Creative template-dependent synthesis by human polymerase mu

Andrea F. Moon, Rajendrakumar A. Gosavi, Thomas A. Kunkel, Lars C. Pedersen[1], and Katarzyna Bebenek

Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709

Edited by Graham C. Walker, Massachusetts Institute of Technology, Cambridge, MA, and approved July 7, 2015 (received for review March 23, 2015)

Among the many proteins used to repair DNA double-strand breaks by nonhomologous end joining (NHEJ) are two related family X DNA polymerases, Pol λ and Pol μ. Which of these two polymerases is preferentially used for filling DNA gaps during NHEJ partly depends on sequence complementarity at the break, with Pol λ and Pol μ repairing complementary and noncomplementary ends, respectively. To better understand these substrate preferences, we present crystal structures of Pol μ on a 2-nt gapped DNA substrate, representing three steps of the catalytic cycle. In striking contrast to Pol λ, Pol μ "skips" the first available template nucleotide, instead using the template base at the 5′ end of the gap to direct nucleotide binding and incorporation. This remarkable divergence from canonical 3′-end gap filling is consistent with data on end-joining substrate specificity in cells, and provides insights into polymerase substrate choices during NHEJ.

DNA polymerase mu | DNA polymerase lambda | DNA repair | nonhomologous end joining

DNA double-strand breaks (DSBs) result from exposure to endogenous and exogenous factors including reactive oxygen species, physical or mechanical stress, ionizing radiation, or activities of nuclear enzymes on DNA (1). Another source of DSBs is programmed DNA breakage during meiotic or mitotic recombination, immunoglobin gene rearrangement in V(D)J, and class-switch recombination (1, 2). Because of the largely random nature of accidental double-strand breakage, the broken ends can have widely varying sequences, structures, or modifications. Nonhomologous end joining (NHEJ) is the predominant form of DSB repair in higher eukaryotes, and requires a certain degree of flexibility from the many factors involved in this complex repair process.

DSB substrates lacking microhomology at the break site are unsuitable for immediate rejoining by ligase IV, and a polymerase is usually required to fill the gaps to generate ends that can be efficiently ligated (3). In higher eukaryotes, this role is performed by family X polymerase λ (Pol λ) and polymerase μ (Pol μ) (4–6). Pol λ has a strong preference for substrates with complementary template-strand pairing opposite the primer terminus. Pol μ can also use such complementary DSB substrates, but is uniquely active in template-dependent synthesis on DSBs entirely lacking complementarity, where the primer terminus is unpaired (7).

Given partial overlap in substrate use in vitro by Pol μ and Pol λ, how does the NHEJ machinery select which enzyme to repair specific substrates with the highest fidelity? Recent work by Pryor et al. (8) (companion article in this issue) has demonstrated a strong preference for Pol λ over Pol μ in repairing the majority of complementary DSBs. This is advantageous, because Pol λ displays higher fidelity of synthesis (9, 10). However, Pol λ cannot use noncomplementary DSB substrates with unpaired primer termini. Pol μ is the only known polymerase that can use noncomplementary substrates, such that in Pol μ knockout cells, these ends are largely unrepaired (figures 1B and 3A in ref. 8, the companion article). To better understand the substrate preferences of Pol λ and Pol μ, here we present crystal structures of the human Pol μ (hPol μ) catalytic domain in complex with a 2-nt gapped DNA substrate, from DNA binding through nucleotide binding and incorporation. These structures are consistent with the in vivo NHEJ assays from Pryor et al. (8) and support a "1-nt gap" spacing rule for Pol μ in which Pol μ prefers to engage all substrates in a similar fashion, using the 5′ unpaired template base in the gap as though it were a single-nucleotide (1-nt) gap. Additionally, this study provides insights into Pol μ substrate preference during in vitro polymerization reactions and how this enzyme might function during NHEJ in vivo.

Results

Kinetic Behavior of Human Pol μ Versus Human Pol λ on 1- and 2-nt Gapped DNA Substrates. hPol μ operates most effectively on 1-nt gaps, with efficiency decreasing as gap length increases (11). By comparison, Pol λ displays less length specificity and can fill longer gaps (12, 13). In the context of DSB substrates lacking any sequence complementarity, namely substrates with an unpaired primer terminus, Pol μ uses 1- and 2-nt gaps equally well (figures 1B and 3A in ref. 8, the companion article). Interestingly, however, Pol μ appears to interact with short gaps of any length as though they were 1-nt gaps, preferentially filling the gap in a template-dependent manner, using the 5′ unpaired template base in the gap rather than the first available 3′ unpaired base proximal to the primer terminus, which results in deletion frameshifts. This behavior could accurately be described as "skipping ahead." In contrast, Pol λ fills gaps using each template base sequentially, starting from the base closest to the primer terminus at the 3′ end of the gap.

To better understand this unique substrate utilization, the kinetic parameters of Pol μ and Pol λ were compared for 1- and 2-nt

Significance

Template-dependent DNA polymerases usually add nucleotides to the 3′ end of a primer, using the first available template-strand nucleotide as a guide. This behavior holds true for all polymerases, except one, DNA polymerase μ. When presented with 2-nt single- or double-strand gaps, polymerase μ (Pol μ) engages the substrate with the last available template-strand nucleotide closest to the 5′-phosphate on the downstream end of the gap, guiding synthesis. Crystal structures of Pol μ with a 2-nt gapped DNA substrate explain how the unpaired base is accommodated in the active site, and yield insights into the behavior of this polymerase within the context of nonhomologous end joining in DNA double-strand break repair.

Author contributions: A.F.M. and K.B. designed research; A.F.M., R.A.G., and K.B. performed research; A.F.M., T.A.K., L.C.P., and K.B. analyzed data; and A.F.M., T.A.K., L.C.P., and K.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The crystallography, atomic coordinates, and structure factors for human Pol μ Δ2 in binary, precatalytic ternary, and postcatalytic nicked complexes with bound DNA substrates reported in this paper have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 4YCX, 4YD1, and 4YD2, respectively).

[1]To whom correspondence should be addressed. Email: pederse2@niehs.nih.gov.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1505798112/-/DCSupplemental.

gap-filling synthesis (Fig. 1 and Table S1). Pol λ has higher catalytic efficiency on the 1- and 2-nt gapped substrates than does Pol μ, and is slightly (twofold) more proficient in filling the 2-nt gap than the 1-nt gap. These results are consistent with published kinetic constants for Pol λ in filling of short-gap single-strand breaks (14). Conversely, Pol μ prefers to fill 1-nt rather than 2-nt gaps (12-fold preference of 1- to 2-nt gaps). For Pol μ, this difference is explained by a 9.5-fold decrease in nucleotide binding affinity and a slight decrease in catalytic rate (Table S1). As suggested by Pryor et al. (figures 3A and 4B in ref. 8, the companion article), Pol μ addresses 2-nt gapped substrates by a single polymerization event, using the templating nucleotide at the 5′ end of the gap even when supplied with an excess of incoming nucleotide complementary to the 3′ templating nucleotide (up to 18-fold preference) (Fig. 1*A*). By contrast, Pol λ exhibited similar catalytic rates and binding affinities for both substrates. Conversely to Pol μ, Pol λ appears to be incapable of skipping ahead, even when sequence context and reaction conditions strongly favor that outcome (Fig. 1*A*).

Structural Characterization of hPol μ Δ2 in Complex with a 2-nt Gapped Substrate. A truncated variant of the catalytic domain of hPol μ (Pro132–Ala494) lacking the large, flexible loop between β-strands 4 and 5 (Pro398–Pro410, loop 2; henceforth referred to as hPol μ Δ2) was crystallized as a binary complex with a 2-nt gapped DNA substrate (Fig. 2 *A*–*C*). This variant has enzymatic properties indistinguishable from those of the wild-type catalytic domain, and has a higher propensity to crystallize (15). Using hPol μ Δ2, the catalytic cycle of synthesis on a 2-nt gapped DNA substrate has been structurally characterized, from the binary complex [Protein Data Bank (PDB) ID code 4YCX] through nucleotide incorporation. Crystals of the binary complex were soaked with an incoming nonhydrolyzable analog, dUMPNPP, to create a precatalytic ternary complex (Fig. 2*D*; PDB ID code 4YD1), or with an incoming hydrolyzable dTTP to allow in crystallo incorporation to create a postcatalytic nicked complex (Fig. 2*E*; PDB ID code 4YD2) (Table S2).

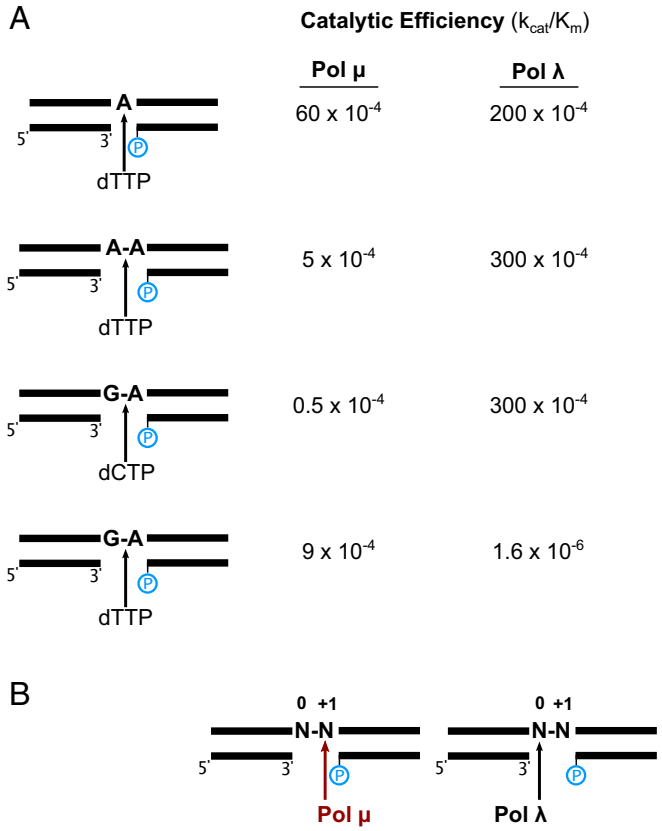

Fig. 1. Substrate utilization and kinetic parameters for Pols μ and λ. (*A*) Comparison of the catalytic efficiencies ($s^{-1}\cdot\mu M^{-1}$) for Pols μ and λ (*Right*) on varying 1- and 2-nt gapped DNA substrates (*Left*). Data are represented as mean ± SD from $n = 3$–4 replicates per experiment. (*B*) Line drawings of single-strand break DNA substrates show the locations of the template nucleotides in the gap relative to the 5′-phosphate (blue). Pols μ (*Left*; red) and λ (*Right*; black) address nucleotide insertion differently in 2-nt gapped substrates.

In the binary complex, hPol μ Δ2 exhibits the "left-hand" configuration characteristic of family X polymerases (5), with the 8-kDa subdomain (Trp137–Arg229) interacting with the duplex DNA downstream of the 2-nt gap (Fig. 2*B*). The fingers and palm subdomains position the upstream primer and template strands, respectively. Interactions between the thumb and the template strand shape the nascent base-pair binding site. Similar to the structure of hPol μ Δ2 in binary complex with a 1-nt gapped DNA substrate (15), both the protein and the DNA are present in the "closed" conformation, with the primer terminus adopting a C3′-*endo* sugar pucker (Fig. 2*C*). One divalent magnesium ion occupies the metal A site, and is coordinated by three catalytic aspartates (Asp330, Asp332, and Asp418). The primer terminal sugar and its 3′-OH are not in a catalytically relevant conformation (Fig. 3*A*); instead, the 3′-OH interacts indirectly with the magnesium ion in the metal A site, via a water molecule. Surprisingly, the enzyme has positioned the bound DNA substrate such that the 5′ unpaired base in the 2-nt gap (Fig. 2*A*, +1 position) occupies the nascent base-pair binding site. This structure suggests an alternative mechanism for generating frameshift deletions, in addition to those previously proposed (i.e., Streisinger slippage or nucleotide-stabilized misalignment) (16, 17).

Soaking the hPol μ Δ2 binary complex with nonhydrolyzable dUMPNPP generates a precatalytic ternary complex (Fig. 2*D*). There is clear density for the incoming nucleotide in the nascent base-pair binding site, which forms a correct base pair with the 5′ unpaired base (+1 position), rather than the 3′ unpaired template base closest to the primer terminus (Fig. 2*A*, 0 position). The primer terminal base correctly pairs with the template base at the −1 position (Fig. 3 *A* and *B*). The primer terminal sugar is shifted 2.2 Å toward the incoming nucleotide, which is aided by a concomitant 1.4 Å shift of the complementary template base toward the primer terminus. The primer terminal base maintains correct Watson–Crick base pairing with its complementary nucleotide through an ∼54° rotation of the base relative to the C1′–N9 bond, compared with the same base in the binary complex. The 3′-OH lies 3.6 Å from the α-phosphate of the dUMPNPP (Fig. 3 *A* and *B*), a distance consistent with that observed in the 1-nt gapped ternary complex (15). As in the binary complex, the primer terminal sugar adopts a C3′-*endo* sugar pucker and contributes to coordination of the ion in the metal A site. Metal A and metal B sites are both occupied by divalent magnesium ions with canonical octahedral geometry. The unpaired base (0 position) is accommodated within the active site, a position likely stabilized by stacking interactions with the flanking template-strand bases (Fig. 3*B*). The extent of such stabilization may change depending on the identity of the templating base and those of its neighbors. This structure provides an explanation for Pol μ's propensity to skip ahead, using the templating base closest to the 5′ end of the gap in a template-dependent manner.

Correct nucleotide incorporation was monitored by soaking the hPol μ Δ2 binary crystals with incoming dTTP (Fig. 2*E*). Clear electron density is observed for the newly formed bond. Some electron density is also present for the pyrophosphate leaving group, but not enough to allow fully accurate modeling.

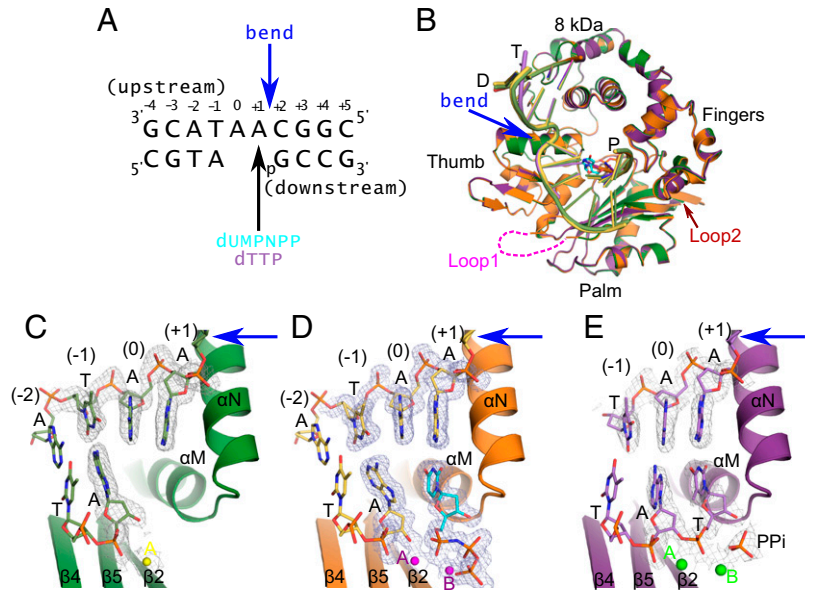

Fig. 2. Structural characterization of hPol μ Δ2 in complex with a 2-nt gapped DNA substrate. (*A*) Structure and sequence of the 2-nt gapped DNA substrate crystallized with hPol μ Δ2. The continuous template strand is drawn 5′→3′ (right to left), with the upstream primer (*Left*) and downstream primer (*Right*) beneath. The black arrow indicates the location of the incoming nucleotide insertion (nonhydrolyzable dUMPNPP in cyan; hydrolyzable dTTP in lavender). (*B*) Superposition of the binary (green; DNA in green), precatalytic ternary (orange; DNA in yellow; dUMPNPP in cyan), and postcatalytic nicked (purple; DNA in lavender) hPol μ Δ2 complexes. Locations of loop 1 (dashed magenta) and loop 2 (red) are noted. The DNA template, upstream primer, and downstream primer strands are annotated as T, P, and D, respectively. (*C*–*E*) Two-nucleotide gapped DNA substrate from *A* bound to hPol μ Δ2 in binary (*C*; green), precatalytic ternary (*D*; orange; incoming dUMPNPP in cyan), and postcatalytic (*E*; purple) complexes. The location of the bend in the DNA is emphasized by blue arrows. $2F_o - F_c$ electron density (contoured at 1σ) is shown for the DNA substrate near the catalytic center of each complex. Divalent magnesium (yellow in binary complex; purple in ternary complex) and monovalent sodium (green) ions are drawn as spheres. A partial structure of the inorganic pyrophosphate leaving group (PPi) remains in the active site of the postcatalytic nicked complex. All structural figures were generated using PyMOL (Schrödinger).

The metal A and B sites remain occupied, but the magnesium ions observed in the ternary complex structure were exchanged for sodium ions (Figs. 2*E* and 3*C*).

Minimal Structural Changes Occur as a Result of Nucleotide Incorporation. Interestingly, nucleotide incorporation in crystallo results in structural changes of the DNA substrate. The template strand becomes disordered past the −1 position. The upstream primer strand remains ordered, but the temperature factors increase from the newly incorporated nucleotide (32 $Å^2$) in the upstream direction (maximum temperature factor of 63 $Å^2$). The newly incorporated base remains paired with the template adenine in the +1 position, but the adenosine that was previously paired with the −1 position now incorrectly interacts with the template base at the 0 position (Fig. 3 *C*–*E*). This generates an A*–A(*syn*) mispair, similar to the Topal–Fresco model (18), where the template adenine at the 0 position is observed in the *anti* conformation and the primer adenine adopts a *syn* conformation (Fig. 3*F*). However, the geometries and distances between the bases are not consistent with optimal hydrogen bonding. Thus, this pairing is more likely stabilized by base stacking with flanking base pairs. The template base (T) at the −1 position makes no hydrogen-bonding interactions, and is out-of-register with the opposing primer-strand residues (Fig. 3 *C* and *D*).

Nucleotide incorporation is accompanied by disordering of loop 1 regions in immediate contact with the template strand (residues Gln364–Cys396 of β-strand 3, and the loop residues rejoining Glu386 of β-strand 4) (Fig. S1). Although there is currently no means of determining a causal relationship between the movement of loop 1 and that of the DNA template, it should be noted that no such movement occurs in loop 1 as a result of nucleotide incorporation for hPol μ Δ2 on a 1-nt gapped substrate (15). Movement of loop 1 in this particular context is likely to be a function of nucleotide incorporation, given that the position of the loop is not constrained by crystal-packing interactions in either the 1- or 2-nt gapped lattices, and might represent attempted translocation or dissociation.

Discussion

Comparing hPol μ Δ2 1- and 2-nt Gapped Complexes Yields Clues to Substrate Binding and Utilization. Superpositions of the 2-nt gapped complexes of hPol μ Δ2 show no large-scale protein domain movements throughout the catalytic cycle (rms deviations of less than 0.3 Å; Fig. 2*B*). This is consistent with the apparent rigidity of the global protein fold observed for hPol μ Δ2 during catalysis on a 1-nt gapped DNA substrate (15). Comparison of the hPol μ Δ2 complexes with 1- and 2-nt gapped substrates similarly demonstrates minimal structural differences in the overall protein fold (rms deviations of less than 0.4 Å; Table S3). In contrast, for the DNA, the upstream region of the template strand is substantially shifted from the location of this strand in the 1-nt gapped ternary complex (Fig. 4 *A* and *B*). The template nucleotides in the nascent base-pair binding site of each structure superimpose well (Fig. 4*B*), as do the nucleosides immediately upstream of that position (0 for the 2-nt gap, and −1 for the 1-nt gap). However, the phosphate backbone begins to deviate at the phosphate immediately 3′ of these nucleosides, with the extent of the shift increasing farther from the nascent base-pair binding site. That such a shift occurs in the 2-nt versus 1-nt gapped complexes is logical, because the polymerase must accommodate a longer template strand in the same physical space as for the shorter template strand. This compaction is primarily accomplished through a slight kink in the phosphate backbone between the −1 and −2 nucleotides (Fig. 4*B*), with the overall effect of this movement being a rotation of the upstream duplex, relative to the palm subdomain, while maintaining the pairing and position of the upstream primer strand (Fig. 4*A*).

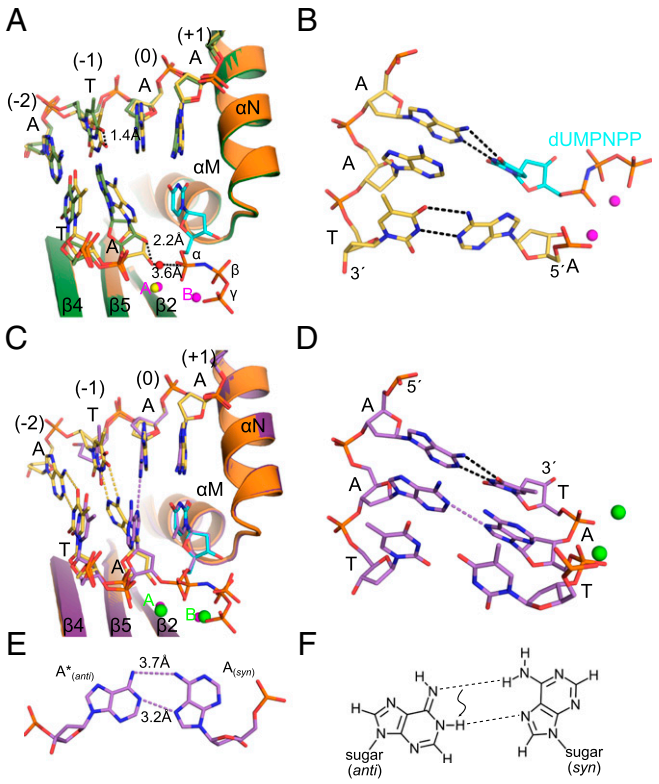

Fig. 3. Protein–DNA interactions throughout the 2-nt gap catalytic cycle. (*A*) Structural superposition of the hPol μ Δ2 binary (green; DNA in light green) and precatalytic ternary (orange; DNA in yellow; dUMPNPP in cyan) complexes. The sequence of the template (*Top*) and primer (*Bottom*) strands is noted and numbered as in Fig. 2*A*. Distances were calculated in PyMOL (Schrödinger). The positions of the divalent magnesium ions (yellow for binary; magenta for ternary) are shown as spheres. (*B*) Stick diagram of the 3′ (unpaired) and 5′ (paired) template-strand nucleotides in the precatalytic ternary complex. Putative hydrogen-bonding interactions are indicated by black dashed lines. (*C*) Superpositions of the hPol μ Δ2 precatalytic ternary (orange; DNA in yellow; dUMPNPP in cyan) and postcatalytic nicked (purple; DNA in lavender) complexes. Putative base-pairing interactions in the ternary complex are shown as yellow dashed lines, whereas the mispair interaction in the nicked complex is shown in lavender. Positions of the magnesium (magenta spheres) and sodium ions (green spheres) in the ternary and nicked complexes, respectively, are shown. (*D*) Stick diagram of the changes in base pairing that occur as a result of nucleotide incorporation. (*E* and *F*) Putative A*–A(*syn*) mispair at the −1 position (*E*), with the ideal hydrogen-bonding pattern drawn as a line diagram (*F*). The template-strand base at the −1 position remains in the *anti* conformation, whereas the primer-strand base flips to the *syn* conformation.

Movement of the DNA substrate is accompanied by an alternate configuration of loop 1, which adopts very different conformations when bound in a ternary complex with a 2-nt versus 1-nt gapped DNA substrate (Fig. 4*C*). Although loop 1 is not entirely ordered in the presence of either substrate, it clearly has a different trajectory, depending on the structure of the substrate encountered. Loop 1 is more ordered in the presence of the 2-nt gapped substrate and adopts a conformation that extends β-strands 3 and 4, increasing the area of the DNA-binding cleft in that region (Fig. 4*D*). Of the ordered residues of loop 1, only the side chain of His367 lies within putative hydrogen-bonding distance of the DNA substrate, at the phosphate between T9 and T10. The backbone carbonyls of Ser368 and Asp383 make water-mediated interactions with the phosphate of T9 and the sugar moiety of T8, respectively. These interactions are not present in the 1-nt gapped ternary complex structure. Additionally, Phe385 resides in a pocket between the thumb and palm subdomains, a position similar to that observed in the hPol μ Δ2 apoprotein conformation [PDB ID code 4LZD (15)]. Kinetic and biochemical assays with an hPol μ F385A mutant have suggested that F385A may help stabilize the position of loop 1, which could ultimately affect the ability of the enzyme to use DNA substrates with specific structures. The observed conformation of loop 1 "cradles" the altered conformation of the kinked template strand in the 2-nt gapped substrate using van der Waals interactions rather than direct electrostatic interactions with the phosphate backbone. However, it should be noted that although loop 1 may stabilize the DNA conformation adopted by this substrate, it is not essential, because loop 1 deletion does not abolish the template skipping-ahead behavior of Pol μ (16).

The crystal structures of hPol μ Δ2 in complex with a 2-nt gapped DNA substrate yield insights into the unique substrate specificity of this enzyme. In single-strand break repair, there is a distinct preference by the enzyme for smaller gap lengths (Fig. 1 and Table S1), which is consistent with the crystal structures bound to a 2-nt gap. Distortion of the template strand is required to accommodate the unpaired template base in the active site (Fig. 4 *A* and *B*), which could create a mechanical strain in that strand. This strain is then transmitted to the primer terminus, and is likely increased by the slight movement of this residue upon binding of the incoming nucleotide (Fig. 3*A*). Upon nucleotide incorporation, formation of the new phosphodiester bond results in an additional 1.6-Å shift of this residue, toward the newly incorporated nucleotide (Fig. 3*C*). This shift is nearly identical to that observed in the 1-nt gapped postcatalytic complex (15), and is likely a normal characteristic of incorporation by Pol μ, rather than being unique to a specific type of substrate. This shift may increase the structural strain in this residue, altering the base pairing from a correct base pair with the template base at the −1 position to a mispair with the template base at the 0 position (Fig. 3 *C*–*F*). Release of the strain upon catalysis appears to dramatically affect the stability of both the template strand upstream of that position and that of loop 1, because both regions become disordered in the crystal (Fig. S1).

In striking contrast to the gap-length preference exhibited by Pol μ in single-strand break repair, there appears to be no such preference in repair of noncomplementary double-strand DNA breaks. Recovery of 1- and 2-nt gapped DSBs repaired by Pol μ is nearly identical, the only difference being in repair fidelity (figures 1B and 3A in ref. 8, the companion article). For a 2-nt gap, insertion by Pol μ occurs in a template-dependent manner opposite the unpaired base on the 5′ end of the gap. Together with 1-nt gaps, this accounts for the majority of biologically relevant Pol μ substrates, because end structures aligned during cellular NHEJ to generate larger gaps were typically first remodeled by nuclease activity, before the polymerase could act (figure S4 in ref. 8, the companion article). Using this mechanism of skip-ahead synthesis would ultimately result in frameshift deletions at this site, but would also generate a substrate that could be efficiently repaired by DNA ligase IV (3). Using the available structural information on DNA binding by Pol μ, we can extrapolate the general appearance of such NHEJ synapses. NHEJ by Pol μ is generally preferred where the primer terminus is unpaired, which would prevent accumulation of the kind of template-strand distortion observed in the single-strand DNA break crystallized here (Fig. 4 *A* and *B*). Because Pol μ positions the unpaired template base closest to the 5′ end of the gap in the nascent base-pair binding site, gaps of longer than a single nucleotide would then have an overhang that is not physically connected to the duplex upstream of the break. Such an overhang could hypothetically be accommodated by the position of the thumb subdomain, with further stabilization from loop 1 (Fig. 5).

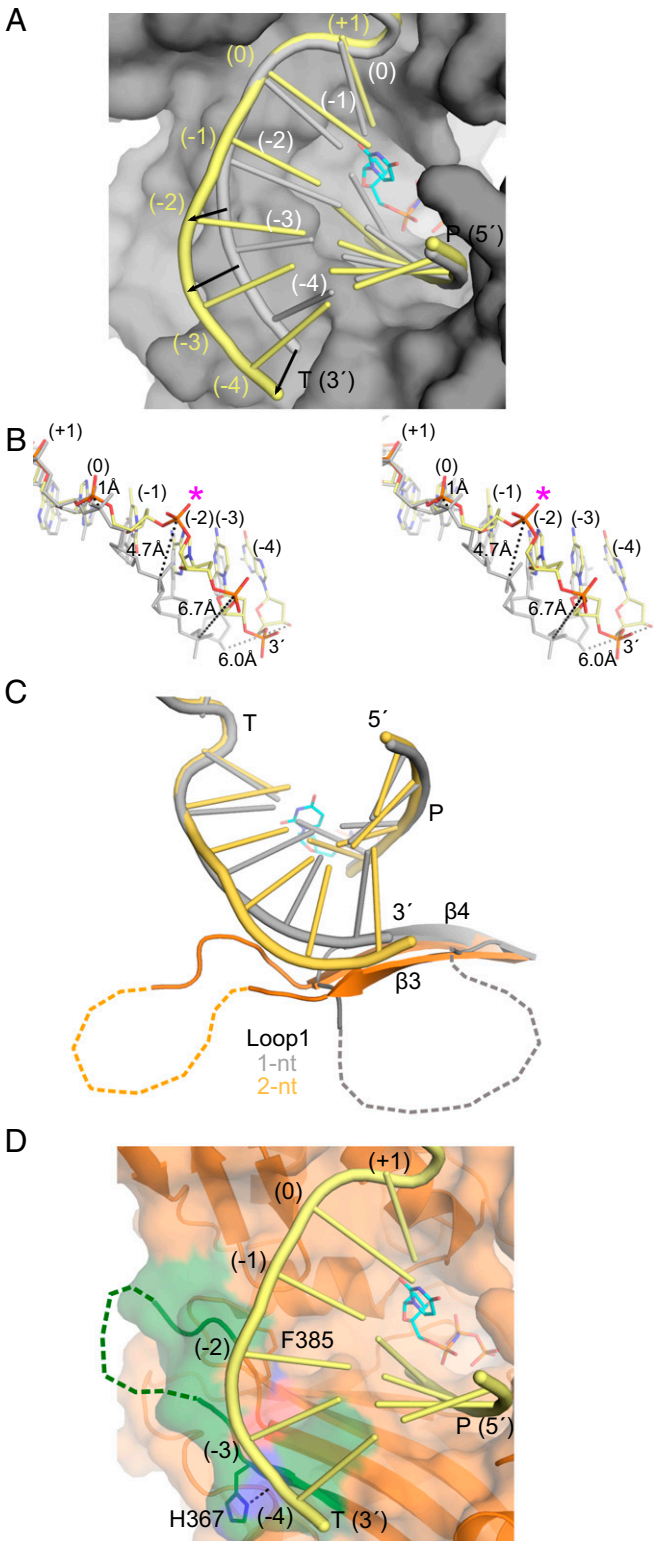

Fig. 4. Structural comparison of substrate binding by hPol μ Δ2 on 1- and 2-nt (yellow) gapped DNA substrates. (*A*) Surface representation of hPol μ Δ2 in ternary complex with a 1-nt gapped DNA substrate [light gray; PDB ID code 4M04 (15)]. DNA from the 2-nt gapped ternary complex (yellow) is superimposed to show the positional differences of the DNA template (T) and upstream primer (P) strand in each complex, as indicated by black arrows. Template-strand residues are numbered as in Fig. 2*A*. (*B*) Stereo diagram of the 1- (light gray) and 2-nt (yellow) gapped DNA substrates superimposed in *A*. Distances are indicated by black dashed lines. The location of the slight kink in the DNA backbone is emphasized by magenta asterisks. (*C*) Differences in loop 1 position in the 1- (gray; DNA in light gray) and 2-nt gapped (orange; DNA in yellow) ternary complexes. Hypothetical trajectories of the disordered regions are indicated by dashed lines. (*D*) Position of the 2-nt gapped DNA substrate (yellow) bound to hPol μ Δ2 (orange surface) in the ternary complex. In this structure, more residues of loop 1 (green) are ordered than in the 1-nt gapped structure. Hypothetical trajectory of disordered residues indicated by dashed green line.

Pol μ Engages Gapped DNA Substrates Differently from Other Family X Polymerases. Multiple structures exist for the related human Pol λ in the presence of single-strand break substrates (19–21). In each case, Pol λ positions the 3′ unpaired template base in the gap proximal to the nascent base-pair binding site using scrunching and strand-realignment mechanisms (Figs. 1*B* and 6 *A* and *B*). Superposition of hPol μ Δ2 in complex with a 2-nt gapped structure shows that the second unpaired base is clearly positioned for pairing with the incoming nucleotide, rather than being in the "scrunch" pocket between the thumb and the 8-kDa subdomains (Fig. 6*A*) (21), or extrahelical, as was observed in a frameshift intermediate complex (Fig. 6*B*) (20). More importantly, Pol λ cannot productively engage a 2-nt gapped substrate in the absence of an incoming nucleotide. In such a binary complex, the primer strand is located 1 nt out-of-register from the catalytically relevant conformation, and moves in concert with the template strand upon binding of the incoming nucleotide (Fig. S2).

Although Pol λ and Pol β are more active and more processive on gapped DNA substrates containing a 5′-phosphate on the downstream end of the gap (22–24), Pol μ is the first gap-filling polymerase, to our knowledge, to engage single- and double-strand break intermediate substrates such that the template base at the 5′ end of the gap directs incoming nucleotide incorporation. Given the strong dependence of Pol μ polymerase activity on 5′-phosphate–containing gapped DNA (11, 25), we postulate that interaction between the 8-kDa domain and the 5′-phosphate on the downstream end of the gap may represent a major driving force for substrate binding. Like other template-dependent family X polymerases, Pol μ enforces a bend in the DNA template strand immediately 3′ of the template base paired with the 5′-phosphorylated downstream primer nucleotide (5). Because Pol μ exhibits a high degree of structural rigidity and does not appear to require an open–closed transition to correctly position the DNA template strand (15, 26), the default position for the template base closest to the 5′-phosphate would then lie in the nascent base-pair binding site. Several interactions among residues in the fingers subdomain and the upstream primer assist in correctly positioning the primer terminus in or near a catalytically relevant conformation (26, 27). Unlike related family X polymerases Pol β and Pol λ, which use stronger electrostatic interactions with the phosphate backbone of the template strand to correctly assemble the active site for catalysis, Pol μ places less emphasis on the template strand, particularly in the region opposite the upstream primer (Figs. S3 and S4). This is consistent with the in vivo role of Pol μ in NHEJ of substrates containing discontinuous structure and noncomplementary sequence at the break (6–8).

Materials and Methods

Expression and Purification of Pol μ. Full-length wild-type hPol μ or hPol μ Δ2 constructs were expressed and purified as described (15). Purified full-length Pol μ was concentrated in 25 mM Tris (pH 8), 100 mM NaCl, 5% (vol/vol) glycerol, 1 mM DTT, 10 mM $MgCl_2$, and hPol μ Δ2 was concentrated in 25 mM Tris (pH 8), 80 mM NaCl, 5% (vol/vol) glycerol, 1 mM DTT.

Expression and Purification of Full-Length Human Pol λ. Human Pol λ was expressed and purified as described (28) and concentrated in 25 mM Tris (pH 7.5), 100 mM NaCl, 1 mM DTT, 10 mM $MgCl_2$.

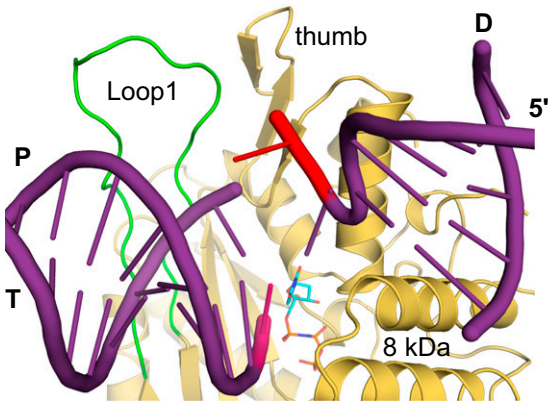

Fig. 5. Proposed model for accommodation of 2-nt gapped noncomplementary DSB substrates by Pol μ. The protein structure shown is that of the hPol μ Δ2 precatalytic ternary complex (khaki; incoming nucleotide in cyan). A possible loop 1 conformation (green) has been built manually, and the DNA-binding cleft contains a bound hypothetical DSB substrate (purple). In this structure, the unpaired template base closest to the 5′-phosphate on the downstream primer is positioned in the nascent base-pair binding site, correctly pairing with the incoming nucleotide, as in the 2-nt gapped structure. The primer terminus (magenta) is unpaired. The unpaired template base overhang (red) is accommodated by the thumb subdomain and a manually built model of loop 1.

Kinetic Analysis of Nucleotide Incorporation. The steady-state measurements of nucleotide incorporation were performed as described (15). Reaction conditions and quantitation are found in *Supporting Information*.

Crystallization of hPol μ Δ2 in Complex with a 2-nt Gapped DNA Substrate. DNA oligonucleotides used to create the binary complex were a template (5′-CGGCAATACG-3′), upstream primer (5′-CGTA-3′), and 5′-phosphorylated downstream primer (5′-pGCCG-3′). Oligonucleotides were mixed in equimolar ratios in 100 mM Tris (pH 7.5), 40 mM $MgCl_2$ and annealed in a thermal cycler by denaturation at 94 °C, followed by slow cooling to 4 °C. The annealed DNA was then mixed in a 4:1 molar ratio with hPol μ Δ2 (10.7 mg/mL) and incubated on ice at 4 °C for 1 h. Crystals of the binary complex were grown at room temperature by mixing 1 μL protein–DNA complex with 1 μL mother liquor [100 mM Tris, pH 8.5, 0.2 M $MgCl_2$, 20% (vol/vol) PEG8000], using the sitting-drop vapor-diffusion technique (29). Crystals were transferred to a cryoprotectant solution containing 0.1 M Tris (pH 8.5), 50 mM NaCl, 0.2 M $MgCl_2$, 20% (vol/vol) PEG8000, 5% (vol/vol) glycerol, and 10% (vol/vol) ethylene glycol. The precatalytic ternary and postcatalytic nicked complexes were generated by soaking binary complex crystals at room temperature in cryoprotectant solution containing 1.07 mM nonhydrolyzable analog dUMPNPP (overnight incubation) or 10.7 mM dTTP (7-min incubation), respectively. All crystals were flash-frozen in liquid nitrogen and placed into a stream of nitrogen gas cooled to –180 °C for data collection.

Data Collection and Structure Solution. All datasets were collected at a wavelength of 1.54 Å on an in-house rotating anode source, using a MicroMax 007 HF X-ray generator (Rigaku) and a Saturn 92 CCD detector (Rigaku). The data were indexed, integrated, and scaled using HKL2000 (30). All three DNA-bound crystal structures were observed in space group $P2_12_12_1$, with a single protein–DNA complex per asymmetric unit (Table S2). The structure of the binary complex was solved by molecular replacement [hPol μ Δ2 apoprotein starting model PDB ID code 4LZD (15)], using Phaser (31) in the PHENIX suite (32). The binary complex structure was then used as the starting model for refinement of the precatalytic ternary complex structure, which was in turn used as the starting model for the postcatalytic nicked complex. The same R_{free} test reflections were used for all datasets. All structures were refined using iterative cycles of manual model building and refinement in Coot (33, 34) and PHENIX (32). Data refinement statistics are listed in Table S2.

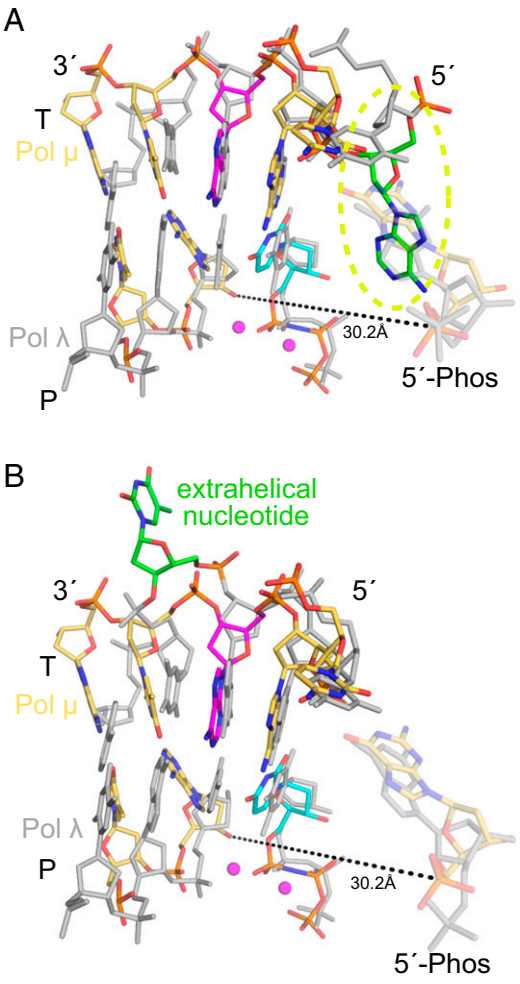

Fig. 6. Pol μ and Pol λ engage gapped 2-nt DNA substrates in different ways. (*A*) Superposition of the 2-nt gapped DNA substrates bound in ternary complex to hPol μ Δ2 (yellow) and Pol λ (gray) [PDB ID code 3HWT (21)]. For Pol λ, the 3′ unpaired template base is bound in the nascent base-pair binding site opposite the incoming nucleotide, and the 5′ unpaired template base (green) is positioned in the scrunch site (dashed yellow circle). In contrast, Pol μ positions the 5′ unpaired template base (magenta) in the nascent base-pair binding site, which directs binding of the incoming nucleotide (cyan). The 3′ template base (magenta) is unpaired. Position of the 5′-phosphate on the downstream primer is shown (transparent), with the distance to the primer terminal 3′-OH labeled. (*B*) Superposition of the 2-nt gapped DNA substrate bound to hPol μ Δ2 (yellow) and a frameshift intermediate (gray) bound to Pol λ [PDB ID code 2BCV (20)]. When bound to hPol μ Δ2, the 3′ unpaired template base (magenta) is accommodated in the active site, rather than being flipped extrahelically (green) as observed for Pol λ in a different sequence context. Position of the 5′-phosphate on the downstream primer is shown (transparent), with the distance to the primer terminal 3′-OH labeled.

ACKNOWLEDGMENTS. We thank R. Williams and W. Beard for critical reading of the manuscript. This research was supported by Division of Intramural Research of the National Institute of Environmental Health Sciences, National Institutes of Health Grants 1ZIA ES102645-03 (to L.C.P.) and Z01 ES065070 (to T.A.K.).

1. Lieber MR (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* 79:181–211.
2. Andersen SL, Sekelsky J (2010) Meiotic versus mitotic recombination: Two different routes for double-strand break repair: The different functions of meiotic versus mitotic DSB repair are reflected in different pathway usage and different outcomes. *BioEssays* 32(12):1058–1066.
3. Waters CA, et al. (2014) The fidelity of the ligation step determines how ends are resolved during nonhomologous end joining. *Nat Commun* 5:4286.
4. Lee JW, et al. (2004) Implication of DNA polymerase lambda in alignment-based gap filling for nonhomologous DNA end joining in human nuclear extracts. *J Biol Chem* 279(1):805–811.

5. Moon AF, et al. (2007) The X family portrait: Structural insights into biological functions of X family polymerases. *DNA Repair (Amst)* 6(12):1709–1725.
6. Nick McElhinny SA, Ramsden DA (2004) Sibling rivalry: Competition between Pol X family members in V(D)J recombination and general double strand break repair. *Immunol Rev* 200:156–164.
7. Nick McElhinny SA, et al. (2005) A gradient of template dependence defines distinct biological roles for family X polymerases in nonhomologous end joining. *Mol Cell* 19(3):357–366.
8. Pryor JM, et al. (2015) Essential role for polymerase specialization in cellular nonhomologous end joining. *Proc Natl Acad Sci USA* 112:E4537–E4545.
9. Fiala KA, Abdel-Gawad W, Suo Z (2004) Pre-steady-state kinetic studies of the fidelity and mechanism of polymerization catalyzed by truncated human DNA polymerase lambda. *Biochemistry* 43(21):6751–6762.
10. Roettger MP, Fiala KA, Sompalli S, Dong Y, Suo Z (2004) Pre-steady-state kinetic studies of the fidelity of human DNA polymerase mu. *Biochemistry* 43(43):13827–13838.
11. Davis BJ, Havener JM, Ramsden DA (2008) End-bridging is required for pol mu to efficiently promote repair of noncomplementary ends by nonhomologous end joining. *Nucleic Acids Res* 36(9):3085–3094.
12. Bebenek K, Garcia-Diaz M, Blanco L, Kunkel TA (2003) The frameshift infidelity of human DNA polymerase lambda. Implications for function. *J Biol Chem* 278(36): 34685–34690.
13. García-Díaz M, et al. (2002) DNA polymerase lambda, a novel DNA repair enzyme in human cells. *J Biol Chem* 277(15):13184–13191.
14. Brown JA, Pack LR, Sanman LE, Suo Z (2011) Efficiency and fidelity of human DNA polymerases λ and β during gap-filling DNA synthesis. *DNA Repair (Amst)* 10(1):24–33.
15. Moon AF, et al. (2014) Sustained active site rigidity during synthesis by human DNA polymerase μ. *Nat Struct Mol Biol* 21(3):253–260.
16. Juárez R, Ruiz JF, Nick McElhinny SA, Ramsden D, Blanco L (2006) A specific loop in human DNA polymerase mu allows switching between creative and DNA-instructed synthesis. *Nucleic Acids Res* 34(16):4572–4582.
17. Tippin B, Kobayashi S, Bertram JG, Goodman MF (2004) To slip or skip, visualizing frameshift mutation dynamics for error-prone DNA polymerases. *J Biol Chem* 279(44): 45360–45368.
18. Topal MD, Fresco JR (1976) Complementary base pairing and the origin of substitution mutations. *Nature* 263(5575):285–289.
19. Garcia-Diaz M, et al. (2004) A structural solution for the DNA polymerase lambda-dependent repair of DNA gaps with minimal homology. *Mol Cell* 13(4):561–572.
20. Garcia-Diaz M, Bebenek K, Krahn JM, Pedersen LC, Kunkel TA (2006) Structural analysis of strand misalignment during DNA synthesis by a human DNA polymerase. *Cell* 124(2):331–342.
21. Garcia-Diaz M, et al. (2009) Template strand scrunching during DNA gap repair synthesis by human polymerase lambda. *Nat Struct Mol Biol* 16(9):967–972.
22. Duym WW, Fiala KA, Bhatt N, Suo Z (2006) Kinetic effect of a downstream strand and its 5′-terminal moieties on single nucleotide gap-filling synthesis catalyzed by human DNA polymerase lambda. *J Biol Chem* 281(47):35649–35655.
23. Prasad R, Beard WA, Wilson SH (1994) Studies of gapped DNA substrate binding by mammalian DNA polymerase beta. Dependence on 5′-phosphate group. *J Biol Chem* 269(27):18096–18101.
24. Singhal RK, Wilson SH (1993) Short gap-filling synthesis by DNA polymerase beta is processive. *J Biol Chem* 268(21):15906–15911.
25. Nick McElhinny SA, Ramsden DA (2003) Polymerase mu is a DNA-directed DNA/RNA polymerase. *Mol Cell Biol* 23(7):2309–2315.
26. Moon AF, et al. (2007) Structural insight into the substrate specificity of DNA polymerase mu. *Nat Struct Mol Biol* 14(1):45–53.
27. Martin MJ, Juarez R, Blanco L (2012) DNA-binding determinants promoting NHEJ by human Polμ. *Nucleic Acids Res* 40(22):11389–11403.
28. Gosavi RA, Moon AF, Kunkel TA, Pedersen LC, Bebenek K (2012) The catalytic cycle for ribonucleotide incorporation by human DNA Pol λ. *Nucleic Acids Res* 40(15):7518–7527.
29. Chayen NE (1998) Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques. *Acta Crystallogr D Biol Crystallogr* 54(Pt 1):8–15.
30. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
31. McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr* 63(Pt 1):32–41.
32. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
33. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
34. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
35. Luscombe NM, Laskowski RA, Thornton JM (1997) NUCPLOT: A program to generate schematic diagrams of protein-nucleic acid interactions. *Nucleic Acids Res* 25(24): 4940–4945.
36. Batra VK, et al. (2006) Magnesium-induced assembly of a complete DNA polymerase catalytic complex. *Structure* 14(4):757–766.
37. Garcia-Diaz M, Bebenek K, Krahn JM, Pedersen LC, Kunkel TA (2007) Role of the catalytic metal during polymerization by DNA polymerase lambda. *DNA Repair (Amst)* 6(9):1333–1340.
38. Strong M, et al. (2006) Toward the structural genomics of complexes: Crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA* 103(21):8060–8065.