a template-directed NHEJ polymerase: a flexible Loop 1, held but not constrained by several other modules in the protein (the thumb loop, the arginine helix), helps to stabilize gaps in the template strand without blocking the use of the templating base. Also, a specific arginine residue (Arg[387]), present only in Polμ, acts as a "brake" during the terminal transferase catalytic cycle [93], limiting the number of untemplated additions and keeping the polymerase in a "stand-by" mode for a longer time, awaiting the arrival of the templating base.

Taking advantage of the Dr. Jekyll & Mr. Hyde duality of Polμ as a template-directed and also template-free polymerase, its appearance in the phylogenetic tree of the X family probably was the starter's pistol shot to the process of generating variability during development of the adaptive immune system response, without losing a DNA repair function. In fact, it has been demonstrated that Polμ still participates in the DJ_H rearrangements in mice embryos, where TdT is still not expressed [86]. Based on its DNA-dependent polymerization ability, which TdT lacks, Polμ also fills-in small sequence gaps at the coding ends and contributes to the ligation of highly processed ends, frequently found in the embryo, by pairing two internal microho-mology sites. Also, Polμ is involved in V(D)J recombination at immunoglobulin k light-chain loci, after synthesis of the N-regions [85]. The lack of Polμ leads to alterations that induce a profound defect in the peripheral B cell compartment which results in an average 40% reduction in the splenic B cell fraction in Polμ knock-out mice. Polμ appears, therefore, as a key element contributing to the relative homogeneity in size of light chain CDR3 and taking part in Ig gene rearrangement at a stage where TdT is not expressed [85]. Polμ has also been shown to be up regulated in germinal centers after immunization, and although it is not a critical partner, Polμ modulates the *in vivo* somatic hypermutation (SHM) process [105]. The role of Polμ in this process was proposed some time ago [52], and further supported by studies of Polμ overexpression in a Burkitt's lymphoma cell line (with constitutive SHM), in which the SHM rate was increased [53].

10. From Polλ to Polß: Losing the BRCT and evolving base excision repair

The similarity between yeast Pol4 and Polλ, which share the same additional domains (Fig. 2), together with the extraordinary evolutionary conservation of the versions of Polλ present in various higher eukaryotes and in plants (*Arabidopsis thaliana*, *Wisteria max*, *Oryza sativa*) suggests that this is the X family member closest to the common ancestor from which all

members of the family derived. This could account for the multiple functions of Polλ, since the common ancestor necessarily carried out various processes of DNA synthesis. In this sense, the presence of the Ser/Pro domain is of special relevance, as it could regulate the participation of Polλ in different processes, such as repair by BER, NHEJ and V(D)J recombination.

Members of the human X family of DNA polymerases have specialized in different processes of DNA synthesis associated with repair. Such processes are basically three: 1) base excision repair (BER), carried out mainly by Polß, although Polλ seems to have a role in specific situations; 2) non-homologous end joining (NHEJ), in which, according to the type of substrate generated, Polλ or Polμ could be involved; 3) V(D)J recombination, involving Polλ, Polμ and TdT, with different roles. Subtle differences in the biochemical properties of X family members seem to be crucial for performing one role and not other. Therefore, the members of this family have diversified to be able to carry out non-redundant tasks, achieving a high degree of specialization that has resulted in a high degree of efficiency of each polymerase on its specific function.

Polλ, as the member of the family more closely related to the common ancestor, bears many of the specific modifications needed to perform a high number of functions: it has a BRCT domain needed for interactions with the NHEJ components, and it harbors an 8 kDa domain that acts both as the main DNA binding domain through the 5'-P pocket and as the container of the dRP-lyase activity needed for an efficient performance during BER. Moreover, it contains a long *nail* motif that helps the polymerase to deal with misaligned substrates and might allow scrunching to occur. It has a brooch (WxCxQ motif) that maintains the Polß-like core in a closed conformation throughout the catalytic cycle possibly helping to correctly orient discontinuous NHEJ substrates [92], and finally it has a mid-length Loop 1 that may have a similar role to that proposed for Polμ Loop 1 during NHEJ, but with the limitation of needing some degree of complementarity between the two DNA ends, probably due to the position occupied by this loop in Polλ at the -2 to -4 positions of the template strand.

As a younger member of the family, Polß is the polymerase that has lost the majority of these features, to be focused on enhancing the efficiency of just one reaction: the filling-in of short gaps during BER. For that, it has strengthened the interactions with the DNA substrate through the 5'-P binding pocket, being the most positively charged in this region of the four human enzymes, and it has maintained the dRP-lyase activity and gained an AP-lyase activity, precious for its dedicated job as a BER polymerase. It also maintains a long *nail* that helps locating the DNA substrate on its final catalytic position, and probably helps to "count" the templating nucleotides when filling-in a long gap. It also has the capacity of changing from an "open" to a "closed conformation" since it has lost the brooch at the N-terminal portion of the core, and thus the space between the 8 kDa domain and the thumb subdomain can be expanded to accommodate the yet-to-be-copied templating nucleotides more easily. On the other hand, the loss of this "closing" motif probably meant that its role as a NHEJ polymerase was greatly impaired, together with the complete loss of the Loop 1, which is now merely a turn connecting two ß-strands. The disappearance of this flexible structure probably also led to an improvement of the polymerization on template-containing substrates such as the

ones produced during BER. Congruently, Polß lost the BRCT domain so it does not get recruited to DNA DSBs where it cannot act, and has in turn gained a new set of protein-protein interactions with other BER factors as XRCC1 through specific residues on the surface of its catalytic domain that are required for an efficient repair [106-108]. The Ser/Pro domain located between the BRCT and the catalytic domains in Polλ is also missing in Polß, and this, together with the total absence of CDK phosphorylation sites, unique in the human X family, indicate the lack of a cell-cycle dependent regulation that correlates with its function as a housekeeping gene. Whereas short-patch BER in mammalian cells plays an important role in the maintenance of genomic stability [109-111], it is unlikely that a similar repair path-way is present in many phylogenetically divergent organisms. Plants do not contain a homolog of DNA ligase III, which is required for mammalian short-patch BER, or a Polß homolog [112]. Additionally, the plant XRCC1 protein lacks the Polß binding domain (N-terminal domain; [113]). In contrast, all enzymes needed for long-patch BER are encoded in the genomes of *A. thaliana* and *O. sativa*, suggesting that plants utilize the long-patch BER pathway [112]. Similarly, no protostomic organism possesses the short-patch BER system [9, 114], and a short-patch BER-like pathway is present in yeast but it differs from the mamma-lian pathway [115]. From the data described above, we hypothesize that short-patch BER is an advanced repair pathway present only in mammals (Fig. 9). Polß, the primary DNA polymerase of this pathway, is highly expressed in brain tissue [116], and would be required mainly to minimize the accumulation of DNA damage in neuronal cells [117] that suffer from a high level of oxidative lesions [118, 119].

11. *In vivo* deficiency models for the X family polymerases: Non–redundant roles in DNA repair and immune system development

The biochemical characteristics of the four members of the X family of polymerases provide strong hints as to what physiological roles they might be performing. To obtain direct evidence of their *in vivo* functions, mouse models were developed for each of the four polymerases individually and in several combinations. In this section we will briefly recapitulate the phenotypes observed with these animals and the conclusions derived from these works.

Initially, two deficiency models were generated for Pol β. The first one eliminated the enzyme from T cells but no differences could be observed between Pol β -deficient and wild-type animals [120]. In the second case, a complete knock-out was generated but the homozygous embryos were unviable due to apoptosis of post-mitotic neurons, as a consequence of defective DNA SSB repair [117]. *In vitro* assays performed with Pol β -deficient cell extracts indicated that this polymerase bears the essential dRP-lyase activity involved in repair of oxidative base lesions [121]. The main mediator of the neuronal apoptosis observed in the Pol β$^{-/-}$ background is p53, as indicated by the combined deletion of both genes in the mouse [122]. However, these animals were still unviable, and the data suggested another role of Pol β in the development of certain neuronal cell types. Heterozygous mice displayed a higher risk of cancer develop-ment than wild-type mice, although no effect on the lifespan was detected [123]. These animals

had normal levels of apoptosis and normal levels of BER enzymes and BER activity, except in spermatogenic cells. These results are in agreement with data showing elevated levels of mutagenesis in this compartment [124] and meiosis failure at prophase I due to defective resolution of DSBs and synapsis at this stage [125]. The sperm cells produced by these animals contained an increased level of transversion mutations. In contrast, Pol β -/- mice displayed lower levels of mutagenesis in the embryonic brain than wild-type animals [126], but this can be explained as a result of the apoptotic elimination of neurons with high levels of unrepaired DNA. Very recently, a knock-in mouse model for a natural allele of the human Pol β was reported [127]. This Y265C variant is a mutator polymerase with slower catalysis [128, 129]. The homozygous mutant mice show slower cellular proliferation and increased apoptosis, as well as deficient gap-filling during BER, with DSBs and chromosomal aberrations as a consequence. All these studies show the clear importance of Pol β in meiosis, neuronal development, DNA repair and genomic stability.

In the case of Pol λ, again two mouse models were reported at the same time. One of them showed a very dramatic phenotype of male infertility due to cilia immobility [130], which was later attributed to disruption of a neighboring gene rather than to deletion of Pol λ itself [131]. The second deficiency model was tested initially for somatic hypermutation and this process was not affected [132], but it was later shown that Pol λ -/- mice lack diversity in their antibody pools, specifically regarding the N-additions at the junctions in the heavy chain of the TCR receptors [95]. The data indicate that Pol λ might act before TdT during heavy chain rearrangement, suggesting a non-redundant role for Pol λ during V(D)J recombination. Using fibroblasts from the Pol λ -/- mice it was shown that this polymerase has a role in the BER pathway to protect cells from oxidative damage [133], and that it can act as a back-up in the absence of Pol β [134]. Moreover, Pol λ is responsible for the majority of the error-free gap-filling in the presence of the 8oxoG lesion in DNA [135].

In 1993 two independent groups published two deficiency mouse models for TdT, reaching very similar conclusions: the TCR receptors of B- and T-lymphocytes had fewer or none N-additions and thus the antibody repertoire was less diverse, maintaining the fetal phenotype in the adult animal [136, 137]. Furthermore, in the absence of TdT, homology-directed repair was detected during V(D)J recombination. Later it was shown that TdT is responsible for 90% of the diversity of the $\alpha\beta$ TCR receptor repertoire [138].

Mice deficient for Polμ have been also studied, and they are viable and fertile [132]. These mice are defective in immunoglobulin light chain rearrangements and thus development of the bone marrow and B cell differentiation are compromised [85]. A different mouse model was reported with a normal immune response but impaired centroblast development, due to defects in somatic hypermutation and V(D)J recombination [105]. These mice are hypersensitive to γ-irradiation due to a defective DSB repair also in non-hematopoietic tissues [139]. Studies of the embryonic stage, when TdT is still not expressed, indicated that Polμ is responsible for the observed N-additions at the post-gastrulation DJ$_H$ joints during immunoglobulin gene rearrangements [86]. These results support the roles of Polμ during hematopoietic development and the processes of somatic hypermutation and class-switch recombination, during the generation of extra diversity in the immune system and, finally, its contribution to genomic stability through repair of DSBs *via* the NHEJ pathway.

12. A case of convergent evolution: Comparison of the characteristics shared by bacterial and eukaryotic NHEJ polymerases

Conventional replicative and lesion bypass DNA polymerases extend off dsDNA substrates, containing both primer and template strands, in a 5' to 3' direction. In contrast, polymerases involved in DSB repair must be capable of binding and extending off non-canonical DNA substrates, including 3' over-hanging termini lacking continuous primer and template strands. Recent studies on the bacterial NHEJ polymerases have revealed some of the unusual activities associated with these repair enzymes that enable DNA extension under the most extreme conditions. For example, a homodimeric arrangement of the mycobacterial NHEJ polymerases can facilitate the association of two incompatible 3'-protruding DNA ends, *via* microhomology-mediated synapsis, forming a stable end-joining intermediate [140]. This synaptic complex reflects an intermediate bridging stage of the NHEJ process, prior to end processing and ligation. In this way, the polymerase restores the continuity of the dsDNA helix, catalyzing a conventional 5´-3´ extension reaction occurring on one DNA end, but templated *in trans* by a second (synapsed) DNA end. This structure showed an intrinsic difference with the eukaryotic system: working as a dimer *versus* a monomer, a two-handed *versus* a one-handed way of fixing broken DNA (Fig. 10). Despite this, and the different origins of the prokaryotic and eukaryotic NHEJ polymerases (AEP family of primases versus X family of DNA polymerases, respectively), we will discuss how these two systems share an unexpected amount of functional and structural features, making it a striking example of convergent evolution.

Mycobacterium tuberculosis PolDom is a unique polymerase with a variety of activities on different NHEJ DNA substrates, displaying terminal transferase activity on blunt and ssDNA substrates and templated polymerization: directed *in cis* on gapped and 5'-protruding substrates [22, 141, 142], and *in trans* on 3'-protruding substrates [103, 140]. The architecture of the bacterial NHEJ polymerases is different to that of the eukaryotic NHEJ polymerases from the X family, although the triad of metal-chelating aspartates is conserved and structurally over-imposable (Fig. 11A), a suggestion of the convergent evolution leading to similar catalytic mechanisms. But the convergence does not stop there: in all the activities tested, PolDom shows a marked preference for the insertion of ribonucleotides over deoxynucleotides. This preference, a consequence of the origins of PolDom from the AEP family of primases, reflects a catalytic plasticity that is maintained during evolution on other unrelated NHEJ polymerases such as Polµ [55, 56], and now serves a different purpose: to take advantage of the most abundant substrates during a laborious reaction. And, like the eukaryotic NHEJ ligase, the bacterial LigD ligates DNA containing ribonucleotides at the 3'-OH terminus [142, 143].

Another example of the common characteristics of the prokaryotic and eukaryotic NHEJ polymerases is the presence of a binding pocket for the 5'-P group of the downstream piece of DNA (Fig. 11B). This pocket, which contains residues Lys[16] and Lys[26], is missing in AEPs from *Archaea* and *Eukarya,* and is the major determinant for the specific binding of PolDom to its substrates, as the interaction significantly enhances its activity [22]. While Polµ or Polλ use a specific HhH motif at the 8 kDa domain to bind the phosphate, PolDom lacks this HhH and must therefore utilize a novel structural element to facilitate this interaction.
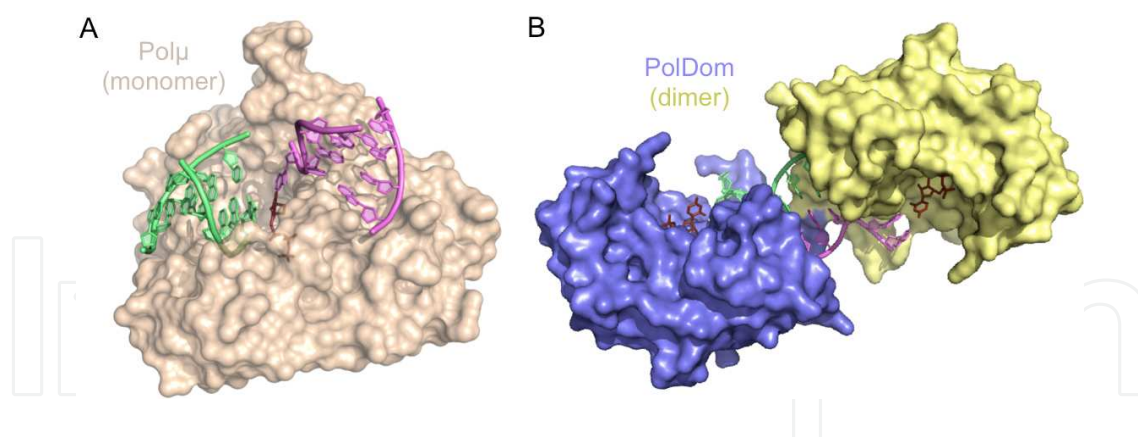
Figure 10. Different solutions for the NHEJ polymerases: monomers and dimers. A) Surface representation of a monomer of Polμ holding two pieces of DNA (green and mauve). B) Surface representation of a dimeric arrangement of *Mt*-PolDom (yellow and blue monomers) bridging two DNA ends (green and mauve).

Although recent studies have provided unique insights into polymerase-mediated orchestration of break synapsis, the order of substrate binding events and mechanism by which these NHEJ polymerases catalyze end-extension is still poorly understood. To address this question, in collaboration with Prof. Doherty (GDSC, University of Sussex), we elucidated the functional meaning of a novel crystal structure of a pre-ternary intermediate of *Mt*-PolDom bound to DNA, showing that this complex is relevant for specific DSB repair processing events [103]. This catalytically competent complex consists of a PolDom monomer, containing two metal ions and a templated nucleotide (UTP) in its active site, bound to a dsDNA end with a 3′ overhang but, significantly, lacking a primer strand. To our knowledge, this structure represents a unique example of a polymerase-DNA complex captured in a pre-ternary intermediate state, relevant for NHEJ.

Is the pre-ternary complex physiologically relevant for prokaryotic NHEJ polymerase extension reactions? Although the pre-ternary complex lacks an incoming primer strand, which provides the attacking nucleophile (3′-OH), a comparison of the positioning of the nucleotide base, phosphate tail, active site ligands and divalent metal ions to those in the active site of a polymerase ternary complex (Polλ) provides compelling evidence that the PolDom pre-ternary complex is catalytically competent (Fig. 11A). The possibility of preforming a pre-ternary complex in solution by incubating the necessary components (PolDom, DNA end, complementary nucleotide and activating metal ions) in the absence of a primer, allowed us to demonstrate its physiological relevance in accelerating NHEJ reactions, probably by providing a "ready to use" primer binding site. By testing the activity of the pre-ternary PolDom complex with different ssDNA primers, we concluded that the minimal primer utilizable by these enzymes is a dinucleotide, as PolDom was not proficient at polymerizing off a single nucleotide "primer". This fact indicates that, although PolDom is evolutionarily related to replicative AEPs, its physiological activity as a primase has effectively been lost and, instead, these polymerases have evolved to have a more restricted capacity to bind short incoming DNA termini, enabling them to perform more specialized roles in NHEJ break repair processes. The innate ability of AEPs to accept short primers may have influenced evolutionary selection of these enzymes by prokaryotes to become the NHEJ polymerase. Indeed, many

bacteria encode additional AEP orthologues whose physiological roles have yet to be determined. Is pre-ternary complex formation also relevant for eukaryotic NHEJ polymerases? It has been demonstrated that human Polμ can catalyze NHEJ extensions on very short and non-complementary DNA ends [29, 144], a reaction that can take advantage of a limited terminal transferase activity [93], and that can occur with both dNTPs and NTPs [145]. It is likely that formation of a Polμ pre-ternary complex, triggered by the strong recognition of a 5′-recessive phosphate and a reinforced avidity for the incoming nucleotide (both properties also intrinsic to Polμ), would be beneficial to carry out non-complementary NHEJ of minimally processed ends in eukaryotes, although this remains to be proven.
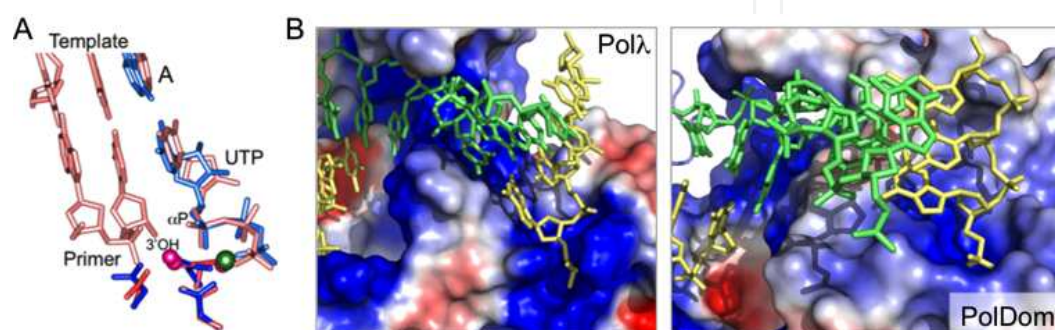
Figure 11. Similarities among the eukaryotic and prokaryotic NHEJ polymerases. A) Superimposition of the ternary complex of Polλ (1XSN) and the pre-ternary complex of Mt-PolDom (3PKY). B) Electrostatic surface of the Polλ and Mt-PolDom 5′-P binding pockets. DNA substrates are shown in green (template strand) and yellow (primer and downstream strands).

From a mechanistic point of view, our study of PolDom identified a conserved loop (loop 2), which plays a prominent role in the activation of the catalytic center. The conformation of loop 2 changes significantly, upon the templated-binding of the correct incoming nucleotide, which induces the rotation of Arg^{220} side-chain (~180°) away from the active site in the pre-ternary complex. Mutation of this invariant residue abolished the extension activity but, significantly, did not alter enzyme binding to other DNA substrates, such as gapped DNA. A comparison of the structures of the PolDom-DNA binary *versus* the pre-ternary complexes reveals the sequential movements that occur in the active site, induced by the binding of both a templating base and an incoming nucleotide. The invariant active site residue Phe^{64}, which stacks against the base of the incoming nucleotide in the PolDom-GTP binary complex, now stacks against the base of the templating nucleotide both in PolDom-DNA binary and pre-ternary complexes, orienting this base and also maintaining (together with Phe^{63}) the major kink in the template strand (~105º). In replicative DNA polymerases, aromatic tyrosine residues are commonly employed as a part of a fidelity mechanism that scrutinizes pairing of the correct incoming base with the templating base, thus acting as a molecular gatekeeper to limit the incorporation of an incorrect/mismatched base during elongation [146]. We propose that an analogous fidelity mechanism involving the two invariant phenylalanine residues also occurs in the bacterial NHEJ polymerases, but in the absence of the primer strand, thus ensuring that the correctly templated incoming base is bound in the active site prior to the encounter with the incoming end/primer providing the attacking 3′-OH.
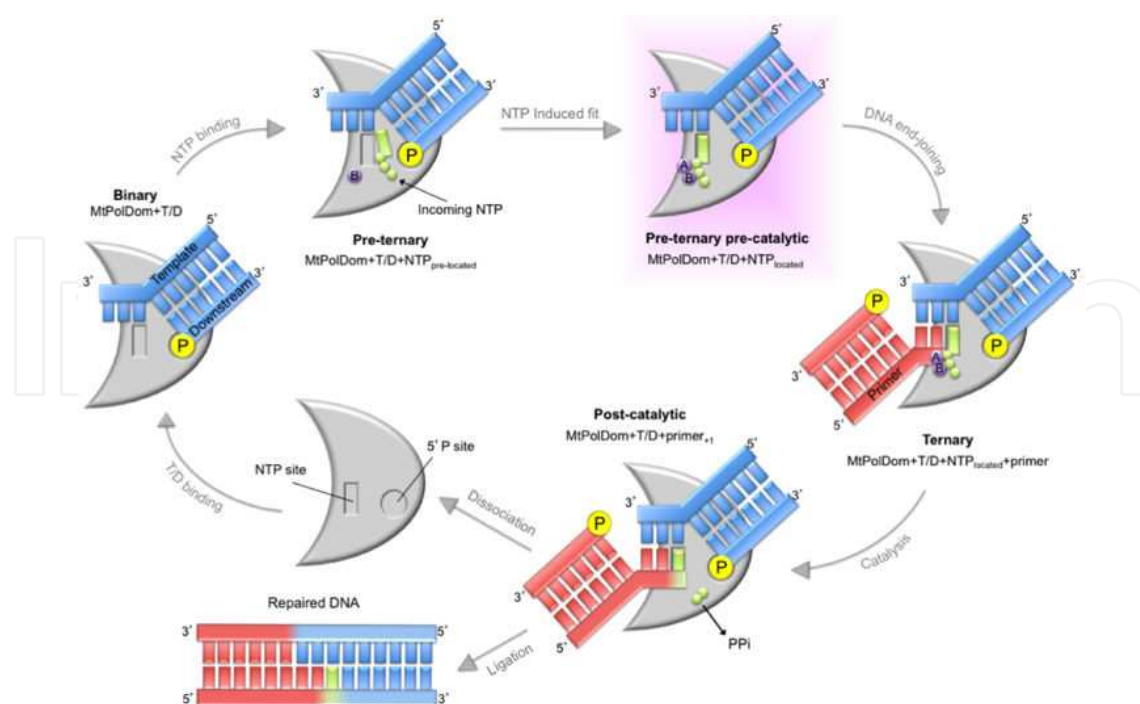
Figure 12. Catalytic Cycle of a Prokaryotic NHEJ Polymerase. Initially, a binary complex between the PolDom enzyme (gray crescent) and DNA (T/D; blue) is formed, mainly stabilized via interactions with the 5'-P. Binding of an incoming NTP (green) forms a preternary complex, still incompetent for catalysis as it lacks metal A. Upon template selection and relocation of the complementary NTP and the two metals, A and B, at the correct site (representing a primer-independent NTP induced-fit step) a preternary precatalytic complex is formed. This activated complex is ready for DNA end joining, allowing the 3'-OH of the incoming primer strand (red) to bind in the active site to form the ternary complex. Further steps of extension, PPi release, dissociation, and ligation (performed by the ligase domain of LigD), complete the DNA repair process.

This phenylalanine-mediated (Phe[64]) stacking interaction with the templating base in the preternary complex also promotes the movement of the incoming nucleotide (UTP) into the active site and, together with the loss of specific contacts (e.g. Arg[246], Lys[175], Lys[52]) promotes the correct repositioning of the α-phosphate group of the incoming nucleotide for catalysis. This reoriented α-phosphate moiety, together with Asp[139], forms a second metal binding site (A) not present in the binary structure, which is required for the two metal catalytic mechanism common to all DNA polymerases [147]. The binding of the second metal, in turn, promotes breakage of the salt bridge between Arg[220] and Asp[139], repositioning this aspartate into a catalytically favorable alignment with the other catalytic aspartates, the α-phosphate group and the two bound metal ions, to form an activated pre-ternary intermediate awaiting the arrival of the nucleophile (3'-OH of the primer strand). The catalytic incompetence of the R220A mutant highlights the importance of the interaction of Arg[220] with Asp[139]. We propose that the maintenance of this amino acid pairing provides a significant barrier to catalysis until the enzyme becomes optimally bound to DNA, metals, and the correct incoming templated nucleotide. Once these are bound within the active site, a sequence of structural rearrangements promotes the binding of a second metal ion (A). The affinity of Asp[139] for this second metal promotes the loss of interaction with Arg[220], leading to expulsion of loop 2 from the active site, which results in full activation of the catalytic center. The movement of loop 2 away from

the active site, most likely, promotes this activation step in two ways. The first consequence is that breaking the salt bridge is irreversible, leading to the release of the acidic side-chain of Asp[139], which is involved in the binding of the second metal (A) within the active site, ensuring that it is optimally poised for catalysis. The second notable consequence, induced by the reorientation of loop 2, is a significant change in the ridge that surrounds the active site, which most likely allows the 3'-OH group of the incoming primer strand to bind in the active site and form the complete ternary complex. Further steps of catalysis, PPi release, and ligation would lead to the conclusion of the NHEJ process. A scheme of the different complexes formed during the whole NHEJ cycle is depicted in figure 12. It is remarkable how, despite the different origins of PolDom and Polß, a similar mechanism of prevention of catalysis exists in both of them: an arginine residue contacts one of the catalytic aspartates, keeping it in an unproductive conformation that does not allow catalysis until binding of the nucleotide.

We have intensively studied the loops and flexible elements in Polμ, and examined the structure and the mutagenesis studies we have performed on PolDom, reaching the conclusion that both enzymes rely on those movable pieces to perform their most specific activities. As an even more striking example of convergent evolution, PolDom possesses a prominent surface ß-hairpin structure, loop 1, which is specific to NHEJ AEPs. Conserved residues in loop 1 interact with the 3' protrusion of NHEJ substrates and orient the synapsis of the ends [140]. Mutation of the apical residues of loop 1 to alanine did not affect binding to a primer-containing (gap) substrate, but abolished the ability of PolDom to form a synaptic complex [140] and, consequently, to catalyze trans-directed additions. Loop 1 in Polμ is also specific for binding and activity on NHEJ substrates [80, 92], through its function in the stabilization of the synapsis of two DNA ends.

13. Conclusion

In recent years, structural genomics has given rise to a vast array of knowledge, which nonetheless needs to be interpreted correctly as a range of still snapshots of a movie that, if seen, would show the highly complex and ever-moving machines that polymerases are. Helped by the biochemistry, and placed in context by the *in vivo* data, this structural approach has been used here to better understand the unique properties of each of the human DNA polymerases of the X family, and also of their bacterial counterparts. Thorough analysis of these structures has provided us with a deeper understanding of the unique abilities attributed to each polymerase.

Acknowledgements

pleasant years of parallel and coordinated research, and all the members of the Blanco lab for their dedicated work.

Author details

Maria Jose Martin and Luis Blanco[*]

Centro de Biología Molecular Severo Ochoa (CSIC-UAM), Madrid, Spain

References

[1] Bebenek, K, & Kunkel, T. A. *Functions of DNA polymerases.* Adv Protein Chem, (2004). , 137-165.

[2] Burgers, P. M, et al. *Eukaryotic DNA polymerases: proposal for a revised nomenclature.* J Biol Chem, (2001). , 43487-43490.

[3] Hubscher, U, Maga, G, & Spadari, S. *Eukaryotic DNA polymerases.* Annu Rev Biochem, (2002). , 133-163.

[4] Pavlov, Y. I, Shcherbakova, P. V, & Rogozin, I. B. *Roles of DNA polymerases in replication, repair, and recombination in eukaryotes.* Int Rev Cytol, (2006). , 41-132.

[5] Ohmori, H, et al. *The Y-family of DNA polymerases.* Mol Cell, (2001). , 7-8.

[6] Garcia-diaz, M, et al. *DNA polymerase lambda (Pol lambda), a novel eukaryotic DNA polymerase with a potential role in meiosis.* J Mol Biol, (2000). , 851-867.

[7] Dominguez, O, et al. *DNA polymerase mu (Pol mu), homologous to TdT, could act as a DNA mutator in eukaryotic cells.* EMBO J, (2000). , 1731-1742.

[8] Oliveros, M, et al. *Characterization of an African swine fever virus 20-kDa DNA polymerase involved in DNA repair.* J Biol Chem, (1997). , 30899-30910.

[9] Takeuchi, R, et al. *Drosophila DNA polymerase zeta interacts with recombination repair protein 1, the Drosophila homologue of human abasic endonuclease 1.* J Biol Chem, (2006). , 11577-11585.

[10] Uchiyama, Y, et al. *Distribution and roles of X-family DNA polymerases in eukaryotes.* Biochimie, (2009). , 165-170.

[11] Demogines, A, et al. *Ancient and recent adaptive evolution of primate non-homologous end joining genes.* PLoS Genet, (2010). , e1001169.

[12]   Kelley, J. L, et al. *Targeted resequencing of two genes, RAGE and POLL, confirms findings from a genome-wide scan for adaptive evolution and provides evidence for positive selection in additional populations.* Hum Mol Genet, (2009). , 779-784.

[13]   Sawyer, S. L, & Malik, H. S. *Positive selection of yeast nonhomologous end-joining genes and a retrotransposon conflict hypothesis.* Proc Natl Acad Sci U S A, (2006). , 17614-17619.

[14]   Bruton, R. K, et al. *C-terminal-binding protein interacting protein binds directly to adenovirus early region 1A through its N-terminal region and conserved region 3.* Oncogene, (2007). , 7467-7479.

[15]   Evans, J. D, & Hearing, P. *Relocalization of the Mre11-Rad50-Nbs1 complex by the adenovirus E4 ORF3 protein is required for viral replication.* J Virol, (2005). , 6207-6215.

[16]   Jayaram, S, et al. *E1B 55k-independent dissociation of the DNA ligase IV/XRCC4 complex by E4 34k during adenovirus infection.* Virology, (2008). , 163-170.

[17]   Kilzer, J. M, et al. *Roles of host cell factors in circularization of retroviral dna.* Virology, (2003). , 460-467.

[18]   Li, L, et al. *Role of the non-homologous DNA end joining pathway in the early steps of retroviral infection.* EMBO J, (2001). , 3272-3281.

[19]   Lin, C. W, & Engelman, A. *The barrier-to-autointegration factor is a component of functional human immunodeficiency virus type 1 preintegration complexes.* J Virol, (2003). , 5030-5036.

[20]   Stracker, T. H, Carson, C. T, & Weitzman, M. D. *Adenovirus oncoproteins inactivate the Mre11-Rad50-NBS1 DNA repair complex.* Nature, (2002). , 348-352.

[21]   Weitzman, M. D, et al. *Interactions of viruses with the cellular DNA repair machinery.* DNA Repair (Amst), (2004). , 1165-1173.

[22]   Pitcher, R. S, et al. *Structure and function of a mycobacterial NHEJ DNA repair polymerase.* J Mol Biol, (2007). , 391-405.

[23]   Garcia-escudero, R, et al. *DNA polymerase X of African swine fever virus: insertion fidelity on gapped DNA substrates and AP lyase activity support a role in base excision repair of viral DNA.* J Mol Biol, (2003). , 1403-1412.

[24]   Taladriz, S, et al. *Nuclear DNA polymerase beta from Leishmania infantum. Cloning, molecular analysis and developmental regulation.* Nucleic Acids Res, (2001). , 3822-3834.

[25]   Bebenek, K, et al. *Biochemical properties of Saccharomyces cerevisiae DNA polymerase IV.* J Biol Chem, (2005). , 20051-20058.

[26]   Gonzalez-barrera, S, et al. *Characterization of SpPol4, a unique X-family DNA polymerase in Schizosaccharomyces pombe.* Nucleic Acids Res, (2005). , 4762-4774.

[27] Tseng, H. M, & Tomkinson, A. E. *A physical and functional interaction between yeast Pol4 and Dnl4-Lif1 links DNA synthesis and ligation in nonhomologous end joining.* J Biol Chem, (2002). , 45630-45637.

[28] Tseng, H. M, & Tomkinson, A. E. *Processing and joining of DNA ends coordinated by interactions among Dnl4/Lif1, Pol4, and FEN-1.* J Biol Chem, (2004). , 47580-47588.

[29] Nick McElhinnyS.A., et al., *A gradient of template dependence defines distinct biological roles for family X polymerases in nonhomologous end joining.* Mol Cell, (2005). , 357-366.

[30] Lee, J. W, et al. *Implication of DNA polymerase lambda in alignment-based gap filling for nonhomologous DNA end joining in human nuclear extracts.* J Biol Chem, (2004). , 805-811.

[31] Manke, I. A, et al. *BRCT repeats as phosphopeptide-binding modules involved in protein targeting.* Science, (2003). , 636-639.

[32] Yu, X, et al. *The BRCT domain is a phospho-protein binding domain.* Science, (2003). , 639-642.

[33] Derose, E. F, et al. *Solution structure of polymerase mu's BRCT Domain reveals an element essential for its role in nonhomologous end joining.* Biochemistry, (2007). , 12100-12110.

[34] Martin, M. J, Juarez, R, & Blanco, L. *DNA-binding determinants promoting NHEJ by human Polmu.* Nucleic Acids Res, (2012). , 11389-11403.

[35] Matsumoto, T, et al. *BRCT domain of DNA polymerase mu has DNA-binding activity and promotes the DNA polymerization activity.* Genes Cells, (2012). , 790-806.

[36] Mueller, G. A, et al. *A comparison of BRCT domains involved in nonhomologous end-joining: introducing the solution structure of the BRCT domain of polymerase lambda.* DNA Repair (Amst), (2008). , 1340-1351.

[37] 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). , 1328-1333.

[38] Prasad, R, Beard, W. A, & Wilson, S. H. *Studies of gapped DNA substrate binding by mammalian DNA polymerase beta. Dependence on 5'-phosphate group.* J Biol Chem, (1994). , 18096-18101.

[39] Pelletier, H, et al. *Crystal structures of human DNA polymerase beta complexed with DNA: implications for catalytic mechanism, processivity, and fidelity.* Biochemistry, (1996). , 12742-12761.

[40] Moon, A. F, et al. *Structural insight into the substrate specificity of DNA Polymerase mu.* Nat Struct Mol Biol, (2007). , 45-53.

[41] Garcia-diaz, M, et al. *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). , 34659-34663.

[42] Prasad, R, et al. *Human DNA polymerase beta deoxyribose phosphate lyase. Substrate specificity and catalytic mechanism.* J Biol Chem, (1998). , 15263-15270.

[43] Garcia-diaz, M, et al. *A structural solution for the DNA polymerase lambda-dependent repair of DNA gaps with minimal homology.* Mol Cell, (2004). , 561-572.

[44] Singhal, R. K, & Wilson, S. H. *Short gap-filling synthesis by DNA polymerase beta is processive.* J Biol Chem, (1993). , 15906-15911.

[45] Garcia-diaz, M, et al. *DNA polymerase lambda, a novel DNA repair enzyme in human cells.* J Biol Chem, (2002). , 13184-13191.

[46] Doherty, A. J, Serpell, L. C, & Ponting, C. P. *The helix-hairpin-helix DNA-binding motif: a structural basis for non-sequence-specific recognition of DNA.* Nucleic Acids Res, (1996). , 2488-2497.

[47] Mullen, G. P, & Wilson, S. H. *DNA polymerase beta in abasic site repair: a structurally conserved helix-hairpin-helix motif in lesion detection by base excision repair enzymes.* Biochemistry, (1997). , 4713-4717.

[48] Sobol, R. W, et al. *The lyase activity of the DNA repair protein beta-polymerase protects from DNA-damage-induced cytotoxicity.* Nature, (2000). , 807-810.

[49] Beard, W. A, & Wilson, S. H. *Structure and mechanism of DNA polymerase Beta.* Chem Rev, (2006). , 361-382.

[50] Garcia-diaz, M, et al. *Structure-function studies of DNA polymerase lambda.* DNA Repair (Amst), (2005). , 1358-1367.

[51] Prasad, R, et al. *Structural insight into the DNA polymerase beta deoxyribose phosphate lyase mechanism.* DNA Repair (Amst), (2005). , 1347-1357.

[52] Ruiz, J. F, et al. *DNA polymerase mu, a candidate hypermutase?* Philos Trans R Soc Lond B Biol Sci, (2001). , 99-109.

[53] Ruiz, J. F, et al. *Overexpression of human DNA polymerase mu (Pol mu) in a Burkitt's lymphoma cell line affects the somatic hypermutation rate.* Nucleic Acids Res, (2004). , 5861-5873.

[54] Zhang, Y, et al. *Highly frequent frameshift DNA synthesis by human DNA polymerase mu.* Mol Cell Biol, (2001). , 7995-8006.

[55] Nick McElhinnyS.A. and D.A. Ramsden, *Polymerase mu is a DNA-directed DNA/RNA polymerase.* Mol Cell Biol, (2003). , 2309-2315.

[56] Ruiz, J. F, et al. *Lack of sugar discrimination by human Pol mu requires a single glycine residue.* Nucleic Acids Res, (2003). , 4441-4449.

[57] Lieber, M. R, et al. *The mechanism of vertebrate nonhomologous DNA end joining and its role in V(D)J recombination.* DNA Repair (Amst), (2004). , 817-826.

[58] Ferguson, D. O, et al. *The nonhomologous end-joining pathway of DNA repair is required for genomic stability and the suppression of translocations.* Proc Natl Acad Sci U S A, (2000). , 6630-6633.

[59] Heidenreich, E, et al. *Non-homologous end joining as an important mutagenic process in cell cycle-arrested cells.* EMBO J, (2003). , 2274-2283.

[60] Walker, J. R, Corpina, R. A, & Goldberg, J. *Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair.* Nature, (2001). , 607-614.

[61] Dvir, A, et al. *Ku autoantigen is the regulatory component of a template-associated protein kinase that phosphorylates RNA polymerase II.* Proc Natl Acad Sci U S A, (1992). , 11920-11924.

[62] Gottlieb, T. M, & Jackson, S. P. *The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen.* Cell, (1993). , 131-142.

[63] Dynan, W. S, & Yoo, S. *Interaction of Ku protein and DNA-dependent protein kinase catalytic subunit with nucleic acids.* Nucleic Acids Res, (1998). , 1551-1559.

[64] Chen, L, et al. *Interactions of the DNA ligase IV-XRCC4 complex with DNA ends and the DNA-dependent protein kinase.* J Biol Chem, (2000). , 26196-26205.

[65] Defazio, L. G, et al. *Synapsis of DNA ends by DNA-dependent protein kinase.* EMBO J, (2002). , 3192-3200.

[66] Yaneva, M, Kowalewski, T, & Lieber, M. R. *Interaction of DNA-dependent protein kinase with DNA and with Ku: biochemical and atomic-force microscopy studies.* EMBO J, (1997). , 5098-5112.

[67] Kramer, K. M, et al. *Two different types of double-strand breaks in Saccharomyces cerevisiae are repaired by similar RAD52-independent, nonhomologous recombination events.* Mol Cell Biol, (1994). , 1293-1301.

[68] Moore, J. K, & Haber, J. E. *Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in Saccharomyces cerevisiae.* Mol Cell Biol, (1996). , 2164-2173.

[69] Roth, D. B, & Wilson, J. H. *Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction.* Mol Cell Biol, (1986). , 4295-4304.

[70] Wilson, T. E, & Lieber, M. R. *Efficient processing of DNA ends during yeast nonhomologous end joining. Evidence for a DNA polymerase beta (Pol4)-dependent pathway.* J Biol Chem, (1999). , 23599-23609.

[71] Ma, Y, et al. *Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination.* Cell, (2002). , 781-794.

[72] Hefferin, M. L, & Tomkinson, A. E. *Mechanism of DNA double-strand break repair by non-homologous end joining.* DNA Repair (Amst), (2005). , 639-648.

[73] Karimi-busheri, F, et al. *Molecular characterization of a human DNA kinase.* J Biol Chem, (1999). , 24187-24194.

[74] Chappell, C, et al. *Involvement of human polynucleotide kinase in double-strand break repair by non-homologous end joining.* EMBO J, (2002). , 2827-2832.

[75] Grawunder, U, et al. *Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells.* Nature, (1997). , 492-495.

[76] Schar, P, et al. *A newly identified DNA ligase of Saccharomyces cerevisiae involved in RAD52-independent repair of DNA double-strand breaks.* Genes Dev, (1997). , 1912-1924.

[77] Teo, S. H, & Jackson, S. P. *Identification of Saccharomyces cerevisiae DNA ligase IV: involvement in DNA double-strand break repair.* EMBO J, (1997). , 4788-4795.

[78] Wilson, T. E, Grawunder, U, & Lieber, M. R. *Yeast DNA ligase IV mediates non-homologous DNA end joining.* Nature, (1997). , 495-498.

[79] Ramsden, D. A. *Polymerases in nonhomologous end joining: building a bridge over broken chromosomes.* Antioxid Redox Signal, (2011). , 2509-2519.

[80] Juarez, R, et al. *A specific loop in human DNA polymerase mu allows switching between creative and DNA-instructed synthesis.* Nucleic Acids Res, (2006). , 4572-4582.

[81] Ma, Y, et al. *A biochemically defined system for mammalian nonhomologous DNA end joining.* Mol Cell, (2004). , 701-713.

[82] Ahnesorg, P, Smith, P, & Jackson, S. P. *XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining.* Cell, (2006). , 301-313.

[83] Hentges, P, et al. *Evolutionary and functional conservation of the DNA non-homologous end-joining protein, XLF/Cernunnos.* J Biol Chem, (2006). , 37517-37526.

[84] Papavasiliou, F, et al. *V(D)J recombination in mature B cells: a mechanism for altering antibody responses.* Science, (1997). , 298-301.

[85] Bertocci, B, et al. *Immunoglobulin kappa light chain gene rearrangement is impaired in mice deficient for DNA polymerase mu.* Immunity, (2003). , 203-211.

[86] Gozalbo-lopez, B, et al. *A role for DNA polymerase mu in the emerging DJH rearrangements of the postgastrulation mouse embryo.* Mol Cell Biol, (2009). , 1266-1275.

[87] Bentolila, L. A, et al. *The two isoforms of mouse terminal deoxynucleotidyl transferase differ in both the ability to add N regions and subcellular localization.* EMBO J, (1995). , 4221-4229.

[88] Bentolila, L. A, et al. *Extensive junctional diversity in Ig light chain genes from early B cell progenitors of mu MT mice.* J Immunol, (1999). , 2123-2128.

[89] Delarue, M, et al. *Crystal structures of a template-independent DNA polymerase: murine terminal deoxynucleotidyltransferase.* EMBO J, (2002). , 427-439.

[90] Romain, F, et al. *Conferring a template-dependent polymerase activity to terminal deoxynu-cleotidyltransferase by mutations in the Loop1 region.* Nucleic Acids Res, (2009). , 4642-4656.

[91] Bebenek, K, et al. *Loop 1 modulates the fidelity of DNA polymerase lambda.* Nucleic Acids Res, (2010). , 5419-5431.

[92] Martin, M. J. *Exclusive Polymerases Repairing Double Strand Breaks. The same magics from bacteria to man.,* (2011). Universidad Autonoma de Madrid.

[93] Andrade, P, et al. *Limited terminal transferase in human DNA polymerase mu defines the required balance between accuracy and efficiency in NHEJ.* Proc Natl Acad Sci U S A, (2009). , 16203-16208.

[94] Bebenek, K, et al. *The frameshift infidelity of human DNA polymerase lambda. Implications for function.* J Biol Chem, (2003). , 34685-34690.

[95] Bertocci, B, et al. *Nonoverlapping functions of DNA polymerases mu, lambda, and terminal deoxynucleotidyltransferase during immunoglobulin V(D)J recombination in vivo.* Immuni-ty, (2006). , 31-41.

[96] Bollum, F. J. *Terminal deoxynucleotidyl transferase as a hematopoietic cell marker.* Blood, (1979). , 1203-1215.

[97] Coleman, M. S, Hutton, J. J, & Bollum, F. J. *Terminal riboadenylate transferase in human lymphocytes.* Nature, (1974). , 407-409.

[98] Kunkel, T. A, et al. *Rearrangements of DNA mediated by terminal transferase.* Proc Natl Acad Sci U S A, (1986). , 1867-1871.

[99] Bollum, F. J. *Mammalian enzymes of desoxyribonucleic acid synthesis.* Ann N Y Acad Sci, (1959). , 792-793.

[100] Bollum, F. J. *Chemically Defined Templates and Initiators for Deoxypolynucleotide Synthe-sis.* Science, (1964). , 560.

[101] Benkovic, S. J, & Cameron, C. E. *Kinetic analysis of nucleotide incorporation and misincor-poration by Klenow fragment of Escherichia coli DNA polymerase I.* Methods Enzymol, (1995). , 257-269.

[102] Deibel, M. R, & Jr, M. S. Coleman, *Biochemical properties of purified human terminal deoxynucleotidyltransferase.* J Biol Chem, (1980). , 4206-4212.

[103] Brissett, N. C, et al. *Structure of a preternary complex involving a prokaryotic NHEJ DNA polymerase.* Mol Cell, (2011). , 221-231.

[104] Roychoudhury, R. *Enzymic synthesis of polynucleotides. Oligodeoxynucleotides with one 3'-terminal ribonucleotide as primers for polydeoxynucleotide synthesis.* J Biol Chem, (1972). , 3910-3917.

[105] Lucas, D, et al. *Polymerase mu is up-regulated during the T cell-dependent immune response and its deficiency alters developmental dynamics of spleen centroblasts.* Eur J Immunol, (2005). , 1601-1611.

[106] Dianova, I. I, et al. *XRCC1-DNA polymerase beta interaction is required for efficient base excision repair.* Nucleic Acids Res, (2004). , 2550-2555.

[107] Gryk, M. R, et al. *Mapping of the interaction interface of DNA polymerase beta with XRCC1.* Structure, (2002). , 1709-1720.

[108] Marintchev, A, et al. *Domain specific interaction in the XRCC1-DNA polymerase beta complex.* Nucleic Acids Res, (2000). , 2049-2059.

[109] Fortini, P, et al. *Different DNA polymerases are involved in the short- and long-patch base excision repair in mammalian cells.* Biochemistry, (1998). , 3575-3580.

[110] Sobol, R. W, et al. *Requirement of mammalian DNA polymerase-beta in base-excision repair.* Nature, (1996). , 183-186.

[111] Sobol, R. W, & Wilson, S. H. *Mammalian DNA beta-polymerase in base excision repair of alkylation damage.* Prog Nucleic Acid Res Mol Biol, (2001). , 57-74.

[112] Kimura, S, & Sakaguchi, K. *DNA repair in plants.* Chem Rev, (2006). , 753-766.

[113] Uchiyama, Y, Suzuki, Y, & Sakaguchi, K. *Characterization of plant XRCC1 and its interaction with proliferating cell nuclear antigen.* Planta, (2008). , 1233-1241.

[114] Radford, S. J, et al. *Heteroduplex DNA in meiotic recombination in Drosophila mei-9 mutants.* Genetics, (2007). , 63-72.

[115] Alseth, I, et al. *Biochemical characterization and DNA repair pathway interactions of Mag1-mediated base excision repair in Schizosaccharomyces pombe.* Nucleic Acids Res, (2005). , 1123-1131.

[116] Hirose, F, et al. *Difference in the expression level of DNA polymerase beta among mouse tissues: high expression in the pachytene spermatocyte.* Exp Cell Res, (1989). , 169-180.

[117] Sugo, N, et al. *Neonatal lethality with abnormal neurogenesis in mice deficient in DNA polymerase beta.* EMBO J, (2000). , 1397-1404.

[118] Nakamura, J, & Swenberg, J. A. *Endogenous apurinic/apyrimidinic sites in genomic DNA of mammalian tissues.* Cancer Res, (1999). , 2522-2526.

[119] Wilson, D. M, & Rd, D. R. McNeill, *Base excision repair and the central nervous system.* Neuroscience, (2007). , 1187-1200.

[120] Gu, H, et al. *Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting.* Science, (1994). , 103-106.

[121] Allinson, S. L, Dianova, I. I, & Dianov, G. L. *DNA polymerase beta is the major dRP lyase involved in repair of oxidative base lesions in DNA by mammalian cell extracts.* EMBO J, (2001). , 6919-6926.

[122] Sugo, N, et al. *p53 deficiency rescues neuronal apoptosis but not differentiation in DNA polymerase beta-deficient mice.* Mol Cell Biol, (2004). p. 9470-7.

[123] Cabelof, D. C, et al. *Haploinsufficiency in DNA polymerase beta increases cancer risk with age and alters mortality rate.* Cancer Res, (2006). , 7460-7465.

[124] Allen, D, et al. *Mutagenesis is elevated in male germ cells obtained from DNA polymerase-beta heterozygous mice.* Biol Reprod, (2008). , 824-831.

[125] Kidane, D, et al. *DNA polymerase beta is critical for genomic stability of sperm cells.* DNA Repair (Amst), (2011). , 390-397.

[126] Niimi, N, et al. *Decreased mutant frequency in embryonic brain of DNA polymerase beta null mice.* Mutagenesis, (2006). , 55-59.

[127] Senejani, A. G, et al. *Y265C DNA polymerase beta knockin mice survive past birth and accumulate base excision repair intermediate substrates.* Proc Natl Acad Sci U S A, (2012). , 6632-6637.

[128] Washington, S. L, et al. *A genetic system to identify DNA polymerase beta mutator mutants.* Proc Natl Acad Sci U S A, (1997). , 1321-1326.

[129] Opresko, P. L, Sweasy, J. B, & Eckert, K. A. *The mutator form of polymerase beta with amino acid substitution at tyrosine 265 in the hinge region displays an increase in both base substitution and frame shift errors.* Biochemistry, (1998). , 2111-2119.

[130] Kobayashi, Y, et al. *Hydrocephalus, situs inversus, chronic sinusitis, and male infertility in DNA polymerase lambda-deficient mice: possible implication for the pathogenesis of immotile cilia syndrome.* Mol Cell Biol, (2002). , 2769-2776.

[131] Zariwala, M, et al. *Investigation of the possible role of a novel gene, DPCD, in primary ciliary dyskinesia.* Am J Respir Cell Mol Biol, (2004). , 428-434.

[132] Bertocci, B, et al. *Cutting edge: DNA polymerases mu and lambda are dispensable for Ig gene hypermutation.* J Immunol, (2002). , 3702-3706.

[133] Braithwaite, E. K, et al. *DNA polymerase lambda protects mouse fibroblasts against oxidative DNA damage and is recruited to sites of DNA damage/repair.* J Biol Chem, (2005). , 31641-31647.

[134] Braithwaite, E. K, et al. *DNA polymerase lambda mediates a back-up base excision repair activity in extracts of mouse embryonic fibroblasts.* J Biol Chem, (2005). , 18469-18475.

[135] Maga, G, et al. *Replication protein A and proliferating cell nuclear antigen coordinate DNA polymerase selection in 8-oxo-guanine repair.* Proc Natl Acad Sci U S A, (2008). , 20689-20694.

[136] Gilfillan, S, et al. *Mice lacking TdT: mature animals with an immature lymphocyte repertoire.* Science, (1993). , 1175-1178.

[137] Komori, T, et al. *Lack of N regions in antigen receptor variable region genes of TdT-deficient lymphocytes.* Science, (1993). , 1171-1175.

[138] Cabaniols, J. P, et al. *Most alpha/beta T cell receptor diversity is due to terminal deoxynucleotidyl transferase.* J Exp Med, (2001). , 1385-1390.

[139] Lucas, D, et al. *Altered hematopoiesis in mice lacking DNA polymerase mu is due to inefficient double-strand break repair.* PLoS Genet, (2009). , e1000389.

[140] Brissett, N. C, et al. *Structure of a NHEJ polymerase-mediated DNA synaptic complex.* Science, (2007). , 456-459.

[141] Zhu, H, & Shuman, S. *Substrate specificity and structure-function analysis of the 3'-phosphoesterase component of the bacterial NHEJ protein, DNA ligase D.* J Biol Chem, (2006). , 13873-13881.

[142] Della, M, et al. *Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine.* Science, (2004). , 683-685.

[143] Yakovleva, L, & Shuman, S. *Nucleotide misincorporation, 3'-mismatch extension, and responses to abasic sites and DNA adducts by the polymerase component of bacterial DNA ligase D.* J Biol Chem, (2006). , 25026-25040.

[144] Davis, B. J, Havener, J. M, & Ramsden, D. A. *End-bridging is required for pol mu to efficiently promote repair of noncomplementary ends by nonhomologous end joining.* Nucleic Acids Res, (2008). , 3085-3094.

[145] Martin, M. J, et al. *Ribonucleotides and manganese ions improve non-homologous end joining by human Polmu.* Nucleic Acids Res, (2012).

[146] Johnson, K. A. *The kinetic and chemical mechanism of high-fidelity DNA polymerases.* Biochim Biophys Acta, (2010). , 1041-1048.

[147] Brautigam, C. A, & Steitz, T. A. *Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes.* Curr Opin Struct Biol, (1998). , 54-63.