Recent Advances in Understanding the Structures of Translesion Synthesis DNA Polymerases

Justin A. Ling †, Zach Frevert † and M. Todd Washington *

Department of Biochemistry, University of Iowa College of Medicine, Iowa City, IA 52242-1109, USA; justin-ling@uiowa.edu (J.A.L.); zach-frevert@uiowa.edu (Z.F.)
* Correspondence: todd-washington@uiowa.edu; Tel.: +1-319-335-7518
† These authors contributed equally to this work.

Abstract: DNA damage in the template strand causes replication forks to stall because replicative DNA polymerases are unable to efficiently incorporate nucleotides opposite template DNA lesions. To overcome these replication blocks, cells are equipped with multiple translesion synthesis polymerases that have evolved specifically to incorporate nucleotides opposite DNA lesions. Over the past two decades, X-ray crystallography has provided a wealth of information about the structures and mechanisms of translesion synthesis polymerases. This approach, however, has been limited to ground state structures of these polymerases bound to DNA and nucleotide substrates. Three recent methodological developments have extended our understanding of the structures and mechanisms of these polymerases. These include time-lapse X-ray crystallography, which allows one to identify novel reaction intermediates; full-ensemble hybrid methods, which allow one to examine the conformational flexibility of the intrinsically disordered regions of proteins; and cryo-electron microscopy, which allows one to determine the high-resolution structures of larger protein complexes. In this article, we will discuss how these three methodological developments have added to our understanding of the structures and mechanisms of translesion synthesis polymerases.

Keywords: DNA replication; DNA polymerases; DNA repair; DNA damage; mutagenesis; genome stability

1. Introduction

DNA damage, which results from a variety of endogenous and exogenous sources, interferes with normal DNA replication and is a leading cause of genome instability. When the replication fork encounters DNA damage in the template strand, the replicative polymerases such as DNA polymerase delta (pol δ) and DNA polymerase epsilon (pol ε) stall. This is because these replicative polymerases are unable to efficiently incorporate nucleotides opposite the template DNA lesion. Thus, cells are equipped with multiple translesion synthesis (TLS) polymerases that have evolved specifically to incorporate nucleotides opposite DNA lesions [1–12]. While a variety of DNA polymerases are capable of carrying out TLS to a limited extent, five DNA polymerases appear to be the predominant TLS polymerases in eukaryotes. Among these are DNA polymerase eta (pol η), DNA polymerase iota (pol ι), DNA polymerase kappa (pol κ), Rev1, and DNA polymerase zeta (pol ζ). Pol η, pol ι, pol κ, and Rev1 are members of the Y-family of DNA polymerases, and pol ζ is a member of the B-family of DNA polymerases [13].

Depending on the type of DNA lesion at the stalled replication fork, TLS may require the catalytic function of either one TLS polymerase or two TLS polymerases [2]. In the case of TLS by one polymerase, the single TLS polymerase both inserts a nucleotide opposite the template lesion and extends from the resulting damaged primer–terminal base pair by incorporating subsequent nucleotides. In this capacity, the single TLS polymerase functions as both an “inserter” and an “extender”. An example of this is the bypass of 8-oxoguanine.
by pol η, which functions as both an inserter and an extender [14,15]. In the case of TLS by two polymerases, the first polymerase functions as an “inserter” by incorporating a nucleotide opposite the DNA lesion. The second polymerase functions as an “extender” by incorporating subsequent nucleotides. Only by the combined action of both polymerases can the DNA damage be bypassed. An example of this is the bypass of an abasic site by Rev1 (that functions as an inserter) and pol ζ (that functions as an extender) [16].

Each TLS polymerase that functions as an inserter has one or more cognate lesions opposite that it has evolved to incorporate nucleotides efficiently [1,6,7,11,17]. For example, the cognate lesions of pol η include thymine dimers and 8-oxoguanines [14,15,18–20]. The cognate lesions of pol ι include minor-groove purine adducts and exocyclic purine adducts [21,22]. The cognate lesions of pol κ include minor-groove guanine adducts and abasic sites [16,17,26,27].

X-ray crystallography has contributed enormously to our understanding of why certain types of DNA damage are cognate lesions for different polymerases. For example, X-ray crystallography has shown that pol η efficiently incorporates nucleotides opposite lesions such as thymine dimers and 8-oxoguanines by having a more spacious active site, which accommodates these geometrically distorted lesions [28–31]. This polymerase then utilizes Watson–Crick base pairing between the incoming nucleotide and the template base to direct nucleotide incorporation. X-ray crystallography has also shown that pol ι efficiently incorporates nucleotides opposite minor-groove purine adducts and exocyclic purine adducts by rotating the template purine base from the normal anti conformation to the syn conformation [22,32]. This polymerase then utilizes Hoogsteen base pairing between the incoming nucleotide and the template to direct incorporation. Crystallography has also shown that Rev1 efficiently incorporates nucleotides opposite minor-groove guanine adducts and abasic sites using a remarkable mechanism [33–35]. The template base is evicted from the double helix and an arginine side chain acts as the template for incorporation of dCTP.

While these past structural studies have provided a wealth of information about the mechanism of TLS polymerases, they have been limited to ground state structures of the polymerase domains. Three recent methodological developments have extended our understanding of the structures of these polymerases. These include time-lapse X-ray crystallography, which allows one to identify novel reaction intermediates; full-ensemble hybrid methods, which allow one to combine molecular simulations with small-angle X-ray scattering (SAXS) data; and cryo-electron microscopy (cryo-EM), which allows one to determine the high-resolution structures of large protein complexes. In this article, we will discuss how these three methodological developments have added to our understanding of the structures and mechanisms of TLS polymerases.

2. Time-Lapse X-ray Crystallography

While traditional X-ray crystallography has provided a wealth of information about the mechanisms of TLS polymerases, these structures have been limited to the protein alone, the binary complex formed by the protein and the DNA substrate, and the ternary complex formed by the protein, the DNA substrate, and the incoming dNTP substrate. What these studies have not been able to provide are structures of complexes formed by the protein, the DNA product, and the pyrophosphate product. These studies have also not provided the structures of any intermediates formed during the nucleotide-incorporation reaction. These limitations have been overcome recently by the use of time-lapse crystallography.

In time-lapse crystallography, a ternary complex of the protein, the DNA substrate, and the incoming dNTP substrate are crystallized in the presence of Ca$^{2+}$ [36–39]. The Ca$^{2+}$ metal ions allow for the formation of this substrate ternary complex without allowing the reaction to proceed. Protein crystals are harvested and transferred to a solution containing Mg$^{2+}$. The Mg$^{2+}$ metal ions diffuse into the protein crystal and replace the Ca$^{2+}$ ions thereby initiating the reaction. After various reaction times, often ranging from approximately 30 s
to several minutes, the crystals are flash frozen in liquid nitrogen. X-ray diffraction data are then collected with the frozen crystals and the X-ray crystal structure is determined. This yields a series of structural snapshots at different stages of the nucleotide-incorporation reaction, and this has led to the discovery of several important and unexpected reaction intermediates. These new findings will be discussed below.

2.1. DNA Polymerase Eta

Time-lapse X-ray crystallography was first applied to the catalytic core of human pol $\eta$ [38] to observe the incorporation of an incoming dATP opposite a non-damaged thymine template (Figure 1). X-ray crystal structures were determined after various reaction times ranging from 40 to 300 s. The structure after 40 s of reaction time showed the ternary substrate complex of pol $\eta$, DNA, and dNTP with no evidence of product formation. The structure after 80 s of reaction time shows that some of the substrate has been converted to product. The structure after 230 s of reaction time shows that approximately 60% of the substrate has been converted to product and the other 40% remains substrate. Overall, the reaction proceeds 20 to 100-fold slower in the crystal than it does in solution. This is likely due to differences in the crystal and solution environments such as reduced thermal motions.

These structural snapshots show that the incoming nucleotide and the template base maintain Watson–Crick base pairing during the course of the reaction. Moreover, the overall global structure of the protein and the DNA does not change during the nucleotide incorporation reaction. However, notable local changes within the active site are visible. For example, the sugar pucker of the incoming nucleotide changes from 2′-endo in the substrate complex to 3′-endo in the product complex. This makes the minor groove shallower and causes the double-stranded region of the DNA substrate to adopt an A form conformation once the nucleotide has been incorporated.

The most striking feature observed during the formation of product is the appearance of a third Mg$^{2+}$ ion that coordinates the non-bridging oxygens of the $\alpha$- and $\beta$-phosphates. This is an important discovery because up to this time, polymerases were believed to utilize
a two-metal ion mechanism. The presence of this third metal ion has directly challenged this view. The authors suggest that this third metal ion may stabilize the accumulating charge during the chemical step of phosphodiester bond formation and that the binding of the third metal ion occurs prior to and is necessary for the chemical step. It should be noted that a third metal ion was also found to be present in the active site of the DNA repair polymerase pol β [37].

Controversy about the role of this third metal ion intermediate persists in the field, because the kinetics of the appearance of this ion in the time-lapse crystallography is identical to the kinetics of phosphodiester bond formation. Thus, based on these structural studies alone, it is not possible to differentiate between the two mechanisms. In the first possible mechanism, the third metal ion binds first, which leads to an immediate phosphodiester bond formation. In such a mechanism, favored by the authors of the pol η study [38], the third metal ion would play an essential role in catalyzing nucleotide incorporation. In the second possible mechanism, the phosphodiester bond is formed first, which leads to an immediate third metal ion binding. In such a mechanism, supported by computational and crystallographic work with pol β [40], binding the third metal ion does not facilitate nucleotide incorporation. In fact, a recent comprehensive review of the literature has argued that the binding of the third metal ion prior to nucleotide incorporation actually inhibits the catalytic step [41]. More work is needed to resolve this important controversy in our mechanistic understanding of DNA polymerases.

2.2. Rev1

Time-lapse X-ray crystallography was recently applied to the catalytic core of yeast Rev1 [39] to observe the incorporation of an incoming dCTP opposite a non-damaged guanine template (Figure 2). X-ray crystal structures were determined after various reaction times ranging from 5 to 30 min. The structure after 5 min of reaction time shows that approximately 30% of the substrate has been converted to product and the other 70% remains as substrate. The structure after the 10 min reaction time shows that all of the substrate has been converted to product. Like pol η, the reaction proceeds significantly slower in the crystal than it does in solution. Unlike pol η, the structural snapshots of Rev1 provided no evidence for the existence of a third metal ion.

The structural snapshots of Rev1 show that the guanine template remains extrahelical and binds within the template-binding pocket in the active site during the entire nucleotide-incorporation reaction. It does not leave this pocket after the reaction is complete. In the structure of the substrate complex, the incoming dNTP binds to the sidechain of arginine-324. The atoms of the guanidino group of the arginine side chain and the atoms of the cytosine ring are all within the same plane allowing for the formation of two hydrogen bonds between the arginine side chain and the cytosine base. In the structure of the product complex, the incorporated cytosine shifts nearly 2 Å toward its neighboring base in the primer strand. This results in the loss of planarity between the arginine side chain and the cytosine base and a weakening of the two hydrogen bonds between them.

The most striking feature observed during the nucleotide-incorporation reaction is the presence of two inorganic phosphates in the active site instead of the usual pyrophosphate product. These unexpected phosphate intermediates are not associated with any metal ions. The presence of these phosphate ions imply that Rev1 possesses a novel pyrophosphatase activity, which cleaves the pyrophosphate product into two phosphates immediately after nucleotide incorporation. Presumably this activity thermodynamically drives the nucleotide-incorporation reaction forward by the removal of the pyrophosphate product. Moreover, the production of inorganic phosphate by Rev1 was also shown to occur during nucleotide incorporation in solution.

X-ray crystal structures of Rev1 were determined after 15 and 30 min to better understand the fate of the two phosphate ion products. The structure after the 15 min reaction time showed that the β phosphate was released while the γ phosphate remained bound in the active site. The structure after the 30 min reaction time showed that the γ phosphate
was also released. These structures show an ordered mechanism of phosphate ion release with the β phosphate releasing before the γ phosphate. Several unresolved issues remain regarding the mechanism of Rev1’s novel pyrophosphatase activity. For example, the identity of the general base that activates water for a nucleophilic attack on the pyrophosphate has yet to be determined. In addition, is the pyrophosphatase reaction necessary for efficient nucleotide incorporation? More work is needed to address these important questions.

![Figure 2](image_url)  
**Figure 2.** Time-lapse X-ray crystallography of the Rev1 catalytic core. Shown are the overall structure of the Rev1 product complex bound to DNA (a) and a close-up of the active site with the positions of the two inorganic phosphate products ($P_i$) and a key amino acid residue (R324) highlighted (b). For the products, carbon atoms are shown in yellow, nitrogen atoms are shown in blue, oxygen atoms are shown in red, and phosphorus atoms are shown in orange. This figure was generated with 6X72.pdb [39].

3. Full-Ensemble Hybrid Methods

SAXS is an experimental technique in which one places a solution of a protein in an X-ray beam and measures the continuous pattern of scattered X-rays as a function of distance from the center of the detector. The analysis of this scattering data, which have been described elsewhere [42–49], allows one to determine the radius of gyration of the protein ($R_g$), the maximal distance between pairs of atoms in the protein ($D_{max}$), and the distribution of distances between every pairwise combination of atoms in the protein. This latter information is depicted in the pairwise ($P(r)$) distribution plot. In the case of highly structured proteins with little conformational flexibility, the information in the $P(r)$ plot can be used for ab initio shape predictions, which provide the overall shape of the protein at a medium to low resolution (about 10 to 20 Å).

In the case of proteins with a large amount of intrinsic disorder or a high degree of conformational flexibility, SAXS data are typically analyzed by carrying out a minimal ensemble search. A common strategy for performing minimal ensemble searches is BILBOMD [50]. This entails running a molecular dynamics simulation where the conformations of individual domains are largely modeled as rigid bodies with flexible regions between them, allowing these domains to adopt a variety of orientations relative to one another. These molecular dynamics simulations typically utilize timesteps of 1 to 2 fs with a structure recorded every 0.5 ps. Thus, a 5 ns-simulation will yield a pool of about 10,000 structures. BILBOMD then uses a genetic algorithm to find the minimum ensemble (usually a weighted set of 2 to 5 structures from the larger pool) that best fits the data. By seeking the minimal ensemble that best accounts for the experimental SAXS data, one can attempt to avoid the overfitting problem that is often encountered in SAXS analyses.
The primary drawback of minimal-ensemble methods, which represent conformationally flexible proteins using a weighted set of only 2 to 5 static structures, is that it results in highly unrealistic models of such proteins. The vast number of the intermediate states between these few static structures are not represented in the final ensemble. Recently, a full-ensemble method has been developed to overcome this problem [49,51]. By using coarse-grained Langevin dynamics simulations, one can generate a full ensemble of tens of thousands of structures. Instead of searching for a minimal ensemble that best fits the experimental data, the entire ensemble is directly compared to the SAXS experimental data. The $R_g$ and $D_{\text{max}}$ values obtained from the entire ensemble are directly compared to the experimentally determined $R_g$ and $D_{\text{max}}$ values. Moreover, the $P(r)$ curve obtained from the full ensemble is directly compared to the experimentally determined $P(r)$ curve. This allows one to overcome the overfitting problem by not relying on curve fitting at all. In addition, this method, which represents conformationally flexible proteins by an ensemble of over ten thousand structures, results in more realistic models of highly flexible proteins.

Yeast pol $\eta$ was predicted to possess a folded catalytic core (residues 1 to 510) as well as a C-terminal intrinsically ordered region (residues 510 to 632), which contains several $\alpha$-helices and a small, folded ubiquitin-binding, zinc-binding (UBZ) motif (residues 550 to 580) [52] and two distinct proliferating cell nuclear antigen (PCNA)-interacting protein (PIP) motifs [53]. Full-ensemble hybrid methods were used to test this prediction and to assess the conformational flexibility of the C-terminal region of pol $\eta$ [51] (Figure 3). Langevin dynamics simulations were performed using a coarse-grained model of full-length pol $\eta$, whereby each amino acid residue was represented by one to four coarse-grained beads. Timesteps of 125 fs were used, and a structure was recorded every 1 ns over the duration of a 10-μs simulation generating an ensemble of 10,000 structures. $R_g$ and $D_{\text{max}}$ values were obtained for the entire ensemble, and these theoretical values were directly compared to those obtained using SAXS. The theoretical and experimental $R_g$ and $D_{\text{max}}$ values and the theoretical and experimental $P(r)$ curves were in close agreement.

**Figure 3.** Full-ensemble hybrid methods with full-length pol $\eta$. Shown is an overlay of ten randomly selected structures from the full ensemble generated for pol $\eta$ as described [51]. The catalytic core and the DNA of these ten structures are superimposed, and the intrinsically disordered C-terminal regions of these ten structures are depicted in different colors as dashed lines.
Analysis of this experimentally validated ensemble showed that the C-terminal region of pol η is indeed mostly unstructured and samples a wide range of conformational space. Approximately 5% of the time, the C-terminal region is compact and remains near the catalytic core. In another approximate 5% of the time, the C-terminal region is highly extended, stretching far from the catalytic core. The remaining 90% of the time, the C-terminal region occupies a wide range of conformations in between these two extremes. These findings support a model in which the C-terminal region acts as a flexible tether linking the catalytic core of pol η to ubiquitin-modified PCNA, a core component of bypass complexes. As pol ι, pol κ, and Rev1 also possess similar disordered C-terminal regions [10,54], this tethering likely facilitates polymerase sampling, which would be needed to ensure that the most appropriate TLS polymerase is selected to bypass a given type of DNA damage in accordance with the kinetic selection model, which is described elsewhere [11].

4. Cryo-Electron Microscopy

In recent years, cryo-EM has emerged as a major method for obtaining high-resolution structural information of large protein complexes. Compared to X-ray crystallography, cryo-EM requires small amounts of purified protein. These complexes are frozen in a thin layer of ice, and cryo-EM images are taken. Images of particles are averaged and 2D class averages are obtained. Ultimately, a 3D reconstitution of the electron density of the complex is determined. Two cryo-EM studies of TLS polymerases have led to significant advances in our understanding of the mechanisms of these proteins. These two studies will be discussed below.

4.1. DNA Polymerase Zeta

In many cases, TLS requires the catalytic function of two polymerases [2]. The first polymerase functions as an “inserter” to catalyze the incorporation of a nucleotide opposite the template lesion, and the second polymerase functions as an “extender” to catalyze the incorporation of an additional nucleotide following the resultant primer–terminal base pair with the lesion in the template strand. Pol ζ functions as an “extender” during the bypass of a wide range of DNA lesions [16,55]. Pol ζ is a member of the B-family of DNA polymerases along with replicative polymerases pol α, pol δ, and pol ε [13]. Yeast pol ζ was initially believed to be comprised of four subunits: the Rev3 catalytic subunit and the Rev7, Pol31, and Pol32 non-catalytic subunits [56–58]. The Rev3 subunit is large (173 kDa), while the Rev7, Pol31, and Pol32 subunits are small (29 kDa, 55 kDa, and 40 kDa, respectively). Interestingly, the Pol31 and Pol32 subunits are also subunits of replicative pol δ. Attempts to determine the X-ray crystal structure of the entire pol ζ protein have been unsuccessful in part due to the difficulty of obtaining sufficient quantities of the protein for crystallization. Instead, X-ray crystal structures have only been determined for small portions of pol ζ, such as the Rev7 non-catalytic subunit bound to a peptide containing 28 amino acid residues of the Rev3 catalytic subunit [59].

Cryo-EM was recently used to determine the structures of the entire yeast pol ζ protein alone and bound to a DNA substrate and an incoming dNTP [60] (Figure 4). The most surprising feature of these structures is that pol ζ contains two Rev7 subunits, denoted Rev7A and Rev7B, rather than the expected one Rev7 subunit. These five subunits—Rev3, Rev7A, Rev7B, Pol31, and Pol32—form a pentameric ring-like structure. The two Rev7 subunits bind to an extended, disordered region of the Rev3 subunit, and the DNA substrate binds exclusively to the Rev3 catalytic subunit.
The Rev7\textsubscript{A} subunit binds the extended loop of the Rev3 subunit and binds the Rev7 subunit. The two Rev7 subunits are shown in green and light green. The Pol3\textsubscript{1} subunit is shown in purple, and the Pol3\textsubscript{2} subunit is shown in pink. Shown are a top view (a) and a side view (b) of pol\(\zeta\). The catalytic Rev3 subunit binds the DNA substrate. The Pol3\textsubscript{1} subunit binds to Rev3 and to the non-catalytic Pol3\textsubscript{2} subunit, thereby closing the ring-like structure.

The overall structure of pol\(\zeta\) resembles the structure of pol\(\delta\), which was also determined by cryo-EM [61] (Figure 4). In the case of pol\(\delta\), the catalytic Pol3 subunits bind the DNA substrate. The non-catalytic Pol3\textsubscript{1} subunit binds to Pol3 and to the non-catalytic Pol3\textsubscript{2} subunit, but Pol 3 and Pol3\textsubscript{2} do not directly interact with each other. In the case of pol\(\zeta\), the catalytic Rev3 subunit binds the DNA substrate. The Pol3\textsubscript{1} subunit binds to Rev3 and to the Pol3\textsubscript{2} subunit, but Rev3 and Pol3\textsubscript{2} do not directly interact with each other. The Rev7\textsubscript{A} subunit binds the extended loop of the Rev3 subunit and binds the Rev7\textsubscript{B} subunit. The Rev7\textsubscript{B} subunit binds the extended loop of the Rev3 subunit and binds the Rev7\textsubscript{A} and Pol3\textsubscript{2} subunit, thereby closing the ring-like structure.

The structure of the catalytic Rev3 subunit of pol\(\zeta\) resembles the structure of the Pol3 subunit of pol\(\delta\). These subunits have two domains: an N-terminal exonuclease domain and a C-terminal polymerase domain. The C-terminal polymerase domains contain three subdomains found in most DNA polymerases: palm, fingers, and thumb. In both pol\(\zeta\) and pol\(\delta\), the palm subdomain contains several conserved acid amino acid residues that coordinate the metal ions necessary for catalyzing the nucleotide incorporation reaction. In

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4}
\caption{Cryo-EM structure of the full pol\(\zeta\) complex. Shown are a top view (a) and a side view (b) of pol\(\zeta\). The polymerase domain of the Rev3 subunit is shown in blue, and the exonuclease domain of the Rev3 subunit is shown in light blue. The two Rev7 subunits are shown in green and light green. The Pol3\textsubscript{1} subunit is shown in purple, and the Pol3\textsubscript{2} subunit is shown in pink. Shown are a top view (c) and a side view (d) of pol\(\delta\). The catalytic Pol3 subunit is shown in blue, the Pol3\textsubscript{1} subunit is shown in purple, and the Pol3\textsubscript{2} subunit is shown in pink. This figure was generated with 7S0T.pdb [60] and 6P1H.pdb [61].}
\end{figure}
both proteins, the fingers subdomain contacts the nascent base-pair ensuring that nucleotide incorporation occurs with high fidelity. In the case of pol δ, the N-terminal exonuclease domain contains conserved acid amino acids that coordinate the catalytically essential metal ions. In the case of pol ζ, the N-terminal exonuclease domain is not active, which contributes to its ability to carry out the extension step of TLS.

The most important difference observed between the structures of the catalytic domains of pol ζ and pol δ is the conformation of the extended loop connecting the N-terminal exonuclease domains and the C-terminal polymerase domains. In the case of pol δ, this loop contacts the primer–terminal base pair. These contacts are likely the structural basis for the inability of pol δ to extend from aberrant primer–terminal base pairs containing mismatches or template DNA lesions. In the case of pol ζ, this loop is in a different conformation and does not contact the primer–terminal base pair. This lack of contacts is likely the structural basis for the remarkable ability of pol ζ to extend from primer–terminal base pairs containing mismatches or template lesions. Thus, it is likely that this structural difference between these two proteins is what allows pol ζ to fulfill its role as an “extender” during TLS.

4.2. DNA Polymerase Kappa and PCNA

Pol κ has two functions in TLS. First, it functions as an “inserter” to incorporate nucleotides opposite minor-groove purine adducts [23–25]. Second, it functions as an “extender” during the bypass of a wide range of DNA lesions [62–64]. Pol κ is a member of the Y-family of DNA polymerases. Multiple X-ray crystal structures of human pol κ have been solved [64–66]. These structures show that the polymerase domain has the usual fingers, thumb, and palms subdomains. Like pol η, the incoming nucleotide and the templating base form Watson–Crick base pairs. Unlike pol η, however, the active site of pol κ is too small to accommodate geometrically distorting DNA lesions. Instead, there is a pocket between the fingers and little fingers subdomains into which minor groove adducts can be accommodated [66]. This is likely the structural basis for pol κ functioning as an “inserter”. In addition, pol κ contains a region called the “N clasp”, which is not found in other Y-family polymerases. The “N clasp” allows pol κ to completely encircle the DNA substrate, which may increase the processivity of pol κ on aberrant primer termini thereby providing at least a partial structural explanation of its “extender” function [64].

While X-ray crystallography has focused on the polymerase domain of pol κ, little effort has been devoted to studies of the C-terminal unstructured region of the protein, which is approximately 300 amino acid residues in length. This unstructured region contains motifs for binding the trimeric DNA sliding clamp PCNA, ubiquitin, and Rev1 [10,54]. Recently, the cryo-EM structures of full-length pol κ bound to DNA and PCNA and to DNA and ubiquitin-modified PCNA (Ub-PCNA), a post-translationally modified form of PCNA that functions in TLS, have been determined [67] (Figure 5). These structures both show that pol κ binds to the DNA substrate on the front face of the PCNA ring. In the structure of pol κ bound to DNA and unmodified PCNA, the C-terminal unstructured region of pol κ is not visible, except for the PIP motif, an eight amino acid motif that binds within a hydrophobic pocket on the front face of the PCNA ring.

In the structure of pol κ bound to DNA and Ub-PCNA, little electron density is visible for the ubiquitin moieties attached to the two PCNA subunits that are not directly binding the pol κ. This suggests that these ubiquitin moieties are flexible and can occupy a large number of conformations. The ubiquitin moiety attached to the PCNA subunit that directly binds pol κ appears to have a stronger density. This suggests that pol κ limits the conformational flexibility of this ubiquitin moiety. Interestingly there is no electron density for the UBZ motif of pol κ, which supposedly binds the ubiquitin moiety on Ub-PCNA. The authors suggest the lack of density for the UBZ motif is likely due to the transient nature of the interaction between the UBZ and ubiquitin.
5. Conclusions

In this review, we discussed several recent approaches to improving our understanding of the structures and mechanisms of TLS polymerases. Time-lapse X-ray crystallography has led to the identification of new features of the nucleotide incorporation reaction, including the discovery of a third metal ion in pol η [38] and a novel pyrophosphatase activity in Rev1 [39]. Full-ensemble hybrid methods have led to better description of the conformational flexibility in the intrinsically disordered region of pol η [51]. Cryo-EM has led to the determination of structures of pol ζ [60] and the complex of pol κ on DNA and PCNA [67]. The use of these new approaches is in its infancy, and much more structural and mechanistic information will undoubtedly be forthcoming in the years ahead.

In addition to these studies of TLS polymerases, similar studies are being carried out with classical DNA polymerases—those involved in normal DNA replication and repair. For example, time-lapse crystallography has been applied to pol β, a classical polymerase involved in base excision repair [36,37]. Similarly, cryo-EM has been used to determine the structure of pol δ [61] and pol ε [68]. Carefully comparing and contrasting the structures and mechanisms of classical polymerases and TLS polymerases will lead to an integrated understanding of how DNA polymerases synthesize DNA—a process necessary for the maintenance of genome stability.
Author Contributions: Conceptualization, J.A.L., Z.F. and M.T.W. Writing—Original draft preparation, J.A.L., Z.F. and M.T.W. Writing—Review and editing, J.A.L., Z.F. and M.T.W. Supervision, M.T.W. Funding Acquisition, M.T.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by award GM081433 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

Acknowledgments: We thank Christine Kondratick, Brittany Ripley, Melissa Gildenberg, Devin Reusch, Lynne Dieckman, and Bret Freudenthal for discussions.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

Cryo-EM: cryo-electron microscopy; D_{max}: maximum distance; PCNA: proliferating cell nuclear antigen; PIP: PCNA-interacting protein; Pol δ: DNA polymerase delta; Pol ε: DNA polymerase epsilon; Pol η: DNA polymerase eta; Pol κ: DNA polymerase kappa; Pol ζ: DNA polymerase zeta; P(r): pairwise distribution; R_g: radius of gyration; SAAS: small-angle X-ray scattering; TLS: translesion synthesis; Ub-PCNA: ubiquitin-modified PCNA; UBZ: ubiquitin-binding/zinc-binding.

References

1. Friedberg, E.C.; Wagner, R.; Radman, M. Specialized DNA polymerases, cellular survival, and the genesis of mutations. Science 2002, 296, 1627–1630. [CrossRef] [PubMed]
2. Prakash, S.; Prakash, L. Translesion DNA synthesis in eukaryotes: A one- or two-polymerase affair. Genes Dev. 2002, 16, 1872–1883. [CrossRef] [PubMed]
3. Lehmann, A.R. Replication of damaged DNA. Cell Cycle 2003, 2, 300–302. [CrossRef] [PubMed]
4. Prakash, S.; Johnson, R.E.; Prakash, L. Eukaryotic translesion synthesis DNA polymerases: Specificity of structure and function. Annu. Rev. Biochem. 2005, 74, 317–353. [CrossRef]
5. Lehmann, A.R. Replication of damaged DNA by translesion synthesis in human cells. FEBS Lett. 2005, 579, 873–876. [CrossRef]
6. Lehmann, A.R.; Niimi, A.; Ogi, T.; Brown, S.; Sabbioneda, S.; Wing, J.F.; Kannouche, P.L.; Green, C.M. Translesion synthesis: Y-family polymerases and the polymerase switch. DNA Repair 2007, 6, 891–899. [CrossRef]
7. Waters, L.S.; Minesinger, B.K.; Wiltrout, M.E.; D’Souza, S.; Woodruff, R.V.; Walker, G.C. Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. Microbiol. Mol. Biol. Rev. 2009, 73, 134–154. [CrossRef]
8. Washington, M.T.; Carlson, K.D.; Freudenthal, B.D.; Pryor, J.M. Variations on a theme: Eukaryotic Y-family DNA polymerases. Biochim. Biophys. Acta 2010, 1804, 1113–1123. [CrossRef]
9. Sale, J.E.; Lehmann, A.R.; Woodgate, R. Y-family DNA polymerases and their role in tolerance of cellular DNA damage. Nat. Rev. Mol. Cell Biol. 2012, 13, 141–152. [CrossRef]
10. Pryor, J.M.; Dieckman, L.M.; Boehm, E.M.; Washington, M.T. Eukaryotic Y-Family Polymerases: A Biochemical and Structural Perspective. In Nucleic Acid Polymerases; Murakami, K.S., Trakselis, M.A., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; Volume 30, pp. 85–108.
11. Powers, K.T.; Washington, M.T. Eukaryotic translesion synthesis: Choosing the right tool for the job. DNA Repair 2018, 71, 127–134. [CrossRef]
12. Marians, K.J. Lesion Bypass and the Reactivation of Stalled Replication Forks. Annu. Rev. Biochem. 2018, 87, 217–238. [CrossRef] [PubMed]
13. Burgers, P.M.; Koonin, E.V.; Bruford, E.; Blanco, L.; Burtis, K.C.; Christman, M.F.; Copeland, W.C.; Friedberg, E.C.; Hanaoka, F.; Hinkle, D.C.; et al. Eukaryotic DNA polymerases: Proposal for a revised nomenclature. J. Biol. Chem. 2001, 276, 43487–43490. [CrossRef] [PubMed]
14. Haracska, L.; Yu, S.L.; Johnson, R.E.; Prakash, L.; Prakash, S. Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase eta. Nat. Genet. 2000, 25, 458–461. [CrossRef] [PubMed]
15. Carlson, K.D.; Washington, M.T. Mechanism of efficient and accurate nucleotide incorporation opposite 7,8-dihydro-8-oxoguanine by Saccharomyces cerevisiae DNA polymerase eta. Mol. Cell Biol. 2005, 25, 2169–2176. [CrossRef] [PubMed]
16. Haracska, L.; Unk, L.; Johnson, R.E.; Johansson, E.; Burgers, P.M.; Prakash, S.; Prakash, L. Roles of yeast DNA polymerases delta and zeta and of Rev1 in the bypass of abasic sites. Genes Dev. 2001, 15, 945–954. [CrossRef] [PubMed]
17. Pryor, J.M.; Washington, M.T. Pre-steady state kinetic studies show that an abasic site is a cognate lesion for the yeast Rev1 protein. DNA Repair 2011, 10, 1138–1144. [CrossRef]
18. Johnson, R.E.; Prakash, S.; Prakash, L. Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase. Pol. **Sci. 1999**, *283*, 1001–1004. [CrossRef]

19. Washington, M.T.; Johnson, R.E.; Prakash, S.; Prakash, L. Accuracy of thymine-thymine dimer bypass by Saccharomyces cerevisiae DNA polymerase eta. **Proc. Natl. Acad. Sci. USA 2000**, *97*, 3094–3099. [CrossRef]

20. Washington, M.T.; Prakash, L.; Prakash, S. Mechanism of nucleotide incorporation opposite a thymine-thymine dimer by yeast DNA polymerase eta. **Proc. Natl. Acad. Sci. USA 2003**, *100*, 12093–12098. [CrossRef]

21. Washington, M.T.; Minko, I.G.; Johnson, R.E.; Wollle, W.T.; Harris, T.M.; Lloyd, R.S.; Prakash, S.; Prakash, L. Efficient and error-free replication past a minor-groove DNA adduct by the sequential action of human DNA polymerases iota and kappa. **Mol. Cell. Biol. 2004**, *24*, 5687–5693. [CrossRef]

22. Pence, M.G.; Blans, P.; Zink, C.N.; Hollis, T.; Fishbein, J.C.; Perrino, F.W. Lesion bypass of N2-ethylguanine by human DNA polymerase iota. **J. Biol. Chem. 2005**, *281*, 21062–21072. [CrossRef] [PubMed]

23. Freudenthal, B.D.; Beard, W.A.; Wilson, S.H. New structural snapshots provide molecular insights into the mechanism of DNA polymerase η. **Proc. Natl. Acad. Sci. USA 2010**, *107*, 9868–9873. [CrossRef] [PubMed]

24. Zhang, Y.; Wu, X.; Guo, D.; Rechkoblit, O.; Wang, Z. Activities of human DNA polymerase kappa in response to the major DNA adducts. **Mol. Cell. Biol. 2006**, *26*, 281, 1044–1048. [CrossRef] [PubMed]

25. Jha, V.; Ling, H. Structural basis of accurate replication beyond a bulky major benzo[a]pyrene adduct by human DNA polymerase η. **DNA Repair 2017**, *49*, 43–50. [CrossRef] [PubMed]

26. Washington, M.T.; Minko, I.G.; Johnson, R.E.; Haracска, L.; Harris, T.M.; Lloyd, R.S.; Prakash, S.; Prakash, L. Efficient and error-free replication past a minor-groove DNA adduct by the sequential action of yeast Rev1 and DNA polymerase zeta. **Mol. Cell. Biol. 2004**, *24*, 6900–6906. [CrossRef]

27. Silverstein, T.D.; Jain, R.; Johnson, R.E.; Prakash, S.; Aggarwal, A.K. Structural basis for error-free replication of oxidatively damaged DNA by yeast DNA polymerase eta. **Structure 2010**, *18*, 1463–1470. [CrossRef]

28. Patra, A.; Nagy, I.D.; Zhang, Q.; Su, Y.; Muller, L.; Guengerich, F.P.; Egli, M. Kinetics, structure, and mechanism of 8-Oxo-7,8-dihydro-2′-deoxyguanosine bypass by human DNA polymerase eta. **J. Biol. Chem. 2014**, *289*, 16867–16882. [CrossRef]

29. Silverstein, T.D.; Johnson, R.E.; Jain, R.; Prakash, L.; Prakash, S.; Aggarwal, A.K. Structural basis for the suppression of skin cancers by DNA polymerase eta. **Nature 2010**, *465*, 1039–1043. [CrossRef]

30. Biertrümpfel, C.; Zhao, Y.; Kondo, Y.; Ramón-Maqués, S.; Gregory, M.; Lee, J.Y.; Masutani, C.; Lehmann, A.R.; Hanaoka, F.; Yang, W. Structure and mechanism of human DNA polymerase eta. **Nature 2010**, *465*, 1044–1048. [CrossRef]

31. Nair, D.T.; Johnson, R.E.; Prakash, S.; Prakash, L. Replication by human DNA polymerase-ιota occurs by Hoogsteen base-pairing. **Nature 2004**, *430*, 377–380. [CrossRef] [PubMed]

32. Nair, D.T.; Johnson, R.E.; Prakash, S.; Prakash, L. Aggarwal, A.K. Rev1 employs a novel mechanism of DNA synthesis using a protein template. **Science 2005**, *309*, 2219–2222. [CrossRef] [PubMed]

33. Nair, D.T.; Johnson, R.E.; Prakash, S.; Prakash, L. Aggarwal, A.K. Protein-template-directed synthesis across an acrolein-derived DNA adduct by yeast Rev1 DNA polymerase. **Structure 2008**, *16*, 239–245. [CrossRef] [PubMed]

34. Nair, D.T.; Johnson, R.E.; Prakash, L.; Prakash, S.; Aggarwal, A.K. DNA synthesis across an abasic lesion by yeast REV1 DNA polymerase. **Mol. Cell. Biol. 2011**, *406*, 18–28. [CrossRef] [PubMed]

35. Freedomthall, B.D.; Beard, W.A.; Wilson, S.H. New structural snapshots provide molecular insights into the mechanism of high-fidelity DNA synthesis. **DNA Repair 2015**, *3*, 3–9. [CrossRef]

36. Freedomthall, B.D.; Beard, W.A.; Shock, D.D.; Wilson, S.H. Observing a DNA polymerase choose right from wrong. **Cell 2013**, *154*, 157–168. [CrossRef]

37. Nakamura, T.; Zhao, Y.; Yamagata, Y.; Hua, Y.J.; Yang, W. Watching DNA polymerase eta make a phosphodiester bond. **Nature 2012**, *487*, 196–201. [CrossRef]

38. Weaver, T.M.; Cortez, L.M.; Khoang, T.H.; Washington, M.T.; Aggarwal, P.K.; Freedomthall, B.D. Visualizing Rev1 catalyze protein-template DNA synthesis. **Proc. Natl. Acad. Sci. USA 2020**, *117*, 25494–25504. [CrossRef]

39. Perera, L.; Freedomthall, B.D.; Beard, W.A.; Pedersen, L.G.; Wilson, S.H. Revealing the role of the product metal in DNA polymerase β catalysis. **Nucleic Acids Res. 2017**, *45*, 2736–2745. [CrossRef]

40. Wang, J.; Königsberg, W.H. Two metal-ion catalysis: Inhibition of DNA polymerase activity by a third divalent metal ion. **Front. Mol. Biosci. 2022**, *9*, 824794. [CrossRef]

41. Malaby, A.W.; Chakravarthy, S.; Irving, T.C.; Kathuria, S.V.; Bilssel, O.; Lambright, D.G. Methods for analysis of size-exclusion chromatography-small-angle X-ray scattering and reconstruction of protein scattering. **J. Appl. Crystallogr. 2015**, *48*, 1102–1113. [CrossRef] [PubMed]

42. Rambo, R.P.; Tainer, J.A. Bridging the solution divide: Comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering. **Curr. Opin. Struct. Biol. 2010**, *20*, 128–137. [CrossRef] [PubMed]

43. Boldon, L.; Laliberte, F.; Liu, L. Review of the fundamental theories behind small angle X-ray scattering, molecular dynamics simulations, and relevant integrated application. **Nano Res. 2015**, *6*, 25661. [CrossRef] [PubMed]
45. Allison, J.R. Using simulation to interpret experimental data in terms of protein conformational ensembles. *Curr. Opin. Struct. Biol.* 2017, 43, 79–87. [CrossRef]
46. Kikhney, A.G.; Svergun, D.I. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett.* 2015, 589, 2570–2577. [CrossRef]
47. Hub, J.S. Interpreting solution X-ray scattering data using molecular simulations. *Curr. Opin. Struct. Biol.* 2018, 49, 18–26. [CrossRef]
48. Rambo, R.P.; Tainer, J.A. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* 2013, 496, 477–481. [CrossRef]
49. Powers, K.T.; Gildenberg, M.S.; Washington, M.T. Modeling conformationally flexible proteins with X-ray scattering and molecular simulations. *Comput. Struct. Biotechnol. J.* 2019, 17, 570–578. [CrossRef]
50. Pelikan, M.; Hura, G.L.; Hammel, M. Structure and flexibility within proteins as identified through small angle X-ray scattering. *Gen. Physiol. Biophys.* 2009, 28, 174–189. [CrossRef]
51. Powers, K.T.; Elcock, A.H.; Washington, M.T. The C-terminal region of translesion synthesis DNA polymerase η is partially unstructured and has high conformational flexibility. *Nucleic Acids Res.* 2018, 46, 2107–2120. [CrossRef]
52. Bienko, M.; Green, C.M.; Crosetto, N.; Rudolf, F.; Zapart, G.; Coull, B.; Kannouche, P.; Wider, G.; Peter, M.; Lehmann, A.R.; et al. Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. *Science* 2005, 310, 1821–1824. [CrossRef] [PubMed]
53. Ripley, B.M.; Reusch, D.T.; Washington, M.T. Yeast DNA polymerase η possesses two PIP-like motifs that bind PCNA and Rad6-Rad18 with different specificities. *DNA Repair* 2020, 95, 102968. [CrossRef] [PubMed]
54. Ohmori, H.; Hanafusa, T.; Ohashi, E.; Vaziri, C. Separate roles of structured and unstructured regions of Y-family DNA polymerases. *Adv. Protein Chem. Struct. Biol.* 2009, 78, 99–146. [PubMed]
55. Johnson, R.E.; Washington, M.T.; Haracska, L.; Prakash, S.; Prakash, L. Eukaryotic polymerases ηa and zeta act sequentially to bypass DNA lesions. *Nature* 2000, 406, 1015–1019. [CrossRef]
56. Baranovsky, A.G.; Lada, A.G.; Siebler, H.M.; Zhang, Y.; Pavlov, Y.I.; Tahir, T.H. DNA polymerase delta and zeta switch by sharing accessory subunits of DNA polymerase delta. *J. Biol. Chem.* 2012, 287, 17281–17287. [CrossRef]
57. Johnson, R.E.; Prakash, L.; Prakash, S. Pol31 and Pol32 subunits of yeast DNA polymerase delta are also essential subunits of DNA polymerase zeta. *Proc. Natl. Acad. Sci. USA* 2012, 109, 12455–12460. [CrossRef]
58. Baranovsky, A.G.; Lada, A.G.; Siebler, H.M.; Zhang, Y.; Pavlov, Y.I.; Tahir, T.H. DNA polymerase delta and zeta switch by sharing accessory subunits of DNA polymerase delta. *J. Biol. Chem.* 2012, 287, 17281–17287. [CrossRef]
59. Johnson, R.E.; Prakash, L.; Prakash, S. Pol31 and Pol32 subunits of yeast DNA polymerase delta are also essential subunits of DNA polymerase zeta. *Proc. Natl. Acad. Sci. USA* 2012, 109, 12455–12460. [CrossRef]
60. Malik, R.; Kopylov, M.; Gomez-Llorente, Y.; Jain, R.; Johnson, R.E.; Prakash, L.; Prakash, S.; Ubarretxena-Belandia, I.; Aggarwal, A.K. Structure and mechanism of B-family DNA polymerase zeta specialized for translesion DNA synthesis. *Nat. Struct. Mol. Biol.* 2020, 27, 913–924. [CrossRef]
61. Jain, R.; Rice, W.J.; Malik, R.; Johnson, R.E.; Prakash, L.; Prakash, S.; Ubarretxena-Belandia, I.; Aggarwal, A.K. Cryo-EM structure and dynamics of eukaryotic DNA polymerase delta holoenzyme. *Nat. Struct. Mol. Biol.* 2019, 26, 955–962. [CrossRef]
62. Washington, M.T.; Johnson, R.E.; Prakash, L.; Prakash, S. Human DNA polymerase zeta is a promiscuous extender of mispaired primer termini. *Proc. Natl. Acad. Sci. USA* 2002, 99, 1910–1914. [CrossRef] [PubMed]
63. Harras, L.W.; Prakash, L. Role of human DNA polymerase kappa as an extender in translesion synthesis. *Proc. Natl. Acad. Sci. USA* 2002, 99, 16000–16005. [CrossRef] [PubMed]
64. Lone, S.; Townson, S.A.; Uljon, S.N.; Johnson, R.E.; Brahma, A.; Nair, D.T.; Prakash, S.; Prakash, L.; Aggarwal, A.K. Human DNA polymerase kappa encircles DNA: Implications for mismatch extension and lesion bypass. * Mol. Cell* 2007, 25, 601–614. [CrossRef] [PubMed]
65. Irimia, A.; Eoff, R.L.; Guengerich, F.P.; Egli, M. Structural and functional elucidation of the mechanism promoting error-prone synthesis by human DNA polymerase kappa opposite the 7,8-dihydro-8-oxo-2′-deoxyguanosine adduct. *J. Biol. Chem.* 2009, 284, 22467–22480. [CrossRef] [PubMed]
66. Jha, V.; Bian, C.; Xing, G.; Ling, H. Structure and mechanism of error-free replication past the major benzo[a]pyrene adduct by human DNA polymerase kappa. *Nucleic Acids Res.* 2016, 44, 4957–4967. [CrossRef]
67. Lacey, C.; Tehseen, M.; Bakshi, S.; Percival, M.; Takahashi, M.; Sobhy, M.A.; Raducanu, V.S.; Blair, K.; Musckett, F.W.; Ragan, T.J.; et al. Cryo-EM structure of human Pol kappa bound to DNA and mono-ubiquitylated PCNA. *Nat. Commun.* 2021, 12, 6095. [CrossRef] [PubMed]
68. Yuan, Z.; Georgescu, R.; Schauer, G.D.; O’Donnell, M.E.; Li, H. Structure of the polymerase epsilon holoenzyme and atomic model of the leading strand replisome. *Nat. Commun.* 2020, 11, 3156. [CrossRef] [PubMed]