

Investigating Current Mechanistic Models of DNA Replication and Repair

THESIS

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in  
the Graduate School of The Ohio State University

By

Petra C Wallenmeyer

Graduate Program in Chemistry

The Ohio State University

2017

Master's Examination Committee:

Dr. Zucai Suo, Advisor

Dr. James A. Cowan

Dr. Thomas J. Magliery

Copyrighted by  
Petra Cheyenne Wallenmeyer  
2017

## Abstract

DNA polymerases (pols) play a pivotal role in both the replication and the repair of genomic DNA. Replicative pols are highly accurate and processive, synthesizing long stretches of DNA in a single binding event, while repair and bypass pols are error-prone and only able to ~~synthesize~~ incorporate a few nucleotides before dissociation. During replication, the pol may encounter DNA modifications ~~induced~~ caused by endogenous and exogenous factors such as oxidative metabolites, UV radiation, or epigenetic additions. These modifications may alter the local structure of DNA, resulting in inhibition of the replicative pol and stalling of the replication machinery. When the replicative pol stalls, a repair or bypass pol can take over and perform translesion synthesis (TLS). During TLS, a nucleotide is inserted opposite a ~~lesion-containing base~~ lesion on the template DNA strand before the replicative pol can continue DNA synthesis. The mechanistic details of DNA replication, bypass, and repair are areas of ongoing research and are important to other areas of research such as drug design, cancer research, metabolism, and aging.

The overarching goal of my research was to contribute to the mechanistic understanding of how pols perform DNA synthesis and bypass of DNA lesions. With this goal in mind, one of my main projects was to investigate the bypass kinetics of a common epigenetic signal, modification of the C5-position on cytosine (5xC). I used a specialized pol, human pol  $\iota$ , to conduct this investigation. Using pre-steady state kinetic

methods, I determined the dissociation constant ( $K_d$ ) and maximum incorporation rate ( $k_{pol}$ ) for each deoxynucleoside (dNTP) opposite each 5xC modification. I also attempted to determine the structural details of the two accessory subunits of human pol  $\epsilon$ , a replicative DNA pol that carries out leading strand synthesis, via X-ray crystallography to propose a hypothesis for its intrinsically high fidelity and processivity. Other projects I was involved in included investigating the biochemical properties of human PrimPol, the second pol found to operate in the mitochondria, using pre-steady state kinetics; using X-ray crystallography and pre-steady state kinetics to characterize human pol  $\beta$ ; and determining the best assay conditions to perform pre-steady state kinetics experiments with the Zika Virus NS5, an RNA-dependent RNA polymerase; which replicates viral RNA genome.

## Dedication

I would like to dedicate this document to the following people:

To those of you who feel like impostors:

You fill a unique space in the world.

To those of you who feel like you're not good enough:

Most standards are arbitrary—do not allow them to define you.

To those of you fighting to be respected simply as a human being:

Do not give up.

## Acknowledgments

I would like to acknowledge the many people whose help, both personal and professional, made this document and the journey leading up to it possible. I would like to thank my mother, Cindy Evers, for supporting me mentally, emotionally and physically for my entire academic career; for helping edit this document; and for teaching me by example the many characteristics I needed to embody to get this far: strength, wisdom, patience, confidence, grace, kindness, and respect. I would like to thank each of my former lab mates (and good friends) who helped me learn the basics of a biochemistry lab, helped me design experiments, and engaged me in scientific discussion: Varun Gadkari, Vimal Parkash, Austin Raper, Andrew Reed, Anthony Stephenson, Jack Tokarsky, and Walter Zahurancik, I would also like to acknowledge my PI, Dr. Suo, for accepting me in his lab, giving me interesting projects, and supporting me financially. Last, but not least, I would like to acknowledge The Ohio State University for accepting me into the Department of Chemistry and Biochemistry PhD program and providing me with funding and excellent academic and career opportunities.

## Vita

- March 1993 .....Born
- May 2009 .....Homeschool
- December 2011 .....Dean’s Honor Graduate, A.S., Johnson  
County Community College
- November 2013..... “ZnS nanoparticles as photocatalyst for  
pollution removal,” *Women in Physical  
Sciences Conference*, University of  
Nebraska-Lincoln
- May 2014 .....B.S. Chemistry, Magna Cum Laude,  
University of Missouri-Kansas City
- March 2017 .....Panel speaker, *Eat and Explore*, Biological  
Sciences Scholars Program, 2017, The Ohio  
State University
- August 2014-present .....Graduate Teaching Associate, Department  
of Chemistry and Biochemistry, The Ohio  
State University

## Awards and Honors

1. President's Honor Roll, Johnson County Community College, Fall 2010, Spring 2010, Spring and Fall 2011
2. Dean's Honor Roll, Johnson County Community College, Fall 2010 semester
3. Chancellors Transfer Scholarship, University of Missouri-Kansas City, 2012-2013
4. Curators Transfer Scholarship, University of Missouri-Kansas City, 2012-2013
5. Phi Theta Kappa Scholarship, 2012-2013
6. Dean's List, University of Missouri-Kansas City, 2013-2014
7. James and Neoma Kilway Scholarship, University of Missouri-Kansas City, 2013-2014
8. KL Cheng Physics Scholarship, University of Missouri-Kansas City, 2013-2014
9. LaVerne Noyes Scholarship, University of Missouri-Kansas City, 2013-2014
10. Great Lakes National Scholarship, \$2500, 2013-2014
11. Students Engaged in the Arts and Research (SEARCH) grant (\$1040), University of Missouri-Kansas City, "Preparation of nanocrystalline ZnS for investigation of performance in photocatalytic hydrogen generation and dye decomposition," 2013

## Publications

1. Y. Zhang, C. X. Harris, P. Wallenmeyer, J. Murowchick, X. Chen "Asymmetric lattice vibrational characteristics of rutile TiO<sub>2</sub> as revealed by laser power dependent

Raman spectroscopy" *Journal of Physical Chemistry*, 2013, (117): 24015 – 24022.

2. Y. Zhang, T. Xia, P. Wallenmeyer, C. X. Harris, A. A. Peterson, G. A. Corsiglia, J. Murowchick, X. Chen "Photocatalytic hydrogen generation from pure water with silicon carbide nanoparticles" *Energy Technology*, 2014, (2): 183 – 187.

3. Y. Zhang, T. Xia, M. Shang, P. Wallenmeyer, D. Katelyn, A. Peterson, J. Murowchick, L. Dong, X. Chen "Structural evolution from TiO<sub>2</sub> nanoparticles to nanosheets and their photocatalytic performance in hydrogen generation and environmental pollution removal" *RSC Advances* 2014, 4 (31): 16146 – 16152.

4. Y. Zhang, M. Shang, Y. Mi, T. Xia, P. Wallenmeyer, J. Murowchick, L. Dong, Q. Zhang, X. Chen "Influence of the amount of HF on the formation of (001) faceted TiO<sub>2</sub> nanosheets and their photocatalytic hydrogen generation performance" *ChemPlusChem* 2014, (79): 1159-1166.

5. T. Xia, P. Wallenmeyer, A. Anderson, J. Murowchick, L. Liu, X. Chen "Hydrogenated Black ZnO Nanoparticles with Enhanced Photocatalytic Performance" *RSC Advances* 2014, 4 (78): 41654 - 41658.

6. L. Tian, X. Yan, J. Xu, P. Wallenmeyer, J. Murowchick, L. Liu, X. Chen "Effect of hydrogenation on the microwave absorption properties of BaTiO<sub>3</sub>

nanoparticles” *Journal of Materials Chemistry A*, 2015, (3): 12550-12556

7. E. Tokarsky, P. Wallenmeyer, K. Phi, Z. Suo, “Significant impact of divalent metal ions on the fidelity, sugar selectivity, and drug incorporation efficiency of human PrimPol” *DNA Repair* 2017, (49): 51–59.

8. R. Vyas, A. Reed, A. Raper, W. Zahurancik, P. Wallenmeyer, Z. Suo “Structural basis for the D-stereoselectivity of human DNA polymerase  $\beta$ ” *Nucleic Acids Research*. 2017, 45 (10): 6228-6237.

#### Fields of Study

Major Field: Chemistry

## Table of Contents

Abstract .....	ii
Dedication .....	iv
Acknowledgments .....	v
Vita .....	vi
List of Tables .....	xiv
List of Figures .....	xv
Chapter 1: Introduction .....	1
1.1 DNA Replication and Repair .....	1
1.2 Types and Sources of DNA Damage .....	2
1.3 DNA Polymerase Structure and Mechanism .....	4
1.3.1 Structure of the Polymerase Domain .....	4
1.3.2 Mechanism of Nucleotide Incorporation .....	5
1.4 Replicative Polymerases .....	6
1.4.1 Polymerase $\alpha$ .....	6
1.4.2 Polymerase $\delta$ .....	7

1.4.3 Polymerase $\epsilon$ .....	7
1.5 Specialized Polymerases .....	8
1.5.1 X-family Polymerases .....	9
1.5.2 Y-family Polymerases .....	9
1.5.3 Other Specialized Polymerases .....	11
1.6 Significance and Impact .....	11
Chapter 2: Theory .....	14
2.1 Kinetics.....	14
2.2 X-ray Crystallography.....	18
2.2.1 Growing a Crystal.....	18
2.2.2 X-Ray Diffraction.....	19
Chapter 3: The Potential Role of Polymerase $\iota$ in Bypassing Epigenetic Modifications .	21
3.1 Introduction .....	21
3.2 Materials and Methods .....	23
3.2.1 Materials .....	23
3.2.2 Preparation of DNA Substrates: .....	24
3.2.3 Expression and Purification of Human Polymerase $\iota$ .....	24
3.2.4 Single-Turnover Polymerase Assays:.....	26
3.3 Results .....	27

3.4 Discussion .....	28
Chapter 4: Crystallization of p12/p17 Complex of Polymerase Epsilon for Determination of Structure.....	35
4.1 Introduction .....	35
4.2 Materials and Methods .....	37
4.2.1 Materials .....	37
4.2.2 Expression and Purification of Human Polymerase $\epsilon$ Subunits p12 and p17 ..	37
4.2.3 Crystallization.....	38
4.3 Results .....	39
4.4 Discussion .....	40
Chapter 5: The Effect of Divalent Metal Ions on the Activity of Human PrimPol .....	44
5.1 Individual Contribution.....	44
5.2 Introduction .....	44
5.3 Materials and Methods .....	47
5.3.1 Materials .....	47
5.3.2 Radiolabeling and annealing DNA substrates: .....	47
5.3.3 Expression and Purification of Human PrimPol.....	47
5.3.4 Single Turnover Polymerase Assays: .....	49
5.3.5 Fluorescence Anisotropy Assays .....	49

5.4 Results .....	50
5.5 Discussion .....	52
Chapter 6: Other Projects.....	58
6.1 D-Stereoselectivity of Human Polymerase $\beta$ .....	58
6.1.1 My Contribution .....	58
6.1.2 Materials and Methods .....	58
6.1.3 Results and Discussion .....	62
6.2 Zika Virus NS5 Polymerase .....	64
6.2.1 My Contribution .....	64
6.2.2 Introduction and Background .....	65
6.2.3 Materials and Methods .....	68
6.2.4 Results and Discussion .....	69
References.....	75

## List of Tables

Table 1: DNA Substrates for dNTP Incorporation by Polymerase $\iota$ Opposite Modified C .....	33
Table 2: Pre-Steady State Kinetic Parameters for dNTP Incorporation by Polymerase $\iota$ Opposite Modified C.....	34
Table 3: Conditions that Yielded Crystal Growth for p12/p17.....	43
Table 4: DNA Substrates for Pre-Steady Kinetics and DNA Binding Assays with PrimPol .....	55
Table 5: Pre-Steady State Kinetic Parameters for dNTP and rNTP Incorporation by PrimPol in the Presence of 5 mM $Mn^{2+}$ .....	56
Table 6: Pre-Steady State Kinetic Parameters for dNTP Incorporation by PrimPol in the Presence of 5 mM $Mg^{2+}$ .....	57
Table 7: Pre-Steady Kinetic Parameters for Cytosine Analog Incorporation by PrimPol in the Presence of $Mg^{2+}$ or $Mn^{2+}$ .....	57
Table 8: DNA Substrate for Pre-Steady State Kinetic Assay with Polymerase $\beta$ .....	63
Table 9: Pre-Steady State Kinetics of Nucleotide Incorporation by Polymerase $\beta$ .....	64
Table 10: RNA Substrate Used in Buffer Optimization Assays and EMSA with ZIKV NS5.....	74

## List of Figures

Figure 1: Architecture of the Catalytic Subunit of Polymerase $\epsilon$ .....	12
Figure 2: Scheme of Polymerase Activity .....	13
Figure 3: Assumptions for Steady- and Pre-Steady State Kinetics.....	20
Figure 4: Schematic Diagram of the Methylation and Demethylation Pathway of the C5 Position of Cytosine .....	31
Figure 5: G:X and T:X Basepairs .....	32
Figure 6: Pre-Steady State Kinetics of dGTP incorporation opposite 5caC.....	33
Figure 7: Unique P-Subdomain of Polymerase $\epsilon$ .....	41
Figure 8: Cryo-EM Structure of <i>Saccharomyces cerevisiae</i> Polymerase $\epsilon$ Holoenzyme .	42
Figure 9: Photos of p12/p17 Crystals.....	42
Figure 10: Buffer and Detergent Optimization Assays for ZIKV NS5 .....	71
Figure 11: Metal Ion Optimization Assay for ZIKV NS5 .....	72
Figure 12: Salt Optimization Assay for ZIKV NS5.....	73
Figure 13: EMSA for ZIKV NS5 with RNA Substrate .....	74

## Chapter 1: Introduction

### 1.1 DNA Replication and Repair

The propagation and sustaining of life depends upon the replication of DNA. The human genome contains 23 chromosomes which in turn contain a total of approximately 3 billion base pairs. This genetic information is contained within the nucleus of all human cells and a highly-orchestrated process takes place to copy this information when needed, such as when a cell divides or a gene is activated.

DNA polymerases (pols) play a pivotal role in both replication and repair of genomic DNA. Of the seven DNA pol families (A, B, C, D, X, Y, RT), replicative pils make up four (A-D). Replicative pils are highly accurate and processive, synthesizing long stretches of DNA in a single binding event, while specialized pils mainly repair and bypass DNA damage and are only able to synthesize a few nucleotides before dissociation.<sup>1</sup> During replication, a replicative pil may encounter DNA lesions or strand breaks which are induced by exposure to a variety of endogenous and exogenous factors such as oxidative metabolites or UV radiation. Many DNA lesions distort the structure of DNA, stalling the progression of replicative pils. Consequently, a specialized pil must take over and perform translesion synthesis (TLS) to bypass the DNA lesion. Because they have relatively flexible active sites, bypass pils are capable of inserting a dNTP

opposite the damaged base before the replicative pol can regain access to the 5'-primer terminus and continue DNA synthesis. Other specialized pols can repair single and double DNA strand breaks via base excision repair (BER) and nonhomologous end-joining (NHEJ) so the replicative pol can continue.<sup>2,3</sup> Replicative pols must also interact with specifically modified bases that are essential for epigenetics.<sup>4-8</sup>

## 1.2 Types and Sources of DNA Damage

DNA damage can be endogenous (due to cellular metabolism), or exogenous (due to environmental factors). Endogenous sources include reactive nitrogen and oxygen species and other metabolic products and byproducts that contribute to the oxidation, hydrolysis, and alkylation of DNA. Mismatched base pairs, shortening of telomeres, and misregulation of epigenetic markers are other damaging events that can happen due to natural cellular processes such as division and aging. Exogenous sources are varied and include UV-light damage, ionizing radiation damage, and a whole host of chemical agents that we ingest, inhale, or absorb through our skin in our day-to-day activities.<sup>9-11</sup>

Both sources of DNA damage can lead to different types of damage which require different repair pathways to be activated. Mismatched bases result from an insertion or deletion of bases during replication, and is repaired via the mismatch repair (MMR) pathway.<sup>12</sup> UV radiation and mutagenic chemicals can cause bulky DNA lesions that are repaired through the nucleotide excision repair (NER) pathway, during which a section of nucleotides that includes the damaged site is excised from the DNA. This process cuts the phosphodiester backbone of the DNA strand in two places. The gap is subsequently filled

in with undamaged nucleotides and ligated at both ends to form a new, lesion-free section of DNA.<sup>9,12</sup>

Base damage, abasic sites, and single strand breaks (SSBs) are common forms of DNA damage and are caused by oxidation, alkylation, hydrolysis, radiation, and some DNA adducts such as 8-oxo-guanine. This type of damage can be repaired in the BER pathway using either the short-patch or long-patch methods. During BER, the damaged base is first removed, leaving the phosphodiester backbone intact and a resulting abasic site in the DNA. The DNA backbone is then cut at the 3' end of the abasic site via deoxyribosephosphatase (DRPase), producing a nick in the DNA strand. A new base is inserted by X-family pol  $\beta$ , and the nick is sealed by a DNA ligase, LigIII/XRCC1. Short patch BER removes and replaces one nucleotide, while long-patch BER can involve up to approximately 13 nucleotides.<sup>13</sup>

Double strand breaks (DSB) can also be caused by oxidation and various types of radiation and mutagenic chemicals. This type of DNA damage must be repaired via NHEJ, homologous recombination (HR), or microhomology-mediated end-joining (MMEJ). In mammals, DSBs are repaired via NHEJ the vast majority of the time. This process requires two blunt ends to the broken DNA, so any single stranded overhang must either be filled in by a pol or trimmed back by a nuclease before the blunt ends can be ligated to other DNA strands, fixing the break. NHEJ does not require a homologous DNA template for repair.<sup>14</sup> In contrast, HR requires a homologous template for DNA repair and does not work with blunt-end DNA breaks. Instead, there must be an overhanging 3' end. If the overhang does not exist in the original DSB, this pathway

utilizes enzyme that will trim back the 5' end to create this overhang in a process called resectioning.<sup>15</sup> Contrasting to both of these DSB repair pathways is MMEJ, whereby two broken strands of dsDNA with overhangs on either the 5' or 3' end will be ligated together by base pairing a region of 5-25 homologous nucleotides. This repair pathway results in deletions at the junction of the base pairing, thereby leading to chromosomal instability, abnormalities, and complex rearrangements of the chromosome structure. However, this repair pathway also helps rescue collapsed replication forks.<sup>16,17</sup>

### 1.3 DNA Polymerase Structure and Mechanism

Most pols have the same general structure of the catalytic domain that enables them to add nucleotides to the growing DNA strand, the main exception being PrimPol. Additionally, each pol has unique structural characteristics within and outside the polymerase active site that help confer specific activity to the pol. Here I will introduce the general structural characteristics of DNA pols, focusing on the polymerase domain, and mechanism of nucleotide incorporation. Keep in mind there is an incredible range of nuances in pol structure that enable each pol to have a distinct role in the cell.

#### *1.3.1 Structure of the Polymerase Domain*

The catalytic domain of most pols has a conserved shape consisting of 3 subdomains: palm, fingers, and thumb, labeled as if looking at a right hand. The exception to this nomenclature is that X-family pols have a topologically distinct arrangement of the subdomains, so they are viewed as left-handed instead of right-handed. The polymerization activity is located specifically in the palm subdomain and contains three conserved acidic residues. For example, in pol X- and B-families the

catalytic residues are all aspartates, but in A- and Y-families there are two aspartates and one glutamate. These three residues coordinate two divalent metal ions, generally agreed to be  $Mg^{2+}$ , which are necessary for the pol to incorporate the incoming dNTP. A recently discovered third metal ion may also play a role in stabilizing the transition state of the reaction and may therefore aid in catalysis.<sup>18</sup> The fingers subdomain has the function of nucleotide binding and selection of correct nucleotide while excluding incorrect nucleotides, and the thumb subdomain is responsible for binding the DNA substrate (Figure 1).

Y-family pols, which are primarily responsible for catalyzing TLS processes in the cell, have shorter fingers and thumb subdomains and more solvent-exposed active site. These characteristics make the Y-family pols more amenable to bypassing a bulky DNA adduct than pols from other families, which have longer thumb and fingers subdomains and a channel for incoming nucleotides leading from the surface of the enzyme to the active site.<sup>19-21</sup>

### *1.3.2 Mechanism of Nucleotide Incorporation*

There is a consensus minimal mechanism of nucleotide incorporation consisting of six steps, including two steps where a conformation change occurs. The first step involves the pol binding to the primer terminus of the DNA substrate, followed by binding a deoxyribonucleotide triphosphate (dNTP). Subsequently, the pol undergoes a conformation change before performing the chemistry step that incorporates a deoxyribonucleotide monophosphate (dNMP) and results in a pyrophosphate byproduct. The fifth step is a conformational change back to the original state of the pol, followed by

release of the pyrophosphate. At this point, the pol could stay associated to the DNA substrate, now with a primer longer by one dNMP, and carry out another incorporation, or the pol could dissociate from the newly formed substrate (Figure 2).<sup>21-25</sup>

## 1.4 Replicative Polymerases

The human genome encodes 3 replicative pols: pol  $\alpha$ , pol  $\delta$ , and pol  $\epsilon$ . Each pol has a unique job in the replication of DNA and all three coordinate to efficiently and accurately copy DNA at rates of 1-2 kb/min.<sup>26</sup> The very stable pol  $\alpha$ -primase complex synthesizes single-stranded RNA primers and then extends those primers with dNTPs. Pol  $\epsilon$  synthesizes the leading strand and pol  $\delta$  synthesizes the lagging strand. Pols  $\epsilon$  and  $\delta$  are recruited to their respective strands of DNA via various protein interactions and are similarly prevented from acting on the other strand, in effect being regulated to a specific strand during replication to ensure the most efficient carrying out of the process possible.<sup>27</sup>

### 1.4.1 Polymerase $\alpha$

Pol  $\alpha$ -primase is necessary to prime both the leading and lagging strands. To carry out this function, it requires association with the CMG (Cdc45-Mcm2-7-GINS) helicase, a holoenzyme that unwinds double-stranded DNA (dsDNA) into two single strands.<sup>26</sup> Pol  $\alpha$  consists of 4 subunits—one with DNA polymerase activity, two that coordinate to form RNA primers, and one subunit with unspecified function, called the B subunit. In the presence of pols  $\delta$  and  $\epsilon$ , pol  $\alpha$ 's polymerase activity is curbed quickly, which is one way for the cell to maintain accurate DNA replication, since pol  $\alpha$  does not have a 3'-5' exonuclease proofreading activity like the other two human replicative pols

and therefore cannot correct any mistakes it makes during replication. Pol  $\alpha$ -primase generates primers about 10 ribonucleotides long before its polymerase activity takes over and it adds on deoxyribonucleotides to a total primer length of about 25-30 nucleotides before pol  $\delta$  or  $\epsilon$  take over replication.<sup>27,28,29</sup>

#### *1.4.2 Polymerase $\delta$*

Pol  $\delta$  has several cellular functions. It has been known for many years that it performs lagging-strand synthesis during DNA replication.<sup>30-32</sup> Pol  $\delta$  is closely associated with proliferating cell nuclear antigen (PCNA) while carrying out replication on the lagging strand. PCNA has a ring-like structure that is loaded onto and can subsequently slide along a DNA strand. Since pol  $\delta$  binds PCNA more tightly than pol  $\epsilon$  does, it can outcompete pol  $\epsilon$  for PCNA binding spots on the lagging strand and therefore carries out the bulk of lagging strand synthesis.<sup>27</sup> It utilizes the short primers left by pol  $\alpha$  to create Okazaki fragments which are processed and ligated into a continuous DNA strand.

However, recent studies have pointed to additional roles that pol  $\delta$  plays. Pol  $\delta$  has been found to have activity on the leading strand in the absence and presence of pol  $\epsilon$ . If the polymerase subunit of pol  $\epsilon$  is mutated out, pol  $\delta$  can synthesize the leading strand DNA, albeit at a much slower pace than pol  $\epsilon$  can. However, even when all 3 fully functional replicative pols are present, pol  $\delta$  still shows some activity on the leading strand.<sup>33-35</sup>

#### *1.4.3 Polymerase $\epsilon$*

Pol  $\epsilon$  primarily performs leading strand DNA synthesis.<sup>31,32,36</sup> Pol  $\epsilon$  is relegated to the leading strand by its strong association with the CMG helicase during replication. Pol  $\epsilon$  also associates with PCNA during replication of the leading strand. Competition with pol  $\delta$  for binding to PCNA is circumvented by the fact that pol  $\epsilon$  is also tightly bound to CMG, pol  $\delta$  has no observable interaction with CMG, and CMG is bound to the leading strand of DNA.<sup>27</sup> Pol  $\epsilon$  cannot extend Okazaki fragments and has no synthesis activity on the lagging strand, leaving that duty completely to pol  $\delta$ . Besides replication, pol  $\epsilon$  is essential for general genome stability and CMG activation. Pol  $\epsilon$  helps recruit all factors necessary for CMG formation, may stimulate CMG activity, helps repair UV damage, and is needed to repair long sections of double-strand breaks in the DNA. Mutations to pol  $\epsilon$  lead to increased DNA damage, higher risks for cancer, and defects in chromatin segregation and methylation of replicated DNA.<sup>26,31,37-39</sup>

Similar to pol  $\alpha$  and pol  $\delta$ , the pol  $\epsilon$  holoenzyme also consists of 4 subunits. Just like pol  $\delta$ , the largest subunit contains both the polymerase and exonuclease activities. The three smaller subunits have less defined functions, although some roles have been determined and others have been suggested. One of the smaller subunits, p17, is identical to a protein used in the chromatin accessibility complex (CHRAC), which indicates a role for pol  $\epsilon$  in heterochromatin replication. Additionally, it is the non-catalytic subunits that facilitate the interaction between pol  $\epsilon$  and CMG.<sup>38,39</sup>

## 1.5 Specialized Polymerases

Unlike replicative pols, specialized pols usually possess no exonuclease activity. These pols are not used in the normal copying of DNA, but are designed to bypass DNA lesions, mismatched base pairs, and other forms of DNA damage. Specialized pols belong mainly to X- and Y-families but can also be found in A- and B-families and include pols  $\beta$ ,  $\lambda$ ,  $\mu$ ,  $\kappa$ ,  $\iota$ ,  $\eta$ ,  $\zeta$ , and  $\theta$ , among others. While specialized pols have some overlapping functions, they each fill a niche role and have specific characteristics. These pols can operate in one or more of the possible DNA repair pathways, such as NER, BER, MMR, HR, and NHEJ. Deregulation, misregulation, and mutants of specialized pols are implicated in many types of cancer.<sup>23,40–43</sup>

#### *1.5.1 X-family Polymerases*

This class includes pols  $\beta$ ,  $\lambda$ , and  $\mu$ . Pol  $\beta$  takes part in BER and has been strongly implicated to partake in TLS.<sup>2,29,43–46</sup> Pol  $\lambda$  also performs BER, but likely in a back-up role to pol  $\beta$ . Cells lacking both pols  $\beta$  and  $\lambda$  cannot carry out BER.<sup>45</sup> Pol  $\lambda$  has also been implicated in additional roles in nonhomologous end-joining (NHEJ) and homologous recombination (HR) activities in the cell.<sup>43,45–48</sup> Pol  $\mu$  has roles in both TLS and NHEJ. In fact, pols  $\lambda$  and  $\mu$  are the only two pols that have NHEJ capability.<sup>2,47–50</sup>

#### *1.5.2 Y-family Polymerases*

All pols within the Y family are capable of TLS and have relatively low fidelity ( $10^0$ - $10^{-4}$ ) compared to pols  $\delta$  and  $\epsilon$  ( $10^{-6}$ - $10^{-8}$ ).<sup>2,3,51</sup> Each pol has unique misincorporation patterns, preferred substrates, and different bypass abilities. This family includes pols  $\kappa$ ,  $\iota$ ,  $\eta$ , and Rev1. Pol  $\eta$  is the first pol with TLS activity discovered and only pol that is causally linked to cancer development. A variant of the rare autosomal

disorder *Xeroderma pigmentosum* (XPV) is caused by a lack of pol  $\eta$  activity. Pol  $\eta$  therefore plays a very important role in bypassing UV-induced damage. It is also the only pol able to bypass cisplatin adducts, double-base lesions caused to DNA after being treated with the common antitumor drug.<sup>2,3,43,51,52</sup> Pol  $\iota$  may also play a backup role in bypassing UV damage to DNA. Human cells with downregulated pol  $\iota$  show hypersensitivity to oxidative damage and decreased BER activity. Pol  $\iota$  does get recruited to chromatin after oxidative damage and interacts with known BER pathway enzymes, indicating a backup role for pol  $\iota$  in BER to pols  $\beta$  and  $\lambda$ . Pol  $\iota$  is also specialized in its role in the mutation of immunoglobulin which is critical for the adaptive process of somatic hypermutation. This may be due to the fact that pol  $\iota$  is the most error-prone of the Y-family pols.<sup>29,43,53,54</sup> Pol  $\kappa$  is unique in that it is the TLS pol with the highest expression and most conserved sequence identity across all domains of life. It has the special ability to extend past a lesion or a mismatched primer terminus for about 20 nucleotides before a replicative pol takes over. This extension activity is probably one of the largest roles pol  $\kappa$  plays in the cell, instead of the actual lesion-bypass step of TLS. Human cells and live mice with pol  $\kappa$  deletions have very mildly negative phenotypes, and pol  $\kappa$  has weak TLS activity compared to other Y-family pols. However, pol  $\kappa$  can bypass and extend past a certain type of bulky lesions that other TLS pols cannot easily bypass: lesions caused by bulky, polycyclic compounds derived from combustion products or chemotherapeutic agents. It is hypothesized that pol  $\kappa$  may play a role in spontaneous mutating of the

genome due to its extension ability. Pol  $\kappa$  is similar to pols  $\iota$  and  $\eta$  in that it plays a role in bypassing UV damage to DNA.<sup>3,55-57</sup>

### *1.5.3 Other Specialized Polymerases*

There are two more specialized pols worth mentioning that do not belong to the X- or Y- families, pol  $\theta$  and pol  $\zeta$ . Pol  $\theta$  is an A-family pol that likely coordinates with pol  $\beta$  in the BER pathway to bypass abasic sites. While pol  $\theta$  is not necessary for cell viability, mammalian cells lacking pol  $\theta$  display higher levels of DNA damage from ionizing radiation and have increased radiosensitivity when being treated for cancer with radiation therapy.<sup>2,58-62</sup> Pol  $\zeta$  belongs to the B-family and has the unique role of extending past damage sites once another TLS pol has inserted a nucleotide opposite the lesion, especially in the case of UV damage. Pol  $\zeta$  and pol  $\kappa$  may have overlapping roles, since both have the extension capability and deletion of pol  $\kappa$  is not lethal to mammalian cells. However, expression of pol  $\zeta$  is tightly regulated and irregular expression of pol  $\zeta$  leads to chromosomal instability and mutagenesis while deletion of pol  $\zeta$  or of just the catalytic activity of the pol is lethal to mice embryos.<sup>2,43,63,64</sup>

## 1.6 Significance and Impact

Because DNA pols are essential to life, it is imperative that these enzymes and the cellular processes they are involved in are researched thoroughly. Basic research gives insight into the replication and repair pathways and various disease states which can help application-focused research develop new drugs and therapies for cancer, genetic-based diseases and disorders, and age-related diseases. The more we understand how the body naturally repairs, protects, and passes on its genomic information, the better we can craft

methods to correct the damage when any of those processes become misregulated or deregulated and prevent further damage from happening.

Figure 1: Architecture of the Catalytic Subunit of Pol2 from *S. cerevisiae*, a homolog of human Pol  $\epsilon$

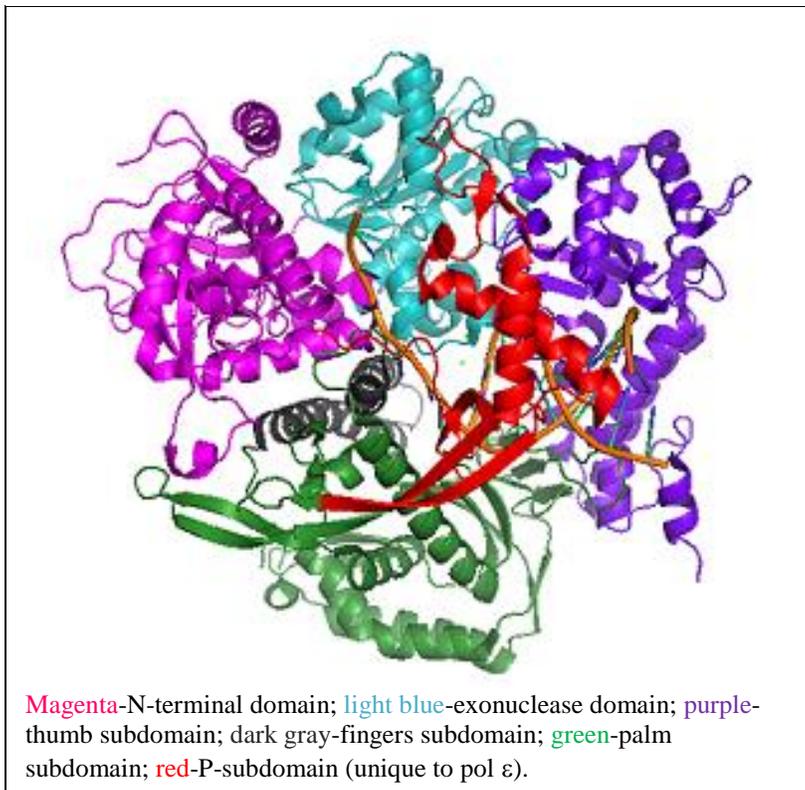
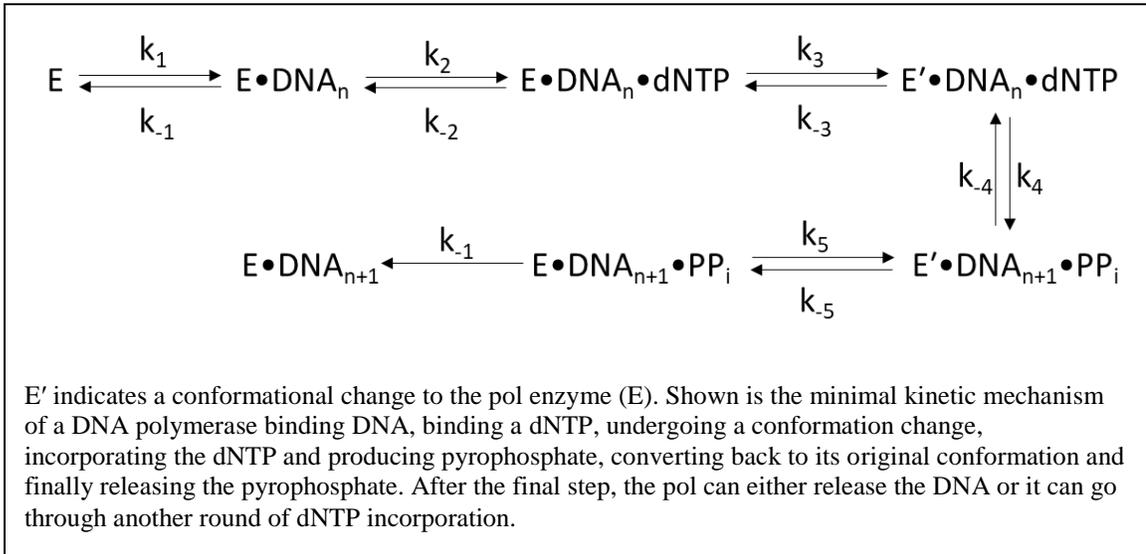


Figure 2: Scheme of Polymerase Activity



## Chapter 2: Theory

### 2.1 Kinetics

The kinetics of a chemical reaction can shed light on the mechanism and pathway of the reaction by studying the rate at which various steps of the reaction happen. Kinetics experiments can tell us what the rate-limiting step is. Is it ligand binding, product formation, or release of product? These types of experiments, together with knowledge of the structure of the enzyme being studied or of a similar enzyme, can also help us determine whether the rate-limiting step is due to chemistry (at the enzyme active site) or conformational change (of the enzyme).

There are two states of a reaction that we can look at: steady-state and pre-steady state. The steady-state frame looks at an enzyme-catalyzed reaction under the reagent and time conditions when reaction intermediates are being formed and deteriorated (or used up) at the same rate. At this point, since the overall concentration of the reaction intermediate is neither increasing nor decreasing, it is said to be in a steady state. Under steady-state reaction conditions, the substrate is in excess over the enzyme so that multiple catalytic turnovers do not greatly affect the substrate concentration and little product is made in the first turnover. In addition to this experimental setup, two assumptions must be made: there is only one reaction intermediate, and there is negligible reversal of the intermediate to product step (Figure 3).<sup>65,66</sup>

The plot of [product] vs. time under these conditions yields a straight line. The slope of the line is the steady-state rate of product release ( $v_{ss}$ , concentration/time), otherwise known as the turnover number. By extrapolating the slope to  $y=0$ , the active enzyme concentration ( $[E_0]$ ) can be found and the intrinsic kinetic rate constant,  $k_{off}$ , can be calculated:

$$k_{off} = v_{ss}/[E_0]$$

where  $k_{off}$  has units of  $\text{time}^{-1}$ .<sup>65,66</sup>

Two more pieces of data can be gathered from a steady-state reaction:  $K_m$  and  $k_{cat}$ . The  $k_{cat}$  value is the maximum product formation rate from the enzyme-substrate intermediate. It is also the lower limit of any first-order rate constant of any single step in the reaction.  $K_m$  is the ratio of the rate of decomposition of the enzyme-substrate intermediate (including to product and back to unbound substrate) to the rate of productive enzyme-substrate binding. It can also be defined as the substrate concentration at which  $k_{cat}$  is  $1/2$  the maximal value. This value can give a rough estimate of the intrinsic equilibrium constant  $K_d$  (the dissociation constant). A lower  $K_m$  (and  $K_d$ ) indicates less dissociation of the substrate from the enzyme, which in turn indicates tighter binding and higher stability of the enzyme-substrate complex. Additionally, the value  $k_{cat}/K_m$  can give insight into the specificity of the substrate for a particular enzyme. The higher this ratio is, the better the substrate binding to the enzyme is.<sup>65-67</sup>

The  $k_{off}$ ,  $k_{cat}$  and  $K_m$  can be found by fitting of the data to the appropriate equations. First, product concentration vs. time data for each concentration of dNTP would be plotted and fit to the general line equation:

$$[\text{Product}] = v_{ss}t + E_0$$

where  $E_0$  is the y-intercept indicating the initial amount of enzyme-substrate complex and  $v_{ss}$  is the slope of the line indicating the steady-state rate of the enzymatic reaction. This data would be compiled to calculate several  $v_{ss}/E_0$  values which would subsequently be plotted vs. the corresponding dNTP concentrations and using a nonlinear regression software such as Kaleidagraph or GraphPad Prism, fit to the line:

$$k_{off} = (k_{cat}[S])/(K_m + [S])$$

where  $[S]$  is the dNTP concentration. The software will then solve for two values,  $k_{cat}$  and  $K_m$ .<sup>66,67</sup>

However, the limitations of the steady-state frame of reference is that we cannot determine the rate constants of the individual steps of a reaction. The values we determine are only indicative of the overall reaction, and cannot give us insight into the specific steps of a mechanism of reaction, such as what the rate-limiting step is and what it is due to, or the order of substrate binding at the enzyme active site. To determine these specifics, pre-steady state kinetic methods must be applied to the reaction.

Pre-steady state reaction conditions must use higher amounts of enzyme than in steady-state reactions so that a significant amount of product is formed in the first turnover. The time course for a pre-steady state reaction is usually shorter than for a steady-state reaction, too. The assumptions for pre-steady state kinetics are the same as for steady-state. Under these conditions, a “burst” phase is seen, which is a state of the reaction when the intermediates are being formed quickly before being used up to form the final product. The state of the system under these conditions is not steady; instead, the

reaction intermediate is forming faster than it is being consumed. The amplitude of the burst phase should be the same or close to the value determined by extending the steady-state slope to  $y=0$ . However, after the burst phase, a second, linear phase should be apparent, and this slope of this linear phase should equal the slope of the linear steady-state phase.<sup>65-67</sup>

A specific assay of pre-steady state kinetics is the single-turnover assay. In this as, the enzyme is in excess over the substrate so that a burst phase will definitely be seen, but the second linear phase will not be seen. Again, the time course of this type of experiment is generally shorter than would be used for a steady-state kinetic experiment, perhaps in the order of ms-min. Under these conditions, all substrate should be bound by the enzyme and the first turnover will produce a significant amount of product. Again, a sharp buildup of reaction intermediates is evidenced by the burst phase, but the time course chosen limits the experiment so the steady state phase is not reached.

Pre-steady state kinetics gives much more detailed insight into the individual steps of a reaction. The kinetic parameters of  $k_{obs}$ ,  $K_d$ , and  $k_{pol}$  can be determined. The parameter  $k_{obs}$ , which is the rate at which the reaction approaches equilibrium, can be determined by plotting [product] vs time and using nonlinear regression to fit the data using the equation:

$$[\text{Product}] = E_0 \bullet [1 - \exp(-k_{obs}t)] + v_{ss}t.$$

However, keep in mind that under single-turnover conditions, the linear part of this equation will not be used since we do not allow the reaction to reach a steady state, but the rest of the equation remains the same.<sup>65-67</sup>

The first part of the pre-steady state equation describing the appearance of product before the steady-state phase is in the form of a single exponential and characteristic of a first-order reaction or a pseudo-first-order reaction. Enzyme-substrate reactions are second-order reactions, meaning two different molecular species interact to form a product. However, since the amount of dNTP substrate that is used in the reaction is usually much higher than the amount of enzyme used, the reaction is pseudo-first order with respect to the dNTP concentration, which is why we can use a first-order single exponential curve to fit the [product] vs. time data. If the reaction were truly first-order, the  $k_{obs}$  should not change with changes in dNTP concentration; however, we see that with pols this is not true.<sup>65,67</sup> The  $k_{obs}$  generally increases as the dNTP concentration increases. The  $k_{obs}$  is important to determine because this parameter is what is used to calculate the maximum rate of dNTP incorporation ( $k_{pol}$ ) and the dissociation equilibrium constant of the dNTP ( $K_d$ ).

Once all the [product] vs. time plots have been fit to the pre-steady state equation and the  $k_{obs}$  determined for all time courses, we then plot the  $k_{obs}$  values vs. the appropriate [dNTP] values. Using nonlinear regression software, the data are then fit to a hyperbolic equation:

$$k_{obs} = k_{pol}[\text{dNTP}]/(K_d + [\text{dNTP}]).^{65,67}$$

## 2.2 X-ray Crystallography

### 2.2.1 Growing a Crystal

The first hurdle to performing X-ray crystallography is growing a crystal. The protein sample used must be very pure. The most common methods to grow crystals are

sitting-drop and hanging-drop vapor diffusion. In sitting-drop, the protein sample sits in a drop on the bottom of the sample chamber and in hanging-drop, the protein is contained within a drop at the top of the sample chamber. Though the location of the protein sample in the crystallization chamber may change, crystals in both settings grow via vapor diffusion. The drop contains a mixture of the protein sample along with some buffer, salt, and one or more precipitants. The reservoir solution contains a higher concentration of the drop components, but no protein. Water therefore diffuses from the drop and into the reservoir solution, slowly concentrating and lowering the solubility of the protein. The goal is for the protein molecules to have time to pack together in an orderly manner and crystallize, instead of quickly aggregating and precipitating out of solution. This process involves much trial and error. Typically, at least one large screening of many different crystallization conditions is done first and if any conditions show crystal growth or precipitation, smaller screens are set up using conditions that vary slightly from the promising ones seen in the larger screen. In this manner, several rounds of screens may be set up before one is found to grow crystals.

### *2.2.2 X-Ray Diffraction*

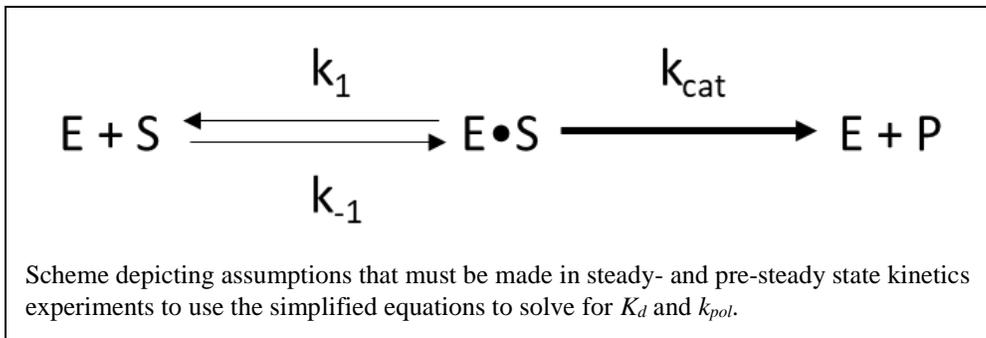
Once crystals are grown, they can be subject to beams of X-ray radiation (high-energy photons) to gain information on the structure. The planes of atoms within a crystal act as reflecting surfaces. When an X-ray hits an atomic plane, it is reflected. This is due to the electric field vector of the photon interacting with the electric charge of the electrons in the atoms of the material and scattering. If the reflections off multiple planes within a crystal are in phase, the result is constructive interference and a bright spot of

radiation is observed on the detector (X-ray film, CCD camera, etc). If the reflections are out of phase, there is destructive interference and a weak spot or no spot shows up on the detector. Bragg's Law defines the conditions under which constructive interference happens:

$$n \cdot \lambda = d \cdot \sin \theta$$

where  $n$  is an integer,  $\lambda$  is the wavelength of the X-rays,  $d$  is the distance between the atomic planes, and  $\theta$  is the angle between the incoming X-ray and the surface of the crystal. By correlating all reflections to the angle of the X-ray that produced them, the spacing of all atoms within a crystal can be found.

Figure 3: Assumptions for Steady- and Pre-Steady State Kinetics



## Chapter 3: The Potential Role of Polymerase $\epsilon$ in Bypassing Epigenetic Modifications

### *3.1 Introduction*

Epigenetic modifications are any inherited alterations to DNA bases that do not change the DNA sequence. These DNA modifications regulate access to chromatin and the shape and stability of chromatin, thereby regulating gene expression.<sup>68</sup> Replicative forks will regularly interact with epigenetically modified bases during cell division. The most common modification encountered is 5-methylcytosine (5mC) which regulates numerous biological processes including gene expression, retrovirus silencing, X chromosome inactivation, and normal development of embryonic stem cells<sup>6,69,70</sup>. Altered patterns in the methylation state of DNA can trigger cancer, psychiatric disorders, myeloid malignancies, embryonic lethality, and ultimately cell death in both embryos and adult mammals.<sup>5,6,8,69,71–73</sup> 5-methylcytosine was assumed to be a stable epigenetic marker until the recent discovery of several oxidative derivatives including 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC). Oxidation of 5mC is catalyzed by TET proteins (Tet1, Tet2, Tet3) which convert 5mC to 5hmC, then to 5fC, and finally to 5caC.<sup>8,69,73–75</sup> Additionally, these oxidative derivatives could arise through reaction with reactive oxygen species (ROS) in the cell.<sup>8,69</sup> The development of bisulfite and TAB-assisted bisulfite sequencing has allowed for single-

base-resolution analysis of 5mC and 5hmC sites in the genome, respectively, providing insight into role of these modified bases in gene regulation and other cellular processes.<sup>5,74</sup> Conversely, 5fC and 5caC are more difficult to identify and their precise roles in regulation of cellular processes, if any, are poorly understood. However, it is known that 5fC and 5caC can be repaired through base excision repair (BER) in which the modified bases are recognized and removed by thymine DNA glycosylase (TDG) (Figure 4).<sup>5,6,8,69,70,73–76</sup> Interestingly, TDG removes 5fC and 5caC from the DNA approximately 40% faster than its natural substrate, a thymine base mispaired with guanine, possibly due to the weakened N-glycosidic bond between the nucleobase of the modified cytosine and the deoxyribose sugar.<sup>5</sup> Furthermore, mismatch repair (MMR) proteins such as the human MutS $\alpha$  complex also recognize 5caC sites.<sup>8,75,77</sup> Though MutS $\alpha$  specifically recognizes G:T base pairs, it will also act upon *correct* G:5caC base pairs<sup>8,75</sup>. It has been theorized that G:5caC base pairs mimic G:T mismatches because the 5caC can form the imino tautomer and adopt G:T-like geometry when base paired to G, though this has not been confirmed structurally (Figure 5A).<sup>75,77</sup> Both repair pathways result in restoration of unmodified cytosine and may serve as alternative pathways for demethylation and gene regulation.

Intriguingly, it has been shown that Pols with exonuclease activities, such as the eukaryotic replicative Pol  $\delta$ , will excise a correctly paired 5caC:G base pair as if it were a mispair.<sup>75</sup> Low concentrations of Klenow fragment and Pol  $\delta$  containing their respective exonuclease domains show slight stalling when incorporating opposite 5caC, and Pol  $\delta$  shows slight stalling one to two nucleotides downstream of the modified templating base.

At higher enzyme concentrations, the exonuclease activity is stimulated, resulting in both continued stalling and DNA fragments smaller than the original template. However, pol  $\eta$ , a Y-family pol with no exonuclease domain, did not stall when incorporating opposite 5caC.<sup>75</sup> Other Y-family DNA pols such as  $\kappa$  and  $\iota$  may display the same bypass ability for the 5caC modification and play a role in TLS for this and other 5xC lesions *in vivo*.

These results bring to light two striking characteristics of the 5caC modification:

1) 5caC acts as an oxidative lesion, triggering BER pathways or even cell death through the MMR pathway, and 2) even when correctly base paired with G, 5caC is recognized as a mismatch by replicative polymerases and mismatch repair proteins. These observations may suggest a mechanism of repetitive incorporation and removal of a nucleotide opposite 5caC by a replicative pol with endogenous exonuclease activity, ultimately resulting in stalled replication. Pol  $\iota$ , a Y-family DNA pol, is characterized by having lower fidelity than replicative pols and having TLS capabilities. These characteristics made it a good candidate for studying the kinetics of bypassing an epigenetic adduct to a base. Pol  $\iota$  is error-prone and known to preferentially form G:T mismatches, and therefore may be an *in vivo* candidate for bypassing the 5caC modification and rescuing stalled replication.<sup>78,79</sup> Here, we have investigated the kinetics of nucleotide incorporation opposite modified cytosines using Y-family DNA Pol  $\iota$ .

### 3.2 Materials and Methods

#### 3.2.1 Materials

Chemicals used for experiments were purchased from the following companies:

[ $\gamma$ -<sup>32</sup>P]ATP from PerkinElmer Life Sciences (Boston, MA), Optinkinase from USB

Corp. (Cleveland, OH), dNTPs from Bioline (Taunton, MA), and oligonucleotide primers from Integrated DNA Technologies (IDT; Coralville, IA). Oligonucleotide templates were HPLC-purified and provided graciously by Professor Chuan He from the Department of Chemistry at the University of Chicago.

### *3.2.2 Preparation of DNA Substrates:*

The 26-mer DNA primer was purified using denaturing polyacrylamide gel electrophoresis (PAGE). It was radiolabeled on the 5'-end by incubating with [ $\gamma$ - $^{32}$ P]ATP and 2  $\mu$ L Optikinase for 3 h at 37 °C. The excess [ $\gamma$ - $^{32}$ P]ATP was removed using a Bio-spin 6 column (Bio-Rad). A labeled primer was mixed with the 5.4-fold molar excess of labeled modified-C template or control template, heated to 72 °C if using a modified-C template or 95 °C if using the control template for 5 min, and cooled slowly to room temperature (Table 1). The concentration of each DNA substrate was confirmed using a UV-Vis spectrometer by measuring the absorbance at 260 nm. The DNA templates provided to us by Professor Chuan He were not re-purified in our lab using gel purification since they were already HPLC-purified and we were concerned with major loss of template when using gel purification. All concentrations were verified using UV-Vis spectroscopy.

### *3.2.3 Expression and Purification of Human Polymerase $\epsilon$*

Plasmid DNA encoding for the N-terminal 420 residues of human pol  $\epsilon$  with an N-terminal GST tag and a 6X-His tag (GST-h $\Delta$ Pol $\epsilon$ ) was transformed into Rosetta cells. An overnight culture was grown in sterile LB media containing 34  $\mu$ g/mL chloramphenicol (Cam) and 50  $\mu$ g/mL ampicillin (Amp) at 37°C and 224 RPM. One-L

expression cultures containing the same concentration of Cam/Amp as the starter culture were inoculated with 1% volume of the starter culture each and grown at 37 °C and 224 RPM until the culture reached OD<sub>600</sub> of 0.6-0.8. At that point, each culture was induced with 0.1 mM IPTG and grown overnight at 16 °C and 224 RPM until an OD<sub>600</sub> of 1.7-1.9 was reached.

The cultures were centrifuged at 40k RPM at 4 °C for 40 min and the cells were resuspended in lysis buffer (50mM HEPES, pH 7.5, 0.5 M NaCl, 10% glycerol, 1 mM EDTA, 0.1% BME). Prior to lysis, a Roche protease inhibitor cocktail tablet and PMSF to 2 mM was added. Cells were passed through a French press 3x at 1500 PSI. Lysate was ultracentrifuged as before and the supernatant containing the soluble fraction of GST-hΔPolI was applied to a gravity GST sepharose column. The column was washed with 10 CV of wash buffer 1 (50mM HEPES, pH 7.5, 0.1 M NaCl, 10% glycerol, 0.1% BME) and eluted with 4 CV of GST elution buffer (50 mM Tris-HCl, pH 8.0, 10 mM Reduced Glutathione, 1 mM EDTA). All protein-containing fractions were dialyzed with heparin buffer A (50 mM TRIS, pH 7.5 at 4°C, 200 mM NaCl, 1mM EDTA, 10% Glycerol, 0.1% BME) and digested with His-TEV (1 mg TEV/100 mg protein) at 4°C overnight.

Cleaved hΔPolI was then passed through tandem HiTrap Q/HiTrap Heparin columns (GE Lifesciences), washed with 2 CV heparin buffer A and eluted in a linear gradient to 100% heparin buffer B (50 mM TRIS, pH 7.5 at 4°C, 1 M NaCl, 1mM EDTA, 10% Glycerol, 0.1% BME) over 5 CV. Protein purity was assessed using a 17% SDS polyacrylamide gel and the concentration was determined using a UV-Vis spectrometer.

The purified protein was dialyzed against 1 L of storage buffer (50 mM Tris-Cl, pH 7.5, 140 mM NaCl, 1 mM DTT, 0.5 mM EDTA, 50% Glycerol) overnight at 4 °C and stored at -20 °C until use.

#### 3.2.4 Single-Turnover Polymerase Assays:

All pre-steady-state kinetic assays were performed in reaction buffer containing 50 mM HEPES, pH 7.5 at 37 °C, 5 mM MgCl<sub>2</sub>, 50 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 10% glycerol, and 0.1 mg/mL bovine serum albumin. All given concentrations are final after mixing all solutions. Single-turnover kinetic assays were employed to obtain the  $k_{pol}$  and  $K_d$  for single nucleotide incorporation. Briefly, a pre-incubated solution of 5'-[<sup>32</sup>P]-labeled DNA (10 nM) and hΔPolI (60 nM) in reaction buffer was mixed with increasing concentrations of an incoming dNTP at 37 °C and quenched at various times with 0.37 M EDTA using a rapid chemical-quench flow apparatus (KinTek). Reaction products were separated using denaturing PAGE, visualized using Typhoon TRIO (GE) and quantified using ImageQuant (GE) (Figures 6A and 6B). Each time course of product formation was fit to a single-exponential equation:

$$[\text{product}] = E_o \bullet [1 - \exp(-k_{obs}t)]$$

using Kaleidagraph (Synergy Software) to yield a reaction amplitude ( $E_o$ ) 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:

$$k_{obs} = k_{pol}[\text{dNTP}]/([\text{dNTP}] + K_d)$$

to yield an apparent equilibrium dissociation constant of enzyme for the nucleotide ( $K_d$ ) and a maximum nucleotide incorporation rate constant ( $k_{pol}$ ).

### 3.3 Results

Pre-steady-state kinetic assays were carried out to measure the kinetic parameters for incorporation of each dGTP by Pol  $\epsilon$  onto the four primer/template DNA substrates to study the ability of pol  $\epsilon$  to bypass 5xC epigenetic modifications and infer if it could have an *in vivo* role in TLS opposite the 5caC lesion. For the control DNA substrate with a templating dC, the  $k_{pol}$  and  $K_d$  for dGTP incorporation was determined to be  $0.22 \pm 0.014$  s<sup>-1</sup> and  $37 \pm 9$   $\mu$ M, respectively, and the efficiency ( $k_{pol}/K_d$ ) was calculated to be  $5.9 \times 10^{-3}$   $\mu$ M/s (Table 2). For the correct incorporation of dGTP opposite each modified cytosines, the  $k_{pol}$  values ranged from 0.23-0.53 s<sup>-1</sup>, similar to the control  $k_{pol}$  (0.22 s<sup>-1</sup>). The largest difference was the incorporation across 5hmC which was 2.4-fold faster than the control (0.53 s<sup>-1</sup>). There was no clear pattern in the  $k_{pol}$  for correct incorporation across the modified cytosines. The  $K_d$  values for dGTP incorporation ranged from 40-348  $\mu$ M, all higher than the  $K_d$  for the control substrate (37  $\mu$ M). The weakest binding was observed for the 5caC substrate, which has a 9.4-fold higher  $K_d$ . In general, the binding affinity decreased as the modification on the base became more oxidized. The efficiency of correct dGTP incorporation ranged from  $9.5 \times 10^{-4}$  to  $5.8 \times 10^{-3}$   $\mu$ M s<sup>-1</sup>. Overall, the modifications had only a modest effect on the efficiency for correct nucleotide incorporation.

The  $k_{pol}$  and  $K_d$  for incorrect incorporations opposite 5mC range from 0.023 to 0.58 s<sup>-1</sup> and 40 to 776  $\mu$ M, respectively. Across from 5hmC the incorporation rate ranged from 0.21 to 0.25 s<sup>-1</sup> and the  $K_d$  from 192 to 1130  $\mu$ M. The  $k_{pol}$  of incorrect incorporation opposite from 5fC ranges from 0.092 to 0.32 s<sup>-1</sup> and the  $K_d$  from 53 to 211  $\mu$ M. For 5caC,

the  $k_{pol}$  ranges from 0.012 to 1.0 s<sup>-1</sup> and the  $K_d$  ranges from 190 to 799  $\mu$ M (Table 2).

Except for the 5caC substrate, the other modifications have incorrect incorporation rates and binding affinities that vary by 10-fold or less, indicating that pol  $\tau$  is equally prone to incorporating any incorrect nucleotide opposite the modified cytosines. Additionally, the probability of incorporating the correct nucleotide steadily decreases as the modification on the cytosine becomes more oxidized. Pol  $\tau$  has the highest fidelity for the 5mC substrate, followed by 5hmC, then 5fC, and lastly, 5caC. The wider range of  $k_{pol}$  and  $K_d$  for 5caC compared to the other modified cytosine substrates may be an indication that this modification significantly perturbs the active site of pol  $\tau$ . Overall, pol  $\tau$  becomes increasingly more error-prone as the cytosine modification becomes more oxidized

### 3.4 Discussion

The 5caC modification most significantly affected nucleotide incorporation and lowered efficiency of correct dGTP incorporation by 10-fold. This is due entirely to the near-10-fold increase in  $K_d^{dGTP}$ . The distribution of the probability of incorrect nucleotide incorporation was not significantly affected by the identity. This indicates that pol  $\tau$  is not discriminatory in its choice of incorrect nucleotide.

However, as the cytosine modification become increasingly oxidized, the fidelity of Pol  $\tau$  decreases. Curiously, for the substrate containing a templating 5fC, Pol  $\tau$  exhibits a remarkably low fidelity (range). The largest variations in  $k_{pol}/K_d$ , and  $k_{pol}/K_d$  for the four dNTPs were for incorporation opposite 5caC. Consequently, dNTP efficiency opposite 5caC varies more widely than for any of the other modifications. We find that incorrect

dTTP incorporation is actually slightly preferred instead of the correct dGTP incorporation (44% vs. 32% probability of incorporation). This raises a few interesting points. First of all, if 5caC closely resembled a templating dT as hypothesized above, we would expect that dGTP would have the highest incorporation efficiency since Pol  $\epsilon$  is well known to preferentially incorporate dGTP opposite a templating dT. This raises the question of whether the G:5caC pair does actually mimic a G:T mismatch and if not, why it activates TDG and other mismatch repair enzymes.

All nucleic acids primarily exist in a single tautomeric form at physiological pH, which encourages strict adherence to Watson-Crick pairing rules. However, the transient existence of other tautomeric forms of the nucleic acids may be the primary cause for spontaneous mutations and misincorporations in DNA and RNA replication.<sup>80-84</sup> Addition of chemical groups to the nucleobases can stabilize a tautomeric form.<sup>85,86</sup> In the case of 5caC, the addition of the carboxyl group likely stabilizes the imino tautomer of cytosine. When 5caC is in the imino form, it could resemble a templating dT; therefore, Pol  $\epsilon$  would preferentially incorporate dGTP, and the resulting G:5caC pair would mimic the geometry of a G:T mismatch. With the kinetic parameters established with this study, a previously unknown characteristic of 5caC base pairing has been discovered. When 5caC is in the neutral tautomeric form, the hydrogen bond donor and acceptor pattern resembles that of an adenosine, and Pol  $\epsilon$  is known to have a modestly high fidelity for incorporating dTTP opposite a templating dA.<sup>87-89</sup> Since 5caC would remain primarily in the neutral tautomeric form, Pol  $\epsilon$  may preferentially incorporate dTTP and form a mismatch that resembles a canonical T:A base pair, which could later be excised in a

BER pathway not involving TDG (Figure 5B). However, structural studies in addition to more kinetic studies need to be carried out to confirm this hypothesis.

The presence of cytosines modified at the C5 position (5xC) in the genome is crucial to epigenetic control of various cellular processes. However, the placement of a modification within the genomic DNA and the prevalence of a modification must be tightly regulated. Abnormal patterns of 5xC lead to irregularities in gene expression.<sup>6,8,69,75–77,90,91</sup> While some 5xC modifications, such as 5mC, have been studied extensively, other 5xC modifications are less prevalent and more difficult to detect by current sequencing methods, making studies into the physiological roles of these modifications difficult to carry out.

5caC is that it is the least abundant of the known 5xC modifications, present only 1,000-100,000 times in an entire diploid human genome, which consists of  $3 \times 10^9$  basepairs.<sup>90</sup> However, it is known to be present in a higher abundance in cells of certain breast cancers, brain tumors, gliomas, and brain cells of patients with Schizophrenia and other related neurological disorders. The higher abundance can sometimes be correlated to a decrease in 5hmC levels, an increase in expression of TET enzymes, or a decrease in TDG expression; but sometimes the greater 5caC abundance is not linked to any of these observations and it is still unclear how to explain this phenomenon.<sup>73,92</sup> This modification also accumulates in the genome with increasing age and is associated with perturbations in other DNA demethylation enzymes.<sup>93</sup> Other studies have shown that 5caC is transiently accumulated in neural stem cells as they undergo lineage specification.<sup>94,95</sup> It

has also been shown that 5caC is a stable modification and that it transiently stalls RNA Pol II during transcription.<sup>90,96</sup>

These observations and others indicate that 5caC may play a dual role in the cell, functioning as both DNA demethylation intermediate activating TDG-initiated BER to replace the 5caC with a natural cytosine and as a stable epigenetic marker regulating RNA transcription and DNA replication by inducing transient stalling of the replicative polymerase. Pol  $\tau$  may fill an important niche role in the cell by being able to bypass 5caC, rescue stalled replication, and mark the 5caC for later removal via BER.

Figure 4: Schematic Diagram of the Methylation and Demethylation Pathway of the C5 Position of Cytosine

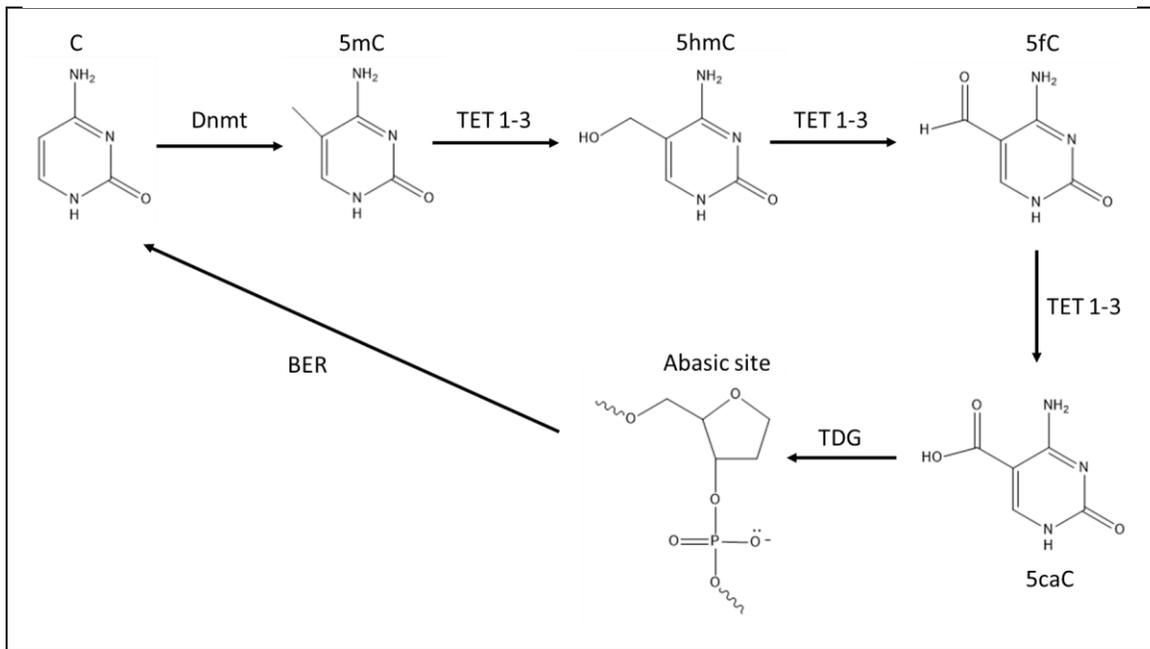


Figure 5: G:X and T:X Basepairs

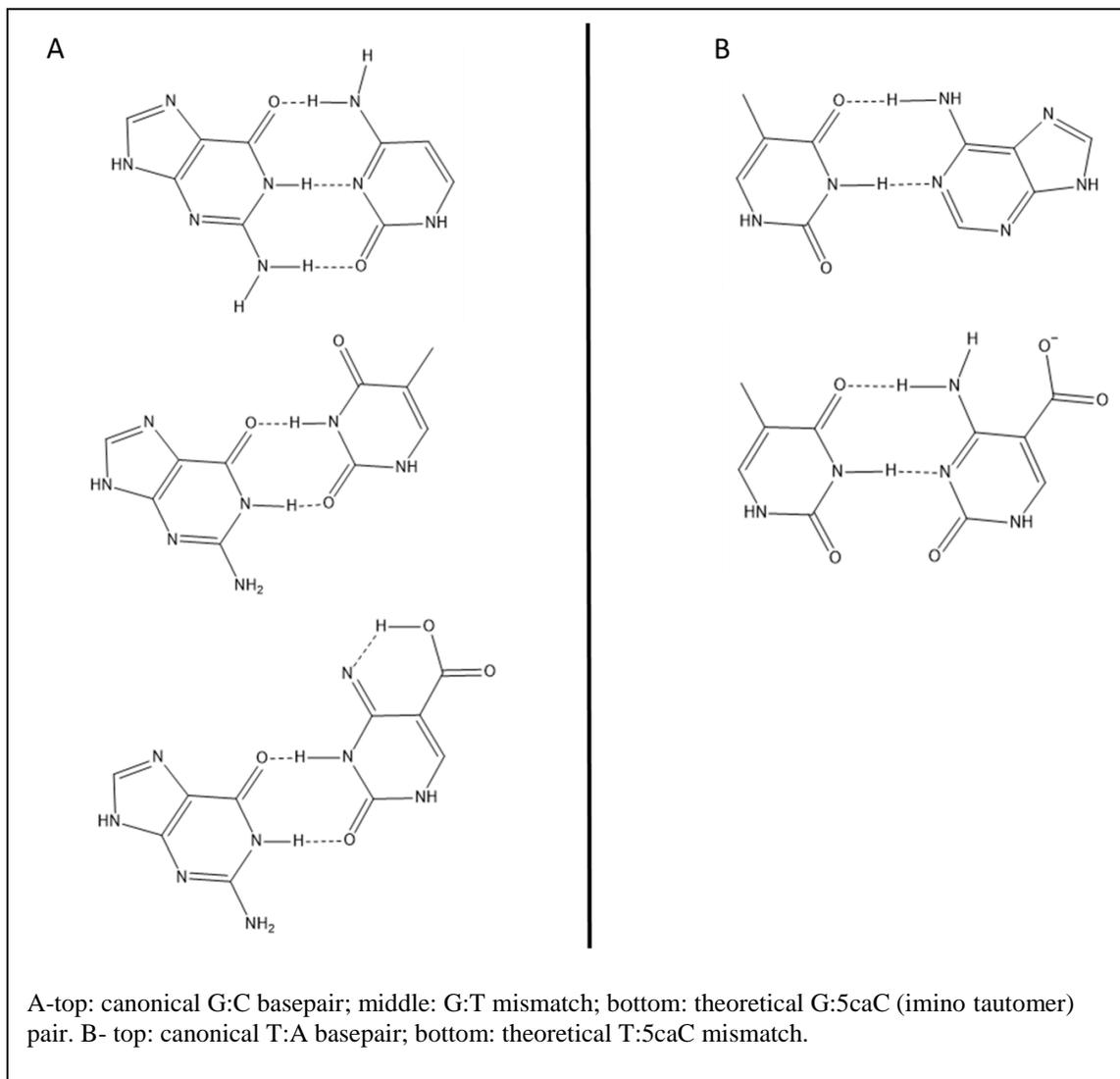
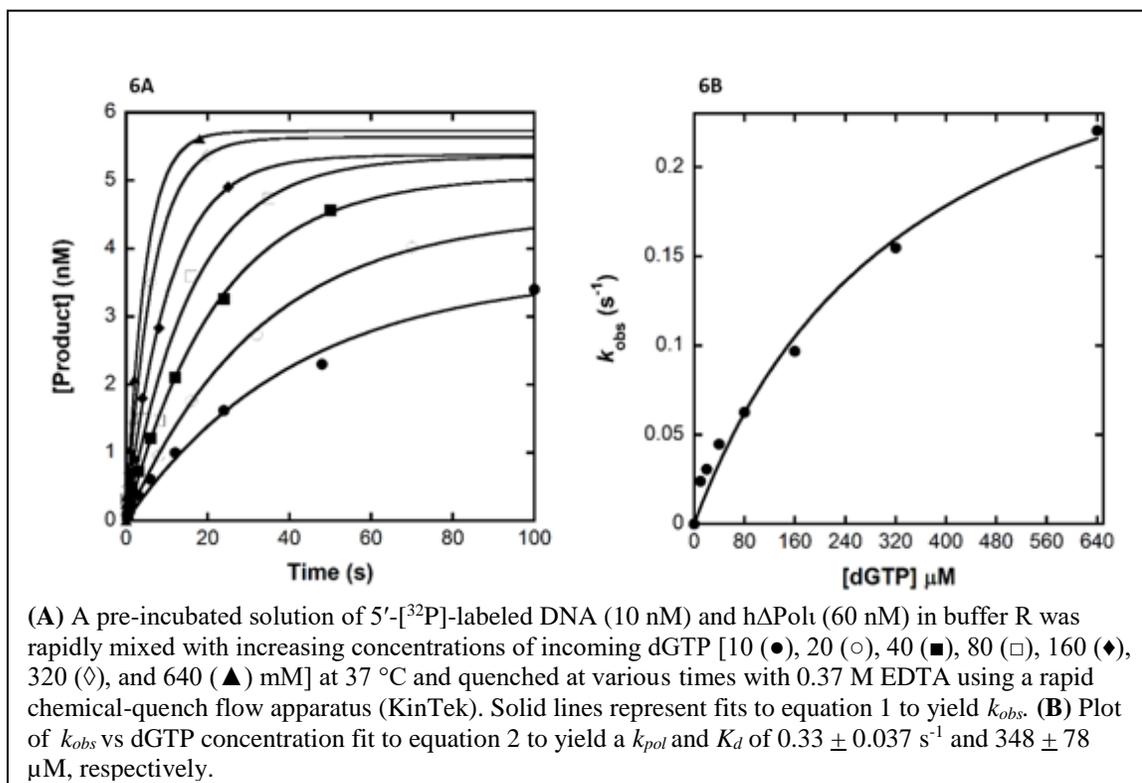


Figure 6: Pre-Steady State Kinetics of dGTP incorporation opposite 5caC



5' -pCCGGTGCC\*GAGG**C**GCTCAATTGGTCGTAGGACTATCCTCACCTCCACCGTTTCA-3'  
 3' -CGAGTTAACCAGCATCCTGATAG-5'

The 5'-p indicates a phosphate group. The bolded templating cytosine (C) indicates the site of the natural or modified C (C, 5mC, 5hmC, 5fC, 5caC). C\* indicates the presence of a second 5fC site in the template, only present when the first modification is also a 5fC. This second 5fC site does not affect the kinetics in our experiments since only single-nucleotide incorporation assays were performed.

Table 1: DNA Substrates for dNTP Incorporation by Polymerase  $\iota$  Opposite Modified C

Table 2: Pre-Steady State Kinetic Parameters for dNTP Incorporation by Truncated Human DNA Polymerase  $\epsilon$  Opposite Modified C

dNTP	$k_{pol}$ ( $s^{-1}$ )	$K_d$ ( $\mu M$ )	$k_{pol}/K_d$ ( $\mu M^{-1}s^{-1}$ )	Efficiency ratio <sup>a</sup>	Fidelity <sup>b</sup>	Probability % <sup>c</sup>
23/54-mer-C						
dGTP	$0.22 \pm 0.014$	$37 \pm 9$	$5.9 \times 10^{-3}$	-	-	-
23/54-mer-5mC						
dGTP	$0.23 \pm 0.0125$	$40 \pm 11$	$5.8 \times 10^{-3}$	-	-	61.6
dATP	$0.58 \pm 0.10$	$776 \pm 230$	$7.5 \times 10^{-4}$	7.7	$1.1 \times 10^{-1}$	8.0
dTTP	NA	NA	$9.6 \times 10^{-4}$	6.0	$1.4 \times 10^{-1}$	10.2
dCTP	$0.023 \pm 0.0013$	$119 \pm 25$	$1.9 \times 10^{-3}$	3.1	$2.5 \times 10^{-1}$	20.2
23/54-mer-5hmC						
dGTP	$0.53 \pm 0.064$	$162 \pm 42$	$3.3 \times 10^{-3}$	-	-	56.0
dATP	$0.25 \pm 0.010$	$192 \pm 22$	$1.3 \times 10^{-3}$	2.5	$2.8 \times 10^{-1}$	22.1
dTTP	NA	NA	$1.1 \times 10^{-3}$	3.0	$2.5 \times 10^{-1}$	18.7
dCTP	$0.21 \pm 0.033$	$1130 \pm 280$	$1.9 \times 10^{-4}$	17.4	$5.4 \times 10^{-2}$	3.2
23/54-mer-5fC						
dGTP	$0.29 \pm 0.038$	$112 \pm 43$	$2.6 \times 10^{-3}$	-	-	35.1
dATP	$0.24 \pm 0.039$	$211 \pm 94$	$1.1 \times 10^{-3}$	2.3	$3.0 \times 10^{-1}$	14.9
dTTP	$0.32 \pm 0.083$	$158 \pm 89$	$2.0 \times 10^{-3}$	1.3	$4.3 \times 10^{-1}$	27.0
dCTP	$0.092 \pm 0.0089$	$53 \pm 16$	$1.7 \times 10^{-3}$	1.5	$4.0 \times 10^{-1}$	23.0
23/54-mer-5caC						
dGTP	$0.33 \pm 0.037$	$348 \pm 78$	$9.5 \times 10^{-4}$	-	-	32.5
dATP	$0.14 \pm 0.017$	$226 \pm 56$	$6.2 \times 10^{-4}$	1.5	$4.0 \times 10^{-1}$	21.1
dTTP	$1.0 \pm 0.20$	$799 \pm 210$	$1.3 \times 10^{-3}$	0.7	$5.8 \times 10^{-1}$	44.3
dCTP	$0.012 \pm 0.0012$	$190 \pm 57$	$6.3 \times 10^{-5}$	15.1	$6.2 \times 10^{-2}$	2.1
<sup>a</sup> Efficiency ratio is calculated as $(k_{pol}/K_d)_{correct}/(k_{pol}/K_d)_{incorrect}$ . <sup>b</sup> Fidelity is calculated as $(k_{pol}/K_d)_{incorrect}/[(k_{pol}/K_d)_{correct} + (k_{pol}/K_d)_{incorrect}]$ . <sup>c</sup> Probability is calculated as $[(k_{pol}/K_d)_{NTP}/(\sum(k_{pol}/K_d)_{all\ NTPs})] \times 100\%$						

## Chapter 4: Crystallization of p12/p17 Complex of Polymerase Epsilon for Determination of Structure

### 4.1 Introduction

In eukaryotic cells, leading strand synthesis during DNA replication is carried out by pol  $\epsilon$ , a B-family DNA polymerase that shares sequence homology to the catalytic subunit of pol II in *E. coli*.<sup>97</sup> Beyond its major role in genome replication, pol  $\epsilon$  has been implicated in many other processes in the cell, such as cell cycle regulation, chromatin state regulation, gene regulation, and DNA damage repair.<sup>97-100</sup> Human pol  $\epsilon$  contains four subunits: p261, the catalytic subunit, and three accessory subunits, p59, p17, and p12. The N-terminal half of the catalytic subunit is highly conserved between yeast and humans (63% sequence identity), and contains both the polymerase and the exonuclease active sites. The catalytic domain shares the general subdomain structure of all other DNA polymerases, consisting of palm, fingers, and thumb subdomains folding to form the polymerase active site.<sup>97</sup> The other three subunits have largely unspecified roles in the function of the holoenzyme, although these subunits are implicated in protein-protein interactions, stabilization of the holoenzyme, and regulation of DNA binding and the cell cycle.<sup>97,98,100</sup> For example, it is thought that the p59 subunit interacts with both the C-terminal portion of p261 and PCNA to form a stable pol  $\epsilon$ •PCNA complex to keep pol  $\epsilon$

tethered to the DNA.<sup>97</sup> To date, the only structural study of human pol  $\epsilon$  is a solution structure of a 75-amino acid N-terminal segment of the p59 subunit.<sup>101</sup> In the past few years, two pol•DNA•dNTP ternary structures of the catalytic domain of the yeast pol  $\epsilon$  homolog, Pol2, and a low-resolution cryo-EM structure of yeast pol  $\epsilon$  holoenzyme were solved.<sup>102,103</sup> The crystal structure of yeast pol  $\epsilon$  shows some interesting differences compared to other polymerases, including insertions in the palm domain and an extra helix located above the thumb domain. The insertions in the palm domain constitute a novel protein fold not previously identified in a B family polymerase, termed the P domain (Figure 7). The P domain can encircle double-stranded DNA and has been shown to be essential for the processivity of the catalytic core of pol  $\epsilon$ . This unique domain structure may be a factor in conferring pol  $\epsilon$  with higher intrinsic processivity than other replicative pols. Furthermore, this domain has been shown to be essential to the polymerization activity of pol  $\epsilon$  and has a role in recognizing incorrect base pairs. This domain also exhibits a metal-binding site that likely coordinates a  $Zn^{2+}$  ion.<sup>103</sup> The cryo-EM structure of the holoenzyme indicates that there is an extended tail domain which is comprised of the three non-catalytic subunits and that the two smallest subunits (homologs to human subunits p12 and p17) are present in an extended conformation (Figure 8).<sup>102</sup> Interestingly, the link between the two largest subunits appears to be quite flexible, giving Dpb2 (the yeast homolog of p59) a large amount of mobility, especially in the absence of the smallest two subunits.<sup>102</sup> However, none of these characteristics have yet to be identified in human pol  $\epsilon$  due to the lack of structural data.

While the yeast structures of pol  $\epsilon$  are very informative, we anticipated identifying vital differences for human pol  $\epsilon$  including altered subunit structure and architecture, which likely lend to the complex functions of the enzyme in the eukaryotic system.

## 4.2 Materials and Methods

### 4.2.1 Materials

The materials for this project were purchased from the following companies: JCSG+ crystallization screen kit from Molecular Dimensions (Maumee, OH) and all chemicals used for hanging- drop crystallization conditions from Sigma-Aldrich (St. Louis, MO) at  $\geq 99\%$  purity.

### 4.2.2 Expression and Purification of Human Polymerase $\epsilon$ Subunits p12 and p17

The cDNAs for the p12 and p17 subunits were cloned into 11A vectors (AmpR) by Dr. Alan Tomkinson (New Mexico). Using multiple rounds of PCR, restriction digests, and ligation, three tags that were hindering expression were cloned out of the original p12 vector and a 6X-His tag was cloned in. The new p12 sequence was cloned into a pFastBac1 vector, as was the p17 sequence.

The pFastBac1 vectors encoding for the 6X-His p12 and untagged p17 subunits were each transformed into DH10Bac *E. coli* cells, the bacmid DNA was prepped and transfected into Sf9 insect cells, and the recombinant viruses were produced according to the Bac-to-Bac Baculovirus Expression System (ThermoFisher). After the second generation of virus expression, insect cell media was co-infected with the p12 and p17

viruses. The cultures were grown for 65 hrs at 26 °C and 114 RPM before centrifuging at 2,000 RPM. The cell pellets were resuspended in nickel wash buffer (20 mM Tris-Cl, pH 8.0 at 4 °C, 200 mM NaCl, 0.1% BME, 10 mM imidazole, 5% glycerol), mixed with 1 Roche protease inhibitor tablet and 1 mM PMSF, and lysed by passing through a French press 3x at 10,000-15,000 PSI. The lysate was cleared by ultracentrifugation at 40k RPM for 40 min at 4 °C and the supernatant filtered using 45-micron syringe filters. The cleared lysate was run over a nickel affinity column and washed with 10 CV nickel wash buffer and eluted with a gradient to 100% nickel elution buffer (nickel wash buffer + 500 mM imidazole) over 20 CV. The fractions were processed using 12% SDS-PAGE to locate protein. The sample-containing fractions were pooled and dialyzed against 1 L of heparin wash buffer (20 mM Tris-Cl, pH 8.0, 200 mM NaCl, 0.1 % BME, 5% glycerol) before being processed by heparin affinity chromatography. The protein was applied to a 5 mL heparin HiTrap HP column and washed with heparin wash buffer for 3 CV before being eluted with a gradient to 100% heparin elution buffer (heparin wash buffer + 1 M NaCl) over 6 CV. A 12% SDS-PAGE was run to locate protein-containing fractions, which were pooled and concentrated to 0.5 mL before applying to a Superdex 200 size exclusion column. The protein was eluted in size exclusion buffer (20 mM Tris-Cl, pH 8.0, 200 mM NaCl, 1 mM DTT, 5% glycerol). The purity was checked using 12% SDS-PAGE and the concentration determined using UV-Vis.

#### *4.2.3 Crystallization*

The protein solution for Hp12/p17 that was purified and concentrated (2.2.2) needed no additional modifications before setting up crystal trays. Initial crystallization

conditions were screened using 96-well sitting drop trays with the common commercially available screen JCSG+ using the Mosquito Crystallization Robot available in the Campus Chemical Instrument Center at The Ohio State University. The reservoir solution and p12/p17 complex (17 mg/mL) were mixed in 1:1 and 2:1 ratios by volume. Appropriate conditions based on promising results were expanded upon in 24-well hanging drop trays.

The hanging drop trays were set up by hand. The reservoir solution in each well was 1 mL total volume. The hanging drops were formed by mixing the reservoir solution and p12/p17 complex in a 1:1 or 1:2 volume ratio. Most crystal trays were set up and grown at 25 °C, although one was grown at 19 °C and one at 4 °C.

Crystals for all proteins and protein complexes were harvested by using a mounted cryo-loop, flash-frozen, and shipped to the Advanced Photon Source, LRL-CAT beamline (Argonne National Lab, Lemont, IL) for X-ray diffraction.

#### 4.3 Results

Crystals were obtained in several different conditions. From the 96-well screening tray using the JCSG+ kit, crystals were seen in various conditions approximately 2 weeks after setting up. There was a mix of rod and plate crystals, as well as some abnormal crystal shapes. The first shipment to APS had crystals from conditions in the JCSG+ tray in addition to a crystal from a unique condition I had set up. (Table 3). No crystals diffracted well enough to give meaningful data or ice rings were seen. Ice rings indicate ice crystals have formed over the protein crystal. They were most likely due to the lack of cryoprotectant used before flash-freezing the harvested crystals. An SDS PAGE was run

on smaller crystals from the same conditions and presence of the p12/p17 complex was confirmed.

A second round of crystals was harvested and sent to APS from self-made conditions as well as crystals that had grown in the JCSG+ tray almost 2 years after initial setup (Table 3). This time, each crystal was soaked in cryoprotectant 10-30 seconds prior to being flash-frozen to prevent ice from forming on the surface of the protein crystal. However, despite the use of cryoprotectant, many crystals showed ice rings or too little diffraction to receive any useful data (Figure 9).

#### 4.4 Discussion

The p12 and p17 proteins can be expressed and purified to sufficient quantities and high purity using the baculovirus expression system in insect cells. The two subunits also interact tightly and are very stable, which is important to allowing well-packed crystal growth. The crystallization was the bottleneck of this project, with crystal growth seen anywhere from 2 weeks to 2 years after conditions were set up. Ultimately, even nicely formed crystals soaked in cryoprotectant did not diffract well enough to give any usable data.

Several techniques can be used to grow difficult crystals. Lower temperatures cause crystals to grow more slowly, thus increasing the chances for symmetrical, well-packed crystals to form, which will diffract better. This technique was used to set up crystal trays at 19 and 4 °C, to no avail. Streak seeding can be done by taking crystals too small to be used to diffract from one growth condition and streaking them into another

condition to provide micronuclei “seeds” about which larger crystals can grow. This technique was used in trays at 25, 19, and 4 °C but ultimately yielded no results.

More drastic measures can be taken to try to get high-quality crystals, such as cleaving the 6X-His tag of the p12 subunit by incorporating a TEV cleavage sequence between the tag sequence and protein sequence in the vector containing the hp12 sequence; the 6X-His tag could be too flexible to allow symmetric packing as crystals grow.

Figure 7: Unique P-Subdomain of Polymerase  $\epsilon$

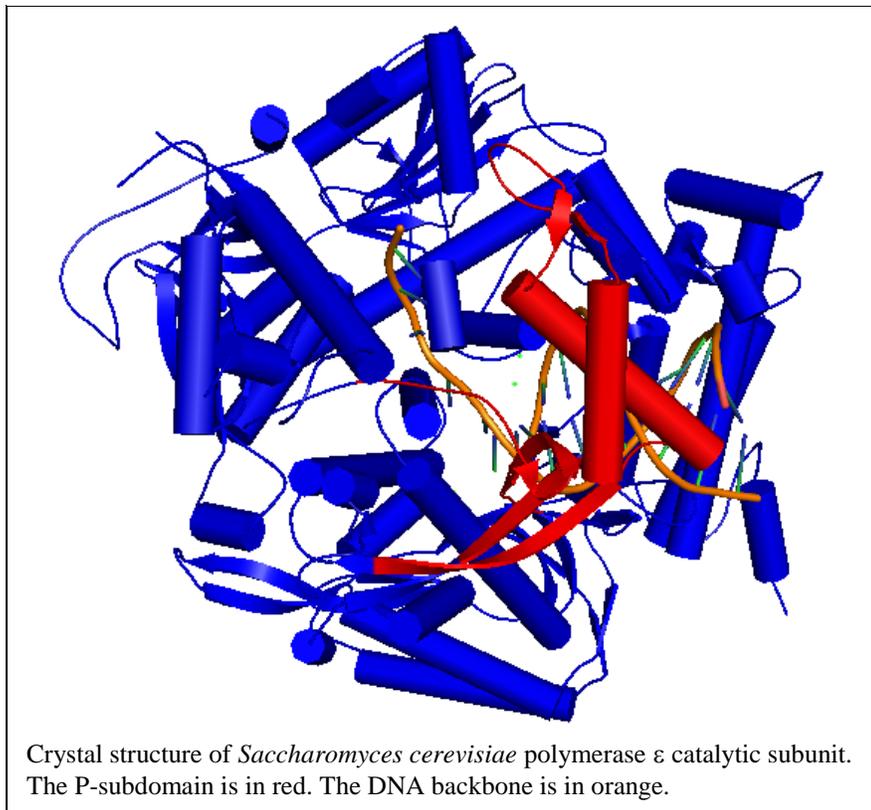


Figure 8: Cryo-EM Structure of *Saccharomyces cerevisiae* Polymerase  $\epsilon$  Holoenzyme

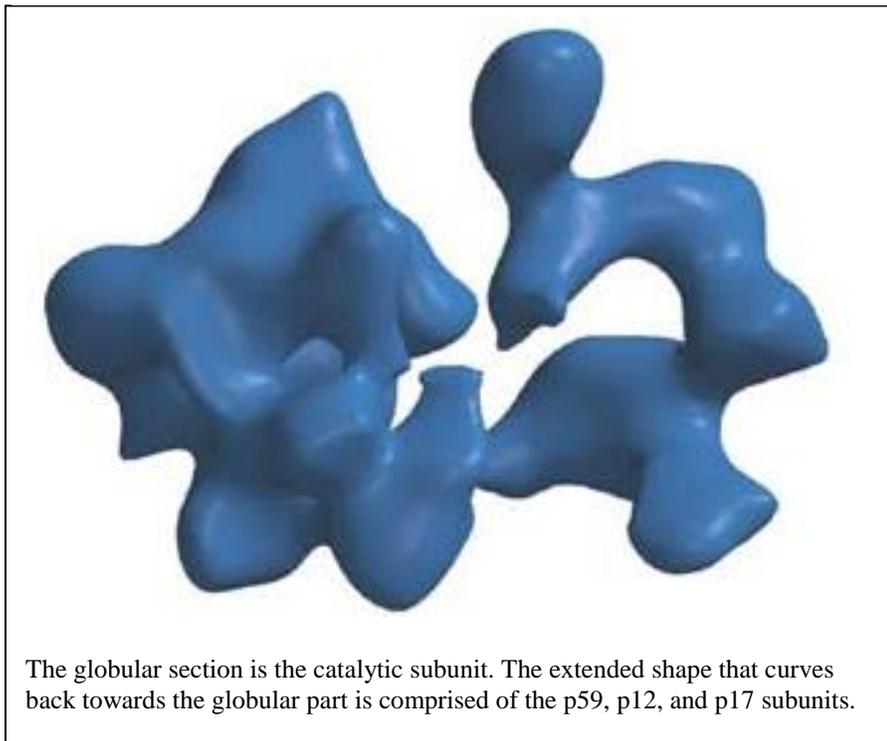


Figure 9: Photos of p12/p17 Crystals

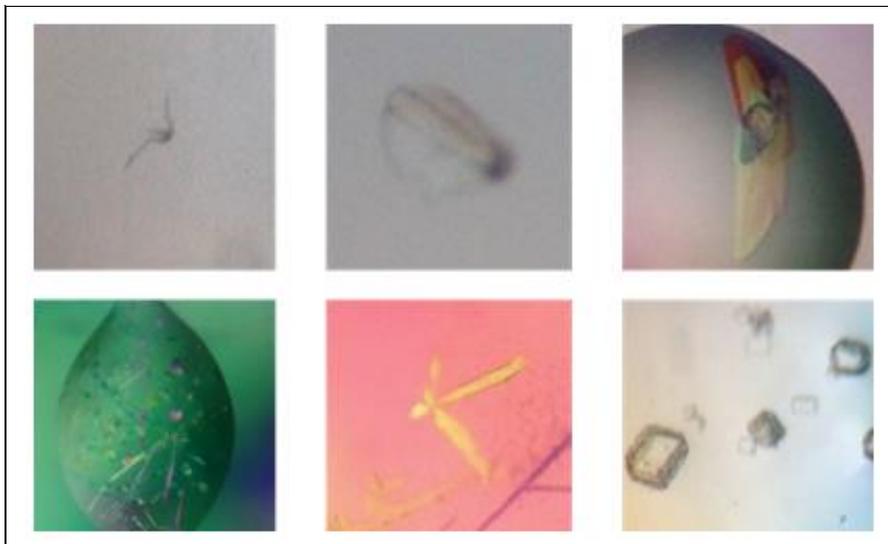


Table 3: Conditions that Yielded Crystal Growth for p12/p17

Reservoir Solution	Reservoir Solution:Protein (volume)	Time Passed Until Growth Seen
10% w/v PEG 3350, 10% w/v PEG 8000, 24% w/v PEG 1500, 20% v/v glycerol	1:1	1 week
100 mM HEPES, pH 7.0, 30% w/v Jeffamine ED-2003	1:1	2 weeks
100 mM Bis-Tris, pH 5.5, 300 mM magnesium formate dehydrate	-	1 week
100 mM Bis-Tris, pH 5.5, 200 mM NaCl, 25% w/v PEG 3350	1:1	1 week
100 mM Bis-Tris, pH 5.5, 200 mM ammonium acetate, 25% w/v PEG 3350	1:1	10 days
100 mM Bis-Tris, pH 7.0, 200 mM ammonium sulfate, 46% w/v PEG 3350	1:2	3 days
100 mM Bis-Tris, pH 7.0, 100 mM NaCl, 80% w/v PEG 3350	1:2	2 months
100 mM sodium acetate, pH 4.5, 200 mM lithium sulfate, 50% w/v PEG 400	1:1, 1:2	18 months
100 mM Tris, pH 8.5, 200 mM lithium sulfate, 40% w/v PEG 400	1:2	18 months

## Chapter 5: The Effect of Divalent Metal Ions on the Activity of Human PrimPol

### 5.1 Individual Contribution

This project resulted in a published manuscript.<sup>104</sup> I performed many of the single-turnover assays as well as carrying out all the associated product quantification and data analysis. I shared the responsibility of radiolabeling primers, annealing substrates. I also performed fluorescence anisotropy assays to determine the DNA binding affinity of PrimPol, though I did not perform the data analysis of those. I created a table of kinetic parameters and updated it as the project progressed. For the manuscript, I contributed writing to the “Introduction” and “Materials and Methods” sections, as well as formatting of the original manuscript for *JACS*, though the manuscript was eventually re-formatted by the first author for publication in *DNA Repair*. The information contained in this chapter is drawn largely from the published manuscript, but the writing is my own.

### 5.2 Introduction

PrimPol is relatively newly discovered enzyme (2013).<sup>105</sup> It is the second enzyme with pol activity discovered in the mitochondria, the first being pol  $\gamma$ . It is also the second primase discovered in the human body besides pol  $\alpha$ -primase.<sup>105-108</sup> These two characteristics make PrimPol a unique and intriguing enzyme. The specific role of

PrimPol in the human cell and its mechanism of activity are areas of research still being explored.

PrimPol has two enzymatic activities and two domains. First of all, PrimPol can prime single-stranded DNA. Unique to it, it can utilize dNTPs to prime as well as rNTPs, while pol  $\alpha$ -primase and RNA pols can only use rNTPs. In fact, PrimPol prefers to prime using dNTPs. Secondly, PrimPol can extend primers via its nucleotidyl transfer activity.<sup>105,106,109</sup> Both the activities of PrimPol are highly activated by  $Mn^{2+}$  compared to  $Mg^{2+}$ . The enzyme is more responsive to as little as 10  $\mu M$   $Mn^{2+}$  than it is to a physiological concentration of 10 mM  $Mg^{2+}$ .<sup>105,110–112</sup> PrimPol contains an N-terminal archaeo-eukaryotic primase (AEP) domain as well as a C-terminal UL52-like zinc finger subdomain, both of which are essential to the primase activity while only the N-terminal domain is needed for pol activity.<sup>105,109,113</sup>

PrimPol plays a role in mitochondrial DNA (mtDNA) synthesis and maintaining genomic integrity, as well as performing TLS, bypassing products of UV damage, and repriming stalled replication forks past damage sites or sites containing G4 quadruplexes in both the mitochondria and nucleus.<sup>105,108–110,112,114–118</sup> It has been shown that in PrimPol<sup>-/-</sup> mice, mtDNA synthesis is severely arrested, though no obvious phenotypical difference were observed.<sup>105</sup> Additionally, PrimPol can bypass very common types of oxidative damage such as abasic sites and 8-oxo-guanine.<sup>112</sup> However, the main replicative pol in the mitochondria, pol  $\gamma$ , can bypass 8-oxo-guanine better than PrimPol, indicating that this bypass capability may not be the most useful aspect of PrimPol. PrimPol can also bypass typical UV damage like 6-4 photoproducts and cyclopyrimidine

dimers.<sup>106,116</sup> Mammalian and avian cells lacking expression of PrimPol are more sensitive to UV radiation and chain-terminating nucleoside analogs and display more breaks in chromatin and slower replication fork progress.<sup>106,111,116,118</sup> Replication protein A (RPA) and PrimPol interact to target PrimPol to chromatin following UV radiation damage, whereupon RPA stimulates PrimPol's primase activity.<sup>108,114</sup> The cumulation of evidence suggests that PrimPol's main use is to reprime downstream of stalled replication forks. Although cellular and live mice studies suggest PrimPol is not *essential* to mammalian cell life, it *is* essential for maintaining optimum genetic integrity and cellular health.

Interestingly, as mentioned briefly above, the activity of PrimPol is more sensitive to  $Mn^{2+}$  than  $Mg^{2+}$ .<sup>105,110-112</sup> While pols can use  $Mg^{2+}$  as the metal ion cofactor to ensure enzymatic activity and faithful replication, they can also use  $Mn^{2+}$ , which usually serves to increase pol activity but also decrease fidelity. For this reason and the fact that physiological  $Mg^{2+}$  concentrations are higher than that of  $Mn^{2+}$ , it is thought that the typical *in vivo* cofactor for pols is  $Mg^{2+}$ . However, since PrimPol is so sensitive to even small amounts of  $Mn^{2+}$  and because it primes and extends relatively short lengths of DNA, there is a possibility that  $Mn^{2+}$  may be used *in vivo* as the preferred metal ion cofactor. We determined the differential effects of these two metal ions on the ability of PrimPol to bind DNA and incorporate dNTPs using pre-steady state kinetics and fluorescence anisotropy assays. We also used pre-steady state kinetics to characterize the ability of PrimPol to incorporate chain-terminating nucleoside analog drugs and discriminate against rNTPs.

## 5.3 Materials and Methods

### 5.3.1 Materials

Materials were purchased from the following companies: Optinkinase from USB Corp. (Cleveland, OH), [ $\gamma$ - $^{32}$ P]ATP from PerkinElmer Life Sciences (Boston, MA), dNTPs and rNTPs from Bioline (Taunton, MA), 2'-deoxy-2',2'-difluorodeoxycytidine 5'-triphosphate (GemCTP) and 2'-aracytidine 5'-triphosphate (AraCTP) from TriLink BioTechnologies (San Diego, CA), 5'-triphosphate of lamivudine ((-)-3TC-TP) and emtricitabine ((-)-FTC-TP) from Gilead Sciences (Foster City, CA), and oligonucleotides from IDT (Coralville, IA).

### 5.3.2 Radiolabeling and annealing DNA substrates:

The 21-mer and 22-mer substrates used for the single turnover assays were purified via denaturing PAGE and radiolabeled using [ $\gamma$ - $^{32}$ P]ATP and 2  $\mu$ L Optinkinase and heating to 37 °C for 3 h. The primers were annealed to 41-mer DNA templates (Table 4) in a 1:1.35 molar ratio by heating to 95 °C for 5 min and slowly cooling to room temperature. Concentration of substrates was verified using UV-Vis spectrometer at 260 nm.

### 5.3.3 Expression and Purification of Human PrimPol

Human PrimPol containing an N-terminal 6x-His tag was subcloned and transformed into *E. coli* Rosetta (DE3) competent cells. A starter culture of LB media containing 30  $\mu$ g/mL Kan and 34  $\mu$ g/mL Cam was grown overnight at 37 °C. When it reached OD<sub>600</sub> of 1.5, overexpression cultures were set up using 1 L flasks of LB media (30  $\mu$ g/mL Kan, 34  $\mu$ g/mL Cam) that were inoculated with 1% volume of the starter

culture. The overexpression cultures were grown at 37 °C to an OD<sub>600</sub> of 0.8 followed by rapid cooling on ice. After reaching an OD<sub>600</sub> of 1.0, the cultures were induced with 0.1 mM IPTG and grown at 16 °C for an additional 15 h. The cultures were then ultracentrifuged and the cell pellets were resuspended in buffer A (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 10% glycerol, 0.1% BME, 10 mM imidazole, 0.01 mM EDTA, and 0.1% IGEPAL). The mixture was supplemented with EDTA-free Protein Inhibitor Cocktail tablets (Roche) and 1 mM PMSF prior to lysis. Cells were lysed by passing through French press 3x at 15,000 psi. The lysate was cleared by ultracentrifugation at 40,000 rpm for 40 min. The soluble fraction was incubated with charged Ni-NTA resin for 3 h at 4 °C. The Ni-NTA beads were packed into a tricorn FPLC column and were washed with 20 CV of lysis buffer and further washed with 10 CV of 4% buffer B (buffer A + 500 mM imidazole). Protein was eluted with a linear gradient of 4–100% buffer B over 15 CV and fractions were analyzed by SDS-PAGE. Protein-containing fractions were pooled and loaded onto a HiTrap Heparin HP column (GE Healthcare). The column was washed with 5 CV buffer C (buffer A with no imidazole) followed by 10 CV of 10% buffer D (buffer C containing 1 M NaCl). The protein was eluted with a gradient of 10–100% buffer D over 10 CV. Stepwise dialysis was performed overnight at 4 °C to decrease the NaCl concentration of the eluted protein solution from 700 mM to 125 mM. The purest samples were pooled and concentrated to 500 µL using an Amicon Ultra-15 Centrifugal filter (Millipore). The protein sample was further purified in a Superdex 200 size exclusion chromatography column (GE Healthcare) to isolate full-length human PrimPol (66.5 kDa). Fractions were analyzed via SDS-PAGE and the purest samples

were pooled and dialyzed against storage buffer (50 mM Tris-HCl, pH 7.5, 400 mM NaCl, 50% glycerol, 1 mM DTT, and 0.1 mM EDTA). The final concentration of purified human PrimPol was determined using UV-vis spectroscopy at 280 nm and the calculated extinction coefficient of  $\epsilon_{280} = 77,655 \text{ M}^{-1}\text{cm}^{-1}$ .

#### 5.3.4 Single Turnover Polymerase Assays:

Pre-steady state, single turnover kinetic assays were performed as generally described in section 2.3.2. Specifically, PrimPol (300 nM) and 5'-[<sup>32</sup>P]-labeled DNA (30 nM) were pre-incubated at 37 °C for 5 min in reaction buffer (50 mM Tris-HCl [pH 7.8 at 37 °C], 50 mM NaCl, 0.1 mM EDTA, 5 mM DTT and 0.1 µg/mL BSA) before mixing with increasing concentrations of dNTPs. The reaction mixture (10 µL) was quenched at various time points with EDTA to a final concentration of 0.37 M. Reaction products were separated using denaturing PAGE, visualized using Typhoon TRIO (GE) and quantified using ImageQuant (GE). Each time course of product formation was fit to a single-exponential equation:

$$[\text{product}] = E_o \bullet [1 - \exp(-k_{obs}t)]$$

using Kaleidagraph (Synergy Software) to yield a reaction amplitude ( $E_o$ ) 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:

$$k_{obs} = k_{pol}[dNTP]/([dNTP] + K_d)$$

to yield an apparent equilibrium dissociation constant of enzyme for the nucleotide ( $K_d$ ) and a maximum nucleotide incorporation rate constant ( $k_{pol}$ ).

#### 5.3.5 Fluorescence Anisotropy Assays

The Cy3-labeled DNA substrate 17/30-mer (Table 4) was titrated with increasing amounts of human PrimPol and the anisotropy was monitored using a FluoroMax-4 (Horiba). Assays were carried out at 25 °C in titration buffer (50 mM Tris–HCl (pH 7.5 at 25 °C), 50 mM NaCl, 0.01 mM EDTA) either without divalent metal ions or in the presence of 5 mM MnCl<sub>2</sub> or 5 mM MgCl<sub>2</sub> where indicated. Excitation and emission for the Cy3 fluorophore were set to 540 and 568 nm, respectively, with a 10 nm slit width and 2 s integration time. The data obtained from anisotropy measurements were fit to the following equation:

$$\Delta A = (\Delta A_T / 2D_0) \times \{ (K_d^{\text{DNA}} + D_0 + E_0) - [(K_d^{\text{DNA}} + D_0 + E_0)^2 - 4 E_0 D_0]^{1/2} \}$$

where  $\Delta A$  is the change in anisotropy,  $\Delta A_T$  is the maximum anisotropy change,  $D_0$  and  $E_0$  are the initial concentrations of DNA and PrimPol, respectively, and  $K_d^{\text{DNA}}$  is the equilibrium dissociation constant of the PrimPol•DNA binary complex.

#### 5.4 Results

We determined the differential effect of metal ion cofactors on the DNA binding affinity and dNTP incorporation kinetics of PrimPol. Using fluorescence anisotropy assays, we demonstrated that PrimPol binds to a primer/template DNA substrate with 29 nM affinity in the presence of Mn<sup>2+</sup> and 979 nM affinity in the presence of Mg<sup>2+</sup>, thus showing that PrimPol binds DNA with almost 34-fold higher affinity with Mn<sup>2+</sup> compared to Mg<sup>2+</sup>. This difference could be due to the Mg<sup>2+</sup> ion competing with the binding site of Zn<sup>2+</sup> or generally perturbing the interaction between Zn<sup>2+</sup> and the coordinating residues in the zinc finger of the UL52-like domain of PrimPol. This domain is necessary for PrimPol's primase activity and interacts with the single-stranded

template of a DNA substrate. In the absence of any metal ion cofactor, human PrimPol binds the substrate with 41 nM affinity. By employing pre-steady state kinetic techniques, namely, single turnover assays with a primer/template DNA substrate, we were able to determine the  $K_d^{dNTP}$  and  $k_{pol}$  of PrimPol for all 16 possible dNTP incorporation possibilities in the presence of 5 mM  $Mn^{2+}$  (Table 5). For correct dNTP incorporations, the  $K_d^{dNTP}$  is 11-17  $\mu M$ , the  $k_{pol}$  is 0.036-0.096  $s^{-1}$ , and a substrate specificity (also known as the catalytic efficiency,  $k_p/K_d^{dNTP}$ ) of  $2.3-6.0 \times 10^{-3} \mu M^{-1} s^{-1}$ . In the presence of 5 mM  $Mn^{2+}$ , PrimPol incorporates correct dNTPs with a  $K_d^{dNTP}$  of 262-895  $\mu M$ , a  $k_{pol}$  of 0.010-0.020  $s^{-1}$  and a specificity of  $2.2-5 \times 10^{-5} \mu M^{-1} s^{-1}$  (Table 6). Therefore, the efficiency is about 100-fold lower in the presence of  $Mg^{2+}$  and this is mostly due to the approximately 100-fold increase of the  $K_d^{dNTP}$ , since the  $k_{pol}$  rates in the presence of each metal ion are comparable. By comparing the  $k_{pol}/K_d^{dNTP}$  of correct incorporation to incorrect incorporations for each templating base, we determined the fidelity of PrimPol in the presence of 5 mM  $Mn^{2+}$  to be  $3.4 \times 10^{-2}$  to  $3.8 \times 10^{-1}$ . We also found the fidelity for templating base dT in the presence of 5 mM  $Mg^{2+}$  to be  $6.1 \times 10^{-4}$  to  $1.1 \times 10^{-2}$ . The data indicate that PrimPol is a relatively error-prone pol, especially in the presence of  $Mn^{2+}$ .

We also determined the sugar selectivity of human PrimPol by using single turnover assays to find the pre-steady state kinetic parameters for rNTP incorporation in the presence of either 5 mM  $Mn^{2+}$  or  $Mg^{2+}$ . By taking the ratio of catalytic efficiencies of the correct dNTP to matching rNTP, we could calculate the selectivity of PrimPol against rNTPs. We found the sugar selectivity to be 57-1800 and 150-4500 with  $Mn^{2+}$  and  $Mg^{2+}$ ,

respectively (Tables 5 and 6). This selectivity against rNTPs is due to both a decrease in  $k_{pol}$  and an increase in  $K_d^{rNTP}$  relative to the matching dNTP.

Finally, chain-terminating cytidine analog drugs commonly used as chemotherapeutic agents or antivirals were used in single turnover kinetic assays with PrimPol in the presence of either  $Mg^{2+}$  or  $Mn^{2+}$  to see if they could be incorporated into the primer. PrimPol could incorporate GemCTP and AraCTP, which are used as chemotherapeutic agents in certain cancers, with either metal ion. However, when (-)3TC-TP or (-)FTC-TP were used—cytidine analogs with *L*-stereochemistry used as antiviral medication against HIV and Hepatitis B—they could only be incorporated if  $Mn^{2+}$  was present, and so inefficiently that no individual pre-steady state parameters could be determined. Only the  $k_{pol}/K_d$  could be determined. No incorporation was detected when  $Mg^{2+}$  was used in the assays. (Table 7).

## 5.5 Discussion

We used single turnover kinetics and fluorescence anisotropy to investigate the kinetics of PrimPol incorporation of correct and incorrect dNTPs, its inherent sugar selectivity, and whether it can incorporate certain important chain-terminating cytidine analogue drugs, all in the presence of either  $Mn^{2+}$  or  $Mg^{2+}$ . As expected, we found that the choice of metal ion used in the assay changed the kinetics of incorporation and the ability of PrimPol to incorporate some substrates, though the disparity in catalytic activities is heightened compared to other pols.

It is possible that PrimPol uses  $Mn^{2+}$  as its cellular cofactor instead of  $Mg^{2+}$ , at least in the mitochondria. While this is opposite what other pols use *in vivo*, there are several lines of evidence that make this switch be more plausible for PrimPol specifically. PrimPol has about 100-fold higher efficiency for incorporating correct dNTPs and binds DNA with 34-fold higher affinity using  $Mn^{2+}$  compared to  $Mg^{2+}$ . *In vivo*, PrimPol may not even be able to access DNA if it uses  $Mg^{2+}$  because its DNA binding affinity is too low compared to other DNA pols in the presence of  $Mg^{2+}$ , which would consequently mean PrimPol is outcompeted for binding the DNA. Even if PrimPol could bind the DNA while using  $Mg^{2+}$ , it would not be able to prime or bypass efficiently enough to have a biologically relevant role. Physiological concentration of  $Mg^{2+}$  in most mammalian cell types is between 15-20 mM, and PrimPol has no observable pol activity past 10 mM  $Mg^{2+}$ . Its primase activity is also nearly nonexistent when using  $Mg^{2+}$  at 10 mM.<sup>105,119,120</sup> However, if we look at the activity of PrimPol in the presence of  $Mn^{2+}$ , it becomes a much more viable player in maintaining genomic stability, bypassing DNA lesions, and repriming the replication fork. With  $Mn^{2+}$  concentration as low as 10  $\mu$ M, it has comparable activity to other specialized pols and its primase activity is stimulated.<sup>105</sup>  $Mn^{2+}$  is a known neurotoxin at 60-160  $\mu$ M in brain cells, but normal  $Mn^{2+}$  concentration in the brain is kept between 20-53  $\mu$ M, and this range is likely similar in other cell types.<sup>121,122</sup>  $Mn^{2+}$  is also concentrated to the mitochondria in the cell by the  $Ca^{2+}$  uniporter.<sup>123</sup> *In vivo*, PrimPol is not processive, synthesizing only 2-12 nucleotides at a time during (re)priming and extension. This characteristic probably helps offset the fact that PrimPol has a fidelity on par with Y-family TLS pols in the presence of  $Mg^{2+}$  and

even lower fidelity in the presence of  $Mn^{2+}$ . With all data taken into consideration, it is possible that  $Mn^{2+}$  is the physiologically relevant metal ion cofactor, even though it makes PrimPol moderately less faithful.

PrimPol is also a little choosy when it comes to nucleosides, preferring to incorporate dNTPs rather than rNTPs. This phenomenon is common for most DNA pols but not for primases. The preference was demonstrated when using either  $Mg^{2+}$  (150-4500) or  $Mn^{2+}$  (57-1800). Selectivity could be due to steric clash between the 2'-hydroxyl group of the incoming rNTP and the backbone carbonyl of Asn289, which was shown in a crystal structure of PrimPol with rATP modeled to replace the native dATP in the active site of crystallized enzyme bound to DNA.<sup>113</sup>

Finally, we established that PrimPol can incorporate the chain-terminating cytidine analogue drugs used as chemo therapeutic agents, GemCTP and AraCTP. However, it showed no incorporation activity on the *L*-stereochemistry cytidine analogs used as antiviral medications, (-)3TC-TP or (-)FTC-TP.

Table 4: DNA Substrates for Pre-Steady Kinetics and DNA Binding Assays with PrimPol

D-1
5' -CGCAGCCGTCCAACCAACTCA-3'
3' -GCGTCGGCAGGTTGGTTGAGTAGCAGCTAGGTTACGGCAGG-5'
D-6
5' -CGCAGCCGTCCAACCAACTCA-3'
3' -GCGTCGGCAGGTTGGTTGAGTGGCAGCTAGGTTACGGCAGG-5'
D-7
5' -CGCAGCCGTCCAACCAACTCA-3'
3' -GCGTCGGCAGGTTGGTTGAGTTGCAGCTAGGTTACGGCAGG-5'
D-8
5' -CGCAGCCGTCCAACCAACTCA-3'
3' -GCGTCGGCAGGTTGGTTGAGTCGCAGCTAGGTTACGGCAGG-5'
Anisotropy Substrate (Cy3-labeled)
5' -GCCTCGCTGCCGTCGCC-3'
3' -CGGAGCGACGGCAGCGGTTTTTTTTTTTTTT-Cy3-5'

Table 5: Pre-Steady State Kinetic Parameters for dNTP and rNTP Incorporation by PrimPol in the Presence of 5 mM Mn<sup>2+</sup>

dNTP	$k_{pol}$ (s <sup>-1</sup> )	$K_d$ (μM)	$k_{pol}/K_d$ (μM <sup>-1</sup> s <sup>-1</sup> )	Fidelity <sup>a</sup>	Sugar Selectivity <sup>b</sup>
D-1					
dTTP	0.096 ± 0.004	17 ± 2	5.8 × 10 <sup>-3</sup>	–	
dATP*	0.0083 ± 0.0001	4.1 ± 0.3	2.0 × 10 <sup>-3</sup>	2.6 × 10 <sup>-1</sup>	
dCTP*	0.014 ± 0.0009	13 ± 4	1.1 × 10 <sup>-3</sup>	1.7 × 10 <sup>-1</sup>	
dGTP*	0.0047 ± 0.0007	13 ± 4	3.6 × 10 <sup>-4</sup>	6.0 × 10 <sup>-2</sup>	
UTP	0.0018 ± 0.0002	561 ± 160	3.2 × 10 <sup>-6</sup>	5.5 × 10 <sup>-4</sup>	1800
D-6					
dCTP*	0.060 ± 0.005	16 ± 4	3.8 × 10 <sup>-3</sup>	–	
dATP*	0.011 ± 0.00005	20 ± 3	5.5 × 10 <sup>-4</sup>	1.3 × 10 <sup>-1</sup>	
dGTP*	0.035 ± 0.003	27 ± 8	1.3 × 10 <sup>-3</sup>	2.6 × 10 <sup>-1</sup>	
dTTP*	0.0058 ± 0.0005	14 ± 6	4.2 × 10 <sup>-4</sup>	9.9 × 10 <sup>-2</sup>	
rCTP	0.0092 ± 0.0002	136 ± 7	6.7 × 10 <sup>-5</sup>	1.7 × 10 <sup>-2</sup>	57
D-7					
dATP	0.066 ± 0.003	11 ± 1	6.0 × 10 <sup>-3</sup>	–	
dCTP*	0.013 ± 0.002	17 ± 9	7.6 × 10 <sup>-4</sup>	1.1 × 10 <sup>-1</sup>	
dGTP	0.017 ± 0.001	22 ± 4	7.7 × 10 <sup>-4</sup>	1.1 × 10 <sup>-1</sup>	
dTTP*	0.0029 ± 0.0002	3 ± 1	1.0 × 10 <sup>-3</sup>	1.4 × 10 <sup>-1</sup>	
rATP	0.0047 ± 0.0006	744 ± 234	6.3 × 10 <sup>-6</sup>	1.0 × 10 <sup>-3</sup>	950
D-8					
dGTP	0.036 ± 0.002	16 ± 3	2.3 × 10 <sup>-3</sup>	–	
dATP	0.0049 ± 0.0003	16 ± 3	3.0 × 10 <sup>-4</sup>	1.2 × 10 <sup>-1</sup>	
dCTP*	0.014 ± 0.001	10 ± 3	1.4 × 10 <sup>-3</sup>	3.8 × 10 <sup>-1</sup>	
dTTP*	0.0024 ± 0.0004	30 ± 7	8.0 × 10 <sup>-5</sup>	3.4 × 10 <sup>-2</sup>	
rGTP	0.0015 ± 0.0001	110 ± 29	1.4 × 10 <sup>-5</sup>	6.1 × 10 <sup>-3</sup>	160
<sup>a</sup> Defined as $(k_{pol}/K_d)_{incorrect}/[(k_{pol}/K_d)_{correct} + (k_{pol}/K_d)_{incorrect}]$ .					
<sup>b</sup> Defined as $(k_{pol}/K_d)_{dNTP}/(k_{pol}/K_d)_{rNTP}$ .					
*Denotes data from experiments I carried out					

Table 6: Pre-Steady State Kinetic Parameters for dNTP Incorporation by PrimPol in the Presence of 5 mM Mg<sup>2+</sup>

dNTP	$k_{pol}$ (s <sup>-1</sup> )	$K_d$ (μM)	$k_{pol}/K_d$ (μM <sup>-1</sup> s <sup>-1</sup> )	Fidelity <sup>a</sup>
D-1				
dTTP	0.020 ± 0.002	526 ± 99	3.8 × 10 <sup>-5</sup>	
D-6				
dCTP	0.013 ± 0.0008	262 ± 55	5.0 × 10 <sup>-5</sup>	
D-7				
dATP	0.011 ± 0.0007	388 ± 65	2.8 × 10 <sup>-5</sup>	–
dCTP	Not determined	Not determined	3.1 × 10 <sup>-7</sup>	1.1 × 10 <sup>-2</sup>
dGTP	Not determined	Not determined	1.3 × 10 <sup>-8</sup>	4.6 × 10 <sup>-4</sup>
dTTP	Not determined	Not determined	1.7 × 10 <sup>-8</sup>	6.1 × 10 <sup>-4</sup>
D-8				
dGTP	0.020 ± 0.0006	895 ± 60	2.2 × 10 <sup>-5</sup>	
<sup>a</sup> Defined as $(k_{pol}/K_d)_{incorrect}/[(k_{pol}/K_d)_{correct} + (k_{pol}/K_d)_{incorrect}]$ .				

Table 7: Pre-Steady Kinetic Parameters for Cytosine Analog Incorporation by PrimPol in the Presence of Mg<sup>2+</sup> or Mn<sup>2+</sup>

NTP	$k_{pol}$ (s <sup>-1</sup> )	$K_d$ (μM)	$k_{pol}/K_d$ (μM <sup>-1</sup> s <sup>-1</sup> )	Discrimination <sup>a</sup>
In the presence of 5 mM Mn <sup>2+</sup>				
AraCTP	0.0057 ± 0.0004	21 ± 4	2.7 × 10 <sup>-4</sup>	14
GemCTP	0.058 ± 0.0006	45 ± 2	1.3 × 10 <sup>-3</sup>	2.9
(-)3TC-TP	Not determined	Not determined	1.2 × 10 <sup>-6</sup>	3200
(-)FTC-TP	Not determined	Not determined	9.0 × 10 <sup>-7</sup>	4200
In the presence of 5 mM Mg <sup>2+</sup>				
AraCTP	0.0076 ± 0.0003	316 ± 44	2.4 × 10 <sup>-5</sup>	2.1
GemCTP*	0.0071 ± 0.0006	1380 ± 255	5.1 × 10 <sup>-6</sup>	9.8
(-)3TC-TP	No observed incorporation			
(-)FTC-TP	No observed incorporation			
<sup>a</sup> Defined as $(k_{pol}/K_d)_{dCTP}/(k_{pol}/K_d)_{analog}$ with the $(k_{pol}/K_d)_{dCTP}$ values from tables 5 and 6.				
*Denotes data from experiments I carried out				

## Chapter 6: Other Projects

### 6.1 *D*-Stereoselectivity of Human Polymerase $\beta$

#### *6.1.1 My Contribution*

For this project, which resulted in a manuscript, I expressed and purified the R283A human pol  $\beta$  mutant and performed the pre-steady state kinetic assay for the R283A mutant with *L*-dCTP and did the quantification and data analysis.<sup>124</sup> I will briefly outline the methods I performed and the results I had a hand in.

#### *6.1.2 Materials and Methods*

**Expression and Purification of human pol  $\beta$  R282A mutant:** Human pol  $\beta$  R283A mutant was subcloned into a pETx plasmid (Kan resistant) and transformed into Rosetta cells. Cells were plated onto on LB-agar plates containing Kan and Cam at 34 and 50  $\mu\text{g}/\text{mL}$ , respectively. A colony was selected and grown in LB media (34  $\mu\text{g}/\text{ml}$  Kan, 50  $\mu\text{g}/\text{mL}$  Cam) overnight to  $\text{OD}_{600}$  of 1.8, shaking at 220 RPM. A secondary culture was grown from 1% overnight culture, again to  $\text{OD}_{600}$  of 1.8. The expression cultures (40  $\text{mg}/\text{L}$  Kan, 50  $\mu\text{g}/\text{L}$  Cam) were inoculated with 1% volume of the overnight culture and grown to  $\text{OD}_{600}$  of 0.6 at 37 °C and 220 RPM before inducing with 0.1 mM IPTG.

When the cultures reached OD<sub>600</sub> of 1.6, the cells were pelleted by centrifugation at 4 °C at 4,000 RPM for 25 minutes. The cell pellet was resuspended in lysis buffer (50 mM Tris, pH 7.5, 0.5 M NaCl, 5% glycerol, 0.1% beta- BME, 1 mM EDTA) and immediately combined with PMSF (1mM final) and 1 Roche protease inhibitor cocktail tablet and lysed by passing 3x in the French press at 15,000 PSI. The lysate was clarified by ultracentrifugation at 4 °C and 40,000 RPM for 40 minutes.

The cleared lysate was incubated with IMAC Fast flow resin charged with Ni<sup>2+</sup> overnight with gentle rocking at 4 °C. The resin was centrifuged at 2,500 RPM for 10 minutes to collect the beads and packed into a gravity column. The protein was washed using 15 CV of 6% buffer B (50mM Tris-Cl, pH 7.5, 0.5M NaCl, 5% glycerol, 300 mM Imidazole, and 0.1% BME) and eluted using a stepwise gradient from 50% to 100% buffer B. Fractions containing human pol β, based on SDS-PAGE, were combined and dialyzed in dialysis buffer (50 mM Tris-Cl, pH 7.5 at 4°C, 100 mM NaCl, 5% glycerol, 0.1% BME). Dialyzed samples were loaded onto a tandem 5 mL HiTrap Q HP Sepharose Fast Flow column/5 mL HiTrap Heparin HP Sepharose column equilibrated with buffer C (50 mM Tris-Cl, pH 7.5, 5% glycerol, 100mM NaCl, 0.1% BME). After loading the protein sample the HiTrap Q HP Sepharose Fast Flow was removed and the bound protein was washed and eluted with buffer D (50 mM Tris, pH 7.5, 5% glycerol, 1 M NaCl, 0.1% BME) over 10 CVs. The eluted fractions containing pol β were checked on SDS-PAGE for purity and then combined according to purity. The pooled samples were concentrated to 1 mL and loaded on to a Superdex 200 size exclusion column equilibrated with buffer E (50 mM Tris, pH 7.5, 5% glycerol, 100mM NaCl,

0.1% BME) The eluted fractions containing pol  $\beta$  were checked on SDS-PAGE for purity and protein-containing fractions were pooled and dialyzed overnight into dialysis buffer (50 mM Tris-Cl, pH 8.4 at 37 °C, 100 mM NaCl, 0.1 mM EDTA, 5 mM DTT, 10% glycerol). Dialysis buffer containing 70% glycerol was added to the 10% glycerol the protein was currently in so the final volume had 50% glycerol. The protein was then concentrated to 40 mg/mL using a Vivaspin centrifugal filter, cutoff 10KDa, at 3000x g. Concentrated protein was aliquoted in 30  $\mu$ l volume and flash-frozen in liquid nitrogen and stored in -80 °C.

**Pre-Steady State Kinetics:** All fast reactions were performed by using a rapid chemical quench-flow apparatus (KinTek). A pre-incubated solution of full-length wild-type human pol  $\beta$  or its R283A mutant (300 nM) and 30 nM [ $^{32}$ P]-labeled-21-19A-41GT-mer (Table 8) was mixed with varying concentrations of a nucleotide in buffer L (50 mM Tris-HCl, pH 8.4, 5 mM MgCl<sub>2</sub>, 100 mM NaCl, 0.1 mM EDTA, 5 mM dithiothreitol (DTT), 10% glycerol and 0.1 mg/ml bovine serum albumin) at 37°C. After various times, the reaction was stopped with 0.37 M EDTA and analyzed by sequencing gel electrophoresis. Reaction products were visualized using Typhoon TRIO (GE) and quantified using ImageQuant (GE). Each time course of product formation was fit to a single-exponential equation

$$[\text{product}] = E_o \bullet [1 - \exp(- k_{obs}t)]$$

using Kaleidagraph (Synergy Software) to yield a reaction amplitude ( $E_o$ ) 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

$$k_{obs} = k_{pol}[dNTP]/([dNTP] + K_d)$$

to yield an apparent equilibrium dissociation constant of enzyme for the nucleotide ( $K_d$ ) and a maximum nucleotide incorporation rate constant ( $k_{pol}$ ).

### 6.1.3 Results and Discussion

Human pol  $\beta$ , like other DNA pols, has strong *D*-stereoselectivity compared to viral reverse transcriptases. Structural determination of pol  $\beta$  in complex with gapped DNA substrate and an incoming *L*-dCTP, combined with pre-steady state kinetic assays of wild-type pol  $\beta$  and R283A pol  $\beta$  mutant with *L*-analogs of cytosine, were performed to determine the mechanism of *D*-stereoselectivity of pol  $\beta$  and compare the mechanism to pol  $\lambda$ , another X-family pol that it shares high sequence homology with. The R283A mutant was used because it is analogous to R517 in pol  $\lambda$ , which is essential for discrimination against *L*-nucleotides.<sup>125</sup>

The results indicated that the mechanism of *D*-stereoselectivity for wild-type human pol  $\beta$  is unique even compared to pol  $\lambda$ . Pol  $\lambda$  has two ways to incorporate *L*-nucleotides. One way is to bind the *L*-nucleotide while the triphosphate is in the incorrect conformation for catalysis. The R517 sidechain will then form a hydrogen bond with the nucleotide and helps it transition into a catalytically competent conformation. The second is to directly bind the correct conformation of the *L*-nucleotide, though this path is much less efficient, which can be seen in the incorporation kinetics of the R517A pol  $\lambda$  mutant.<sup>125,126</sup> Pol  $\beta$ , however, only binds *L*-nucleotides in the correct conformation, and the R283 sidechain analogous to R517 of pol  $\lambda$  is located too far from the incoming base to form any hydrogen bonds. This means the path pol  $\lambda$  prefers to take when incorporating *L*-nucleotides is not the same that pol  $\beta$  takes.

The kinetics of nucleotide incorporation for pol  $\beta$  also point to a unique incorporation mechanism of *L*-nucleotides. Each *L*-nucleotide used in the kinetic assays had a slower  $k_{pol}$ , a higher  $K_d$ , and a 2- to 5-fold lower efficiency ( $k_{pol}/K_d$ ) than the incorporation of the *D*-dCTP did. However, the R283A mutant displayed relaxed *D*-stereoselectivity. The  $k_{pol}$  of *D*-dCTP dropped by 13-fold compared to wild-type, but the  $k_{pol}$  for the *L*-nucleotides dropped by 4-fold or less. Additionally, the binding affinity for *D*-dCTP and all *L*-nucleotides decreased. Due mainly to the drop in  $K_d$ , the efficiency of incorporation for both *D*- and *L*-nucleotides decreased. Structurally, this can be explained by the loss of anchoring that R283 provides to the thumb subdomain in the closed conformation. Abolishing R517 activity for pol  $\lambda$  only resulted in a decrease in incorporation efficiency for *L*-nucleotides, indicating that R283 in pol  $\beta$  does not play the same role in *D*-stereoselectivity as R517 in pol  $\lambda$ .<sup>126</sup>

Table 8: DNA Substrate for Pre-Steady State Kinetic Assay with Polymerase  $\beta$

<p>3' -GCGTCGGCAGGTTGGTTGAGT<b>GT</b>CAGCTAGGTTACGGCAGG-5'</p> <p>5' -CGCAGCCGTCCAACCAACTCA*<b>A</b>GTCGATCCAATGCCGTCC-3'</p>
<p>The * indicates the gap in the substrate</p>

Table 9: Pre-Steady State Kinetics of Nucleotide Incorporation by Polymerase  $\beta$

Nucleotide	$k_p$ ( $s^{-1}$ )	$K_d$ ( $\mu M$ )	$k_p/K_d$ ( $\mu M^{-1}s^{-1}$ )	$D$ -Stereoselectivity <sup>†</sup>	$R_{D\text{-stereoselectivity}}$ <sup>§</sup>
Catalyzed by wild-type hPol $\beta$					
$D$ -dCTP*	$5.02 \pm 0.07$	$0.71 \pm 0.04$	7.1		
$L$ -dCTP	$0.00059 \pm 0.00002$	$22 \pm 2$	$2.7 \times 10^{-5}$	$2.6 \times 10^5$	
(-) $\beta$ TC-TP*	$0.0039 \pm 0.0001$	$0.18 \pm 0.02$	$2.2 \times 10^{-2}$	323	
(-) $\beta$ TC-TP*	$0.027 \pm 0.001$	$11 \pm 2$	$2.5 \times 10^{-3}$	$2.9 \times 10^3$	
Catalyzed by the R283A mutant of hPol $\beta$					
$D$ -dCTP	$0.39 \pm 0.02$	$41 \pm 6$	$9.5 \times 10^{-3}$		
$L$ -dCTP	$0.00036 \pm 0.00006$	$1110 \pm 370$	$3.2 \times 10^{-7}$	$3.0 \times 10^4$	8.7
(-) $\beta$ TC-TP	$0.0068 \pm 0.0003$	$17 \pm 2$	$4.0 \times 10^{-4}$	24	13
(-) $\beta$ TC-TP	$0.0063 \pm 0.0002$	$24 \pm 3$	$2.6 \times 10^{-4}$	37	78
The * denotes the results I obtained from the experiments I performed.					

## 6.2 Zika Virus NS5 Polymerase

### 6.2.1 My Contribution

Buffer optimization experiments for biochemical characterization of Zika Virus NS5 (ZIKV NS5) have been carried out. Interestingly, some of the results seem to be contradictory. Some characterization experiments were done with no buffer optimization reported and no reasoning for why buffer conditions were chosen.<sup>127–130</sup> Therefore, we

took it upon ourselves to optimize buffer conditions before carrying out kinetic experiments with the ZIKV NS5 of the MR766 strain. The purified protein was graciously provided to us by Jikui Song from the Department of Biochemistry at University of California, Riverside.

My goal in the very beginning stages of the ZIKV NS5 project was to determine the optimum buffer conditions for ribonucleotide incorporation into the primer strand of a self-annealing primer/template substrate by the RNA pol by varying the metal ion, buffer, detergent, and salt used in the reaction conditions. I performed multiple-incorporation assays to test the buffer and detergent components and single-nucleotide incorporation assays to test the metal ion and salt conditions. I also performed an EMSA to estimate the binding affinity of ZIKV NS5 for the non-native RNA substrate used in the incorporation assays. As another part of this project, I write a general review for my PI about ZIKV, the ZIKV NS5, and any kinetics done with other *Flavivirus* RNA pols.

### 6.2.2 Introduction and Background

Zika virus (ZIKV) belongs to the *Flavivirus* genus of the *Flaviviridae* family of viruses.<sup>131–134</sup> This genus includes other pathogenic human viruses such as West Nile (WNV), Dengue (DENV), Japanese encephalitis (JEV), yellow fever (YFV), tick-borne encephalitis (TBEV). Millions of people each year suffer from *Flaviviridae* infection. ZIKV was first isolated in 1947 in Uganda, Africa from a Rhesus monkey and the genome sequenced in 2006. From Africa, it has spread to Asia and the Americas.<sup>133–135</sup>

The main interest in ZIKV comes from the correlation between recent outbreaks of the virus in the Americas (Mexico, Brazil, Haiti, etc.) with heightened rates of

microcephaly and Guillain-Barré Syndrome (GBS).<sup>131,133-138</sup> Because of the association between the recent ZIKV outbreaks and a sharp rise in the occurrence of these neurological disorders, and because no vaccine or treatment for ZIKV yet exists, the World Health Organization declared ZIKV to be a global public health emergency in February 2016.<sup>133,139</sup> ZIKV, like many Flaviviruses, is transmitted to humans by infected mosquitos.<sup>135,140</sup> However, the recent strains cropping up in the Americas have evidenced that ZIKV can be sexually transmitted and that it can cross the placental barrier to infect developing humans, increasing rates of ZIKV infection.<sup>140,141</sup>

ZIKV, like all Flaviviruses, has its genome encoded by single-stranded positive-sense strand RNA that serves as the initial template for minus-strand polymerization.<sup>127,128,135,141</sup> After the minus strand has been synthesized, it serves as a template for multiple rounds of positive-strand synthesis. Ten times more positive-strand RNA will be produced than minus-strand RNA in the replication process of the viral genome. The positive strand RNA encodes for a polypeptide that is co- and post-translationally cleaved into a total of ten proteins: three structural (C, prM, and E) and seven nonstructural (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5).<sup>127,129,135,141,142</sup>

The NS5 protein has two domains: a methyltransferase (MT) domain located on the N-terminal that is responsible for capping the 5'-end of the positive-strand RNA to aid in translation of the polypeptide, and an RNA-dependent RNA polymerase (RdRp) domain on the C-terminus that is responsible for ~~further~~ viral genome replication.<sup>129,130,141,142</sup> Both domains are essential for viral replication. The MT domain catalyzes both the methylation of the N-7 of the 5' guanosine triphosphate and the 2'-O-

hydroxyl of the following nucleotide.<sup>141,143,144</sup> In humans, there are two separate MTs, each of which specifically catalyzes one of the end-capping reactions. Since a single MT is used by Flaviviruses to perform both reactions, inhibition on the MT domain halts viral RNA replication in cell, making this domain a promising drug target. Since the RdRp domain catalyzes *de novo* RNA synthesis and elongation, it is a classic drug target for inhibition of viral replication. The requirement of both domains of NS5 for replication of the viral genome and the spread of the disease in the host makes this enzyme a valuable drug target for treatment of the ZIKV.

The structure of the full-length NS5 protein of the strain MR766, the original strain isolated in Uganda, Africa, was solved to 3.0 Å resolution and 3.3 Å resolution.<sup>129,130</sup> A low-resolution solution SAXS structure of full-length NS5 and from the French Polynesian strain was also solved recently.<sup>145</sup> All structures showed a typical “right-handed” polymerase domain, consisting of the finger, palm, and thumb subdomains. One zinc ion in the fingers subdomain and one near the base of the thumb subdomain were bound.<sup>129,130,146</sup> The SAXS structure implies that in solution, NS5 may exist as a homodimer with the majority of the contacts existing along the faces of the MT domains. The SAXS data also suggests that large conformational changes can occur in the orientation of the MT domain to the RdRp domain.<sup>145</sup> This flexibility is needed since both RNA capping and synthesis are performed by the same protein; the RNA must be handed off from the RdRp domain to the MT domain for capping of the 5' end following *de novo* RNA synthesis.

The different strains of ZIKV have only slight variations in amino acid sequence, but that is somehow sufficient in determining whether a strain is associated with epidemic symptoms or not. The Brazilian strain has 35 amino acids substitutions in the NS5 enzyme compared to the original Ugandan strain. When these residues are mapped onto the structure of Ugandan ZIKV NS5, they all appear on the surface of the enzyme. These substitutions, therefore, should not directly affect polymerization activity.<sup>130</sup> They may, however, affect viral infection efficiency via protein-protein interactions in the cell.

### *6.2.3 Materials and Methods*

**Materials:** Chemicals and reactants were purchased from the following companies: [ $\gamma$ -<sup>32</sup>P]ATP from PerkinElmer Life Sciences (Boston, MA), Optikinase from USB Corp., RNA substrates from Trilink BioTechnologies (San Diego, CA), and rNTPs from Bioline (Taunton, MA).

**Incorporation Assays for Buffer Optimization:** A pre-incubated mixture of ZIKV NS5 (2  $\mu$ M) and R-4 RNA substrate (10 nM) in various buffers was incubated on ice, at room temperature, and at 37°C for 5 min each and mixed with rCTP (100  $\mu$ M) before quenching at various time points with 0.37 EDTA solution. For the buffer optimization assay, buffers contained 50 mM Tris-Cl or HEPES, pH 7.5 at 37°C, 10% glycerol, 0.05% Triton-X 100, 5 mM DTT, and 5 mM MnCl<sub>2</sub>. For triton optimization, the same buffer conditions were used but without Triton-X 100. For salt concentration optimization assays, buffers contained 50 mM Tris-HCl, pH 7.5 at 37°C, 10% glycerol, 0.05% Triton-X 100, 5 mM MgCl<sub>2</sub>, 5 mM DTT, and 10 mM NaCl. Metal ion optimization assay buffers consisted of 50 mM Tris-Cl, pH 7.5 at 37°C, 10% glycerol, 5

mM DTT, and 1-20 mM MgCl<sub>2</sub> or 0.2-5 mM MnCl<sub>2</sub>. All incorporation products were visualized using 1% denaturing agarose gel and Typhoon TRIO (GE) and quantified using ImageQuant (GE).

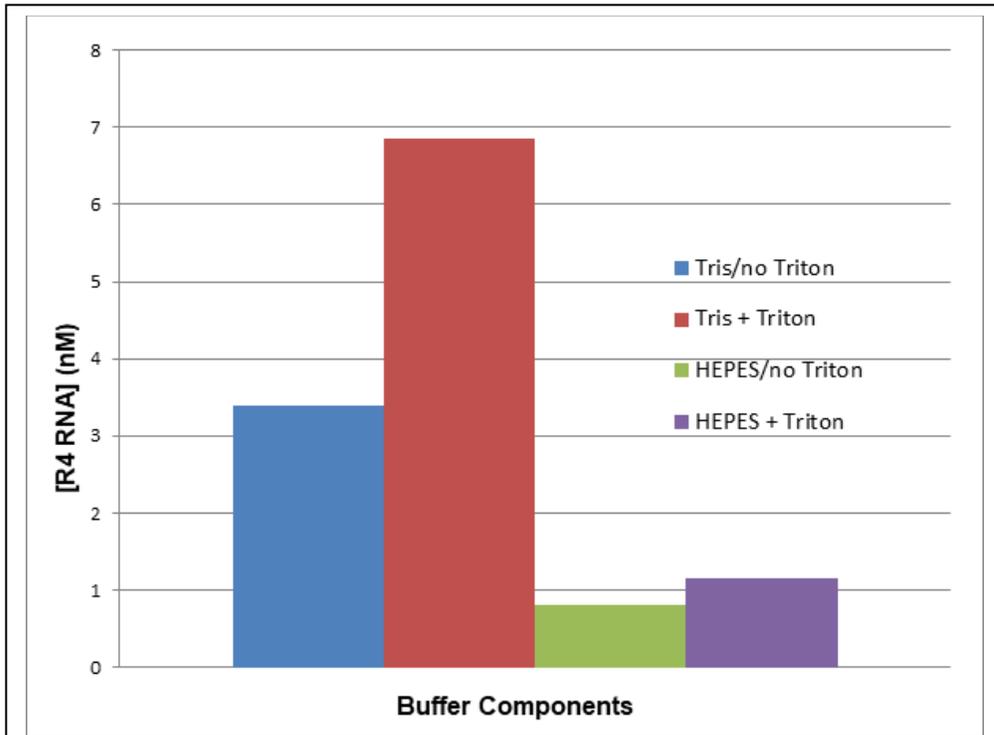
**RNA Binding Assay:** Binding buffer containing 50 mM Tris-Cl, pH 7.5 at 37°C, 10% glycerol, 0.05% Triton-X 100, 10 mM NaCl, 5 mM DTT, and 1 nM R-4 RNA was mixed with increasing concentrations of ZIKV NS5 from 10 nM to 5 μM ZIKV NS5. After a 30 min incubation at 37 °C, samples were run on a native PAGE, visualized using a Typhoon TRIO (GE), and quantified using ImageQuant (GE). Results were graphed using Kaleidagraph (Synergy Software) and fit to the quadratic binding equation to estimate the  $K_d$ .

#### *6.2.4 Results and Discussion*

The presence of Triton-X 100 at 0.05% is beneficial to the polymerase activity. Additionally, using Tris-Cl as the buffering agent appears to be better for polymerase activity than HEPES (Figure 10). Our experiments with increasing the metal ion concentration have not shown product saturation (Figure 11). We have also found that ZIKV NS5 is very sensitive to monovalent salt ions. A low amount of salt (10 mM) added to the reaction assay greatly increases product formation. Adding as little as 5 mM or as much as 100 mM NaCl increased the activity of the polymerase than adding no salt at all (Figure 12). The EMSA revealed that the binding affinity of ZIKV NS5 to the RNA substrate is higher than estimated (21 nM) and that such large concentrations of protein and RNA may not be needed in later kinetic assays.

When considering replicative polymerases, an argument can be made that  $Mg^{2+}$  is the physiological metal catalyst used compared to  $Mn^{2+}$  because  $Mn^{2+}$  decreases polymerase fidelity and is present in much lower concentrations in the cell compared to  $Mg^{2+}$ —and replicative polymerases should be both processive and have high fidelity. However, since NS5 is a viral polymerase, and viruses depend upon mutations to continue evading the host immune system, the fidelity part of the metal catalyst argument may not apply. It remains to be determined whether ZIKV NS5 makes use of predominantly  $Mn^{2+}$  or  $Mg^{2+}$  in the cell, or whether one metal may be used for forming the elongation complex while the other is used during the elongation process. In our assays testing metal ion concentrations, no salt or detergent was added, which we now know would increase the polymerase activity. Repeating that set of optimization experiments using salt and detergent could result in product saturation and therefore reveal an optimal metal ion concentration to use for either  $Mg^{2+}$  or  $Mn^{2+}$  going forward.

Figure 10: Buffer and Detergent Optimization Assays for ZIKV NS5



Bar graph showing the final R-4 RNA concentration at 30 min after mixing for each of the 4 conditions. The final buffer conditions for each were as follows:

Tris/no Triton- 50 mM Tris-Cl, pH 7.5 at 37°C, 10% glycerol, 5 mM DTT, 5 mM Mn<sup>2+</sup>, 100 uM CTP, 2 uM ZIKV NS5, 10 nM R-4 RNA.

Tris + Triton: added 0.05% Triton-X 100

HEPES/no Triton: 50 mM HEPES, pH 7.5 at 37°C, 10% glycerol, 5 mM DTT, 5 mM Mn<sup>2+</sup>, 100 uM CTP, 2 uM ZIKV NS5, 10 nM R-4 RNA.

HEPES + Triton: added 0.05% Triton X-100

Figure 11: Metal Ion Optimization Assay for ZIKV NS5

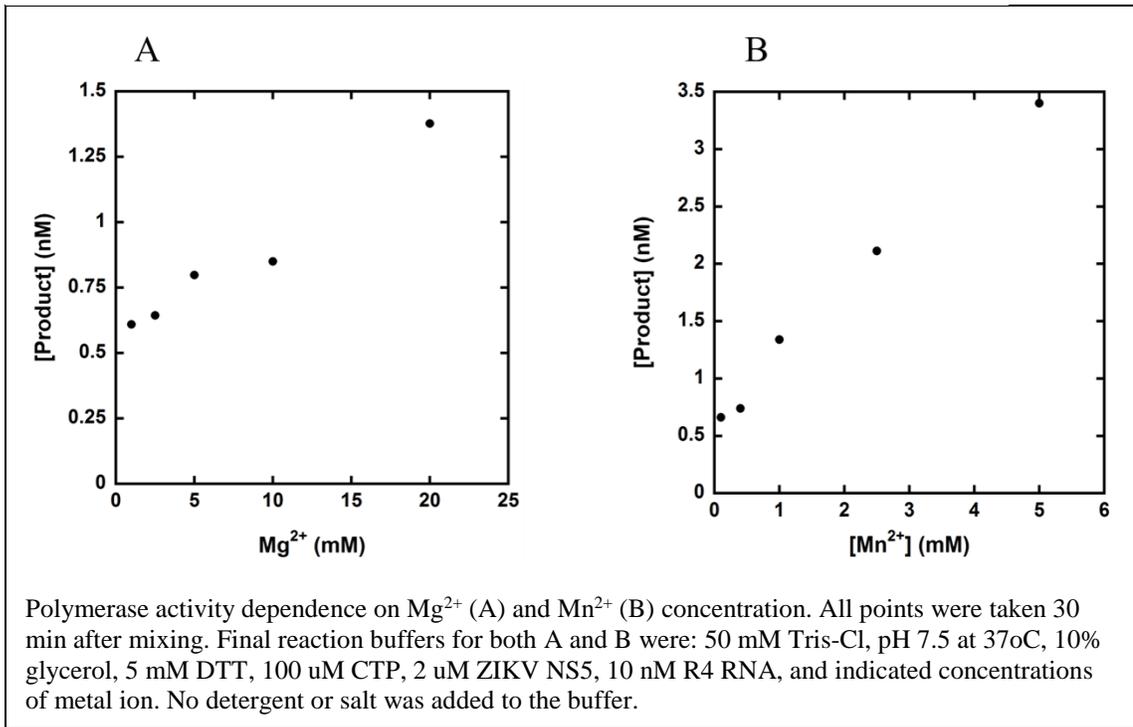


Figure 12: Salt Optimization Assay for ZIKV NS5

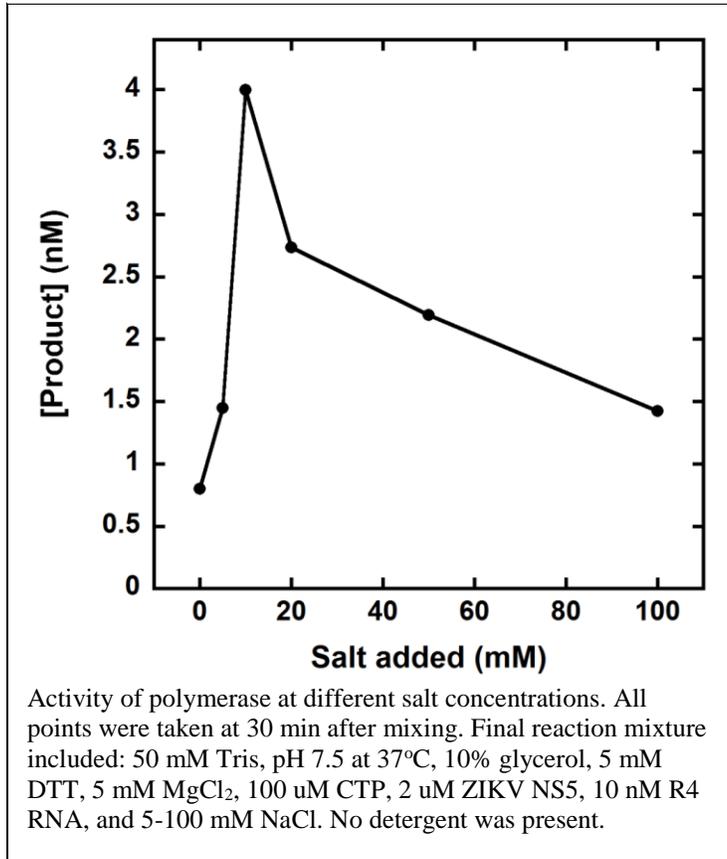


Figure 13: EMSA for ZIKV NS5 with RNA Substrate

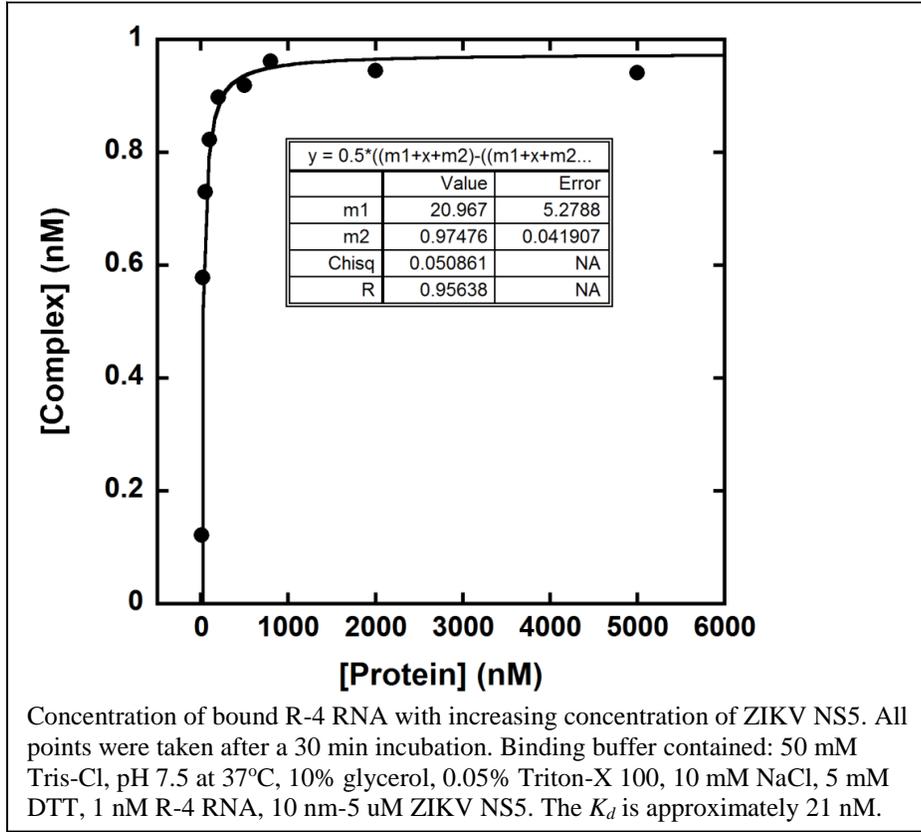


Table 10: RNA Substrate Used in Buffer Optimization Assays and EMSA with ZIKV NS5

3' -CAACCGGUUGGCUA-5'
5' -AUCGGUUGGCCAAC-3'
Self-annealing R-4 RNA substrate.

## References

1. Johansson E, Dixon N. Replicative DNA polymerases. *Cold Spring Harb Perspect Biol.* 2013;5(6). doi:10.1101/cshperspect.a012799.
2. Parsons JL, Nicolay NH, Sharma RA. Biological and Therapeutic Relevance of Nonreplicative DNA Polymerases to Cancer. *Antioxid Redox Signal.* 2013;18(8):851-873. doi:10.1089/ars.2011.4203.
3. Waters LS, Minesinger BK, Wiltrout ME, D'Souza S, Woodruff R V, Walker GC. Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. *Microbiol Mol Biol Rev.* 2009;73(1):134-154. doi:10.1128/MMBR.00034-08.
4. He X-J, Chen T, Zhu J-K. Regulation and function of DNA methylation in plants and animals. *Cell Res.* 2011;21(3):442-465. doi:10.1038/cr.2011.23.
5. Yu M, Hon GC, Szulwach KE, et al. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell.* 2012;149(6):1368-1380. doi:10.1016/j.cell.2012.04.027.
6. Moore LD, Le T, Fan G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology.* 2013;38(1):23-38. doi:10.1038/npp.2012.112.

7. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat Rev Genet.* 2013;14(3):204-220. doi:10.1038/nrg3354.
8. Ito S, Kuraoka I. Epigenetic modifications in DNA could mimic oxidative DNA damage: A double-edged sword. *DNA Repair (Amst).* 2015;32:52-57. doi:10.1016/j.dnarep.2015.04.013.
9. Hakem R. DNA-damage repair; the good, the bad, and the ugly. *EMBO J.* 2008;27(4):589-605. doi:10.1038/emboj.2008.15.
10. Behrens A, van Deursen JM, Rudolph KL, Schumacher B. Impact of genomic damage and ageing on stem cell function. *Nat Cell Biol.* 2014;16(3):201-207. doi:10.1038/ncb2928.
11. Vermeij WP, Hoeijmakers JH, Pothof J. Aging: not all DNA damage is equal. *Curr Opin Genet Dev.* 2014;26:124-130. doi:10.1016/j.gde.2014.06.006.
12. Hanawalt PC. Historical perspective on the DNA damage response. *DNA Repair (Amst).* 2015;36:2-7. doi:10.1016/j.dnarep.2015.10.001.
13. Wilson DM, Bohr VA. The mechanics of base excision repair, and its relationship to aging and disease. *DNA Repair (Amst).* 2007;6(4):544-559. doi:10.1016/j.dnarep.2006.10.017.
14. Davis AJ, Chen DJ. DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res.* 2013;2(3):130-143. doi:10.3978/j.issn.2218-676X.2013.04.02.
15. Li X, Heyer W-D. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res.* 2008;18(1):99-113. doi:10.1038/cr.2008.1.

16. Truong LN, Li Y, Shi LZ, et al. Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proc Natl Acad Sci U S A*. 2013;110(19):7720-7725. doi:10.1073/pnas.1213431110.
17. McVey M, Lee SE. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet*. 2008;24(11):529-538. doi:10.1016/j.tig.2008.08.007.
18. Vyas R, Reed AJ, Tokarsky EJ, Suo Z. Viewing Human DNA Polymerase  $\beta$  Faithfully and Unfaithfully Bypass an Oxidative Lesion by Time-Dependent Crystallography. *J Am Chem Soc*. 2015;137(15):5225-5230. doi:10.1021/jacs.5b02109.
19. Wu S, Beard WA, Pedersen LG, Wilson SH. Structural comparison of DNA polymerase architecture suggests a nucleotide gateway to the polymerase active site. *Chem Rev*. 2014;114(5):2759-2774. doi:10.1021/cr3005179.
20. Yamtich J, Sweasy JB. DNA polymerase family X: function, structure, and cellular roles. *Biochim Biophys Acta*. 2010;1804(5):1136-1150. doi:10.1016/j.bbapap.2009.07.008.
21. Rothwell PJ, Waksman G. Structure and mechanism of DNA polymerases. *Adv Prot Chem*. 2005;71:401-440. doi:10.1016/S0065-3233(04)71011-6.
22. Maxwell BA, Suo Z. Recent insight into the kinetic mechanisms and conformational dynamics of Y-Family DNA polymerases. *Biochemistry*. 2014;53(17):2804-2814. doi:10.1021/bi5000405.

23. Brown JA, Suo Z. Elucidating the Kinetic Mechanism of DNA Polymerization Catalyzed by *Sulfolobus solfataricus* P2 DNA Polymerase B1. *Biochemistry*. 2009;48:7502-7511. doi:10.1021/bi9005336.
24. Zahurancik WJ, Klein SJ, Suo Z. Kinetic mechanism of DNA polymerization catalyzed by human DNA polymerase  $\epsilon$ . *Biochemistry*. 2013;52(40):7041-7049. doi:10.1021/bi400803v.
25. Reed AJ, Suo Z, Vyas R, Raper AT. Structural Insights into the Post-Chemistry Steps of Nucleotide Incorporation Catalyzed by a DNA Polymerase. *J Am Chem Soc*. 2016;139(1):465-471. doi:10.1021/jacs.6b11258.
26. Yeeles JTP, Janska A, Early A, Diffley JFX. How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication. *Mol Cell*. 2017;65(1):105-116. doi:10.1016/j.molcel.2016.11.017.
27. Schauer GD, O'Donnell ME. Quality control mechanisms exclude incorrect polymerases from the eukaryotic replication fork. *Proc Natl Acad Sci*. 2017;201619748. doi:10.1073/pnas.1619748114.
28. Georgescu RE, Schauer GD, Yao NY, et al. Reconstitution of a eukaryotic replisome reveals suppression mechanisms that define leading/lagging strand operation. *Elife*. 2015;2015(4):1-20. doi:10.7554/eLife.04988.
29. Hübscher U, Maga G, Spadari S. Eukaryotic DNA Polymerases. *Annu Rev Biochem*. 2002;71(1):133-163. doi:10.1146/annurev.biochem.71.090501.150041.
30. Nick McElhinny SAN, Gordenin DA, Garrett-Engle CM, Burgers PMJ, Kunkel TA. Division of labor at the eukaryotic replication fork. *Mol Cell*. 2008;30(2):137-

144. doi:10.1016/j.molcel.2008.02.022.
31. Shinbrot E, Henninger EE, Weinhold N, et al. Exonuclease mutations in DNA polymerase epsilon reveal replication strand specific mutation patterns and human origins of replication. *Genome Res.* 2014;24(11):1740-1750. doi:10.1101/gr.174789.114.
32. Miyabe I, Kunkel TA, Carr AM. The major roles of DNA polymerases epsilon and delta at the eukaryotic replication fork are evolutionarily conserved. *PLoS Genet.* 2011;7(12):e1002407. doi:10.1371/journal.pgen.1002407.
33. Baranovskiy AG, Babayeva ND, Liston VG, et al. X-ray structure of the complex of regulatory subunits of human DNA polymerase delta. *Cell Cycle.* 2008;7(19):3026-3036. doi:10.4161/cc.7.19.6720.
34. Zhou Y, Meng X, Zhang S, Lee EYC, Lee MYWT. Characterization of human DNA polymerase delta and its subassemblies reconstituted by expression in the multibac system. *PLoS One.* 2012;7(6). doi:10.1371/journal.pone.0039156.
35. Lee MYWT, Zhang S, Hua S, et al. The tail that wags the dog. *Cell Cycle.* 2014;13(1):23-31. doi:http://dx.doi.org/10.4161/cc.27407.
36. Pursell ZF, Isoz I, Lundström E-B, Johansson E, Kunkel TA. Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science.* 2007;317(5834):127-130. doi:10.1126/science.1144067.
37. Yeeles JTP, Deegan TD, Janska A, Early A, Diffley JFX. Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature.* 2015;519(7544):431-435. doi:10.1038/nature14285.

38. Henninger EE, Pursell ZF. DNA polymerase  $\epsilon$  and its roles in genome stability. *IUBMB Life*. 2014;66(5):339-351. doi:10.1002/iub.1276.
39. Zhou JC, Janska A, Goswami P, et al. CMG–Pol epsilon dynamics suggests a mechanism for the establishment of leading-strand synthesis in the eukaryotic replisome. *Proc Natl Acad Sci*. 2017;201700530. doi:10.1073/pnas.1700530114.
40. Kane DP, Shcherbakova P V. A common cancer-associated DNA polymerase ?? mutation causes an exceptionally strong mutator phenotype, indicating fidelity defects distinct from loss of proofreading. *Cancer Res*. 2014;74(7):1895-1901. doi:10.1158/0008-5472.CAN-13-2892.
41. Friedberg EC. Suffering in silence: the tolerance of DNA damage. *Nat Rev Mol Cell Biol*. 2005;6(12):943-953. doi:10.1038/nrm1781.
42. Broughton BC, Cordonnier A, Kleijer WJ, et al. Molecular analysis of mutations in DNA polymerase eta in xeroderma pigmentosum-variant patients. *Proc Natl Acad Sci U S A*. 2002;99(2):815-820. doi:10.1073/pnas.022473899.
43. Lange SS, Takata K, Wood RD. DNA polymerases and cancer. *Nat Rev Cancer*. 2011;11(2):96-110. doi:10.1038/nrc2998.
44. Allinson SL, Dianova II, Dianov GL. DNA polymerase  $\beta$  is the major dRP lyase involved in repair of oxidative base lesions in DNA by mammalian cell extracts. *EMBO J*. 2001;20(23):6919-6926. doi:10.1093/emboj/20.23.6919.
45. Braithwaite EK, Kedar PS, Stumpo DJ, et al. DNA polymerases beta and lambda mediate overlapping and independent roles in base excision repair in mouse embryonic fibroblasts. Santos J, ed. *PLoS One*. 2010;5(8):e12229.

- doi:10.1371/journal.pone.0012229.
46. Belousova EA, Lavrik OI. DNA polymerases  $\beta$  and  $\lambda$  and their roles in cell. *DNA Repair (Amst)*. 2015;29:112-126. doi:10.1016/j.dnarep.2015.02.001.
  47. Pryor JM, Waters CA, Aza A, et al. Essential role for polymerase specialization in cellular nonhomologous end joining. *Proc Natl Acad Sci U S A*. 2015;112(33):E4537-45. doi:10.1073/pnas.1505805112.
  48. Mueller GA, Moon AF, Derose EF, et al. A comparison of BRCT domains involved in nonhomologous end-joining: introducing the solution structure of the BRCT domain of polymerase lambda. *DNA Repair (Amst)*. 2008;7(8):1340-1351. doi:10.1016/j.dnarep.2008.04.018.
  49. Nick McElhinny SA, Havener JM, Garcia-Diaz M, et al. A Gradient of Template Dependence Defines Distinct Biological Roles for Family X Polymerases in Nonhomologous End Joining. *Mol Cell*. 2005;19(3):357-366. doi:10.1016/j.molcel.2005.06.012.
  50. Mahajan KN, Mcelhinny SAN, Mitchell BS, Ramsden DA. Association of DNA Polymerase mu (pol mu) with Ku and Ligase IV: Role for pol mu in End-Joining Double-Strand Break Repair. *Mol Cell Biol*. 2002;22(14):5194-5202. doi:10.1128/MCB.22.14.5194-5202.2002.
  51. Yang W. An overview of Y-Family DNA polymerases and a case study of human DNA polymerase  $\eta$ . *Biochemistry*. 2014;53(17):2793-2803. doi:10.1021/bi500019s.
  52. Brabec V, Malina J, Margiotta N, Natile G, Kasparkova J. Thermodynamic and

- mechanistic insights into translesion DNA synthesis catalyzed by Y-family DNA polymerase across a bulky double-base lesion of an antitumor platinum drug. *Chemistry*. 2012;18(48):15439-15448. doi:10.1002/chem.201202117.
53. Petta TB, Nakajima S, Zlatanou A, et al. Human DNA polymerase iota protects cells against oxidative stress. *EMBO J*. 2008;27(21):2883-2895. doi:10.1038/emboj.2008.210.
54. Faili A, Aoufouchi S, Flatter E, Gu?ranger Q, Reynaud C-A, Weill J-C. Induction of somatic hypermutation in immunoglobulin genes is dependent on DNA polymerase iota. *Nature*. 2002;419(6910):944-947. doi:10.1038/nature01117.
55. Stancel JNK, McDaniel LD, Velasco S, Richardson J, Guo C, Friedberg EC. Polk mutant mice have a spontaneous mutator phenotype. *DNA Repair (Amst)*. 2009;8(12):1355-1362. doi:10.1016/j.dnarep.2009.09.003.
56. Washington MT, Johnson RE, Prakash L, Prakash S. Human DINB1-encoded DNA polymerase kappa is a promiscuous extender of mispaired primer termini. *Proc Natl Acad Sci U S A*. 2002;99(4):1910-1914. doi:10.1073/pnas.032594399.
57. Ohashi E, Bebenek K, Matsuda T, et al. Fidelity and processivity of DNA synthesis by DNA polymerase kappa, the product of the human DINB1 gene. *J Biol Chem*. 2000;275(50):39678-39684. doi:10.1074/jbc.M005309200.
58. Seki M, Masutani C, Yang LW, et al. High-efficiency bypass of DNA damage by human DNA polymerase Q. *EMBO J*. 2004;23(22):4484-4494. doi:10.1038/sj.emboj.7600424.
59. Higgins GS, Prevo R, Lee Y-F, et al. A small interfering RNA screen of genes

- involved in DNA repair identifies tumor-specific radiosensitization by POLQ knockdown. *Cancer Res.* 2010;70(7):2984-2993. doi:10.1158/0008-5472.CAN-09-4040.
60. Yoshimura M, Kohzaki M, Nakamura J, et al. Vertebrate POLQ and POL $\beta$  Cooperate in Base Excision Repair of Oxidative DNA Damage. *Mol Cell.* 2006;24(1):115-125. doi:10.1016/j.molcel.2006.07.032.
61. Prasad R, Longley MJ, Sharief FS, Hou EW, Copeland WC, Wilson SH. Human DNA polymerase theta possesses 5'-dRP lyase activity and functions in single-nucleotide base excision repair in vitro. *Nucleic Acids Res.* 2009;37(6):1868-1877. doi:10.1093/nar/gkp035.
62. Hogg M, Seki M, Wood RD, Doublié S, Wallace SS. Lesion bypass activity of DNA polymerase  $\theta$  (POLQ) is an intrinsic property of the pol domain and depends on unique sequence inserts. *J Mol Biol.* 2011;405(3):642-652. doi:10.1016/j.jmb.2010.10.041.
63. O-Wang J, Kajiwara K, Kawamura K, et al. An essential role for REV3 in mammalian cell survival: absence of REV3 induces p53-independent embryonic death. *Biochem Biophys Res Commun.* 2002;293(3):1132-1137. doi:10.1016/S0006-291X(02)00341-8.
64. Jansen JG, Tsaalbi-Shtylik A, Hendriks G, et al. Mammalian Polymerase  $\zeta$  Is Essential for Post-Replication Repair of UV-Induced DNA Lesions. *DNA Repair.* 2009;8(12):1444-1451. doi:10.1016/j.dnarep.2009.09.006.
65. Sassa A, Beard WA, Shock DD, Wilson SH. Steady-state, pre-steady-state, and

- single-turnover kinetic measurement for DNA glycosylase activity. *J Vis Exp*. 2013;(78):e50695. doi:10.3791/50695.
66. Johnson KA. 1 Transient-State Kinetic Analysis of Enzyme Reaction Pathways. In: *The Enzymes, Vol. XX.* ; 1992:1-61. doi:10.1016/S1874-6047(08)60019-0.
67. Carruthers A. Enzyme Kinetics. *eBook*. 2015:1-81.  
<http://glutxi.umassmed.edu/grad/GradKinetics.pdf>. Accessed June 28, 2017.
68. Handy DE, Castro R, Loscalzo J. Epigenetic modifications: basic mechanisms and role in cardiovascular disease. *Circulation*. 2011;123(19):2145-2156.  
doi:10.1161/CIRCULATIONAHA.110.956839.
69. Meng H, Cao Y, Qin J, et al. DNA methylation, its mediators and genome integrity. *Int J Biol Sci*. 2015;11(5):604-617. doi:10.7150/ijbs.11218.
70. He XJ, Chen T, Zhu JK. Regulation and function of DNA methylation in plants and animals. *Cell Res*. 2011;21(3):442-465. doi:10.1038/cr.2011.23.
71. Grayson DR, Guidotti A. The Dynamics of DNA Methylation in Schizophrenia and Related Psychiatric Disorders. *Neuropsychopharmacology*. 2012;38(1):138-166. doi:10.1038/npp.2012.125.
72. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat Rev Genet*. 2013;14(3):204-220. doi:10.1038/nrg3354.
73. Eleftheriou M, Pascual AJ, Wheldon LM, et al. 5-Carboxylcytosine levels are elevated in human breast cancers and gliomas. *Clin Epigenetics*. 2015;7(1):88.  
doi:10.1186/s13148-015-0117-x.
74. Peng J, Xia B, Yi C. Single-base resolution analysis of DNA epigenome via high-

- throughput sequencing. *Sci China Life Sci.* 2016;59(3):219-226.  
doi:10.1007/s11427-016-5013-x.
75. Shibutani T, Ito S, Toda M, et al. Guanine-5-carboxylcytosine base pairs mimic mismatches during DNA replication. *Sci. Rep.* 2014;4(5220).  
doi:10.1038/srep05220.
76. He Y-F, Li B-Z, Li Z, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science.* 2011;333(6047):1303-1307.  
doi:10.1126/science.1210944.
77. Hashimoto H, Hong S, Bhagwat AS, Zhang X, Cheng X. Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation. *Nucleic Acids Res.* 2012;40(20):10203-10214. doi:10.1093/nar/gks845.
78. Frank EG, Tissier A, McDonald JP, et al. Altered nucleotide misinsertion fidelity associated with poliota-dependent replication at the end of a DNA template. *EMBO J.* 2001;20(11):2914-2922. doi:10.1093/emboj/20.11.2914.
79. Kirouac KN, Ling H. Structural basis of error-prone replication and stalling at a thymine base by human DNA polymerase iota. *EMBO J.* 2009;28(11):1644-1654.  
doi:10.1038/emboj.2009.122.
80. Watson JD, Crick FHC. The Structure of DNA. *Cold Spring Harb Symp Quant Biol.* 1953;18:123-131 doi: 10.1101/SQB.1953.018.01.020
81. Kimsey IJ, Petzold K, Sathyamoorthy B, Stein ZW, Al-Hashimi HM. Visualizing transient Watson–Crick-like mispairs in DNA and RNA duplexes. *Nature.*

- 2015;519(7543):315-320. doi:10.1038/nature14227.
82. Topal MD, Fresco JR. Complementary base pairing and the origin of substitution mutations. *Nature*. 1976;263(5575):285-289. doi:10.1038/263285a0.
  83. Harris VH, Smith CL, Cummins WJ, et al. The effect of tautomeric constant on the specificity of nucleotide incorporation during DNA replication: Support for the rare tautomer hypothesis of substitution mutagenesis. *J Mol Biol*. 2003;326(5):1389-1401. doi:10.1016/S0022-2836(03)00051-2.
  84. Bebenek K, Pedersen LC, Kunkel TA. Replication infidelity via a mismatch with Watson–Crick geometry. *Proc Natl Acad Sci*. 2011;108(5):1862-1867. doi:10.1073/pnas.1012825108.
  85. Warren JJ, Forsberg LJ, Beese LS. The structural basis for the mutagenicity of O(6)-methyl-guanine lesions. *Proc Natl Acad Sci U S A*. 2006;103(52):19701-19706. doi:10.1073/pnas.0609580103.
  86. Sowers LC, Goodman MF, Eritja R, Kaplan B, Fazakerley GV. Ionized and wobble base-pairing for bromouracil-guanine in equilibrium under physiological conditions. *J Mol Biol*. 1989;205(2):437-447. doi:10.1016/0022-2836(89)90353-7.
  87. Washington MT, Johnson RE, Prakash L, Prakash S. Human DNA polymerase iota utilizes different nucleotide incorporation mechanisms dependent upon the template base. *Mol Cell Biol*. 2004;24(2):936-943. doi:10.1128/mcb.24.2.936-943.2004.
  88. Zhang Y, Yuan F, Wu X, Wang Z. Preferential Incorporation of G Opposite Template T by the Low-Fidelity Human DNA Polymerase iota. *Mol Cell Biol*.

- 2000;20(19):7099-7108. doi:10.1128/MCB.20.19.7099-7108.2000.
89. Tissier a, McDonald JP, Frank EG, Woodgate R. poliota, a remarkably error-prone human DNA polymerase. *Genes Dev.* 2000;14(13):1642-1650. doi:10.1101/gad.14.13.1642.
90. Kellinger MW, Song C-X, Chong J, Lu X-Y, He C, Wang D. 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. *Nat Struct Mol Biol.* 2012;19(8):831-833. doi:10.1038/nsmb.2346.
91. Vaisman A, Woodgate R. Unique misinsertion specificity of polt may decrease the mutagenic potential of deaminated cytosines. *EMBO J.* 2001;20(22):6520-6529. doi:10.1093/emboj/20.22.6520.
92. Ramsawhook A, Lewis L, Coyle B, Ruzov A. Medulloblastoma and ependymoma cells display increased levels of 5-carboxylcytosine and elevated TET1 expression. *Clin Epigenetics.* 2017;9:18. doi:10.1186/s13148-016-0306-2.
93. Valentini E, Zampieri M, Malavolta M, et al. Analysis of the machinery and intermediates of the 5hmC-mediated DNA demethylation pathway in aging on samples from the MARK-AGE Study. *Aging (Albany NY).* 2016;8(9):1896-1922. doi:10.18632/aging.101022.
94. Wheldon LM, Abakir A, Ferjentsik Z, et al. Transient Accumulation of 5-Carboxylcytosine Indicates Involvement of Active Demethylation in Lineage Specification of Neural Stem Cells. *Cell Rep.* 2014;7(5):1353-1361. doi:10.1016/j.celrep.2014.05.003.

95. Wang Z, Tang B, He Y, Jin P. DNA methylation dynamics in neurogenesis. *Epigenomics*. 2016;8(3):401-414. doi:10.2217/epi.15.119.
96. Shen L, Song C-X, He C, Zhang Y. Mechanism and function of oxidative reversal of DNA and RNA methylation. *Annu Rev Biochem*. 2014;83:585-614. doi:10.1146/annurev-biochem-060713-035513.
97. Pursell ZF, Kunkel TA. Chapter 4 DNA Polymerase  $\epsilon$ . A Polymerase of Unusual Size (and Complexity). *Prog Nucleic Acid Res Mol Biol*. 2008;82:101-145. doi:10.1016/S0079-6603(08)00004-4.
98. Tsubota T, Tajima R, Ode K, et al. Double-stranded DNA binding, an unusual property of DNA polymerase epsilon, promotes epigenetic silencing in *Saccharomyces cerevisiae*. *J Biol Chem*. 2006;281(43):32898-32908. doi:10.1074/jbc.M606637200.
99. Bermudez VP, Farina A, Raghavan V, Tappin I, Hurwitz J. Studies on human DNA polymerase epsilon and GINS complex and their role in DNA replication. *J Biol Chem*. 2011;286(33):28963-28977. doi:10.1074/jbc.M111.256289.
100. Vaara M, Itkonen H, Hillukkala T, et al. Segregation of replicative DNA polymerases during S phase: DNA polymerase  $\epsilon$ , but not DNA polymerases  $\alpha/\delta$ , are associated with lamins throughout S phase in human cells. *J Biol Chem*. 2012;287(40):33327-33338. doi:10.1074/jbc.M112.357996.
101. Nuutinen T, Tossavainen H, Fredriksson K, et al. The solution structure of the amino-terminal domain of human DNA polymerase epsilon subunit B is homologous to C-domains of AAA+ proteins. *Nucleic Acids Res*.

- 2008;36(15):5102-5110. doi:10.1093/nar/gkn497.
102. Asturias FJ, Cheung IK, Sabouri N, Chilkova O, Wepplo D, Johansson E. Structure of *Saccharomyces cerevisiae* DNA polymerase epsilon by cryo-electron microscopy. *Nat Struct Mol Biol.* 2006;13(1):35-43. doi:10.1038/nsmb1040.
  103. Hogg M, Osterman P, Bylund GO, et al. Structural basis for processive DNA synthesis by yeast DNA polymerase  $\epsilon$ . *Nat Struct Mol Biol.* 2014;21(1):49-55. doi:10.1038/nsmb.2712.
  104. Tokarsky EJ, Wallenmeyer PC, Phi KK, Suo Z. Significant impact of divalent metal ions on the fidelity, sugar selectivity, and drug incorporation efficiency of human PrimPol. *DNA Repair (Amst).* 2017;49. doi:10.1016/j.dnarep.2016.11.003.
  105. García-Gómez S, Reyes A, Martínez-Jiménez MI, et al. PrimPol, an Archaic Primase/Polymerase Operating in Human Cells. *Mol Cell.* 2013;52(4):541-553. doi:10.1016/j.molcel.2013.09.025.
  106. Bianchi J, Rudd SG, Jozwiakowski SK, et al. PrimPol bypasses UV photoproducts during eukaryotic chromosomal DNA replication. *Mol Cell.* 2013;52(4):566-573. doi:10.1016/j.molcel.2013.10.035.
  107. Bailey LJ, Doherty AJ. Mitochondrial DNA replication: a PrimPol perspective. *Biochem Soc Trans.* 2017;45(2):513-529. doi:10.1042/BST20160162.
  108. Guillian TA, Brissett NC, Ehlinger A, et al. Molecular basis for PrimPol recruitment to replication forks by RPA. *Nat Commun.* 2017;8:15222. doi:10.1038/ncomms15222.
  109. Wan L, Lou J, Xia Y, et al. hPrimpol1/CCDC111 is a human DNA primase-

- polymerase required for the maintenance of genome integrity. *EMBO Rep.* 2013;14(12):1104-1112. doi:10.1038/embor.2013.159.
110. Martínez-Jiménez MI, García-Gómez S, Bebenek K, et al. Alternative solutions and new scenarios for translesion DNA synthesis by human PrimPol. *DNA Repair (Amst)*. 2015;29:127-138. doi:10.1016/j.dnarep.2015.02.013.
  111. Keen BA, Jozwiakowski SK, Bailey LJ, Bianchi J, Doherty AJ. Molecular dissection of the domain architecture and catalytic activities of human PrimPol. *Nucleic Acids Res.* 2014;42(9):5830-5845. doi:10.1093/nar/gku214.
  112. Zafar MK, Ketkar A, Lodeiro MF, Cameron CE, Eoff RL. Kinetic analysis of human PrimPol DNA polymerase activity reveals a generally error-prone enzyme capable of accurately bypassing 7,8-dihydro-8-oxo-2'-deoxyguanosine. *Biochemistry.* 2014;53(41):6584-6594. doi:10.1021/bi501024u.
  113. Rechkoblit O, Gupta YK, Malik R, et al. Structure and mechanism of human PrimPol , a DNA polymerase with primase activity. *Sci Adv.* 2016;2(10):e1601317. doi:10.1126/sciadv.1601317.
  114. Guillian TA, Jozwiakowski SK, Ehlinger A, et al. Human PrimPol is a highly error-prone polymerase regulated by single-stranded DNA binding proteins. *Nucleic Acids Res.* 2015;43(2):1056-1068. doi:10.1093/nar/gku1321.
  115. Keen BA, Bailey LJ, Jozwiakowski SK, Doherty AJ. Human PrimPol mutation associated with high myopia has a DNA replication defect. *Nucleic Acids Res.* 2014;42(19):12102-12111. doi:10.1093/nar/gku879.
  116. Kobayashi K, Guillian TA, Tsuda M, et al. Repriming by PrimPol is critical for

- DNA replication restart downstream of lesions and chain-terminating nucleosides. *Cell Cycle*. 2016;15(15):1997-2008. doi:10.1080/15384101.2016.1191711.
117. Schiavone D, Jozwiakowski SK, Romanello M, et al. PrimPol Is Required for Replicative Tolerance of G Quadruplexes in Vertebrate Cells. *Mol Cell*. 2016;61(1):161-169. doi:10.1016/j.molcel.2015.10.038.
118. Mourón S, Rodriguez-Acebes S, Martínez-Jiménez MI, et al. Repriming of DNA synthesis at stalled replication forks by human PrimPol. *Nat Struct Mol Biol*. 2013;20(12):3-8. doi:10.1038/nsmb.2719.
119. Romani AM, Scarpa A. Regulation of cellular magnesium. *Front Biosci*. 2000;5(1):D720-34. doi:10.1016/0003-9861(92)90086-C.
120. Romani AMP. Cellular magnesium homeostasis. *Arch Biochem Biophys*. 2011;512(1):1-23. doi:10.1016/j.abb.2011.05.010.
121. Bowman AB, Aschner M. Considerations on manganese (Mn) treatments for in vitro studies. *Neurotoxicology*. 2014;41:141-142. doi:10.1016/j.neuro.2014.01.010.
122. Kumar KK, Lowe EW, Aboud AA, et al. Cellular manganese content is developmentally regulated in human dopaminergic neurons. *Sci Rep*. 2014;4:6801. doi:10.1038/srep06801.
123. Tuschl K, Mills PB, Clayton PT. Manganese and the brain. *Int Rev Neurobiol*. 2013;110:277-312. doi:10.1016/B978-0-12-410502-7.00013-2.
124. Vyas R, Reed AJ, Raper AT, Zahurancik WJ, Wallenmeyer PC, Suo Z. Structural basis for the D-stereoselectivity of human DNA polymerase  $\beta$ . *Nucleic Acids Res*.

- 2017;45(10). doi:10.1093/nar/gkx252.
125. Vyas R, Zahurancik WJ, 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*. 2014;111(30):E3033-42. doi:10.1073/pnas.1401286111.
  126. Brown JA, Pack LR, Fowler JD, 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*. 2011;55(1):276-283. doi:10.1128/AAC.01229-10.
  127. Lu G, Bluemling GR, Collop P, et al. Analysis of Ribonucleotide 5'-Triphosphate Analogs as Potential Inhibitors of Zika Virus RNA-Dependent RNA Polymerase by Using Nonradioactive Polymerase Assays. *Antimicrob Agents Chemother*. 2017;61(3):e01967-16. doi:10.1128/AAC.01967-16.
  128. Potisophon S, Ferron F, Fattorini V, Selisko B, Canard B. Substrate selectivity of Dengue and Zika virus NS5 polymerase towards 2'-modified nucleotide analogues. *Antiviral Res*. 2017;140:25-36. doi:10.1016/j.antiviral.2016.12.021.
  129. Wang B, Tan X-F, Thurmond S, et al. The structure of Zika virus NS5 reveals a conserved domain conformation. *Nat Commun*. 2017;8:14763. doi:10.1038/ncomms14763.
  130. Zhao B, Yi G, Du F, et al. Structure and function of the Zika virus full-length NS5 protein. 2017;8:14762. doi:10.1038/ncomms14762.
  131. Duffy MR, Chen T-H, Hancock WT. Zika Virus Outbreak on Yap Island, Federated States of Micronesia. *N Engl J Med*. 2009;360(24):2536-2543.

132. Barba-Spaeth G, Dejnirattisai W, Rouvinski A, et al. Structural basis of potent Zika–dengue virus antibody cross-neutralization. *Nature*. 2016;536. doi:10.1038/nature18938.
133. Peterson LR, Jamieson DJ, Powers AM, Honein MA. Zika Virus. *N Engl J Med*. 2016;(374):1552-1563. doi:10.1056/NEJMra1602113.
134. Cao-Lormeau V-M, Blake A, Mons S, et al. Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *Lancet (London, England)*. 2016;387(10027):1531-1539. doi:10.1016/S0140-6736(16)00562-6.
135. Faria NR, Azevedo R do S da S, Kraemer MUG, et al. Zika virus in the Americas: Early epidemiological and genetic findings. *Science*. 2016;352(6283):345-349. doi:10.1126/science.aaf5036.
136. de Siqueira IC, Rodrigues SG, Martins LC, et al. Guillain–Barré Syndrome After Zika Virus Infection in Brazil. *Am J Trop Med Hyg*. 2016;95(5):1157-1160. doi:10.4269/ajtmh.16-0306.
137. Rasmussen SA, Jamieson DJ, Honein MA, Petersen LR. Zika Virus and Birth Defects — Reviewing the Evidence for Causality. *N Engl J Med*. 2016;20:1-7. <http://www.nejm.org/doi/pdf/10.1056/NEJMs1604338>.
138. Mcgrath EL, Rossi SL, Gao J, et al. Differential Responses of Human Fetal Brain Neural Stem Cells to Zika Virus Infection. *Stem Cell Reports*. 2017;8(3):715-727. doi:10.1016/j.stemcr.2017.01.008.
139. Zhang C, Feng T, Cheng J, et al. Structure of the NS5 methyltransferase from Zika

- virus and implications in inhibitor design. *Biochem Biophys Res Commun.* 2016.  
doi:10.1016/j.bbrc.2016.11.098.
140. Li Y, He L, He RL, Yau SS-T. Zika and Flaviviruses Phylogeny Based on the Alignment-Free Natural Vector Method. *DNA and Cell Biology.* 2017;36(2):109-116. doi:10.1089/dna.2016.3532.
141. Hercik K, Brynda J, Nencka R, Boura E. Structural basis of Zika virus methyltransferase inhibition by sinefungin. *Arch Virol.* 2017;(162):2091-2096. doi:10.1007/s00705-017-3345-x.
142. Grant A, Ponia SS, Tripathi S, et al. Zika Virus Targets Human STAT2 to Inhibit Type I Interferon Signaling. *Cell Host Microbe.* 2016;19(6):882-890. doi:10.1016/j.chom.2016.05.009.
143. Zhou H, Wang F, Wang H, et al. The conformational changes of Zika virus methyltransferase upon converting SAM to SAH. *Oncotarget.* 2017;8(9):14830-14834. doi:10.18632/oncotarget.14780.
144. Coutard B, Barral K, Lichière J, et al. Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives. *J Virol.* 2017;91(5):e02202-16. doi:10.1128/JVI.02202-16.
145. Saw WG, Pan A, Subramanian Manimekalai MS, Grüber G. Structural features of Zika virus non-structural proteins 3 and -5 and its individual domains in solution as well as insights into NS3 inhibition. *Antiviral Res.* 2017;141:73-90. doi:10.1016/j.antiviral.2017.02.005.
146. Godoy AS, Lima GMA, Oliveira KIZ, et al. ARTICLE Crystal structure of Zika

virus NS5 RNA-dependent RNA polymerase. *Nat Commun.* 2017;8.

doi:10.1038/ncomms14764.