A Structural and Biochemical Investigation of Human DNA Polymerase Beta

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Andrew James Reed, B.S.

The Ohio State University Biochemistry Graduate Program

The Ohio State University

2018

Dissertation Committee

Dr. Zucai Suo, Advisor

Dr. Jane Jackman

Dr. Kotaro Nakanishi

Dr. Marcos Sotomayor

Abstract

Adenosine, guanosine, cytidine, and thymidine nucleotides are the building blocks of life and are arranged in distinct combinations to give unique genomic DNA sequences. These DNA encode for all biological molecules and processes and therefore the preservation of genomic integrity is essential for cell growth and viability. To do so the cell has evolved specialized enzymes for efficient and faithful DNA replication. However, genomic DNA is continually damaged by reactive agents occurring from host processes or those encountered through the environment. The resulting DNA damage can act as mutagens causing errors in DNA replication and ultimately leading to disease states or can cause blocks to DNA replication causing replication fork collapse and potentially cell death. To counteract the onslaught of DNA damage cells have evolved a multitude of DNA damage repair mechanisms that can directly revert damaged DNA bases back to the canonical bases, remove and replace single DNA bases (base excision repair), or remove and resynthesize segments of DNA containing damage (nucleotide excision repair and mismatch repair).

Base excision repair (BER) in humans is initiated by a damage specific DNA glycosylase that recognizes and removes a single damaged base from DNA, resulting in a product abasic site. This action is followed by cleavage of the abasic site containing DNA strand by apyrimidinic/apurinic endonuclease 1 generating a single-nucleotide gapped

DNA substrate with a deoxyribophosphate-aducted 5′-end (5′-dRP). This gap is then filled and the 5′-dRP removed by DNA polymerase β (Polβ) resulting in a nicked DNA substrate. Finally, this nick is ligated by DNA LigaseIII/XRCC1 to complete repair.

Here I have investigated the structure and function of Polβ to better define its role in DNA repair. Through time-resolved X-ray crystallography and pre-steady-state gel-based kinetics, I have identified and characterized a third divalent metal ion utilized in the synthesis of DNA, determined the mechanism of nucleotide stereoselectivity, and elucidated distinct mechanistic steps following nucleotide incorporation.

Acknowledgments

I would like to thank my doctoral advisor Dr. Suo for his belief in my abilities and his continual encouragement to pursue new discoveries in our field of research. I am thankful for his guidance and support throughout my graduate career and his advice as I look for new positions following graduation. I would like to thank Dr. Jane Jackman for taking on the role of academic advisor following the departure of Dr. Suo from the university. Her support and guidance during this transition was invaluable and greatly appreciated. In addition to Dr. Jackman, I would like to thank my other committee members, Drs. Kotaro Nakanishi and Marcos Sotomayor. These successful and talented researchers have been instrumental to my success as a scientist and have provided great insight to my various projects throughout my graduate career. Finally, I would like to thank the past members of Dr. Suo's lab, Drs. Varun Gadkari, Austin Raper, Anthony Stephenson, Jack Tokarsky, Walter Zahurancik, Vimal Prakash, and Rajan Vyas. These labmates made graduate school and research fun, even on the most unsuccessful days. They were always available for guidance during struggles and celebrations in success. It was an honor to work with these men and I can honestly say without them I would not be where I am today.

Vita

2009–2013……....B.S. Biochemistry
University of Mount Union, Alliance, Ohio

2013–2018……Ph.D. Biochemistry
The Ohio State University, Columbus, Ohio

2013–2014……Ohio State Biochemistry Program Fellow
The Ohio State University, Columbus, Ohio

2014–2015……Graduate Teaching Associate
Center for Life Sciences Education
The Ohio State University, Columbus, Ohio

2014–2017……Graduate Teaching Associate,
Department of Chemistry and Biochemistry,
The Ohio State University, Columbus, Ohio

2017–2018……Pelotonia Graduate Research Fellow
The Ohio State University, Columbus, Ohio

Publications

***denotes co-first authorship**

1. Raper, A. T., **Reed, A. J.***, and Suo, Z. (2018) Kinetic Mechanism of DNA Polymerases: Contribution of Conformational Dynamics and a Third Divalent Metal Ion. Chem. Rev. 118 (12), 6000-6025.

2. **Reed, A.J.**, and Suo, Z. (2017) Time-Dependent Extension from an 8‑Oxoguanine Lesion by Human DNA Polymerase Beta. J. Am. Chem. Soc. 139, 9684-9690.
3. Vyas R., **Reed, A.J.***, Raper, A.T., Suo Z. (2017) Structural basis for the D-stereoselectivity of Human DNA polymerase β. Nucleic Acids Res., 45 (10), 6228-6237.
4. Raper, A.T., **Reed, A.J.***, Gadkari, V.V., and Suo, Z. (2017) Advances in Structural and Single-Molecule Methods for Investigating DNA Damage Tolerance and Repair. Chem. Res. Toxicol., 30, 260-269.
5. **Reed, A.J.***, Vyas R., Raper, A.T., Suo Z. (2017) Structural Insights into the Post-Chemistry Steps of Nucleotide Incorporation Catalyzed by a DNA Polymerase. J. Am. Chem. Soc., 139, 465-471.
6. Vyas, R., **Reed, A.J.**, Tokarksy, E.J., and Suo, Z. (2015) Viewing human DNA polymerase β faithfully and unfaithfully bypass an oxidative lesion by time-dependent crystallography. J. Am. Chem. Soc., 137, 5225-5230.

Fields of Study

Major Field: The Ohio State University Biochemistry Graduate Program

Table of Contents

List of Schemes

List of Tables

List of Figures

Chapter 1. Introduction to DNA Replication and DNA Polymerases

Kinetic Mechanism of DNA Polymerases: Contributions of Conformational Dynamics and a Third Divalent Metal Ion

Reproduced in part with permission from Raper, Austin T., Reed, A.J., and Suo, Z. (2018) Kinetic Mechanism of DNA Polymerases: Contributions of Conformational Dynamic and a Third Divalent Metal Ion. *Chem. Rev.* DOI: 10.1021/acs.chemrev.7b00685.

Author Contributions

Austin T. Raper and Andrew J. Reed equally contributed as co-first authors. Dr. Zucai Suo provided some conceptual direction and helped edit the manuscript.

1.1 Abstract

Faithful transmission and maintenance of genetic material is primarily fulfilled by DNA polymerases. During DNA replication, these enzymes catalyze incorporation of deoxynucleotides into a DNA primer strand based on Watson-Crick complementarity to the DNA template strand. Through the years, research on DNA polymerases from every family and reverse transcriptases, has revealed structural and functional similarities, including a conserved domain architecture and purported two-metal-ion mechanism for nucleotidyltransfer. However, it is equally clear that DNA polymerases possess distinct differences that often prescribe a particular cellular role. Indeed, a unified kinetic mechanism to explain all aspects of DNA polymerase catalysis, including DNA binding, nucleotide binding and incorporation, and metal-ion-assisted nucleotidyltransfer (*i.e.* chemistry), has been difficult to define. In particular, the contributions of enzyme conformational dynamics to several mechanistic steps and their implications for replication fidelity are complex. Moreover, recent time-resolved X-ray crystallographic studies of DNA polymerases have uncovered a third divalent metal ion present during DNA synthesis, the function of which is currently unclear and debated within the field. In this review, we survey past and current literature describing the structures and kinetic mechanisms of DNA polymerases from each family to explore every major mechanistic step while emphasizing the impact of enzyme conformational dynamics on DNA synthesis and replication fidelity. This also includes brief insight into the structural and kinetic techniques utilized to study DNA polymerases and RTs. Furthermore, we present the evidences for the two-metal-ion mechanism for DNA polymerase catalysis prior to

interpreting the recent structural findings describing a third divalent metal ion. We conclude by discussing the diversity of DNA polymerase mechanisms and suggest future characterization of the third divalent metal ion to dissect its role in DNA polymerase catalysis.

1.2 Introduction

It is well-known that enzymes evolved for catalysis on nucleic acid substrates often undergo conformational dynamics and engage metal ion cofactors to achieve remarkable catalytic efficiency and reaction specificity.[1-14] In fact, replication of valuable genetic material is entrusted to DNA polymerases, which utilize divalent metal ions to catalyze DNA synthesis. Since their initial discovery in 1950s,[15,16] many DNA polymerases have been identified and phylogenetically classified into distinct A, B, C, D, X, Y, and reverse transcriptase (RT) families based on sequence homology as well as functional and structural analyses.[8,17,18]

As DNA polymerases catalyze the same fundamental reaction (*i.e.* incorporation of deoxyribonucleotide (dNTP) into a nascent DNA primer strand), one could expect these enzymes to share a unified kinetic mechanism describing DNA binding, nucleotide binding, and nucleotide incorporation. However, functional studies have revealed that each polymerase family is often suited to a particular cellular role[19,20] as evident through the utilization of distinct DNA substrates (*i.e.* primer-template DNA, gapped DNA, damage-containing DNA, single-stranded DNA, *etc.*) and wide-ranging nucleotide substrate specificities, which result in varying DNA replication efficiency and fidelity.[8,20-26] Indeed, while some mechanistic steps remain common among DNA polymerases,

researchers have uncovered several events that seem unique to a particular polymerase, or more broadly, a polymerase family. These events are often related to conformational dynamics and may prescribe distinct properties to the polymerase, which dramatically influence DNA and nucleotide binding as well as nucleotide incorporation. In fact, there is substantial debate about the involvement of a particular conformational change in the rate-limiting step of single-nucleotide incorporation and how this step may influence the fidelity of DNA polymerization.[4-6,20,27-29] In this review, we will describe the minimal kinetic mechanism for single-nucleotide incorporation determined by extensive structural and functional studies of DNA polymerases performed by our lab and many others over the last 30 years and we will emphasize the importance of DNA polymerase dynamics to the mechanism of DNA polymerization. Altogether, our comprehensive analysis of DNA polymerase kinetics has led us to the conclusion that a common kinetic mechanism, encompassing all DNA polymerases, likely does not exist and each enzyme should be considered independently.

A two-metal ion mechanism for enzymes that can act on the phosphodiester backbone of DNA or RNA was first postulated by Beese and Steitz in 1991 based on crystal structures of exonucleolytic substrate and product complexes within the active site of the 3′-5′ exonuclease domain of *Escherichia coli* DNA polymerase I (Pol I).[30,31] They postulated that the mechanism of exonucleolytic cleavage would extend to DNA polymerization with each divalent metal ion coordinating essential active site residues and substrate groups as well as providing necessary transition-state stabilization for DNA synthesis thereby reducing the activation energy and facilitating successful

nucleotidyltransfer onto a DNA primer strand. Thus far, the putative roles of the two divalent metal ions during the DNA polymerase-catalyzed reaction have been well-established empirically through biochemical and structural investigations.[1-3,8,24-26,30-37]

Notably, this proposed two-metal ion mechanism (Figure 1.1A) has been heralded as "a mechanism for all polymerases" and draws support from the fact that many crystal structures of DNA or RNA polymerases in complex with nucleic acid and incoming nucleotide (E•DNA•dNTP or E•RNA•rNTP, ternary complex) contain two divalent metal ions in the polymerase active site.[2,26] In a striking example of evolutionary conservation, DNA polymerases from all families have been characterized to follow the same two-metal ion mechanism based on mutational analysis, structural studies, and kinetic investigation.[2,8,25,33,34] However, recent time-resolved (also known as time-lapse, time-dependent, or soak-trigger-freeze) crystallographic studies of the Y- and X-family DNA polymerases, including human DNA polymerases η (hPolη)[38,39], β (hPolβ)[40-45], and μ (hPolμ)[46], have provided substantial evidence to compel an expansion of the two-metal ion mechanism to include a transient (*i.e.* not observed in all time-resolved partial reaction structures), third divalent metal ion, the precise role of which is currently debated (Figures 1.1B and 1.1C).[47,48]

In this review, we aim to briefly summarize the extensive evidence supporting the two-metal ion mechanism for DNA polymerization while highlighting the possibility of a third divalent metal ion and evaluating its involvement in catalysis as well as its biological purpose and significance. This will include a detailed synopsis of the seminal time-resolved X-ray crystallography findings over the last five years that have sparked

renewed interest in the metal ion mechanism including discussion about the evidence, timing, and dynamic nature of the third divalent metal ion. As a result of its transient character, there is some inconsistency with the time at which the third divalent metal ion appears during the reaction with some groups reporting its occupancy during nucleotidyltransfer[38,39,42,45] and others reporting its appearance only in the product complex.[40,41,44,46] Thus, it is unclear if the third divalent metal ion serves a role in transition-state stabilization (Figure 1.1B), product release, catalysis of the reverse reaction (*i.e.* pyrophosphorolysis, Figure 1.1C), or in modulating the chemical equilibrium of nucleotidyltransfer through product-state stabilization. A recent computational analysis of the third divalent metal ion with hPolη[49] supports roles in transition-state stabilization during the forward and reverse reactions. Similarly, our work with hPolβ[42,45] suggests a possible role in transition-state stabilization, while other structural and computational studies completed with hPolβ[40,41,44,50,51] and hPolμ[46] provide evidence for perturbation of the chemical equilibrium by inhibition of pyrophosphorolysis by the third divalent metal ion. Thus, the role of the third divalent metal ion is yet to be fully delineated, and may be unique for each polymerase or polymerase family. We will conclude with a short discussion of the implications that the third divalent metal ion has for the polymerase field including its potential role in the mechanisms of replicative polymerases or as a potential target for antiviral therapies.

1.3 DNA Polymerases and DNA Polymerization

For 5′–3′ nucleic acid synthesis, the terminal 3′-hydroxyl group of a DNA or RNA strand serves as a nucleophile to attack the α-phosphate of a dNTP or ribonucleotide

(rNTP) to form a phosphodiester bond while releasing pyrophosphate (PP_i) as a byproduct (Figure 1.1). In effect, a phosphodiester bond is transferred from the nucleotide to the nascent nucleic acid strand (*i.e.* nucleotidyltransfer). This reaction is catalyzed by enzymes termed DNA/RNA polymerases which bind both DNA/RNA and nucleotide substrates. As DNA and RNA polymerases share certain structural and functional similarities, much of the mechanistic discussion of DNA polymerases focused on in this review may also apply to RNA polymerases. However, for more detailed evaluations of RNA polymerase structure and mechanism, we point the interested readers to several insightful reviews.[52-56]

DNA polymerases take advantage of the specific shape and hydrogen bonding patterns of nucleobase pairs (*i.e.* A:T, G:C) to faithfully recognize and incorporate correct nucleotides during DNA synthesis.[23] In addition to following a conserved two-metal-ion mechanism for nucleotide incorporation (Figure 1.1A),[2,26] DNA polymerases of all families adopt a "right-hand" architecture (with the exception of the X-family members which are left-handed: hPolβ, hPolλ, hPolμ, and terminal deoxynucleotidyl transferase (TdT))[57] consisting of fingers, palm, and thumb subdomains (Figure 1.2).[8,33] Along with these core domains, DNA polymerases may possess auxiliary domains (Figure 1.2) which often help in the execution of a specific biological function. For example, i) members of the A- and B- families often demonstrate high base substitution fidelity during DNA synthesis partially due to their accessory 3′-5′ exonuclease domain, which removes the small number of incorrect nucleotides incorporated during DNA replication; ii) members of the X-family may contain a deoxyribophosphate lyase (dRPase) domain for processing

DNA ends during DNA repair; and iii) members of the Y-family contain a little finger subdomain (also named polymerase associated domain (PAD)) thought to serve a role in damaged DNA binding. These unique accessory subdomains and biochemical characteristics outfit polymerases from a particular family for a specific biological function.[19] Thus, the faithful, efficient, and processive A- and B-family polymerases perform the bulk of leading and lagging DNA strand replication.[58] In contrast, the error-prone and distributive X-family and Y-family DNA polymerases function in DNA repair and DNA damage response, respectively. Thus, the cell has evolved specialized DNA polymerases to perform an array of diverse functions and activities.[19]

To catalyze nucleotidyltransfer, DNA polymerases require divalent metal ion cofactors. The roles of these metal ions in catalysis were elucidated in early structures of the Klenow fragment of *E. coli* DNA polymerase I with single-stranded DNA and dTMP product bound to the 3′-5′ exonuclease domain.[30,31,37] As the enzymatic synthesis and decomposition of nucleic acid molecules are closely related processes, the two-metal-ion mechanism proposed for 3′-5′ exonuclease degradation was extended to DNA polymerization (Figure 1.1A). In the exonuclease active site, one divalent metal ion was coordinated by several carboxylate side chains of surrounding amino acids (Asp355, Glu357, and Asp501), a water molecule, and the 5′-phosphate of the primer terminus. An additional divalent metal ion was coordinated by Asp355, the 5′-phosphate of dTMP, and several water molecules. Through mutation of the coordinating residues to alanine it was determined that these metal ions serve distinct mechanistic roles. Interestingly, it was later discovered that the catalytic subunit of HIV-1 reverse transcriptase[59] (HIV-1 RT)

shares the same "right-hand" domain architecture of Klenow fragment with finger, palm, and thumb domains arranged to form the DNA binding cleft. Moreover, three conserved carboxylate amino acids identified in Klenow fragment were found in HIV-1 RT and their mutation to Ala also abolished catalytic activity.[59] Together, these data strongly supported the two-metal-ion mechanism for phosphoryltransfer reactions, including phosphodiester bond formation and degradation (Figure 1.1A).[60]

This two-metal-ion mechanism for DNA synthesis was later exemplified through the structure of rat DNA polymerase β, an X-family member, bound to primer-template DNA and dideoxy-terminated nucleotide (ddNTP).[61] From this structure, and those of rat DNA polymerase β bound to Mn^{2+} and dATP,[62] a common nucleotidyltransfer reaction mechanism involving two divalent metal ions for all DNA polymerases was postulated (Figure 1.1A).[61] Following polymerase binding at the primer-template junction of a DNA substrate, an incoming nucleotide is bound and positioned in the active site by i) Watson-Crick base-pairing with the templating base; ii) intermolecular contacts between the base, sugar, and phosphates with amino acid residues; and iii) coordination of two divalent metal ions by the three carboxylate residues. One metal ion binds between the primer terminal O3′ atom and the α-phosphate of the incoming dNTP and is often referred to as the A-site (M_A) or catalytic metal ion (Figure 1.1A). The second metal ion is coordinated by the incoming dNTP through the non-bridging oxygen atoms of the α-, β-, and γ-phosphates and is often referred to as the B-site or nucleotide binding metal ion (M_B) as its appearance coincides precisely with the binding of nucleotide (Figure 1.1A). During catalysis, M_A serves as a Lewis acid to lower the pK_a of the primer hydroxyl proton for

abstraction and subsequent in-line nucleophilic attack on the α-phosphate of the dNTP to form a pentacoordinated transition-state with the 3′-oxygen of the primer terminus and four oxygen atoms of the α-phosphate, including one from the PP_i leaving group, occupying each position of the trigonal bipyramid. On the other hand, M_B acts to orient the triphosphate moiety of the bound nucleotide for catalysis and destabilizes the ground state ternary complex of the polymerase to promote catalysis. Furthermore, following nucleophilic attack, M_B stabilizes the pentacoordinated transition-state and neutralizes the developing negative charge on the PP_i leaving group (Figure 1.1A).[2]

Support for the two-metal-ion mechanism of DNA polymerization exists for structurally characterized DNA polymerases from all major families including A,[63-66] B,[67-70] C,[71-74] X,[40,61,62,75-79] and Y[38,80-86] as demonstrated through ternary complex structures of enzyme, DNA, and dNTP with bound divalent metal ions (Figure 1.3). For example, the structures of bacteriophage T7 DNA polymerase[63] as well as *Thermus aquaticus* (*Taq*) DNA polymerase I[64] of the A-family, with DNA, ddNTP, and both divalent metal ions bound, were solved and are consistent with the earlier structural and mechanistic findings with rat DNA polymerase β.[61,62] Crystal structures of the replicative B-family DNA polymerases, including the bacteriophage polymerases T4[67] and RB69[68] also support the two-metal-ion mechanism. Consistently, two metal ions are also found in the active sites of repair and damage bypass DNA polymerases as demonstrated through crystal structures of rat DNA polymerase β,[36,37] discussed above, human DNA polymerase λ (hPolλ),[76] hPolβ,[75,78] and hPolμ[77] of the X-family, as well as hPolη[80] and *Sulfolobus solfataricus* DNA polymerase IV (Dpo4)[87] of the Y-family. Interestingly, RTs

also engage two divalent metal ions for catalysis as demonstrated by the ternary crystal structure of HIV-1 RT.[88] Limited structural evidence from the C- and D-families of DNA polymerases is available as these enzymes are under-represented in the protein data bank. However, a ternary complex crystal structure of a C-family polymerase from *Geobacillus kaustophilus*, as well as a lower resolution (4.6 Å) structure of *E. coli* PolIIIα, suggest a two-metal ion mechanism for nucleotide incorporation.[71,73]

1.4 Kinetic and Structural Mechanism of DNA Polymerases

Throughout the years, mechanistic studies of DNA polymerases from many diverse families, as well as reverse transcriptases, have culminated in a comprehensive kinetic pathway for nucleotide incorporation (Scheme 1.1A).[5,8-12,14,21-23,27,29,31,43,59,61,62,89-112] While particular details of this model may vary between DNA polymerases or systems (*i.e.* kinetically obligated removal or inclusion of elementary steps, see Scheme 1.1), we attest that the polymerase-catalyzed addition of correct nucleotides into a growing DNA primer strand occurs through ten steps (Scheme 1.1A). A DNA polymerase first binds a DNA substrate (Step 1, Scheme 1.1A) containing a primer-template junction to form the binary complex (E•DNA_n*, Scheme 1.1A), Initial DNA binding may place the terminal base pair of the DNA substrate within the polymerase active site (*i.e.* pre-insertion state) thereby occluding dNTP binding. However, DNA translocation (Step 2, Scheme 1.1A) by one nucleotide (E•DNA_n, Scheme 1.1A) to an insertion state correctly positions the templating base and creates the necessary space to bind an incoming dNTP in the subsequent step (Step 3, Scheme 1.1A). Notably, Step 3 includes the association of M_B and possibly M_A. Upon formation of this ground-state or loose ternary complex

(E•DNA_n•dNTP, Scheme 1.1A), many polymerases then undergo a conformational change (Step 4, Scheme 1.1A) of the finger subdomain (or the thumb subdomain for the X-family DNA polymerases) which encloses the newly-formed base pair of the templating nucleotide and the incoming dNTP to form the tight ternary complex (E′•DNA_n•dNTP, Scheme 1.1A). A second conformational change (Step 5, Scheme 1.1A) within the polymerase active site generates the activated ternary complex (E″•DNA_n•dNTP, Scheme 1.1A) wherein reactive groups, including divalent metal ions, catalytic carboxylate residues, 3′-OH of the primer strand, and the α-phosphate of the bound nucleotide, are properly aligned for subsequent nucleotidyltransfer (Step 6, Scheme 1.1A), conventionally referred to as the chemistry step, which extends the primer strand by one nucleotide (E″•DNA_{n+1}•PP_i, Scheme 1.1A). The nucleotide-binding induced conformational changes (Steps 4 and 5, Scheme 1.1A) are reversed in Steps 7 and 8 (Scheme 1.1A) before PP_i is released (Step 9, Scheme 1.1A) from the polymerase active site. Following the reverse conformational changes and PP_i dissociation, the polymerase may translocate by one base pair along the DNA (Step 10_a, Scheme 1.1A) for additional cycles of nucleotide incorporation (*i.e.* processive DNA synthesis) or may dissociate (Step 10_b, Scheme 1.1A) from the DNA substrate (*i.e.* distributive DNA synthesis).

1.4.1 DNA binding and associated polymerase dynamics

The inclusion and order of the elementary steps in Scheme 1.1A are strongly supported by kinetic, structural, and/or biophysical evidence. Logically, DNA binding (Step 1, Scheme 1.1A) occurs before dNTP binding as the templating information

required for faithful replication is encoded in the DNA. This assertion is supported by inhibitor studies of nucleotide incorporation using PP_i,[24,113] as well as processivity assays, wherein DNA polymerases are observed to incorporate more than one nucleotide per DNA binding event.[90-92,94] In addition, ^{32}P-partioning experiments with *E. coli* Pol I indicated that the reaction followed a specific order in which the polymerase first associated with the DNA then bound dNTP.[114] Lastly, the relative affinity (*i.e.* K_d^{DNA}) of many polymerases for DNA is often in the sub-nanomolar concentration range, while binding affinities for correct or incorrect dNTPs (*i.e.* K_d^{dNTP}) often range from micromolar to millimolar concentrations. Accordingly, DNA polymerases likely spend disproportionately more time bound to DNA than to dNTP, increasing the likelihood of a strict order of substrate binding events. However, a recent structural and biochemical study of the X-family member hPolλ shows a preformed nucleotide binding pocket and reports relatively high affinities for dNTPs with a slight preference for dATP (3.3 μM for dATP and 15-45 μM for the other three dNTPs) in the absence of DNA. This suggests that the hPolλ may in fact bind Mg^{2+}-associated dNTP before DNA,[115] and helps to explain the higher base substitution frequency of hPolλ relative to hPolβ, a close X-family homolog.[109,116,117] While the ability of hPolλ to bind dNTPs prior to DNA is unusual, it has been structurally observed before[62,64,118-120] with the nucleotides often bound in a non-productive conformation. However, dNTP bound crystal structures of truncated hPolλ have shown productive binding of dNTP at the polymerase active site, but the global conformation of the protein has yet to reach the catalytically active state in the absence of DNA.[115] Similarly, dNTP bound crystal and solution NMR structures

(E•dNTP) of African swine fever virus (ASFV) Pol X, an X-family homolog, have revealed dGTP bound in a productive conformation that allows formation of *syn*-dGTP:dG Hoogsteen base pairs upon subsequent DNA binding,[118,121] which is different from the binding of nucleotide in multiple conformations shown for *Thermus thermophilus* Pol X.[122] Importantly, the structural results regarding this unique substrate binding order for ASFV Pol X have been confirmed by steady-state inhibition assays and nucleotide trapping assays.[123] Additionally, modeling based on chemical shift perturbations suggests that nucleotide binding to ASFV Pol X induces a conformational change in the absence of DNA, which further substantiates that dNTP binds first for this viral polymerase.[124] Nevertheless, the binding of nucleotide prior to DNA is likely a rare occurrence and may contribute to the low fidelity of Pol X.[125]

In addition, there may be certain scenarios in which the outcome of nucleotide incorporation is not influenced by the identity of the templating base. For example, a DNA polymerase may prefer to almost exclusively incorporate one particular nucleotide or catalyze template independent nucleotide incorporation such as the Y-family DNA polymerase Rev1, which is known to preferentially incorporate dCTP regardless of the templating base through a "protein template" arginine residue,[83,126-129] or the X-family DNA polymerase TdT, which prefers single-stranded DNA over double-stranded DNA and is incapable of replicating a DNA template,[23] respectively. Moreover, damage to the DNA may result in an unreadable templating base in which case a Y-family DNA polymerase is recruited for non-templated lesion bypass. Accordingly, while certain exceptions to the order of DNA and dNTP substrate binding to a DNA polymerase exist

for specific biological contexts or for specialized DNA polymerases, it is widely accepted that DNA binding typically occurs prior to dNTP binding.

DNA polymerases have been observed to bind to a variety of DNA substrates and the substrate specificity seems to depend on the polymerase family as well as the particular biological function. In general, it is understood that DNA polymerases bind primer-template DNA substrates wherein the 3′-end of the primer strand is recessed relative to the 5′-end of the template strand. However, specialized DNA polymerases such as hPolβ and hPolλ of the X-family prefer to act on gapped DNA substrates containing an upstream primer along with a 5′-phosphorylated or 5′-deoxyribophosphate adducted downstream primer,[79,116,117,130,131] and the Y-family polymerases can tolerate binding to and replicating on damage-containing DNA substrates,[132-139] unlike their replicative polymerase counterparts.[140-143] Several pre-steady-state kinetic assays exist to measure the equilibrium dissociation constant (*i.e.* K_d^{DNA}) for DNA binding by a polymerase including the active site titration. During the active site titration, a fixed amount of a DNA polymerase is titrated with varying amounts of a radiolabeled DNA substrate before being rapidly mixed with correct dNTP to initiate nucleotide incorporation.[144] A burst of product formation is observed at each DNA substrate concentration as dNTP is rapidly bound and incorporated by the pre-formed E•DNA (Scheme 1.1) complex. An important consideration for successful execution of the active site titration is that nucleotide binding and incorporation must be much faster than the binding equilibration of a polymerase and DNA (E + DNA ⇌ E•DNA), otherwise the burst of product formation will not be observed. The concentration of the E•DNA

complex is given by the amplitude of the burst phase which varies as a function of DNA substrate concentration. A quadratic binding equation can then be used to determine both the K_d^{DNA} as well as the active concentration of the polymerase. For example, the DNA binding affinity of the catalytic subunit (*i.e.* p261) of human DNA polymerase ε (hPolε) of the B-family, which is responsible for leading strand DNA replication, was measured to be 79 nM and the enzyme was determined to be only ~16% active.[91] Interestingly, the assay was later repeated with the holoenzyme of hPolε (*i.e.* p261, p59, p17, p12) and the binding affinity increased 2.4-fold to 33 nM, while the enzyme did not gain appreciable activity (18% active).[145] In addition, active site titrations have been useful in determining the DNA binding affinities of *Sulfolobus solfataricus* DNA polymerase B1 (PolB1)[92] and Dpo4[94] as both polymerase demonstrate clear burst phase kinetics. However, when a small or indeterminate burst phase is present, as observed for hPolβ,[110,146] the necessary conditions prescribed for an active site titration are not met (*i.e.* nucleotide binding and/or incorporation is not faster than the binding equilibration of a polymerase and DNA) and therefore a different method must be used to accurately measure DNA binding affinity. An alternative strategy to determine polymerase affinity to DNA involves measuring the microscopic rate constants of k_{off} and k_{on} comprising K_d^{DNA} (*i.e.* $k_{off}/k_{on} = K_d^{DNA}$) through assays designed to monitor the kinetics of polymerase dissociation from or association to a DNA substrate, respectively. Indeed, directly measured values for k_{off} and k_{on} often strongly agree with the measured K_d^{DNA} of a DNA polymerase and provide additional insight into mechanistic steps which may kinetically limit multiple rounds of DNA synthesis.[10,90-92,94,110,145,147,148]

Binding of a DNA substrate by a polymerase is often accompanied with conformational dynamics of the enzyme as well as nucleic acids (Figure 1.2). One striking example of protein dynamics upon DNA binding is demonstrated by *S. solfataricus* Dpo4 of the Y-family (Figure 1.4A). While in the apo state, a crystal structure of Dpo4 reveals that the auxiliary little finger domain interacts with the thumb domain and occupies the DNA binding cleft, thereby occluding the binding of a DNA substrate.[149] Consequently, a major 131° rotation and 1.7 Å translation of the little finger domain is observed in the binary complex crystal structure as the little finger breaks contact with the thumb and establishes new contacts with the finger to vacate the necessary space for DNA to bind (Figure 1.4A).[149] Additionally, high affinity of DNA binding[9,94,134] (K_d^{DNA} = 10 nM) is ensured as both the thumb and finger domains rotate 10° to better contact the DNA in the binary complex structure.[149] This dramatic structural transition was further investigated in fluorescence[149] and stopped-flow Förster resonance energy transfer (FRET) studies monitoring distance changes between[22] and within[10] individual Dpo4 domains. Contrary to the initial hypothesis that Dpo4 must follow an induced fit mechanism for DNA binding, wherein the DNA substrate would induce the conformational change of the little finger domain, it was found that in the absence of DNA, Dpo4 exists in conformational equilibrium between the structurally distinct apo and binary complex configurations (Figure 1.4A) and DNA binding selects for the DNA bound state (*i.e.* shifts conformational equilibrium toward DNA bound state). This was concluded as the rate of conformational transition between apo and binary complex upon DNA binding, monitored through relative distance change between interdomain FRET

probes positioned in the little finger and palm domains, was independent of DNA concentration over two orders of magnitude.[22] This assertion was supported by a nuclear magnetic resonance (NMR) study assigning the backbone chemical shifts of full length Dpo4 in the absence and presence of DNA, which suggested that a minor conformation of apo Dpo4 existed in a conformation consistent with that observed in the Dpo4 binary complex crystal structure.[150] Importantly, it was later shown that the little finger mediates initial DNA binding of Dpo4 through a stopped-flow FRET system reporting on intradomain distance changes of the little finger,[10] as predicted by a previous computational investigation.[151] A similar dramatic structural rearrangement is observed for the related Y-family member human DNA polymerase κ (hPolκ) during DNA binding as its little finger domain moves ~50 Å to intimately contact the DNA major groove and the N-clasp, a unique N-terminal extension, helps encircle the DNA substrate.[152] Conformational dynamics during DNA binding can also be observed for the X-family DNA polymerases hPolβ[61,62] and hPolλ,[76,115] but not hPolμ,[77,153] as the 8 kDa N-terminal dRPase domain of both hPolβ and hPolλ move to engage gapped-DNA substrates (Figure 1.2). Notably, subtle differences in the dRPase domain dynamics between hPolβ and hPolλ may help explain the higher affinity of hPolβ (0.077 – 22 nM)[146,154] for gapped-DNA relative to hPolλ (110 nM),[155] and its role as the primary polymerase for short-patch base excision repair (BER).[21] In contrast to the limited domain motion demonstrated by the X-family polymerases upon DNA binding, the gapped- (or nicked) DNA duplex undergoes a dramatic structural change involving a 90° kink occurring at the 5′-phosphodiester bond of the templating base.[79,153,156] Importantly, this unusual DNA

structure is necessary for thumb domain closure during single-nucleotide gap-filling DNA synthesis by hPolβ and therefore ensures that an important fidelity checkpoint is maintained during nucleotide incorporation. Examination of product complex structures from enzymes involved in BER reveals that the DNA becomes progressively bent as it is sequentially bound and processed by most enzymes of the DNA repair pathway (DNA glycosylase → AP endonuclease → X-family DNA polymerase).[157] Accordingly, DNA repair enzymes may recognize and preferentially bind the bent DNA to facilitate rapid and efficient repair of DNA damage. The higher-fidelity A- and B-family DNA polymerases have also been observed to undergo conformational dynamics upon DNA binding as exemplified by comparison of the apo[158-160] and binary or ternary[64,65,68,161] crystal structures of bacteriophage RB69 DNA polymerase, *Pyrococcus furiosus* DNA polymerase, and *Taq* DNA polymerase I (Figure 1.2). In general, beyond the occasional structuring of disordered regions, DNA binding is typically accompanied by movement of the thumb (or the fingers for the X-family DNA polymerases) domain towards the palm domain in order to wrap around the DNA substrate.

1.4.2 DNA translocation and divalent metal-ion binding to the A- and B-sites

Following formation of the E•DNA binary complex (Step 1, Scheme 1.1A), nucleotide incorporation into the primer strand of the bound DNA substrate commences upon binding of a dNTP. Importantly, DNA polymerases in the binary complex may exist in non-productive or productive configurations depending on whether the polymerase active site is bound in the pre-insertion or insertion state, respectively. Indeed, a binary complex crystal structure of the Y-family member Dpo4 showed the polymerase in the

pre-insertion state, while a ternary complex structure revealed the polymerase to have translocated by one base pair along the DNA to the insertion state in order to accommodate the incoming correct dNTP.[162] This essential DNA translocation event was later validated and measured to be rapid (>150 s^{-1} at 20 °C) by stopped-flow[11,12] and single-molecule[9,163] FRET studies monitoring distance changes between a DNA substrate and various domains of Dpo4. Interestingly, the single-molecule FRET studies revealed that the polymerase dynamically fluctuates between the pre-insertion and insertion states on the DNA but exclusively populates the insertion state in the presence of correct nucleotide.[9,163] Similar repositioning of the DNA polymerase from the pre-insertion state to the insertion state via DNA translocation is proposed from crystal structures of the Klenow fragment of *Taq* DNA polymerase I[65] and the large fragment of DNA polymerase I from *Bacillus stearothermophilus*.[164] Indeed, as the polymerase transitions between these two states while bound to DNA (Step 2, Scheme 1.1), nucleotide may directly bind to the polymerase-DNA complex at the insertion site or may induce DNA translocation from the pre-insertion site depending on which state is favored at equilibrium.[112]

Once the polymerase has translocated along the DNA to the insertion state, nucleotide binding can commence (Step 3, Scheme 1.1). Binding of dNTP coincides with association of the A- and B-site divalent metal ions. Time-resolved crystallographic studies (Figure 1.5, see Section 1.5.1) of bacteriophage N4 RNA polymerase indicate that M_B binding occurs simultaneously with nucleotide binding and that M_A binding occurs shortly after.[165] This explicit order for divalent metal ion binding to the polymerase (*i.e.*

M_B followed by M_A) is supported by the crystal structure of DNA polymerase ι (hPolι) wherein M_B is clearly associated with the bound dNTP but M_A has yet to bind.[82] However, those authors suggest that M_A may not be necessary for nucleotidyltransfer, as abstraction of the 3′-hydroxyl proton may occur by an active site carboxylate (Glu 127) positioned unusually near to the primer terminus.[82] In addition to binding after the M_B-associated nucleotide, M_A binding is proposed to occur following the conformational change (Step 4, Scheme 1.1A) to the tight ternary complex, as association of M_B-dNTP, not M_A, is sufficient to elicit the conformational change.[112,166-168] While this implies a defined order for the binding of each metal ion during the kinetic mechanism of DNA polymerase-catalyzed nucleotide incorporation (Scheme 1.1A), explicit evidence for the defined sequence of events is lacking and M_A could associate or dissociate at different step(s) of the mechanism.[112] Nevertheless, time-resolved crystallographic studies of hPolβ[40-42,44,45] and hPolμ[46] (see Sections 1.5.4 and 1.5.5, respectively) have demonstrated that following nucleotide incorporation, M_A dissociates prior to M_B indicating that the relative affinity for the divalent metal ion at the A-site is weaker, and further suggesting that M_A likely associates after M_B.

1.4.3 Nucleotide binding, incorporation, and polymerase fidelity

The apparent affinity of the DNA polymerase binary complex for dNTP (*i.e.* K_d^{dNTP}) and the maximum rate constant of single-nucleotide incorporation (*i.e.* k_{pol}) can be experimentally measured by pre-steady-state kinetic assays.[144] Briefly, DNA polymerase and DNA substrate can be pre-incubated under single-turnover reaction conditions (*i.e.* [E] >> [DNA]) before mixing with various concentrations of correct or

incorrect dNTP. Reactions are quenched at increasing amounts of time and the data are fit to a single-exponential equation ([product] = $A[1 - \exp(- k_{obs}t)]$) to obtain an observed rate constant (k_{obs}) at each concentration of dNTP. The k_{obs} values are then plotted as a function of dNTP concentration and fit to a hyperbolic equation (*i.e.* $k_{obs} = k_{pol}[\text{dNTP}]/(K_d^{dNTP} + [\text{dNTP}])$ to obtain the desired kinetic parameters.[144] Importantly, the measured values of k_{pol} and K_d^{dNTP} for all 16 possible nucleotide incorporations are extremely useful metrics of polymerase efficiency (k_{pol}/K_d^{dntp}), fidelity (calculated as $(k_{pol}/K_d^{dNTP})_{\text{incorrect}} / [((k_{pol}/K_d^{dNTP})_{\text{correct}} + (k_{pol}/K_d^{dNTP})_{\text{incorrect}}]$), and processivity. This latter metric can be calculated as the ratio of k_{pol} to the rate of DNA dissociation (k_{off}, see Section 1.4.1) and describes the average number of bases incorporated by the DNA polymerase during a single DNA binding event, which can be more than 1,500 as observed for highly-processive T7 DNA polymerase (bound to the processivity factor *E. coli* thioredoxin)[90] and human mitochondrial DNA polymerase γ holoenzyme,[169] or less than 20 as observed for the poorly-processive X-family member hPolβ[116] and the Y-family member Dpo4.[94] Importantly, processivity values often increase when the DNA polymerase is associated with accessory subunits or processivity factors (*e.g.* proliferating cell nuclear antigen (PCNA) and β-clamp).[90,108,169-171]

Some of the fastest and most faithful-DNA polymerases exhibit k_{pol} values >200 s^{-1} and bind correct nucleotide with relatively high affinity ($K_d^{dNTP} < 10$ μM), while incorrect nucleotides are bound with ~10- to 100-fold lower affinities and are generally incorporated 100- to 10,000-fold more slowly. Consequently, high fidelity polymerases typically make only one error per ~1 x 10^6 incorporations.[112] For example, the p261

catalytic subunit of hPolε was shown to incorporate correct nucleotides at a rate of 219 – 275 s^{-1} with a high base substitution fidelity of 10^{-4} – 10^{-7} (*i.e.* one error per 10^4 – 10^7 incorporations).[172] More impressively, it was further shown that the 3′-5′ exonuclease activity of hPolε bolstered the overall *in vitro* polymerization fidelity to 10^{-6} – 10^{-11} (*i.e.* one error per 10^6 – 10^{11} incorporations), which unprecedentedly translates to 0.1 – 1 misincorporations per round of human genome replication.[172] This enhancement in overall *in vitro* polymerization fidelity afforded by the exonuclease activity was greater than that observed for the related B-family DNA polymerase PolB1, which demonstrated a two orders of magnitude improvement (*i.e.* 10^{-4} – 10^{-6} to 10^{-6} – 10^{-8}).[173]

For many years, the mechanism by which a DNA polymerase recognizes a mismatch and switches between polymerization and exonuclease modes in order to correct the mismatch remained unclear. It was thought that the exonuclease domain must proofread the nascent DNA for mistakes and, upon identification of a mismatch, must transfer the DNA duplex from the polymerase active site to the exonuclease active site.[95,174,175] This was hypothesized to be a dynamic conformational change between polymerization and editing modes of the polymerase as the distance between the active sites is relatively large (~60 Å *E. coli* Pol IIIα, ~40 Å for *Saccharomyces cerevisiae* Pol ε).[174] Nevertheless, a recent cryo-electron microscopy (cryo-EM) study has revealed the structural basis for mismatch correction by *E. coli* Pol IIIα.[176] Rather than serving an active role as a "proofreader", the exonuclease domain is actually passive, with a terminal mismatch causing the DNA substrate to fray (as supported by NMR analysis of the DNA duplex) resulting in a distorted DNA conformation.[176] Accordingly, the mismatch is essentially

self-correcting as the primer strand from the frayed DNA duplex travels ~55 Å to the exonuclease active site for passive nucleotide excision.[176] This passive mechanism of exonucleolytic cleavage is corroborated by biochemical studies of *S. cerevisiae* Pol ε wherein an extended β-hairpin loop motif, originally thought to serve an active role in mediating a switch between polymerization and editing modes, was shown to have no such effect.[177] Interestingly, relative to matched primer-template termini, the rate of primer extension from a mismatched terminus is slow relative to the rate of exonuclease excision allowing for efficient mismatch removal.[134,172,173,178,179] The structural basis for inefficient polymerization beyond a mismatch stems from a myriad of active site and DNA distortions that misalign reactive groups[180-184] even when the mismatch is several base pairs removed from the primer-template junction (*i.e.* mismatch position *n*-1 to *n*-4).[180] Moreover, binding of a correct nucleotide when the terminal base pair is a mismatch induces distinct structural alterations which ultimately deter nucleotidyltransfer.[184] Together, these structural determinants prevent misincorporations and subsequent extension and push the equilibrium to exonucleolytic removal of the errantly incorporated nucleotide. In contrast to high-fidelity DNA polymerases, moderate-fidelity DNA polymerases[185] such as the X-family members hPolβ and hPolλ lack exonuclease domains (Figure 1.2) and demonstrate poor base substitution fidelity on both non-gapped and gapped DNA substrates ranging from 10^{-2} – 10^{-5} (*i.e.* one error per 10^{2} – 10^{5} incorporations).[109,117,186,187] Similarly, the error-prone lesion-bypass Y-family DNA polymerases including hPolη[147,188] and Dpo4[93] also lack exonuclease domains

(Figure 1.2) and display comparably poor base substitution fidelities of 10^{-2} – 10^{-4} on undamaged DNA substrates.

1.4.4 Kinetic basis for polymerase fidelity and the rate-limiting step of single-nucleotide incorporation

The mechanistic basis by which DNA polymerases achieve their remarkable base substitution fidelity has been thoroughly investigated over the years. Based on the seminal findings of Watson and Crick,[189] it was originally thought that DNA polymerases would achieve high base substitution fidelity from the distinct hydrogen bonding patterns between correct versus incorrect base pairs. However, it was quickly discovered that hydrogen bonding alone could not explain the large difference in efficiency between incorporation of correct and incorrect nucleotides.[190] It was later suggested that both the shape of the nascent base pair within the polymerase active site as well as hydrogen bonding contribute to nucleotide specificity.[191-196] Alternatively, it was hypothesized that the difference in free energy between the chemistry of correct versus incorrect nucleotide incorporation alone could explain polymerase fidelity.[6] However, it has been shown that for many polymerases in which the kinetic mechanism has been thoroughly investigated that chemistry is not the rate-limiting step of correct nucleotide incorporation.[5] Accordingly, research now indicates that many factors including but not limited to hydrogen bonding, free-energy differences, base-pair shape complementarity, and polymerase conformational dynamics contribute to high-fidelity DNA synthesis.[5,6,196-198] In fact, even non-catalytic accessory domains[199] as well as the solvent accessibility and water network of a polymerase active site[200] have been implicated or directly shown to

modulate polymerase fidelity. Furthermore, substrate dynamics have also been hypothesized to effect polymerase fidelity. For example, the rare tautomer hypothesis of polymerase fidelity postulates that replication errors occur at low frequencies due to the formation of high energy tautomers of DNA bases which allow incorrect base pairs to form Watson-Crick-like geometries and mislead the polymerase to catalyze a misincorporation. This hypothesis has gained recent support from crystal structures of the *Bacillus stearothermophilus* DNA polymerase I large fragment bound to a dC:dA mismatch[201] and a mutant of hPolλ bound to a dG:dT mismatch,[202] as well as through NMR spectroscopy of DNA duplexes containing site-specific mismatches, which revealed that sequence-dependent tautomerization and ionization of incoming nucleotides within the polymerase active site leads to misincorporations as originally suspected by Watson and Crick.[189,203,204] Moreover, DNA template dynamics associated with incorrect nucleotide incorporation have also been observed.[205] Lastly, DNA polymerases have been demonstrated to monitor base complementarity through sequence independent minor groove interactions.[206-208] Therefore, it is clear that polymerase fidelity is complex and is achieved through a vast array of polymerase and substrate interactions and dynamics.

1.4.4.1 A two-step binding model for DNA polymerase fidelity

Pre-steady-state kinetic studies coupled with pertinent crystal structures of DNA polymerases and RTs provided the first indication that enzyme conformational dynamics were important for the mechanism of DNA polymerization and polymerase fidelity.[4,5,29,61,63,65,68,88,90,209,210] Together, these studies helped define a two-step nucleotide binding mechanism (Scheme 1.1B) involving rapid equilibrium binding of

dNTP (Step 3_a and 3_b, Scheme 1.1B) followed by an open→closed conformational change of the finger (or the thumb for the X-family DNA polymerases) (Step 4_a and 4_b, Scheme 1.1B), supported by comparison of binary and ternary complex structures for many DNA polymerases (Figure 1.2). If rate-limiting, this conformational change would provide the additional selectivity crucial for discriminating against incorrect dNTPs.[27] For many polymerases, kinetic data from experiments studying the incorporation of a S_p-dNTPαS, a nucleotide analog in which the pro-S_p oxygen of the α-phosphate has been substituted with sulfur, indicated that a conformational change, rather than the chemistry of nucleotidyltransfer, was rate-limiting for single-nucleotide incorporation as the sulfur elemental effect (*i.e.* the decrease in the rate of nucleotide incorporation when using α-thio-dNTP versus normal dNTP) was negligible.[5] Briefly, as the A- and B-site metal ions do not interact with the pro-S_p oxygen of a dNTP, the substitution of this atom to sulfur allows kineticists to infer the identity of the rate-limiting step of nucleotidyltransfer from the magnitude of the change in the observed single-nucleotide incorporation rate. Accordingly, significant elemental effects of 4–11-fold (*i.e.* $k_{obs,\ \mathrm{dNTP}}/k_{obs,\ Sp\text{-}dNTP\alpha S}$) were previously considered to indicate that chemistry is rate-limiting for DNA polymerase catalysis, whereas smaller values (*i.e.* < 2-fold) suggest that pre-chemistry conformational changes are rate-limiting.[5,94,103,211] Notably, while a significant sulfur elemental effect was not often observed for correct nucleotide incorporation, experiments performed with incorrect nucleotide frequently revealed that the chemistry step (Step 5_b, Scheme 1.1B) was rate-determining, presumably as a result of misaligned reactive moieties within the polymerase active site.[5] It is important to mention that the sulfur

elemental effect is no longer considered a reliable diagnostic for the rate-limiting step of nucleotide incorporation as intermediate effects (*i.e.* 2–3-fold) are difficult to interpret, while large effects (*i.e.* >10-fold) are thought to arise from disruption of the geometry of the transition-state (*i.e.* steric effects) and therefore no longer exclusively report on the chemistry step.[5] Moreover, the pro-S_p oxygen of the α-phosphate of the incoming dNTP has been hypothesized to not serve a major role in transition-state stabilization and therefore its substitution with sulfur does not adequately probe the chemistry step.[6] More convincing than the sulfur elemental effect, were results obtained through the pulse-chase/pulse-quench experiment, wherein an increase in reaction amplitude of the pulse-chase compared to the pulse-quench is indicative of a rate-limiting pre-chemistry conformational change.[5,99] During the pulse-quench, a pre-incubated solution of polymerase and DNA is mixed with [α-^{32}P]-radiolabeled dNTP for varying amounts of time before quenching. The pulse-chase proceeds similarly, except that before quenching, an excess of cold dNTP is added to the reaction mixture. Accordingly, if a slow-to-form polymerase complex (E′•DNA$_n$•dNTP, Scheme 1.1B) accumulates before the chemistry step (Step 5$_a$ or 5$_b$, Scheme 1.1B) then the chase with excess cold dNTP should cause an increase in reaction amplitude as the reaction is chased forward (*i.e.* an [α-^{32}P]-dNTP-bound polymerase complex can form additional product rather than the [α-^{32}P]-dNTP dissociating out of the active site). Indeed, results of the pulse-chase/pulse-quench experiments for many DNA polymerases identified a rate-limiting pre-chemistry conformational change and supported the two-step binding mechanism for correct nucleotide incorporation.[4,5]

1.4.4.2 The open→closed conformational change is not rate-limiting

While the initial two-step model for nucleotide selection and incorporation was strongly supported by structural and kinetic data, it was later contested as the measured rate of the open→closed conformational transition (Step 4_a, Scheme 1.1B) for polymerases was too rapid to be considered rate-limiting.[7,27,212,213] As this model depended on the assumption that the rate of the open→closed conformational change (Step 4 , Scheme 1.1A and 1B) must be slow relative to the chemistry step (Step 6, Scheme 1.1A; Step 5, Scheme 1.1B) to afford nucleotide selection specificity, a revised interpretation of the model was necessary.[4] Indeed, studies of T7 DNA polymerase,[27,112] RB69 DNA polymerase,[196,214] and HIV-1 RT[215-217] showed that the rates of the pre-chemistry forward and reverse conformational changes for correct (k_{4a} and k_{-4a}, respectively, Scheme 1.1B) or incorrect nucleotide (k_{4b} and k_{-4b}, respectively, Scheme 1.1B) relative to the rate of chemistry (k_{5a} or k_{5b}, Scheme 1.1B) defined nucleotide specificity. Accordingly, binding of the correct nucleotide rapidly induces a conformational change (*i.e.* large k_{4a}, Scheme 1.1B) to an enzyme complex committed to catalysis (*i.e.* small k_{-4a} relative to k_{5a}, Scheme 1.1B), while binding of the incorrect nucleotide induces a unique conformational change to an enzyme complex ($E^{‡}$•DNA_n•dNTP, Step 3, Scheme 1.1B) which allows rapid release of the incorrect nucleotide (*i.e.* large k_{-4b} relative to k_{5b}, Scheme 1.1B). Altogether, kinetic analysis revealed that the controversy concerning the relative magnitudes of the rates of the pre-chemistry conformational change (k_{4a} and k_{4b}, Scheme 1.1B) versus the chemistry step (k_{5a} and k_{5b}, Scheme 1.1B), and how this could impact nucleotide specificity, was

unfounded if the reverse rate of the pre-chemistry conformational change was slow for correct dNTP (k_{-4a}, Scheme 1.1B), but fast for incorrect dNTP (k_{-4b}, Scheme 1.1B), relative to the chemistry step (k_{5a} and k_{5b}, Scheme 1.1B).[27,112,215,216] This current model suggests that chemistry is fast relative to nucleotide release during correct nucleotide incorporation as the rapid conformational change, promoted by the correct geometry of the base pair within the ground-state ternary complex (E•DNA•dNTP, Scheme 1.1B), ensures the proper alignment of catalytic moieties (E′•DNA_n•dNTP, Scheme 1.1B). On the other hand, chemistry is slow during incorrect nucleotide incorporation as a unique conformational change, prompted by the incorrect geometry of the base pair within the ground-state ternary complex (E•DNA•dNTP, Scheme 1.1B), instigates improper alignment of catalytic groups ($E^{‡}$•DNA_n•dNTP, Scheme 1.1B) and dissociation of nucleotide. In other words, the reversal of the pre-chemistry conformational change is fast relative to chemistry during incorporation of a mismatch, thereby favoring rapid nucleotide dissociation from the polymerase active site prior to nucleotidyltransfer as demonstrated through studies of high-fidelity T7 DNA polymerase[27,112] and RB69 DNA polymerases,[187,203] as well as moderate-fidelity HIV-1 RT.[215-217]

1.4.4.3 Multiple mechanisms of DNA polymerase fidelity

While the latter model (see Section 1.4.4.2)[27,112,196] elegantly explains how high-fidelity DNA polymerases achieve their remarkable substrate specificity for correct nucleotide, it may not extend to the low-fidelity X- and Y-family DNA polymerases. In contrast to the large structural change upon nucleotide binding observed for many A- and B-family DNA polymerases, involving closure of the finger domain,[210] all members of

the Y-family[14,218] and some members of the X-family[57,115,153] do not undergo such a nucleotide-induced conformational change (Figures 1.2 and 1.4B). Despite the structural data, results from the sulfur elemental effect and pulse-chase/pulse-quench experiments for several Y-family DNA polymerases suggest that a pre-chemistry conformational change is rate-limiting for single-nucleotide incorporation.[14,94,103,104] Accordingly, the basis by which error-prone polymerases select for correct nucleotide may involve a distinct mechanism.

Strikingly, stopped-flow fluorescence[219] and FRET[10-12,22] studies of the Y-family polymerase Dpo4 have revealed subtle conformational motions between and within each subdomain (finger, palm, thumb, little finger) during binding and incorporation of a correct nucleotide (Figures 1.4B and 1.4C). While monitoring distance changes between each polymerase domain and the DNA substrate, based on the anti-correlated increases or decreases in the donor and acceptor fluorescent signals, three FRET phases were observed upon mixing the polymerase-DNA binary complex with correct dNTP corresponding to i) rapid DNA translocation by Dpo4 (P_0), ii) synchronized gripping of the DNA substrate by each domain prior to nucleotide incorporation (Figure 1.4B, P_1), and iii) subsequent relaxation of each domain following nucleotidyltransfer (Figure 1.4B, P_2). Interestingly, the slow FRET phase (P_2) vanished during analogous experiments performed using a DNA substrate containing a dideoxy-terminated primer to prevent nucleotide incorporation and therefore must occur following nucleotidyltransfer. The P_1 phase (~15.3 s^{-1}) occurred much faster than the rate-limiting step of single nucleotide incorporation measured by radioactive chemical quench (0.66 s^{-1}). If the rate of

synchronized domain motion (~15.3 s^{-1}) is considered the forward rate for enzyme isomerization (k_{4a}) in Scheme 1.1B, then the rate of chemistry (k_{5a}, Scheme 1.1B) can be calculated as 0.69 s^{-1} from the relationship $k_{5a} \approx k_{4a}k_{pol}/(k_{4a} - k_{pol})$,[27] where k_{pol} is the observed single-turnover rate for correct nucleotide incorporation (0.66 s^{-1}). In contrast to T7 DNA polymerase,[27,112] the forward isomerization rate (k_{4a}, Scheme 1.1B) for Dpo4 is much faster (22-fold) than the surprisingly slow calculated rate of chemistry (k_{5a}, Scheme 1.1B). Consequently, the reverse isomerization rate (k_{-4a}, Scheme 1.1B) must be much slower (0.0017 s^{-1} based on the 410-fold difference between forward and reverse isomerization rates measured for T7 DNA polymerase)[27] than 0.69 s^{-1} in order for Dpo4 to efficiently select the correct nucleotide according to the aforementioned revised model for nucleotide specificity (see Section 1.4.4.2 and Scheme 1.1B). However, given the clear lack of a sulfur elemental effect (1.4) and obvious increase in amplitude (5.5 nM) for the pulse-chase compared to the pulse-quench experiment measured for Dpo4,[94] we hesitate to assign the chemistry step (k_{5a}, Scheme 1.1B) such a slow rate (0.69 s^{-1}) when it appears that some other rate-limiting, pre-chemistry step clearly exists. Indeed, previous kinetic studies of Dpo4 at a range of temperatures (2 – 56 °C) provided four independent lines of kinetic evidence that a pre-chemistry protein conformational change must limit correct nucleotide incorporation.[94,220] Accordingly, nucleotide incorporation for Dpo4 likely proceeds through a mechanism (Scheme 1.1A) requiring two pre-chemistry conformational changes (Steps 4 and 5, Scheme 1.1A). The first conformational change (Step 4, Scheme 1.1A) involves the synchronized domain movements to enhance interaction with the DNA substrate (Figure 1.4B), while the second (Step 5, Scheme

1.1A) is rate-limiting and involves precise alignment of reactive groups achieved through subtle protein motions. Indeed, stopped-flow FRET experiments monitoring distance changes between and within individual domains of Dpo4 during nucleotide binding and incorporation garner support for this model (Figure 1.4C).[10,22] For example, the majority of intradomain FRET pairs demonstrated characteristic P_1 and P_2 phases (*i.e.* anti-correlated phases consistent with interdomain FRET pairs, Figure 1.4B)[11,22] and the P_2 phase was absent during experiments with a dideoxy-terminated primer (Figure 1.4D). However, intradomain FRET pairs positioned within the finger domain showed a unique P_2 phase regardless if nucleotide incorporation was prevented by utility of a dideoxy-terminated primer (Figure 1.4E). Importantly, the unique P_2 phases were in the same direction as P_1 (Figure 1.4E) and the rates of these P_2 phases were on the order of the rate-limiting step of single-nucleotide incorporation measured for Dpo4 by ^{32}P-based assays. We speculate that the observed P_2 phases reflect subtle, collective domain motions necessary to align reactive moieties around the nascent base pair in preparation for rapid nucleotidyltransfer. Consistent with experimental results, these motions should occur whether or not phosphodiester bond formation is prohibited by a terminating primer[10] and may be reflected in the fine adjustments of loops, secondary structural elements, and amino acid side chains near the nucleotide binding pocket as demonstrated through comparison of the binary and ternary crystal structures of Dpo4.[149] Experiments are currently underway to measure the reverse isomerization rate (k_{-4}, Scheme 1.1A) of Dpo4 in order to distinguish between the two competing mechanisms (see Sections 1.4.4.2 and 1.4.4.3) and identify how Dpo4 selects for the correct nucleotide. Future work will also

determine how an incorrect nucleotide may affect the conformational dynamics of Dpo4. Importantly, an additional mechanism for nucleotide specificity by Dpo4 is based on hydrogen-deuterium exchange experiments suggesting that the DNA translocation step may be involved in correct dNTP selection as certain protein motions, speculated to occur during DNA translocation, are only observable in the presence of correct nucleotide.[221] Thus, correct nucleotide binding may stabilize the insertion state relative to incorrect nucleotide by slowing down Dpo4 reverse transition to the pre-insertion state. Similar to the latter model (Scheme 1.1B) for polymerase fidelity (see Section 1.4.4.2),[27,112] this suggests that nucleotide specificity hinges on a reverse step (*i.e.* reverse translocation, k_{-2}) being slow for correct dNTP, but fast for incorrect dNTP, relative to nucleotidyltransfer (Step 5, Scheme 1.1B). This model is supported by single-molecule FRET studies of Dpo4 showing that the correct nucleotide stabilizes the insertion state to a greater extent than incorrect nucleotide.[9,163]

Taken together, it is clear that the mechanisms by which DNA polymerases attain nucleotide specificity are complex and may vary significantly among the polymerase families. As a result, an overarching or unified mechanism to explain these intricate processes for DNA polymerases is likely not possible and we caution that what may appear true for one polymerase may not extend to all. A clear example of polymerase dependent selection mechanisms comes from our recent structural characterization of the inherent *D*-stereoselectivity of several DNA polymerases.[87,222,223] Through structures of the Y-family DNA polymerase Dpo4[87] and the X-family DNA polymerases hPolβ[222] and hPolλ[223] bound to various nucleotide analogs with *L*-stereochemistry we identified several

unique mechanisms by which these polymerases achieve *D*-stereoselectivity. While it was unsurprising that the Y-family polymerase Dpo4 and the X-family polymerases would not have common mechanisms of *D*-stereoselectivity, it was unexpected that hPolβ and hPolλ, which share a high amount of sequence and structural homology, select against *L*-nucleotides in different ways.[87,222,223] Thus, these studies highlight the difficulties in generating a unified mechanism for any aspect of DNA polymerase catalysis and support the necessity to study each polymerase individually.

1.4.5 Post-chemistry steps of nucleotide incorporation

Many biochemical, biophysical, and structural studies have aimed to deduce the kinetic mechanism and molecular bases for single-nucleotide incorporation and polymerase fidelity through characterization of the steps up to and including nucleotidyltransfer (Steps 1-6, Scheme 1.1A). However, post-chemistry steps involving the reverse (Steps 7 and 8, Scheme 1.1A) of the conformational changes observed during nucleotide binding and incorporation (Steps 4 and 5, Scheme 1.1A), as well as PP_i release (Step 9, Scheme 1.1A) have been seldom examined biochemically and/or structurally. Indeed, isolating post-chemistry events has proven to be difficult leading to the lack of sufficient structural and mechanistic characterization. As a consequence, the order in which PP_i release and the post-chemistry conformational changes occur as well as whether or not the events are cooperative (*i.e.* PP_i release triggers the reverse conformational change or vice versa) is unknown. However, recently the slow incorporation of nucleotide analogs, which closely resemble natural nucleotides but possess *L*-stereochemistry, has been utilized to capture *in crystallo* snapshots of post-

chemistry events by hPolβ.[43] In performing time-resolved X-ray crystallography (see Section 1.5.1) with these analogs, the order of events following the chemistry step were unambiguously defined. Interestingly, hPolβ completed the closed→open conformational changes (Steps 7 and 8, Scheme 1.1A) while the product PP_i remained bound to the polymerase active site. In fact, many of the side chain interactions with the PP_i were maintained despite the domain rearrangement, causing the PP_i to move with the thumb domain away from the incorporated nucleotide during the closed→open conformational transition. Presumably this reopening and movement of PP_i away from the reaction center would facilitate PP_i solvation and dissociation. Surprisingly, the third divalent metal ion previously identified in several time-resolved structural investigations[38-42] had already dissociated following domain reopening thereby directly opposing the hypothesis that the third divalent metal ion plays a role in PP_i dissociation (see Sections 1.2 and 1.5.3). Consistently, recent time-resolved X-ray crystallographic experiments with hPolμ in the presence of Mn^{2+} revealed that the third divalent metal ion dissociates prior to PP_i release and surprisingly showed that the B-site metal ion remains bound following PP_i release, rather than concomitantly dissociating with PP_i as previously purported.[46] Moreover, the post-chemistry structures of hPolβ demonstrate that the next correct nucleotide can bind to the open polymerase conformation to aid PP_i dissociation. This is not unexpected considering that if PP_i were to remain bound at the active site, then the incorporated nucleotide could be removed via pyrophosphorolysis. In this instance, one would expect concerted post-catalytic events including PP_i release, DNA translocation, and dNTP binding (*i.e.* Steps 9, 10_a, and 3, Scheme 1.1A). Altogether, rapid domain opening and the

active displacement of PP_i by the incoming nucleotide ensures forward reaction efficiency during processive DNA synthesis.

In a recent study, PP_i mimetic analogs were used to follow the reverse reaction by time-resolved crystallography (see Section 1.5.1).[224] Consistent with the abovementioned structural findings that PP_i dissociation occurs after opening of the thumb domain,[43] this study demonstrated that during pyrophosphorolysis, PP_i binds to the open form of hPolβ and an open→closed conformational change occurs prior to the reaction.[224] Moreover, structures inform that PP_i fails to support binding of catalytic Mg^{2+} at the A-site and is too far from the reaction center to promote efficient pyrophosphorolysis.[40] Consistently, biochemical and structural analyses with an imidodiphosphate PP_i analog demonstrated that a single atom change (*i.e.* bridging oxygen of PP_i substituted to nitrogen) allows optimal binding of catalytic M_A and positions the analog for efficient catalysis.[224] Interestingly, neither PP_i nor imidodiphosphate was efficient at removing mismatched primer termini, suggesting that pyrophosphorolysis does not act as a fidelity checkpoint during DNA synthesis.[224] Together, these studies have dissected the post-chemistry events of DNA polymerization and have shown that domain reopening occurs prior to PP_i release and the reverse reaction is highly disfavored.

1.5 New Paradigm for DNA Synthesis Catalyzed by DNA Polymerases

1.5.1 Time-resolved X-ray crystallography of DNA polymerase-catalyzed DNA synthesis

The ability to follow an enzymatic reaction at atomic resolution has been sought after for many years by biochemists and structural biologists.[225] Static crystal structures

of complexes carefully designed to mimic reactant-, intermediate-, and product-states can, at best, only offer hints of the actual reaction mechanism. With the advent of time-resolved X-ray crystallography (Figure 1.5), much of the ambiguity that accompanies the interpretation of static crystal structures is replaced with clear insight into the chemical mechanism of a particular enzyme-catalyzed reaction. Generally, the technique involves preparation and isolation of a crystal containing an enzyme-substrate complex in a pre-catalytic state, followed by reaction initiation by transferring the crystal to a solution containing the reaction activator(s) and cryo-protectant (Figure 1.5A). Next, the reaction is allowed to proceed for a defined time interval before it is freeze-quenched by transferring the crystal to liquid N_2 (Figure 1.5A) for subsequent diffraction experiments. During generation of a structural model, the proportions (*i.e.* occupancies) of two or more states (*i.e.* reactant and product) are modeled and refined to fit the diffraction data (Figure 1.6). It is important to note that time-resolved crystallography is not a single-molecule technique. Rather, the average behavior of numerous molecules within the crystal, in the reactant- or product-states, contributes to the electron density at each time point. This process is repeated for several crystals, each allowed to react for an increasing amount of time. Finally, after solving the structure of the pre-catalytic complex (zero time point), the reaction progress from beginning to the end can be visualized, with each structure of a particular time point serving as a frame in the reaction film (Figure 1.7). For example, recent time-resolved crystallography with DNA polymerases (Figure 1.5) has been successful in following single-nucleotide incorporation[38-46] with an example of the modeling procedure applied during time-resolved crystallography of nucleotidyltransfer

by hPolβ shown in Figure 1.6 and example of the time-resolved snapshots captured by this technique depicted in Figure 1.7. Mechanistic events were able to be temporally resolved *in crystallo*, as reaction rates for single-nucleotide incorporation are observed to be 20- to 100-fold slower for hPolη[38] and hPolβ,[40] compared to rates measured by pre-steady-state kinetic studies of these enzymes in solution.[38,104,226] To obtain these structures, the non-catalytic divalent metal ion Ca^{2+} was exploited to form a stable pre-catalytic complex (Figure 1.5B). Ca^{2+} was then exchanged for the catalytic divalent metal ions Mg^{2+} or Mn^{2+} to start the reaction (Figure 1.5B). Other unique properties of DNA polymerases made implementation of time-resolved X-ray crystallography successful including the relative ease of crystallizing pre-catalytic complexes, the ability to achieve high resolution diffraction data, and the limited impact of conformational heterogeneity or dynamics on crystal integrity during reaction progression.[225] While still a relatively new method (*e.g.* the first reports for DNA polymerases appeared only ~five years ago), researchers have already enjoyed success in utilizing the time-resolved structural technique to uncover new details of structure and function relationships of DNA polymerases. In the coming years, we expect that more details of the DNA polymerase mechanism, extending beyond just the X- and Y-families, will be uncovered as more investigators adapt this powerful methodology.

1.5.2 A third divalent metal ion during nucleotide incorporation

Despite hundreds of structures of DNA polymerases accumulated through years in the Protein Data Bank, no evidence for a divalent metal ion directly involved in DNA synthesis beyond the well documented M_A and M_B divalent metal ions had been

observed. However, the first time-resolved X-ray crystallographic investigation of hPolη,[38] as well as seven additional studies identify a transient, third divalent metal ion, referred to as the C-site (M_c) or product-associated metal ion, during phosphodiester bond formation (Figure 1.3).[39-46] Thus, traditional pre-catalytic substrate complexes and post-catalytic product complexes of DNA polymerases were not sufficient to structurally capture this apparently dynamic third divalent metal ion. High-resolution diffraction data (~1.5-2.0 Å) permitted for the unambiguous identification of metal ion electron density, coordination geometry, and metal ion-to-ligand coordination distances for M_C (Figure 1.8). Appropriately, M_C was shown to exhibit octahedral coordination geometry, consistent with a bound divalent metal ion, and short metal-to-ligand coordination distances (~2.2 Å), consistent with Mg^{2+} or Mn^{2+}, rather than non-catalytic Ca^{2+} (~2.4 Å). In some cases,[39-42] Mn^{2+} was used to initiate the *in crystallo* reaction, rather than Mg^{2+}, as its stronger signal (*i.e.* Mn^{2+} is more electron rich) allowed for confident assignment of M_C electron density even at 5σ levels (*i.e.* five standard deviations above background) (Figure 1.8D)[40,42] or resulted in anomalous diffraction.[40] M_C was shown to coordinate four water molecules as well as the non-bridging oxygen atom of the α-phosphate and the leaving oxygen atom of the β-phosphate (bridging oxygen between α- and β-phosphates) of the bound nucleotide (Figure 1.1B). In some instances, the number of ligands bound to M_C varied from four to six due to the dynamic nature of coordinating water molecules and presumably the transient nature of M_C. Importantly, these ligands fail to form any protein contacts, but only coordinate to the metal ion, therefore preventing any mutational confirmation of the existence or significance of M_C. Differences in the timing and

occupancy of the third divalent metal ion amongst the time-resolved studies of hPolη,[38,39] hPolβ,[40-42,44,45] and hPolμ[46] are suggestive of its dynamic nature. Furthermore, a third divalent metal ion only appeared with hPolμ when Mn^{2+} was used for metal ion exchange and not with Mg^{2+}.[46] As a consequence of these inconsistencies, the role of the third divalent metal ion in the mechanism of nucleotide incorporation has been highly debated (Figures 1.1B and 1.1C) with hypothesized roles in transition-state stabilization, product release, catalysis of pyrophosphorolysis, or product-state stabilization.[46,47]

1.5.3 Evidence and hypothesized roles for the third divalent metal ion in single-nucleotide incorporation catalyzed by hPolη

In the inceptive time-resolved crystallography study, Nakamura *et al.*[38] followed nucleotide incorporation by hPolη and identified for the first time a third divalent metal ion utilized by a DNA polymerase during catalysis. M_C appeared midway through phosphodiester bond formation (140 s, 60% reactants, 40% products) and remained associated at the active site until the final recorded time point (230 s, 40% reactants, 60% product). Unfortunately, observation of full product and the subsequent release of PP_i was not observed as both the forward DNA synthesis and reverse pyrophosphorolysis reactions became competing at later time points as product occupancy decreased from the penultimate to the final time point. Furthermore, as the competing reactions were occurring simultaneously, it is likely that both activities were aided by M_C through transition-state stabilization and lowering the activation energy barrier for bond formation. Interestingly, appearance of M_C occurred concomitantly with the movement of a positively charged arginine residue (R61), which had flipped away from the α-

phosphate of the bound dNTP, effectively replacing the charge in the active site. This active site configuration led to the hypothesis that the M_C would support chemistry while the subsequent reverse conformational transition of the arginine, or an equivalent positively charged residue in other polymerases, back to its pre-catalytic configuration would act in concert with the third divalent metal ion to actively displace product PP_i. In this way, the side chain is essentially "sweeping out" PP_i from the active site in preparation for DNA translocation and an additional catalytic cycle. Consistently, a molecular dynamics investigation of the pre-catalytic and reaction state side chain conformations of R61 concluded that only the pre-catalytic side chain configuration facilitates nucleotide binding, suggesting that the absence of nucleotide precludes reaction-state configuration. Moreover, following nucleotidyltransfer the side chain must revert back to the pre-catalytic configuration before subsequent rounds of dNTP binding and incorporation.[49] Furthermore, the computational study maintained the notion that M_C, accompanied by the R61 conformational transition, serves as an exit shuttle for PP_i release, and dissociates along with PP_i.[49]

This mechanism of third divalent metal ion-dependent conformational transitions for nucleotide incorporation and product release may be conserved given that a lysine side chain in many polymerases from diverse families,[18,64,68,164,227,228] or an arginine residue in HIV reverse transcriptase,[88] is present at a position similar to R61 of hPolη. Alternatively, this positively charged side chain may be static and therefore occlude the binding of a third divalent metal ion and thus fulfil the roles of transition-state stabilization and/or PP_i release. Hence, it is necessary to investigate these other

polymerase families or RTs through similar methods to determine the precise role of these amino acids and identify whether a third divalent metal ion is utilized. Notably, the X-family repair polymerases, hPolβ, hPolλ, and hPolμ, which typically act on single-nucleotide gapped DNA substrates (see Section 1.4.1), do not possess an analogous positively charged residue to interact with the α-phosphate. This may be a mechanism by which processive DNA synthesis is suppressed to prevent potential misincorporations by these moderate-fidelity enzymes.[79,146-148] While the Y-family polymerases, such as hPolη, are considered low-fidelity enzymes[14,147,229-231] (see Section 1.4.3), the necessity to bypass DNA damage and subsequently extend the DNA primer, which are both difficult tasks for high-fidelity replicative polymerases, may justify the need for the positively charged side chain to aid in processivity, in contrast to the X-family polymerases.

More recently a time-resolved crystallography study of hPolη directly investigated the role of M_C and suggested that it is absolutely essential for catalysis and its binding may kinetically limit the rate of single-nucleotide incorporation.[39] When crystals of the hPolη pre-catalytic ternary complex (E•DNA•dNTP•Ca^{2+}) were soaked with 1 mM Mn^{2+} for varying amounts of time, only the A- and B-site metal ions were occupied by Mn^{2+} and no product formation was detected even after 1,800 s. However, after soaking the crystals in 10 mM Mn^{2+}, the A- and B- sites were readily occupied, while M_C appeared at 30 s coinciding precisely with the appearance of reaction product. The authors were able to determine the relative binding affinities for each metal ion from *in crystallo* metal ion titration experiments, where electron density for each metal ion was examined following crystal soaking at various metal ion concentrations and time points.

It was determined that Mn^{2+} binds to the C-site with an affinity of ~3.2 mM. The apparent weak binding of M_C was further supported by in-solution metal ion titration experiments, wherein the concentration of metal ion necessary to achieve half-maximal reaction rate was determined and yielded a similar affinity (2.7 mM). Importantly, the agreement of the metal ion binding affinities from the two approaches (*i.e.* in-solution and *in crystallo* metal ion titrations) is the first experimental evidence to suggest that M_C is bound by hPol*η* in solution and is not simply a crystal artifact of the time-resolved crystallography technique.

Examination of the interaction of hPol*η* with S_p-dATP*α*S demonstrated that A- and C- site metal ion binding is affected as a direct result of the substitution of the pro-S_p oxygen atom of dATP with a sulfur atom. As the atom at the S_p position is expected to coordinate M_C, the larger atomic radius of sulfur relative to oxygen disrupts the binding and therefore likely explains the observed rate reduction *in crystallo*. In fact, M_C is not observed at all despite product formation (50% product at 600 s), which suggests that M_C may not be absolutely essential for catalysis. However, the authors argue that M_C must be present but is too transient or low occupancy to be observed in the electron density.[39] In addition, mutation of active site residue (R61) of hPol*η* to alanine, resulted in delayed third divalent metal ion binding and misalignment of the bound dNTP relative to the primer 3′-OH. This is somewhat unexpected considering that R61 in the pre-catalytic ternary crystal structures occupies the space where M_C would bind following catalytic metal ion exchange. Based on this single pre-catalytic rotameric conformation, we expect that the exchange between pre- and post-catalytic side chain configurations of R61 is

slow and may partially limit the rate of M_C binding. Therefore, it is reasonable to hypothesize that mutation of R61 to alanine would facilitate more efficient binding of M_C. However, considering that R61 makes important contacts with both the α- and β-phosphates of the incoming dNTP, and the primer 3′-OH is misaligned relative to wild-type structures, their results may simply reflect the negative impact of the mutation on nucleotide binding as suggested by the significant increase (~2-10-fold) in K_m compared to the wild-type enzyme in steady-state kinetic experiments.[80]

Altogether, the authors use these data to suggest that chemistry (Step 6, Scheme 1.1A) is indeed rate-limiting for hPolη, which contrasts with previous kinetic evidence,[103,104] and that M_C binding may provide the free energy needed to overcome the activation energy barrier for nucleotidyltransfer (*i.e.* transition-state stabilization). However, it is important to consider that hPolη is first crystallized in a ternary complex in the presence of non-catalytic Ca^{2+} and therefore only the bond forming chemistry step is observed. Substrate binding (Step 3, Scheme 1.1A) and any associated conformational rate-limiting steps (Steps 4 and 5, Scheme 1.1A) have presumably already occurred. Thus, the reduction in reaction rate caused by the postponement or disruption to third divalent metal ion binding caused by the S_p-dATPαS or R61A mutation suggest that M_C binding limits the rate of the chemistry step, but likely is not rate-limiting for the entire kinetic pathway for single-nucleotide incorporation (Scheme 1.1A, see Section 1.4.4). Altogether, this work confirmed and measured the binding of M_C to hPolη at reasonable metal ion concentrations (in contrast to work with hPolβ, 200 mM Mg^{2+}/Mn^{2+})[40-42,45] and correlates the M_C binding affinities estimated crystallographically and in-solution. However, this

study falls short of unequivocally limiting the role of M_C to transition-state stabilization, as the temporal resolution afforded by the time-resolved crystallography technique[225] is not sufficient to distinguish the order of M_C binding and nucleotidyltransfer.

1.5.4 Evidence and hypothesized roles for the third divalent metal ion in single-nucleotide incorporation catalyzed by hPol*β*

Following the seminal study with hPol*η*,[38] time-resolved crystallography was employed to visualize nucleotide incorporation by hPol*β* on a gapped DNA substrate.[40] Interestingly, a third divalent metal ion at the C-site was also observed for hPol*β*[40] during correct nucleotide incorporation in a similar position to the third divalent metal ion in hPol*η* (Figure 1.3).[38] However, in contrast to the results obtained for hPol*η*,[38] M_C only appeared in the product complex structures and coincided with the loss of M_A. This observation suggested that M_C is solely involved in post-chemistry events and perhaps the diverse polymerase families utilize the third divalent metal ion in distinct ways. In fact, the timing of M_C binding suggested a role for M_C in pyrophosphorolysis, wherein it would stabilize the attacking oxygen atom of PP_i following proton abstraction by a water molecule (Figure 1.1C). Consistent with this hypothesis, open nicked DNA binary complex (E•DNA) crystals soaked with Mg^{2+} and PP_i yielded structures of the polymerase in the closed conformation with M_C and PP_i bound to the active site (*i.e.* reactant-state for reverse reaction, pyrophosphorolysis). However, these complexes failed to initiate pyrophosphorolysis, presumably because M_A was a Na^+ ion rather than the catalytic Mg^{2+}. Interestingly, M_C was not observed during incorrect nucleotide incorporation.[40]

The role of M_C in pyrophosphorolysis was further investigated through use of quantum mechanical/molecular mechanical computational methods.[50] It was determined that M_C was beneficial in the initial stages of the chemical reaction (*i.e.* initiating the attack of Pα by Oβ), but became inhibitory as the two reacting atoms (Pα and Oβ) approached a distance of 2.3 Å, likely due to the strict coordination distances and geometry of M_C, thus effectively preventing the transition-state from forming.[50] However, replacement of Mg^{2+} with Na^+ at the C-site resulted in a lowered activation energy barrier, suggesting a mechanism where metal ions may exchange during the reaction pathway to favor reaction completion. In addition, the authors demonstrated that Mg^{2+} binding at the A-site is required for catalysis in either the forward or reverse (*i.e.* nucleotidyltransfer or pyrophosphorolysis) direction and rapid exchange with Na^+ following catalysis effectively pushes the reaction to completion. The inability for Mg^{2+} at the C-site to support the reverse reaction but the requirement of Mg^{2+} at the A-site for forward and reverse catalysis is consistent with the hypothesis that M_C may only be involved in post-chemistry events (*i.e.* PP_i release or conformational changes) for hPolβ and explains the inability of PP_i to support pyrophosphorolysis *in crystallo*.[43,50]

In subsequent time-resolved crystallographic investigations of hPolβ during faithful and unfaithful translesion DNA synthesis across from the major oxidative lesion 8-oxo-7,8-dihydro-2′-deoxyguanine (8-oxoG) (Figure 1.7),[42] M_C was observed in the reaction-state (Figures 1.8A and 1.8B) and post-chemistry (Figure 1.8C and 1.8D) structures with its occupancy similar or equivalent to the product-state occupancy. Notably, as the reaction progressed, M_C moved towards its final position wherein it was

fully coordinated with the reaction products and water molecules during bypass of 8-oxoG with dATP, highlighting the dynamic nature of the third divalent metal ion. This observation is similar to hPolη wherein M_C binding occurs prior to or immediately following reaction initiation,[38,39] and suggests that M_C may diffuse into the hPolβ active site following reaction initiation to associate with reaction intermediates as a means to stabilize the transition-state. This result was further supported in a follow-up time-resolved crystallographic study following hPolβ-catalyzed extension from 8-oxoG containing base pairs which were generated during the bypass[45] (*i.e.* dC:8-oxoG or dA:8-oxoG was the primer-template junction pair). Importantly, during this study,[45] M_C was observed as early as 15% product formation, suggesting an early role in catalysis (*i.e.* transition-state stabilization). Furthermore, the investigation of post-chemistry events for hPolβ-catalyzed nucleotide incorporation, revealed that M_C is not involved in PP_i release (see Section 1.4.5), thus limiting the third divalent metal ion to function in chemical events such as transition-state stabilization or product-state stabilization.

Additional time-resolved studies following the incorporation of 8-oxo-dGTP opposite a template dC or dA[41] or the extension from these incorporation products (*i.e.* 8-oxoG:dC or 8-oxoG:dA were the primer-template junction pair)[44] similarly demonstrated that M_C appears in reaction- and product-state structures. Unexpectedly, the third divalent metal ion appeared in the pre-catalytic structure for incorporation of 8-oxo-dGTP opposite dC presumably as a result of favorable M_C coordination facilitated by the optimal position of the O8 modification of the damaged nucleotide. The inclusion of a third coordinating ligand offered by the O8 atom, in addition to the non-bridging oxygen

atoms on the α- and β-phosphates, likely makes M_C binding more favorable in this structure.[41] These damage-specific interactions coupled with the absence of M_C in all other pre-catalytic structures from time-resolved crystallographic investigations of hPolβ suggest that pre-catalytic M_C binding is likely unique to this damage DNA context and therefore does not represent a common mechanistic feature. Nevertheless, despite the appearance of M_C in reaction intermediate structures, the authors argue that the third divalent metal ion is only involved in post-chemistry events and does not provide transition-state stabilization.[41,44]

To further investigate the role of M_C in the forward nucleotidyltransfer reaction a similar computational investigation as that completed for hPolβ-catalyzed pyrophosphorolysis[50] was performed.[51] As coordination of M_C by the bridging oxygen (O$\alpha\beta$) between the Pα and Pβ of the incoming nucleotide can only occur after the phosphodiester bond (*i.e.* Pα–O$\alpha\beta$) is broken, molecular dynamics was used to determine the position of the Mg^{2+} prior to nucleotidyltransfer. The calculated position of modeled pre-catalytic M_C was similar to that experimentally observed for M_C. However, the modeled pre-catalytic M_C is coordinated by the Pα pro-S_p oxygen of the incoming nucleotide and five water molecules, rather than the experimentally observed coordination by the Pα pro-S_p oxygen of the incorporated nucleotide, O$\alpha\beta$ of PP_i, and four water molecules. Within this system, the simulated activation energy barrier for nucleotidyltransfer was calculated to be 16.6 to 18.1 kcal/mol and, except for a slight repositioning of the product PP_i, very few differences in active site structure were observed relative to the time-resolved studies.[40-42,44,45] Interestingly, a two Mg^{2+} system

in which M_C was omitted and only M_A and M_B were used gave a very similar activation energy barrier of 17.5 to 18.6 kcal/mol, suggesting that M_C does not appreciably aid nucleotidyltransfer. Consistent with these computational predictions, the incorporation of a phosphorothioate nucleotide analog, S_p-dCTPαS, in which the sulfur substitution should ablate M_C binding, was only 3-fold slower than incorporation of dCTP.[146] Moreover, time-resolved crystallography of dCTPαS incorporation did not reveal the presence of a third divalent metal ion following nucleotide incorporation.[51] Taken together, these evidences suggest that the third divalent metal ion does not aid the forward reaction by significantly lowering the activation barrier,[50,51] which is in contrast to hPolη.[39] Alternatively, in a two Mg^{2+} (M_A and M_B) and one Na^+ (M_C) system the activation energy barrier is significantly lowered to 11.6 to 13.2 kcal/mol, suggesting a possible mechanism wherein a Na^+ is initially bound at the C-site to assist nucleotidyltransfer and is subsequently exchanged with Mg^{2+} following incorporation in order to prevent pyrophosphorolysis. This hypothesized metal ion exchange at the C-site is akin to the structurally observed exchange of the A-site Mg^{2+} for Na^+ following nucleotidyltransfer,[40-42,44,45] which also prevents pyrophosphorolysis. These metal ion exchanges within the hPolβ active site may act to favor nucleotidyltransfer while disfavoring pyrophosphorolysis.

While the first time-resolved structural study of hPolβ documenting a third divalent metal ion showed the appearance of M_C only after full product formation,[40] the latter four investigations showed M_C binding coinciding exactly with product formation,[41,42,44,45] which is consistent with the reports of the third divalent metal ion for

hPolη.[38,39] Accordingly, it is unclear for hPolβ if M_C binds prior to and supports nucleotidyltransfer, or binds following nucleotidyltransfer to stabilize the product complex (*i.e.* preventing pyrophosphorolysis).[50] Altogether, conflicting results from the time-resolved[40-42,44,45] and computational studies,[50,51] as well as the aforementioned investigation of hPolβ post-chemistry events[43] (see Section 1.4.5) limits the potential roles of M_C to either transition-state stabilization during nucleotidyltransfer (Figure 1.1B) or suppression of pyrophosphorolysis through stabilization of the product complex. However, similar concentration dependent divalent metal ion soaks to those with hPolη[39] to determine the effect of M_C on enzymatic rate are necessary to experimentally demonstrate whether or not M_C aids nucleotidyltransfer through transition-state stabilization.

1.5.5 Evidence and hypothesized roles for the third divalent metal ion in single-nucleotide incorporation catalyzed by hPolμ

Similar to studies with hPolβ,[40-45] time-resolved crystallography was used to follow nucleotide incorporation into single-nucleotide gapped DNA by hPolμ,[46] wherein the pre-catalytic ternary complex formed in the presence of non-catalytic Ca^{2+} was soaked with either Mg^{2+} or Mn^{2+} to initiate metal ion exchange and catalysis. In contrast to the time-resolved structural findings with hPolβ[40-42,44,45] and hPolη,[38,39] a third divalent metal ion bound at the C-site could not be observed with Mg^{2+} even after extensive soaking at a high concentration of Mg^{2+} (100 mM) and despite full product formation. As expected, the Mg^{2+} bound at the A-site was eventually replaced by Na^+ at longer time points showing complete product formation. On the other hand, soaking with Mn^{2+}

resulted in appearance of M_C at time points coincident with 40% product formation and beyond, as well as sustained presence of Mn^{2+} at the A- and B-sites at every time point therefore suggesting that Mn^{2+} may be the physiological catalytic divalent metal ion for hPolμ and perhaps other DNA polymerases that exhibit low activity with Mg^{2+}. Notably, the position of M_C and the coordination of M_C by the reaction products, as well as the timing of M_C, are consistent with that observed for hPolβ[40-45] and hPolη.[38,39] Based on the computational work with hPolβ,[50,51] occupancy of the A-site by a divalent metal ion is essential for pyrophosphorolysis whereas the presence of M_C is inhibitory (see Section 1.5.4). As hPolβ and hPolμ share significant structural similarity, the roles of M_A and M_C for the reverse reaction may be conserved between these two polymerases. Accordingly, by analogy to hPolβ (see Section 1.5.4), the A-site metal ion exchange to Na^+ in the presence of Mg^{2+} observed with hPolμ likely precludes pyrophosphorolysis and therefore a third divalent metal ion may not be necessary to suppress this reaction when Mg^{2+} is supplied as the catalytic metal ion. However, the persistence of M_A in the presence of Mn^{2+} necessitates binding of M_C to prevent pyrophosphorolysis.

Single-nucleotide incorporation experiments with S_p-dTTPαS (see Section 1.4.4.1) suggest that M_C could serve a role in nucleotidyltransfer as experiments in the presence of Mg^{2+}, wherein M_C should not be bound, demonstrated a strong elemental effect, whereas this effect was lost in the presence of Mn^{2+} and presumably M_C.[46] However, in their publication, the authors suggest that absence of an elemental effect with Mn^{2+} was not due to binding of M_C as i) C-site divalent metal ion binding occurred following nucleotidyltransfer in the time-dependent structures (*i.e.* occupancy

corresponded exactly with product accumulation), ii) Mn^{2+} is generally considered thiophobic,[232] and iii) the sulfur substitution of the pro-S_p oxygen would likely disrupt M_C coordination. Thus, similar to their work with hPolβ (see Section 1.5.4),[51] the authors suggest that M_C is not involved in transition-state stabilization during nucleotidyltransfer but rather serves to stabilize the product state to prevent pyrophosphorolysis. Nevertheless, the simultaneous appearance of M_C and reaction products could just as easily imply that the third divalent metal ion is critical for nucleotidyltransfer. Moreover, the expected disruption of M_C binding by the longer P–S bond distance would likely be alleviated by the flexibility of the other coordinating ligands, as four of six are water molecules. Fittingly, it is possible that M_C provides the necessary transition-state stabilization to accelerate the chemistry step in the presence of Mn^{2+} thereby eliminating the elemental effect. In this scenario, the results would be consistent with those of hPolη (see Section 1.5.3).[39]

Similar to our recent investigation of hPolβ-catalyzed post-chemistry events,[43] during time-dependent crystallography of hPolμ, M_C was also observed to dissociate before PP_i.[46] While hPolβ displays a large open→closed conformational change of the thumb domain (Figure 1.2) during nucleotide binding (see Section 1.4.4), such a large change is not observed for hPolμ, which may explain why time-resolved structural capture of the order of M_C and PP_i dissociation was difficult for hPolβ (*i.e.* rapid domain motion of hPolβ results in loss of synchronization of *in crystallo* events and associated electron density, see Section 1.4.5), but readily possible for hPolμ. Altogether, the time-resolved structural study of hPolμ,[46] featuring the third divalent metal ion, parallel those of

hPolβ[40-45] and suggest a conserved role for M_C during X-family polymerase-catalyzed DNA synthesis. However, more work is needed to explicitly delineate the mechanistic function of M_C in transition-state stabilization or preventing pyrophosphorolysis.

1.5.6 Future characterization of the third divalent metal ion

The role of M_C in single-nucleotide incorporation is not well-defined with compelling evidence to support its involvement in i) stabilizing the transition-state of nucleotidyltransfer, ii) supporting pyrophosphate release, and/or iii) promoting or suppressing pyrophosphorolysis. It is clear from the limited work investigating the third divalent metal ion that its function in DNA polymerase catalysis is complex and may differ between the X- and Y-family polymerases. For example, the complementary time-resolved crystallographic[40-45] and computational studies[50,51] completed for hPolβ (see Section 1.5.4) support the proposed roles of M_C and provide an argument for analogous functions in hPolμ (see Section 1.5.5).[46] However, a potential role in transition-state stabilization for M_C in the X-family DNA polymerases cannot be completely ruled out. Conversely, for hPolη, it has been proposed that M_C is directly involved in transition-state stabilization during nucleotide incorporation.[39,198] Nevertheless, additional work must be completed with other Y-family polymerases to validate the proposed roles of M_C suggested for hPolη (see Section 1.5.3) and to determine if M_C function is conserved for the Y-family polymerases. In addition, computational investigations, such as those performed for hPolβ,[50,51] must be undertaken for hPolμ and hPolη to better substantiate the proposed roles of M_C in these polymerases. Similarly, concentration dependent metal ion soaking as performed for hPolη[39] must be performed for hPolβ and hPolμ to

determine if M_C also acts in transition-state stabilization for these polymerases as argued for hPolη,[39,198] especially considering that high metal ion concentrations of ~200 mM were used for the studies of hPolβ[40-45] and could negatively affect polymerase activity.[93,109,233,234]

Furthermore, the apparent differences between how M_C is utilized between the X- and Y-family polymerases advocates for future research on the A- or B-family replicative polymerases as well as RTs to determine if a third divalent metal ion is used at all, and if so, what apparent role does it serve, and how does this compare to results of hPolη,[38,39] hPolβ,[40-45] and hPolμ.[46] Importantly, if a third divalent metal ion is observed for viral DNA polymerases or RTs and serves a purpose in catalysis (*i.e.* transition-state stabilization as with hPolη), then it may be a potential therapeutic target. For example, the active site of HIV-1 RT is very comparable to that of hPolη and contains an equivalent arginine residue (see Section 1.5.3) that may function similarly with the third divalent metal ion to facilitate nucleotidyltransfer and pyrophosphorolysis (Figure 1.3). As HIV-1 RT is known to remove chain-terminating nucleotide analogs by pyrophosphorolysis,[235] design of antiviral small molecules to specifically block the third divalent metal ion binding may prove to be an effective treatment strategy. Finally, as it stands, the only experimental evidence for the third divalent metal ion comes from X-ray structures capturing *in crystallo* reaction progression. Accordingly, it is possible that these findings may represent an artifact of the structural technique. Therefore, we expect that advanced spectroscopic methods such as electroparamagnetic resonance (EPR) spectroscopy,[236] will be necessary to fully validate and further elucidate the function of

the third divalent metal ion during polymerase catalysis under a more physiological context.

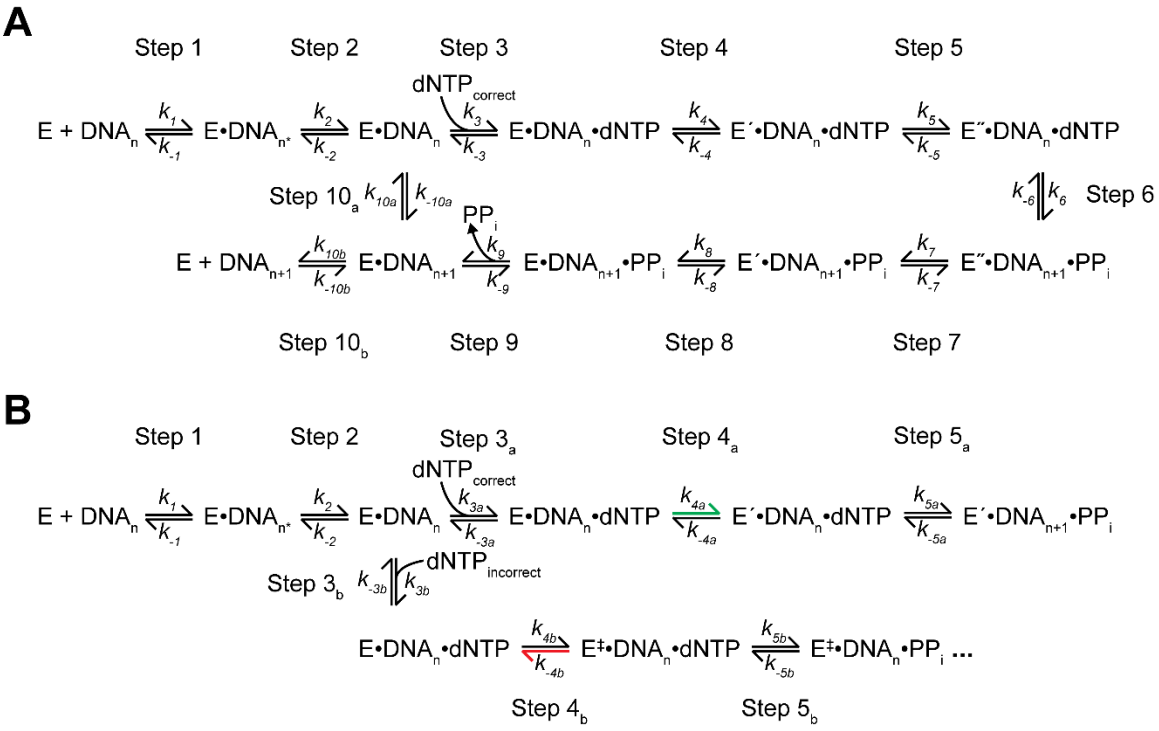
1.6 Concluding Remarks

Despite thousands of published studies investigating the structure and mechanism of DNA polymerases and RTs, it is abundantly clear that there is so much more to learn. Indeed, as the kinetics and conformational dynamics of each step of the DNA synthesis mechanism have been rigorously investigated for many polymerases from all of the diverse families, it is becoming evident that a singular, unified mechanism to describe every unique aspect of polymerase catalysis, including polymerase fidelity, is unrealistic. Thus far, it is evident that conformational dynamics differentially impact various aspects of the catalytic and kinetic mechanism between DNA polymerase and RT families, or even within a family. In the coming years, research to better understand the contributions of polymerase conformational dynamics during DNA binding and translocation, nucleotide binding, selectivity, and incorporation, pyrophosphate binding, and pyrophosphorolysis to the mechanism of DNA polymerization will be paramount.

The skillful application of time-resolved X-ray crystallography to study DNA polymerases has enabled the discovery of a third divalent metal ion during single-nucleotide incorporation. Remarkably, this third divalent metal ion may be important for DNA polymerase and RT catalysis which shifts the long-standing paradigm of two-metal-ion catalysis for DNA polymerization. We are excited at the prospect of identifying the third divalent metal ion in other families of DNA polymerases and RTs to determine if a three-metal-ion mechanism is conserved for DNA synthesis. To date, the third

divalent metal ion has only been captured in two X-family members and one Y-family member, but the significant differences in the proposed function of the third divalent metal ion, already apparent between these two families, implore future research of other polymerase families in hopes of exploiting potential drug targets for developing novel antiviral and antibiotic small molecule therapeutics.

1.7 Schemes



Scheme 1.1: Minimal kinetic Mechanism for nucleotide incorporation.

(A) Kinetic mechanism of nucleotide binding and incorporation with E, E′, and E″ representing different conformations of the DNA polymerase with Step 5 representing an essential, rate-limiting conformational change. (B) Alternative kinetic mechanism wherein incorrect nucleotide is selected against by binding in a unique DNA polymerase conformation designated by $E^{‡}$. Steps 4_a and 5_a occur during correct nucleotide incorporation. Steps 4_b and 5_b occur during incorrect nucleotide incorporation. The green arrow in Step 4_a signifies that the forward rate is highly favored in the presence of correct

(Continued)

Scheme 1.1: Continued

nucleotide, where E and E′ represent a conformational change upon nucleotide binding. In the bottom branch, the red arrow in Step 4_b indicates that the reverse rate is highly favored in the presence of incorrect nucleotide. Following Step 5 the mechanism proceeds as in (A) for both correct and incorrect nucleotides. For (A) and (B) DNA_n^* signifies that the polymerase is bound to the DNA at the pre-insertion site (*i.e.* pre-translocated state).

1.8 Figures

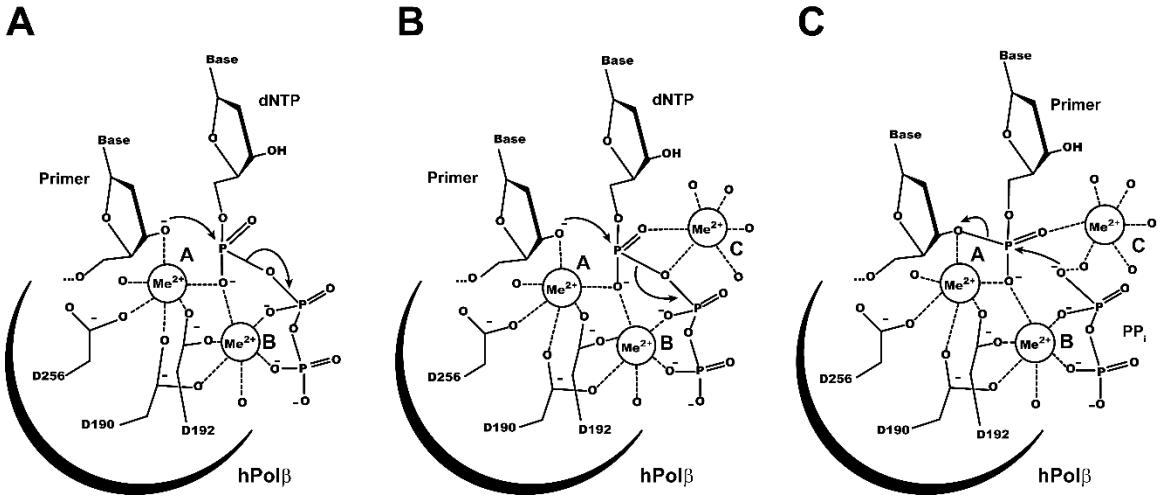


Figure 1.1: Two- vs. three-metal-ion mechanism for DNA polymerase-catalyzed nucleotidyltransfer and third-metal-ion assisted pyrophosphorolysis.

The active site of the well-studied hPolβ was selected to depict the metal ion-based chemical mechanisms. (A) Two-metal-ion mechanism. The 3′-OH of the primer is activated (*i.e.* deprotonated) for an in-line nucleophilic attack on the α-phosphate of the incoming dNTP. The α-phosphate is coordinated by two divalent metal ions (Me^{2+}). The catalytic metal ion at the A-site is also coordinated by the 3′-OH of the primer, active site carboxylate groups (Asp 190, 192, and 256), and a water molecule. The metal ion at the B-site is coordinated by active site carboxylates (Asp 190 and 192), a water molecule, and non-bridging oxygen atoms of the β- and γ-phosphates, to complete the α,β,γ-tridentate coordination of the dNTP. The A-site ion is suggested to activate the primer 3′-OH nucleophile and the B-site ion stabilizes the negative charge of the pentacoordinated

(Continued)

Figure 1.1: Continued

transition state. (B) Three-metal-ion mechanism. The reaction proceeds as in (A) except that a third divalent metal ion at the C-site appears to perhaps stabilize the transition state, serve as counter-ion to the oxyanion of the PP_i leaving group to aid product release, or participate in the reverse reaction, pyrophosphorolysis. The C-site ion is coordinated by water molecules as well as non-bridging oxygen atom of the α-phosphate and the bridging oxygen between α- and β-phosphates. (C) Third-metal-ion assisted pyrophosphorolysis. The third divalent metal ion may serve a similar role as the A-site metal ion in (A) and (B) to assist in the deprotonation and subsequent stabilization of the O_1 of PP_i. This atom would then attack the nascent phosphodiester bond of the DNA backbone, and the primer 3′-hydroxyl would be protonated to restore the pre-catalytic active site of nucleotide incorporation.

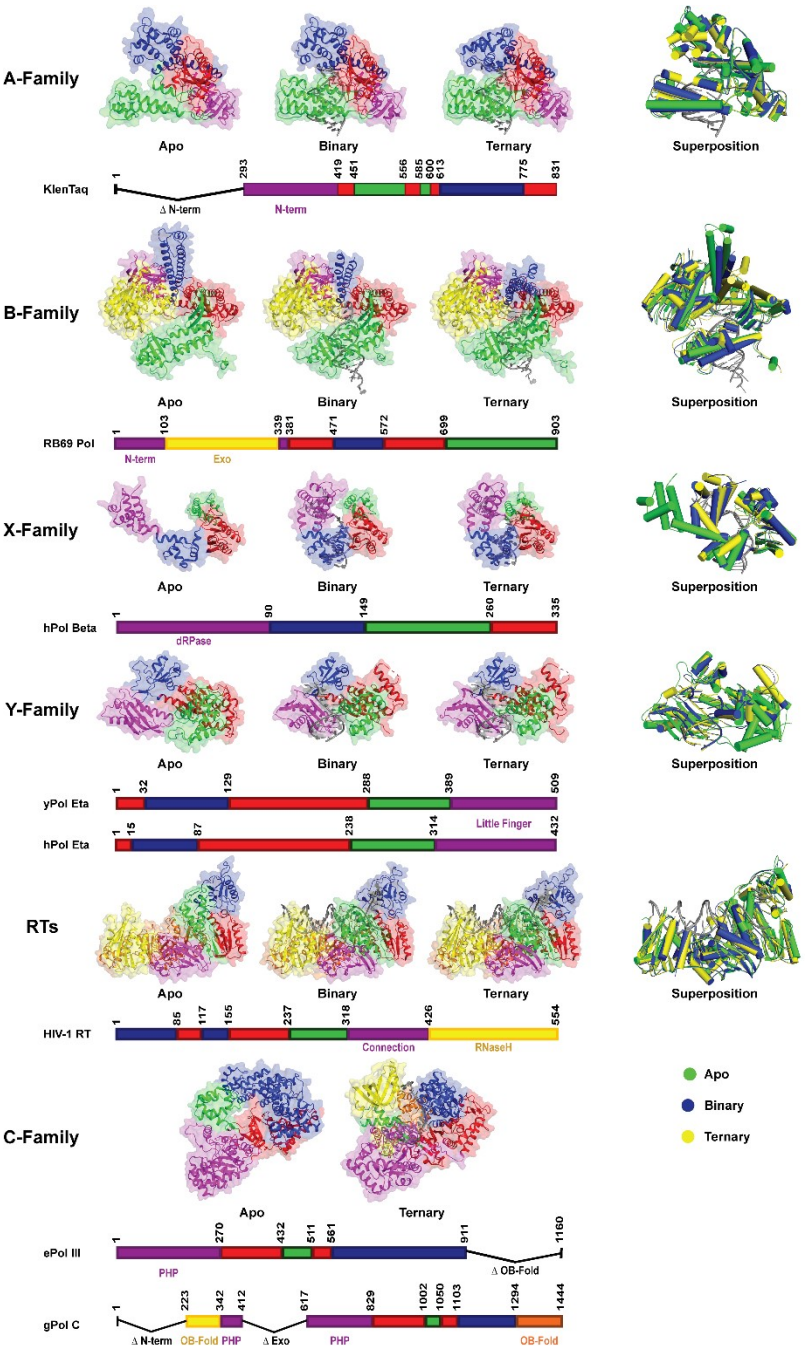


Figure 1.2: Structural comparison of DNA polymerase families.

(Continued)

Figure 1.2: Continued

binary (DNA bound, E•DNA), and ternary (DNA and nucleotide bound, E•DNA•dNTP) enzyme forms of representative polymerases from each family and superposition of all three forms. The Klenow fragment of Taq DNA polymerase I (KlenTaq) was used for A-family (1KTQ, 4KTQ, and 3KTQ), RB69 DNA polymerase (RB69 Pol) was used for B-family (1IH7, 2P5O, and 3NCI), rat DNA polymerase β (rPol Beta, apo) and human DNA polymerase β (hPol Beta, binary and ternary) were used for X-family (1BPD, 1BPX, and 4KLG), yeast DNA polymerase η (yPol Eta, apo) and human DNA polymerase η (hPol Eta, binary and ternary) were used for Y-family (1JIH, 3TQ1, 4ECX), HIV-1 reverse transcriptase (HIV-1 RT) was used for RTs (1DLO, 3KJV, 3KK2), and E. coli DNA polymerase III (ePol III, apo) and Geobacillus kaustophilus PolC (gPol C, ternary) were used for C-family (4JOM and 3F2D). Each structure is shown as cartoon with transparent surface rendering and individual domains colored. For all structures the thumb, palm, and finger domains are green, red, and blue, respectively. Accessory domains are uniquely colored and named in the associated line diagrams. For the binary and ternary structures, the DNA is shown as gray cartoon. In the ternary structures, the nucleotide is omitted for clarity. The superpositions are shown with cylindrical helices for simplicity of comparison with apo, binary, and ternary structures colored green, blue, and yellow, respectively.

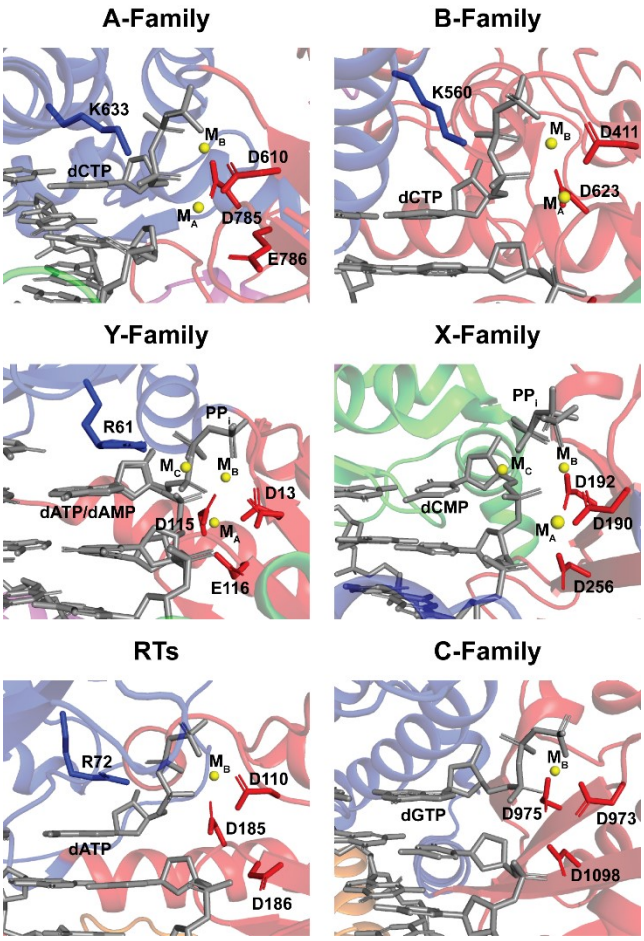


Figure 1.3: Active site comparison of DNA polymerases.

Zoomed views of ternary structures of representative DNA polymerases from the A-family (KlenTaq, 3KTQ), B-family (RB69 DNA polymerase, 3NCI), the X-family (hPolβ, 4KLG), the Y-family (hPolη, 4ECX), the RTs (HIV-1 RT, 3KK2), and the C-family (Geobacillus kaustophilus PolC, 3F2D). The incoming/incorporated nucleotide (dNTP/dNMP+PPi), DNA primer 3′-nucleotide, and active site carboxylates are shown as sticks. Metal ions bound at the active site are shown as yellow spheres. Importantly, in

(Continued)

Figure 1.3: Continued

addition to the typical A- and B-site metal ions (MA and MB), the X- and Y-family structures have a third divalent metal ion bound (MC). Many polymerases have positively charged residue side chains in the area where a third metal ion may bind and are shown as sticks in blue. Notably, Geobacillus kaustophilus PolC does not have a positively charged residue in this location.

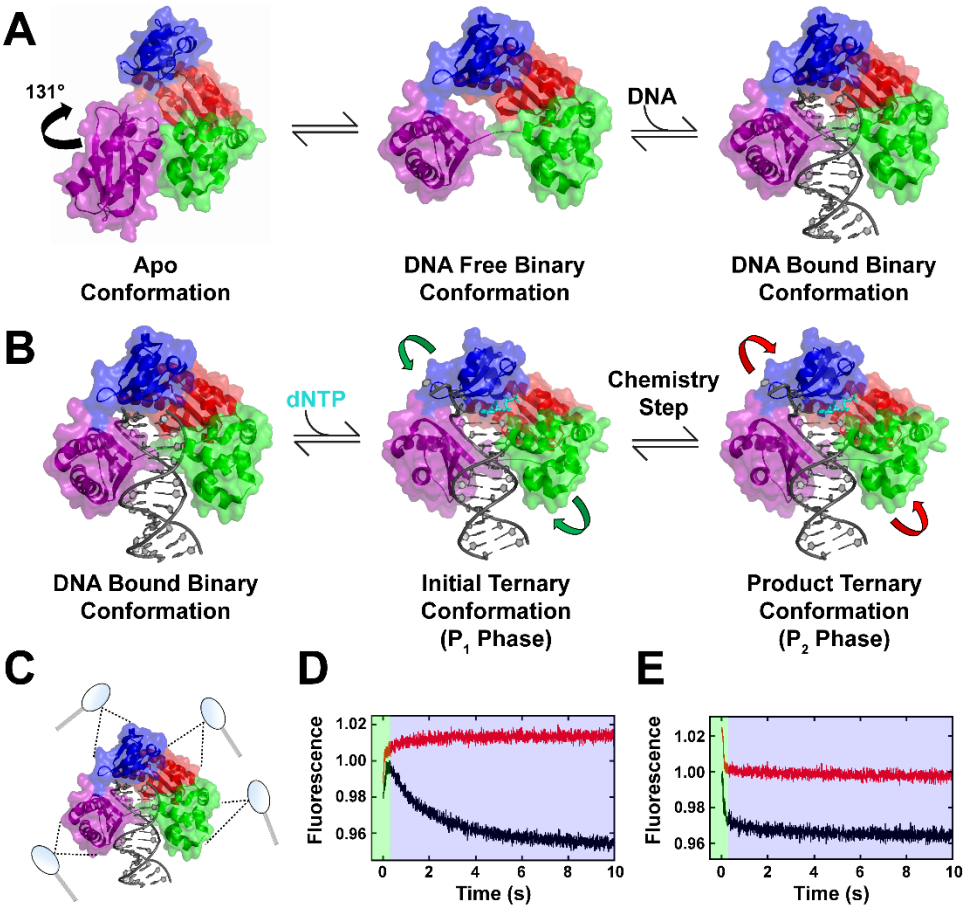


Figure 1.4: Conformational dynamics of Dpo4.

Finger, palm, thumb, and little finger are colored blue, red, green, and purple, respectively. (A) Dynamics of DNA binding. Dpo4 exists in equilibrium between the apo and binary conformations in the absence of DNA. DNA binds to the free binary conformation which may be mediated by the little finger domain.[10,13] (B) Dynamics during nucleotide binding and incorporation. In contrast to A-, B-, and some X-family DNA polymerases, comparison of binary and ternary crystal structures of Dpo4

(Continued)

Figure 1.4: Continued

demonstrates a lack of significant nucleotide binding associated protein dynamics. However, stopped-flow FRET analyses have uncovered subtle motions for each domain of Dpo4. Green arrows indicate the concerted movement of domains upon nucleotide binding during P_1 phase to grip the DNA substrate. Red arrows depict the relaxation of domains during the P_2 phase (*i.e.* opposite direction of P_1) following nucleotide incorporation.[10-13] (C) Pictorial representation of the intradomain FRET approach to investigate Dpo4 conformational dynamics within each domain (represented by magnifying glasses). Trp residues were site-specifically introduced into each domain to serve as FRET donors, while Cys residues modified with 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin) were site-specifically introduced into each domain to serve as FRET acceptors.[10] (D) Stopped-flow trace of little finger intradomain FRET construct (Y274W-K329C^{CPM}). Black trace shows correct nucleotide binding and incorporation on a natural DNA primer and demonstrates characteristic, anti-correlated P_1 (green shaded area) and P_2 (blue shaded area) phases. Red trace shows correct nucleotide binding with a dideoxy-terminated primer.[10] (E) Stopped-flow trace of finger intradomain FRET construct (S22W-K56C^{CPM}) colored as in (D). Note the similar direction of P_1 and P_2 phases regardless of natural or dideoxy-terminated primer.[10] PDBs 2RDI and 2RDJ were used to generate the structural figures in (A), (B), and (C).[149]

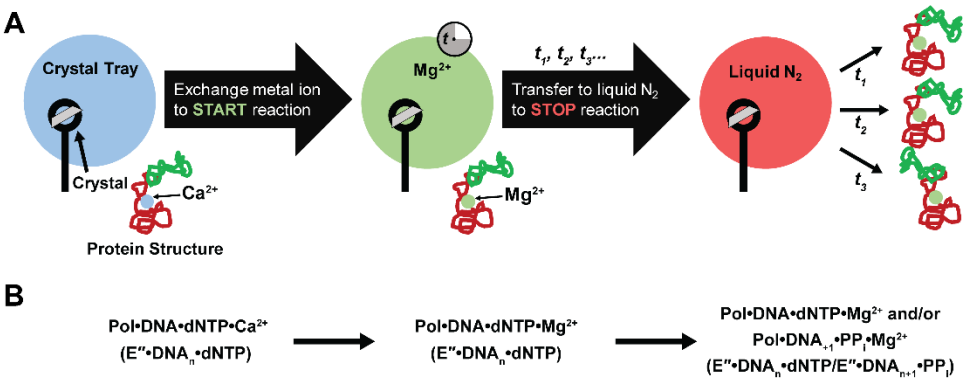


Figure 1.5: Time-resolved crystallography technique.

(A) A crystal in a pre-reactive state is isolated and transferred to a cryo-solution containing the catalytic metal ion. This initiates the reaction *in crystallo* and after varying periods of time, the reaction can be quenched by transferring the crystals to liquid N_2. Diffraction experiments are then performed on the crystals and the diffraction data are used to determine the three dimensional structures. During this process, the electron density of the bond forming and of the bond breaking is modeled as percent occupancy. The F_o-F_c difference map is then used to evaluate how well the model satisfies the experimental electron density.[225] (B) For time-resolved crystallography of a DNA polymerase-catalyzed nucleotidyltransfer reaction, a crystal of the ternary complex formed in the presence of non-catalytic Ca^{2+} is transferred to a cryo-solution containing the catalytic divalent metal ion, Mg^{2+} or Mn^{2+}. The polymerase complexes relevant to panel (A) are depicted and the corresponding enzyme forms relevant to those shown in Scheme 1.1A are shown in parentheses.

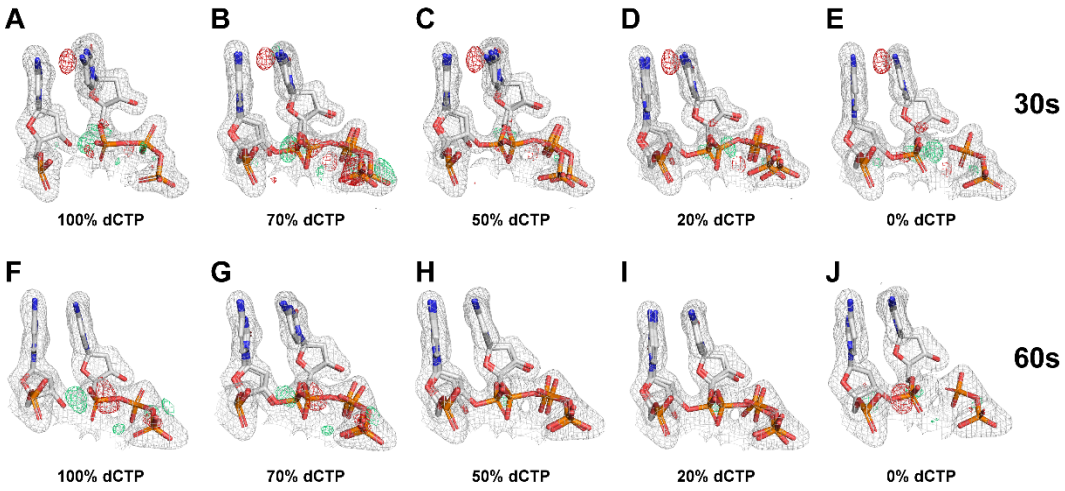


Figure 1.6: Modeling of phosphodiester bond formation.

Modeling of phosphodiester bond formation after 30 or 60 s of Mg^{2+}/Ca^{2+} ion-exchange during time-resolved X-ray crystallography of hPolβ.[42] The $2F_o$-F_c (light blue) maps contoured at 1σ and the F_o-F_c omit maps contoured at either 3σ (green) or -3σ (red) are presented for the primer 3′-terminal nucleotide, incoming dCTP, incorporated dCMP, and PP_i. The modeled occupancy of the reactants is listed below each structure. Strong positive (green) and negative (red) electron density mesh between the primer 3'-OH and the α-phosphate group of dCTP or between the α- and β-phosphate groups of dCTP indicate unsatisfactory modeling, *e.g.* the modeling of the reactants at 100% (A), 70% (B), 20% (D) and 0% (E) occupancies for the 30 s structure and 100% (F), 70% (G), 50% (H), 20% (I), and 0% (J) occupancies for the 60 s structure. In contrast, the absence of any positive or negative electron density with the modeling of the reactants at 50% (C) and 20% (I) occupancies suggests satisfactory modeling for the 30 and 60 s structures, respectively.[42]

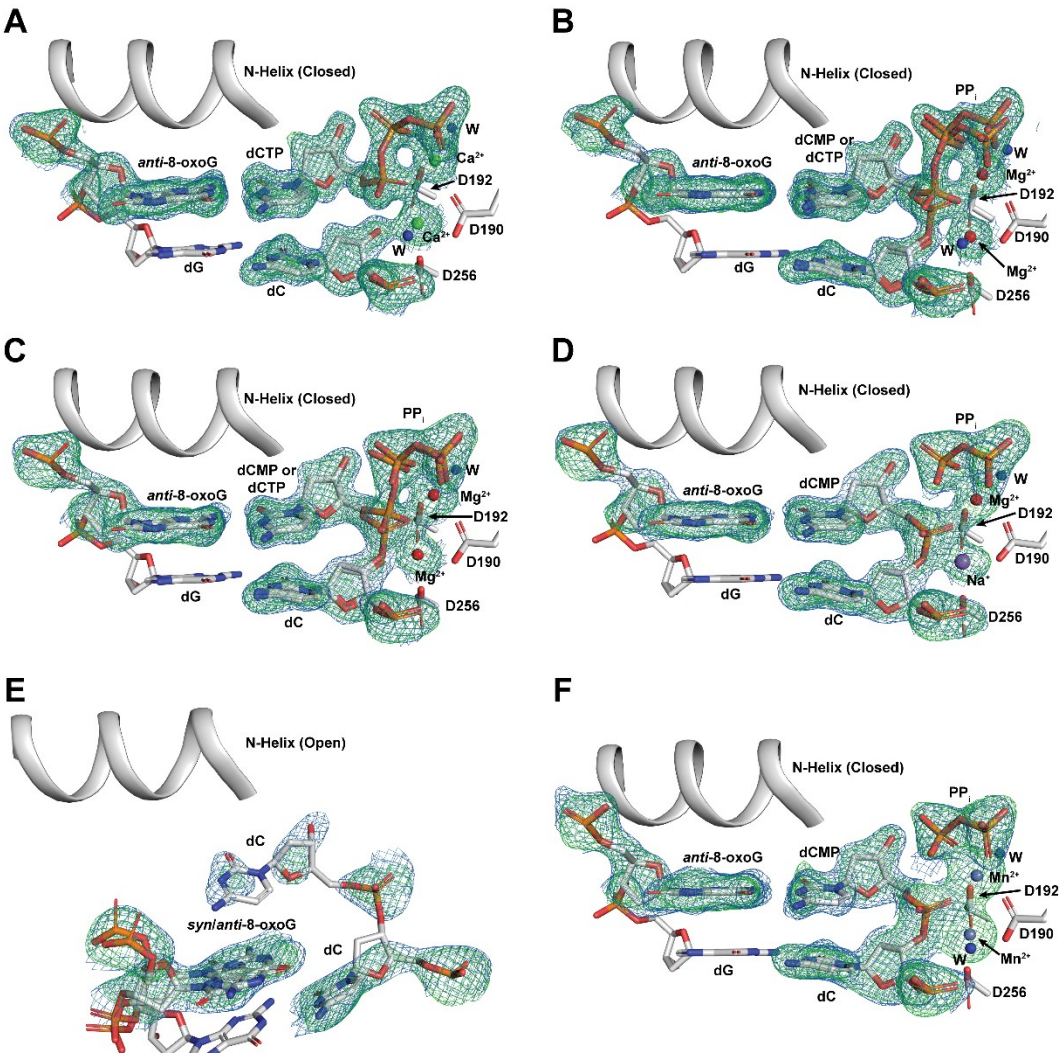


Figure 1.7: *In crystallo* phosphodiester bond formation by hPolβ.

Zoomed active sites show the incorporation of dCTP opposite 8-oxoG. The $2F_o$-F_c (blue mesh) and F_o-F_c (green mesh) maps are shown for the templating 8-oxoG, incoming dCTP, incorporated dCMP, pyrophosphate (PP_i), metal ions at the A- and B-sites, and the primer 3′-terminal nucleotide (dC), contoured to 1σ and 3σ levels, respectively. Water

(Continued)

Figure 1.7: Continued

molecules are shown as blue spheres. Ca^{2+}, Mg^{2+}, Mn^{2+}, and Na^{+} are shown as green, red, light blue, and purple spheres, respectively. Structures of the hPolβ pre-catalytic ternary complex (A), and structures of the hPolβ reaction intermediate or product ternary complexes following crystal-soaking with either 200 mM Mg^{2+} for 30 s (B), 60 s (C), 80 s (D), and 1 h (E), or 200 mM Mn^{2+} for 35 s (F). In (E), 8-oxoG was modeled in both *anti*- and *syn*-conformations.[42]

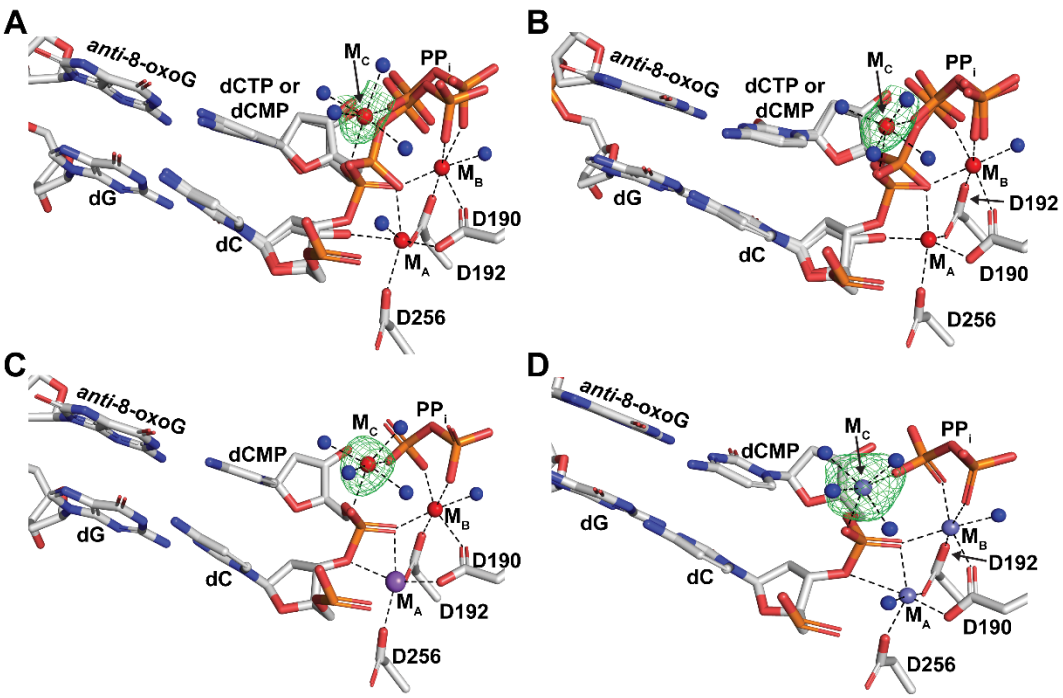


Figure 1.8: Appearance of the third divalent metal ion *in crystallo*

The third divalent metal ion is present in the reaction-state and post-catalytic structures of hPolβ during incorporation of dCTP opposite 8-oxoG. After soaking crystals of the pre-catalytic ternary complex of hPolβ (hPolβ•8-oxoG-DNA•dCTP) with Mg^{2+} for 30, 60, and 80 s, dCTP incorporation had completed by 50% (A), 80% (B), and 100% (C), respectively. (D) Following 35 s soaking with Mn^{2+}, dCTP was 100% incorporated. The F_o-F_c omit maps were contoured at 3σ (A-C) or 5σ (D) to show the electron density (green) of Mg^{2+} or Mn^{2+}, respectively, at the C-site. Coordinating aspartate side chains are shown as stick models, while water molecule ligands (blue), Mg^{2+} (red), Mn^{2+} (light blue), and Na^{+} (purple) are shown as spheres.[42]

Chapter 2. Introduction to Time-dependent X-ray Crystallography

Advances in Structural and Single-Molecule Methods for Investigating DNA Lesion Bypass and Repair Polymerases

Reproduced in part with permission from Raper, Austin T., Reed, A.J., Gadkari, V.V., and Suo, Z. (2017) Advances in Structural and Single-Molecule Methods for Investigating DNA Damage Tolerance and Repair. *Chem. Res. Toxicol.* 30, 260-269. DOI: 10.1021/acs.chemrestox.6b00342.

Author Contributions

Austin T. Raper and Andrew J. Reed equally contributed as co-first authors to the writing of the manuscript. Varun V. Gadkari provided initial ideas and helped edit the manuscript. Dr. Zucai Suo provided some conceptual direction and helped edit the manuscript.

2.1 Abstract

Innovative advances in X-ray crystallography and single-molecule biophysics have yielded unprecedented insight into the mechanisms of DNA lesion bypass and damage repair. Time-dependent X-ray crystallography has been successfully applied to view the bypass of 8-oxo-7,8-dihydro-2'-deoxyguanine (8-oxoG), a major oxidative DNA lesion, and the incorporation of the triphosphate form, 8-oxo-dGTP, catalyzed by human DNA polymerase β (hPolβ). Significant findings of these studies are highlighted here and their contributions to the current mechanistic understanding of mutagenic translesion DNA synthesis (TLS) and base excision repair (BER) are discussed. In addition, single-molecule Förster resonance energy transfer (smFRET) techniques have recently been adapted to investigate nucleotide binding and incorporation opposite undamaged dG and 8-oxoG by *Sulfolobus solfataricus* DNA polymerase IV (Dpo4), a model Y-family DNA polymerase. The mechanistic response of Dpo4 to a DNA lesion and the complex smFRET technique are described here. In this perspective, we also describe how time-dependent X-ray crystallography and smFRET can be used to achieve the spatial and temporal resolutions necessary to answer some of the mechanistic questions that remain in the fields of TLS and DNA damage repair.

2.2 Introduction

Over the last few decades, more emphasis has been placed on structural evidence of biochemical events than ever before. While many enzyme mechanisms can be inferred from the results yielded through typical biochemical methods, convincing structural evidence is often necessary to validate a particular model. For example, by solving the

crystal structures of the apo form, the binary complex (E•DNA), the pre-catalytic ternary complex (E•DNA•dNTP), and the product ternary complex (E•DNA•pyrophosphate) of numerous DNA polymerases across all six phylogenetic families (A, B, C, D, X, and Y),[237] which function in DNA replication, DNA damage repair, and translesion DNA synthesis (TLS),[23,61,63,64,67,68,76-78,80-82,84,87,223,238-240] the mechanism of DNA synthesis at the atomic-level was largely determined. Accordingly, more than 1,000 such crystal structures of DNA polymerases provide valuable information on the binding and positioning of substrates (DNA and dNTP) and products (DNA and pyrophosphate), the positioning and coordination of divalent metal ions, the binding conformations of active site residues, and putative protein conformational changes. Although spatially informative, traditional structural techniques are limited to static snapshots of stable states, therefore certain mechanistic details of the transition between reactant and product complexes of polymerase-catalyzed nucleotide incorporation may still be ambiguous or unidentified.

To overcome the disadvantages of conventional X-ray crystallography, time-dependent X-ray crystallography was recently developed and applied to follow polymerase-catalyzed reactions.[38-42,165] This relatively inexpensive and simple to execute technique has allowed several research laboratories to temporally capture the structures of reaction intermediates at atomic level resolutions.[38-42,165] As this elegant technique will likely change the research landscape for structurally characterizing biological processes such as DNA replication, DNA damage repair, and translesion DNA synthesis (TLS), we will highlight several notable time-dependent X-ray crystallographic studies of DNA

polymerases and give our opinions for new research directions in this perspective. Although useful, time-dependent X-ray crystallography is limited to enzyme reactions that can be readily activated *in crystallo*, and may only illuminate specific steps of a multi-step enzyme mechanism. For DNA polymerases, this has been limited to the chemistry step of nucleotide incorporation (Scheme 2.1, Step 6). Therefore, various mechanistic events, such as DNA binding, DNA translocation, nucleotide binding, protein conformational changes (Scheme 2.1, Steps 1, 4, 5, 7, and 8), and product release (Scheme 2.1, Step 9) cannot be observed. Furthermore, while time-dependent crystallography is possible through slowed enzyme catalysis *in crystallo*, any measured rates cannot be reliably compared to solution-based kinetics. Thus, the high-spatial resolution afforded by this technique comes at the cost of temporal resolution.

Consequently, a more versatile approach, capable of accessing faster, more relevant time scales and multiple enzymatic steps, is necessary to fully probe a catalytic mechanism. In recent years, advances in microscope optics, high resolution cameras, and state-of-the-art data acquisition and analysis software have spurned the advent and rapid growth of single-molecule methodologies by which to study biological processes *in vitro*. One particularly informative method, single-molecule Förster resonance energy transfer (smFRET),[241] reports on distance changes between site-specifically placed donor and acceptor fluorophores covalently linked to nucleic acid substrates or enzymes of interest. In this way, the solution behavior of individual molecules can be tracked in real-time to give unparalleled insight into complex enzyme mechanisms as well as protein-protein and protein/DNA interactions. Ensemble steady-state and pre-steady-state kinetic

methodologies,[22] while very informative and valuable, can only be utilized to report on the collective or average behavior of all molecules during a reaction.[241] As a result of this limitation, information on critical minor species or sub-populations of molecules during an enzyme-catalyzed reaction, that may perform distinct functions and/or exist in different conformational states, is lost. Furthermore, by following the reaction trajectories of individual molecules, smFRET has the remarkable ability to show the order of events for a given enzyme-catalyzed reaction without the need for synchronization, which is difficult or impossible for some systems. Thus, smFRET allows for simultaneous spatial and temporal characterization of an enzyme-catalyzed reaction. We have recently applied this technique to study *Sulfolobus solfataricus* DNA polymerase IV (Dpo4), a model Y-family DNA polymerase,[242] as it encountered and bypassed an 8-oxo-7,8-dihydro-2'-deoxyguanine (8-oxoG) lesion, a major oxidative DNA lesion, on a DNA template during DNA synthesis.[9] Structural studies have established that 8-oxoG does not significantly distort the structure of DNA, however its mutagenicity results from its ability to readily form two possible base pairing conformations within the polymerase active site.[243-245] Furthermore, published studies have shown that 8-oxoG forms specific interactions with polymerase active site residues to directly affect lesion bypass activities.[246-251]

Due to the wealth of data that single-molecule techniques can provide, we will give a more thorough description of our recent article in this perspective, as well as comment on future applications of this methodology to the study of DNA replication, TLS, and DNA damage repair. Previously, our laboratory has published many articles on the bypass of a variety of DNA lesions catalyzed by X- and Y-family DNA

polymerases.[11-13,133,134,138,139,226,240,252,253] Since most of these articles have been reviewed recently,[22] here we mainly discuss our work on 8-oxoG bypass catalyzed by Dpo4.[9,134]

2.3 Time-Resolved X-ray Crystallography

DNA polymerases bind DNA and dNTP sequentially, in addition to binding and coordinating divalent metal ions (*e.g.* Mg^{2+} or Mn^{2+}) to three conserved carboxylate residues at the active site in order to catalyze a nucleotidyl transfer reaction. Beginning in 2005, our laboratory attempted to perform time-dependent nucleotide incorporation in the crystals of Dpo4 through soaking its binary complex crystals (E•DNA) with an incoming correct or incorrect dNTP and Mg^{2+} for various times. However, this method resulted in the deterioration of crystal morphology and prevented us from obtaining high-resolution structures of reaction intermediate states. Instead, we collaborated with the laboratory of Paul Carey and utilized low-resolution single crystal Raman crystallography to obtain useful mechanistic information on nucleotide incorporation and protein conformational changes *in crystallo*.[252] In contrast to our misfortune, the laboratory of Katsuhiko Murakami employed similar time-dependent soak-trigger-freeze X-ray crystallography and solved the high-resolution crystal structures of pre-catalytic complex and two reaction intermediate states of bacteriophage N4 RNA polymerase.[165] Their structures reveal that the nucleotide-binding divalent metal ion at the B-site binds first followed by the catalytic divalent metal ion at the A-site. Following A-site divalent metal ion binding a phosphodiester bond was formed between the two bound rNTPs producing a 2-mer RNA transcript at the active site of bacteriophage N4 RNA polymerase, after which the A-site divalent metal ion quickly dissociates.[165] However, they did not observe any

reaction intermediates for bond formation or the third divalent metal ion bound at the C-site (see below).

Concurrently, the laboratory of Wei Yang pioneered a new method for time-dependent X-ray crystallographic studies of polymerases.[38] Rather than the aforementioned soak-trigger-freeze method where nucleotide and catalytic divalent metal ions were soaked simultaneously, their methodology takes advantage of the fact that the binding of Ca^{2+} at a polymerase active site allows for proper binding and positioning of DNA and nucleotide but does not support catalysis.[38] First, a pre-catalytic ternary complex (Scheme 2.1, E′′•DNA•dNTP), that has gone through all elementary steps in a kinetic mechanism (Scheme 2.1, Steps 1-5) prior to phosphodiester bond formation (Step 6)[5,9,11,22] was crystalized in the presence of non-catalytic Ca^{2+}. The Yang laboratory then initiated phosphodiester bond formation *in crystallo* via Ca^{2+} to Mg^{2+} ion exchange. This step was accomplished by removing crystals of the pre-catalytic ternary complex from the crystallization solution and transferring them to a cryo-protectant solution containing an excess of catalytic Mg^{2+} for a variable amount of time (seconds to minutes) before freeze-quenching the reaction in liquid nitrogen (Figure 2.1A). This methodology affords visualization of reaction intermediates in which mechanistic intricacies, such as chemical bond formation and breakage, steric inversion of the SN2 reaction center, changes in side chain rotamer conformations, DNA and dNTP binding conformations, and divalent metal ion coordination can be directly observed.[38] More significantly, the Yang laboratory unprecedentedly observed the appearance of a third Mg^{2+} ion during phosphodiester bond formation (Step 6, Scheme 2.1).[38]

Very recently, the Yang methodology was successfully adapted by Samuel Wilson's group and our laboratory[40-42] to investigate phosphodiester bond formation catalyzed by full-length human DNA polymerase β (hPolβ), an X-family polymerase. Wilson's group employed time-dependent X-ray crystallography to watch how hPolβ incorporated undamaged correct and incorrect dNTPs[40] as well as a damaged nucleotide (8-oxo-dGTP),[41] while we were the first to view both faithful (Figure 2.1B-2.1D) and unfaithful TLS across a DNA lesion (8-oxoG).[42] Notably, the mutagenic bypass of the major oxidative DNA lesion 8-oxoG during base excision repair (BER) *in vivo* could lead to cancer formation.[254-257]

Rotational freedom about the glycosidic bond (*anti/syn*-conformations) and the additional hydrogen bonding capabilities of 8-oxo-dGTP and 8-oxoG on the Hoogsteen face (oxidation of C8 and protonation of N7) result in the dual coding potential of the oxidized nucleotide and the damaged templating base (Figure 2.2). Therefore, *anti*-8-oxo-dGTP and *anti*-8-oxoG can Watson-Crick base pair with cytosine (Figures 2.2A and 2.2C) and *syn*-8-oxo-dGTP and *syn*-8-oxoG can Hoogsteen base pair with adenosine (Figures 2.2B and 2.2D). In the time-dependent studies, both Wilson and our laboratories identified key side chain interactions that support the previously reported dNTP and templating base preference for 8-oxoG bypass and 8-oxo-dGTP incorporation, respectively (Figure 2.2).[226] Notably, R283 of the thumb domain acts as a negative selection factor against 8-oxo-dGTP incorporation opposite dC through electrostatic repulsion between the side chain of R283 and the N2 position of 8-oxo-dGTP (Figure 2.2C).[41] R283 was also shown to stabilize the *syn*-conformation of a templating 8-oxoG

during dATP misincorporation (Figure 2.2B).[42] Interestingly, this structural information is consistent with our previous kinetic characterization of 8-oxoG bypass or 8-oxo-dGTP incorporation by hPolβ, and suggests that R283 modulates the mutagenicity of these processes.[41,42,226]

In addition, in the structures where partial dCTP (Figure 2.1C) or dATP incorporation opposite 8-oxoG was observed the primer 3′-OH moves 0.8-1.3 Å relative to its position in the pre-catalytic ternary structures and initiates the nucleophilic attack on the Pα of the incoming dNTP, resulting in a 1.0-1.1 Å shift of the Pα towards the 3′-OH and the steric inversion of the nonbridging oxygen atoms, as expected for an SN2 reaction.[42] Furthermore, key details for the post-catalytic process were identified including a stacking interaction between the newly incorporated nucleotide and the templating 8-oxoG, suggesting a role in preventing successive nucleotide incorporation and strand displacement on the nicked DNA product.[42]

More interestingly, as with the first time-dependent crystallography study of phosphoryl transfer by hPolη,[38] all three studies on hPolβ revealed three divalent metal ions (Mg^{2+} or Mn^{2+}) at the polymerase active site during phosphodiester bond formation.[40-42] The third divalent metal ion was unambiguously detected when Ca^{2+} was exchanged *in crystallo* with Mn^{2+},[40,42] which yielded anomalous scattering, and suggests that polymerases use a "three-metal-ion mechanism" (Figures 2.1C – 2.1D). These time-dependent X-ray crystallographic studies[38-42,165] clearly established that the binding of the divalent metal ions before phosphodiester bond formation follows the order of the B-, A-, and C-site while their dissociation order after the chemistry step is the A-, C-, and B-site.

These results challenge the long-standing "two-metal-ion mechanism" for the nucleotidyl transfer reaction proposed by Lauren Beese and Tom Steitz nearly two decades ago.[31] The Steitz laboratory solved the first structures of the Klenow fragment of *E. coli* DNA polymerase I and discovered that the exonuclease activity appears to require two Mg^{2+} ions[31] and therefore it was hypothesized that all phosphoryl transfer reactions including polymerase-catalyzed nucleotide incorporation[26,258] likely follow a two-divalent metal-ion mechanism. Their results have since been supported by hundreds of pre-catalytic ternary structures of other DNA polymerases[26,59,61,63,158,207] in which nucleotide and two divalent metal ions are bound at the polymerase active site. In the "two-metal-ion mechanism", the A- and B-site divalent metal ions, are associated with the deprotonation of the primer terminus and the binding of incoming nucleotide, respectively.

However, through the time-dependent X-ray crystallography studies,[38,40-42] the observed third Mg^{2+} ion at the C-site appears in a position bridging the α- and β-phosphates of the incoming nucleotide following reaction initiation (Figure 2.1C) and persisting through completion of nucleotide incorporation (Figure 2.1D), with its role in catalysis still debated and several competing hypotheses proposed. The positioning and timing of the third divalent metal ion have led Wilson's group to initially hypothesize a role in the reverse reaction, pyrophosphorolysis,[40] while others[38,39,41,259] and our laboratory[42] have suggested the third metal ion is involved in transition state stabilization and may aid in pyrophosphate release (Figures 2.1C and 2.1D). In any case, the necessity of a third divalent metal ion for catalysis is hard to discredit with five recent structural studies[38-42] and two computational analyses that detail its presence.[49,50]

While further biochemical or biophysical characterization is necessary to elucidate the role of the third divalent metal ion, it does not coordinate protein residues and therefore it is difficult to definitively assay its role in catalysis through simple mutational studies. However, if the third divalent metal ion serves a catalytic role in the forward reaction, then it should also participate in the reverse reaction. Consistently, the Wilson laboratory was able to crystallize the ground state of the reverse, pyrophosphorolysis reaction with the third divalent metal ion bound, but did not capture any reaction state structures (presumably because the reverse reaction is unfavorable), which lead them to hypothesize its importance in either pyrophosphate release or pyrophosphorolysis.[40] The Yang group was first to attempt to assay the functional role of the third divalent metal ion via measuring metal ion saturation *in crystallo* and then correlating this with the amount of divalent metal ion necessary for nucleotidyl transfer in solution.[39] While this was a reasonable approach, it would be perhaps more convincing to determine its functional role through techniques other than time dependent X-ray crystallography. For example, our laboratory is utilizing electroparamagnetic resonance (EPR) spectroscopy to investigate the functional roles of divalent metal ions on the activity of hPolβ.

Notably, the above mechanistic details as well as others presented in the time-dependent X-ray crystallographic studies[38-42] are unobservable with conventional X-ray crystallography. It will be interesting to see if some of the new mechanistic aspects, such as the “three-metal-ion mechanism”, will hold true for other DNA polymerases, especially those in the A, B, C, and D families as well as reverse transcriptases. It will

also be exciting to watch how other types of DNA lesions, *e.g.* bulky single-base lesions, double-base lesions, and tandem lesions, are accommodated and bypassed by the X- and Y-family DNA polymerases. Furthermore, while the mechanistic steps pertaining to the reverse conformational change, pyrophosphate release, and DNA translocation or release (Scheme 2.1, Steps 7-9) have remained elusive to structural characterization, time-dependent X-ray crystallography has the potential to probe these processes to further define the polymerase-catalyzed mechanism of nucleotide incorporation (Scheme 2.1).

2.4 Single-Molecule Förster Resonance Energy Transfer

smFRET can be applied to study distance changes between donor and acceptor fluorescent dyes covalently attached to one or more interacting biomolecules. The donor and acceptor fluorophores must meet several photophysical criteria,[241] and be positioned such that the inter-fluorophore distance is limited to $R = R_0 \pm 0.5R_0$, where R is the inter-fluorophore distance and R_0 is the calculated Förster radius, which is unique to the selected dye pair and depends on several important parameters,[241] not discussed here. When properly designed, a FRET system can be very sensitive to distances changes, yielding low-resolution structural information for distance changes occurring between the donor and acceptor within 10-75 Å.[260]

To apply this technique for the monitoring of single molecules, especially DNA polymerases, one very popular strategy involves the surface immobilization of a DNA substrate that is chemically modified to contain a donor fluorophore and 5′- or 3′-biotin group (Figure 2.3A). Quartz slides and glass coverslips are chemically functionalized and passivated with a lawn of polyethylene glycol molecules (some of which are modified

with biotin bound by NeutrAvidin) as well as sandwiched together and sealed to form a reaction chamber. The modified DNA is then flowed into the chamber and immobilized via the very tight biotin-NeutrAvidin interaction. Prism-based total internal reflection fluorescence (TIRF)[241] is then utilized to greatly reduce fluorescent background in order to achieve signal-to-noise ratios sufficient to image single DNA molecules. Briefly, the donor molecules are specifically excited by a laser of appropriate wavelength that is incident on a Pellin Broca prism at the critical angle necessary to achieve TIRF (Figure 2.3A).[241] Polymerase molecules labeled with an acceptor fluorophore are then introduced into the imaging chamber to bind to the immobilized DNA. Upon polymerase binding to DNA, if the donor and acceptor fluorophores are sufficiently near to one another, energy transfer (i.e. FRET) from the donor to the acceptor can occur (Figure 2.3A). The fluorescence emission of the donor is then spectrally separated from the acceptor by the utility of dichroic mirrors to form distinct donor and acceptor emission channels (Figure 2.3B). Subsequently, both donor and acceptor emissions are incident on an electron multiplying charged coupled device (EMCCD) camera to record the fluorescence intensities over time (Figure 2.3C). The FRET efficiency (E) is then calculated as $E = I_A/(I_A+I_D)$, where I_D and I_A are the fluorescence intensities of the donor and acceptor, respectively. Accordingly, polymerase binding events and subsequent protein conformational changes are reflected as anti-correlated decreases in the donor emission intensity and increases in the acceptor emission intensity. Notably, smFRET has been applied to study various biochemical systems, including DNA replication.[9,111,163,213,261-266]

However, of the many existing DNA polymerase smFRET studies, only a few have explored TLS.[9,265,266]

A recent publication from our laboratory utilized smFRET to study the lesion bypass capabilities of Dpo4.[9] Dpo4, the lone Y-family DNA polymerase in *Sulfolobus solfataricus*, has been extensively studied due to its structural and functional similarities to the human Y-family polymerases.[84,242] We sought to explore the effects of a templating 8-oxoG lesion on the dynamics of Dpo4 during DNA binding as well as nucleotide binding. By examination of single-molecule FRET trajectories of Dpo4 binding to immobilized DNA substrates with a dG base at the templating position, we first observed the dynamic behavior of the polymerase on undamaged DNA. Interestingly, Dpo4 was shown to interconvert between three primary, non-zero FRET efficiency states (Figure 2.4A). In any smFRET analysis, assigning physical significance to observed FRET states is challenging. For our study, we had to carefully analyze all control and experimental data, as well as apply our knowledge of Dpo4 from previous stopped-flow FRET investigations,[11,12,22,267] before developing the proposed model. We assigned the mid-FRET efficiency state (FRET ≈ 0.6) to the polymerase bound to the DNA with the primer/template junction base pair occupying the position of nascent base pair in the active site (pre-insertion state), while in the low-FRET efficiency state (FRET ≈ 0.4), Dpo4 had translocated along the DNA substrate by one base pair to permit the binding of an incoming nucleotide (insertion state). Beyond the location of the polymerase at the primer/template terminus, the low- and mid-FRET states are also associated with distinct conformations of the polymerase as suggested by our dwell time analysis of the low- and

mid-FRET states, wherein the duration of time the polymerase remains at a particular FRET efficiency value before transitioning to another is quantified. We found that the rate of interconversion (2.8-4 s^{-1}) between the low-and mid-FRET states was far too slow to be reporting on the actual DNA translocation event itself, as mistakenly suggested by another single-molecule research group,[163] which has been estimated to occur at rates of >100 s^{-1}.[11,12,268-270] The rate of interconversion actually reports on domain motions, that occur concomitant with translocation along the DNA substrate, as observed through the Cy5-labeled Finger domain in our smFRET study.[9] Accordingly, our FRET analysis was also reporting on the dynamics of the Finger domain during DNA binding as well as DNA translocation, which is consistent with the dynamic behavior of the Finger domain, as well as other domains of Dpo4, previously elucidated by our research group through stopped-flow FRET studies.[11,12,22,151,267] Interestingly, our experiments also detected the existence of a high-FRET state (FRET ≈ 0.8) which we expect to be a unique binding mode of Dpo4 involving the Little Finger domain, although more work is necessary to better characterize this unique finding.

When Dpo4 was bound to DNA containing the 8-oxoG lesion at the templating position (Figure 2.4A), the results were largely unchanged compared to those with the undamaged DNA substrate (Figure 2.4B). However, we did detect subtle differences between the experiments using 8-oxoG-containing DNA compared to those with undamaged DNA including a shift in the number of FRET events at the low-FRET state (50% with the 8-oxoG-containing DNA (Figure 2.4C), versus 35% with the undamaged DNA (Figure 2.4B) as well as increased polymerase binding times and slower

dissociation kinetics from the lesion-containing DNA substrate. Therefore, more polymerase molecules were observed to be in the productive insertion state, ready to accept an incoming nucleotide, as well as spent longer periods of time on the 8-oxoG-containing DNA substrate without dissociating. Accordingly, we suggested that the 8-oxoG lesion modestly stabilized the formation of productive binary complexes. This assertion was further supported by experiments wherein the correct nucleotide, dCTP, was included in the imaging chamber in the presence of non-catalytic Ca^{2+}, to form ternary complexes (E′′•DNA•dCTP, Scheme 2.1) of either Dpo4 with undamaged or 8-oxoG-containing DNA. Interestingly, the presence of the correct nucleotide completely shifted all FRET events to the low-FRET state and significantly enhanced the FRET event durations. In fact, Dpo4 remained bound to DNA 5-10-fold longer in presence of a correct over an incorrect dNTP. Furthermore, we observed that Dpo4, in the presence of dCTP, remained bound to 8-oxoG-containing DNA nearly 2-fold longer than undamaged DNA. In this way, we were able to observe the nucleotide selectivity of a Y-family DNA polymerase on undamaged as well as damage-containing DNA, and preference for the 8-oxoG lesion at the templating position. Notably, Dpo4 has been shown to replicate more efficiently across from 8-oxoG than normal dG,[134] and our data suggest that this occurs as a result of stabilization of the ternary complex of the polymerase. Consistently, these findings are in agreement with previous structural studies that have established that Arg332 in the Dpo4 active site forms specific interactions with 8-oxoG.[162] Our single-molecule investigation of the binding modes and kinetics of Dpo4 is largely consistent with a previous smFRET study of Dpo4 by the Rueda and Romano research laboratories,

wherein Dpo4 binds to undamaged DNA in at least two FRET states, with the low-FRET state stabilized by correct nucleotide binding.[163] Another smFRET study by these groups investigated the effects of the bulky DNA adducts 2-aminofluoene and N-acetyl-2-aminofluorene on Dpo4 binding conformations and activity when the lesions were located at the templating or junction base pair positions.[266] Their results suggest that these bulky lesions significantly distort the structure of the binary complex as compared to results with unmodified DNA or our results with the non-helix distorting lesion, 8-oxoG. Consistent with our results with 8-oxoG, the addition of the correct nucleotide again induced a stable ternary complex (E′′•DNA•dCTP, Scheme 2.1) despite the DNA damage in the template strand. Collectively, their study and ours illuminate key details of the response of Dpo4 to two bulky DNA adducts and a small, major oxidative lesion, respectively, and help to explain their mechanisms of mutagenicity. It is clear from these studies that Dpo4 is better suited to faithfully replicate opposite 8-oxoG, whereas the bulkier amine adducts more significantly affect Dpo4 binary and ternary complexes to contribute to misincorporation events.

For future investigation, it will be interesting to see how the extension step of TLS by Dpo4, which has been shown to be more significantly disrupted by DNA damage,[138,139] is affected by the 8-oxoG lesion at the single-molecule level as well as how the inclusion of bulkier, DNA-helix distorting lesions, such as cyclobutane pyrimidine dimers or nitropolyaromatic hydrocarbon-generated lesions, affect the single-molecule behavior of Dpo4. Furthermore, it remains to be seen if other Y-family polymerases act similar to Dpo4 when they bypass a lesion and then subsequently extend the lesion

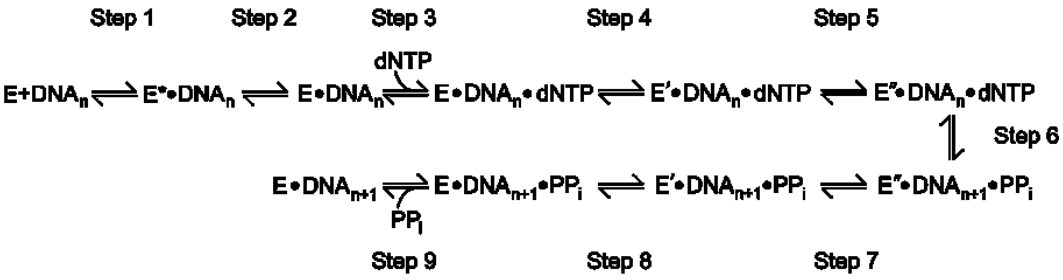
bypass product. Moreover, we are excited at the prospect of utilizing single-molecule spectroscopy to investigate the complexities of polymerase switching during TLS. When a replicative polymerase is stalled by a DNA lesion during DNA replication, it is replaced by a Y-family polymerase, which is expected to bypass the damage, in a process known as polymerase switching.[230,271-273] This event is thought to be mediated by the ring-like clamp protein, proliferating cell nuclear antigen (PCNA) which may simultaneously bind to both the replicative and Y-family DNA polymerases.[274-276] Although polymerase switching is a vital, damage tolerance response for the cell, the details of the process are not well-defined and several competing mechanisms exist to describe its functioning.[230,271-273] Application of smFRET to study polymerase switching will undoubtedly yield key mechanistic insight and clarity into the protein-protein and protein-DNA dynamics at the stalled replication fork.

2.5 Conclusion

Taken together, our detailed analyses of DNA damage tolerance and repair by cutting-edge time-dependent crystallographic and single-molecule methodologies have significantly increased our understanding of these vital life processes. In the coming years, it will be interesting to see how these techniques will continue to inform on the mechanisms of BER enzymes, as well as the putative coordination and interactions between them. Moreover, we are excited about the prospect of further clarifying the underlying mechanisms of mutagenic translesion DNA synthesis by X- and Y-family DNA polymerases by these methods. In all, the technological advances in enzymology described in this review have informed ground-breaking research to yield paradigm-

shifting results, and we expect that future applications will be equally impactful to the field.

2.6 Schemes



Scheme 2.1: Proposed mechanism of nucleotide incorporation catalyzed by a DNA polymerase.

E, E′, and E″ represent three different conformations of the DNA polymerase sampled during correct nucleotide incorporation. E* represents the non-productive, pre-translocated binary complex.

2.7 Figures

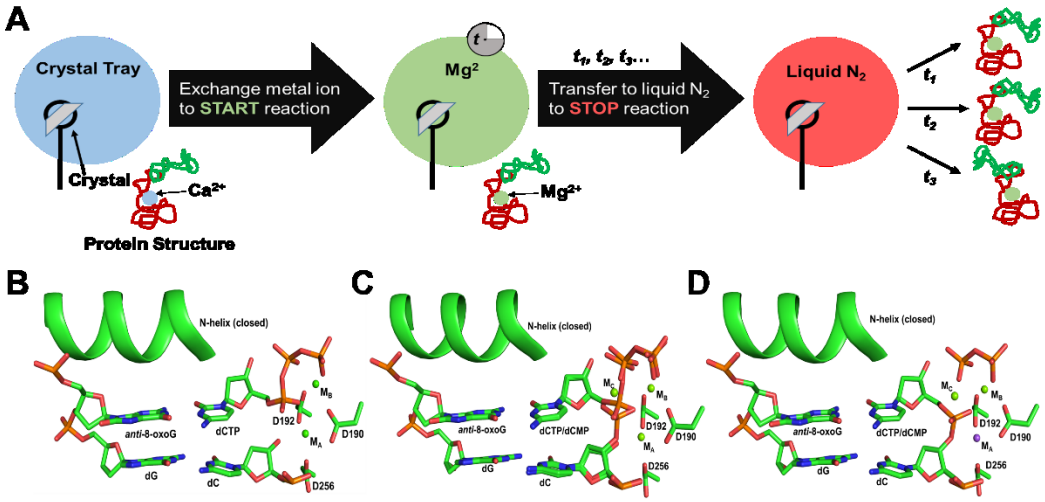


Figure 2.1: Time-dependent X-ray crystallography of DNA lesion bypass.

(A) Cartoon representation of time-dependent X-ray crystallography method. Crystal trays are set up in the necessary conditions to grow protein crystals of pre-catalytic ternary polymerase complexes in the presence of Ca^{2+}. The isolated crystals are transferred to a cryo-solution that contains magnesium (Mg^{2+}) to initiate the reaction *in crystallo* by exchanging the bound Ca^{2+} with Mg^{2+}. The crystals are soaked in the Mg^{2+} cryo-solution for varying times (t_1, t_2, t_3…) and then transferred to liquid N_2 to freeze-quench the reaction before X-ray diffraction experiments. (B-D) Structural progression of single-nucleotide incorporation catalyzed by hPolβ (PDB codes 4RPX, 4RPZ, and 4RQ0).[42] (B) Pre-catalytic ternary complex of hPolβ with correct incoming dCTP paired with templating 8-oxoG with bound A- and B-site divalent metal ions (green spheres, M_A and M_B, respectively). (C) Reaction state structure of hPolβ during phosphoryl transfer

(Continued)

Figure 2.1: Continued

with the incoming nucleotide modeled partially as dCTP and dCMP to match the increasing electron density (not shown) between the primer 3′-OH and the α-phosphate of dCTP and the decreasing electron density (not shown) between the α- and β-phosphates of dCTP. The third divalent metal ion occupies the C-site (M_C) and is coordinated by the nonbridging oxygen atoms of the α- and β-phosphates and water molecules (not shown). (D) Post-catalytic complex of hPolβ following complete incorporation of dCTP opposite 8-oxoG with M_C and pyrophosphate (PP_i) still bound. Notably, M_A is modeled as Na^+ (purple sphere). (B-D) Each panel shows a cartoon model of the N-Helix of the thumb domain in the closed conformation.

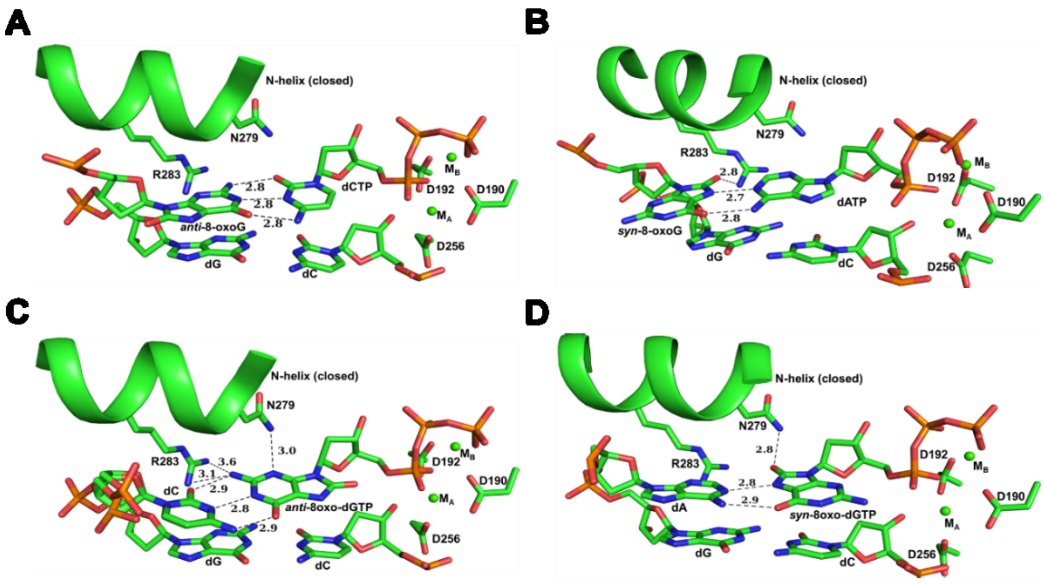


Figure 2.2: Zoomed active site views of hPolβ with templating 8-oxoG or incoming 8-oxo-dGTP.

(A) Templating *anti*-8-oxoG with incoming dCTP (PDB code 4RPX).[42] (B) Templating *syn*-8-oxoG with incoming dATP (PDB code 4RQ3).[42] (C) Incoming *anti*-8-oxo-dGTP with templating dC (PDB code 4UBC).[41] (D) Incoming *syn*-8-oxo-dGTP with templating dA (PDB code 4UAW).[41]

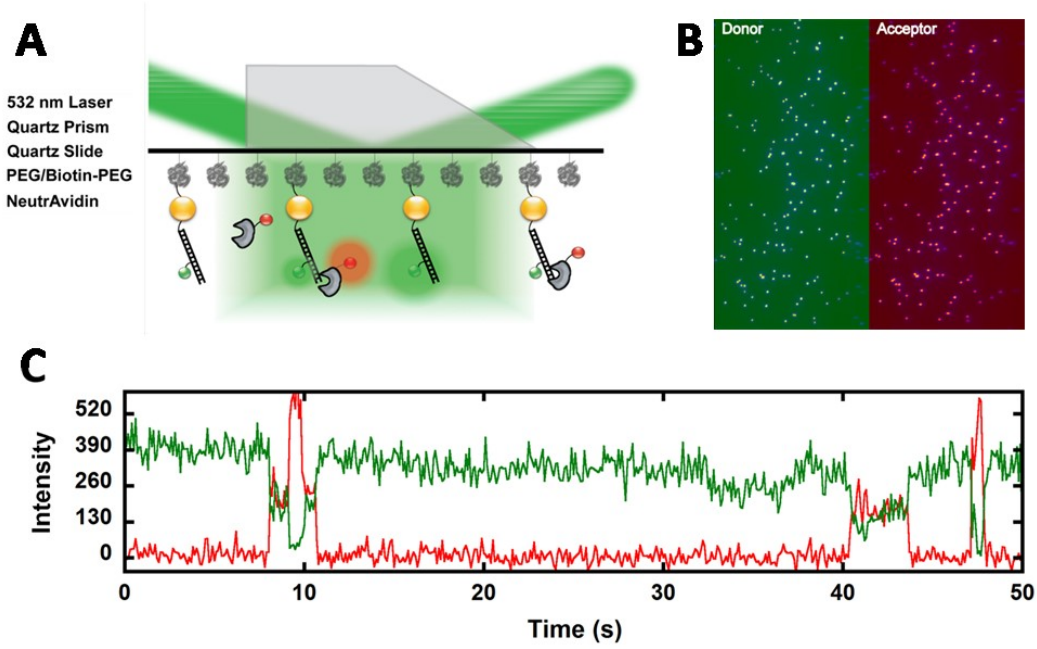


Figure 2.3: Single-molecule FRET.

(A) Modified, donor labeled DNA is immobilized via the biotin-NeutrAvidin interaction. Donor molecules (green spheres) are specifically excited by a laser of appropriate wavelength using total internal reflection fluorescence (TIRF). Acceptor labeled polymerase (red spheres) binding to DNA, results in energy transfer (i.e. FRET). (B) The resulting fluorescence emissions of the donor and acceptor are spectrally separated to form separate donor and acceptor emission channels. (C) Movies of single molecules are recorded and colocalized fluorescence intensities in the donor and acceptor emission channels are extracted by customized software. Anti-correlated changes in the donor and acceptor fluorescence intensities, shown in green and red, respectively, indicate FRET events.

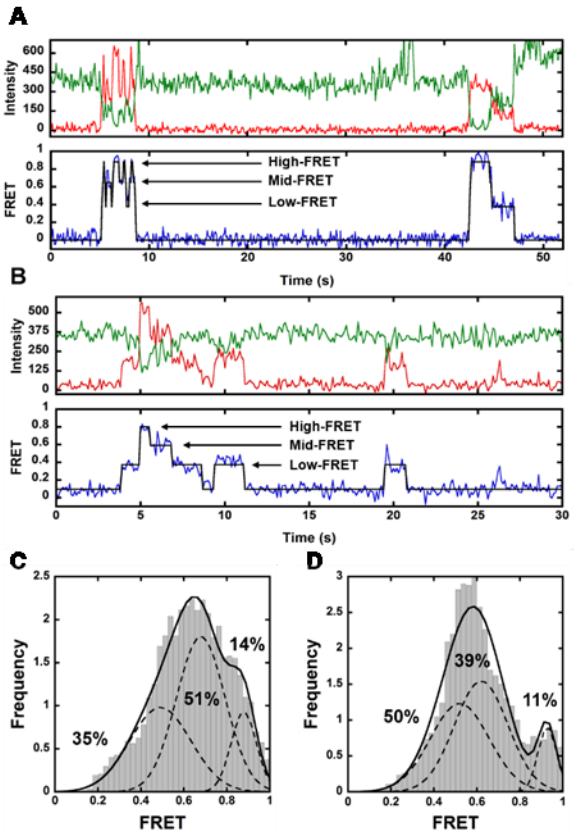


Figure 2.4: Single-molecule analysis of Dpo4 binding to undamaged or 8-oxoG-containing DNA.

(A, B)[9] The top panels show representative single-molecule trajectories of Dpo4 binding to undamaged DNA (A) or 8-oxoG-containing DNA (B) with donor and acceptor fluorescence intensities shown in green and red, respectively. The bottom panels display the corresponding, calculated FRET efficiency in blue and the idealized FRET efficiency from a Hidden Markov analysis in black. Three non-zero FRET states are clearly evident in each trajectory, High-, Mid-, and Low-FRET. FRET efficiency histograms of Dpo4

(Continued)

Figure 2.4: Continued

binding to undamaged (B) or 8-oxoG-containing DNA (C) with individual Gaussian peak fits for each FRET state shown as overlapping, black dashed lines and the Gaussian sum shown as the solid black line. The respective population percentages for each FRET state are indicated in black.

Chapter 3. Viewing Human DNA Polymerase β Faithfully and Unfaithfully Bypass an Oxidative Lesion by Time-Dependent Crystallography

Reproduced in part with permission from Vyas, R., Reed, A.J., Tokarsky, E.J., and Suo, Z. (2015) Viewing Human DNA Polymerase β Faithfully and Unfaithfully Bypass an Oxidative Lesion by Time-Dependent Crystallography. *J. Am. Chem. Soc.* 137 (15), 5225-5230. DOI: 10.1021/jacs.5b02109.

Author Contributions

Rajan Vyas and Andrew J. Reed crystallized the protein DNA complexes, performed time-dependent crystallography experiments, solved crystal structures (R.V. solved 8, A.J.R. solved 4), prepared figures, and wrote/edited the manuscript. E. John Tokarsky performed kinetic experiments and helped edit the manuscript. Zucai Suo conceived the research, provided guidance for time-dependent crystallography experiments, and wrote/edited the manuscript.

3.1 Abstract

One common oxidative DNA lesion, 8-oxo-7,8-dihydro-2′-deoxyguanine (8-oxoG), is highly mutagenic *in vivo* due to its *anti*-conformation forming a Watson-Crick base pair with correct dCTP and its *syn*-conformation forming a Hoogsteen base pair with incorrect dATP. Here, we utilized time-resolved X-ray crystallography to follow 8-oxoG bypass by human DNA polymerase β (hPolβ). In the 12 solved structures, both Watson-Crick (*anti*-8-oxoG:*anti*-dCTP) and Hoogsteen (*syn*-8-oxoG:*anti*-dATP) base pairing were clearly visible and were maintained throughout the chemical reaction. Additionally, a third Mg^{2+} appeared during the process of phosphodiester bond formation and was located between the reacting α- and β-phosphates of the dNTP, suggesting its role in stabilizing reaction intermediates. After phosphodiester bond formation, hPolβ reopened its conformation, pyrophosphate was released, and the newly incorporated primer 3′-terminal nucleotide stacked, rather than base paired, with 8-oxoG. These structures provide the first real-time pictures, to our knowledge, of how a polymerase correctly and incorrectly bypasses a DNA lesion.

3.2 Introduction

Aerobic respiration, inflammatory responses, phagocytosis, ionizing radiation, pollutants, and xenobiotics generate intracellular reactive oxygen species which damage genomic DNA and form various oxidative lesions including 8-oxoG,[277,278] 8-oxoG lesions can form in genomic double stranded DNA by oxidation of guanine bases or by the incorporation of oxidized guanine nucleotides (8-oxoGTP). The incorporation of 8-oxoGTP is minimal due to the action of MTH1, a hydrolase that specifically eliminates

oxidized dNTPs and rNTPs.[279] In humans, most 8-oxoG lesions occur in double stranded DNA, where 8-oxoG:dC base pairs are first removed by 8-oxoG DNA glycosylase OGG1, and then repaired by the sequential actions of AP endonuclease 1, hPolβ, and DNA ligase I or III/XRCC1 in the base excision repair (BER) pathway (Figure 3.1A).[246] For an unrepaired, templating 8-oxoG lesion, its *anti*-conformation forms a Watson-Crick base pair with correct dCTP while its *syn*-conformation forms a Hoogsteen base pair with incorrect dATP. Such incorrect base pairing results in G→T transversion during subsequent rounds of DNA replication (Figure 3.1A and B). Repair of an 8-oxoG:dA mispair in genomic DNA begins with excision of the undamaged adenine by the adenine DNA glycoslyase MYH. The resulting abasic site is then processed similarly by the remaining steps of the BER pathway (Figure 3.1A).[246] *In vivo* studies suggest Polβ is the major polymerase involved in this alternative BER pathway, and bypasses 8-oxoG by predominantly inserting either dCTP or dATP[246,280].

Previous kinetic studies have determined that opposite 8-oxoG, hPolβ incorporates dCTP with a 2-fold, $1x10^4$-fold and $2x10^4$-fold higher efficiency than dATP, dGTP and dTTP, respectively.[226] Consistent with these data, hPolβ was found to selectively incorporate dCTP over dATP opposite 8-oxoG in a single-nucleotide gapped DNA substrate (8-oxoG-DNA) by ~2-fold while misincorporation of dGTP and dTTP was barely observable in the presence of catalytic Mg^{2+} (Figure 3.1D). In contrast, hPolβ was more efficient with mutagenic Mn^{2+} [(75)] than with Mg^{2+} and incorporated dCTP and dATP onto 8-oxoG-DNA to similar extents (Figure 3.1D). To reveal a structural basis for the preferential binding of correct dCTP and incorrect dATP, Wilson *et, al.* have used

either dideoxy-terminated primers[247] or non-hydrolyzable nucleotide analogs[281] to obtain the crystal structures of the pre-catalytic ternary complexes of hPolβ with either the *anti*-8-oxoG:*anti*-dCTP Watson-Crick base pair or the *syn*-8-oxoG:*anti*-dATP Hoogsteen base pair in the closed conformations. Although instructive, these structures do not show the process of nucleotide incorporation during 8-oxoG bypass. Recently, Basu and Murakami attempted to catch elusive intermediates formed during bacteriophage N4 RNA polymerase-catalyzed transcription by time-dependent soak-trigger-freeze X-ray crystallography.[165] Similarly, Yang *et al.* developed a time-resolved crystallization method capable of capturing structural snapshots at specific time points of phosphodiester bond formation during correct nucleotide incorporation onto undamaged DNA by human DNA polymerase η (hPolη).[38] Their methods were adapted by Wilson *et al.* to capture several structural intermediates formed in the processes of nucleotide incorporation onto undamaged DNA by hPolβ.[40] Here, we employed a similar time-resolved crystallization method to structurally visualize the bypass of a DNA lesion for the first time, to our knowledge.

3.3 Materials and Methods

3.3.1 Preparation of protein and DNA

Human DNA polymerase β (hPolβ) was over expressed and purified as reported earlier.[226] After purification, the protein was washed four times with several volumes of the 20 mM Bis-Tris (pH 7.0) buffer, concentrated to 15 mg/ml and stored at -80 °C. The single-nucleotide gapped DNA substrate 8-oxoG-DNA (Figure S1C, 1 mM) was prepared by mixing equimolars of a 16-mer template (5′-CCGACG$_o$GCGCATCAGC-3′),

where G_o represents 8-oxoG, a 10-mer upstream primer (5′-GCTGATGCGC-3′) and a 5′-phosphorylated 5-mer downstream primer (5′-P-GTCGG-3′) and heating to 80 °C and then cooling to room temperature.

3.3.2 Crystallization and structure determination

To form the binary complex of hPolβ and a single-nucleotide gapped DNA substrate, the purified protein (15 mg/ml) and 8-oxoG-DNA (1 mM) were mixed in a 1:1 volume ratio. The resulting mixture was then heated to 40 °C and then slowly cooled to room temperature as reported previously.[247] The binary complex crystals were obtained using the hanging drop vapor diffusion method by equilibrating against a reservoir buffer containing 50 mM imidazole (pH 7.5), 16%–18% PEG3350, and 350 mM sodium acetate[247] at room temperature and streak seeded after 24 hrs. To form the pre-catalytic ternary complex crystal, the binary complex crystals were transferred to a cryosolution containing 50 mM imidazole (pH 7.5), 16%–18% PEG3350, 90 mM sodium acetate, 15% ethylene glycol, 50 mM $CaCl_2$ and an incoming nucleotide (dCTP or dATP). The pre-catalytic ternary crystals were flash-frozen in liquid N_2. To initiate nucleotide incorporation reaction in the crystals, the pre-catalytic ternary complex crystals were transferred to another cryosolution containing 50 mM imidazole (pH 7.5), 16%–18% PEG3350, 90 mM sodium acetate, 15% ethylene glycol, and 200 mM $MgCl_2$ or $MnCl_2$, and soaked for various times. The reactions were stopped by flash freezing the crystals in liquid N_2. X-ray diffraction data were collected using LRL-CAT beamline facilities at Advance Photon Source (APS), Argonne National Laboratory. The X-ray diffraction data were processed using MOSFLM.[282] The structures were solved through the molecular

replacement method by PHASER[283] using a previously determined structure (PDB code: 4KLD)[40] as the initial model in the absence of ligands and solvent molecules. Initial models for all structures were generated by performing simulated annealing using Phenix.[284] Further iterations of refinement were carried out using REFMAC5.[285] COOT[286] was used for visualization and model building. Quality of the models was assessed using PROCHECK.[287] All electron density omit maps were generated using Phenix.[284] Figures were created using PYMOL.[288]

3.3.3 ^{32}P-labeled primer extension assay

A preincubated solution of hPolβ (10 nM) and 5′-[^{32}P]-labeled 8-oxoG-DNA (30 nM) was reacted for 15 s with a dNTP (100 μM) in the buffer (50 mM Tris-HCl (pH 7.8 at 25 °C), 5 mM $MgCl_2$ or $MnCl_2$, 50 mM NaCl, 0.1 mM EDTA, 5 mM DTT, 10% glycerol, and 0.1 mg/ml of BSA) at 25 °C. The reaction was quenched with 0.37 M EDTA. Reaction products were resolved using sequencing gel electrophoresis (17% (w/v) acrylamide, 8 M urea).

3.4 Results and Discussion

3.4.1 Pre-catalytic ternary structures

The binary complex of hPolβ and 8-oxoG-DNA (hPolβ•8-oxoG-DNA) was first crystallized and the resulting crystals were then soaked with a cryosolution containing non-catalytic Ca^{2+} and either dCTP or dATP in order to form pre-catalytic ternary complex crystals. The pre-catalytic ternary complex crystals of hPolβ•8-oxoG-DNA•dATP (Table 3.1) and hPolβ•8-oxoG-DNA•dCTP (Table 3.2) diffracted to 1.90 Å and 2.00 Å, respectively, and their structures were solved through molecular replacement.

Both F_o-F_c and $2F_o$-F_c electron density maps of the pre-catalytic ternary structures clearly show the closed conformation of hPolβ relative to the open binary structure of hPolβ•8-oxoG-DNA as seen previously with undamaged DNA.[40] Also visible, are the *anti*-8-oxoG:*anti*-dCTP Watson-Crick or *syn*-8-oxoG:*anti*-dATP Hoogsteen base pairing and two hexacoordinated Ca^{2+} ions at the A-site and B-site, liganded by the primer 3′-OH, the triphosphate of dCTP or dATP, water, and the carboxylates of D190, D192, and D256 (Figures 3.2A, 3.3A, and 3.4). The pre-catalytic structures of hPolβ•8-oxoG-DNA•dCTP and hPolβ•8-oxoG-DNA•dATP are nearly superimposable with an RMSD of 0.30 Å and only a slight repositioning of the nascent base pair (Figure 3.4A). Notably, other than the Hoogsteen base pairing, the *syn*-conformation of 8-oxoG is further stabilized by a hydrogen bond (2.8 Å) between the O8 atom and the side chain of R283 in Helix N which occupies the minor groove (Figure 3.4H). The Ca^{2+} ions were positioned at the same A-site and B-site as the two Mg^{2+} ions in the pre-catalytic ternary structures of hPolβ [281] and other polymerases.[26,36,63,65,289-294]

3.4.2 Reaction State Structures

To trigger phosphodiester bond formation, the pre-catalytic hPolβ•8-oxoG-DNA•dCTP and hPolβ•8-oxoG-DNA•dATP crystals were transferred to a cryosolution containing 200 mM catalytic Mg^{2+} to initiate the exchange with non-catalytic Ca^{2+} ions at the A-site and B-site of hPolβ. The reaction was stopped after 30, 60 and 80 s by flash freezing the crystals in liquid nitrogen. The crystals diffracted to 1.90-2.20 Å for the dCTP reaction intermediates (Table 3.2) and 2.00-2.32 Å for the dATP reaction intermediates (Table 3.1) and the structures were solved by molecular replacement. The

structures after 30 and 60 s metal ion-exchange, partial occupancies of the reactants (primer 3′-nucleotide and dCTP or dATP) and products (pyrophosphate and dCMP or dAMP) were modeled to account for the observed electron density gain associated with phosphodiester bond formation between the primer 3′-OH and the α-phosphate (Pα) of dNTP, and electron density loss associated with bond breakage between the α- and β-phosphates of the dNTP (Figure 3.5). Using this approach, both F_o-F_c and $2F_o$-F_c electron density maps for 30 and 60 s metal ion-exchange were appropriately modeled at 50% and 80% dCTP incorporation and 30% and 70% dATP misincorporation, respectively (Figure 3.5). For partial dCTP incorporation, the primer 3′-OH moved 0.8-1.1 Å upwards relative to its position in the pre-catalytic state structures and launched the in-line nucleophilic attack on Pα, leading to 1.0 Å downward movement of Pα and the steric inversion of its nonbridging oxygens (Figure 3.2B and 3.2C). Similarly, for partial dATP incorporation, the primer 3′-OH and Pα moved toward each other by 0.9-1.3 Å and 1.0-1.1 Å, respectively (Figure 3.3B and 3.3C). Surprisingly, besides the two typical hexacoordinated Mg^{2+} ions at the A-site and B-site as observed in the pre-catalytic ternary structures of other polymerases,[26,36,63,65,289-294] additional electron density near Pα and Pβ was observed (Figure 3.6) and modeled as a third Mg^{2+} at 0.5-0.8 occupancy (Tables 3.1 and 3.2). This C-site Mg^{2+} was liganded by the non-bridging oxygen atoms of Pα and Pβ and water molecules (Figure 3.6 and Table 3.3).

3.4.3 Post-catalytic ternary structures

In comparison, the structures after 80 s metal ion-exchange display 100% phosphodiester bond formation and no electron density between Pα and Pβ of either

dCTP (Figure 3.2D) or dATP (Figure 3.3D), indicating the nucleotide incorporation was slower in the crystals than in solution (Figure 3.1D). The slower rate was possibly caused by decreased thermal motion *in crystallo* and by the viscous cryosolution.[38,40] Furthermore, the F_o-F_c omit maps contoured at 3σ show the presence of the C-site ion which was modeled as Mg^{2+} at full occupancy (Figure 3.6C-3.6G). Although the B-site Mg^{2+} persisted, the A-site ion was modeled as Na^+ based on the number of coordinating ligands and their distances from the ion[40] (Figure 3.7A and 3.7B). Such Mg^{2+} to Na^+ exchange presumably facilitated the subsequent closed to open protein conformational change and pyrophosphate release.

3.4.4 Appearance of the 3rd divalent metal ion at the active site

To provide additional evidence for the existence of the C-site ion, the pre-catalytic crystals of hPolβ•8-oxoG-DNA•dCTP and hPolβ•8-oxoG-DNA•dATP were soaked with 200 mM Mn^{2+} to exchange for Ca^{2+} for 35 s to initiate nucleotide incorporation. The resulting crystals diffracted to 2.2 (Table 3.2) and 2.0 Å (Table 3.1), respectively. In the refined structures (Figures 3.2F and 3.3F), the F_o-F_c and $2F_o$-F_c electron density maps clearly show 100% product formation and two hexacoordinated Mn^{2+} at the A-site and B-site. The presence of the C-site Mn^{2+} in both structures is validated by the F_o-F_c electron density map contoured at 5σ, and this Mn^{2+} is coordinated by water molecules and non-bridging oxygen atoms of the newly formed phosphodiester bond and pyrophosphate (Figure 3.6D-3.6H and Table 3.3). Moreover, overlaying the product-state structures from 80 s Ca^{2+} to $Mg^{2+}$and 35 s Ca^{2+} to Mn^{2+} ion-exchange shows the divalent metal ions at each site were nearly superimposable (Figure 3.7C and

3.7D). Previously, with undamaged DNA and correct dNTP, the third Mg^{2+} was observed in both the reaction-state and product-state structures with hPolη,[38] but only in the product-state structure with hPolβ.[40] More recently during submission of this manuscript, Wilson *et al.* published a study to show a third divalent metal ion in the pre-catalytic, reaction-state, and post-catalytic structures for *anti*-8-oxoGTP incorporation opposite cytosine and only in the post-catalytic structure for *syn*-8-oxoGTP incorporation opposite adenine.[295] This is the first instance, to our knowledge, of a pre-catalytic structure containing a third divalent metal ion but is likely unique to the 8-oxoGTP substrate as its O8 position stabilizes two bridging water molecules which act as coordinating ligands for this third metal ion.[295] Together, these previous and our current time-resolved crystallization studies strongly suggest that DNA polymerases likely use a "three-metal-ion mechanism" for catalysis,[38] rather than the "two-metal-ion mechanism" proposed by L. Beese and T. Steitz,[31] which is based on *E.* coli DNA polymerase I and has been supported by numerous pre-catalytic ternary structures of other polymerases.[26,59,61,63,158,207] The potential role of the C-site Mg^{2+} is to neutralize the negative charge in the transition state and facilitate phosphodiester bond formation.[38] Notably, the number of ligands for the C-site ion varied from four to six due to different numbers of water ligands (Figure 3.6), which is likely caused by the labile nature of the water ligands.

3.4.5 Comparison of the pre-catalytic, reaction-state and post-catalytic ternary structures

Interestingly, overlaying the structures of the pre-catalytic ternary complex with the reaction-state and post-catalytic ternary complexes after metal ion-exchange for 30, 60 and 80 s (Figure 3.8) indicates that the *anti*-8-oxoG:*anti*-dC Watson-Crick and *syn*-8-oxoG:*anti*-dA Hoogsteen base pairing were maintained throughout the chemistry step. In addition, all structures of hPolβ were in the closed conformation and nearly superimposable with RMSDs of 0.13-0.49 Å. Other than the primer 3′-nucleotide and incoming nucleotide involved in product formation described above, the DNA substrate and the active site residues barely changed their positions. Notably, among the three metal ions, only the C-site ion substantially altered its position (1.4 Å) during dATP incorporation (Figure 3.8B), indicating its dynamic nature. Failure to observe this notable change in the C-site ion position during correct incorporation by both hPolβ (Figure3.8A) and hPolη[38] is possibly due to the inability to capture earlier reaction-state structures.

3.4.6 Post-catalytic binary structures

To monitor changes in the product-state structure caused by pyrophosphate release, the pre-catalytic crystals of hPolβ•8-oxoG-DNA•dCTP and hPolβ•8-oxoG-DNA•dATP were soaked with $MgCl_2$ for longer than 80 s and several structures were solved. However, only the 1 hr structures of 2.70 (Table S1) and 2.00 Å (Table S2) for dCTP and dATP incorporation, respectively, show significant differences from the corresponding 80 s structures (Figures 3.2D and 3.3D). In the 1 hr post-catalytic binary structures, both pyrophosphate and the three metal ions had dissociated, and hPolβ had

transitioned from closed to open conformation as indicated by the movement of Helix N away from the active site, and the repositioning of the active site residues Y271, F272, D192, D190 and D256 (Figures 3.2E, 3.3E, and 3.9). Such post-catalytic active site rearrangement and protein conformational change are consistent with those projected through kinetic studies.[5,6,11,296] The pyrophosphate dissociation from the crystals observed here was not seen previously after correct nucleotide incorporation onto undamaged DNA by hPolβ *in crystallo*.[40] Furthermore, 8-oxoG in the nicked DNA product after dCTP incorporation was best modeled in both *anti*- and *syn*-conformations (Figure 3.10A), and the incorporated dCMP had poor electron density and stacked, rather than base paired, with 8-oxoG (Figure 3.2E). Similarly, in the post-catalytic binary product structure after dATP incorporation (Figure 3.3E), 8-oxoG was better modeled in the *anti*-conformation (Figure 3.10B) and stacked with the incorporated dAMP. Moreover, there were two separate electron densities for the phosphate portion of dAMP, leading to the modeling of alternative conformations for the last three nucleotides of the primer 3′-terminus (Figure 3.3E). These base-stacking patterns after dCTP and dATP incorporation are not suitable for further DNA synthesis and thereby may facilitate the transition from the closed to open conformation of hPolβ. Alternatively, following nucleotide incorporation, hPolβ may first transition from the closed to the open binary conformation, allowing the observed stacking interactions to occur. Either mechanism would prevent further upstream primer elongation and downstream primer displacement.

3.5 Conclusions

In summary, our time-resolved crystallization studies have revealed that both Watson-Crick (*anti*-8-oxoG:*anti*-dCTP) and Hoogsteen (*syn*-8-oxoG:*anti*-dATP) base pairing modes were persistent throughout phosphodiester bond formation. This bond formation was facilitated by a third divalent metal ion which has never been observed in any pre-catalytic ternary structure of a polymerase. Transition from the closed ternary to open binary conformation facilitated by the stacking interactions between the incorporated nucleotide (dAMP or dCMP) and the templating 8-oxoG, suggesting that the open binary conformation is primed for the transfer of the nicked DNA substrate to a DNA ligase for continuous BER. Through the snapshots of nucleotide incorporation obtained by time resolved crystallography, this study provides structural details important for the understanding of error-free and error-prone DNA lesion bypass.

3.6 Tables

	Pre-catalytic ternary complex	30 s	60 s	80 s	1 hr	35 s ($MnCl_2$)
Data collection*						
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell dimensions						
a, *b*, *c* (Å)	49.3, 79.6, 55.4	49.1, 79.6, 55.3	49.2, 79.7, 55.3	50.4, 80.4, 55.6	55.3, 79.6, 55.8	50.2, 82.5, 54.7
α, β, γ (°)	90, 106.5, 90	90, 106.4, 90	90, 106.6, 90	90, 107.7, 90	90, 109.5, 90	90, 110.8, 90
Resolution (Å)	53.16-2.00 (2.05-2.00)	40.59-2.10 (2.16-2.10)	79.70-2.32 (2.45-2.32)	44.28-2.25 (2.32-2.25)	52.61-2.00 (2.05-2.00)	41.27-2.00 (2.11-2.00)
R_{merge}†	0.103 (0.472)	0.076 (0.382)	0.081 (0.644)	0.082 (0.542)	0.104 (0.562)	0.120 (0.499)
$I/\sigma I$	7.3 (2.2)	7.8 (2.5)	11.6 (2.0)	8.7 (2.0)	6.8 (2.2)	5.0 (2.4)
Completeness (%)	95.6 (95.0)	99.7 (99.8)	95.3 (88.2)	98.8 (98.0)	92.7 (90.8)	99.5 (98.7)
Redundancy	3.5 (3.4)	3.2 (3.2)	3.3 (2.7)	3.7 (3.5)	3.6 (3.2)	3.4 (3.2)
Refinement						
Resolution (Å)	47.29-2.00	40.59–2.10	50.0-2.32	44.28–2.25	45.40–2.00	40.79–2.00
No. reflections	25340	22700	16065	18900	27220	26717
R_{work}/R_{free}‡	0.191/0.245	0.189/0.243	0.196/0.265	0.197/0.252	0.206/0.260	0.206/0.270
No. atoms						
Protein	2636	2626	2614	2633	2610	2709
DNA	633	652	652	654	716	654
Water	242	195	77	71	202	231
B-factors (Å²)						
Protein	26.7	33.7	38.8	43.3	34.1	27.6
DNA/dCTP/PPi	34.14/25.1/-	40.5/24.6/19.8	44.1/28.8/31.8	51.9/-/34.6	31.3/-/-	32.06/-/23.7
Water	32.7	22.7	29.8	34.7	33.5	30.8
Metal A/B/C**	26.6/24.6/-	13.0/19.4/38.2	26.4/27.7/32.7	28.6/28.9/36.7	-/-/-	18.3/17.9/38.7
R.m.s deviations						
Bond lengths (Å)	0.007	0.008	0.009	0.007	0.007	0.009
Bond angles (°)	1.282	1.342	1.413	1.232	1.259	1.359
Reaction ratio						
Ratio of RS/PS***	1.0/0.0	0.7/0.3	0.3/0.7	0.0/1.0	0.0/1.0	0.0/1.0
Occupancy						
Metal A/B/C**	1.0/1.0/0	1.0/1.0/0.8	1.0/1.0/0.8	1.0/1.0/0.8	-/-/-	1.0/1.0/1.0
Pyrophosphate	-	0.3	0.7	1.0	-	1.0
PDB ID	4RQ3	4RQ4	4RQ5	4RQ6	4RQ7	4RQ8

*Highest resolution shell is shown in parenthesis.

** Metal A/B/C refers to the metal ions at the A-, B- and C-site, respectively. Values indicate the occupancy at which the metal ion was modeled in the given structure.

***RS and PS are abbreviations for the reactant state and product state.

†$R_{merge} = \Sigma|I-\langle I\rangle|/\Sigma I$, where I is the integrated intensity of each reflection.

‡R value = $\Sigma||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c are observed and calculated structure factor amplitudes, respectively.

Table 3.1: Data collection and refinement statistics of the pre-catalytic ternary complex of hPolβ●8-oxoG-DNA●dATP.

Statistics for the hPolβ●8-oxoG-DNA●dATP complex after soaking crystals with Mg^{2+} for 30 s, 60 s, 80 s and 1 hr, or with Mn^{2+} for 35 s.

	Pre-catalytic ternary complex	30 s	60 s	80 s	1 hr	35 s ($MnCl_2$)
Data collection*						
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell dimension						
a, *b*, *c* (Å)	49.6, 79.3, 55.6	49.4, 79.5, 55.5	49.4, 79.3, 55.4	49.6, 79.5, 55.4	55.4, 80.3, 55.8	49.8, 82.2, 54.6
α, β, γ (°)	90, 106.3, 90	90, 106.6, 90	90, 106.2, 90	90, 106.6, 90	90, 110.2, 90	90, 110.6, 90
Resolution (Å)	31.85-1.90 (1.94-1.90)	31.85-1.90 (1.94-1.90)	44.19-2.19 (2.26-2.19)	47.52-2.20 (2.27-2.20)	52.34-2.70 (2.83-2.70)	46.64-2.20 (2.27-2.20)
R_{merge}†	0.083 (0.370)	0.092 (0.506)	0.122 (0.548)	0.154 (0.587)	0.079 (0.426)	0.097 (0.604)
$I/\sigma I$	6.8 (2.0)	7.7 (2.2)	6.4 (2.1)	4.7 (2.1)	8.3 (2.1)	7.7 (2.3)
Completeness (%)	95.4 (91.6)	99.7 (99.8)	97.6 (96.6)	99.3 (99.4)	92.9 (92.5)	90.8 (90.1)
Redundancy	3.3 (3.1)	3.3 (3.2)	3.5 (3.1)	3.2 (3.2)	3.2 (3.0)	3.6 (3.5)
Refinement						
Resolution (Å)	31.43-1.90	31.23–1.90	44.19–2.19	44.18–2.20	45.63–2.70	46.64–2.20
No. reflections	29702	30816	19738	19866	11202	18209
R_{work}/R_{free}‡	0.188/0.226	0.189/0.229	0.195/0.251	0.207/0.268	0.203/0.280	0.194/0.268
No. atoms						
Protein	2641	2636	2627	2628	2598	2692
DNA	633	671	671	652	716	652
Water	246	286	158	126	10	155
B-factors (Å²)						
Protein	23.7	21.9	28.0	27.7	62.4	35.9
DNA/dCTP/PPi	25.1/15.9/-	21.5/14.3/15.6	26.5/20.4/28.3	29.1/-/29.3	53.4/-/-	36.9/-/33.6
Water	28.7	29.0	27.8	25.1	33.1	32.9
Metal A/B/C**	18.3/18.7/-	11.3/14.5/17.5	22.4/24.3/25.5	22.1/23.0/30.6	-/-/-	26.3/23.7/35.4
R.m.s deviations						
Bond lengths (Å)	0.006	0.006	0.007	0.007	0.008	0.010
Bond angles (°)	1.283	1.327	1.303	1.320	1.367	1.403
Reaction ratio						
Ratio of RS/PS***	1.0/0.0	0.5/0.5	0.2/0.8	0.0/1.0	0.0/1.0	0.0/1.0
Occupancy						
Metal A/B/C**	1.0/1.0/0	1.0/1.0/0.5	1.0/1.0/0.5	1.0/1.0/1.0	-/-/-	1.0/1.0/0.8
Pyrophosphate	-	0.5	0.8	1.0	-	1.0
PDB ID	4RPX	4RPY	4RPZ	4RQ0	4RQ1	4RQ2

*Highest resolution shell is shown in parenthesis.

*** Metal A/B/C refers to the metal ions at the A-, B- and C-site, respectively. Values indicate the occupancy at which the metal ion was modeled in the given structure.

**RS and PS are abbreviations for the reactant state and product state.

†$R_{merge} = \Sigma|I-\langle I\rangle|/\Sigma I$, where I is the integrated intensity of each reflection.

‡R value = $\Sigma||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c are observed and calculated structure factor amplitudes, respectively.

Table 3.2: Data collection and refinement statistics of the pre-catalytic ternary complex of hPolβ●8-oxoG-DNA●dCTP.

Statistics for the hPolβ●8-oxoG-DNA●dCTP complex after soaking crystals with Mg^{2+} for 30 s, 60 s, 80 s and 1 hr, or with Mn^{2+} for 35 s.

	Metal Ion	50% Incorporation (Mg^{2+} soak)	80% Incorporation (Mg^{2+} soak)	100% Incorporation (Mg^{2+} soak)	100% Incorporation (Mn^{2+} soak)
dCTP	**A-site**	Mg^{2+}: D256 (1), D192 (1), D190 (1), 3′-OH (1), dCTP/dCMP (1), W (1)	Mg^{2+}: D256 (1), D192 (1), D190 (1), 3′-OH (1), dCTP/dCMP (1)	Na^{+}: D256 (1), D192 (1), D190 (1), 3′-OH (1), dCMP (1)	Mn^{2+}: D256 (1), D192 (1), D190 (1), 3′-OH (1), dCMP (1), W (1)
	B-site	Mg^{2+}: D192 (1), D190 (1), dCTP/dCMP (1), PPi (2), W (1)	Mg^{2+}: D192 (1), D190 (1), dCTP/dCMP (1), PPi (2), W (1)	Mg^{2+}: D192 (1), D190 (1), dCMP (1), PPi (2), W (1)	Mn^{2+}: D192 (1), D190 (1), dCMP (1), PPi (2), W (1)
	C-site	Mg^{2+}: dCTP/dCMP (1), PPi (1), W (4)	Mg^{2+}: dCTP/dCMP (1), PPi (1), W (4)	Mg^{2+}: dCMP (1), PPi (1), W (4)	Mn^{2+}: dCMP (1), PPi (1), W (4)
dATP	**Metal Ion**	**30% Incorporation (Mg^{2+} soak)**	**70% Incorporation (Mg^{2+} soak)**	**100% Incorporation (Mg^{2+} soak)**	**100% Incorporation (Mn^{2+} soak)**
	A-site	Mg^{2+}: D256 (1), D192 (1), D190 (1), 3′-OH (1), dAMP /dATP (1), W (1)	Mg^{2+}: D256 (1), D192 (1), D190 (1), 3′-OH (1), dATP/dAMP (1), W (1)	Na^{+}: D256 (1), D192 (1), D190 (1), 3′-OH (1), dAMP (1))	Mn^{2+}: D256 (1), D192 (1), D190 (1), 3′-OH (1), dAMP (1), W (1)
	B-site	Mg^{2+}: D192 (1), D190 (1), dATP/dAMP (1), PPi (2), W (1)	Mg^{2+}: D192 (1), D190 (1), dATP/dAMP (1), PPi (2), W (1)	Mg^{2+}: D192 (1), D190 (1), dAMP (1), PPi (2), W (1)	Mn^{2+}: D192 (1), D190 (1), dAMP (1), PPi (2), W (1)
	C-site	Mg^{2+}: dATP/dAMP (1), PPi (1), W (3)	Mg^{2+}: dATP/dAMP (1), PPi (1), W (2)	Mg^{2+}: dAMP (1), PPi (1), W (4)	Mn^{2+}: dAMP (1), PPi (1), W (3)

Table 3.3: Coordinating ligands for each metal-ion site during dCTP or dATP incorporation.

The table lists the identity of the ligands coordinating each metal ion, *e.g.* the primer 3′-OH (3′-OH), pyrophosphate (PPi), water molecule (W), and the oxygen atoms of the phosphate groups of bound dCTP/dATP or incorporated dCMP/dAMP, and their coordination numbers in parentheses.

3.7 Figures

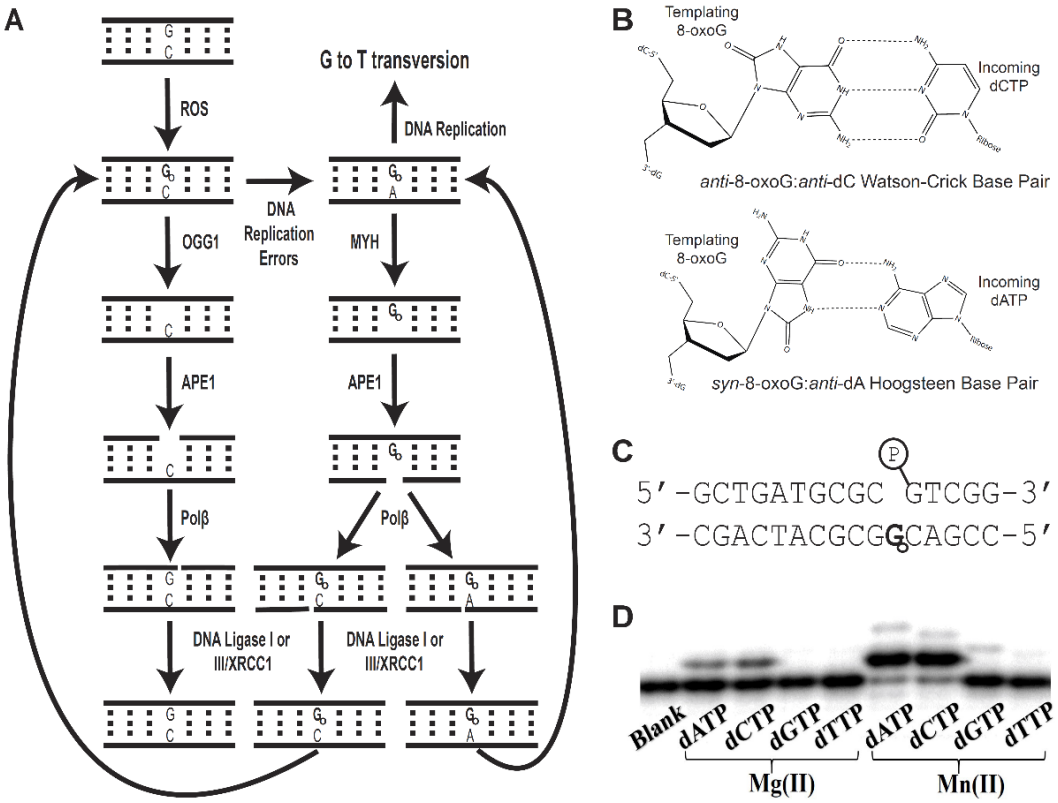


Figure 3.1: 8-oxoG and Base Excision Repair (BER).

Generation and repair of 8-oxoG by BER in eukaryotes is shown. Intracellular reactive oxygen species (ROS) damage genomic DNA and form various lesions including 8-oxoG (**G_o**). Most *anti*-8-oxoG:*anti*-dC base pairs are repaired by removal of the damaged guanine base via an oxoguanine DNA glycosylase OGG1. The backbone of the resulting apurinic site is then cleaved by apurinic endonuclease 1 (APE1). DNA polymerase β (Polβ) then performs gap-filling synthesis and deoxyribose cleavage. Finally, DNA ligase I or III/XRCC1 complex ligates the nicked DNA substrate to complete BER. Any

(Continued)

Figure 3.1: Continued

anti-8-oxoG:*anti*-dC pairs that go unrepaired may undergo erroneous DNA replication and result in *syn*-8-oxoG:*anti*-dA mispairs in genomic DNA. Repair of a *syn*-8-oxoG:*anti*-dA mispair begins with the removal of the undamaged adenine base via an adenine DNA glycosylase MYH. Subsequently, the apurinic site is processed by APE1, Polβ and DNA ligase I or III/XRCC1 in BER. Notably, Polβ-catalyzed gap-filling synthesis can result in either *syn*-8-oxoG:*anti*-dA mispairs or *anti*-8-oxoG:*anti*-dC correct pairs due to the dual coding potential of 8-oxoG. The *anti*-8-oxoG:*anti*-dC pairs can be repaired through the OGG1 branch of the BER pathway while the *syn*-8-oxoG:*anti*-dA pairs can reenter the MYH branch of the pathway. Alternatively, both base pairs could enter DNA replication where *syn*-8-oxoG:*anti*-dA base pairs would result in G to T transversion. (B) The *anti*-8-oxoG:*anti*-dCTP Watson-Crick base pair and the *syn*-8-oxoG:*anti*-dATP Hoogsteen base pair. (C) A single-nucleotide gapped DNA substrate (8-oxoG-DNA) for our crystallization and ^{32}P-labeled primer extension assay. The downstream primer 5-mer was 5′-phosphorylated while the upstream primer 10-mer was not. The templating 8-oxoG is shown as $\mathbf{G_o}$. (D) Gel image of the ^{32}P-labeled primer extension assay. A preincubated solution of hPolβ (10 nM) and 5′-[^{32}P]-labeled 8-oxoG-DNA (30 nM) was reacted with the indicated dNTP (100 μM) for 15 s in the presence of either Mg^{2+} (left lanes) or Mn^{2+} (right lanes) at 25 °C and then quenched with 0.37 M EDTA. “Blank” indicates no nucleotides added in the reaction.

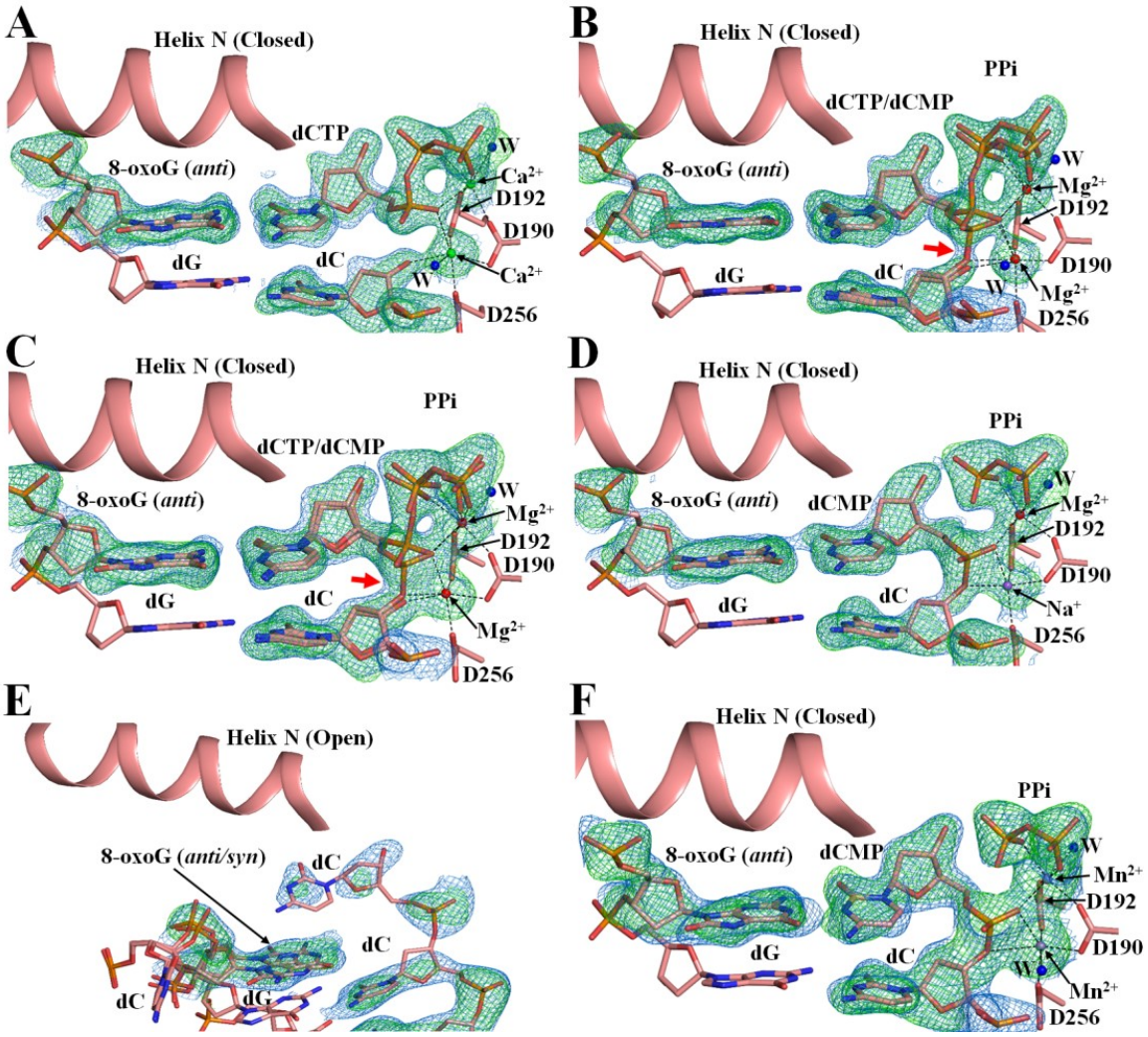


Figure 3.2: Zoomed active sites to illustrate the binding and time-dependent incorporation of dCTP opposite 8-oxoG by hPolβ.

Both $2F_o$-F_c (blue, 1σ) and F_o-F_c (green, 3σ) maps are shown for the templating 8-oxoG, dCTP or incorporated dCMP, PPi, the metal ions at the A-site and B-site, and the primer 3′-terminal nucleotide. Each red arrow indicates the electron density for the newly formed phosphodiester bond. Water molecules are shown as blue spheres. The displayed

(Continued)

Figure 3.2: Continued

structures are of the pre-catalytic ternary complex (A) and those after the pre-catalytic ternary complex crystals were soaked with either 200 mM Mg^{2+} for 30 s (B), 60 s (C), 80 s (D), and 1 hr (E), or with 200 mM Mn^{2+} for 35 s (F). In (E), the last three nucleotides at the primer 3′-terminus were modeled in two different conformations.

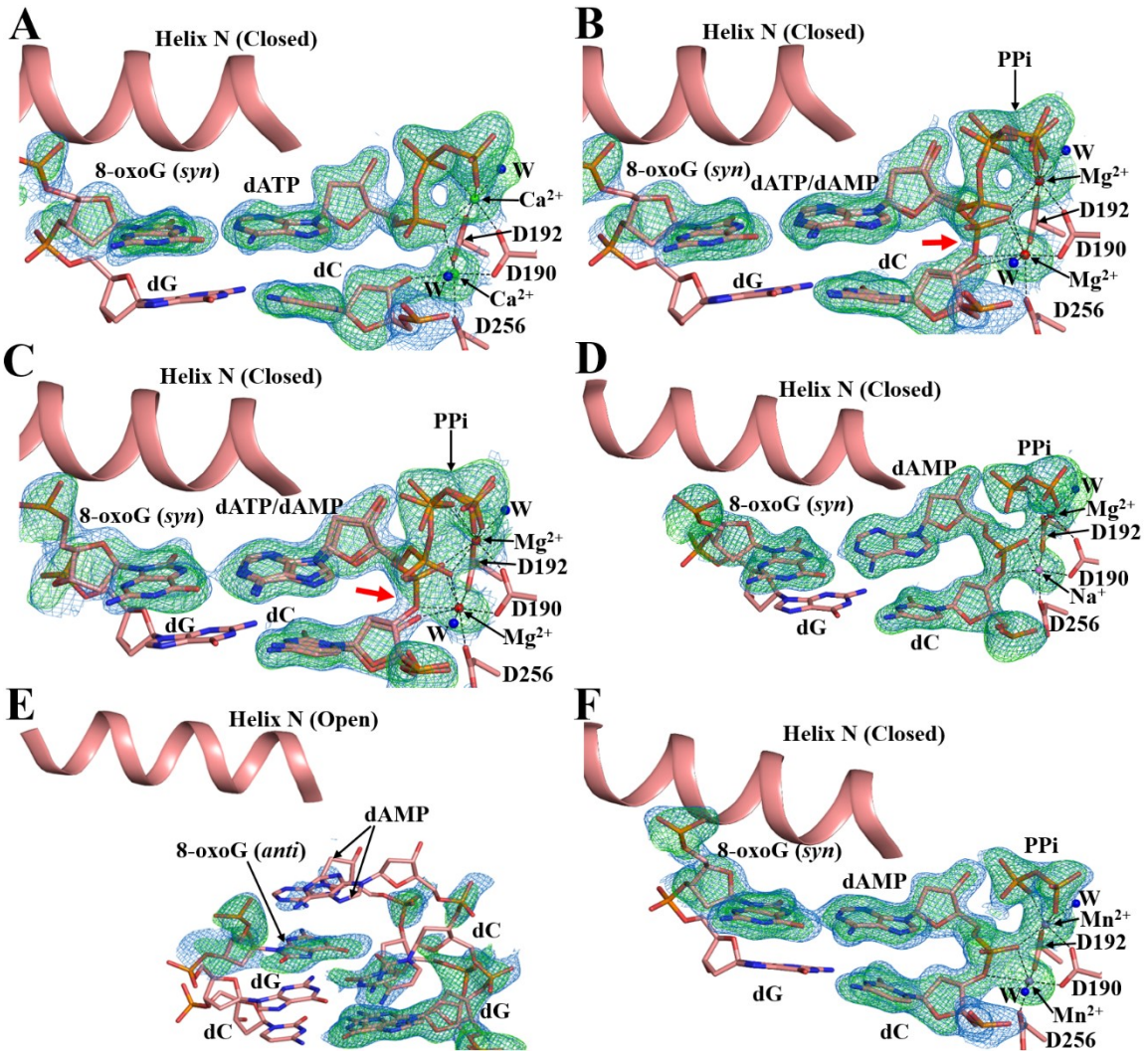


Figure 3.3: Zoomed active sites to illustrate the binding and time-dependent incorporation of dATP opposite 8-oxoG by hPolβ.

Both $2F_o$-F_c (blue, 1σ) and F_o-F_c (green, 3σ) maps are shown for the templating 8-oxoG, dATP or incorporated dAMP, PPi, the metal ions at the A-site and B-site, and the primer 3′-terminal nucleotide. Each red arrow indicates the electron density for the newly

(Continued)

Figure 3.2: Continued

formed phosphodiester bond. Water molecules are shown as blue spheres. The displayed structures are of the pre-catalytic ternary complex (A) and those after the pre-catalytic ternary complex crystals were soaked with either 200 mM Mg^{2+} for 30 s (B), 60 s (C), 80 s (D), and 1 hr (E), or with 200 mM Mn^{2+} for 35 s (F). In (E), the last three nucleotides at the primer 3′-terminus were modeled in two different conformations.

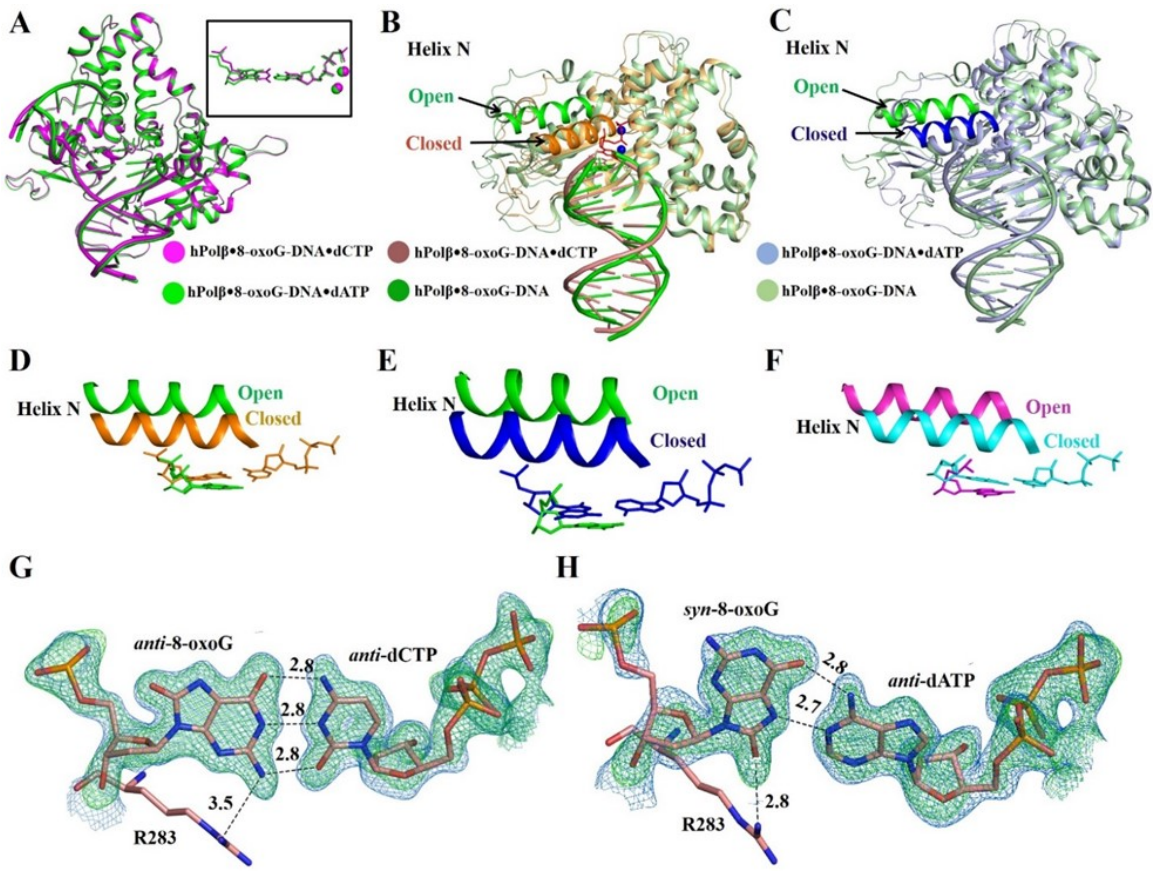


Figure 3.4: Structural differences between various ternary and binary complexes and the binding conformations of 8-oxoG at the active site of hPolβ in the pre-catalytic ternary complex of hPolβ●8-oxoG-DNA●dCTP or hPolβ●8-oxoG-DNA●dATP.

(A) Superposition of the pre-catalytic ternary complexes of hPolβ●8-oxoG-DNA●dCTP and hPolβ●8-oxoG-DNA●dATP in the presence of Ca^{2+}. The close-up view shows the overlaying of the nascent base pair and Ca^{2+} ions at the A- and B-site. (B) Superposition of the binary structure of hPolβ●8-oxoG-DNA (3RJE) and the pre-catalytic ternary

(Continued)

Figure 3.4: Continued

structure of hPolβ●8-oxoG-DNA●dCTP with Ca^{2+}. (C) Superposition of the binary structure of hPolβ●8-oxoG-DNA (3RJE) and the pre-catalytic ternary structure of hPolβ●8-oxoG-DNA●dATP with Ca^{2+}. In (B) and (C), Helix N is in the same but darker color relative to the rest of the corresponding structure. (D) Zoomed view of Helix N and the nascent base pair in (B). (E) Zoomed view of Helix N and the nascent base pair in (C). (F) Zoomed view of Helix N and the nascent base pair in the overlaid structures of the binary complex of hPolβ●undamaged DNA (magenta, 3ISB) and the ternary complex of hPolβ●undamaged DNA●dCTP (cyan, 4KLD). (G) The electron density map for the nascent *anti*-8-oxoG:*anti*-dCTP Watson-Crick base pair in the pre-catalytic ternary complex of hPolβ●8-oxoG-DNA●dCTP with Ca^{2+}. (H) The electron density map for the nascent *syn*-8-oxoG:*anti*-dATP Hoogsteen base pair in the pre-catalytic ternary complex of hPolβ●8-oxoG-DNA●dATP with Ca^{2+}. In (G) and (H), the $2F_o$-F_c (blue) and F_o-F_c (green) maps were contoured at 1σ and 3σ, respectively.

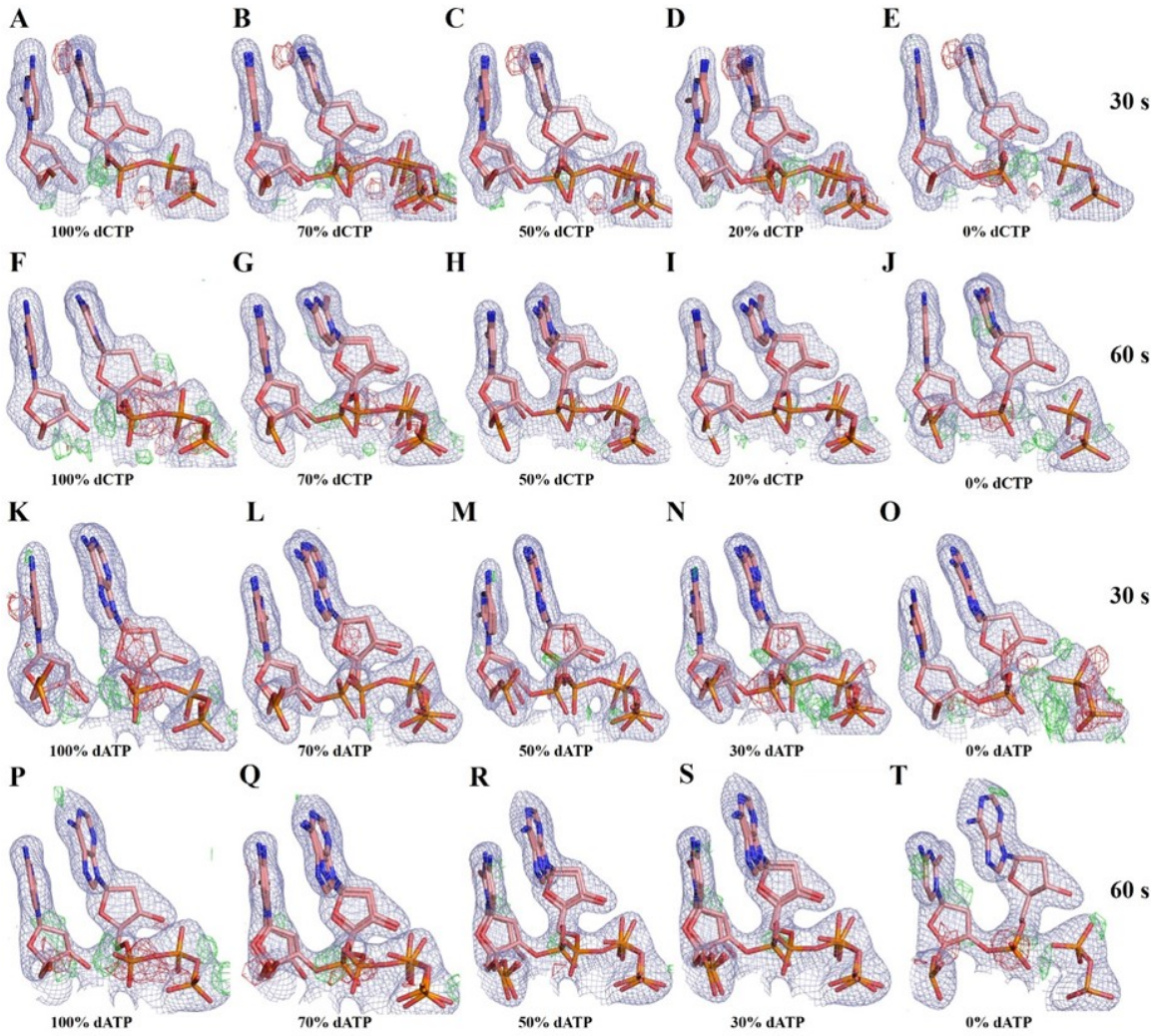


Figure 3.5: Modeling of reactant-, reaction- and product-states during phosphodiester bond formation after 30 or 60 s of Mg^{2+}/Ca^{2+} ion-exchange.

The $2F_o$-F_c (light blue) maps contoured at 1σ and the F_o-F_c omit maps contoured at either 3σ (green) or -3σ (red) are presented for the primer 3′-terminal nucleotide, incoming nucleotide, incorporated nucleotide, and pyrophosphate. Only the occupancy of the reactant dCTP/dATP is listed below each structure. Strong positive (green) and negative

(Continued)

Figure 3.5: Continued

(red) electron density blobs between the primer 3'-OH and the α-phosphate group of dCTP/dCMP, between the α and β-phosphate groups of dCTP, or between the phosphates of dCMP and pyrophosphate indicate unsatisfactory modeling, *e.g.* the modeling of the reactants at 100% (A), 70% (B), 20% (D), and 0% (E) occupancies for the 30 s structure and 100% (F), 70% (G), 50% (H), and 0% (J) occupancies for the 60 s structure. In contrast, the absence of any positive or negative electron density with the modeling of the reactants at 50% (C) and 20% (I) occupancies suggests satisfactory modeling for the 30 and 60 s structures, respectively. Similarly, for dATP incorporation, the modeling of the reactants at 100% (K), 50% (M), 30% (N), and 0% (O) occupancies for the 30 s structure and 100% (P), 70% (Q), 50% (R), and 0% (T) occupancies for the 60 s structure is unsatisfactory while the modeling of the reactants at 70% (L) and 30% (S) occupancies for the 30 and 60 s structures, respectively, is satisfactory.

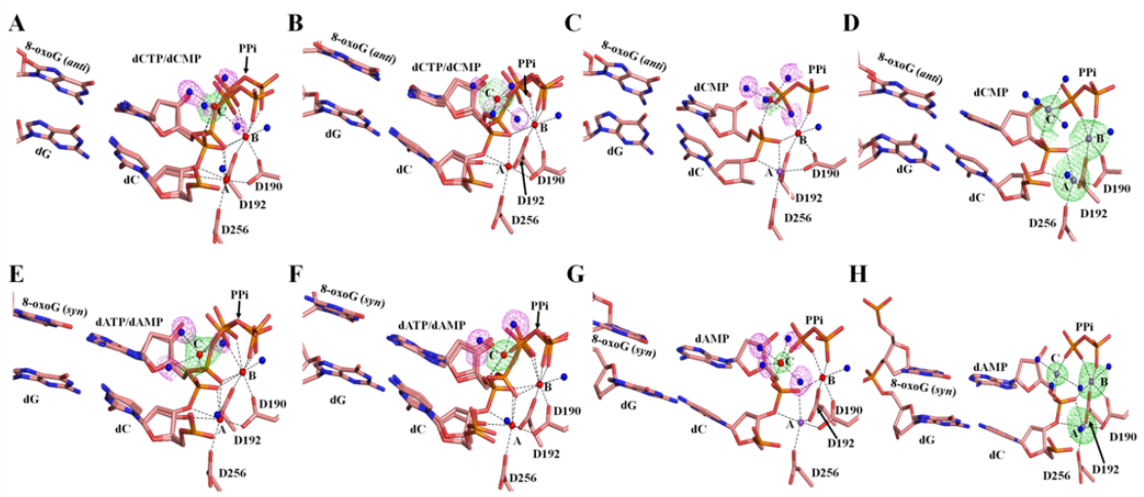


Figure 3.6: Three divalent metal ion binding sites at the active site of hPolβ during dCTP or dATP incorporation opposite 8-oxoG.

After the crystals of the pre-catalytic ternary complexes of hPolβ●8-oxoG-DNA●dCTP and hPolβ●8-oxoG-DNA●dATP were soaked with Mg^{2+} for 30, 60 and 80 s, dCTP incorporation had completed by 50% (A), 80% (B) and 100% (C), respectively, while dATP incorporation had finished by 30% (E), 70% (F) and 100% (G), respectively. When the crystals of the pre-catalytic ternary complexes of hPolβ●8-oxoG-DNA●dCTP and hPolβ●8-oxoG-DNA●dATP were soaked with Mn^{2+} for 35 s, both dCTP (D) and dATP (H) were 100% incorporated. For panels (A) to (C) and (E) to (G), the F_o-F_c omit maps were contoured at 3σ to show the electron density (green) of Mg^{2+} at the C-site and the electron density (pink) for associated water molecule ligands. The F_o-F_c omit maps contoured at 5σ show the electron densities (green) of the three Mn^{2+} ions after dCTP (D) and dATP (H) incorporation. The side chains of the aspartate residues at the active

(Continued)

Figure 3.6: Continued

site are shown as stick models while the water molecule ligands (blue) and the metal ions of Mg^{2+} (red), Mn^{2+} (grey), and Na^{+} (purple) are denoted as spheres. For description of coordinating ligands and distances, please see Table 3.3.

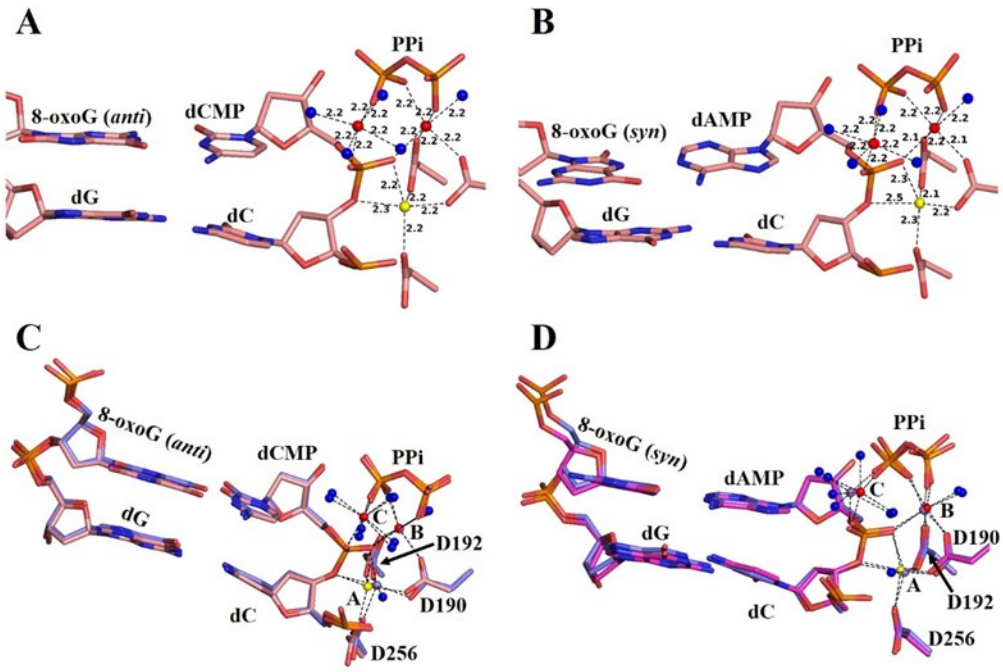


Figure 3.7: Comparison of structures captured at different time points while incorporating either dCTP or dATP opposite 8-oxoG.

(A) Active site metal ion coordination distances for the dCTP incorporation product-state structure after Mg^{2+}/Ca^{2+} ion-exchange for 80 s. (B) Active site metal ion coordination distances for the dATP incorporation product-state structure after Mg^{2+}/Ca^{2+} ion-exchange for 80 s. (C) Superposition of the dCTP incorporation product-state structures after soaking crystals with Mg^{2+} for 80 s (salmon) and with Mn^{2+} for 35 s (blue). (D) Superposition of the dATP incorporation product-state structures after soaking crystals with Mg^{2+} for 80 s (pink) and with Mn^{2+} for 35 s (blue). From (A) to (D), active site metal ions, including Mg^{2+} (red), Mn^{2+} (purple) and Na^{+} (yellow), and water molecule ligands (blue) are represented as spheres. Note that the spheres for the metal ions are not representative of their relative ionic radii.

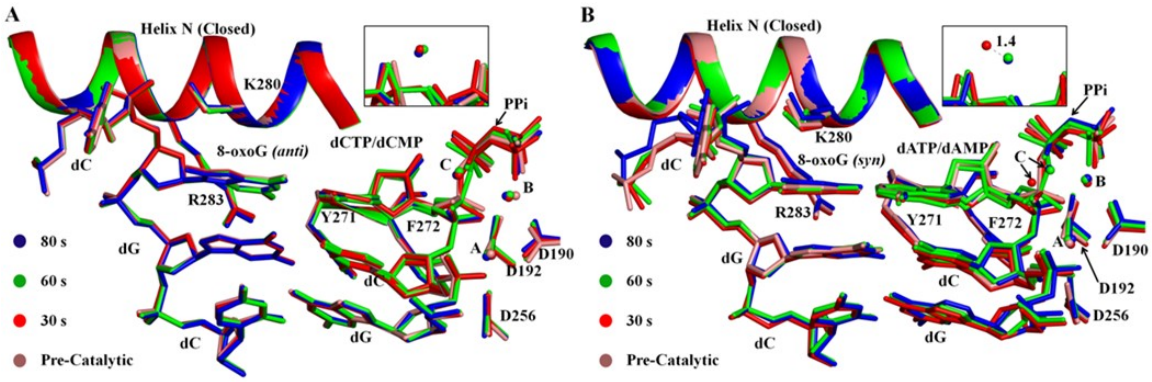


Figure 3.8: Zoomed active sites in the ternary complexes with either Ca^{2+} or Mg^{2+} at the A-, B-, and C-site.

Superposition of the pre-catalytic (Ca^{2+}) ternary structure of hPolβ●8-oxoG-DNA●dCTP (A) or hPolβ●8-oxoG-DNA●dATP (B) with its corresponding structures after soaking the pre-catalytic ternary complex crystals with Mg^{2+} for 30, 60 and 80 s shows the differences at the active site of hPolβ. The three metal ion sites are marked as A, B and C and shown as spheres. The RMSD values for the hPolβ backbone in the 30, 60 and 80 s structures relative to the pre-catalytic complex structure are 0.13, 0.17 and 0.25 Å in (A), and 0.14, 0.29 and 0.49 Å in (B). The close-up views show the location of the C-site metal ion at a different angle. In (A), the C-site metal ion is at the same position. In contrast, the C-site metal ion in (B) moved 1.4 Å from its position in the 30 s structure (red sphere) to its position in the 60 (green sphere) and 80 s (blue sphere) structures.

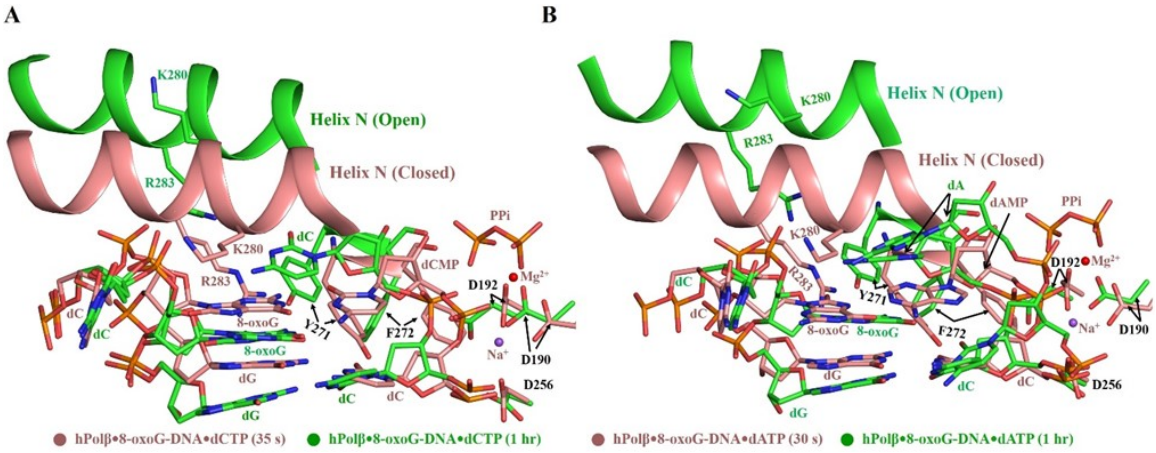


Figure 3.9: Structural overlay of the active sites of the product-state ternary complex and the post-catalytic binary complex.

(A) Superposition of the closed ternary structure of the dCTP incorporation product-state complex (hPolβ●8-oxoG-DNA$_{+1}$●PP$_i$, brown) after Mg^{2+}/Ca^{2+} ion-exchange for 80 s and the open binary structure of the post-catalytic complex (hPolβ●8-oxoG-DNA$_{+1}$, green) after 1 hr Mg^{2+}/Ca^{2+} ion-exchange. (B) Superposition of the closed ternary structure of the dATP incorporation product-state complex (hPolβ●8-oxoG-DNA$_{+1}$●PP$_i$, brown) after Mg^{2+}/Ca^{2+} ion-exchange for 80 s and the open binary structure of the post-catalytic complex (hPolβ●8-oxoG-DNA$_{+1}$, green) after 1 hr Mg^{2+}/Ca^{2+} ion-exchange. For the 80 s product-state structures (hPolβ●8-oxoG-DNA$_{+1}$●PP$_i$), the A-site metal ion is modeled as a Na^+ ion (purple sphere) while the B-site metal ion is modeled as an Mg^{2+} ion (red sphere). 8-oxoG-DNA$_{+1}$ indicates that the 8-oxoG-DNA substrate was elongated by one nucleotide as a result of either dCTP or dATP incorporation and became a nicked DNA substrate.

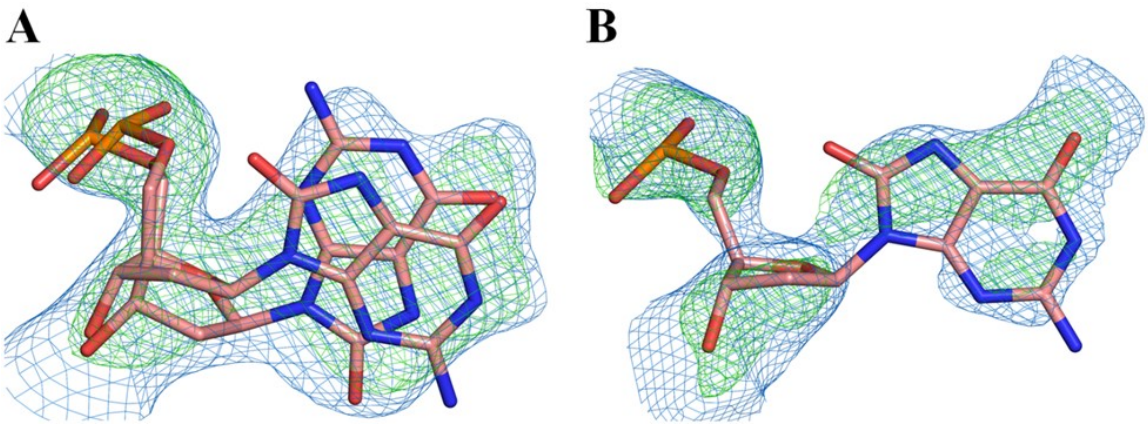


Figure 3.10: Electron density maps for the modeling of 8-oxoG conformation in the open binary complex of hPolβ●8-oxoG-DNA$_{+1}$ after 1 hr Mg^{2+}/Ca^{2+} ion-exchange. (A) Both *anti*- and *syn*-conformations of the templating nucleotide 8-oxoG were modeled after dCTP incorporation. (B) Only the *anti*-conformation of the templating nucleotide 8-oxoG was modeled after dATP incorporation. For (A) and (B), both $2F_o$-F_c (blue, 1σ) and F_o-F_c (green, 3σ) maps are shown.

Chapter 4: Time-Dependent Extension from an 8-Oxoguanine Lesion by Human DNA Polymerase Beta

Author Contributions

Andrew J. Reed partially conceived the work and performed all experiments, interpreted all data, and wrote/edited the manuscript. Zucai Suo partially conceived the work and helped with editing the manuscript.

4.1 Abstract

The oxidative DNA lesion, 7,8-dihydro-2′-deoxyguanine (8-oxoG), often occurs in double-stranded DNA and poses a threat to genomic integrity due to the ability of 8-oxoG to form stable Watson-Crick base pairs with deoxycytidine (8-oxoG:dC) and Hoogsteen base pairs with deoxyadenosine (8-oxoG:dA). In humans, short-patch base excision repair of 8-oxoG:dA base pairs requires human DNA polymerase β (hPolβ) to bypass 8-oxoG. Previously, we have shown hPolβ-catalyzed 8-oxoG bypass to exhibit low fidelity and identified a unique stacking interaction between the newly incorporated nucleotide (dCMP or dAMP) and the templating 8-oxoG. The effect of this stacking on the ability of hPolβ to extend from 8-oxoG during long-patch base excision repair was unknown. Here we report pre-steady-state kinetics and time-dependent crystal structures to demonstrate that extension from both 8-oxoG:dC and 8-oxoG:dA base pairs is 18- to 580-fold less efficient compared to 8-oxoG bypass, and that extension from 8-oxoG:dC over 8-oxoG:dA is favored by 15-fold. The overall decrease in efficiency of extension relative to 8-oxoG bypass is due to an alternative nucleotide binding conformation in the pre-catalytic ternary structures (hPolβ•DNA•dNTP) for both extension contexts, wherein the incoming nucleotide is bound in either the canonical Watson-Crick base pair or in a non-planar base pair. In addition, the decreased stability of the ternary complex of 8-oxoG:dA extension results in further loss of efficiency when compared to 8-oxoG:dC extension. Therefore, we hypothesize that the inefficient extension from 8-oxoG:dA serves as a newly discovered fidelity checkpoint during base excision repair.

4.2 Introduction

Oxidative damage to DNA occurs frequently and most commonly results in oxidation of guanine bases within double-stranded DNA or in nucleotide pools to form 8-oxo-7,8-dihydro-2′-deoxyguanine (8-oxoG) lesions.[277,278] These lesions are particularly mutagenic as both Watson-Crick base pairs with deoxycytidine (dC) and Hoogsteen base pairs with deoxyadenosine (dA) readily form.[246] Often replicative DNA polymerases will encounter an 8-oxoG lesion and incorrectly incorporate dA opposite the lesion, thus resulting in G to T transversions in subsequent rounds of replication. 8-oxoG lesions are repaired through short-patch base excision repair (spBER), wherein a damage specific DNA glycosylase, human oxoguanine DNA glycosylase 1 (hOGG1), removes the lesion base when paired with dC (8-oxoG:dC), followed by the actions of apyrmidinic/apurinic endonuclease 1 (Ape1), human DNA polymerase β (hPolβ), and a DNA ligase to complete repair of the DNA strand.[297,298] To repair 8-oxoG paired with dA (8-oxoG:dA), spBER occurs in a similar manner with the exception that the undamaged dA base is removed by an adenine DNA glycosylase MYH.[297,298] In this repair context hPolβ must faithfully incorporate dCTP opposite a templating 8-oxoG (0 position in the template) to ensure that dATP is not futilely incorporated and removed by spBER.[299] However, we have shown through pre-steady-state kinetics[226] and time-dependent X-ray crystallography[42] that hPolβ exhibits low fidelity opposite 8-oxoG and will efficiently incorporate both dCTP and dATP opposite the lesion. In addition, we have shown hPolβ bypasses 8-oxoG lesions using the three-metal-ion mechanism detailed in several time-dependent crystallography studies.[38-42,51] Interestingly, following lesion bypass, we

observed the newly incorporated nucleotide and the templating 8-oxoG base stacking rather than forming a stable base pair.[42] A transition to long-patch BER (lpBER)[300-302], wherein 2-11 nucleotides are incorporated following the activities of the DNA glycosylase and Ape1, is suggested to occur when hPolβ is unable to process the downstream primer strand[300] or as a response to depletion of cellular ATP levels.[303] If lpBER were to become necessary, we hypothesized that this observed base stacking would inhibit the transition by making further incorporations and strand displacement less efficient. Furthermore, since hPolβ incorporates both dCTP and dATP opposite the lesion to similar extents, the efficiency of extending these 8-oxoG lesion bypass products may act as a secondary fidelity checkpoint wherein slow extension from 8-oxoG:dA base pairs would facilitate completion of spBER and allow a second chance to remove the dA and subsequently incorporate a correct dC. Here we generated a DNA substrate containing an 8-oxoG:dC or 8-oxoG:dA base pair at the primer-template junction (-1 position) and utilized pre-steady-state kinetics and time-dependent crystallography to show that extension from 8-oxoG:dC base pairs is more efficient compared to 8-oxoG:dA base pairs but extension from either pair is less efficient than bypass of 8-oxoG. Moreover, the stacking interaction between the incorporated dCMP or dAMP and the templating 8-oxoG observed following 8-oxoG bypass likely weakens nucleotide binding affinity, demonstrated through the multiple dNTP binding conformations seen in the pre-catalytic ternary structures (hPolβ•DNA•dNTP) for extension from 8-oxoG:dC or 8-oxoG:dA base pairs.

4.3 Materials and Methods

4.3.1 Preparation of hPolβ and 8-oxoG extension DNA substrates

hPolβ was expressed and purified as described previously.[226] hPolβ for pre-steady state kinetic assays was stored in a final buffer (50 mM Tris–HCl (pH 7.5 at 9 °C), 50% (v/v) glycerol, 2 mM dithiothreitol, 100 mM NaCl). hPolβ for crystallization was further buffer exchanged (20 mM Bis-Tris (pH 7.0), at 22 °C, 2 mM dithiothreitol, 100 mM NaCl) and concentrated to 15 mg/mL as described previously.[42] The DNA substrates used consisted of a 16-mer template (5′-CCG ACA 8-oxoG GCG CAT CAG-3′), a 5-mer downstream 5′-phosphorylated primer (5′-^{P}GTC GG-3′), and one of two 10-mer upstream primers for either extension from 8-oxoG:dC (5′-CTA ATG CGC C-3′) or 8-oxoG:dA (5′-CTA ATG CGC A-3′). For crystallization the three oligonucleotides were annealed in a 1:1:1 ratio to a final concentration of 1 mM by mixing and heating to 95 °C and slowly cooling to room temperature. For pre-steady-state kinetic assays the oligonucleotides were similarly annealed but using a 16-mer:5-mer:10-mer ratio of 1:1.15:1.25 at a final concentration of 2 μM. The individual oligonucleotides were purchased from Integrated DNA Technologies and annealed to generate a single-nucleotide gapped DNA substrate with either an 8-oxoG:dC or 8-oxoG:dA base pair at the primer-template junction.

4.3.2 Pre-steady-state single nucleotide incorporation assays

Assays for pre-steady-state kinetic analysis were performed as described previously.[226] Briefly, a pre-incubated solution of hPolβ (300 nM) and 5′-[^{32}P]-labeled 8-oxoG:dC or 8-oxoG:dA extension DNA (30 nM) was reacted with variable concentration of correct nucleotide dTTP in buffer (50 mM Tris-HCl (pH 7.8 at 25 °C), 5 mM $MgCl_2$,

50 mM NaCl, 0.1 mM EDTA, 5 mM DTT, 10% glycerol, and 0.1 mg/ml of BSA) at 37 °C. The reaction was quenched at increasing time points with 0.37 M EDTA using a rapid chemical quench apparatus.[304] Reaction products were resolved using denaturing polyacrylamide gel electrophoresis (17% (w/v) acrylamide, 8 M urea) and quantified using ImageQuant. Product concentration was plotted against time for each nucleotide concentration and fit to a single- (dA extension, [product] = A[1 – exp(- k_{obs}t)]) or double-exponential (dC extension, [product] = A_1[1 – exp(- k_{obs1}t)] + A_2[1 – exp(- k_{obs2}t)]) equation to yield both a reaction amplitude (A) and an observed rate constant of nucleotide incorporation (k_{obs}). k_{obs} was then plotted against nucleotide concentration and fit to a hyperbolic equation ($k_{obs} = k_p$[dNTP]/(K_d + [dNTP])) to yield the maximum nucleotide incorporation rate constant (k_p) and apparent equilibrium dissociation constant (K_d).

4.3.3 Crystallization and time-dependent *in crystallo* nucleotide incorporation

Binary complex crystals of hPolβ and 8-oxoG:dC or 8-oxoG:dA extension DNA were formed as reported previously.[42] Ternary complex crystals were formed by soaking binary complex crystals in cryosolution (50 mM imidazole (pH 7.5 at 22 °C), 16–18% PEG3350, 90 mM sodium acetate, 15% ethylene glycol, 50 mM $CaCl_2$) containing non-catalytic Ca^{2+} ions and correct nucleotide, dTTP (2 mM). Pre-catalytic ternary complex crystals were then frozen in liquid nitrogen for data collection. To initiate nucleotide incorporation, pre-catalytic ternary complex crystals were transferred to a second cryosolution (50 mM imidazole (pH 7.5 at 22 °C), 16–18% PEG3350, 90 mM sodium acetate, 15% ethylene glycol) containing 200 mM $MgCl_2$. Crystals were soaked for

variable times and the reaction was quenched by rapidly freezing the crystals in liquid nitrogen. X-ray diffraction data were collected using LRL-CAT beamline facilities at Advance Photon Source (APS), Argonne National Laboratory. The X-ray diffraction data were processed using XDS[305] or iMosflm.[282] The structures were solved through the Phaser[283] molecular replacement method with 4KLD[40] as a starting model. Model building and refinement were carried out using Coot[286] and REFMAC5[285], respectively. Procheck[287] was used to evaluate the quality of the final models. All electron density omit maps were calculated using Phenix.[284] Figures were generated using PYMOL.[288]

4.4 Results and Discussion

4.4.1 Pre-steady-state kinetics of hPolβ-catalyzed extension from an 8-oxoG lesion

We have previously examined the efficiency and fidelity of 8-oxoG bypass by hPolβ by pre-steady-state kinetic analysis[226] and time-dependent X-ray crystallography.[42] To evaluate how efficiently hPolβ is able to extend from 8-oxoG bypass products we performed single-nucleotide incorporation assays following the extension from both 8-oxoG:dC and 8-oxoG:dA base pairs, the two major products of hPolβ-catalyzed 8-oxoG bypass. From these assays, the kinetic parameters for the maximal nucleotide incorporation rate constant (k_p) and apparent equilibrium dissociation constant (K_d) were determined and are summarized in Table 4.1. Unexpectedly, assays for the extension from 8-oxoG:dC base pairs exhibited biphasic kinetics under single-turnover conditions and were fit to a double-exponential equation (Figure 4.1A and 4.1B). While the first phase showed a dependence on nucleotide concentration, the second phase did not

demonstrate an obvious correlation (Figure 4.1D), suggesting that a subset of the enzyme population binds nucleotide in a non-productive ternary complex ($(E•DNA_n•dNTP)^N$, Scheme 4.1). Therefore, we hypothesize that the second phase represents a slower conformational change (k_1, Scheme 4.1) to a productive ternary complex ($(E•DNA_n•dNTP)^P$, Scheme 4.1) prior to nucleotide incorporation (k_2, Scheme 4.1). Through plotting the observed rate constants (k_{obs}) of the first phase against nucleotide concentration, it was determined that extension from 8-oxoG:dC base pairs with correct dTTP had a similar k_p (3.3 ± 0.4 s^{-1}, Table 4.1) to the bypass of 8-oxoG with dCTP (6.4 ± 0.7 s^{-1}, Table 4.1)[226] or dATP (7.5 ± 0.7 s^{-1}, Table 4.1)[226] but exhibited an 8- to 20-fold decrease in nucleotide binding affinity ($K_d = 220 \pm 70$ μM, Table 4.1). Conversely, the extension from an 8-oxoG:dA base pair demonstrated monophasic kinetics under single-turnover conditions (Figure 4.2A and 4.2B) and resulted in a 103- to 120-fold slower k_p (0.062 ± 0.003 s^{-1}, Table 4.1) and a moderate 2.3- to 5.5-fold loss of nucleotide binding affinity (61 ± 8 μM). Although monophasic kinetics are observed for extension from 8-oxoG:dA, a similar non-productive population ($(E•DNA_n•dNTP)^N$, Scheme 4.1) cannot be ruled out as the average rate of the second phase from the 8-oxoG:dC extension (~0.03 s^{-1}, k_1, Scheme 4.1) is similar to the k_p (k_2, Scheme 4.1) for extension from 8-oxoG:dA. Therefore, at low nucleotide concentrations, the presumed second phase conformational change would be faster than the observed rates (k_{obs}) of nucleotide incorporation (Figure 4.2A and 4.2B) and consequently would not be observed as it would no longer be rate-limiting for 8-oxoG:dA extension. Overall, extension from either lesion-containing primer-template junction base pair is less efficient (18- to 580-fold) than bypass of 8-

oxoG with extension from 8-oxoG:dC base pairs favored by 15-fold over extension from 8-oxoG:dA base pairs (Table 4.1). The significant decrease in extension efficiency from 8-oxoG:dA could potentially serve as a fidelity checkpoint for BER in which, inefficient extension of the mismatched base pair would favor the subsequent steps of spBER, rather than continuous primer extension. Once spBER has completed, the incorrect dA opposite 8-oxoG will be removed by MYH which provides a second opportunity for hPolβ to incorporate dCTP opposite the lesion. On the other hand, the small drop in efficiency for 8-oxoG:dC extension suggests that lpBER could compete against spBER, with either BER pathway resulting in the removal of the 8-oxoG lesion by hOGG1 in the next round of BER. Therefore, the extension preference from 8-oxoG:dC base pairs by hPolβ, shown here, would likely increase the overall fidelity of BER.

4.4.2 Pre-catalytic ternary complex structures for extension from an 8-oxoG lesion

To determine a structural basis for the observed changes in kinetics described above, and to evaluate the structural effects of an 8-oxoG lesion at the -1 position in the template strand, we determined structures of the pre-catalytic ternary complexes for the extension from 8-oxoG:dC (Figure 4.3A and 4.3B, Table 4.2) and 8-oxoG:dA (Figure 4.4A and 4.4B, Table 4.3) base pairs. Interestingly, in crystals for both extension contexts, two binding conformations for the incoming nucleotide, dTTP, were observed. The first nucleotide binding conformation is the canonical Watson-Crick (WC) planar base pair between the incoming dTTP and templating dA and can be associated with the productive ternary complex ($(E•DNA_n•dNTP)^P$) in Scheme 4.1 determined from the

kinetic assays above. The second conformation is a non-planar (NP) base pair between the incoming dTTP and templating dA. Based on the above pre-steady-state kinetic experiments, this conformation likely represents the non-productive ternary complex ($(E•DNA_n•dNTP)^N$) in Scheme 4.1. Therefore, the conformational transition to the WC productive complex from the NP non-productive complex is related to k_1 in Scheme 4.1. Similarly, the WC and NP structures for 8-oxoG:dC extension were solved to 2.20 and 2.58 Å, respectively (Figure 4.3A and 4.3B, Table 4.2), and the WC and NP structures for 8-oxoG:dA extension were solved to 2.33 and 2.50 Å, respectively (Figure 4.4A and 4.4B, Table 4.3). The WC structures for both extension contexts superimpose well with previously determined pre-catalytic ternary structures and show well-defined electron density from the F_o-F_c omit maps for the primer terminal nucleotide (dC or dA) and the incoming nucleotide base and triphosphate moieties (Figures 4.3B and 4.4B). In contrast, the NP structures only have well-defined electron density for the base of primer terminal nucleotide and the triphosphate group of the incoming nucleotide, while the ribose of the primer terminal nucleotide and the base of the incoming nucleotide are less-defined (Figures 4.3A and 4.4A). Surprisingly, in both WC and NP structures for either extension context, the 8-oxoG at the -1 position does not significantly alter the DNA backbone or enzyme structure (Figure 4.5), unlike when 8-oxoG is at the template position and is base paired with an incoming dCTP.[42,246] Therefore, differences in active site structure and the observed decreases in nucleotide incorporation efficiency during extension are not due to either the *anti*- or *syn*-conformation of 8-oxoG at the -1 template position.

In addition to the non-planar base pair geometry, comparison of the NP and WC structures shows that the side chains of residues Y271 and F272 exist in alternate conformations in the NP structures (Figure 4.6A and 4.6D). These changes in the polymerase active site suggest that an active site conformational rearrangement must occur prior to nucleotide incorporation and likely accounts for the nucleotide concentration-independent phase (k_1, Scheme 4.1) in the kinetic studies of extension from 8-oxoG:dC (Figure 4.1). Moreover, since extension from an 8-oxoG:dA base pair also exhibits the same altered conformation and the distance between the primer 3′-OH and α-phosphate of the incoming dTTP is lengthened by 0.2 Å from the WC to NP structure (Figure 4.6), a similar conformational change (k_1, Scheme 4.1) would presumably need to occur for nucleotide incorporation to proceed. However, since biphasic kinetics for extension were not observed (Figure 4.2), nucleotide incorporation in the WC conformation must be comparably slow to this conformational change. Furthermore, the less defined electron density for the incoming nucleotide and the longer and weaker hydrogen bonds in the dTTP:dA base pair suggest that NP conformation may be inherently less stable and would likely account for the observed 2.3- to 20–fold decrease in nucleotide binding affinity for extension from both 8-oxoG:dC and 8-oxoG:dA base pairs compared to 8-oxoG lesion bypass (Table 4.1).

While the different nucleotide binding conformations account for the observed decreased K_d, we hypothesized that the 53-fold lower k_p for 8-oxoG:dA extension compared to 8-oxoG:dC (Table 4.1) likely arises due to a conformationally flexible ternary complex. Accordingly, average B-factors for the incoming nucleotide, templating

base, and primer-template junction base pair are higher in both the WC and NP structures for extension from an 8-oxoG:dA base pair when compared to WC and NP structures for extension from an 8-oxoG:dC base pair (Figure 4.7). This change in B-factors suggests that the 8-oxoG:dA primer-template junction base pair and the base pair between the incoming dTTP and the templating dA are dynamic. However, the similar hydrogen bond distances in the NP and WC structures for 8-oxoG:dA extension (Figure 4.6E and 4.6F), unlike those for 8-oxoG:dC extension, likely results in a significant higher population being bound in the non-productive conformation. Therefore, only a small population of the ternary complex may exist in a catalytically competent conformation, leading to a decrease in incorporation rate during extension from an 8-oxoG:dA base pair. This is consistent with a destabilization of the *syn*-conformation of 8-oxoG in which a hydrogen bond between side chain of R283 of the N-helix and the O8 atom of 8-oxoG during bypass (0 position) by dATP incorporation is significantly longer (0.5 Å) and presumably weaker when the 8-oxoG is at the -1 position (Figure 4.8).

4.4.3 Time-dependent extension from and 8-oxoG lesion

To visualize time-dependent nucleotide incorporation, pre-catalytic ternary crystals for 8-oxoG:dC and 8-oxoG:dA extension were transferred to a cryosolution containing catalytic Mg^{2+} for various times before quenching in liquid nitrogen. Structures for 8-oxoG:dC extension (Table 4.2) were solved at 20 (15%, 1.90 Å), 60 (80%, 2.05 Å), and 80 s (100%, 2.10 Å) of reaction completion and show well-defined electron density at each time point (Figure 4.3C-4.3E). Structures for 8-oxoG:dA extension (Table 4.3) were solved at 90 (80%, 1.70 Å) and 120 s (100%, 1.87 Å) of

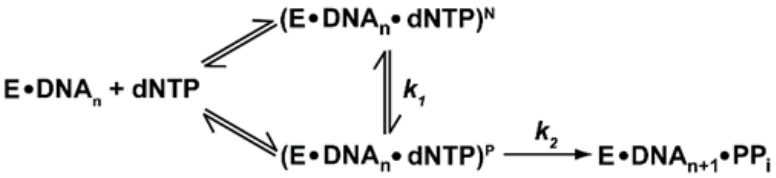
reaction completion with well-defined electron density at both time points (Figure 4.4C and 4.4D). All five structures have WC base pair geometry and show the canonical divalent metal ions at the A- and B-sites and all were modeled as Mg^{2+} (Figures 4.3, 4.4, 4.9, and 4.10). In addition, the *anti*- and *syn*-conformation of 8-oxoG in the 8-oxoG:dC and 8-oxoG:dA extension structures is maintained throughout nucleotide incorporation and all active site residue side chains are in orientations similar to previous time-dependent crystallographic studies with hPolβ.[40-42] Moreover, similar to the bypass of 8-oxoG with dCTP and dATP,[42] a 3rd divalent metal ion is observed during extension from both 8-oxoG:dC (Figure 4.9) and 8-oxoG:dA base pairs (Figure 4.10). During 8-oxoG:dA extension, the electron density for the third divalent metal ions are well-defined after 80 and 100% incorporation, with all 6 coordinating ligands visible, and were modeled as Mg^{2+} ions (Figure 4.10). However, at 15 % incorporation for 8-oxoG:dC extension the electron density is not as well-defined for the third divalent metal ion and for all time points only 4 coordinating ligands are observed (Figure 4.9). Despite lacking the expected 6 coordinating ligands, we chose to model this metal ion as Mg^{2+} based on the consistent identification of a similar divalent metal ion in recent time-dependent crystallography studies.[38-42,51] Furthermore, superposition of the 8-oxoG:dC extension structures with a partial nucleotide incorporation structure from the bypass of 8-oxoG with dCTP (PDB: 4RPY)[42] or superposition of the time-dependent structures following extension of 8-oxoG:dA base pairs with bypass of 8-oxoG with dATP (PDB: 4RQ5)[42] show that the 3rd divalent metal ions occupy the same position (Figure 4.11). Thus, this

study provides further evidence for the utilization of a three-metal-ion mechanism by DNA polymerases.[38-42,51]

4.5 Conclusions

Here we investigated the extension from 8-oxoG bypass products to determine the effect of the previously identified base stacking following lesion bypass[42] and the translocation of the 8-oxoG lesion to the -1 template position on nucleotide incorporation and active site structure. Pre-catalytic ternary structures of 8-oxoG extension show that the incoming nucleotide was bound through both Watson-Crick and non-planar base pairing with the latter likely leading to the observed decreased nucleotide binding affinity. We hypothesize that the non-planar base pairing is not competent for catalysis and must undergo a conformational change prior to nucleotide incorporation, which is consistent with the nucleotide concentration-independent phase observed in the single-turnover kinetic experiments. In addition, extension from 8-oxoG:dA base pairs was less efficient compared to extension from 8-oxoG:dC base pairs due to the presumed instability of the ternary complex, suggesting that the extension step could act as an additional fidelity checkpoint during BER, wherein extension from 8-oxoG:dA to initiate lpBER would be inefficient and therefore favor termination of spBER, allowing a subsequent round of spBER to repair the incorrect base pair. Conversely, efficient extension from 8-oxoG:dC base pairs suggests that either lp or spBER would be efficient and would result in removal of 8-oxoG in further rounds of BER. Finally, we showed that extension from 8-oxoG utilizes the three-metal-ion mechanism, providing further evidence for the existence of a third divalent metal ion during nucleotide incorporation.

4.6 Schemes



Scheme 4.1: Kinetic pathway for nucleotide binding and incorporation during 8-oxoG extension by hPolβ.

Following formation of the binary complex (E•DNA$_n$), incoming nucleotide (dTTP) is bound in a productive ((E•DNA$_n$•dNTP)^P) or non-productive ((E•DNA$_n$•dNTP)^N) ternary complex. The productive and non-productive ternary complexes can transition between one another (k_1), with only the productive leading to nucleotide incorporation (k_2).

4.7 Tables

Nucleotide	k_p (s^{-1})	K_d (μM)	k_p/K_d ($\mu M^{-1}s^{-1}$)	Efficiency Ratio†
8-oxoG:dC Extension				
dTTP	3.3 ± 0.4	220 ± 70	0.015	1
8-oxoG:dA Extension				
dTTP	0.062 ± 0.003	61 ± 8	1.0×10^{-3}	0.067
8-oxoG Bypass*				
dCTP	6.4 ± 0.7	11 ± 4	0.58	1
dATP	7.5 ± 0.7	27 ± 6	0.28	0.48

An incoming nucleotide was incorporated by hPolβ at 37 °C opposite dA in the single-nucleotide gapped 8-oxoG:dC/dA DNA substrates for the extension assays or opposite 8-oxoG for the bypass assays.
†Defined as $(k_p/K_d)_{8\text{-oxoG:dN}}/(k_p/K_d)_{8\text{-oxoG:dC}}$ for extension or $(k_p/K_d)_{dNTP}/(k_p/K_d)_{dCTP}$ for bypass.
*Reference 226.

Table 4.1: Pre–steady-state kinetic parameters for single-nucleotide incorporation during bypass or extension from an 8-oxoG lesion by hPolβ at 37 °C.

	Non-Planar Pre-catalytic Ternary	Watson-Crick Pre-catalytic Ternary	Extension After 20 s	Extension After 60 s	Extension After 80 s
Data collection*					
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell dimension					
a, *b*, *c* (Å)	50.7, 79.5, 55.3	51.0, 80.1, 55.6	50.9, 80.1, 55.6	50.7, 79.8, 55.3	50.6, 80.3, 55.4
α, β, γ (°)	90, 107.1, 90	90, 107.2, 90	90, 107.2, 90	90, 107.2, 90	90, 107.6, 90
Resolution (Å)	52.85–2.58 (2.60–2.58)	48.72–2.20 (2.40–2.20)	48.62–1.90 (2.20–1.90)	52.83–2.05 (2.10–2.05)	52.81–2.10 (2.20–2.10)
R_{meas}†	0.091 (0.557)	0.081 (0.475)	0.069 (0.317)	0.058 (0.246)	0.092 (0.438)
$I/\sigma I$	11.5 (2.0)	11.5 (2.5)	11.1 (3.4)	11.9 (3.8)	8.8 (2.5)
Completeness (%)	98.0 (100)	94.4 (89.1)	96.5 (98.7)	93.1 (90.8)	96.8 (98.7)
Redundancy	1.98 (1.97)	1.97 (1.88)	1.94 (1.92)	1.91 (1.68)	1.93 (1.91)
Refinement					
Resolution (Å)	52.85–2.58	48.72–2.2	48.62–1.90	52.83–2.05	52.81–2.10
No. reflections	12586	20062	31519	24166	23241
R_{work}/R_{free}‡	0.234/0.306	0.187/0.260	0.203/0.243	0.197/0.242	0.202/0.262
No. atoms					
Protein	2594	2631	2678	2638	2619
DNA	683	661	671	671	661
Water	48	136	305	221	145
B-factors (Å^2)					
Protein	35.32	32.14	26.32	29.69	31.76
DNA	43.44	35.89	32.96	35.24	38.21
Water	28.78	30.99	32.54	34.83	32.26
R.m.s deviations					
Bond lengths (Å)	0.004	0.013	0.005	0.009	0.010
Bond angles (°)	0.893	1.742	1.148	1.428	1.490
Reaction ratio					
Ratio of RS/PS**	1.0/0.0	1.0/0.0	0.85/0.15	0.2/0.8	0.0/1.0
PDB ID	5VRW	5VRX	5VRY	5VRZ	5VS0

*Highest resolution shell is shown in parenthesis.

**RS and PS are abbreviations for the reactant state and product state.

†R_{merge} = Redundancy independent R-factor.

‡R value = $\Sigma||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c are observed and calculated structure factor amplitudes, respectively.

Table 4.2: Structure parameters for extension from 8-oxoG:dC base pairs.

	Non-Planar Pre-catalytic Ternary	Watson-Crick Pre-catalytic Ternary	Extension After 90 s	Extension After 120 s
Data collection*				
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell dimension				
a, *b*, *c* (Å)	50.6, 80.5, 55.3	50.9, 80.5, 55.7	50.4, 79.6, 55.2	50.8, 79.7, 55.3
α, β, γ (°)	90, 106.9, 90	90, 107.3, 90	90, 107.1, 90	90, 107.1, 90
Resolution (Å)	52.91–2.50 (2.60–2.50)	53.15–2.33 (2.41–2.33)	52.76–1.70 (1.90–1.70)	48.55–1.87 (1.90–1.87)
R_{merge}† or R_{meas}‡	0.071 (0.532)‡	0.065 (0.159)†	0.054 (0.498)‡	0.052 (0.532)‡
$I/\sigma I$	13.0 (2.3)	12.4 (5.6)	13.1 (2.3)	13.5 (2.1)
Completeness (%)	96.4 (95.3)	99.7 (99.9)	94.9 (96.9)	93.5 (95.4)
Redundancy	1.98 (1.96)	3.3 (3.3)	1.93 (1.94)	1.92 (1.89)
Refinement				
Resolution (Å)	52.91–2.50	48.64–2.33	52.76–1.70	48.55–1.87
No. reflections	13761	17477	42694	32063
R_{work}/R_{free}'	0.211/0.272	0.187/0.250	0.206/0.245	0.199/0.257
No. atoms				
Protein	2584	2609	2672	2693
DNA	663	663	663	654
Water	50	147	337	340
B-factors (Å²)				
Protein	38.35	24.69	25.87	25.93
DNA	45.64	31.84	34.43	30.00
Water	32.26	24.76	35.08	33.24
R.m.s deviations				
Bond lengths (Å)	0.009	0.012	0.007	0.009
Bond angles (°)	1.417	1.797	1.372	1.598
Reaction ratio				
Ratio of RS/PS**	1.0/0.0	1.0/0.0	0.2/0.8	0.0/1.0
PDB ID	5VS1	5VS2	5VS3	5VS4

*Highest resolution shell is shown in parenthesis.
**RS and PS are abbreviations for the reactant state and product state.
†$R_{merge} = \Sigma|I-\langle I\rangle|/\Sigma I$, where I is the integrated intensity of each reflection.
‡R_{meas} = Redundancy independent R-factor.
′R value = $\Sigma||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c are observed and calculated structure factor amplitudes, respectively.

Table 4.3: Structure parameters for extension from 8-oxoG:dA base pairs.

4.8 Figures

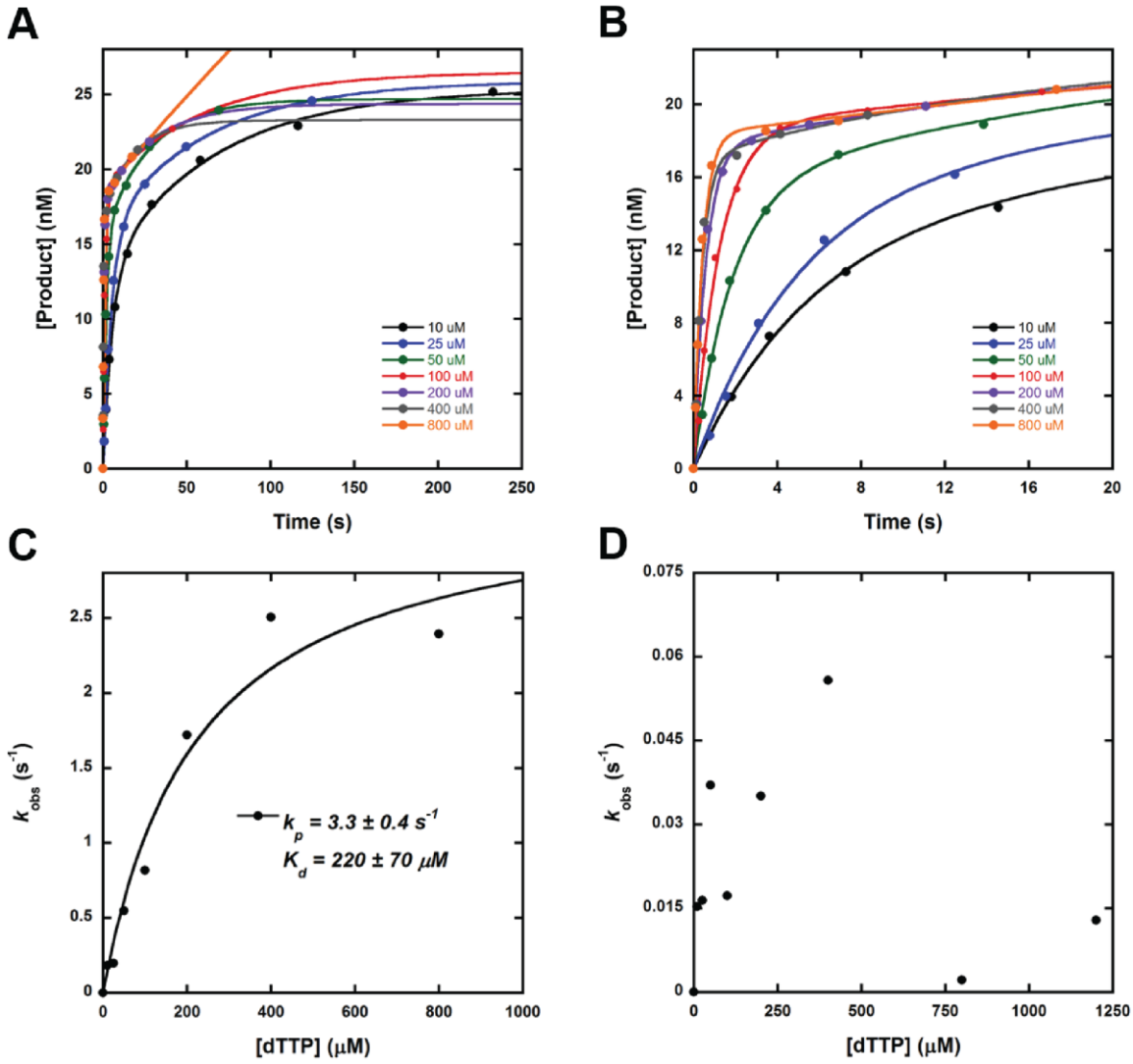


Figure 4.1: Extension from an 8-oxoG:dC base pair under single-turnover conditions.

(A) A pre-incubated solution of hPolβ and [^{32}P]-labeled-8-oxoG:dC extension DNA was mixed with varying concentrations of dTTP at 37 °C. After various times, the reaction

(Continued)

Figure 4.1: Continued

was stopped with 0.37 M EDTA and analyzed polyacrylamide gel electrophoresis. Each time course demonstrated biphasic kinetics of product formation was fit to a double-exponential equation, [product] = $A_1[1 – \exp(- k_{obs1}t)] + A_2[1 – \exp(- k_{obs2}t)]$, to yield a reaction amplitude (A) and an observed rate constant of nucleotide incorporation (k_{obs}) for each phase. (B) Plots from (A) zoomed into fast time points to show the fit of the first phase. (C) The k_{obs} values for the first phase were then plotted against nucleotide concentration and the data were fit to a hyperbolic equation, $k_{obs} = k_p[dNTP]/([dNTP] + K_d)$, to yield an apparent equilibrium dissociation constant ($K_d = 220 \pm 70$ μM) and a maximum nucleotide incorporation rate constant ($k_p = 3.3 \pm 0.4$ s^{-1}). (D) The k_{obs} values for the second phase were then plotted against nucleotide concentration and did not show a dependence.

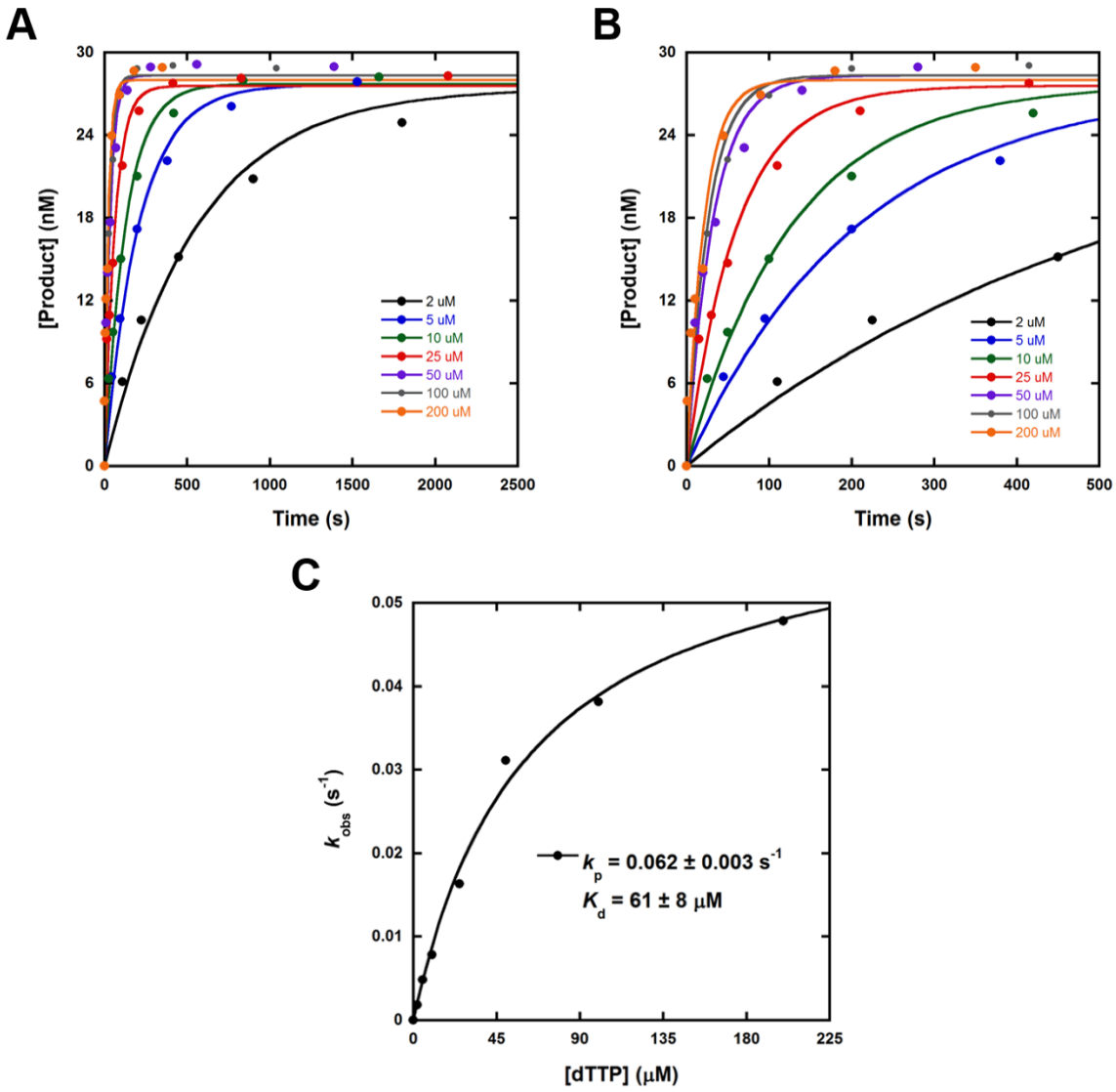


Figure 4.2: Extension from an 8-oxoG:dA base pair under single-turnover conditions.

(A) As in Figure S1, single-nucleotide incorporation assays were performed but with a

(Continued)

Figure 4.2: Continued

pre-incubated solution of hPolβ and [^{32}P]-labeled-8-oxoG:dA extension DNA and the product concentration at variable nucleotide concentration plotted against time. Each time course demonstrated monophasic kinetics and were fit to a single-exponential equation, [product] = A[1 – exp(- $k_{obs}t$)]. (B) Plots from (A) zoomed to fast time points to show the fit early in the time courses. (C) The k_{obs} values were then plotted against nucleotide concentration and the data were fit to a hyperbolic equation, $k_{obs} = k_p$[dNTP]/([dNTP] + K_d), to yield an apparent equilibrium dissociation constant (K_d = 61 ± 8 μM) and a maximum nucleotide incorporation rate constant (k_p = 0.062 ± 0.003 s^{-1}).

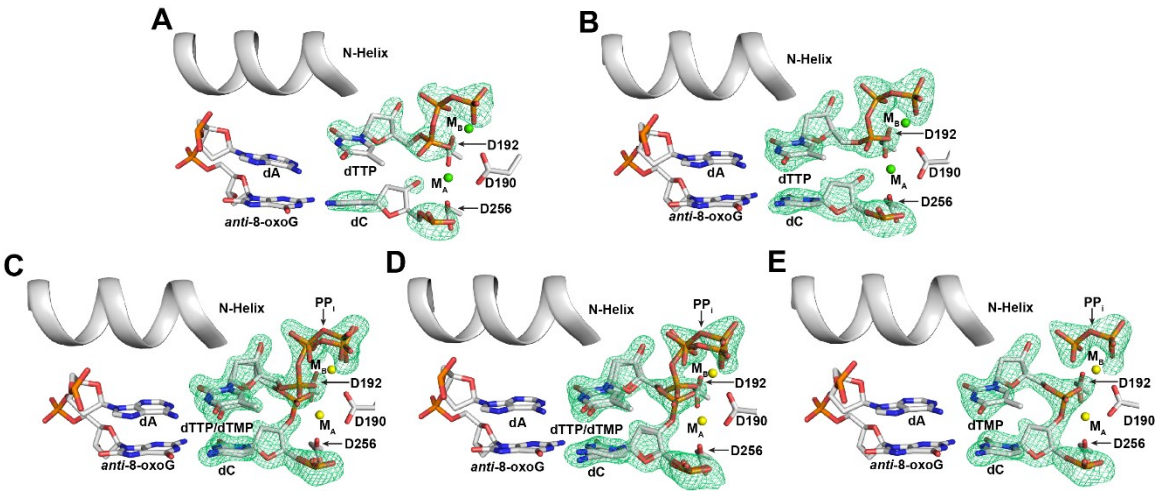


Figure 4.3: Time-dependent extension from an 8-oxoG:dC base pair.

Active site views of hPolβ bound to 8-oxoG:dC extension DNA and correct nucleotide dTTP. (A and B) Pre-catalytic ternary complex with alternative non-planar (A) and Watson-Crick base pair (B). (C-E) Time-dependent incorporation of dTTP after 20 (C), 60 (D), and 80 s (E) showing 15, 80, and 100 % incorporation. The N-helix of the thumb domain is shown as a cartoon and is in the closed conformation in all panels. F_o-F_c omit electron density maps (3σ, green mesh) are shown for the primer terminal base and the incoming/incorporated dTTP. Ca^{2+} and Mg^{2+} ions at the A- (M_A) and B-site (M_B) are shown as green and yellow spheres, respectively.

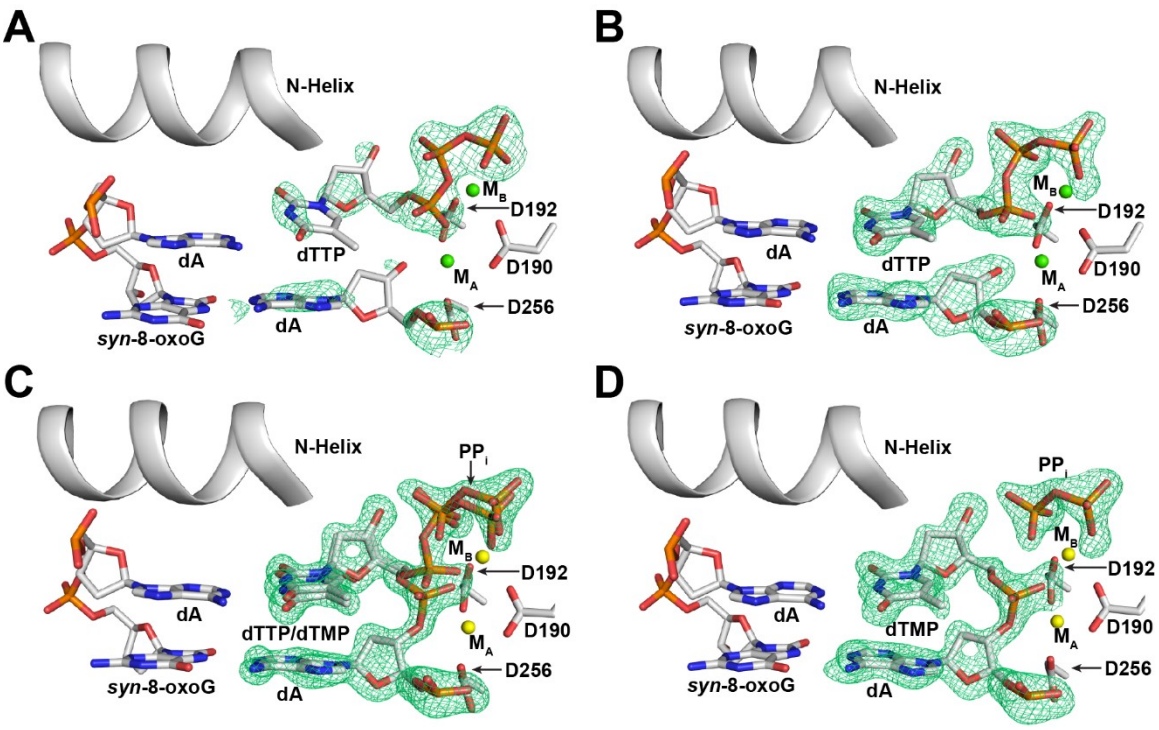


Figure 4.4: Time-dependent extension from an 8-oxoG:dA base pair.

Active site views of hPolβ bound to 8-oxoG:dA extension DNA and correct nucleotide dTTP. (A and B) Pre-catalytic ternary complex with alternative non-planar (A) and Watson-Crick base pair (B). (C and D) Time-dependent incorporation of dTTP after 90 (C) and 120 s (D) showing 80 and 100 % incorporation. The N-helix of the thumb domain is shown as a cartoon and is in the closed conformation in all panels. F_o-F_c omit electron density maps (3σ, green mesh) are shown for the primer terminal base and the incoming/incorporated dTTP. Ca^{2+} and Mg^{2+} ions at the A- (M_A) and B-site (M_B) are shown as green and yellow spheres, respectively.

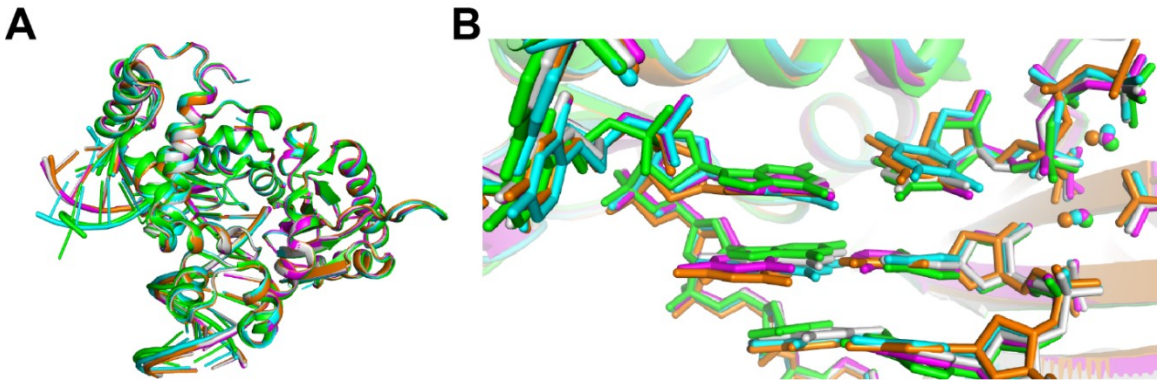


Figure 4.5: Superposition of the pre-catalytic ternary structures for extension from 8-oxoG lesions with undamaged DNA pre-catalytic ternary structure.

(A) Overall structural superposition of WC 8-oxoG:dC (gray), NP 8-oxoG:dC (cyan), WC 8-oxoG:dA (magenta), NP 8-oxoG:dA (orange), and undamaged DNA (4KLD, green). (B) Zoomed view of template strand showing that the 8-oxoG at the -1 position does not alter the local DNA structure.

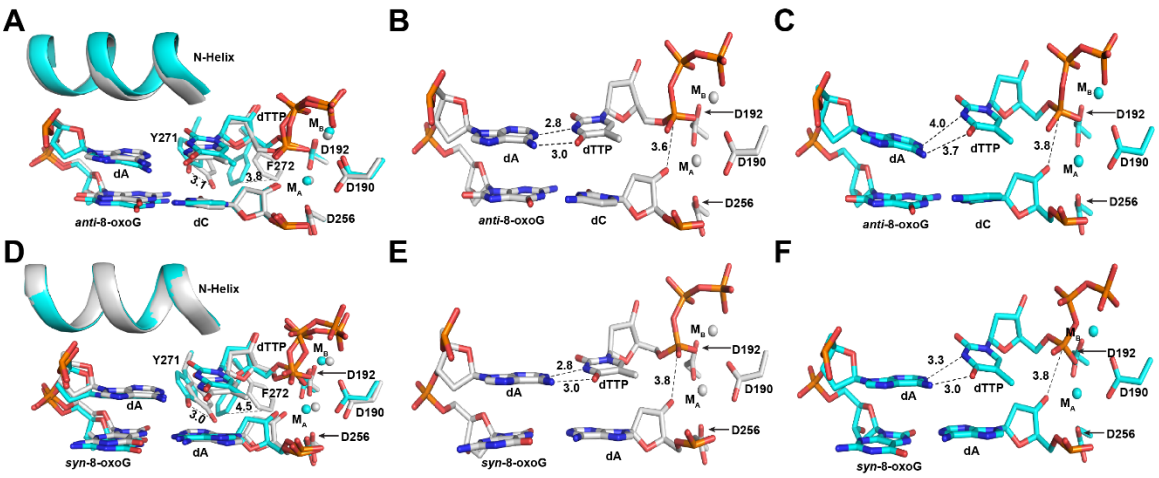


Figure 4.6: Comparison of Watson-Crick (WC) and non-planar (NP) pre-catalytic ternary structures.

(A and D) Superposition of WC (grey) and NP (cyan) active sites for extension from 8-oxoG:dC (A) and 8-oxoG:dA (D) base pairs. Distances between location of Y271 and F272 side chains in WC and NP are shown as a dashed lines. (B and E) Active sites of WC structures for 8-oxoG:dC (B) and 8-oxoG:dA (E) extension. (C and F) Active sites of NP structures for 8-oxoG:dC (C) and 8-oxoG:dA (F) extension. (B, C, E, and F) Base pair hydrogen bonds and distance between primer 3′-OH and incoming nucleotide α-phosphate are shown as dashed lines. All distances are given in Å.

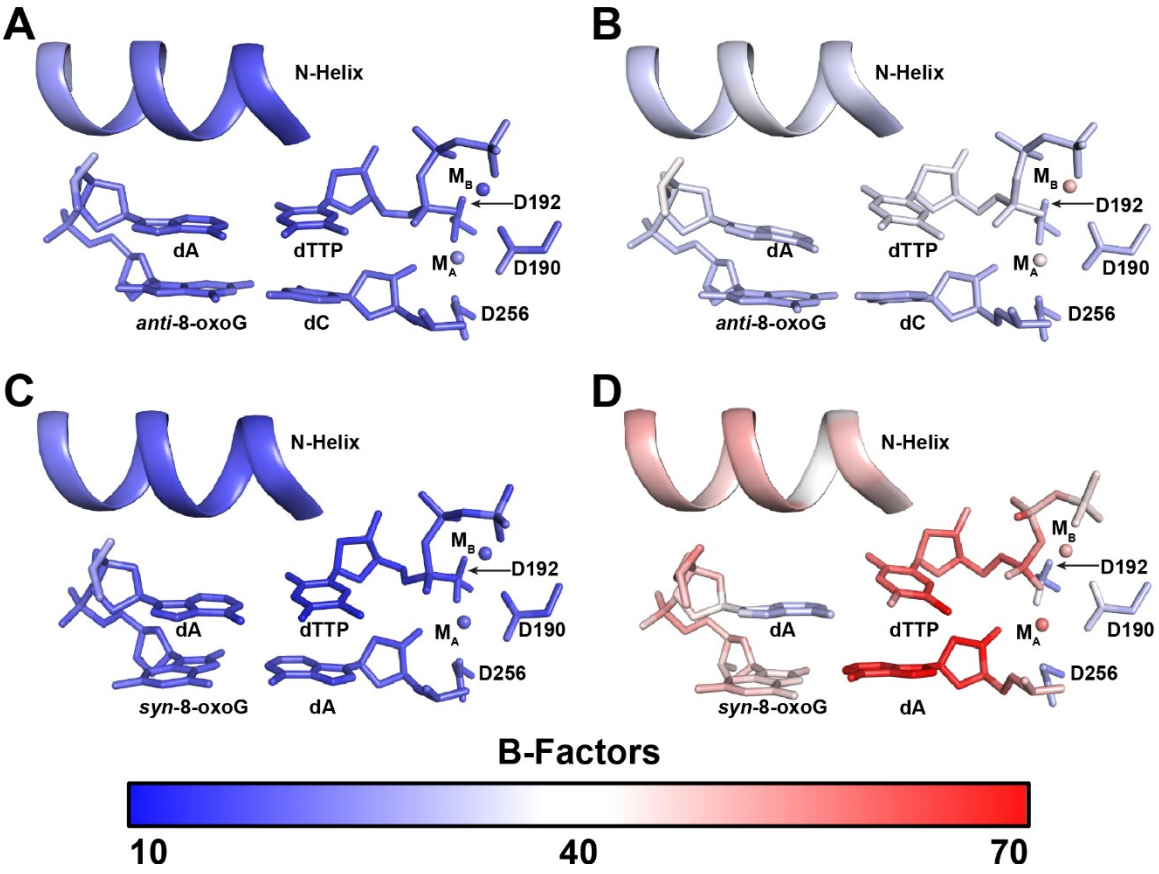


Figure 4.7: Average B-factors of the pre-catalytic ternary structures for extension from 8-oxoG lesions.

B-factors for each structure are shown as a spectrum. (A) WC 8-oxoG:dC. (B) NP 8-oxoG:dC. (C) WC 8-oxoG:dA. (D) NP 8-oxoG:dA. DNA, incoming/incorporated nucleotide, and catalytic aspartate side chains are shown as sticks. The N-helix is shown as cartoon and is in the closed conformation in all structures. Metal ions are shown as spheres at the A- (M_A) and B-sites (M_B).

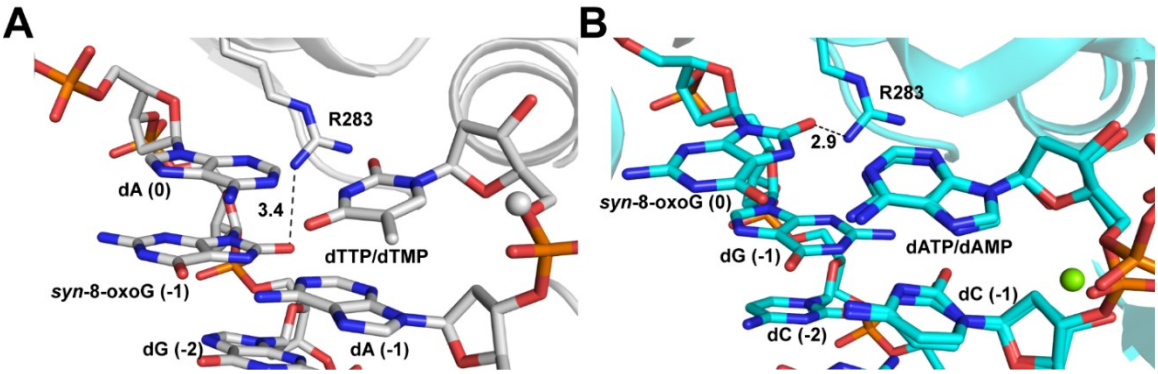


Figure 4.8: Comparison of R283 interaction in 8-oxoG:dA extension and 8-oxoG bypass with dATP.

Active site views of the (A) 90 s time point for 8-oxoG:dA extension and (B) the 60 s time point for dATP bypass of 8-oxoG (4RQ5) showing the interaction of R283 with the template residues at position 0 and -1. DNA, incoming/incorporated nucleotide, and R283 side chains are shown as sticks. Hydrogen bonds are shown as dashed lines and distances are given in Å.

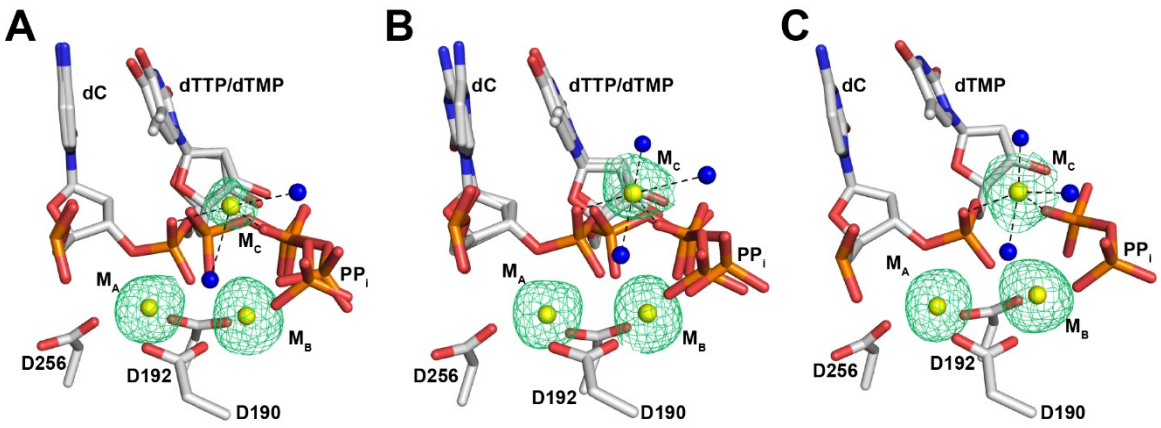


Figure 4.9: Binding of 3rd divalent metal ion during extension from 8-oxoG:dC base pair.

Active site views showing the primer terminal dC, incoming dTTP, incorporated dTMP, PP_i, catalytic aspartate residues, and divalent metal ions for incorporation after 20 s (A), 60 s (B), and 80 s (C). F_o-F_c omit electron density maps are shown for the divalent metal ions (3 σ, green mesh). Coordinating ligands for the 3rd divalent metal ion are shown as dashes. Water molecules are shown as blue spheres.

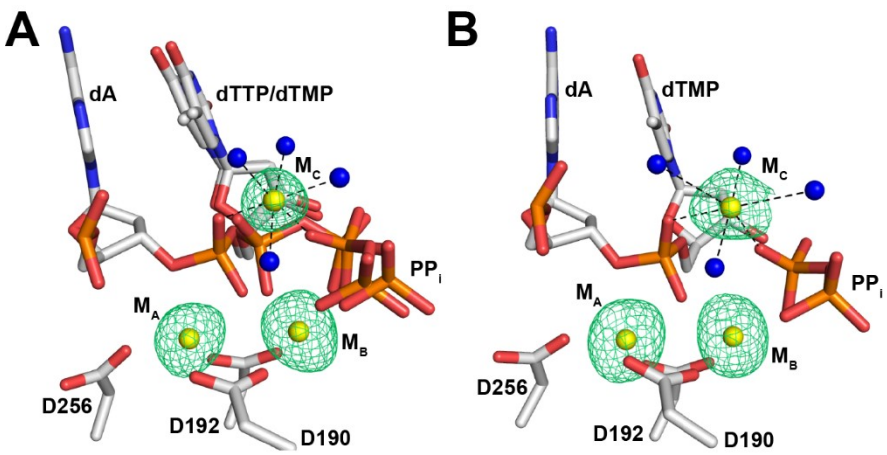


Figure 4.10: Binding of 3rd divalent metal ion during extension from 8-oxoG:dA base pair.

Active site views showing the primer terminal dA, incoming dTTP, incorporated dTMP, PP_i, catalytic aspartate residues, and divalent metal ions for incorporation after 90 s (A) and 120 s (B). F_o-F_c omit electron density maps are shown for the divalent metal ions (3 σ, green mesh). Coordinating ligands for the 3rd divalent metal ion are shown as dashes. Water molecules are shown as blue spheres.

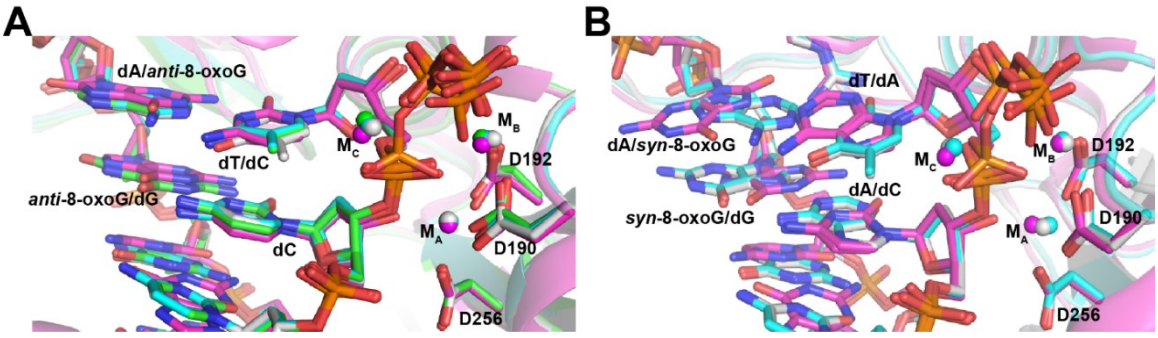


Figure 4.11: Superposition of partial incorporation structure for 8-oxoG extension with 8-oxoG bypass.

(A) Superposition of the 20 (gray), 60 (cyan), and 80 s (green) structures for 8-oxoG:dC extension with the 30 s structure for dCTP bypass of 8-oxoG (4RPY, magenta). (B) Superposition of the 90 (gray) and 120 s (cyan) structures for 8-oxoG:dA extension with the 60 s structure for dATP bypass of 8-oxoG (4RQ5, magenta). DNA, incoming/incorporated nucleotide, and catalytic aspartate side chains are shown as sticks. Metal ions are shown as spheres at the A- (M_A), B- (M_B), and C-sites (M_C) and superimpose well.

Chapter 5: Structural basis for the *D*-stereoselectivity of human DNA polymerase β

Reproduced in part with permission from Vyas, R., Reed, A.J., Raper, A.T., Zahurancik, W.J., Wallenmeyer, P.C., and Suo, Z. (2017) Structural basis for the D-stereoselectivity of human DNA polymerase β. *Nucleic Acids Research*. 45 (10), 6228-6237. DOI: 10.1093/nar/gkx252.

Author Contributions

Rajan Vyas and Andrew J. Reed contributed equally to this manuscript. R.V. crystallized protein complexes and solved the structures. A.J.R. interpreted the structures, developed the model for *D*-stereoselectivity mechanism, prepared mutants for kinetics experiments, prepared figures, and wrote/edited the manuscript. Austin T. Raper helped interpret data, develop the model for *D*-stereoselectivity mechanism, and edit the manuscript. Walter J. Zahurancik and Petra C. Wallenmeyer designed and performed the kinetics experiments as well as aided in editing the manuscript. Zucai Suo conceived the research, helped develop the model for *D*-stereoselectivity mechanism, and helped edit the manuscript.

5.1 Abstract

Nucleoside reverse transcriptase inhibitors (NRTIs) with *L*-stereochemistry have long been an effective treatment for viral infections because of the strong *D*-stereoselectivity exhibited by human DNA polymerases relative to viral reverse transcriptases. The *D*-stereoselectivity of DNA polymerases has only recently been explored structurally and all three DNA polymerases studied to date have demonstrated unique stereochemical selection mechanisms. Here, we have solved structures of human DNA polymerase β (hPolβ), in complex with single-nucleotide gapped DNA and *L*-nucleotides and performed pre-steady-state kinetic analysis to determine the *D*-stereoselectivity mechanism of hPolβ. Beyond a similar 180° rotation of the *L*-nucleotide ribose ring seen in other studies, the pre-catalytic ternary crystal structures of hPolβ, DNA, and *L*-dCTP or the triphosphate forms of antiviral drugs lamivudine ((-)3TC-TP) and emtricitabine ((-)FTC-TP) provide little structural evidence to suggest that hPolβ follows the previously characterized mechanisms of *D*-stereoselectivity. Instead, hPolβ discriminates against *L*-stereochemistry through accumulation of several active site rearrangements that lead to a decreased nucleotide binding affinity and incorporation rate. The two NRTIs escape some of the active site selection through the base and sugar modifications but are selected against through the inability of hPolβ to complete thumb domain closure.

5.2 Introduction

Replication fidelity is partially determined by the preference of DNA polymerases (Pols) to form correct G:C and A:T Watson-Crick base pairs. Many factors including hydrogen bonding with the incoming nucleotide,[306,307] minor groove interactions with active site

residues,[61,308] and base stacking,[192] affect fidelity to varying degrees among different DNA Pols. The structural basis for the inherent *D*-stereoselectivity of DNA Pols is not as well-defined despite wide-spread use of numerous clinically successful antiviral nucleotide analog reverse transcriptase inhibitors (NRTIs) possessing non-natural *L*-stereochemistry. Two of the NRTIs approved to treat HIV infection are the *L*-deoxycytidine analogs, lamivudine ((-)3TC; (-)-β-*L*-2′,3′-dideoxy-3′-thiacytidine) and its 5-fluorinated derivative, emtricitabine ((-)FTC; (-)-β-*L*-2′,3′-dideoxy-5-fluoro-3′-thiacytidine), which are activated by host kinases to their respective triphosphate forms, (-)3TC-TP and (-)FTC-TP (Figure 5.1). Upon nucleotide incorporation catalyzed by viral reverse transcriptases (RTs), viral DNA synthesis is terminated due to the absence of a 3′-OH on the deoxyribose moiety of the NRTIs. Interestingly, both lamivudine and emtricitabine have been shown to be more effective at inhibiting HIV-1 RT and less toxic than their *D*-enantiomers.[309-313]

Despite the high clinical efficacy of NRTIs, there are often various drug toxicities associated with usage of these analogs. This is not surprising considering DNA Pols and RTs share a common mechanism for nucleotide incorporation and therefore, likely share a similar mechanism of nucleotide selection.[5] For example, NRTI-induced genomic instability and mitochondrial toxicity have been associated with inhibition of human B-family DNA Pols (α, δ, and ε)[314-316] and human A-family DNA Pol γ (hPolγ),[317,318] respectively. Furthermore, it is likely that DNA Pols possessing lower fidelity such as those involved in DNA damage repair and lesion bypass are able to incorporate the NRTIs more easily and disrupt DNA replication and therefore, may account for the

observed *in vivo* drug toxicities. Consistently, our recent kinetic investigations of NRTI incorporation by human DNA Pols from the X- (DNA damage repair; hPolβ and λ)[319] and Y-families (DNA lesion bypass; human Pols η, ι, κ, and Rev1)[320] demonstrate that these Pols are more susceptible to inhibition by NRTIs than replicative Pols and in some instances incorporate the NRTIs as efficiently as HIV-1 RT. The increased inhibition and efficiency of NRTI incorporation by the X- and Y-family Pols are likely attributed to the lack of 3′→5′ proofreading capability as well as a more flexible active site in each of these damage repair or lesion bypass Pols.[319,320]

Although NRTIs with *L*-stereochemistry possess known drug toxicities associated with the ability of host Pols to incorporate such drugs, only recently has a structural basis for the mechanism of *L*-nucleotide incorporation been investigated. Our recent publications detailing the mechanism of *D*-stereoselectivity for hPolλ[223] and *Sulfolobus solfataricus* DNA polymerase IV (Dpo4)[87], model X- and Y-family Pols, respectively, are the first structures to demonstrate the binding and incorporation of *L*-nucleotides and NRTIs with *L*-stereochemistry. Interestingly, hPolλ[223] and Dpo4[87] utilize different mechanisms to achieve their strong *D*-stereoselectivity of 1.2×10^4 and 4.4×10^4, respectively. Dpo4 selects against *L*-stereochemistry by driving an *L*-nucleotide to adopt non-productive triphosphate binding conformations which have moderately longer distances (2.0-4.4 Å) between the α-phosphate of the *L*-nucleotide and the reacting primer 3′-OH than the distance with a natural *D*-dNTP.[87] In comparison, hPolλ initially binds an *L*-nucleotide in a non-productive mode via unique hydrogen bonds between the side chain of R517 and the *L*-dNTP, thereby forcing the reacting 3′-OH of the primer and

the α-phosphate of the nucleotide to be much farther apart (9.1-9.3 Å).[223] This non-productive binding mode isomerizes into a productive mode prior to catalysis.[223]

It is unclear if the observed differences in *D*-stereoselectivity between Dpo4 and hPolλ arise due to the evolutionary differences between the Pol families. To examine if Pols from the same family exhibit a conserved mechanism of *D*-stereoselectivity, we have co-crystallized and solved the structures of hPolβ (Figure 5.2), an X-family Pol with high sequence homology to hPolλ, in complex with a single-nucleotide gapped DNA substrate and either *L*-dCTP, (-)3TC-TP, or (-)FTC-TP (Figure 5.1) bound at the active site. Furthermore, we measured the incorporation rates and binding affinities of wild-type hPolβ for the various *L*-nucleotides to directly compare with published kinetic studies of hPolλ[223] and Dpo4.[87] In addition, we performed similar kinetic assays with an hPolβ R283 to alanine mutant (R283A) to determine if R283 in hPolβ exhibits a similar role in nucleotide selection to the analogous R517 residue in hPolλ. Collectively, these structural and kinetic studies reveal a mechanism of *D*-stereoselectivity for hPolβ that is distinct from either of those described for hPolλ or Dpo4.

5.3 Materials and Methods

5.3.1 Preparation of Protein and DNA

Human full-length hPolβ was overexpressed and purified as described previously.[42] DNA oligomers were purchased from Integrated DNA Technologies and were purified by denaturing polyacrylamide gel electrophoresis. Crystallization oligomers consisted of a 16-mer template (5′-CCGACGGCGCATCAGC-3′), a 10-mer upstream primer (5′-GCTGATGCGC-3′), and a 5-mer downstream 5′-phosphorylated

primer (5′-pGTCGG-3′) as described previously.[40] Template, upstream primer, and downstream primer crystallization oligos were mixed in a 1:1:1 ratio and annealed by heating to 95 °C for 5 minutes followed by slowly cooling to 4 °C to form a 1 mM DNA substrate. The 21-19A-41GT-mer DNA substrate reported earlier was used for kinetic studies and consisted of a 41-mer template, a 21-mer upstream primer, and a 19-mer downstream 5′-phosphorylated primer.[319] To visualize product formation the 21-mer primer of the 21-19A-41GT-mer was radiolabeled with [γ-^{32}P]ATP and OptiKinase according to the manufacturer's protocol, and the unreacted [γ-^{32}P]ATP was subsequently removed via a Bio-Spin 6 column. The DNA substrate was annealed at a 1:1.25:1.15 ratio of upstream primer to downstream primer to template. The *L*-nucleotides, *L*-dCTP, (-)FTC-TP, and (-)3TC-TP were obtained from Jena Bioscience.

5.3.2 Pre-Steady-State Kinetic Assays

All fast reactions were performed by using a rapid chemical quench-flow apparatus (KinTek), as previously described.[319] Briefly, a pre-incubated solution of full-length wild-type hPolβ or its R283A mutant (300 nM) and 30 nM [^{32}P]-labeled-21-19A-41GT-mer was mixed with varying concentrations of a nucleotide in buffer L (50 mM Tris-HCl, pH 8.4, 5 mM $MgCl_2$, 100 mM NaCl, 0.1 mM EDTA, 5 mM DTT, 10% glycerol, and 0.1 mg/ml BSA) at 37 °C. After various times, the reaction was stopped with 0.37 M EDTA and analyzed by sequencing gel electrophoresis. Each time course of product formation was fit to a single-exponential equation, [product] = A[1 - exp(- k_{obs}t)], using KaleidaGraph (Synergy Software) to yield a reaction amplitude (A) and an observed rate constant of nucleotide incorporation (k_{obs}). The k_{obs} values were then

plotted against nucleotide concentrations and the data were fit to a hyperbolic equation, $k_{obs} = k_p[dNTP]/([dNTP] + K_d)$, to yield an apparent equilibrium dissociation constant (K_d) and a maximum nucleotide incorporation rate constant (k_p).

5.3.3 hPolβ Crystallization and Structure Determination

Purified hPolβ was dialyzed into a buffer containing 50 mM sodium acetate (pH 5.5), 75 mM NaCl, 5% glycerol and 0.1 mM DTT, and concentrated to 20 mg/ml as described previously.[78] Binary complexes were prepared by mixing purified hPolβ and DNA (1 mM) in a 1:1 ratio at 4 °C followed by heating to 35 °C and subsequent cooling to room temperature.[32] Binary crystals were obtained by hanging drop vapor diffusion against a reservoir solution containing 50 mM imidazole, pH 7.5, 16 to 18% PEG3350, and 350 mM sodium acetate.[32] Binary crystals were then seeded into freshly prepared ternary complex solution (hPolβ:DNA, 10 mM $CaCl_2$ and 2 mM *L*-dNTP) to obtain pre-catalytic ternary crystals at room temperature as described previously [79]. The resulting pre-catalytic ternary crystals were harvested and transferred briefly to a cryosolution containing 15% ethylene glycol, 50 mM imidazole (pH 7.5), 20% PEG3350, 90 mM sodium acetate before they were flash frozen in liquid nitrogen. X-ray diffraction data were collected using the LRL-CAT beamline facilities at Advance Photon Source (APS), Argonne National Laboratory. X-ray diffraction data were processed using MOSFLM[282] and structures were solved using the molecular replacement method by PHASER [283] using a previous structure (PDB code: 4KLD)[40] in the absence of ligands and solvent molecules as the initial model. Structural refinement was carried out using REFMAC5.[285]

COOT[286] was used for visualization and model building. Quality of the models was assessed using PROCHECK.[287] Figures were created using PYMOL.[288]

5.4 Results

5.4.1 Pre-steady-state kinetics of wild-type hPolβ-catalyzed *L*-nucleotide incorporation

To determine how efficiently *L*-nucleotides are bound and incorporated by hPolβ, we performed single-nucleotide incorporation assays to measure a maximum incorporation rate constant (k_p) and an apparent equilibrium dissociation constant (K_d) for *L*-dCTP (Table 5.1, Figure 5.3). This pre-steady state kinetic analysis of hPolβ with *L*-dCTP complements our previous work with hPolβ[319] incorporating (-)3TC-TP and (-)FTC-TP and confirms that each *L*-nucleotide is incorporated slower and with an efficiency (k_p/K_d) two to five orders of magnitude lower than the k_p/K_d value for *D*-dCTP (Table 5.1). Unlike hPolλ, which displays a similarly tight nucleotide binding affinity ($1/K_d$) for *L*-dCTP relative to *D*-dCTP,[223] the *L*-stereochemistry resulted in a 30-fold decrease in nucleotide binding affinity of hPolβ for *L*-dCTP compared to *D*-dCTP (Table 5.1, Figure 5.3). Surprisingly, the chemical modifications (Figure 5.1) in the sugar rings of (-)3TC-TP and (-)FTC-TP result in significantly different effects on nucleotide binding affinity. While (-)3TC-TP was bound approximately 4-fold tighter than *D*-dCTP, the additional fluorination of the base resulted in a 15-fold decrease in nucleotide binding affinity compared to *D*-dCTP (Table 5.1). Furthermore, hPolβ preferentially incorporated *D*-dCTP over *L*-dCTP with a *D*-stereoselectivity, defined as $(k_p/K_d)_{D\text{-dCTP}}/(k_p/K_d)_{L\text{-dCTP}}$, of 2.7×10^5, while the *D*-stereoselectivity was reduced to only 325 and 2,900 for the

incorporation of (-)3TC-TP and (-)FTC-TP, respectively (Table 5.1). Thus, the chemical changes in the ribose of (-)3TC-TP and (-)FTC-TP relaxed the *D*-stereoselectivity of hPolβ by 830- and 100-fold, respectively, making these *L*-nucleotide analogs significantly better substrates than *L*-dCTP.

5.4.2 Pre-steady-state kinetics of the mutant R283A of hPolβ-catalyzed *L*-nucleotide incorporation

To determine if R283 in hPolβ, analogous to R517 in hPolλ, is critical for incorporation of L-nucleotides, we generated the R283A substitution mutant and performed single-nucleotide incorporation assays to investigate the effect of this mutation on the kinetic parameters for incorporation of D-dCTP, L-dCTP, (-)3TC-TP, and (-)FTC-TP (Table 5.1, Figure 5.4). The mutation resulted in a 13-fold drop in incorporation rate for D-dCTP relative to wild-type but only resulted in a small change (0.5- to 4-fold) for the L-nucleotides (Table 5.1, Figure 5.4). Similarly, the R283A mutation resulted in a large 59-fold decrease in nucleotide binding affinity for D-dCTP relative to wild-type (Table 5.1). In addition, the binding of the L-nucleotides was affected to varying degrees with L-dCTP, (-)3TC-TP, and (-)FTC-TP exhibiting 50-, 94-, and 2-fold decreases in binding affinity, respectively. These changes resulted in respective decreases in nucleotide incorporation efficiency (kp/Kd) of 747-, 84-, 55-, and 9.6-fold for D-dCTP, L-dCTP, (-)3TC-TP, and (-)FTC-TP (Table 5.1). As a result, the D-stereoselectivity was relaxed relative to wild-type by 8.7-, 13-, and 78-fold for L-dCTP, (-)3TC-TP, and (-)FTC-TP, respectively (Table 5.1). Although the D-stereoselectivity of the R283A mutant of hPolβ is relaxed, the role of R283 in anchoring the thumb domain (Figures 5.2 and 5.5)

in the closed conformation upon nucleotide binding via a hydrogen bond with the templating base most likely leads to an overall loss in nucleotide selection upon mutation and thus all nucleotides are bound weakly regardless of their D- or L-stereochemistry.

5.4.3 Structure of hPolβ with bound DNA and *L*-dCTP

To capture a pre-catalytic ternary complex of hPolβ, DNA, and *L*-dCTP (hPolβ•DNA•*L*-dCTP), crystallization was performed in the presence of the non-catalytic metal ion, Ca^{2+} (Figure 5.6A). The pre-catalytic ternary structure of hPolβ•DNA•*L*-dCTP was refined to 1.8 Å (Table 5.2), and the active site shows well-defined electron density for the incoming *L*-dCTP (Figures 5.6A and 5.7). Similar to the canonical pre-catalytic ternary structure with *D*-dCTP (hPolβ•DNA•*D*-dCTP) (Figure 5.8A), the base of *L*-dCTP was present in the *anti*-conformation and formed three hydrogen bonds (2.7 – 3.0 Å) with the templating nucleotide dG (Figures 5.6A and 5.8B). Consistent with our earlier structural studies of hPolλ with *L*-nucleotides,[223] the constraint imposed by the *L*-stereochemistry combined with Watson-Crick base pairing caused the sugar ring of *L*-dCTP to rotate 180° relative to *D*-dCTP. Beyond the active site differences, the overall quaternary structure of hPolβ•DNA•*L*-dCTP is very similar to the hPolβ•DNA•*D*-dCTP structure (Fig. 5.9A) with an RMSD of 0.76 Å (Table 5.3). Furthermore, in the hPolβ•DNA•*L*-dCTP structure, the dG:*L*-dCTP base pair is sandwiched between the terminal template-primer junction base pair and Helix N of the thumb domain (Figure 5.6A), as is typical for closed conformation structures with a correct incoming nucleotide (Figures 5.8A and 5.9).[40]

The triphosphate moiety of *L*-dCTP was observed in the productive chair-like conformation, similar to the *D*-dCTP structure[40] and shows similar interactions with the active site divalent metal ions (Figures 5.6A, 5.8A, 5.8B, and 5.9E). Furthermore, *L*-dCTP exhibits several key active site interactions reminiscent of *D*-dCTP. For example, the side chain of R283 from Helix-N forms a hydrogen bond with the template strand, the side chain of N279 hydrogen bonds with the incoming nucleotide, and the side chain of Y271 forms a hydrogen bond with the base of the primer 3′-terminal nucleotide (Figures 5.5, 5.6A, 5.8A, and 5.8B). The only notable structural differences are the side chain rotation of residue F272, the shortening of the hydrogen bond between the backbone carbonyl of F272 and the 3′-OH of *L*-dCTP, the slight repositioning of the divalent metal ions within the active site, and a water molecule that bridges the 3′-OH of *L*-dCTP and the 3′-OH of the primer terminal nucleotide (Figures 5.8B and 5.10A). These interactions are a direct result of the ribose rotation which causes the 3′-OH of *L*-dCTP and the 3′-OH of the primer terminal nucleotide to point towards each other at a distance of 3.9 Å. Surprisingly, the distance between the primer 3'-OH and the α-phosphate of the incoming nucleotide (3.7 Å) is maintained at a distance similar to the canonical structure of hPolβ•DNA•*D*-dCTP (3.8 Å) despite a potential steric clash between the 3′-OH of *L*-dCTP and the primer 3′-OH (Figure 5.10A and 5.10B).

5.4.4 Structures of hPolβ with bound NRTIs

The pre-catalytic ternary complexes of hPolβ, DNA, and the NRTIs, (-)3TC-TP (hPolβ•DNA•(-)3TC-TP) and (-)FTC-TP (hPolβ•DNA•(-)FTC-TP), were crystallized in the presence of non-catalytic Ca^{2+}, as was done for the hPolβ•DNA•*L*-dCTP complex.

The structures of hPolβ•DNA•(-)3TC-TP and hPolβ•DNA•(-)FTC-TP were refined to 2.2 and 1.8 Å, respectively (Table 5.2). Within the active site of the hPolβ•DNA•(-)3TC-TP structure, the electron density difference map (F_O-F_C) for the triphosphate moiety is well-defined but is less evident for the ribose ring and base portion of the nucleotide (Figure 5.6B), while in the active site of the hPolβ•DNA•(-)FTC-TP structure, the electron density difference map (F_O-F_C) for the entire nucleotide is well-defined (Figure 5.6C). However, both *L*-nucleotides agree well with the $2F_O$-F_C electron density map (Figure 5.7B and 5.7C). Surprisingly, both *L*-nucleotide analogs exhibited triphosphate moieties in a productive chair-like conformation similar to the *D*-dCTP[40] and *L*-dCTP structures (Figures 5.2 and 5.9). However, in both structures the thumb domain displays moderate electron density with Helix-N observed in a semi-open conformation (Figure 5.9) that is intermediate to the open and closed conformations observed in the binary[79] and ternary[40] complex structures, respectively (Figure 5.5).

Although many interactions with the incoming nucleotide and triphosphate are maintained in the hPolβ•DNA•(-)3TC-TP and hPolβ•DNA•(-)FTC-TP structures when compared to the *D*-dCTP (Table S2, RMSD of 1.75 and 1.95 Å, respectively) and *L*-dCTP structures, the hydrogen bond between R283 and the templating nucleotide is absent and the side chain of Y271 forms a hydrogen bond with the template nucleotide, rather than the primer 3′-terminal nucleotide (Figure 5.8). In addition to these structural changes, (-)FTC-TP forms a non-planar Watson-Crick base pair with the templating dG while (-)3TC-TP does not form a Watson-Crick base pair with any nucleotide. Rather (-

)3TC-TP is facing the side chain of R283 at a distance too far to form a hydrogen bond (3.6 Å) but maintains a hydrogen bond with N279 (Figure 5.6B).

Unlike the global conformation of the thumb domain (Figure 5.3), which is intermediate to its position in the binary and the ternary structures, the active sites of the (-)3TC-TP and (-)FTC-TP structures exhibit side chain positions that resemble either the binary or ternary structures. For example, the side chain positions of F272 and D190 are similar to the binary complex while those of D192 and R258 resemble the ternary complex. This mixed active site structure, consisting of both binary and ternary features, disrupts metal ion coordination and causes the primer 3′-OH, the A-site metal ion, and α-phosphate to significantly move (2.4-3.9 Å) relative to the *D*-dCTP structure (Figures 5.9F and 5.10). As a result of these movements, the coordination number, distances, and geometries were altered for the A-site metal ion and were inconsistent with what is expected for a Ca^{2+} ion and therefore was rather modeled as a Na^{+} ion (Figure 5.9F). Remarkably, the distance from the primer 3′-OH to the α-phosphate of the nucleotide for (-)FTC-TP (3.4 Å) is comparable to the *L*-dCTP (3.7 Å) and *D*-dCTP (3.8 Å) structures despite the movement of the reacting groups and the altered metal ion geometry (Figures 5.8 and 5.10). However, the distance between reacting groups for the hPolβ•DNA•(-)3TC-TP structure (4.7 Å) is significantly lengthened compared to the other three nucleotides (Figures 5.9 and 5.10).

5.5 Discussion

5.5.1 Comparison to Previously Characterized *D*-stereoselectivity Mechanisms

To date, the *D*-stereoselectivity mechanism of hPolλ[223] and Dpo4[87] have been determined as well as the mechanism of drug selectivity of hPolγ.[321] All three studies have demonstrated that these Pols use unique methods to establish *D*-stereochemical selection. Dpo4 achieves *D*-stereoselectivity by forcing the *L*-nucleotides to adopt several non-productive triphosphate conformations.[87] hPolγ selects between (+) and (-)FTC-TP (equivalent to *D*- and *L*-enantiomers) through an altered Watson-Crick geometry, a direct result of the ribose rotation, and a steric clash between the modified ribose of (-)FTC-TP and the side chain of Y951.[321] Interestingly, Y951 of hPolγ is also a selection factor against rNTPs and is analogous to the "steric gate" residue in DNA Pols.[322,323] Previously, we identified that X-family hPolλ uses two pathways to incorporate nucleotides with *L*-stereochemistry (Figure 5.11).[223] Briefly, in Pathway I, the nucleotide is bound in a catalytically incompetent triphosphate conformation (Figure 5.11, chains I and M) and forms an unusual base pair-like hydrogen bond with the side chain of R517. This is followed by transition to a productive conformation (Figure 5.11, chains A and E) where the canonical Watson-Crick base pair is formed and the triphosphate adopts the productive chair-like conformation. Alternatively, in Pathway II the nucleotide is directly bound in the catalytically competent conformation (Figure 5.11, chains A and E). However, the direct binding of an *L*-nucleotide in a productive conformation is much less efficient, as evident by the severely decreased binding and incorporation of the *L*-nucleotides when Pathway I is abolished by mutation of R517 to alanine (Figure 5.11).[223]

Here, all the structures of hPolβ with the *L*-nucleotides have productive triphosphate binding conformations (Figures 5.6, 5.8, and 5.11) and the equivalent "steric gate" tyrosine residue, Y271, does not clash with the ribose of the *L*-nucleotides (Figure 5.8A and 5.8B). Thus, hPolβ does not utilize either of the selection mechanisms observed with Dpo4[87] or hPolγ.[321] It was expected that hPolβ would follow a mechanism of *D*-stereoselectivity similar to that of hPolλ[223] due to the high degree of sequence, structural, and functional similarity between the C-terminal Polβ-like domain of hPolλ and the full-length hPolβ.[226] However, (-)FTC-TP and *L*-dCTP, in the pre-catalytic ternary structures of hPolβ presented here, bind with Watson-Crick base pairs and productive triphosphate conformations that are very similar to those of hPolλ in chain A and E of Pathway II (Figures 5.6, 5.8, 5.11, and 5.12). Interestingly, the hPolβ•DNA•(-)3TC-TP structure in the presence of Ca^{2+} binds a nucleotide in a position reminiscent to that observed in the Chain I/M hPolλ•DNA•(-)3TC-TP structure (Figures 5.6B and 5.12). Despite the similar (-)3TC-TP binding conformation, the base portion of (-)3TC-TP in the hPolβ structure is too far (3.6-4.1 Å) to form a hydrogen bond with R283 (Figure 5.6B), the residue analogous to the critical R517 in hPolλ. This increased distance and failure to form stable hydrogen bonds between (-)3TC-TP and R283 suggests that this Pathway I-like binding conformation of *L*-nucleotides may be unfavorable for hPolβ (Figures 5.6B and 5.12). To test this hypothesis, we performed an analogous pre-steady-state investigation of the R283A mutation (Table 5.1, Figure 5.4). We determined that the mutation resulted in a decreased incorporation rate and nucleotide binding affinity for *D*-dCTP and all three *L*-nucleotides (Table 5.1) and relaxed the *D*-stereoselectivity by 8.7- to 78-fold, relative to

wild-type. This is suggestive of an alternative binding mode between the R283A side chain and the *L*-nucleotides similar to that of chain I and M of hPolλ, and that removal of this interaction results in weakened binding affinities and selection for the *L*-nucleotides. However, the R517A mutation in hPolλ resulted in a similar 54-fold decrease in binding affinity for *L*-dCTP compared to wild-type hPolλ, while the binding affinity for natural *D*-dCTP was actually increased by 4-fold [223]. Therefore, these contradictory results would suggest that R283 in hPolβ has a different role compared to R517 in hPolλ. Thus, it is likely that the weakened binding affinity and *D*-stereoselectivity observed with the R283A hPolβ mutant are coincidental and arise from an inherent inability to stabilize the closed conformation as evident by the similar drop (50- to 94-fold) in binding affinity for *D*-dCTP, *L*-dCTP, and (-)3TC-TP (Table 5.1). In addition, the absence of binding conformations similar to hPolλ chain I and M in the hPolβ•DNA•(-)FTC-TP and hPolβ•DNA•*L*-dCTP structures suggests that hPolβ uses a mechanism distinct from that of hPolλ. Furthermore, as the major factors allowing hPolλ to select against *L*-stereochemistry are due to the differences between Pathways I and II (Figure 5.11), a unique and more efficient selection mechanism against *L*-nucleotides must be employed by hPolβ to account for the observed 22-, 7- and 24-fold higher selectivity against *L*-dCTP, (-)3TC-TP and (-)FTC-TP, respectively, when compared to hPolλ.[319]

5.5.2 Structural Basis for the *D*-stereoselectivity of hPolβ

Unlike hPolλ,[76] nucleotide incorporation by hPolβ requires a large conformational change that results in the N-helix of the thumb domain clasping down on the nascent base pair, effectively securing the reacting groups and optimally positioning them for catalysis.[21] It

is widely thought that such conformational changes or differences in nucleotide-bound complexes may act as fidelity checkpoints during single-nucleotide incorporation to preferentially select correct over incorrect nucleotides. However, the exact mechanism by which the conformational changes or differences in conformations of bound complexes affect Pol fidelity has long been debated.[4-6,27-29,79,103,213,324-327] Regardless of the mechanism, it is clear that the closure of the thumb domain is important for efficient nucleotide incorporation[14] and therefore, it was expected that the conformation of hPolβ would play a role in selecting against nucleotides with *L*-stereochemistry. Surprisingly, the hPolβ•DNA•*L*-dCTP structure is observed in the closed conformation and seemingly escapes any conformational selection checkpoints. However, several active site rearrangements occur compared to the hPolβ•DNA•*D*-dCTP structure (Figures 5.2, 5.8, and 5.13). Many of the atoms of the incoming *L*-dCTP have been shifted (0.3-1.8 Å, Figure 5.13) compared to the respective atoms in the incoming *D*-dCTP of the hPolβ•DNA•*D*-dCTP structure. In addition, the A- and B-site metal ions shift by 0.79 and 0.73 Å, respectively (Figure 5.13). Moreover, similar to the hPolλ•DNA•*L*-dCTP structure,[223] a water molecule is observed bridging the 3′-OH of *L*-dCTP (2.82 Å) and the primer 3′-OH (2.62 Å), which likely weakens the ability of the primer 3′-OH to act as a nucleophile during nucleotide incorporation (Figures 5.8B and 5.10B). The ribose rotation also positions the 3′-OH of *L*-dCTP near the hydrophobic side chain of F272, which results in the side chain rotating slightly from its position in the hPolβ•DNA•*D*-dCTP structure (Figure 5.9E) and likely further weakens the binding of *L*-dCTP. Together, these active site alterations establish the *D*-stereoselectivity of hPolβ and result

in an overall decreased nucleotide binding affinity and incorporation rate for *L*-dCTP incorporation (24 μM and 0.00063 s^{-1}, respectively, Table 5.1) compared to the incorporation of natural *D*-dCTP (0.71 μM and 5.02 s^{-1}, respectively, Table 5.1).

5.5.3 Conformational Change in the Drug Selectivity of hPolβ

While it was unexpected that the large conformational change exhibited by hPolβ would not contribute to the mechanism of *D*-stereoselectivty, the ability of *L*-dCTP to support domain closure can be attributed to a key hydrogen bond between the 3′-OH of *L*-dCTP and the backbone carbonyl of F272 (Figures 5.8B and 5.10B). This hydrogen bond stabilizes the thumb domain in the closed conformation and is similar to that seen in the hPolβ•DNA•*D*-dCTP structure, where the 3′-OH of *D*-dCTP forms a hydrogen bond with the backbone carbonyl of T273.[40] Notably, nucleotide binding is known to stabilize the rotation of the F272 side chain to its position in the ternary complex.[79] This rotation results in the disruption of the salt bridge formed between D192 and R258 in the binary complex and allows the side chain of D192 to rotate and bind the catalytic A-site metal ion during the formation of a productive ternary complex.[79]

In contrast, as the NRTIs, (-)3TC-TP and (-)FTC-TP, lack a 3′-OH and cannot form a hydrogen bond with the N-helix and would not stabilize the closed conformation to the same extent as *L*- and *D*-dCTP (Figure 5.10C and 5.10D). This is clearly evident by the inability of the thumb domain to fully close in the hPolβ•DNA•(-)3TC-TP and hPolβ•DNA•(-)FTC-TP structures (Figure 5.9D). Furthermore, the geometrical strain induced by the 180° rotation of the ribose results in non-planar ((-)FTC-TP) or absent ((-)3TC-TP) Watson-Crick base pairing between the NRTIs and the templating nucleotide.

Thus, it is likely that a combination of these two aspects prohibit full closure of the thumb domain. Notably, the distance between the α-phosphate of the incoming (-)3TC-TP and the primer 3′-OH (4.7 Å, Figures 5.8C and 5.10C) is too large for catalysis to occur and accounts for the reduced incorporation rate (0.0039 s^{-1}, Table 5.1) compared to *D*-dCTP (5.02 s^{-1}, Table 5.1). On the other hand, despite the inability to close fully, the distance between the α-phosphate of the incoming (-)FTC-TP and the primer 3′-OH (3.4 Å, Figures 5.8D and 5.10D) is very similar to the distances observed for both *L*-dCTP (3.7 Å, Figure 5.8B) and *D*-dCTP (3.8 Å, Figure 5.8A) and likely results in the higher observed incorporation rate (0.027 s^{-1}, Table 5.1) compared to (-)3TC-TP.

The inability of the NRTIs to support domain closure would seemingly make these nucleotides worse substrates but this is not the case considering the observed increased binding affinities and incorporation rates of (-)3TC-TP (0.18 μM and 0.0039 s^{-1}, respectively, Table 5.1) and (-)FTC-TP (11 μM and 0.027 s^{-1}, respectively, Table 5.1) compared to *L*-dCTP (24 μM and 0.00063 ± s^{-1}, respectively, Table 1). The substitution of the C3′ atom of *L*-dCTP with a more electron-rich sulfur in the NRTIs accounts for the increased binding affinity through an enhanced stacking interaction with the active site residues, F272 and Y271 (Figure 5.14). Furthermore, the lack of the steric clash between the NRTI 3′-OH and primer 3′-OH or Y271 may allow optimal alignment of the reacting groups to occur more readily following full domain closure (Figure 5.10). The differences in binding affinities of (-)3TC-TP (0.18 μM, Table 1) and (-)FTC-TP (11 μM, Table 5.1) are likely associated with the addition of the fluorine group to the cytosine base of (-)FTC-TP. As this fluorine group withdraws electrons from the base, the stacking

interaction between the base of (-)FTC-TP and the primer 3′-terminal base may be weakened and may disfavor the binding of (-)FTC-TP to some degree.

Overall, the structures presented here demonstrate that the mechanism of *D*-stereoselectivity by hPolβ consists of the accumulation of several minor active site rearrangements that ultimately disfavor the binding and incorporation of nucleotides that have *L*-stereochemistry rather than through alternative nucleotide and triphosphate binding modes seen with other polymerases.[87,223] Interestingly, hPolβ exhibits a protein conformational selection against NRTIs that have a sulfur substituted ribose. However, this substitution allows (-)3TC-TP and (-)FTC-TP to escape some of the active site alterations associated with the presence of the 3′-OH in *L*-dCTP, and therefore these NRTIs are more efficiently incorporated. Without an evident unifying mechanism for *D*-stereoselectivity, it remains necessary to investigate how replicative DNA polymerases or the target of the NRTIs, HIV RT, select against nucleotides with *L*-stereochemistry. Such an investigation would provide valuable insight into novel drug design and potentially limit the cytotoxicity associated with *L*-nucleotide drugs.

5.6 Tables

Nucleotide	k_p (s^{-1})	K_d (μM)	k_p/K_d ($\mu M^{-1}s^{-1}$)	Selectivity†	Ratio$
Catalyzed by wild-type hPolβ					
D-dCTP*	5.02 ± 0.07	0.71 ± 0.04	7.1	—	—
L-dCTP	0.00059 ± 0.00002	22 ± 2	2.7×10^{-5}	2.6×10^{5}	—
(-)3TC-TP*	0.0039 ± 0.0001	0.18 ± 0.02	2.2×10^{-2}	323	—
(-)FTC-TP*	0.027 ± 0.001	11 ± 2	2.5×10^{-3}	2.9×10^{3}	—
Catalyzed by the R283A mutant of hPolβ					
D-dCTP	0.39 ± 0.02	41 ± 6	9.5×10^{-3}	—	—
L-dCTP	0.00036 ± 0.00006	1110 ± 370	3.2×10^{-7}	3.0×10^{4}	8.7
(-)3TC-TP	0.0068 ± 0.0003	17 ± 2	4.0×10^{-4}	24	13
(-)FTC-TP	0.0063 ± 0.0002	24 ± 3	2.6×10^{-4}	37	78

An incoming nucleotide was incorporation opposite dG in the single-nucleotide gapped DNA substrate 21-19A-41GT-mer catalyzed by wild-type hPolβ and the R283A mutant of hPolβ at 37 °C.
*Reference 319.
†Selectivity = *D*-stereoselectivity = $(k_p/K_d)_{D\text{-dCTP}}/(k_p/K_d)_{L\text{-nucleotide}}$.
$Ratio is defined as (*D*-stereoselectivity of wild-type hPolβ)/(*D*-stereoselectivity of the R283A mutant of hPolβ).

Table 5.1: Pre-steady-state kinetic parameters for single nucleotide incorporation.

	Polβ•DNA•*L*-dCTP	Polβ•DNA•(-)3TC-TP	Polβ•DNA•(-)FTC-TP
Data collection*			
Space group	$P2_1$	$P2_1$	$P2_1$
Cell dimensions			
a, *b*, *c* (Å)	50.9, 80.1, 55.6	54.6, 79.4, 55.1	54.6, 79.3, 54.7
α, β, γ (°)	90.0, 107.7, 90.0	90.0, 106.6, 90.0	90.0, 107.6, 90.0
Resolution (Å)	30.90-1.80 (1.84-1.80)	52.85-2.20 (2.27-2.20)	50.00-1.80 (1.86-1.80)
R_{merge}†	0.112 (0.632)	0.160 (0.442)	0.069 (0.778)
$I/\sigma I$	8.9 (2.5)	6.2 (4.6)	19.7 (2.4)
Completeness (%)	99.8 (100.0)	99.9 (100.0)	99.3 (100.0)
Redundancy	5.4 (5.3)	3.6 (3.4)	5.6 (5.6)
Refinement			
Resolution (Å)	26.96-1.80	44.03-2.30	15.78–1.80
No. reflections	37460	19157	39217
R_{work}/R_{free}‡	0.192/0.229	0.233/0.284	0.202/0.247
No. atoms			
Protein	2652	2612	2614
DNA	632	632	632
Nucleotide	28	27	28
B-factors			
Protein	18.76	42.10	45.26
DNA	27.19	30.66	36.23
Nucleotide	12.04	75.21	51.02
R.m.s deviations			
Bond lengths (Å)	0.006	0.010	0.010
Bond angles (°)	1.148	1.476	1.458
PDBID	5U2R	5U2S	5U2T

*Highest resolution shell is shown in parenthesis.

†$R_{merge} = \Sigma|I-\langle I\rangle|/\Sigma I$, where I is the integrated intensity of each reflection.

‡R value = $\Sigma||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c are observed and calculated structure factor amplitudes, respectively.

Table 5.2: Structural Data collection and refinement statistics

***L*-analog complexes**	**RMSD** (Å) (relative to the binary structure 1BPX)	**RMSD** (Å) (relative to the ternary structure 4KLD)
hPolβ•DNA•*L*-dCTP (Ca^{2+})	1.76	0.76
hPolβ•DNA• (-)3TC-TP (Ca^{2+})	0.91	1.75
hPolβ•DNA• (-)FTC-TP (Ca^{2+})	1.07	1.95

Table 5.3: Root mean-square deviation (RMSD) of each Polβ structure with respect to the earlier published binary hPolβ•DNA (1BPX) and ternary hPolβ•DNA•*D*-dCTP structures (4KLD).

5.7 Figures

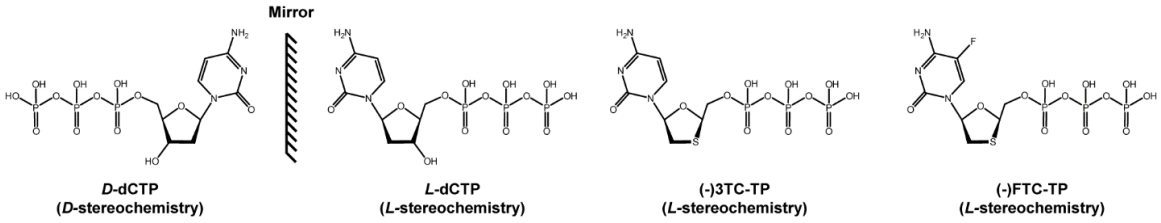


Figure 5.1: Chemical structures of *D*-dCTP, *L*-dCTP, (-)3TC-TP, and (-)FTC-TP. *D*-dCTP and *L*-dCTP are the enantiomers of dCTP. (-)3TC-TP and (-)FTC-TP are the biologically active triphosphate forms of lamivudine [(-)3TC; (-)-β-*L*-2′,3′-dideoxy-3′-thiacytidine] and emtricitabine [(-)FTC; (-)-β-*L*-2′,3′-dideoxy-5-fluoro-3′-thiacytidine], respectively, and analogs of *L*-dCTP.

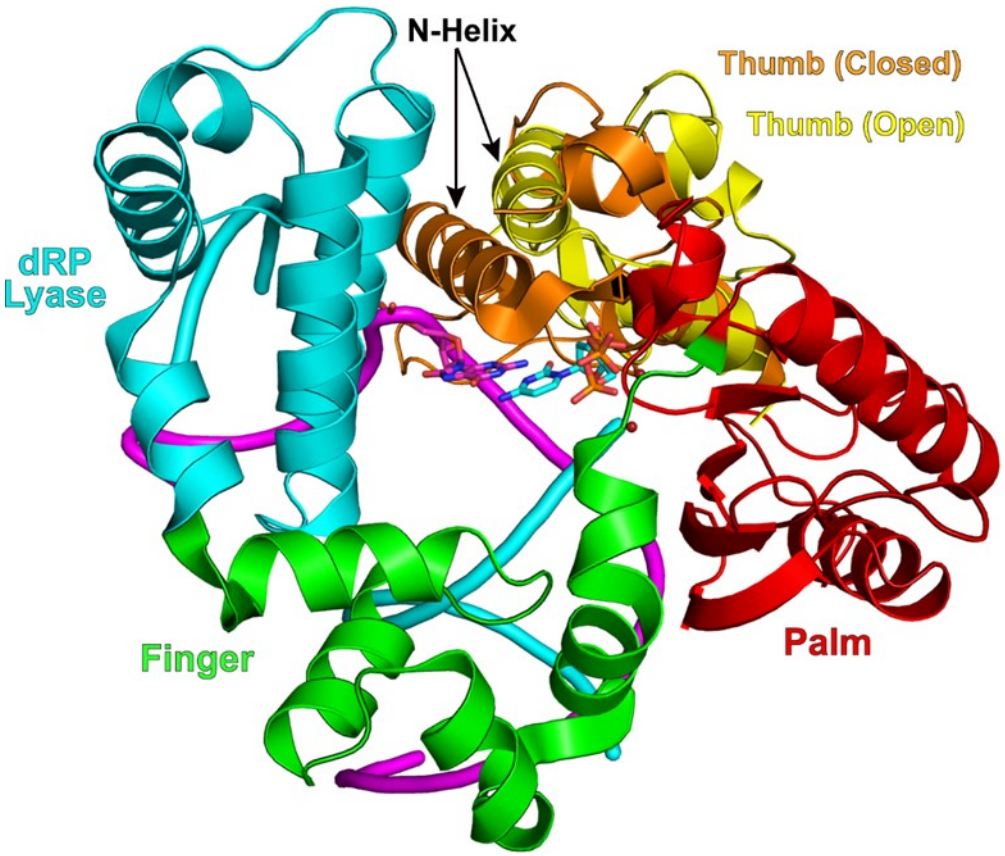


Figure 5.2: Domain architecture of hPolβ.

The crystal structure of hPolβ shown as cartoons. The domains are color coded with the label color and are termed the dRP lyase (cyan), finger (green), palm (red), and thumb (orange/yellow) domains from N- to C-terminus. The thumb domain is depicted in the open (yellow) and closed (orange) conformations with the N-Helix highlighted. The incoming nucleotide and template base are shown as sticks.

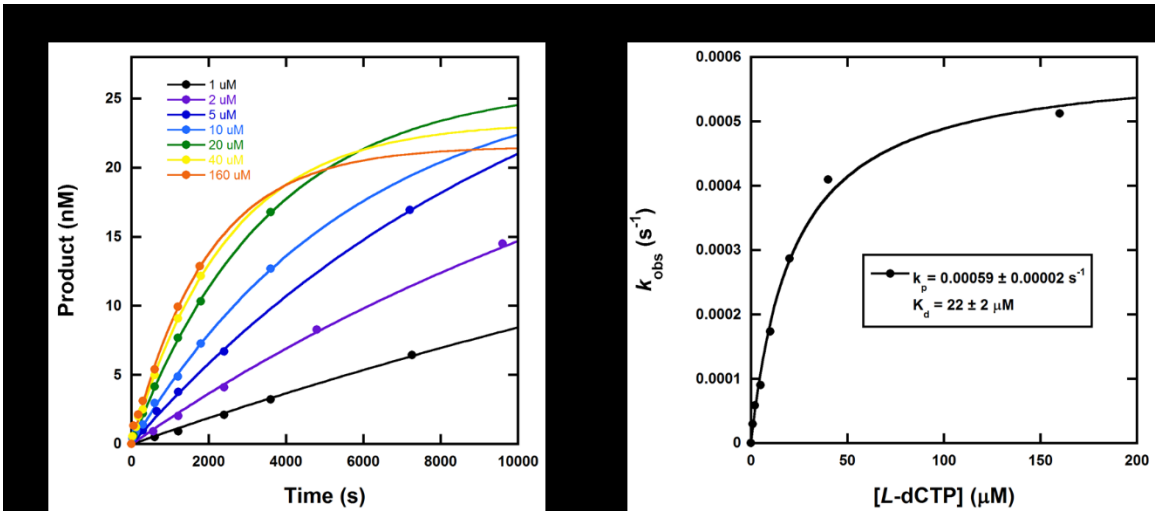


Figure 5.3: Incorporation of *L*-dCTP by wild-type hPolβ under single-turnover conditions.

(A) Briefly, a pre-incubated solution of full-length hPolβ (300 nM) and 30 nM [^{32}P]-labeled-21-19A-41GT-mer was mixed with varying concentrations of a nucleotide in buffer L (50 mM Tris-HCl, pH 8.4, 5 mM $MgCl_2$, 100 mM NaCl, 0.1 mM EDTA, 5 mM DTT, 10% glycerol, and 0.1 mg/ml BSA) at 37 °C. After various times, the reaction was stopped with 0.37 M EDTA and analyzed by sequencing gel electrophoresis. Each time course of product formation was fit to a single-exponential equation, [product] = A[1 - exp(- k_{obs}t)], to yield a reaction amplitude (A) and an observed rate constant of nucleotide incorporation (k_{obs}). (B) The k_{obs} values were then plotted against nucleotide concentration and the data were fit to a hyperbolic equation, $k_{obs} = k_p$[dNTP]/([dNTP] + K_d), to yield an apparent equilibrium dissociation constant (K_d, 22 μM) and a maximum nucleotide incorporation rate constant (k_p 0.00059 s^{-1}).

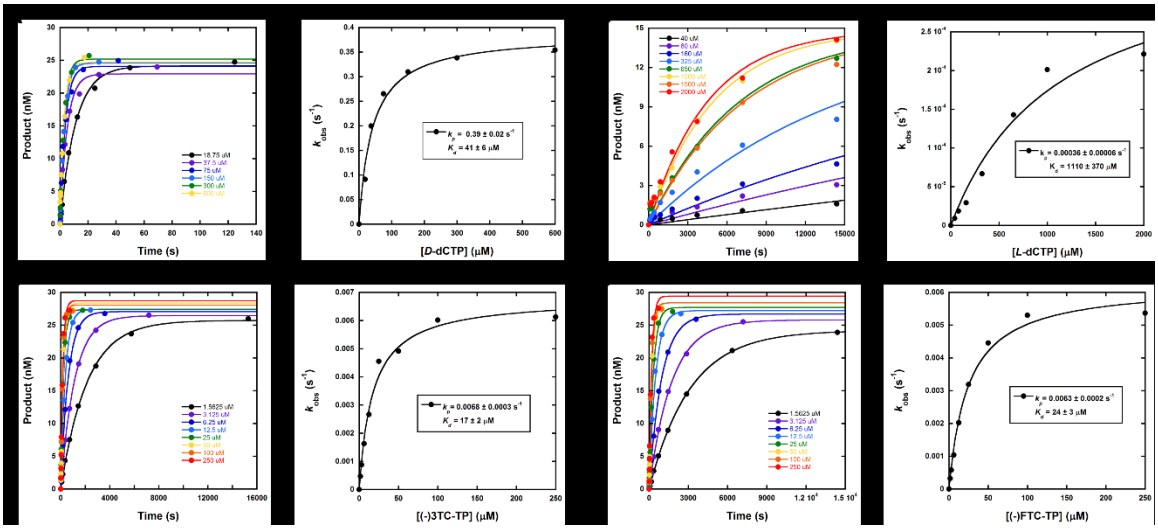


Figure 5.4: *D*- and *L*-nucleotide incorporation catalyzed by the R283A mutant of hPolβ.

Single-nucleotide incorporation assays were performed to measure the kinetic parameters k_p and K_d for incorporation of *D*-dCTP (A), *L*-dCTP (B), (-)3TC-TP (C), and (-)FTC-TP (D). These plots were generated as described in Figure S2 to yield respective kinetic parameters, k_p and K_d, of 0.39 s^{-1} and 41 μM for *D*-dCTP, 0.00036 s^{-1} and 1110 μM for *L*-dCTP, 0.0068 s^{-1} and 17 μM for (-)3TC-TP, and 0.0063 s^{-1} and 24 μM for (-)FTC-TP.

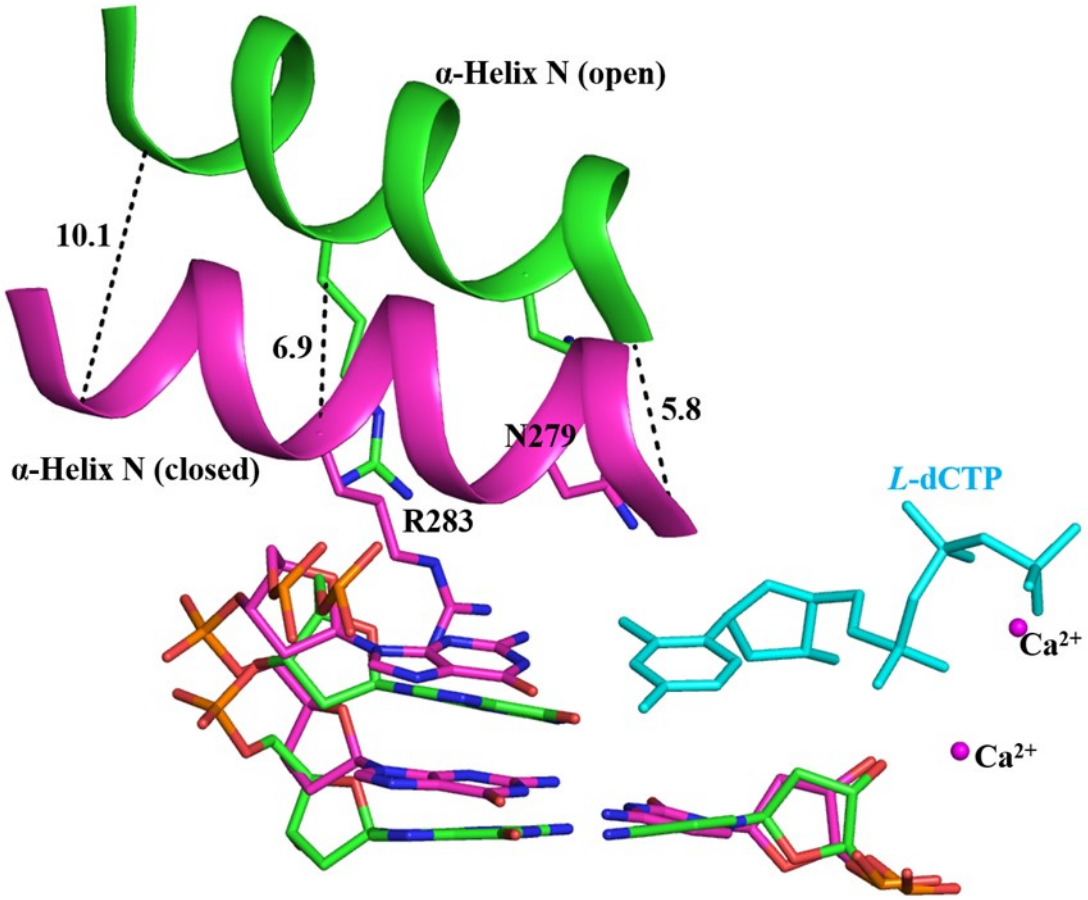


Figure 5.5: Comparison of the open and closed conformations of binary hPolβ•DNA (PDB 1BPX) and ternary hPolβ•DNA•*L*-dCTP structures.

Active site superposition of the binary hPolβ•DNA (green) and ternary hPolβ•DNA•L-dCTP (magenta) structures. Closing of α-Helix N results in the movement of active site residues and the nucleotides in the DNA substrate to the proper catalytic positions in the ternary complex.

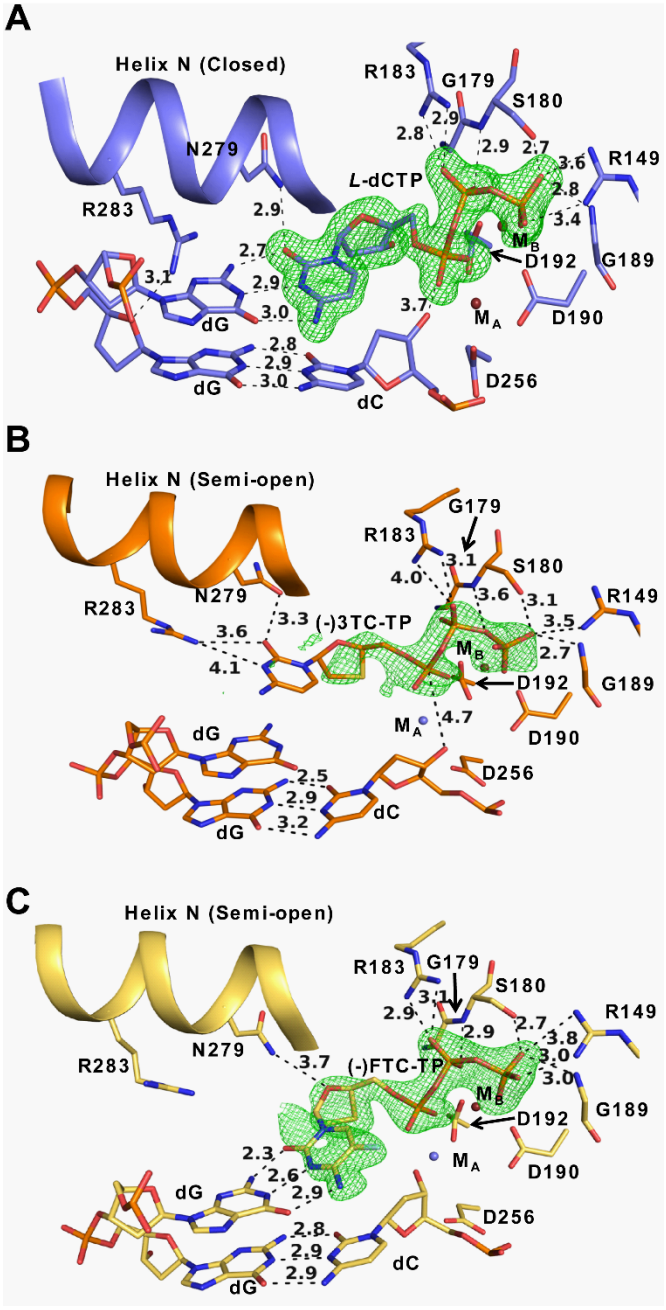


Figure 5.6: Binding of *L*-nucleotides in the presence of Ca^{2+} within the active site of hPolβ.

(Continued)

Figure 5.6: Continued

(A) hPolβ•DNA•*L*-dCTP; (B) hPolβ•DNA•(-)3TC-TP ; and (C) hPolβ•DNA•(-)FTC-TP. The F_o - F_c omit maps (3σ level) for the incoming *L*-nucleotides are illustrated in green. The two template nucleotides, the primer 3'-terminal nucleotide, and active site residues are displayed as sticks. Hydrogen bonds and the distance between the primer 3′-OH group and the α-phosphorus atom of an incoming *L*-nucleotide are displayed as black dashed lines with the numbers representing the distance in Å. The Ca^{2+} and Na^{+} ions are shown as red and light blue spheres, respectively.

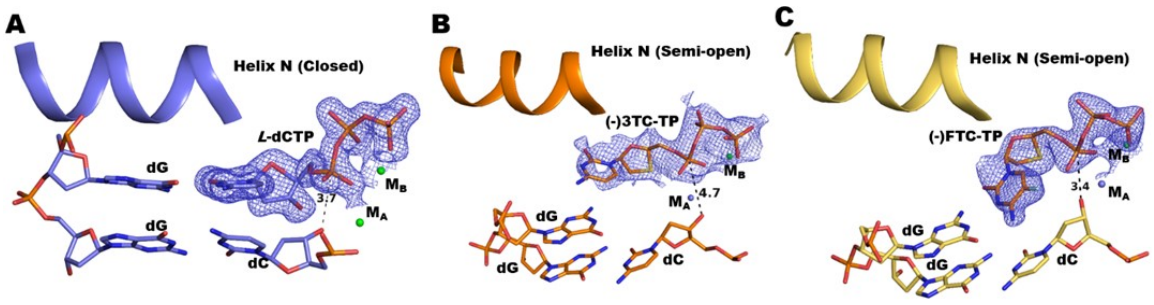


Figure 5.7: Binding of *L*-nucleotides within the active site of hPolβ.

(A) hPolβ•DNA•*L*-dCTP; (B) hPolβ•DNA•(-)3TC-TP; and (C) hPolβ•DNA•(-)FTC-TP. The $2F_o$ - F_c omit maps (1σ level) for the incoming *L*-nucleotides are illustrated in blue. The two template nucleotides, the primer 3'-terminal nucleotide, and active site residues are displayed as sticks. Hydrogen bonds and the distance between the primer 3′-OH group and the α-phosphorus atom of an incoming *L*-nucleotide are displayed as black dashed lines with the numbers representing the distance in Å. The Ca^{2+} and Na^{+} ions are shown as green and light blue spheres, respectively.

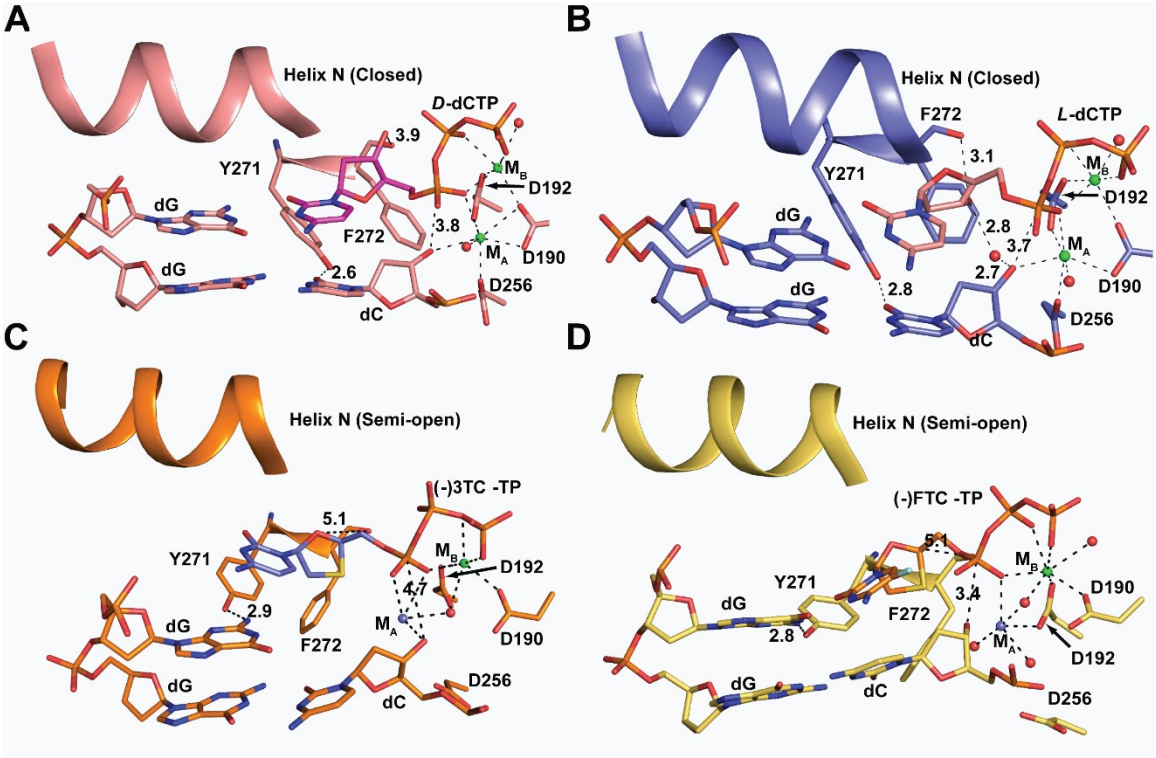


Figure 5.8: Metal coordination in pre-catalytic ternary structures of hPolβ.

(A) hPolβ•DNA•*D*-dCTP (4KLD); (B) hPolβ•DNA•*L*-dCTP; (C) hPolβ•DNA•(-)3TC-TP; and (D) hPolβ•DNA•(-)FTC-TP. Hydrogen bonds between coordinating ligands to the metal ions and the distance between the primer 3′-OH group and the α-phosphate atom of the incoming nucleotide are displayed as black dashed lines with the numbers indicating the distance in Å. Water molecules (red), Ca^{2+} ions (green), and Na^{+} ions (light blue), are shown as spheres.

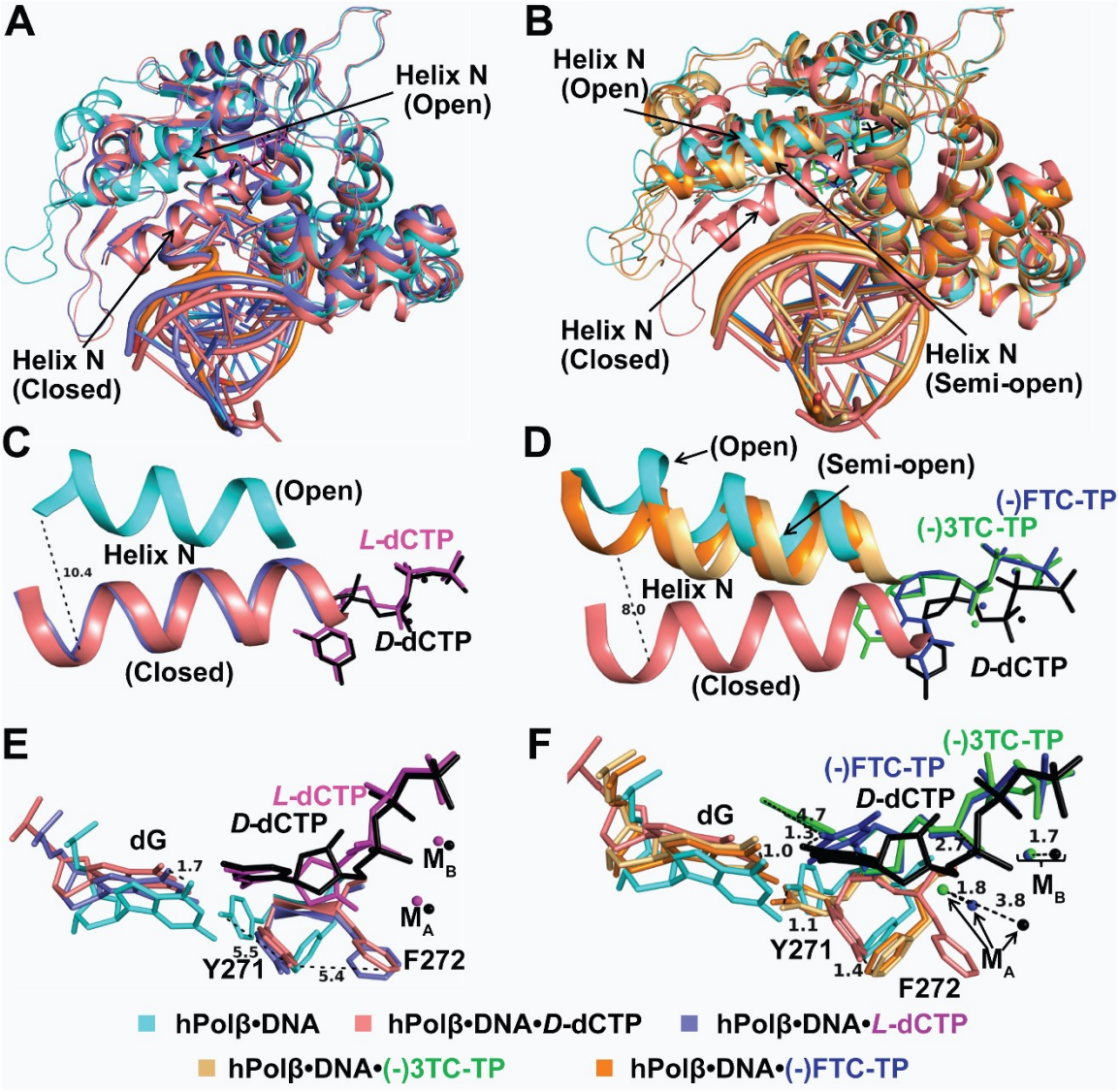


Figure 5.9: Structural differences between the binary and ternary structures complexed with either a *D*- or an *L*-nucleotide in the presence of Ca^{2+}.

(A) Superposition of hPolβ•DNA (1BPX), hPolβ•DNA•*D*-dCTP (4KLD), and hPolβ•DNA•*L*-dCTP structures. (B) Superposition of hPolβ•DNA (1BPX),

(Continued)

Figure 5.9: Continued

hPolβ•DNA•*D*-dCTP (4KLD), hPolβ•DNA•(-)3TC-TP, and hPolβ•DNA•(-)FTC-TP structures. (C) and (D) Zoomed views of Helix-N conformation adopted in (A) and (B), respectively. (E) and (F) Active site differences of structures superimposed in (A) and (B). In (E) and (F) only the templating nucleotide, incoming nucleotide, Y271, F272 and the A-site (M_A) and B-site (M_B) metal ions are presented. In (E) both metal sites are occupied by Ca^{2+} ions. In (F) the M_B site is occupied by Ca^{2+} for all three structures, while the M_A site is occupied by Na^{+} ions for the (-)3TC-TP and (-)FTC-TP structures but is a Ca^{2+} ion in the *D*-dCTP structure.

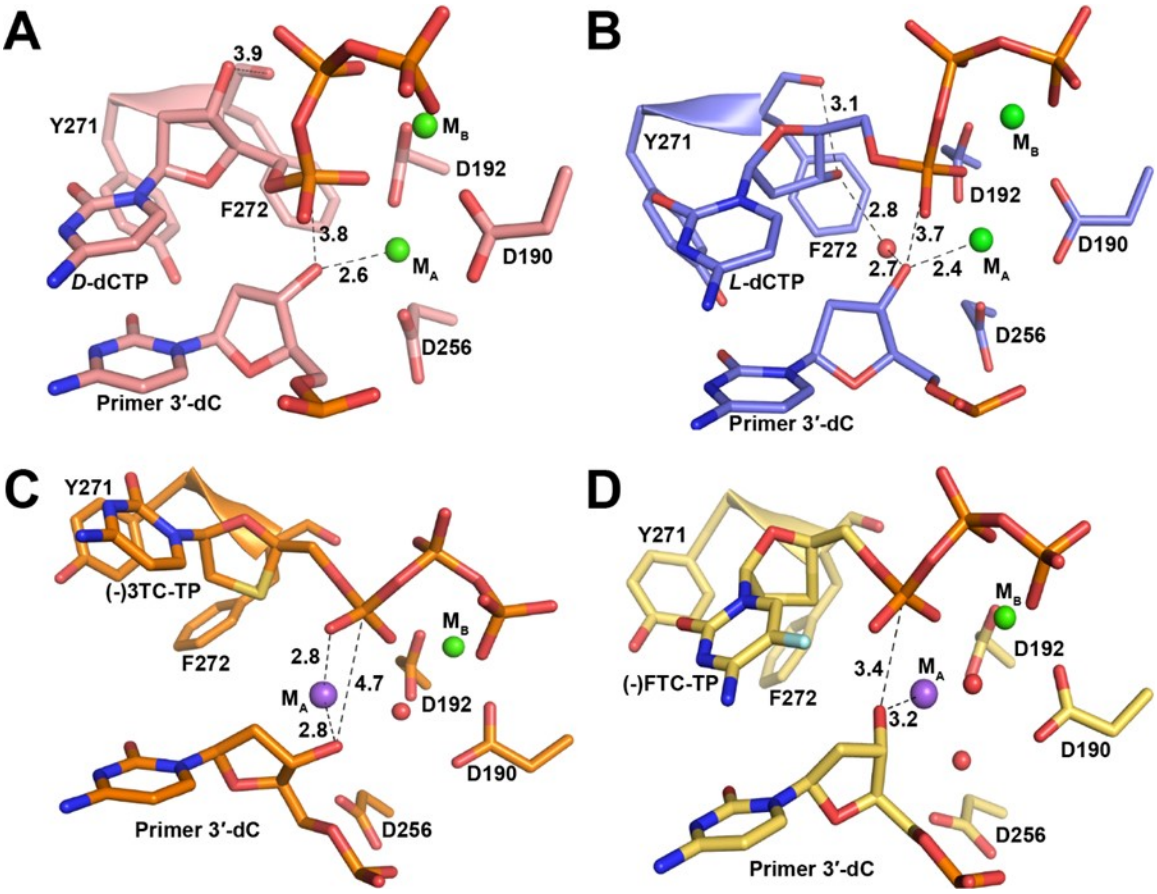


Figure 5.10: Relative orientation of the primer 3′-OH and α-phosphate of the incoming nucleotide.

(A) hPolβ•DNA•*D*-dCTP (4KLD); (B) hPolβ•DNA•*L*-dCTP; (C) hPolβ•DNA•(-)3TC-TP; and (D) hPolβ•DNA•(-)FTC-TP. Hydrogen bonds between the primer 3′-OH group and the A-site metal ion and the distance between the α-phosphate atom of the incoming nucleotide and the primer 3′-OH are displayed as black dashed lines with the numbers indicating the distance in Å. Water molecules (red), Ca^{2+} ions (green), and Na^{+} ions (purple), are shown as spheres.

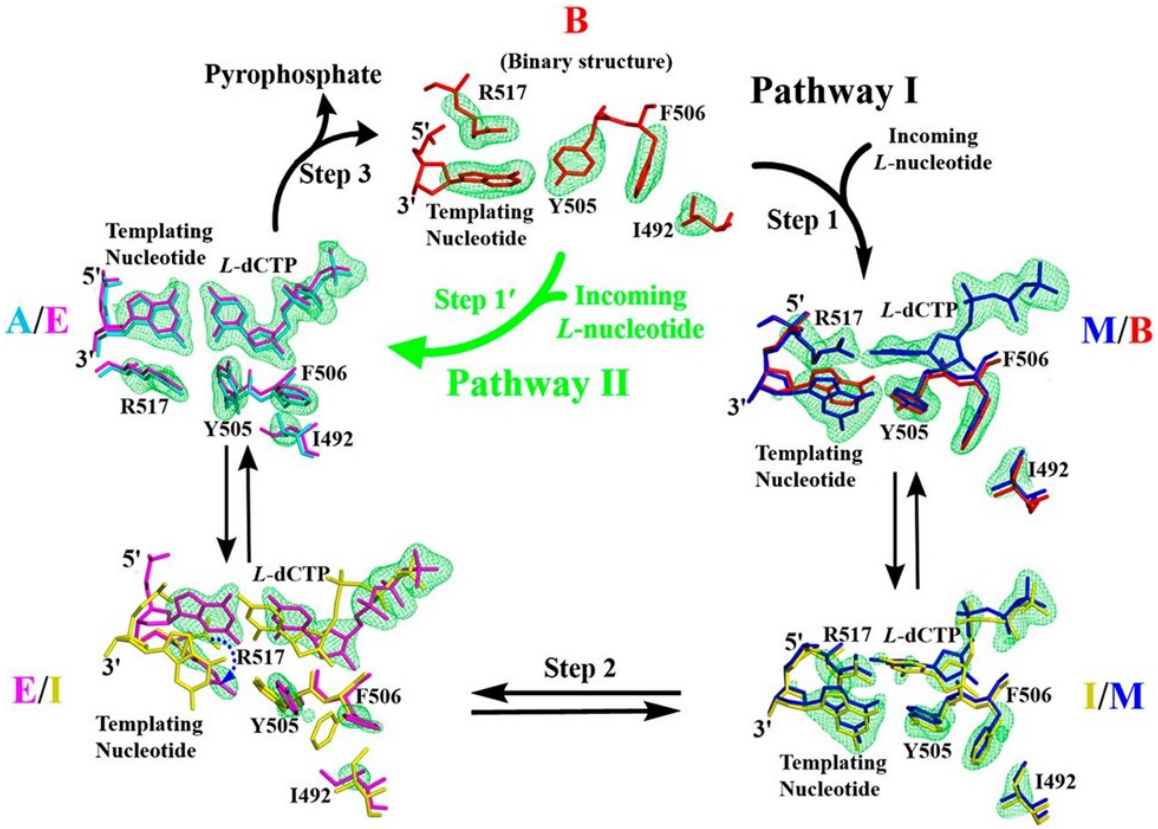


Figure 5.11: Proposed pathways for *L*-nucleotide incorporation catalyzed by human DNA polymerase λ.

Mechanism of D-stereoselectivity for human DNA polymerase λ proposed by Vyas et al.[223]

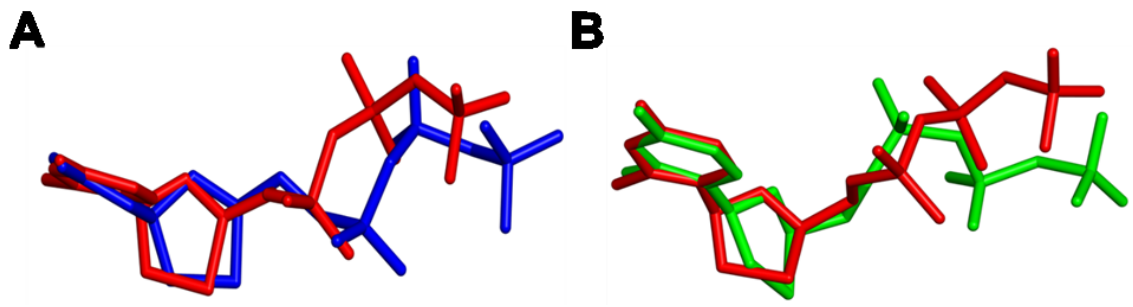


Figure 5.12: Different binding conformations of (-)3TC-TP in the X-family DNA polymerase structures.

(A) Superposition of the chair-like binding conformations of (-)3TC-TP in the hPolβ•DNA•(-)3TC-TP (blue) and Chain A hPolλ•DNA•(-)3TC-TP (red) structures.[223] (B) Comparison of the *N*-shaped binding conformation of (-)3TC-TP in Chain M hPolλ•DNA•(-)3TC-TP structure (green) and the chair-like binding conformation in Chain A hPolλ•DNA•(-)3TC-TP structure (red).[223] Interestingly, both the chair-like (red) and *N*-shaped (green) binding conformations of (-)3TC-TP in the hPolλ•DNA•(-)3TC-TP structures were observed in the presence of Ca^{2+}. Notably only the chair-like triphosphate conformation was observed in the hPolβ structures with *D*-dCTP, *L*-dCTP, and (-)FTC-TP and compare similarly to the Chain A structures of hPolλ.

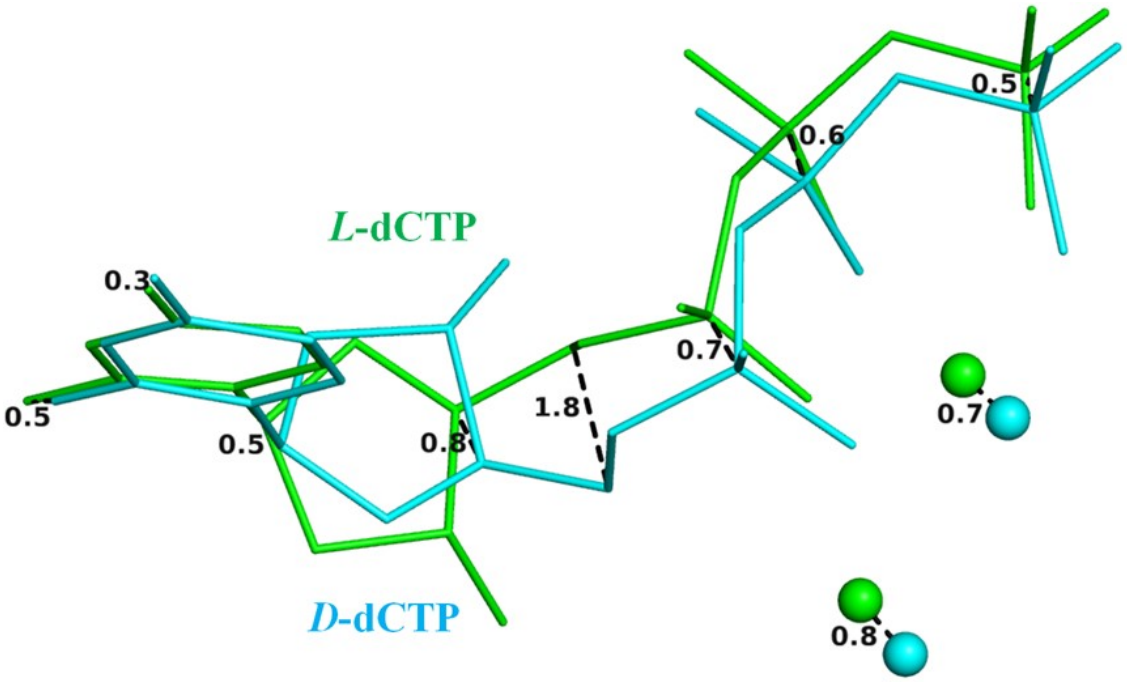


Figure 5.13: Comparison of the binding conformations of *D*-dCTP and *L*-dCTP in their ternary structures.

The hPolβ•DNA•*D*-dCTP[40] and hPolβ•DNA•*L*-dCTP structures were superimposed and the distances (Å) between similar atoms (dashed lines) were measured. *L*-dCTP has shifted considerably compared to *D*-dCTP with the maximal change (1.8 Å) between the C5′ atoms. In addition, both the A- and B-site metal ions change their positions considerably.

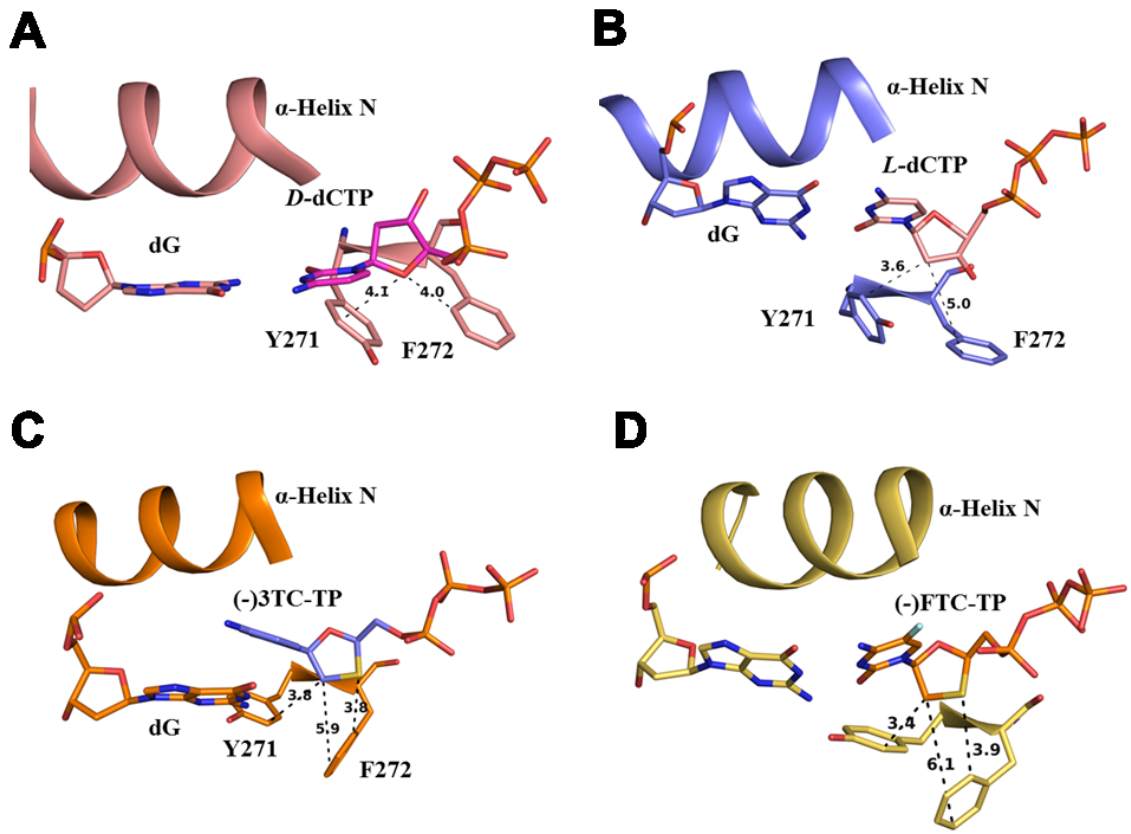


Figure 5.14: Interaction pattern of incoming *L*-nucleotides with active site residues Y271 and F272.

The stacking interactions between the active site residues Y271 and F272 of pre-catalytic ternary hPolβ structures and the ribose of *D*-dCTP[40] (A), *L*-dCTP (B), (-)3TC-TP (C), and (-)FTC-TP (D). The nucleotides and active site residues are presented as sticks. The distances between the atoms of the incoming nucleotide and the atoms of Y271 and F272 are presented as black dashed lines with the numbers representing the distance in Å.

Chapter 6: Structural Insights into the Post-Chemistry Steps of Nucleotide Incorporation Catalyzed by a DNA Polymerase

Reproduced in part with permission from Reed, A.J., Vyas, R., Raper, A.T., and Suo, Z. (2017) Structural Insights into the Post-Chemistry Steps of Nucleotide Incorporation Catalyzed by a DNA Polymerase. *J. Am. Chem. Soc.* **139**, 465-471. DOI: 10.1021/jacs.6b11258.

Author Contributions

Rajan Vyas and Andrew J. Reed contributed equally to this manuscript. R.V. crystallized protein complexes and solved the structures. A.J.R. interpreted the structures, developed the mechanistic model, and wrote/edited the manuscript. Austin T. Raper helped interpret data, develop the mechanistic model, and edit the manuscript. Zucai Suo conceived the research, helped develop the mechanistic model, and helped edit the manuscript.

6.1 Abstract

DNA polymerases are essential enzymes that faithfully and efficiently replicate genomic information.[5,8,109] The mechanism of nucleotide incorporation by DNA polymerases has been extensively studied structurally and kinetically, but several key steps following phosphodiester bond formation remain structurally uncharacterized due to utilization of natural nucleotides. It is thought that the release of pyrophosphate (PP_i) triggers reverse conformational changes in a polymerase in order to complete a full catalytic cycle as well as prepare for DNA translocation and subsequent incorporation events. Here, by using the triphosphates of chain-terminating antiviral drugs lamivudine ((-)3TC-TP) and emtricitabine ((-)FTC-TP), we structurally reveal the correct sequence of post-chemistry steps during nucleotide incorporation by human DNA polymerase β (hPolβ) and provide a structural basis for PP_i release. These post-catalytic structures reveal hPolβ in an open conformation with PP_i bound in the active site, thereby strongly suggesting that the reverse conformational changes occur prior to PP_i release. The results also help to refine the role of the newly discovered third divalent metal ion for DNA polymerase-catalyzed nucleotide incorporation. Furthermore, a post-chemistry structure of hPolβ in the open conformation, following incorporation of (-)3TC-MP, with a second (-)3TC-TP molecule bound to the active site in the absence of PP_i, suggests that nucleotide binding stimulates PP_i dissociation and occurs before polymerase translocation. Our structural characterization defines the order of the elusive post-chemistry steps in the canonical mechanism of a DNA polymerase.

6.2 Introduction

DNA polymerases (Pols) are intricate cellular machines capable of synthesizing genomic DNA in an efficient and faithful manner. In general, Pols from all six families contain three canonical polymerase domains, the fingers, thumb, and palm, arranged in a right hand configuration.[8,23] In addition to structural similarities, Pols catalyze DNA synthesis through a largely conserved mechanism of nucleotide incorporation following DNA binding, wherein a nucleotide is selected, bound, and subsequently incorporated into DNA (Scheme 6.1).[5,8,109] Accordingly, the products of nucleotide incorporation by Pols include an extended DNA primer and pyrophosphate (PP_i). Global conformational changes upon nucleotide binding, such as an open→closed transition of the fingers domain for most Pols, or the thumb domain for human DNA polymerase beta (hPolβ), and a closed→open transition following nucleotide incorporation have been shown to act as fidelity checkpoints for Pols, influencing both nucleotide selectivity and catalytic efficiency.[4,328] Beyond these large conformational changes, more subtle motions that take place within individual Pol domains are suggested to be important for DNA and nucleotide binding, as well as nucleotide incorporation.[328]

While a great deal of biochemical, biophysical, and structural work has focused on characterizing the steps up to and including phosphodiester bond formation,[5,8] little structural work has investigated the steps occurring after catalysis, such as PP_i release, reverse conformational changes, and either DNA translocation or dissociation (Scheme 6.1). Although these steps are key determinants in Pol processivity and likely control the balance between the competing polymerization and editing (i.e. exonuclease) activities of

Pols,[4,5,180] characterization of these events is problematic as these steps are rapid during natural nucleotide incorporation.[90] Accordingly, it remains unclear if PP_i release occurs first, to effectively trigger the reverse conformational changes of the Pol, or the reverse conformational changes occur first, allowing PP_i to simply diffuse out of the active site. Furthermore, it is unclear what roles various active site residues, metal ions, and an incoming nucleotide serve in these post-chemistry steps. Here, we utilized the triphosphate derivatives of lamivudine ((-)3TC-TP) and emtricitabine ((-)FTC-TP) to structurally characterize the elusive post-chemistry steps of DNA polymerase-catalyzed nucleotide incorporation. Despite their *L*-stereochemistry, sulfur-substituted ribose, and lack of a 3′-OH (Figure 6.1), these unnatural nucleotides form natural Watson-Crick base pairs with the templating base, adopt productive triphosphate conformations, and can be utilized by hPolβ, albeit with altered incorporation kinetics.[319] Collectively, the unique properties offered by these nucleotide analogs, including their near natural configurations and interactions within the hPolβ active site, as well as their ability to prevent subsequent primer extension following slow incorporation, permitted us to capture the first glimpse of mechanistic steps following phosphodiester bond formation.

6.3 Materials and Methods

6.3.1 Preparation of Protein and DNA

Full-length hPolβ was overexpressed and purified as described previously.[42] DNA oligomers were purchased from Integrated DNA Technologies and were purified by denaturing polyacrylamide gel electrophoresis. Crystallization oligomers consisted of a 16-mer template (5′-CCGACGGCGCATCAGC-3′), a 10-mer upstream primer (5′-

GCTGATGCGC-3′), and a 5-mer downstream 5′-phosphorylated primer (5′-pGTCGG-3′) as described previously.[40] For crystallization, the template, upstream primer, and downstream primer were mixed in a 1:1:1 ratio and annealed by heating to 95 °C for 5 minutes followed by slowly cooling to 4 °C to form the single-nucleotide gapped DNA substrate (1 mM). The *L*-nucleotides, (-)FTC-TP and (-)3TC-TP, were obtained from Jena Bioscience. If not specifically mentioned, all DNA substrates are single-nucleotide gapped. DNA_{nicked}, DNA–(-)3TC-MP, and DNA–(-)FTC-MP are nicked DNA substrates.

6.3.2 hPolβ Crystallization and Structure Determination

Purified human hPolβ was dialyzed into a buffer containing 50 mM NaOAc (pH 5.5), 75 mM NaCl, 5% glycerol and 0.1 mM DTT, and concentrated to 20 mg/ml as described previously.[78] Binary complexes (hPolβ•DNA) were prepared by mixing purified hPolβ and the single-nucleotide gapped DNA substrate (1 mM) in a 1:1 ratio at 4 °C followed by heating to 35 °C and subsequent cooling to room temperature.[32] Binary crystals were obtained by hanging drop vapor diffusion against a reservoir solution containing 50 mM imidazole, pH 7.5, 16%–18% PEG3350, and 350 mM sodium acetate.[32] Binary crystals were then seeded into freshly prepared ternary complex solution (hPolβ•DNA, 10 mM $CaCl_2$ and 2 mM (-)3TC-TP or (-)FTC-TP) to obtain pre-catalytic ternary crystals (hPolβ•DNA•(-)3TC-TP or hPolβ•DNA•(-)FTC-TP) at room temperature, as described previously.[79] The resulting pre-catalytic ternary crystals were then soaked in a cryosolution (15% ethylene glycol, 50 mM imidazole (pH 7.5), 20% PEG3350, 90 mM NaOAc) containing either Mg^{2+} or Mn^{2+} ions (50 mM) for 5 hrs to achieve thumb domain closure or overnight to allow nucleotide incorporation. Crystals

were flash frozen in liquid nitrogen and X-ray diffraction data were collected using the LRL-CAT beamline facilities at Advance Photon Source (APS), Argonne National Laboratory. The X-ray diffraction data were processed using MOSFLM.[282] Structures were solved by molecular replacement using PHASER.[283] A previous structure (PDB code: 4KLD),[40] devoid of ligands and solvent molecules, was used as the initial model. Structural refinement was carried out using REFMAC5.[285] COOT was used for visualization and model building.[286] Quality of the models was assessed using PROCHECK.[287] Figures were created using PYMOL.[288]

6.4 Results and Discussion

Previous structural studies have been unable to characterize the post-chemistry steps with natural nucleotides due to rapid reverse protein conformational changes and PP_i release.[5,8,38-42] To gain insight into the post-chemistry processes, we crystallized ternary complexes of hPolβ, DNA, and (-)3TC-TP (hPolβ•DNA•(-)3TC-TP) or (-)FTC-TP (hPolβ•DNA•(-)FTC-TP), the triphosphates of two highly successful antiviral drugs, lamivudine and emtricitabine, respectively, in the presence of Ca^{2+} followed by *in crystallo* Mn^{2+} exchange for five hours. Unexpectedly, there was no observed nucleotide incorporation even after complete Ca^{2+} to Mn^{2+} exchange, likely due to slow *L*-nucleotide incorporation catalyzed by hPolβ.[319] These pre-catalytic ternary crystals yielded structures at resolutions of 2.49 and 2.00 Å (Table 6.1) for hPolβ•DNA•(-)FTC-TP (Figure 6.2A) and hPolβ•DNA•(-)3TC-TP (Figure 6.2D), respectively, with similar overall tertiary structures (r.m.s.d. of 0.78 and 0.72 Å, respectively) compared to the ternary structure of hPolβ, DNA, and natural dCTP (PDB code 4KLD)[40] (Figure 6.3),

including a fully closed thumb domain. In addition, the pre-catalytic ternary structures have well-defined electron density for the incoming nucleotide and exhibit similar active site architecture compared to hPolβ•DNA•dCTP, except for a slight repositioning of A-site (M_A) and B-site (M_B) metal ions (Figure 6.3A). The hPolβ•DNA•(-)3TC-TP and hPolβ•DNA•(-)FTC-TP structures with Mn^{2+} likely represent E′′•DNA_n•dNTP in Scheme 6.1.

Crystals of the pre-catalytic ternary complexes were then soaked in a Mn^{2+} or Mg^{2+} solution overnight to obtain post-catalytic structures. The crystals of hPolβ•DNA–(-)FTC-MP•PP_i in the presence of Mg^{2+} (Figure 6.2B) and hPolβ•DNA–(-)3TC-MP•PP_i in the presence of Mn^{2+} (Figure 6.2E) diffracted to a resolution of 2.39 and 2.49 Å, respectively (Table 6.1). The active sites of both structures clearly indicate the formation of a covalent linkage between the terminal primer nucleotide, dC, and the monophosphates of lamivudine and emtricitabine, (-)FTC-MP or (-)3TC-MP (Figure 6.2B and 6.2E). Notably, in the hPolβ•DNA–(-)3TC-MP•PP_i structure (Figure 6.2E), the catalytic A-site metal ion is absent and the nucleotide binding B-site metal ion is modeled as Na^+, while they are modeled as Na^+ and Mn^{2+}, respectively, in the hPolβ•DNA–(-)FTC-MP•PP_i (Figure 6.2B). Further comparison between the active sites of hPolβ•DNA–(-)3TC-MP•PP_i and hPolβ•DNA–(-)FTC-MP•PP_i, reveals that the B-site metal ions are in nearly identical positions and are associated with the newly generated PP_i (Figures 6.2B, 6.2E, and 6.4).

The overall tertiary structures of these post-catalytic ternary complexes are similar, with the N-helix adopting an open protein conformation, and overlay well with

the binary structures, hPolβ•DNA and hPolβ•DNA_{nicked} (PDB codes 1BPX and 1BPZ, respectively),[79] except the base portion of (-)3TC-TP, which forms a non-planar Watson-Crick base pair with the templating nucleotide (Figure 6.4). Surprisingly, the product PP_i is also bound in the active site (Figures 6.2B, 6.2E, 6.5B, and 6.5E) unlike previous post-catalytic structures wherein PP_i had been released.[40-42] Moreover, PP_i forms hydrogen bonds with several palm domain residues that are consistent with those seen in the closed, pre-catalytic ternary complexes (Figure 6.5). Previously, we have shown that following nucleotide incorporation opposite 8-oxoguanine (8-oxoG), hPolβ is in the closed conformation with PP_i bound, but following extended incubation periods, hPolβ is observed in the open conformation with PP_i released.[42] Interestingly, a similar time-dependent structural study observing dCTP incorporation by hPolβ on undamaged DNA could not capture PP_i release or the reverse conformational change even 11 hours after reaction initiation.[40] Furthermore, post-catalytic structures with T7 RNA polymerase[329] were unable to define the sequence of PP_i release and the closed→open conformational change. Clearly, the timing and the rapid kinetics of these post-chemistry processes are ambiguous and have likely contributed to poor structural characterization to date.

Here, the hPolβ•DNA–(-)3TC-MP•PP_i and the hPolβ•DNA–(-)FTC-MP•PP_i structures are observed in an open conformation with PP_i bound (Figures 6.2B, 6.2E, 6.5B, and 6.5E), which is surprising given that computational analyses of DNA polymerase I from *Bacillus stearothermophilus* (BF) have shown that PP_i release triggers post-catalytic conformational changes.[330] These open post-catalytic structures suggest a mechanism where PP_i can stay tightly associated with the polymerase active site

following the reverse conformational change of the N-helix in the absence of the next nucleotide. Furthermore, this mechanism of PP_i release following a closed→open conformational change is supported by the persistent hydrogen bonding network between the β- and γ-phosphates of the nucleotides or the product PP_i and residues in the palm domain (Figure 6.5), which may also impede binding of the next nucleotide.

While examination of the DNA polymerase mechanism with natural substrates is preferred, the rapid kinetics associated with certain steps render structural characterization difficult or impossible. Fortuitously, our use of the chain-terminating nucleotide analogs has permitted the capture of the elusive post-chemistry events to further clarify the catalytic mechanism of hPolβ (Figures 6.2, 6.5, and 6.6, Scheme 6.1). Consistently, a kinetic study of human DNA polymerase γ (hPolγ) incorporating the triphosphate of another chain-terminating nucleotide analog, 3'-azido-2',3'-dideoxythymidine (AZT-TP), demonstrated that PP_i release following AZT-TP incorporation is slow,[235] suggesting the existence of a stable hPolγ•DNA–AZT-MP•PP_i complex.

In terms of the overall kinetic mechanism (Scheme 6.1), the above closed and open post catalytic structures help to define Steps 8 and 9, respectively. The pre-catalytic and chemistry steps (Scheme 6.1, Steps 1-6) of nucleotide incorporation by hPolβ have been previously investigated structurally and kinetically, and we will not discuss them here.[97,296,331] However, we point out that Step 4 in Scheme 6.1 represents the forward open→closed conformational transition of the thumb domain for hPolβ, which already happened in the crystals prior to the metal ion exchange experiments described in this

study. Steps 5-7 of Scheme 6.1 occurred following metal ion exchange but were not observed *in crystallo*. Notably, Step 7 of Scheme 6.1 represents the reverse of the conformational changes (*i.e.* E″→E′) in the active site of hPolβ that occurred during Step 5,[296,332] but our structures here do not inform on these events. Following nucleotide binding and incorporation, our post-catalytic ternary structures of hPolβ (Figures 6.2B, 6.2E, 6.5B, and 6.5E) strongly suggest that the thumb domain must first complete the reverse, closed→open transition (Scheme 6.1,Step 8) before PP_i can be released (Scheme 6.1, Step 9). However, the sequence of the post-chemistry steps following a misincorporation event may be different than described here. For example, PP_i release following a misincorporation is purported to be fast, as shown in a time-resolved crystallographic study of hPolβ wherein only open-product complexes without PP_i, rather than closed-product complexes with bound PP_i, could be isolated following a misincorporation event.[40] This is consistent with kinetic studies of T7 DNA polymerase[90] and φ29 DNA polymerase,[333] wherein a mismatched primer terminus could not be removed via pyrophosphorolysis due to fast PP_i dissociation. In addition, the sequence of post-chemistry steps for natural nucleotide incorporation may or may not follow the sequence described here. However, structural characterization of such events may be unattainable due to the fast kinetics of correct, natural nucleotide incorporation. For example, our previous time-dependent studies viewing dCTP incorporation opposite 8-oxoG during hPolβ catalyzed translesion synthesis were unable to elucidate the order of events following chemistry.[42] To evaluate whether or not incorporation of a natural

nucleotide or misincorporation follows the sequence of post-chemistry steps in Scheme 1, similar structural studies as presented here are required.

Clearly, the molecular determinants for the reverse conformational changes and pyrophosphate release have not been well understood structurally but our results illuminate key details of these events for correct nucleotide incorporation. As proposed previously, it is possible that rather than passive PP_i release and DNA translocation (Scheme 6.1, Steps 9 and 10_a, respectively), the presence of the next correct nucleotide may actively displace PP_i during processive DNA synthesis. Therefore, during a subsequent nucleotide incorporation, PP_i dissociation may occur as late as Step 3 in Scheme 6.1.[148,330,334] To investigate the role of nucleotide binding in PP_i release, crystals of hPolβ•DNA–(-)3TC-TP with Ca^{2+} were soaked in a cryosolution containing Mg^{2+} and (-)3TC-TP. Crystals of hPolβ•DNA–(-)3TC-MP•(-)3TC-TP diffracted to 1.85 Å resolution and showed the binding of the second (-)3TC-TP following the incorporation of the first (-)3TC-TP, but the notable absence of PP_i generated during nucleotide incorporation (Figure 6.7). Surprisingly, (-)3TC-TP in the hPolβ•DNA–(-)3TC-MP•(-)3TC-TP structure forms a similar hydrogen bonding network with several residues (S180, R183, and G189) compared to PP_i in the hPolβ•DNA–(-)3TC-MP•PP_i structure (Figures 6.6B and 6.7B). In addition, (-)3TC-TP forms two hydrogen bonds with the side chain of R149 that further facilitate nucleotide binding and may aid in PP_i displacement (Figures 6.6B and 6.7B). This is consistent with several stopped-flow studies which demonstrated that in the presence of the next correct nucleotide, PPi release is fast, while in the absence of additional nucleotide the release is slow.[148,330,334] Furthermore, the

hPolβ•DNA–(-)3TC-MP•(-)3TC-TP structure may suggest that the binding of the next nucleotide occurs before and provides energy for hPolβ translocation. This is contrary to the sequence of Steps 10_a and 3 in the canonical mechanism (Scheme 6.1). We were able to capture such a meta-stable complex because the nicked DNA–(-)3TC-MP substrate prevents hPolβ translocation, and the incoming (-)3TC-TP mismatch with the next nucleotide (dC) of the template would disfavor the strand-displacement activity of hPolβ.[335]

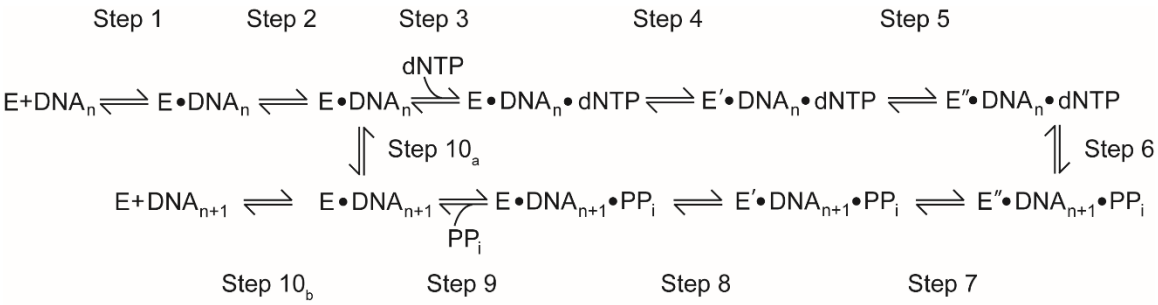
Alternatively, in the absence of the next correct nucleotide, PPi release could be triggered or aided by protein side chain movements or water solvation that disrupts interactions between PPi and protein side chains. Conversely, additional metal ions, such as the third divalent metal ion observed in recent time-resolved X-ray crystallography studies with Pols,[38-42] are unlikely to aid PPi release. In the structures following nucleotide incorporation and reopening of the thumb domain a third metal ion is only observed for the hPolβ•DNA–(-)3TC-MP•PPi structure (Figures 6.2D-6.2F and 6.8) but not with the hPolβ•DNA–(-)FTC-MP•PPi (Figure 6.2A-6.2C). In addition, the third metal ion and PPi remain associated following chemistry, and together, move upon reopening of the thumb domain when compared with the positions of the third divalent metal ion and PPi in the structures of correct nucleotide incorporation with hPolβ (Figure 6.8A and 6.8B). Accordingly, these data limit the role of the third metal ion to catalytic events occurring during the chemistry step (Step 6, Scheme 6.11). Despite the third metal ion remaining associated with PPi immediately following chemistry, it dissociates first. This is further supported by the product structures of hPolβ with correct nucleotide, wherein

the third divalent metal ion is released prior to domain reopening.[40] This is contrary to the initial hypothesis that the third divalent metal acts as an exit shuttle for PPi release.[38,42] However, for hPolη, it cannot be ruled out that the third divalent metal ion aids PPi release due to the lack of post-catalytic structures with the third divalent metal ion released and PPi still bound. Regardless of the mode in which PPi release is triggered, it is clear that the release of this small molecule occurs independent of the major, reverse conformational changes of hPolβ.

6.5 Conclusions

In summary, we report five post-catalytic structures of human DNA polymerase β (hPolβ) following incorporation of the L-nucleotide analogs, (-)3TC-TP and (-)FTC-TP. Contrary to common belief, we observe the first structural evidence, to our knowledge, that the inevitable closed→open thumb domain transition following nucleotide incorporation occurs before PPi release. Furthermore, the ternary structure in which an additional nucleotide is bound with a previously incorporated nucleotide suggests that the nucleotide binding to the complex following the closed→open conformational change may trigger PPi release and initiate the next catalytic cycle.

6.6 Schemes



Scheme 6.1: Canonical mechanism of nucleotide incorporation for DNA polymerases.

E, E′, and E″ represent three different conformations of a polymerase sampled during correct nucleotide incorporation.

6.7 Tables

	hPolβ•DNA•(-)3TC-TP (Mn^{2+})	hPolβ•DNA•(-)FTC-TP (Mn^{2+})	hPolβ•DNA–(-)3TC-MP•PP_i (Mg^{2+})	hPolβ•DNA–(-)FTC-MP•PP_i (Mn^{2+})	hPolβ•DNA–(-)3TC-MP• (-)3TC-TP (Mg^{2+})
Data collection*					
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell dimensions					
a, *b*, *c* (Å)	49.4, 81.4, 53.5	50.1, 82.3, 54.2	54.6, 79.3, 54.9	55.1, 79.3, 55.4	55.3,79.5, 55.3
α, β, γ (°)	90.0, 110.2, 90.0	90.0, 110.9, 90.0	90.0, 107.0, 90.0	90.0, 107.2, 90.0	90.0, 107.4, 90.0
Resolution (Å)	50.25-2.00 (2.05-2.00)	50.68-2.49 (2.59-2.49)	50.00-2.50 (2.59-2.50)	79.33-2.39 (2.48-2.39)	32.76-1.85 (1.89-1.85)
R_{merge}†	0.109 (0.433)	0.107 (0.545)	0.168 (0.708)	0.094 (0.474)	0.088 (0.686)
$I/\sigma I$	6.4 (2.0)	8.3 (2.0)	11.9 (4.30)	7.7 (2.0)	8.4 (2.0)
Completeness (%)	98.2 (97.6)	99.9 (99.9)	99.6 (100.0)	99.9 (100.0)	98.9 (98.7)
Redundancy	3.3 (3.0)	3.6 (3.4)	5.6 (5.8)	3.6 (3.3)	3.8 (3.8)
Refinement					
Resolution (Å)	46.41–2.00	46.87–2.49	25.18-2.50	44.52–2.39	32.76-1.85
No. reflections	25091	13868	14831	17190	36763
R_{work}/R_{free}‡	0.184/0.230	0.208/0.267	0.208/0.273	0.216/0.269	0.230/0.278
No. atoms					
Protein	2642	2631	2631	2368	2354
DNA	651	632	632	632	632
Nucleotide	27	28	-	-	27
B-factors					
Protein	21.37	30.91	30.91	43.71	28.97
DNA	26.53	38.70	38.70	33.59	30.66
Nucleotide	14.39	25.70	-	-	33.34
R.m.s deviations					
Bond lengths (Å)	0.014	0.020	0.010	0.012	0.007
Bond angles (º)	1.696	1.151	1.464	1.535	1.162
PDBID	5TB8	5TB9	5TBA	5TBB	5TBC

Table 6.1: Structural data collection and refinement statistics.

6.8 Figures

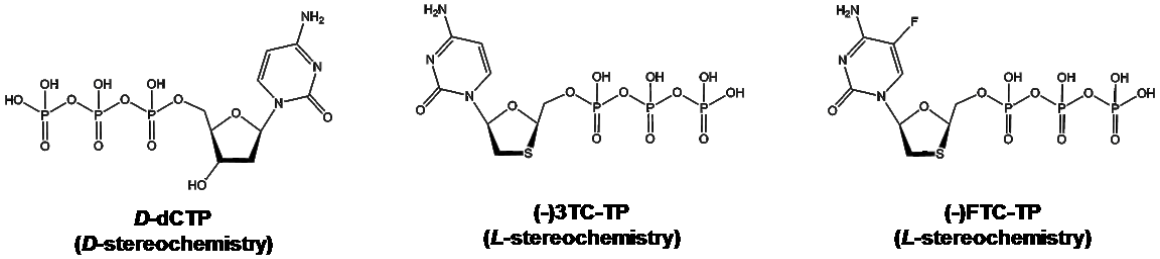


Figure 6.1: Chemical structures of *D*-dCTP, (-)3TC-TP, and (-)FTC-TP.

D-dCTP is the natural *D*-enantiomer of dCTP. (-)3TC-TP and (-)FTC-TP are the biologically active triphosphate forms of lamivudine [(-)3TC; (-)-β-*L*-2′,3′-dideoxy-3′-thiacytidine] and emtricitabine [(-)FTC; (-)-β-*L*-2′,3′-dideoxy-5-fluoro-3′-thiacytidine], respectively, and are analogs of the *L*-enantiomer of dCTP. In this article, we refer to "*D*-dCTP" simply as "dCTP".

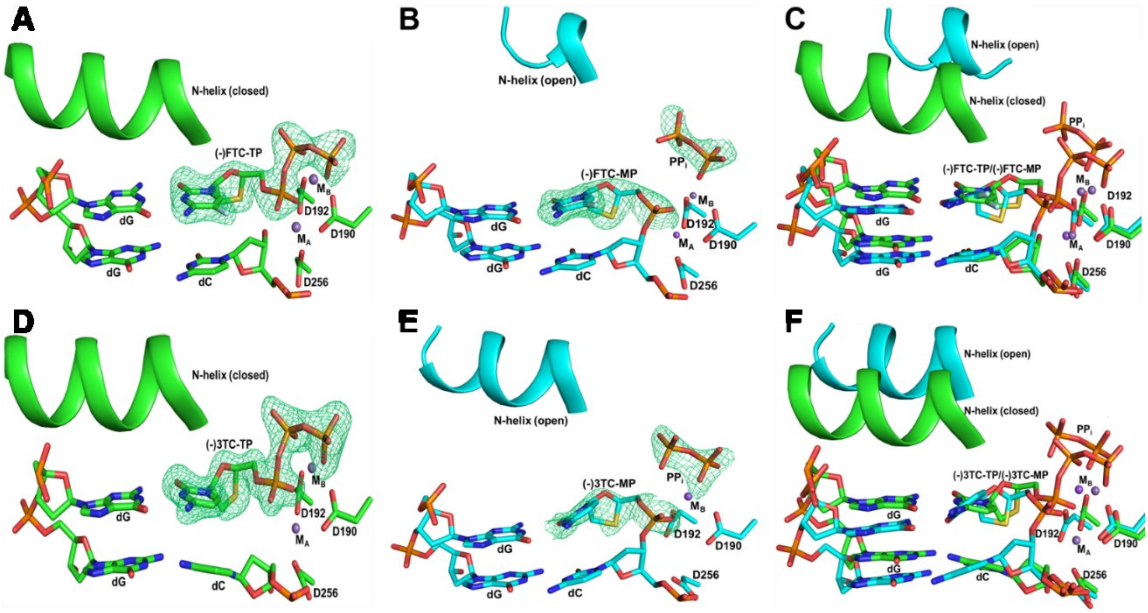


Figure 6.2: Active site views of bound and incorporated *L*-nucleotides.

(A) Closed pre-catalytic ternary complex of hPolβ•DNA•(-)FTC-TP (B) Open post-catalytic ternary complex of hPolβ•DNA•(-)FTC-MP•PP_i. (C) Superposition of structures presented in (A) and (B). (D) Closed pre-catalytic ternary complex of hPolβ•DNA•(-)3TC-TP. (E) Open post-catalytic ternary complex of hPolβ•DNA•(-)3TC-MP•PP_i. (F) Superposition of structures presented in (D) and (E). The N-helix is shown as a cartoon model. The metal ions, represented as spheres, are modeled as Mn^{2+} (M_A) and Mn^{2+} (M_B) in (A) and (D), Na^+ (M_A) and Mn^{2+} (M_B) in (B), and Na^+ (M_B) in (E). The side chains of D190, D192, and D256 are shown as sticks. The F_o-F_c omit map (3σ) is displayed as green mesh.

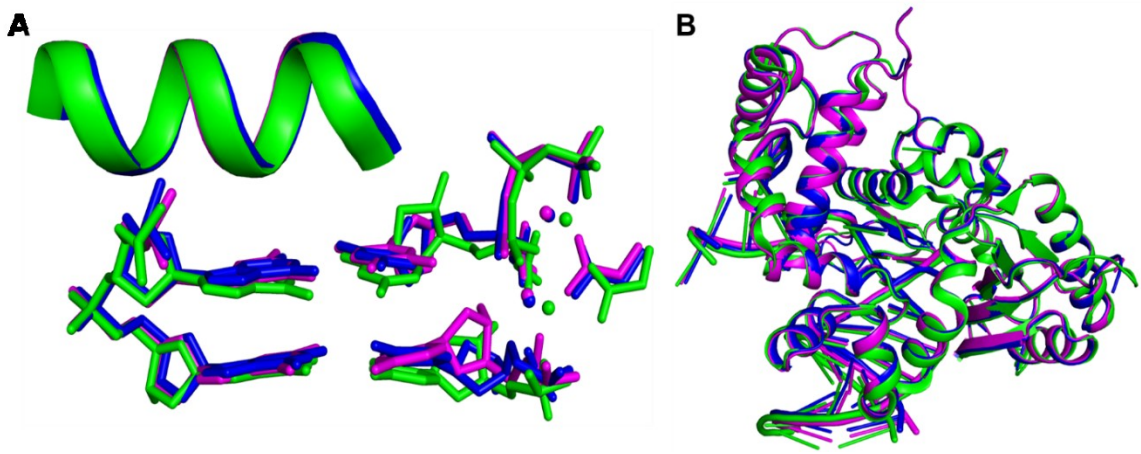


Figure 6.3: Superposition of natural and *L*-nucleotide closed, pre-catalytic ternary structures.

(A) Zoomed view of the superposition of the active sites of hPolβ•DNA•(-)3TC-TP (blue), hPolβ•DNA•(-)FTC-TP (magenta), and hPolβ•DNA•dCTP (PDB code 4KLD, green). Both the A-site and B-site metal ions, shown as spheres, are modeled as Mn^{2+} for hPolβ•DNA•(-)3TC-TP and hPolβ•DNA•(-)FTC-TP, and Ca^{2+} for hPolβ•DNA•dCTP. (B) Superposition of the overall tertiary structures of hPolβ•DNA•(-)3TC-TP (blue), hPolβ•DNA•(-)FTC-TP (magenta), and hPolβ•DNA•dCTP (green).

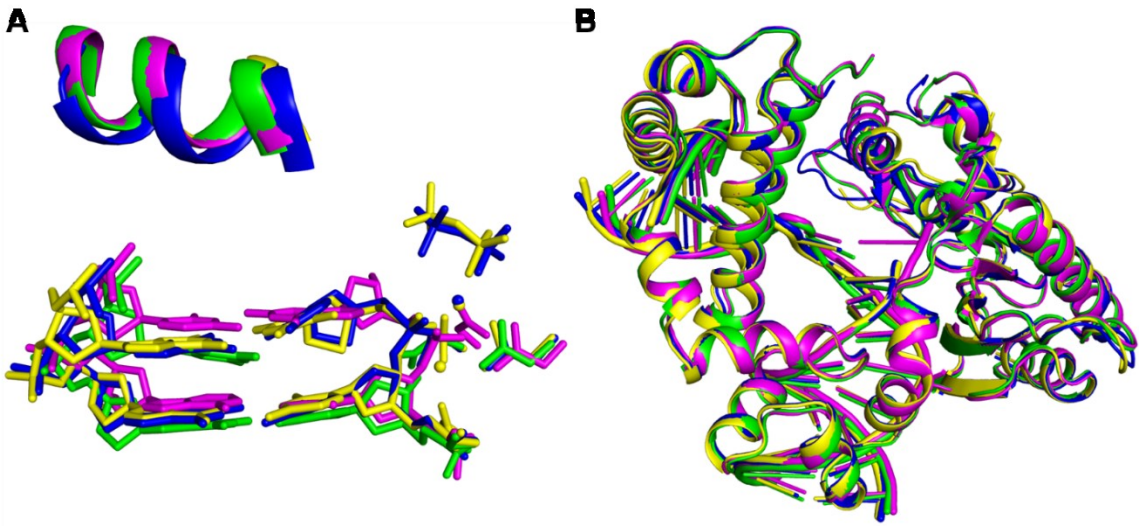


Figure 6.4: Superposition of open, post-catalytic *L*-nucleotide ternary structures and binary structures.

(A) Zoomed view of the superposition of the active sites of hPolβ•DNA–(-)3TC-MP•PP_i (blue), hPolβ•DNA–(-)FTC-MP•PP_i (yellow), hPolβ•DNA_{nicked} (magenta, PDB code 1BPZ), and hPolβ•DNA (green, PDB code 1BPX). Both the A-site and B-site metal ions, shown as spheres, are modeled as Mn^{2+} for hPolβ•DNA–(-)3TC-MP•PP_i and hPolβ•DNA–(-)FTC-MP•PP_i. (B) Superposition of the overall tertiary structures of hPolβ•DNA–(-)3TC-MP•PP_i (blue), hPolβ•DNA–(-)FTC-MP•PP_i (yellow), hPolβ•DNA_{nicked} (magenta), and hPolβ•DNA (green). DNA_{nicked} denotes a nicked DNA substrate.

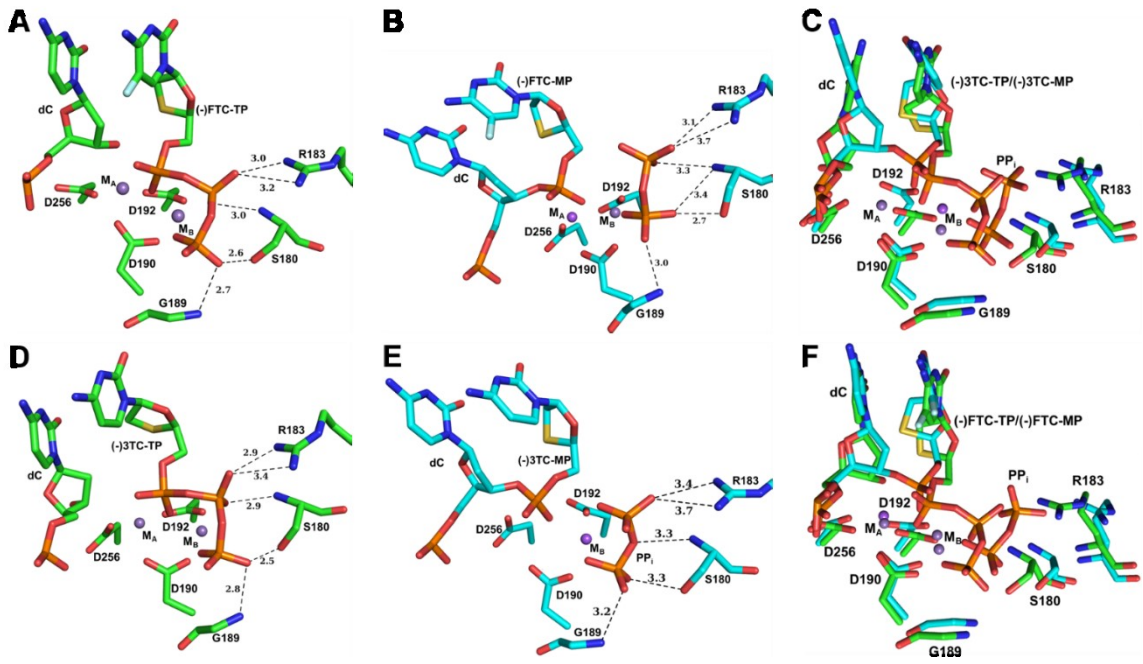


Figure 6.5: Zoomed interactions of the triphosphate of a nucleotide or the product pyrophosphate with palm domain residues.

(A) Closed pre-catalytic ternary complex of hPolβ•DNA•(-)FTC-TP. (B) Open post-catalytic ternary complex of hPolβ•DNA•(-)FTC-MP•PP_i. (C) Superposition of structures presented in (A) and (B). (D) Closed pre-catalytic ternary complex of hPolβ•DNA•(-)3TC-TP. (E) Open post-catalytic ternary complex of hPolβ•DNA•(-)3TC-MP•PP_i. (F) Superposition of structures presented in (D) and (E). In all panels, the triphosphate or pyrophosphate forms hydrogen bonds (dashes) with the backbone nitrogen atoms of G189 and S180 and the side chains of R183 and S180. Distances are given in Å. The metal ions, represented as spheres, are modeled as Mn^{2+} (M_A) and Mn^{2+} (M_B) in (A) and (D), Na^+ (M_A) and Mn^{2+} (M_B) in (B), and Na^+ (M_B) in (E). The active site residues are shown as sticks.

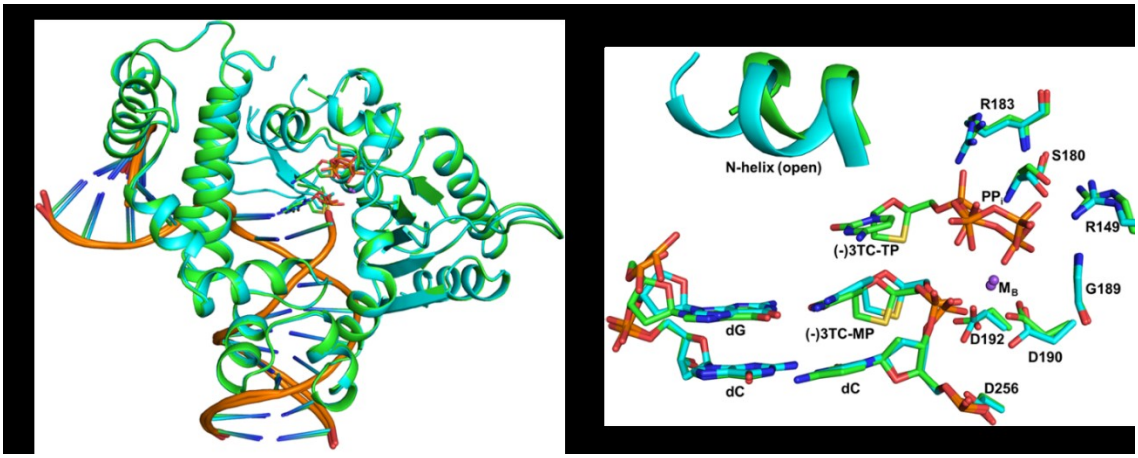


Figure 6.6: Superposition of post-catalytic structures of hPolβ with the phosphate forms of (-)3TC.

Overall superposition of the hPolβ•DNA–(-)3TC-MP•PP_i (cyan) and hPolβ•DNA–(-)3TC-MP•(-)3TC-TP (green) structures. Between the two structures, the DNA substrate and protein backbone overlay well, respectively, with only a slight difference in the position of the N-helix. (B) Active site superposition of the hPolβ•DNA–(-)3TC-MP•PP_i (cyan) and hPolβ•DNA–(-)3TC-MP•(-)3TC-TP (green) structures. The side chains (R183 and S180) and backbone atoms (G189 and S180) have similar interactions with PP_i and the triphosphate of (-)3TC-TP. The incoming (-)3TC-TP also has an additional interaction with R149 due to a slight repositioning of the γ-phosphate of (-)3TC-TP compared to PP_i.

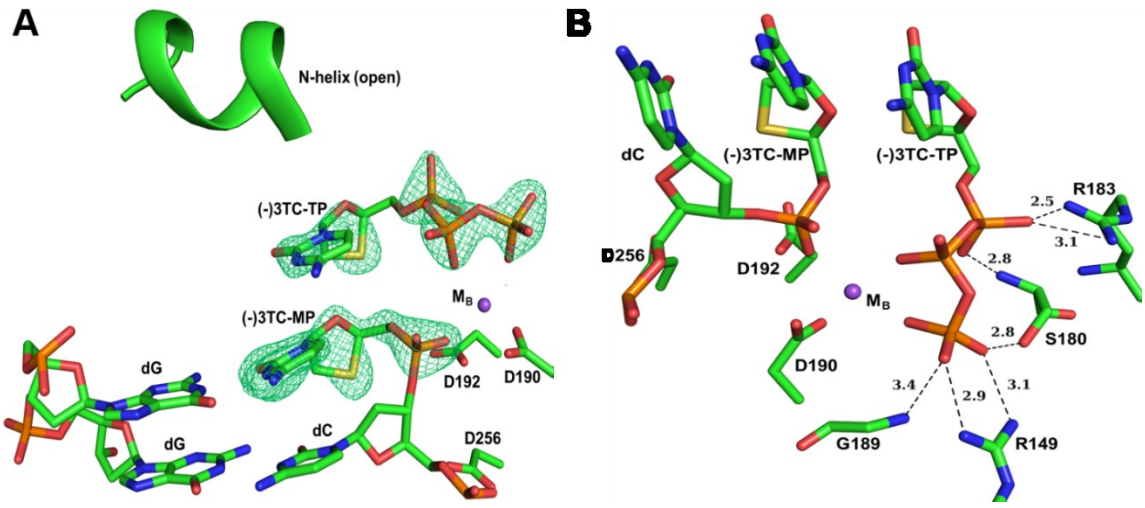


Figure 6.7: Zoomed active site views of the post-catalytic hPolβ•DNA–(-)3TC-MP•(-)3TC-TP structure.

(A) (-)3TC-TP bound in the open conformation following (-)3TC-MP incorporation. The N-helix is shown as a cartoon model. The F_o-F_c map (3σ) is shown as green mesh. (B) Interactions of the bound (-)3TC-TP with palm domain residues. Hydrogen bonds (dashes with distances given in Å) are formed with the backbone nitrogen atoms of G189 and S180 and the side chains of R183 and S180. Two unique hydrogen bonds are formed between the triphosphate of (-)3TC-TP and the side chain of R149. In both panels the B-site (M_B) metal ion, shown as a sphere, is modeled as Mn^{2+}. The active site residues are represented as sticks.

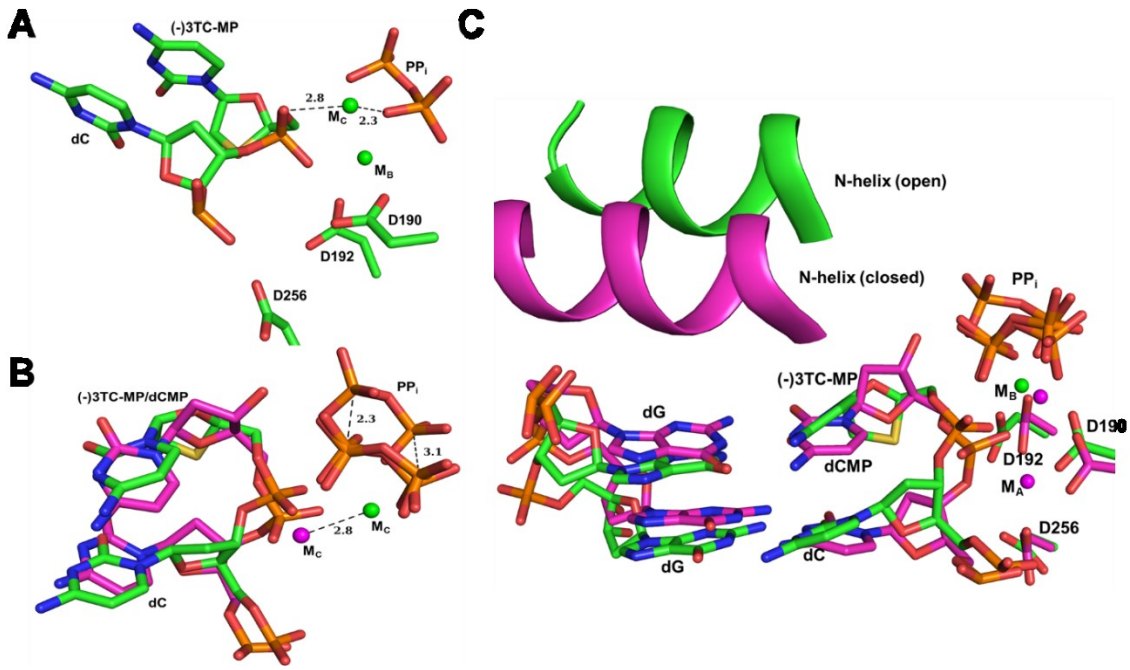


Figure 6.8: Evidence for a third metal ion following (-)3TC-TP incorporation.

(A) Stick model of zoomed hPolβ•DNA–(-)3TC-MP•PP_i structure with both the B-site (M_B) and C-site (M_C) metal ions shown as spheres and modeled as Na^+ ions. Ligands for the C-site metal ion are shown as dashes with distances in Å. (B and C) Superposition of the hPolβ•DNA–(-)3TC-MP•PP_i structure and the structure of hPolβ following normal dCTP incorporation (hPolβ•DNA–dCMP•PP_i, PDB code 4KLG). Metal ions for the hPolβ•DNA–dCMP•PP_i structure are modeled as Mg^{2+} ions. (B) Movement of the C-site metal ion (M_C) and PP_i following domain reopening are shown as dashes with distances in Å. The M_C ion for the hPolβ•DNA–(-)3TC-MP•PP_i structure is modeled as Na^+ based on the ligand distances and the concerted movement of M_C and PP_i. (C) Overall active site superposition demonstrating the change in the N-helix and similar positions of catalytic aspartate residues and the DNA template strand.

References

1. Palermo, G. et al. Catalytic metal ions and enzymatic processing of DNA and RNA. *Acc. Chem. Res.* **48**, 220-8 (2015).
2. Yang, W., Lee, J.Y. & Nowotny, M. Making and breaking nucleic acids: two-Mg2+-ion catalysis and substrate specificity. *Mol Cell* **22**, 5-13 (2006).
3. Steitz, T.A. & Steitz, J.A. A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A* **90**, 6498-502 (1993).
4. Johnson, K.A. Conformational coupling in DNA polymerase fidelity. *Annu Rev Biochem* **62**, 685-713 (1993).
5. Joyce, C.M. & Benkovic, S.J. DNA polymerase fidelity: kinetics, structure, and checkpoints. *Biochemistry* **43**, 14317-24 (2004).
6. Showalter, A.K. & Tsai, M.D. A reexamination of the nucleotide incorporation fidelity of DNA polymerases. *Biochemistry* **41**, 10571-6 (2002).
7. Rothwell, P.J., Mitaksov, V. & Waksman, G. Motions of the fingers subdomain of klentaq1 are fast and not rate limiting: implications for the molecular basis of fidelity in DNA polymerases. *Mol Cell* **19**, 345-55 (2005).
8. Rothwell, P.J. & Waksman, G. Structure and mechanism of DNA polymerases. *Adv Protein Chem* **71**, 401-40 (2005).
9. Raper, A.T., Gadkari, V.V., Maxwell, B.A. & Suo, Z. Single-Molecule Investigation of Response to Oxidative DNA Damage by a Y-Family DNA Polymerase. *Biochemistry* **55**, 2187-96 (2016).
10. Raper, A.T. & Suo, Z. Investigation of Intradomain Motions of a Y-Family DNA Polymerase during Substrate Binding and Catalysis. *Biochemistry* **55**, 5832-5844 (2016).
11. Xu, C., Maxwell, B.A., Brown, J.A., Zhang, L. & Suo, Z. Global conformational dynamics of a Y-family DNA polymerase during catalysis. *PLoS Biol* **7**, e1000225 (2009).
12. Maxwell, B.A., Xu, C. & Suo, Z. DNA lesion alters global conformational dynamics of Y-family DNA polymerase during catalysis. *J Biol Chem* **287**, 13040-7 (2012).
13. Maxwell, B.A., Xu, C. & Suo, Z. Conformational dynamics of a Y-family DNA polymerase during substrate binding and catalysis as revealed by interdomain Forster resonance energy transfer. *Biochemistry* **53**, 1768-78 (2014).
14. Xu, C., Maxwell, B.A. & Suo, Z. Conformational dynamics of Thermus aquaticus DNA polymerase I during catalysis. *J Mol Biol* **426**, 2901-17 (2014).

15. Bessman, M.J., Kornberg, A., Lehman, I.R. & Simms, E.S. Enzymic synthesis of deoxyribonucleic acid. *Biochim Biophys Acta* **21**, 197-8 (1956).
16. Lehman, I.R., Bessman, M.J., Simms, E.S. & Kornberg, A. Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from Escherichia coli. *J Biol Chem* **233**, 163-70 (1958).
17. Joyce, C.M. & Steitz, T.A. Function and structure relationships in DNA polymerases. *Annu Rev Biochem* **63**, 777-822 (1994).
18. Braithwaite, D.K. & Ito, J. Compilation, alignment, and phylogenetic relationships of DNA polymerases. *Nucleic Acids Res* **21**, 787-802 (1993).
19. Garcia-Diaz, M. & Bebenek, K. Multiple functions of DNA polymerases. *CRC Crit. Rev. Plant Sci.* **26**, 105-122 (2007).
20. Yang, W. & Gao, Y. Translesion and Repair DNA Polymerases: Diverse Structure and Mechanism. *Annu Rev Biochem* (2018).
21. Beard, W.A. & Wilson, S.H. Structure and mechanism of DNA polymerase beta. *Biochemistry* **53**, 2768-80 (2014).
22. Maxwell, B.A. & Suo, Z. Recent insight into the kinetic mechanisms and conformational dynamics of Y-Family DNA polymerases. *Biochemistry* **53**, 2804-14 (2014).
23. Fowler, J.D. & Suo, Z. Biochemical, structural, and physiological characterization of terminal deoxynucleotidyl transferase. *Chem Rev* **106**, 2092-110 (2006).
24. Berdis, A.J. Mechanisms of DNA polymerases. *Chem Rev* **109**, 2862-79 (2009).
25. Steitz, T.A. DNA polymerases: structural diversity and common mechanisms. *J Biol Chem* **274**, 17395-8 (1999).
26. Steitz, T.A. A mechanism for all polymerases. *Nature* **391**, 231-2 (1998).
27. Tsai, Y.C. & Johnson, K.A. A new paradigm for DNA polymerase specificity. *Biochemistry* **45**, 9675-87 (2006).
28. Xiang, Y., Goodman, M.F., Beard, W.A., Wilson, S.H. & Warshel, A. Exploring the role of large conformational changes in the fidelity of DNA polymerase beta. *Proteins* **70**, 231-47 (2008).
29. Wong, I., Patel, S.S. & Johnson, K.A. An induced-fit kinetic mechanism for DNA replication fidelity: direct measurement by single-turnover kinetics. *Biochemistry* **30**, 526-37 (1991).
30. Freemont, P.S., Friedman, J.M., Beese, L.S., Sanderson, M.R. & Steitz, T.A. Cocrystal structure of an editing complex of Klenow fragment with DNA. *Proc Natl Acad Sci U S A* **85**, 8924-8 (1988).
31. Beese, L.S. & Steitz, T.A. Structural basis for the 3'-5' exonuclease activity of Escherichia coli DNA polymerase I: a two metal ion mechanism. *Embo J* **10**, 25-33 (1991).
32. Batra, V.K. et al. Magnesium-induced assembly of a complete DNA polymerase catalytic complex. *Structure* **14**, 757-66 (2006).
33. 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* **8**, 54-63 (1998).

34. Nakamura, T.Z., Ye; Yamagata, Yuriko; Hua, Yue-jin; Yang, Wei;. Mechanism of the nucleotidyl-transfer reaction in DNA polymerase revealed by time-resolved crystallography. *Biophysics* **9**, 31-36 (2013).
35. Yang, W. Damage repair DNA polymerases Y. *Curr Opin Struct Biol* **13**, 23-30 (2003).
36. Yang, W. An equivalent metal ion in one- and two-metal-ion catalysis. *Nat Struct Mol Biol* **15**, 1228-31 (2008).
37. Derbyshire, V. et al. Genetic and crystallographic studies of the 3',5'-exonucleolytic site of DNA polymerase I. *Science* **240**, 199-201 (1988).
38. Nakamura, T., Zhao, Y., Yamagata, Y., Hua, Y.J. & Yang, W. Watching DNA polymerase eta make a phosphodiester bond. *Nature* **487**, 196-201 (2012).
39. Gao, Y. & Yang, W. Capture of a third Mg(2)(+) is essential for catalyzing DNA synthesis. *Science* **352**, 1334-7 (2016).
40. Freudenthal, B.D., Beard, W.A., Shock, D.D. & Wilson, S.H. Observing a DNA polymerase choose right from wrong. *Cell* **154**, 157-68 (2013).
41. Freudenthal, B.D. et al. Uncovering the polymerase-induced cytotoxicity of an oxidized nucleotide. *Nature* **517**, 635-9 (2015).
42. Vyas, R., Reed, A.J., Tokarsky, E.J. & Suo, Z. Viewing Human DNA Polymerase beta Faithfully and Unfaithfully Bypass an Oxidative Lesion by Time-Dependent Crystallography. *J Am Chem Soc* **137**, 5225-30 (2015).
43. Reed, A.J., Vyas, R., Raper, A.T. & Suo, Z. Structural Insights into the Post-Chemistry Steps of Nucleotide Incorporation Catalyzed by a DNA Polymerase. *J Am Chem Soc* **139**, 465-471 (2017).
44. Whitaker, A.M., Smith, M.R., Schaich, M.A. & Freudenthal, B.D. Capturing a mammalian DNA polymerase extending from an oxidized nucleotide. *Nucleic Acids Res* **45**, 6934-6944 (2017).
45. Reed, A.J. & Suo, Z. Time-Dependent Extension from an 8-Oxoguanine Lesion by Human DNA Polymerase Beta. *J Am Chem Soc* **139**, 9684-9690 (2017).
46. Jamsen, J.A. et al. Time-lapse crystallography snapshots of a double-strand break repair polymerase in action. *Nat. Commun.* **8**, 253 (2017).
47. Yang, W., Weng, P.J. & Gao, Y. A new paradigm of DNA synthesis: three-metal-ion catalysis. *Cell Biosci* **6**, 51 (2016).
48. Freudenthal, B.D., Beard, W.A. & Wilson, S.H. New structural snapshots provide molecular insights into the mechanism of high fidelity DNA synthesis. *DNA Repair (Amst)* **32**, 3-9 (2015).
49. Genna, V., Gaspari, R., Dal Peraro, M. & De Vivo, M. Cooperative motion of a key positively charged residue and metal ions for DNA replication catalyzed by human DNA Polymerase-eta. *Nucleic Acids Res* **44**, 2827-36 (2016).
50. Perera, L. et al. Requirement for transient metal ions revealed through computational analysis for DNA polymerase going in reverse. *Proc Natl Acad Sci U S A* **112**, E5228-36 (2015).
51. Perera, L., Freudenthal, B.D., Beard, W.A., Pedersen, L.G. & Wilson, S.H. Revealing the role of the product metal in DNA polymerase beta catalysis. *Nucleic Acids Res* **45**, 2736-2745 (2017).

52. Liu, X., Bushnell, D.A. & Kornberg, R.D. RNA polymerase II transcription: structure and mechanism. *Biochim Biophys Acta* **1829**, 2-8 (2013).
53. Svetlov, V. & Nudler, E. Basic mechanism of transcription by RNA polymerase II. *Biochim Biophys Acta* **1829**, 20-8 (2013).
54. Hahn, S. Structure and mechanism of the RNA polymerase II transcription machinery. *Nat Struct Mol Biol* **11**, 394-403 (2004).
55. Kochetkov, S.N., Rusakova, E.E. & Tunitskaya, V.L. Recent studies of T7 RNA polymerase mechanism. *FEBS Lett* **440**, 264-7 (1998).
56. Ebright, R.H. RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *J Mol Biol* **304**, 687-98 (2000).
57. Moon, A.F. et al. The X family portrait: structural insights into biological functions of X family polymerases. *DNA Repair (Amst)* **6**, 1709-25 (2007).
58. Burgers, P.M.J. & Kunkel, T.A. Eukaryotic DNA Replication Fork. *Annu Rev Biochem* **86**, 417-438 (2017).
59. Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A. & Steitz, T.A. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**, 1783-90 (1992).
60. Steitz, T.A. et al. Two DNA polymerases: HIV reverse transcriptase and the Klenow fragment of Escherichia coli DNA polymerase I. *Cold Spring Harb Symp Quant Biol* **58**, 495-504 (1993).
61. Pelletier, H., Sawaya, M.R., Kumar, A., Wilson, S.H. & Kraut, J. Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. *Science* **264**, 1891-903 (1994).
62. Sawaya, M.R., Pelletier, H., Kumar, A., Wilson, S.H. & Kraut, J. Crystal structure of rat DNA polymerase beta: evidence for a common polymerase mechanism. *Science* **264**, 1930-5 (1994).
63. Doublie, S., Tabor, S., Long, A.M., Richardson, C.C. & Ellenberger, T. Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution. *Nature* **391**, 251-8 (1998).
64. Li, Y., Kong, Y., Korolev, S. & Waksman, G. Crystal structures of the Klenow fragment of Thermus aquaticus DNA polymerase I complexed with deoxyribonucleoside triphosphates. *Protein Sci* **7**, 1116-23 (1998).
65. Li, Y., Korolev, S. & Waksman, G. Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of Thermus aquaticus DNA polymerase I: structural basis for nucleotide incorporation. *EMBO J* **17**, 7514-25 (1998).
66. Korolev, S., Nayal, M., Barnes, W.M., Di Cera, E. & Waksman, G. Crystal structure of the large fragment of Thermus aquaticus DNA polymerase I at 2.5-A resolution: structural basis for thermostability. *Proc Natl Acad Sci U S A* **92**, 9264-8 (1995).
67. Wang, J., Yu, P., Lin, T.C., Konigsberg, W.H. & Steitz, T.A. Crystal structures of an NH2-terminal fragment of T4 DNA polymerase and its complexes with single-stranded DNA and with divalent metal ions. *Biochemistry* **35**, 8110-9 (1996).

68. Franklin, M.C., Wang, J. & Steitz, T.A. Structure of the replicating complex of a pol alpha family DNA polymerase. *Cell* **105**, 657-67 (2001).
69. Hogg, M., Wallace, S.S. & Doublie, S. Crystallographic snapshots of a replicative DNA polymerase encountering an abasic site. *EMBO J* **23**, 1483-93 (2004).
70. Wang, M. et al. Insights into base selectivity from the 1.8 A resolution structure of an RB69 DNA polymerase ternary complex. *Biochemistry* **50**, 581-90 (2011).
71. Evans, R.J. et al. Structure of PolC reveals unique DNA binding and fidelity determinants. *Proc Natl Acad Sci U S A* **105**, 20695-700 (2008).
72. Bailey, S., Wing, R.A. & Steitz, T.A. The structure of T. aquaticus DNA polymerase III is distinct from eukaryotic replicative DNA polymerases. *Cell* **126**, 893-904 (2006).
73. Wing, R.A., Bailey, S. & Steitz, T.A. Insights into the replisome from the structure of a ternary complex of the DNA polymerase III alpha-subunit. *J Mol Biol* **382**, 859-69 (2008).
74. Barros, T. et al. A structural role for the PHP domain in E. coli DNA polymerase III. *BMC Struct Biol* **13**, 8 (2013).
75. Pelletier, H., Sawaya, M.R., Wolfle, W., Wilson, S.H. & Kraut, J. A structural basis for metal ion mutagenicity and nucleotide selectivity in human DNA polymerase beta. *Biochemistry* **35**, 12762-77 (1996).
76. Garcia-Diaz, M., Bebenek, K., Krahn, J.M., Kunkel, T.A. & Pedersen, L.C. A closed conformation for the Pol lambda catalytic cycle. *Nat Struct Mol Biol* **12**, 97-8 (2005).
77. Moon, A.F. et al. Structural insight into the substrate specificity of DNA Polymerase mu. *Nat Struct Mol Biol* **14**, 45-53 (2007).
78. Pelletier, H., Sawaya, M.R., Wolfle, W., Wilson, S.H. & Kraut, J. Crystal structures of human DNA polymerase beta complexed with DNA: implications for catalytic mechanism, processivity, and fidelity. *Biochemistry* **35**, 12742-61 (1996).
79. Sawaya, M.R., Prasad, R., Wilson, S.H., Kraut, J. & Pelletier, H. Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* **36**, 11205-15 (1997).
80. Biertumpfel, C. et al. Structure and mechanism of human DNA polymerase eta. *Nature* **465**, 1044-8 (2010).
81. 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* **284**, 22467-80 (2009).
82. Nair, D.T., Johnson, R.E., Prakash, S., Prakash, L. & Aggarwal, A.K. Replication by human DNA polymerase-iota occurs by Hoogsteen base-pairing. *Nature* **430**, 377-80 (2004).
83. Swan, M.K., Johnson, R.E., Prakash, L., Prakash, S. & Aggarwal, A.K. Structure of the human Rev1-DNA-dNTP ternary complex. *J Mol Biol* **390**, 699-709 (2009).

84. Ling, H., Boudsocq, F., Woodgate, R. & Yang, W. Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. *Cell* **107**, 91-102 (2001).
85. Trincao, J. et al. Structure of the catalytic core of S. cerevisiae DNA polymerase eta: implications for translesion DNA synthesis. *Mol Cell* **8**, 417-26 (2001).
86. Ummat, A. et al. Human DNA polymerase eta is pre-aligned for dNTP binding and catalysis. *J Mol Biol* **415**, 627-34 (2012).
87. Gaur, V. et al. Structural and kinetic insights into binding and incorporation of L-nucleotide analogs by a Y-family DNA polymerase. *Nucleic Acids Res* **42**, 9984-95 (2014).
88. Huang, H., Chopra, R., Verdine, G.L. & Harrison, S.C. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **282**, 1669-75 (1998).
89. Kuchta, R.D., Mizrahi, V., Benkovic, P.A., Johnson, K.A. & Benkovic, S.J. Kinetic mechanism of DNA polymerase I (Klenow). *Biochemistry* **26**, 8410-7 (1987).
90. Patel, S.S., Wong, I. & Johnson, K.A. Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant. *Biochemistry* **30**, 511-25 (1991).
91. Zahurancik, W.J., Klein, S.J. & Suo, Z. Kinetic mechanism of DNA polymerization catalyzed by human DNA polymerase epsilon. *Biochemistry* **52**, 7041-9 (2013).
92. Brown, J.A. & Suo, Z. Elucidating the kinetic mechanism of DNA polymerization catalyzed by Sulfolobus solfataricus P2 DNA polymerase B1. *Biochemistry* **48**, 7502-11 (2009).
93. Fiala, K.A. & Suo, Z. Pre-steady-state kinetic studies of the fidelity of Sulfolobus solfataricus P2 DNA polymerase IV. *Biochemistry* **43**, 2106-15 (2004).
94. Fiala, K.A. & Suo, Z. Mechanism of DNA polymerization catalyzed by Sulfolobus solfataricus P2 DNA polymerase IV. *Biochemistry* **43**, 2116-25 (2004).
95. Donlin, M.J., Patel, S.S. & Johnson, K.A. Kinetic partitioning between the exonuclease and polymerase sites in DNA error correction. *Biochemistry* **30**, 538-46 (1991).
96. Kati, W.M., Johnson, K.A., Jerva, L.F. & Anderson, K.S. Mechanism and fidelity of HIV reverse transcriptase. *J Biol Chem* **267**, 25988-97 (1992).
97. Beard, W.A. & Wilson, S.H. Structure and mechanism of DNA polymerase Beta. *Chem Rev* **106**, 361-82 (2006).
98. Eger, B.T. & Benkovic, S.J. Minimal kinetic mechanism for misincorporation by DNA polymerase I (Klenow fragment). *Biochemistry* **31**, 9227-36 (1992).
99. Dahlberg, M.E. & Benkovic, S.J. Kinetic mechanism of DNA polymerase I (Klenow fragment): identification of a second conformational change and evaluation of the internal equilibrium constant. *Biochemistry* **30**, 4835-43 (1991).
100. Frey, M.W., Sowers, L.C., Millar, D.P. & Benkovic, S.J. The nucleotide analog 2-aminopurine as a spectroscopic probe of nucleotide incorporation by the Klenow

fragment of Escherichia coli polymerase I and bacteriophage T4 DNA polymerase. *Biochemistry* **34**, 9185-92 (1995).
101. Hsieh, J.C., Zinnen, S. & Modrich, P. Kinetic mechanism of the DNA-dependent DNA polymerase activity of human immunodeficiency virus reverse transcriptase. *J Biol Chem* **268**, 24607-13 (1993).
102. Capson, T.L. et al. Kinetic characterization of the polymerase and exonuclease activities of the gene 43 protein of bacteriophage T4. *Biochemistry* **31**, 10984-94 (1992).
103. Washington, M.T., Prakash, L. & Prakash, S. Yeast DNA polymerase eta utilizes an induced-fit mechanism of nucleotide incorporation. *Cell* **107**, 917-27 (2001).
104. Washington, M.T., Johnson, R.E., Prakash, L. & Prakash, S. The mechanism of nucleotide incorporation by human DNA polymerase eta differs from that of the yeast enzyme. *Mol Cell Biol* **23**, 8316-22 (2003).
105. Wohrl, B.M., Krebs, R., Goody, R.S. & Restle, T. Refined model for primer/template binding by HIV-1 reverse transcriptase: pre-steady-state kinetic analyses of primer/template binding and nucleotide incorporation events distinguish between different binding modes depending on the nature of the nucleic acid substrate. *J Mol Biol* **292**, 333-44 (1999).
106. Schermerhorn, K.M. & Gardner, A.F. Pre-steady-state Kinetic Analysis of a Family D DNA Polymerase from Thermococcus sp. 9 degrees N Reveals Mechanisms for Archaeal Genomic Replication and Maintenance. *J Biol Chem* **290**, 21800-10 (2015).
107. Lahiri, I., Mukherjee, P. & Pata, J.D. Kinetic characterization of exonuclease-deficient Staphylococcus aureus PolC, a C-family replicative DNA polymerase. *PLoS One* **8**, e63489 (2013).
108. Einolf, H.J. & Guengerich, F.P. Kinetic analysis of nucleotide incorporation by mammalian DNA polymerase delta. *J Biol Chem* **275**, 16316-22 (2000).
109. Fiala, K.A., Abdel-Gawad, W. & Suo, Z. Pre-Steady-State Kinetic Studies of the Fidelity and Mechanism of Polymerization Catalyzed by Truncated Human DNA Polymerase lambda. *Biochemistry* **43**, 6751-62 (2004).
110. Werneburg, B.G. et al. DNA polymerase beta: pre-steady-state kinetic analysis and roles of arginine-283 in catalysis and fidelity. *Biochemistry* **35**, 7041-50 (1996).
111. Maxwell, B.A. & Suo, Z. Single-molecule investigation of substrate binding kinetics and protein conformational dynamics of a B-family replicative DNA polymerase. *J Biol Chem* **288**, 11590-600 (2013).
112. Johnson, K.A. The kinetic and chemical mechanism of high-fidelity DNA polymerases. *Biochim Biophys Acta* **1804**, 1041-8 (2010).
113. McClure, W.R. & Jovin, T.M. The steady state kinetic parameters and non-processivity of Escherichia coli deoxyribonucleic acid polymerase I. *J Biol Chem* **250**, 4073-80 (1975).
114. Bryant, F.R., Johnson, K.A. & Benkovic, S.J. Elementary steps in the DNA polymerase I reaction pathway. *Biochemistry* **22**, 3537-46 (1983).

115. Liu, M.S. et al. Structural Mechanism for the Fidelity Modulation of DNA Polymerase lambda. *J Am Chem Soc* **138**, 2389-98 (2016).
116. Ahn, J., Kraynov, V.S., Zhong, X., Werneburg, B.G. & Tsai, M.D. DNA polymerase beta: effects of gapped DNA substrates on dNTP specificity, fidelity, processivity and conformational changes. *Biochem J* **331 (Pt 1)**, 79-87 (1998).
117. Brown, J.A., Pack, L.R., Sanman, L.E. & Suo, Z. Efficiency and fidelity of human DNA polymerases lambda and beta during gap-filling DNA synthesis. *DNA Repair (Amst)* **10**, 24-33 (2011).
118. Wu, W.J. et al. How a low-fidelity DNA polymerase chooses non-Watson-Crick from Watson-Crick incorporation. *J Am Chem Soc* **136**, 4927-37 (2014).
119. Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G. & Steitz, T.A. Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP. *Nature* **313**, 762-6 (1985).
120. Beese, L.S., Friedman, J.M. & Steitz, T.A. Crystal structures of the Klenow fragment of DNA polymerase I complexed with deoxynucleoside triphosphate and pyrophosphate. *Biochemistry* **32**, 14095-101 (1993).
121. Chen, Y. et al. Unique 5'-P recognition and basis for dG:dGTP misincorporation of ASFV DNA polymerase X. *PLoS Biol* **15**, e1002599 (2017).
122. Nakane, S., Ishikawa, H., Nakagawa, N., Kuramitsu, S. & Masui, R. The structural basis of the kinetic mechanism of a gap-filling X-family DNA polymerase that binds Mg(2+)-dNTP before binding to DNA. *J Mol Biol* **417**, 179-96 (2012).
123. Kumar, S., Bakhtina, M. & Tsai, M.D. Altered order of substrate binding by DNA polymerase X from African Swine Fever virus. *Biochemistry* **47**, 7875-87 (2008).
124. Maciejewski, M.W. et al. Solution structure of a viral DNA repair polymerase. *Nat Struct Biol* **8**, 936-41 (2001).
125. Showalter, A.K. & Tsai, M.D. A DNA polymerase with specificity for five base pairs. *J Am Chem Soc* **123**, 1776-7 (2001).
126. Brown, J.A., Fowler, J.D. & Suo, Z. Kinetic basis of nucleotide selection employed by a protein template-dependent DNA polymerase. *Biochemistry* **49**, 5504-10 (2010).
127. Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. & Aggarwal, A.K. Rev1 employs a novel mechanism of DNA synthesis using a protein template. *Science* **309**, 2219-22 (2005).
128. Nelson, J.R., Lawrence, C.W. & Hinkle, D.C. Deoxycytidyl transferase activity of yeast REV1 protein. *Nature* **382**, 729-31 (1996).
129. Howell, C.A., Prakash, S. & Washington, M.T. Pre-steady-state kinetic studies of protein-template-directed nucleotide incorporation by the yeast Rev1 protein. *Biochemistry* **46**, 13451-9 (2007).
130. Chagovetz, A.M., Sweasy, J.B. & Preston, B.D. Increased activity and fidelity of DNA polymerase beta on single-nucleotide gapped DNA. *J Biol Chem* **272**, 27501-4 (1997).
131. Garcia-Diaz, M. et al. DNA polymerase lambda, a novel DNA repair enzyme in human cells. *J Biol Chem* **277**, 13184-91 (2002).

132. Fiala, K.A., Hypes, C.D. & Suo, Z. Mechanism of abasic lesion bypass catalyzed by a Y-family DNA polymerase. *J Biol Chem* **282**, 8188-98 (2007).
133. Brown, J.A., Newmister, S.A., Fiala, K.A. & Suo, Z. Mechanism of double-base lesion bypass catalyzed by a Y-family DNA polymerase. *Nucleic Acids Res* **36**, 3867-78 (2008).
134. Maxwell, B.A. & Suo, Z. Kinetic basis for the differing response to an oxidative lesion by a replicative and a lesion bypass DNA polymerase from Sulfolobus solfataricus. *Biochemistry* **51**, 3485-96 (2012).
135. Kusumoto, R., Masutani, C., Shimmyo, S., Iwai, S. & Hanaoka, F. DNA binding properties of human DNA polymerase eta: implications for fidelity and polymerase switching of translesion synthesis. *Genes Cells* **9**, 1139-50 (2004).
136. Ohashi, E. et al. Error-prone bypass of certain DNA lesions by the human DNA polymerase kappa. *Genes Dev* **14**, 1589-94 (2000).
137. Woodgate, R. A plethora of lesion-replicating DNA polymerases. *Genes Dev* **13**, 2191-5 (1999).
138. Sherrer, S.M. et al. Mechanistic studies of the bypass of a bulky single-base lesion catalyzed by a Y-family DNA polymerase. *J Biol Chem* **284**, 6379-88 (2009).
139. Gadkari, V.V., Tokarsky, E.J., Malik, C.K., Basu, A.K. & Suo, Z. Mechanistic investigation of the bypass of a bulky aromatic DNA adduct catalyzed by a Y-family DNA polymerase. *DNA Repair (Amst)* **21**, 65-77 (2014).
140. Garcia-Gomez, S. et al. PrimPol, an archaic primase/polymerase operating in human cells. *Mol Cell* **52**, 541-53 (2013).
141. Meng, X. et al. DNA damage alters DNA polymerase delta to a form that exhibits increased discrimination against modified template bases and mismatched primers. *Nucleic Acids Res* **37**, 647-57 (2009).
142. Schmitt, M.W., Matsumoto, Y. & Loeb, L.A. High fidelity and lesion bypass capability of human DNA polymerase delta. *Biochimie* **91**, 1163-72 (2009).
143. Higuchi, K. et al. Fate of DNA replication fork encountering a single DNA lesion during oriC plasmid DNA replication in vitro. *Genes Cells* **8**, 437-49 (2003).
144. Johnson, K.A. *In 1 Transient-State Kinetic Analysis of Enzyme Reaction Pathways*, 1-61 (Academic Press, Inc., San Diego, CA, 1992).
145. Zahurancik, W.J., Baranovskiy, A.G., Tahirov, T.H. & Suo, Z. Comparison of the kinetic parameters of the truncated catalytic subunit and holoenzyme of human DNA polymerase varepsilon. *DNA Repair (Amst)* **29**, 16-22 (2015).
146. Vande Berg, B.J., Beard, W.A. & Wilson, S.H. DNA structure and aspartate 276 influence nucleotide binding to human DNA polymerase beta. Implication for the identity of the rate-limiting conformational change. *J Biol Chem* **276**, 3408-16 (2001).
147. Brown, J.A. et al. Pre-Steady-State Kinetic Analysis of Truncated and Full-Length Saccharomyces cerevisiae DNA Polymerase Eta. *J. Nucleic Acids* **2010**(2010).
148. Zhao, L. et al. Elucidation of kinetic mechanisms of human translesion DNA polymerase kappa using tryptophan mutants. *FEBS J* **281**, 4394-410 (2014).

149. Wong, J.H., Fiala, K.A., Suo, Z. & Ling, H. Snapshots of a Y-family DNA polymerase in replication: substrate-induced conformational transitions and implications for fidelity of Dpo4. *J Mol Biol* **379**, 317-30 (2008).
150. Lee, E., Fowler, J.D., Suo, Z. & Wu, Z. Backbone assignment of the binary complex of the full length Sulfolobus solfataricus DNA polymerase IV and DNA. *Biomol. NMR Assign.* **11**, 39-43 (2017).
151. Chu, X. et al. Dynamic conformational change regulates the protein-DNA recognition: an investigation on binding of a Y-family polymerase to its target DNA. *PLoS Comput Biol* **10**, e1003804 (2014).
152. Lone, S. et al. Human DNA polymerase kappa encircles DNA: implications for mismatch extension and lesion bypass. *Mol Cell* **25**, 601-14 (2007).
153. Moon, A.F. et al. Sustained active site rigidity during synthesis by human DNA polymerase mu. *Nat Struct Mol Biol* **21**, 253-60 (2014).
154. Lang, T., Maitra, M., Starcevic, D., Li, S.X. & Sweasy, J.B. A DNA polymerase beta mutant from colon cancer cells induces mutations. *Proc Natl Acad Sci U S A* **101**, 6074-9 (2004).
155. Brown, J.A. et al. Identification of critical residues for the tight binding of both correct and incorrect nucleotides to human DNA polymerase lambda. *J Mol Biol* **403**, 505-15 (2010).
156. Garcia-Diaz, M. et al. A structural solution for the DNA polymerase lambda-dependent repair of DNA gaps with minimal homology. *Mol Cell* **13**, 561-72 (2004).
157. Wilson, S.H. & Kunkel, T.A. Passing the baton in base excision repair. *Nat Struct Biol* **7**, 176-8 (2000).
158. Wang, J. et al. Crystal structure of a pol alpha family replication DNA polymerase from bacteriophage RB69. *Cell* **89**, 1087-99 (1997).
159. Kim, Y. et al. Crystal structure of Thermus aquaticus DNA polymerase. *Nature* **376**, 612-6 (1995).
160. Kim, S.W., Kim, D.U., Kim, J.K., Kang, L.W. & Cho, H.S. Crystal structure of Pfu, the high fidelity DNA polymerase from Pyrococcus furiosus. *Int J Biol Macromol* **42**, 356-61 (2008).
161. Wynne, S.A., Pinheiro, V.B., Holliger, P. & Leslie, A.G. Structures of an apo and a binary complex of an evolved archeal B family DNA polymerase capable of synthesising highly cy-dye labelled DNA. *PLoS One* **8**, e70892 (2013).
162. Rechkoblit, O. et al. Stepwise translocation of Dpo4 polymerase during error-free bypass of an oxoG lesion. *PLoS Biol* **4**, e11 (2006).
163. Brenlla, A., Markiewicz, R.P., Rueda, D. & Romano, L.J. Nucleotide selection by the Y-family DNA polymerase Dpo4 involves template translocation and misalignment. *Nucleic Acids Res* **42**, 2555-63 (2014).
164. Johnson, S.J., Taylor, J.S. & Beese, L.S. Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. *Proc Natl Acad Sci U S A* **100**, 3895-900 (2003).

165. Basu, R.S. & Murakami, K.S. Watching the bacteriophage N4 RNA polymerase transcription by time-dependent soak-trigger-freeze X-ray crystallography. *J Biol Chem* **288**, 3305-11 (2013).
166. Bakhtina, M., Roettger, M.P., Kumar, S. & Tsai, M.D. A unified kinetic mechanism applicable to multiple DNA polymerases. *Biochemistry* **46**, 5463-72 (2007).
167. Wang, M., Lee, H.R. & Konigsberg, W. Effect of A and B metal ion site occupancy on conformational changes in an RB69 DNA polymerase ternary complex. *Biochemistry* **48**, 2075-86 (2009).
168. Bakhtina, M. et al. Use of viscogens, dNTPalphaS, and rhodium(III) as probes in stopped-flow experiments to obtain new evidence for the mechanism of catalysis by DNA polymerase beta. *Biochemistry* **44**, 5177-87 (2005).
169. Johnson, A.A., Tsai, Y., Graves, S.W. & Johnson, K.A. Human mitochondrial DNA polymerase holoenzyme: reconstitution and characterization. *Biochemistry* **39**, 1702-8 (2000).
170. Graves, S.W., Johnson, A.A. & Johnson, K.A. Expression, purification, and initial kinetic characterization of the large subunit of the human mitochondrial DNA polymerase. *Biochemistry* **37**, 6050-8 (1998).
171. O'Donnell, M.E. & Kornberg, A. Complete replication of templates by Escherichia coli DNA polymerase III holoenzyme. *J Biol Chem* **260**, 12884-9 (1985).
172. Zahurancik, W.J., Klein, S.J. & Suo, Z. Significant contribution of the 3'-->5' exonuclease activity to the high fidelity of nucleotide incorporation catalyzed by human DNA polymerase. *Nucleic Acids Res* **42**, 13853-60 (2014).
173. Zhang, L., Brown, J.A., Newmister, S.A. & Suo, Z. Polymerization fidelity of a replicative DNA polymerase from the hyperthermophilic archaeon Sulfolobus solfataricus P2. *Biochemistry* **48**, 7492-501 (2009).
174. Joyce, C.M. How DNA travels between the separate polymerase and 3'-5'-exonuclease sites of DNA polymerase I (Klenow fragment). *J Biol Chem* **264**, 10858-66 (1989).
175. Reha-Krantz, L.J. DNA polymerase proofreading: Multiple roles maintain genome stability. *Biochim Biophys Acta* **1804**, 1049-63 (2010).
176. Fernandez-Leiro, R. et al. Self-correcting mismatches during high-fidelity DNA replication. *Nat Struct Mol Biol* **24**, 140-143 (2017).
177. Ganai, R.A., Bylund, G.O. & Johansson, E. Switching between polymerase and exonuclease sites in DNA polymerase epsilon. *Nucleic Acids Res* **43**, 932-42 (2015).
178. Johnson, A.A. & Johnson, K.A. Exonuclease proofreading by human mitochondrial DNA polymerase. *J Biol Chem* **276**, 38097-107 (2001).
179. Johnson, A.A. & Johnson, K.A. Fidelity of nucleotide incorporation by human mitochondrial DNA polymerase. *J Biol Chem* **276**, 38090-6 (2001).
180. Johnson, S.J. & Beese, L.S. Structures of mismatch replication errors observed in a DNA polymerase. *Cell* **116**, 803-16 (2004).

181. Xia, S. & Konigsberg, W.H. Mispairs with Watson-Crick base-pair geometry observed in ternary complexes of an RB69 DNA polymerase variant. *Protein Sci* **23**, 508-13 (2014).
182. Trincao, J. et al. Dpo4 is hindered in extending a G.T mismatch by a reverse wobble. *Nat Struct Mol Biol* **11**, 457-62 (2004).
183. Picher, A.J. et al. Promiscuous mismatch extension by human DNA polymerase lambda. *Nucleic Acids Res* **34**, 3259-66 (2006).
184. Batra, V.K., Beard, W.A., Pedersen, L.C. & Wilson, S.H. Structures of DNA Polymerase Mispaired DNA Termini Transitioning to Pre-catalytic Complexes Support an Induced-Fit Fidelity Mechanism. *Structure* **24**, 1863-1875 (2016).
185. McCulloch, S.D. & Kunkel, T.A. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res* **18**, 148-61 (2008).
186. Osheroff, W.P., Jung, H.K., Beard, W.A., Wilson, S.H. & Kunkel, T.A. The fidelity of DNA polymerase beta during distributive and processive DNA synthesis. *J Biol Chem* **274**, 3642-50 (1999).
187. Bebenek, K., Garcia-Diaz, M., Blanco, L. & Kunkel, T.A. The frameshift infidelity of human DNA polymerase lambda. Implications for function. *J Biol Chem* **278**, 34685-90 (2003).
188. Johnson, R.E., Washington, M.T., Prakash, S. & Prakash, L. Fidelity of human DNA polymerase eta. *J Biol Chem* **275**, 7447-50 (2000).
189. Watson, J.D. & Crick, F.H. Genetical implications of the structure of deoxyribonucleic acid. *Nature* **171**, 964-7 (1953).
190. Loeb, L.A. & Kunkel, T.A. Fidelity of DNA synthesis. *Annu Rev Biochem* **51**, 429-57 (1982).
191. Kool, E.T. & Sintim, H.O. The difluorotoluene debate--a decade later. *Chem Commun (Camb)* **0**, 3665-75 (2006).
192. Kool, E.T. Hydrogen bonding, base stacking, and steric effects in dna replication. *Annu Rev Biophys Biomol Struct* **30**, 1-22 (2001).
193. Dzantiev, L., Alekseyev, Y.O., Morales, J.C., Kool, E.T. & Romano, L.J. Significance of nucleobase shape complementarity and hydrogen bonding in the formation and stability of the closed polymerase-DNA complex. *Biochemistry* **40**, 3215-21 (2001).
194. Lee, H.R., Helquist, S.A., Kool, E.T. & Johnson, K.A. Importance of hydrogen bonding for efficiency and specificity of the human mitochondrial DNA polymerase. *J Biol Chem* **283**, 14402-10 (2008).
195. Xia, S., Eom, S.H., Konigsberg, W.H. & Wang, J. Structural basis for differential insertion kinetics of dNMPs opposite a difluorotoluene nucleotide residue. *Biochemistry* **51**, 1476-85 (2012).
196. Xia, S. & Konigsberg, W.H. RB69 DNA polymerase structure, kinetics, and fidelity. *Biochemistry* **53**, 2752-67 (2014).
197. Kunkel, T.A. & Bebenek, K. DNA replication fidelity. *Annu Rev Biochem* **69**, 497-529 (2000).

198. Nat. Commun.CRC Critical Reviews in BiochemistryWu, W.-J., Yang, W. & Tsai, M.-D. How DNA polymerases catalyse replication and repair with contrasting fidelity. *Nat. Rev. Chem.* **1**, 0068 (2017).
199. Fiala, K.A., Duym, W.W., Zhang, J. & Suo, Z. Up-regulation of the fidelity of human DNA polymerase lambda by its non-enzymatic proline-rich domain. *J Biol Chem* **281**, 19038-44 (2006).
200. Qin, Y. et al. Direct probing of solvent accessibility and mobility at the binding interface of polymerase (Dpo4)-DNA complex. *J Phys Chem A* **117**, 13926-34 (2013).
201. Wang, W., Hellinga, H.W. & Beese, L.S. Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis. *Proc Natl Acad Sci U S A* **108**, 17644-8 (2011).
202. Bebenek, K., Pedersen, L.C. & Kunkel, T.A. Replication infidelity via a mismatch with Watson-Crick geometry. *Proc Natl Acad Sci U S A* **108**, 1862-7 (2011).
203. Szymanski, E.S., Kimsey, I.J. & Al-Hashimi, H.M. Direct NMR Evidence that Transient Tautomeric and Anionic States in dG.dT Form Watson-Crick-like Base Pairs. *J Am Chem Soc* **139**, 4326-4329 (2017).
204. Kimsey, I.J. et al. Dynamic basis for dG*dT misincorporation via tautomerization and ionization. *Nature* **554**, 195-201 (2018).
205. Batra, V.K., Beard, W.A., Shock, D.D., Pedersen, L.C. & Wilson, S.H. Structures of DNA polymerase beta with active-site mismatches suggest a transient abasic site intermediate during misincorporation. *Mol Cell* **30**, 315-24 (2008).
206. Minnick, D.T., Astatke, M., Joyce, C.M. & Kunkel, T.A. A thumb subdomain mutant of the large fragment of Escherichia coli DNA polymerase I with reduced DNA binding affinity, processivity, and frameshift fidelity. *J Biol Chem* **271**, 24954-61 (1996).
207. Kiefer, J.R., Mao, C., Braman, J.C. & Beese, L.S. Visualizing DNA replication in a catalytically active Bacillus DNA polymerase crystal. *Nature* **391**, 304-7 (1998).
208. Freudenthal, B.D., Beard, W.A. & Wilson, S.H. DNA polymerase minor groove interactions modulate mutagenic bypass of a templating 8-oxoguanine lesion. *Nucleic Acids Res* **41**, 1848-58 (2013).
209. Spence, R.A., Kati, W.M., Anderson, K.S. & Johnson, K.A. Mechanism of inhibition of HIV-1 reverse transcriptase by nonnucleoside inhibitors. *Science* **267**, 988-93 (1995).
210. Doublie, S., Sawaya, M.R. & Ellenberger, T. An open and closed case for all polymerases. *Structure* **7**, R31-5 (1999).
211. Herschlag, D., Piccirilli, J.A. & Cech, T.R. Ribozyme-catalyzed and nonenzymatic reactions of phosphate diesters: rate effects upon substitution of sulfur for a nonbridging phosphoryl oxygen atom. *Biochemistry* **30**, 4844-54 (1991).
212. Arndt, J.W. et al. Insight into the catalytic mechanism of DNA polymerase beta: structures of intermediate complexes. *Biochemistry* **40**, 5368-75 (2001).

213. Luo, G., Wang, M., Konigsberg, W.H. & Xie, X.S. Single-molecule and ensemble fluorescence assays for a functionally important conformational change in T7 DNA polymerase. *Proc Natl Acad Sci U S A* **104**, 12610-5 (2007).
214. Lee, H.R., Wang, M. & Konigsberg, W. The reopening rate of the fingers domain is a determinant of base selectivity for RB69 DNA polymerase. *Biochemistry* **48**, 2087-98 (2009).
215. Kirmizialtin, S., Nguyen, V., Johnson, K.A. & Elber, R. How conformational dynamics of DNA polymerase select correct substrates: experiments and simulations. *Structure* **20**, 618-27 (2012).
216. Kellinger, M.W. & Johnson, K.A. Nucleotide-dependent conformational change governs specificity and analog discrimination by HIV reverse transcriptase. *Proc Natl Acad Sci U S A* **107**, 7734-9 (2010).
217. Kellinger, M.W. & Johnson, K.A. Role of induced fit in limiting discrimination against AZT by HIV reverse transcriptase. *Biochemistry* **50**, 5008-15 (2011).
218. Yang, W. An overview of Y-Family DNA polymerases and a case study of human DNA polymerase eta. *Biochemistry* **53**, 2793-803 (2014).
219. Beckman, J.W., Wang, Q. & Guengerich, F.P. Kinetic analysis of correct nucleotide insertion by a Y-family DNA polymerase reveals conformational changes both prior to and following phosphodiester bond formation as detected by tryptophan fluorescence. *J Biol Chem* **283**, 36711-23 (2008).
220. Fiala, K.A., Sherrer, S.M., Brown, J.A. & Suo, Z. Mechanistic consequences of temperature on DNA polymerization catalyzed by a Y-family DNA polymerase. *Nucleic Acids Res* **36**, 1990-2001 (2008).
221. Eoff, R.L., Sanchez-Ponce, R. & Guengerich, F.P. Conformational changes during nucleotide selection by Sulfolobus solfataricus DNA polymerase Dpo4. *J Biol Chem* **284**, 21090-9 (2009).
222. Vyas, R. et al. Structural basis for the D-stereoselectivity of human DNA polymerase beta. *Nucleic Acids Res* **45**, 6228-6237 (2017).
223. Vyas, R., Zahurancik, W.J. & Suo, Z. Structural basis for the binding and incorporation of nucleotide analogs with L-stereochemistry by human DNA polymerase lambda. *Proc Natl Acad Sci U S A* **111**, E3033-42 (2014).
224. Shock, D.D., Freudenthal, B.D., Beard, W.A. & Wilson, S.H. Modulating the DNA polymerase beta reaction equilibrium to dissect the reverse reaction. *Nat Chem Biol* **13**, 1074-1080 (2017).
225. Raper, A.T., Reed, A.J., Gadkari, V.V. & Suo, Z. Advances in Structural and Single-Molecule Methods for Investigating DNA Lesion Bypass and Repair Polymerases. *Chem Res Toxicol* **30**, 260-269 (2017).
226. Brown, J.A., Duym, W.W., Fowler, J.D. & Suo, Z. Single-turnover kinetic analysis of the mutagenic potential of 8-oxo-7,8-dihydro-2'-deoxyguanosine during gap-filling synthesis catalyzed by human DNA polymerases lambda and beta. *J Mol Biol* **367**, 1258-69 (2007).
227. Bedford, E., Tabor, S. & Richardson, C.C. The thioredoxin binding domain of bacteriophage T7 DNA polymerase confers processivity on Escherichia coli DNA polymerase I. *Proc Natl Acad Sci U S A* **94**, 479-84 (1997).

228. Beard, W.A. & Wilson, S.H. Structural insights into the origins of DNA polymerase fidelity. *Structure* **11**, 489-96 (2003).
229. Ohmori, H. et al. The Y-family of DNA polymerases. *Mol Cell* **8**, 7-8 (2001).
230. Lehmann, A.R. et al. Translesion synthesis: Y-family polymerases and the polymerase switch. *DNA Repair (Amst)* **6**, 891-9 (2007).
231. Matsuda, T., Bebenek, K., Masutani, C., Hanaoka, F. & Kunkel, T.A. Low fidelity DNA synthesis by human DNA polymerase-eta. *Nature* **404**, 1011-3 (2000).
232. Pecoraro, V.L., Hermes, J.D. & Cleland, W.W. Stability constants of Mg2+ and Cd2+ complexes of adenine nucleotides and thionucleotides and rate constants for formation and dissociation of MgATP and MgADP. *Biochemistry* **23**, 5262-71 (1984).
233. Roettger, M.P., Fiala, K.A., Sompalli, S., Dong, Y. & Suo, Z. Pre-steady-state kinetic studies of the fidelity of human DNA polymerase mu. *Biochemistry* **43**, 13827-38 (2004).
234. Oertell, K. et al. Kinetic selection vs. free energy of DNA base pairing in control of polymerase fidelity. *Proc Natl Acad Sci U S A* **113**, E2277-85 (2016).
235. Hanes, J.W. & Johnson, K.A. A novel mechanism of selectivity against AZT by the human mitochondrial DNA polymerase. *Nucleic Acids Res* **35**, 6973-83 (2007).
236. Nami, F., Gast, P. & Groenen, E.J. Rapid Freeze-Quench EPR Spectroscopy: Improved Collection of Frozen Particles. *Appl Magn Reson* **47**, 643-653 (2016).
237. Hubscher, U., Maga, G. & Spadari, S. Eukaryotic DNA polymerases. *Annu Rev Biochem* **71**, 133-63 (2002).
238. Savino, C. et al. Insights into DNA replication: the crystal structure of DNA polymerase B1 from the archaeon Sulfolobus solfataricus. *Structure* **12**, 2001-8 (2004).
239. Steitz, T.A., Smerdon, S.J., Jager, J. & Joyce, C.M. A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science* **266**, 2022-5 (1994).
240. Vyas, R. et al. Mechanistic Basis for the Bypass of a Bulky DNA Adduct Catalyzed by a Y-Family DNA Polymerase. *J Am Chem Soc* **137**, 12131-42 (2015).
241. Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nat Methods* **5**, 507-16 (2008).
242. Boudsocq, F., Iwai, S., Hanaoka, F. & Woodgate, R. Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4): an archaeal DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic poleta. *Nucleic Acids Res* **29**, 4607-16 (2001).
243. Oda, Y. et al. NMR studies of a DNA containing 8-hydroxydeoxyguanosine. *Nucleic Acids Res* **19**, 1407-12 (1991).
244. McAuley-Hecht, K.E. et al. Crystal structure of a DNA duplex containing 8-hydroxydeoxyguanine-adenine base pairs. *Biochemistry* **33**, 10266-70 (1994).

245. Lipscomb, L.A. et al. X-ray structure of a DNA decamer containing 7,8-dihydro-8-oxoguanine. *Proc Natl Acad Sci U S A* **92**, 719-23 (1995).
246. Beard, W.A., Batra, V.K. & Wilson, S.H. DNA polymerase structure-based insight on the mutagenic properties of 8-oxoguanine. *Mutat Res* **703**, 18-23 (2010).
247. Krahn, J.M., Beard, W.A., Miller, H., Grollman, A.P. & Wilson, S.H. Structure of DNA polymerase beta with the mutagenic DNA lesion 8-oxodeoxyguanine reveals structural insights into its coding potential. *Structure* **11**, 121-7 (2003).
248. Brieba, L.G. et al. Structural basis for the dual coding potential of 8-oxoguanosine by a high-fidelity DNA polymerase. *Embo J* **23**, 3452-61 (2004).
249. Brieba, L.G., Kokoska, R.J., Bebenek, K., Kunkel, T.A. & Ellenberger, T. A lysine residue in the fingers subdomain of T7 DNA polymerase modulates the miscoding potential of 8-oxo-7,8-dihydroguanosine. *Structure* **13**, 1653-9 (2005).
250. Eoff, R.L., Irimia, A., Angel, K.C., Egli, M. & Guengerich, F.P. Hydrogen bonding of 7,8-dihydro-8-oxodeoxyguanosine with a charged residue in the little finger domain determines miscoding events in Sulfolobus solfataricus DNA polymerase Dpo4. *J Biol Chem* **282**, 19831-43 (2007).
251. Rechkoblit, O. et al. Impact of conformational heterogeneity of OxoG lesions and their pairing partners on bypass fidelity by Y family polymerases. *Structure* **17**, 725-36 (2009).
252. Espinoza-Herrera, S.J., Gaur, V., Suo, Z. & Carey, P.R. Following DNA Chain Extension and Protein Conformational Changes in Crystals of a Y-Family DNA Polymerase via Raman Crystallography. *Biochemistry* **52**, 4881−4890 (2013).
253. Tokarsky, E.J. et al. Pre-steady-state kinetic investigation of bypass of a bulky guanine lesion by human Y-family DNA polymerases. *DNA Repair (Amst)* **46**, 20-28 (2016).
254. Kalam, M.A. & Basu, A.K. Mutagenesis of 8-oxoguanine adjacent to an abasic site in simian kidney cells: tandem mutations and enhancement of G-->T transversions. *Chem Res Toxicol* **18**, 1187-92 (2005).
255. Ohno, M. et al. 8-oxoguanine causes spontaneous de novo germline mutations in mice. *Sci Rep* **4**, 4689 (2014).
256. Fraga, C.G., Shigenaga, M.K., Park, J.W., Degan, P. & Ames, B.N. Oxidative damage to DNA during aging: 8-hydroxy-2'-deoxyguanosine in rat organ DNA and urine. *Proc Natl Acad Sci U S A* **87**, 4533-7 (1990).
257. Moriya, M. Single-stranded shuttle phagemid for mutagenesis studies in mammalian cells: 8-oxoguanine in DNA induces targeted G.C-->T.A transversions in simian kidney cells. *Proc Natl Acad Sci U S A* **90**, 1122-6 (1993).
258. Steitz, T.A. DNA- and RNA-dependent DNA polymerases. *Curr Opin Struct Biol* **3**, 31-38 (1993).
259. Freudenthal, B.D., Beard, W.A. & Wilson, S.H. Watching a DNA polymerase in action. *Cell Cycle* **13**, 691-2 (2014).
260. Ha, T. Single-molecule fluorescence resonance energy transfer. *Methods* **25**, 78-86 (2001).

261. Berezhna, S.Y., Gill, J.P., Lamichhane, R. & Millar, D.P. Single-molecule Forster resonance energy transfer reveals an innate fidelity checkpoint in DNA polymerase I. *J Am Chem Soc* **134**, 11261-8 (2012).
262. Christian, T.D., Romano, L.J. & Rueda, D. Single-molecule measurements of synthesis by DNA polymerase with base-pair resolution. *Proc Natl Acad Sci U S A* **106**, 21109-14 (2009).
263. Markiewicz, R.P., Vrtis, K.B., Rueda, D. & Romano, L.J. Single-molecule microscopy reveals new insights into nucleotide selection by DNA polymerase I. *Nucleic Acids Res* **40**, 7975-84 (2012).
264. Santoso, Y. et al. Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *Proc Natl Acad Sci U S A* **107**, 715-20 (2010).
265. Vrtis, K.B., Markiewicz, R.P., Romano, L.J. & Rueda, D. Carcinogenic adducts induce distinct DNA polymerase binding orientations. *Nucleic Acids Res* **41**, 7843-53 (2013).
266. Brenlla, A., Rueda, D. & Romano, L.J. Mechanism of aromatic amine carcinogen bypass by the Y-family polymerase, Dpo4. *Nucleic Acids Res* **43**, 9918-27 (2015).
267. Raper, A.T. & Suo, Z. Investigation of Intradomain Motions of a Y-Family DNA Polymerase during Substrate Binding and Catalysis. *Biochemistry* (2016).
268. Lieberman, K.R., Dahl, J.M., Mai, A.H., Akeson, M. & Wang, H. Dynamics of the translocation step measured in individual DNA polymerase complexes. *J Am Chem Soc* **134**, 18816-23 (2012).
269. Lieberman, K.R. et al. Kinetic mechanism of translocation and dNTP binding in individual DNA polymerase complexes. *J Am Chem Soc* **135**, 9149-55 (2013).
270. Dahl, J.M. et al. Direct observation of translocation in individual DNA polymerase complexes. *J Biol Chem* **287**, 13407-21 (2012).
271. Lovett, S.T. Polymerase switching in DNA replication. *Mol Cell* **27**, 523-6 (2007).
272. Friedberg, E.C., Lehmann, A.R. & Fuchs, R.P. Trading places: how do DNA polymerases switch during translesion DNA synthesis? *Mol Cell* **18**, 499-505 (2005).
273. Jansen, J.G., Fousteri, M.I. & de Wind, N. Send in the clamps: control of DNA translesion synthesis in eukaryotes. *Mol Cell* **28**, 522-9 (2007).
274. Pages, V. & Fuchs, R.P. How DNA lesions are turned into mutations within cells? *Oncogene* **21**, 8957-66 (2002).
275. Becherel, O.J., Fuchs, R.P. & Wagner, J. Pivotal role of the beta-clamp in translesion DNA synthesis and mutagenesis in E. coli cells. *DNA Repair (Amst)* **1**, 703-8 (2002).
276. Indiani, C., McInerney, P., Georgescu, R., Goodman, M.F. & O'Donnell, M. A sliding-clamp toolbelt binds high- and low-fidelity DNA polymerases simultaneously. *Mol Cell* **19**, 805-15 (2005).
277. Cooke, M.S., Evans, M.D., Dizdaroglu, M. & Lunec, J. Oxidative DNA damage: mechanisms, mutation, and disease. *Faseb J* **17**, 1195-214 (2003).

278. Asami, S. et al. Increase of a type of oxidative DNA damage, 8-hydroxyguanine, and its repair activity in human leukocytes by cigarette smoking. *Cancer Res* **56**, 2546-9 (1996).
279. Nakabeppu, Y., Kajitani, K., Sakamoto, K., Yamaguchi, H. & Tsuchimoto, D. MTH1, an oxidized purine nucleoside triphosphatase, prevents the cytotoxicity and neurotoxicity of oxidized purine nucleotides. *DNA Repair (Amst)* **5**, 761-72 (2006).
280. Cabelof, D.C., Raffoul, J.J., Yanamadala, S., Guo, Z. & Heydari, A.R. Induction of DNA polymerase beta-dependent base excision repair in response to oxidative stress in vivo. *Carcinogenesis* **23**, 1419-25 (2002).
281. Batra, V.K., Shock, D.D., Beard, W.A., McKenna, C.E. & Wilson, S.H. Binary complex crystal structure of DNA polymerase beta reveals multiple conformations of the templating 8-oxoguanine lesion. *Proc Natl Acad Sci U S A* **109**, 113-8 (2012).
282. Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R. & Leslie, A.G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **67**, 271-81 (2011).
283. McCoy, A.J. et al. Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).
284. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-21 (2010).
285. Murshudov, G.N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67**, 355-67 (2011).
286. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32 (2004).
287. Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R. & Thornton, J.M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**, 477-86 (1996).
288. DeLano, W.L. The PyMOL Molecular Graphics System. . (DeLano Scientific, San Carlos, CA, USA., 2002).
289. Rodriguez, A.C., Park, H.W., Mao, C. & Beese, L.S. Crystal structure of a pol alpha family DNA polymerase from the hyperthermophilic archaeon Thermococcus sp. 9 degrees N-7. *J Mol Biol* **299**, 447-62 (2000).
290. Garcia-Diaz, M., Bebenek, K., Krahn, J.M., Pedersen, L.C. & Kunkel, T.A. Role of the catalytic metal during polymerization by DNA polymerase lambda. *DNA Repair (Amst)* **6**, 1333-40 (2007).
291. Ding, J. et al. Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA template-primer and an antibody Fab fragment at 2.8 A resolution. *J Mol Biol* **284**, 1095-111 (1998).
292. Vaisman, A., Ling, H., Woodgate, R. & Yang, W. Fidelity of Dpo4: effect of metal ions, nucleotide selection and pyrophosphorolysis. *EMBO J* **24**, 2957-67 (2005).

293. Alt, A. et al. Bypass of DNA lesions generated during anticancer treatment with cisplatin by DNA polymerase eta. *Science* **318**, 967-70 (2007).
294. Silverstein, T.D. et al. Structural basis for error-free replication of oxidatively damaged DNA by yeast DNA polymerase eta. *Structure* **18**, 1463-70 (2010).
295. Freudenthal, B.D. et al. Uncovering the polymerase-induced cytotoxicity of an oxidized nucleotide. *Nature* (2014).
296. Towle-Weicksel, J.B. et al. Fluorescence resonance energy transfer studies of DNA polymerase beta: the critical role of fingers domain movements and a novel non-covalent step during nucleotide selection. *J Biol Chem* **289**, 16541-50 (2014).
297. David, S.S., O'Shea, V.L. & Kundu, S. Base-excision repair of oxidative DNA damage. *Nature* **447**, 941-50 (2007).
298. Dianov, G.L. & Hubscher, U. Mammalian base excision repair: the forgotten archangel. *Nucleic Acids Res* **41**, 3483-90 (2013).
299. Hashimoto, K., Tominaga, Y., Nakabeppu, Y. & Moriya, M. Futile short-patch DNA base excision repair of adenine:8-oxoguanine mispair. *Nucleic Acids Res* **32**, 5928-34 (2004).
300. Klungland, A. & Lindahl, T. Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *Embo J* **16**, 3341-8 (1997).
301. Frosina, G. et al. Two pathways for base excision repair in mammalian cells. *J Biol Chem* **271**, 9573-8 (1996).
302. Wilson, S.H. Mammalian base excision repair and DNA polymerase beta. *Mutat Res* **407**, 203-15 (1998).
303. Petermann, E., Ziegler, M. & Oei, S.L. ATP-dependent selection between single nucleotide and long patch base excision repair. *DNA Repair (Amst)* **2**, 1101-14 (2003).
304. Johnson, K.A. Rapid kinetic analysis of mechanochemical adenosinetriphosphatases. *Methods Enzymol* **134**, 677-705 (1986).
305. Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-32 (2010).
306. Lee, H.R., Helquist, S.A., Kool, E.T. & Johnson, K.A. Base pair hydrogen bonds are essential for proofreading selectivity by the human mitochondrial DNA polymerase. *J Biol Chem* **283**, 14411-6 (2008).
307. Washington, M.T., Helquist, S.A., Kool, E.T., Prakash, L. & Prakash, S. Requirement of Watson-Crick hydrogen bonding for DNA synthesis by yeast DNA polymerase eta. *Mol Cell Biol* **23**, 5107-12 (2003).
308. Bebenek, K. et al. A minor groove binding track in reverse transcriptase. *Nat Struct Biol* **4**, 194-7 (1997).
309. Feng, J.Y. & Anderson, K.S. Mechanistic studies comparing the incorporation of (+) and (-) isomers of 3TCTP by HIV-1 reverse transcriptase. *Biochemistry* **38**, 55-63 (1999).
310. Feng, J.Y., Shi, J., Schinazi, R.F. & Anderson, K.S. Mechanistic studies show that (-)-FTC-TP is a better inhibitor of HIV-1 reverse transcriptase than 3TC-TP. *FASEB J* **13**, 1511-7 (1999).

311. Gumina, G., Chong, Y., Choo, H., Song, G.Y. & Chu, C.K. L-nucleosides: antiviral activity and molecular mechanism. *Curr Top Med Chem* **2**, 1065-86 (2002).
312. Chang, C.N., Skalski, V., Zhou, J.H. & Cheng, Y.C. Biochemical pharmacology of (+)- and (-)-2',3'-dideoxy-3'-thiacytidine as anti-hepatitis B virus agents. *J Biol Chem* **267**, 22414-20 (1992).
313. Schinazi, R.F. et al. Selective inhibition of human immunodeficiency viruses by racemates and enantiomers of cis-5-fluoro-1-[2-(hydroxymethyl)-1,3-oxathiolan-5-yl]cytosine. *Antimicrob Agents Chemother* **36**, 2423-31 (1992).
314. Moyle, G. Toxicity of antiretroviral nucleoside and nucleotide analogues: is mitochondrial toxicity the only mechanism? *Drug Saf* **23**, 467-81 (2000).
315. Olivero, O.A. et al. Incorporation of zidovudine into leukocyte DNA from HIV-1-positive adults and pregnant women, and cord blood from infants exposed in utero. *AIDS* **13**, 919-25 (1999).
316. Wutzler, P. & Thust, R. Genetic risks of antiviral nucleoside analogues--a survey. *Antiviral Res* **49**, 55-74 (2001).
317. Johnson, A.A. et al. Toxicity of antiviral nucleoside analogs and the human mitochondrial DNA polymerase. *J Biol Chem* **276**, 40847-57 (2001).
318. Feng, J.Y. et al. Relationship between antiviral activity and host toxicity: comparison of the incorporation efficiencies of 2',3'-dideoxy-5-fluoro-3'-thiacytidine-triphosphate analogs by human immunodeficiency virus type 1 reverse transcriptase and human mitochondrial DNA polymerase. *Antimicrob Agents Chemother* **48**, 1300-6 (2004).
319. Brown, J.A., Pack, L.R., Fowler, J.D. & Suo, Z. Pre-steady-state kinetic analysis of the incorporation of anti-HIV nucleotide analogs catalyzed by human X- and Y-family DNA polymerases. *Antimicrob Agents Chemother* **55**, 276-83 (2011).
320. Brown, J.A., Pack, L.R., Fowler, J.D. & Suo, Z. Presteady state kinetic investigation of the incorporation of anti-hepatitis B nucleotide analogues catalyzed by noncanonical human DNA polymerases. *Chem Res Toxicol* **25**, 225-33 (2012).
321. Sohl, C.D. et al. Probing the structural and molecular basis of nucleotide selectivity by human mitochondrial DNA polymerase gamma. *Proc Natl Acad Sci U S A* **112**, 8596-601 (2015).
322. DeLucia, A.M., Chaudhuri, S., Potapova, O., Grindley, N.D. & Joyce, C.M. The properties of steric gate mutants reveal different constraints within the active sites of Y-family and A-family DNA polymerases. *J Biol Chem* **281**, 27286-91 (2006).
323. Wang, W., Wu, E.Y., Hellinga, H.W. & Beese, L.S. Structural factors that determine selectivity of a high fidelity DNA polymerase for deoxy-, dideoxy-, and ribonucleotides. *J Biol Chem* **287**, 28215-26 (2012).
324. Moscato, B., Swain, M. & Loria, J.P. Induced Fit in the Selection of Correct versus Incorrect Nucleotides by DNA Polymerase beta. *Biochemistry* **55**, 382-95 (2016).

325. Evans, G.W., Hohlbein, J., Craggs, T., Aigrain, L. & Kapanidis, A.N. Real-time single-molecule studies of the motions of DNA polymerase fingers illuminate DNA synthesis mechanisms. *Nucleic Acids Res* **43**, 5998-6008 (2015).
326. Krahn, J.M., Beard, W.A. & Wilson, S.H. Structural insights into DNA polymerase beta deterrents for misincorporation support an induced-fit mechanism for fidelity. *Structure* **12**, 1823-32 (2004).
327. Beard, W.A., Shock, D.D., Batra, V.K., Prasad, R. & Wilson, S.H. Substrate-induced DNA polymerase beta activation. *J Biol Chem* **289**, 31411-22 (2014).