

Electronic Thesis and Dissertation Repository

12-9-2010 12:00 AM

Structural insights into DNA replication and lesion bypass by Y family DNA polymerases

Kevin N. Kirouac, *The University of Western Ontario*

Supervisor: Dr. Hong Ling, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Biochemistry

© Kevin N. Kirouac 2010

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>

 Part of the [Biochemistry, Biophysics, and Structural Biology Commons](#), and the [Pharmacology, Toxicology and Environmental Health Commons](#)

Recommended Citation

Kirouac, Kevin N., "Structural insights into DNA replication and lesion bypass by Y family DNA polymerases" (2010). *Electronic Thesis and Dissertation Repository*. 59.
<https://ir.lib.uwo.ca/etd/59>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

STRUCTURAL INSIGHTS INTO DNA REPLICATION AND LESION BYPASS BY Y
FAMILY DNA POLYMERASES

(Spine title: DNA replication and lesion bypass by Y family DNA polymerases)

(Thesis format: Integrated Article)

by

Kevin Nicholas Kirouac

Graduate Program in Biochemistry

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Kevin N Kirouac 2010

THE UNIVERSITY OF WESTERN ONTARIO
School of Graduate and Postdoctoral Studies

CERTIFICATE OF EXAMINATION

Supervisor

Examiners

Dr. Hong Ling

Dr. Carole Creuzenet

Supervisory Committee

Dr. Megan Davey

Dr. Eric Ball

Dr. Stan Dunn

Dr. Brian Shilton

Dr. Frank Sicheri

The thesis by

Kevin Nicholas Kirouac

entitled:

**Structural insights into DNA replication and lesion bypass by Y
family DNA polymerases**

is accepted in partial fulfillment of the
requirements for the degree of
Doctor of philosophy

Date

Chair of the Thesis Examination Board

Abstract

Y family DNA polymerases are specialized enzymes for replication through sites of DNA damage in the genome. Although the DNA damage bypass activity of these enzymes is important for genome maintenance and integrity, it is also responsible for DNA mutagenesis due to the error-prone nature of the Y family. Understanding how these enzymes select incoming nucleotides during DNA replication will give insight into their role in cancer formation, aging, and evolution. This work attempts to mechanistically explain, primarily through X-ray crystallography and enzymatic activity assays, how Y family polymerases select incoming nucleotides in various DNA replication contexts.

Initially, I sought to determine how the model Y family polymerase Dpo4 differentiates between ribo and deoxyribonucleotides. Crystal structures were solved of a mutant Dpo4 enzyme (Y12A) deficient in ribonucleotide discrimination, incorporating either deoxy-adenine (dA) or ribo-adenine (rA) nucleotides opposite template thymine DNA. It was revealed that the Dpo4 Y12A mutant allowed rA incorporation by accommodating the 2'-OH group of the ribose sugar. Thus Y family polymerases block ribonucleotide incorporation during DNA replication by clashing with the 2'-OH group of ribose sugars.

Next, I examined how human DNA polymerase iota (polI) prefers to incorporate mismatched nucleotides opposite an undamaged thymine base. Crystal structures of polI in complex with template thymine DNA incorporating either correct adenine (A) or mismatched thymine (T) or guanine (G) revealed the structural basis of error-prone replication. The A nucleotide was destabilized by a narrow polI active site and mis-matched G was preferred by hydrogen bonding with glutamine59 from the finger domain. Domain swapping experiments confirmed the role of the polI finger domain in nucleotide selection opposite T.

I then investigated how polt selects the correct cytosine (C) nucleotide opposite the mutagenic oxidative lesion 8-oxo-guanine. Crystal structures of polt in complex with 8-oxo-guanine DNA incorporating correct C or mis-matched A, T, or G revealed the structural basis of error-free replication. The narrow polt active site destabilizes A and G purine bases while selecting correct C due to the greatest hydrogen bonding potential with the 8-oxo-guanine Hoogsteen edge. We also show how Glu59 from the finger domain is involved in nucleotide selection and bypass activity through site-directed mutagenesis.

Lastly, I examined how polt replicates opposite a bulky lesion produced from environmental pollution: N-[deoxyguanosine-8-yl]-1-amino-pyrene (APG). Crystal structures of polt in complex with APG DNA incorporating C or mis-matched A reveal the structural mechanism of APG replication. The C nucleotide is preferred opposite the lesion due to Watson-Crick base pairing while mis-matched A is incorporated by base stacking above the lesion. I also demonstrate using the model Y family polymerase Dpo4, that the hydrophobic lesion interacts with protein side chains from the little finger domain, which inhibits DNA replication past the lesion site.

Taken together, these results further our understanding of how Y family polymerases select incoming nucleotides and how this selection can result in error-free or error-prone replication depending on the chemical nature of the template base.

Keywords

Y family DNA polymerase, DNA polymerase iota, mutagenesis, cancer, oxidative DNA damage, polycyclic aromatic hydrocarbons

Co-Authorship Statement

Chapter 2: Structural basis of ribonucleotide discrimination by a Y family DNA polymerase

Dr. Hong Ling contributed to structural interpretation and manuscript preparation.

Dr. Zucui Suo provided the Dpo4 Y12A expression plasmid.

Chapter 3: Structural basis of error-prone replication and stalling opposite a template thymine base by human DNA polymerase iota

Dr. Hong Ling contributed to structural interpretation and manuscript preparation.

Chapter 4: A unique active site promotes error-free replication opposite an 8-oxo-guanine lesion by human DNA polymerase iota

Dr. Hong Ling contributed to structural interpretation and manuscript preparation.

Chapter 5: DNA replication opposite a nitrated polycyclic aromatic hydrocarbon lesion by human DNA polymerase iota

Dr. Hong Ling contributed to structural interpretation and manuscript preparation.

Dr. Ashis Basu synthesized the N-[deoxyguanosine-8-yl]-1-amino-pyrene lesion DNA.

Dr. Zucui Suo performed the running start replication assays for pol η , pol κ , and pol ι .

Dedication

This thesis is dedicated to my parents. Without your unconditional love and support, none of this would be possible.

Acknowledgments

I would first like to thank my supervisor Dr. Hong Ling for all her training, guidance and support throughout my PhD. You have always taught me to strive for the best and have made me a better scientist as a result.

I would also like to thank Dr. Guanxin Xing for his help and advice with projects and experimental methods. I thank my committee members Dr. Eric Ball and Dr. Brian Shilton for their help and guidance. I want to acknowledge all the past and present members of the Ling lab and thank beam line support at 24ID-C/E at the Advanced Photon Source (APS) in Argonne National Laboratory. This work was supported by the Canadian Institutes for Health Research (CIHR).

I wish to thank my brother Daniel for being an outstanding role model and showing me nothing but kindness, support, and friendship throughout my life.

I thank Abi and Zahra Madadi for all their love and support and for always treating me as one of their own.

To my beloved Parvaz, your undying love and affection has given me strength and purpose throughout hard times and has guided me to what truly matters in life. This thesis marks the end of one chapter of our lives. May the next be filled with health, happiness, passion, and harmonious growth.

Table of Contents

CERTIFICATE OF EXAMINATION	ii
Abstract	iii
Co-Authorship Statement.....	v
Dedication	vi
Acknowledgments.....	vii
Table of Contents	viii
List of Tables	xi
List of Figures	xii
List of Abbreviations	xiv
List of Appendices	xvi
Chapter 1	1
1 Introduction	1
1.1 DNA damage and replcation.....	1
1.2 Translesion DNA synthesis.....	6
1.3 Y family DNA polymerases.....	8
1.3.1 History.....	8
1.3.2 Structure.....	12
1.3.3 Members and functions.....	16
1.4 Implications.....	21
1.5 Scope of thesis	21
1.5.1 Hypothesis.....	22
1.5.2 Objectives	22
1.6 References.....	23
Chapter 2.....	34
2 Structural basis of ribonucleotide discrimination by a Y family DNA polymerase	34
2.1 Introduction.....	34
2.2 Results.....	36
2.3 Discussion.....	43
2.4 Methods.....	48
2.5 References.....	49
Chapter 3.....	53

3	Structural basis of error-prone replication and stalling opposite a thymine base by human DNA polymerase ι	53
3.1	Introduction.....	53
3.2	Results.....	55
3.3	Discussion.....	76
3.4	Methods.....	79
3.5	References.....	83
	Chapter 4.....	86
4	A unique active site promotes error-free replication opposite an 8-oxo-guanine lesion by human DNA polymerase ι	86
4.1	Introduction.....	86
4.2	Results and Discussion	88
4.3	Methods.....	105
4.4	references	108
	Chapter 5.....	112
5	DNA replication through a nitrated polycyclic aromatic hydrocarbon lesion by human DNA polymerase ι	112
5.1	Introduction.....	112
5.2	Results.....	114
5.3	Discussion.....	130
5.4	Methods.....	134
5.5	References.....	137
	Chapter 6.....	142
6	Discussion	142
6.1	Summary.....	142
6.2	Mechanistic insights.....	143
6.2.1	Incoming nucleotide selection	143
6.2.2	Undamaged DNA replication	144
6.2.3	Lesion bypass.....	145
6.2.4	Rational drug design	147
6.3	Future directions	148
6.3.1	Function	148
6.3.2	Regulation	149
6.4	Conclusions.....	150

6.5 References.....	151
Appendices.....	154
Curriculum Vitae	155

List of Tables

Table 2.1	Summary of crystallographic data for Dpo4 Y12A structures.....	39
Table 3.1	Summary of crystallographic data for polt template thymine structures	56
Table 4.1	Summary of crystallographic data for polt 8-oxo-guanine structures.....	92
Table 5.1	Summary of crystallographic data for aminopyrene lesion structures	119

List of Figures

Figure 1.1 Structure and base pairing of DNA nucleotide bases	2
Figure 1.2 Structural comparison of different DNA nucleotide lesions	4
Figure 1.3 The translesion DNA synthesis pathway	7
Figure 1.4 Structure based sequence alignment of Y family polymerases	9
Figure 1.5 Domain organization of Y family DNA polymerases	11
Figure 1.6 The active site of a Y family DNA polymerase	13
Figure 1.7 Comparison of replicative and Y family polymerases	15
Figure 1.8 Structural comparison of human Y family polymerases	18
Figure 2.1 Incorporation of dNTP and NTP by Dpo4 and Dpo4 Y12A mutant.....	37
Figure 2.2 Superposition of Dpo4 wild type and Dpo4 Y12A ternary complexes.....	40
Figure 2.3 Conformation of replicating base pairs in the Y12A ternary structures.....	42
Figure 2.4 Comparison of incoming nucleotides between Dpo4 and Dpo4 Y12A	45
Figure 3.1 The effect of calcium on polt replication.....	58
Figure 3.2 Comparison of polt-DNA-nucleotide ternary structures	61
Figure 3.3 Structure of T:ddADP showing template DNA ‘U-turn’ stabilization.....	64
Figure 3.4 Polt and Dpo4 active site comparison	68
Figure 3.5 Base stacking and hydrogen bonding of replicating base pairs.....	71
Figure 3.6 The role of the finger domain in nucleotide incorporation specificity	74
Figure 4.1 8-oxo-guanine and nucleotide specificity of Y family DNA polymerases	90

Figure 4.2 8-oxo-G:dNTP positioning in the Polt/DNA/nucleotide structures	94
Figure 4.3 Replicating base pairs in polt/8-oxo-G structures	97
Figure 4.4 Structural comparison of 8-oxo-guanine replication	102
Figure 5.1 1-nitropyrene guanine attachment and Y family DNA polymerase activity	117
Figure 5.2 Superposition of polt-APG ternary structures	121
Figure 5.3 Replicating base pair conformations in the polt-APG ternary structures	123
Figure 5.4 APG lesion conformation and replication stalling	125
Figure 5.5 APG stalling characteristics of different Y family DNA polymerases	127
Figure 5.6 Superposition of Dpo4 APG extension structure	129
Figure 5.7 APG positioning in a stalled Dpo4 extension complex	131

List of Abbreviations

A	adenine
APG	N-[deoxyguanosine-8-yl]-1-aminopyrene
ATP	adenosine triphosphate
BER	base excision repair
BRCT	C-terminal domain of a breast cancer susceptibility protein
BSA	bovine serum albumin
C	cytosine
CTP	cytidine triphosphate
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
ddATP	dideoxyadenosine triphosphate
dGTP	deoxyguanosine triphosphate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
EDTA	ethylenediaminetetraacetic acid
G	guanine
GST	glutathione S-transferase
GTP	guanosine triphosphate
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
MES	2-(N-morpholino)ethanesulfonic acid
MME	monomethyl ether

MUTY	gene encoding adenine glycosylase
NER	nucleotide excision repair
1-NP	1-nitropyrene
NT	nucleotide
NTP	nucleotide triphosphate
5-OHU	5-hydroxy-uracil
8-oxo-G	7,8-dihydro-8-oxo-guanine
PAH	polycyclic aromatic hydrocarbon
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PEG	polyethylene glycol
Pol ι	human DNA polymerase iota
Pol κ	human DNA polymerase kappa
Pol η	human DNA polymerase eta
SUMO	small ubiquitin-like modifier
T	thymine
TEV	tobacco etch virus
TLS	trans lesion DNA synthesis
TRIS	2-Amino-2-(hydroxymethyl)-1,3-propanediol
TT	thymine-thymine
TTP	thymidine triphosphate
U	uracil
UV	ultraviolet
XP-V	xeroderma pigmentosum variant
XRCC1	X-ray repair complementing defective repair in Chinese hamster cells 1

List of Appendices

Appendix 1: Copyright premission for previously published work (Chapter 3)	152
---	-----

Chapter 1

1 Introduction

1.1 DNA damage and replication

The formation of the deoxyribonucleic acid (DNA) molecule was a key event that enabled life forms to exist on planet earth. The chemical nature of DNA allowed its utilization as a storage device for cellular information (Figure 1.1A). However, DNA is susceptible to damage and decay (Lindahl, 1993), which may ultimately lead to detrimental processes, such as aging and disease (de Boer et al., 2002; Hoeijmakers, 2009; Lu et al., 2004). Thus, from the very beginning of life, organisms have been under constant pressure to maintain their genetic information in an unadulterated state. Ironically, the same DNA fragility that threatens the survival of organisms is also responsible for their subsequent evolution (Friedberg, 2003). DNA damage induces mutagenesis, which in turn gives rise to DNA sequence variation, resulting in the evolution of species throughout time. Thus, nature must perform a delicate balancing act between allowing some DNA mutagenesis to occur so that the fitness of a species may be improved, and preventing excessive DNA mutagenesis, which would ultimately kill the organism. If this mutagenic balancing act was skewed to either end of the spectrum, it would inevitably lead to extinction in both scenarios.

DNA molecules are composed of four nucleotide bases: adenine (A), thymine (T), guanine (G), and cytosine (C). These bases pair together to form double stranded helical molecules with characteristic major and minor grooves (Figure 1.1A). Nucleotide pairing

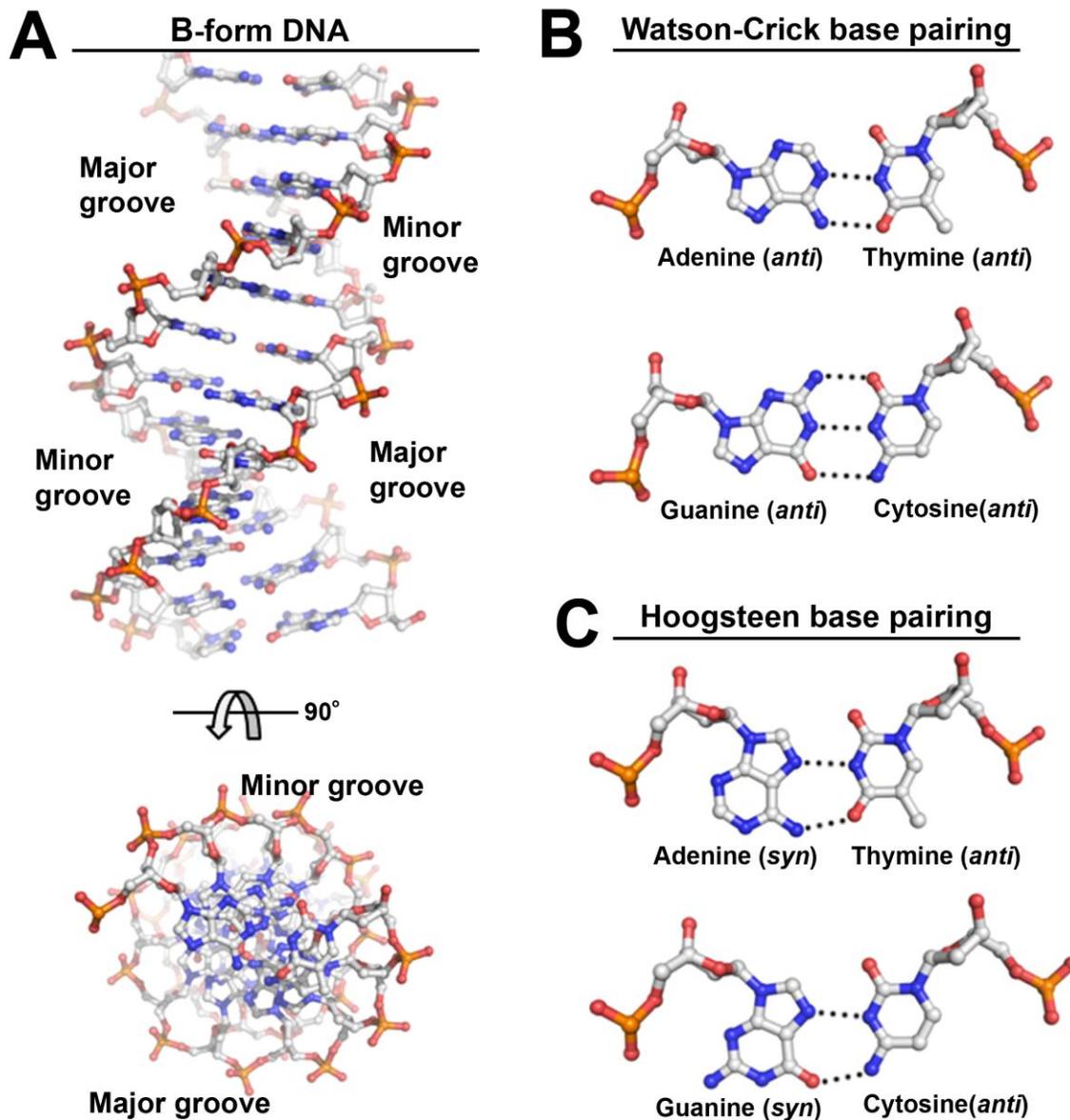


Figure 1.1. Structure and base pairing of DNA nucleotide bases. (A) Side and top view of double stranded B-form DNA with major and minor grooves indicated. (B) Watson-Crick base pairing between adenine and thymine and guanine and cytosine. (C) Hoogsteen base pairing between adenine and thymine and guanine and cytosine. Hydrogen bonds are shown as dotted black lines.

occurs with complementary bases forming strong hydrogen bonding interaction networks (Figure 1.1B). Adenine and guanine are classified as purine bases and pair with thymine and cytosine respectively, which are pyrimidine bases. The standard nucleotide interaction is known as Watson-Crick base pairing, where both purine and pyrimidine bases are in *anti* conformations (Figure 1.1B). Non-standard interactions such as Hoogsteen base pairing can also occur, where purine nucleotides adopt a *syn* conformation (Figure 1.1C)

DNA damage is often referred to as lesions or adducts, where one or multiple DNA nucleotides are covalently modified, leading to chemical and structural alterations in the DNA. These nucleotide alterations can be in forms of chemical compound attachments, nucleotide base linkages, or single atom conversions (Figure 1.2). The most common external DNA damaging agent is ultraviolet radiation (UV) from sunlight exposure, which induces cross linking within DNA (Figure 1.2B) (Freeman et al., 1989; Setlow, 1966). Other external damaging agents include specific chemical compounds that are produced naturally or by human activity. These compounds, typically ingested or inhaled, are absorbed into the body where they can then chemically attach to DNA nucleotide bases (Figure 1.2A) (Braithwaite et al., 1998; Feldman et al., 1978). Internal DNA damaging agents mainly arise from by-products of normal cellular metabolism such as reactive oxygen species from aerobic respiration (Figure 1.2C) (De Bont and van Larebeke, 2004). It is estimated that endogenous oxidative DNA damage alone produces 1000 lesions per day in the average human cell (Collins, 1999). Interestingly, DNA damage can also occur in the absence of damaging agents, in a process known as spontaneous deamination, where a cytosine is converted into a non-DNA nucleotide base

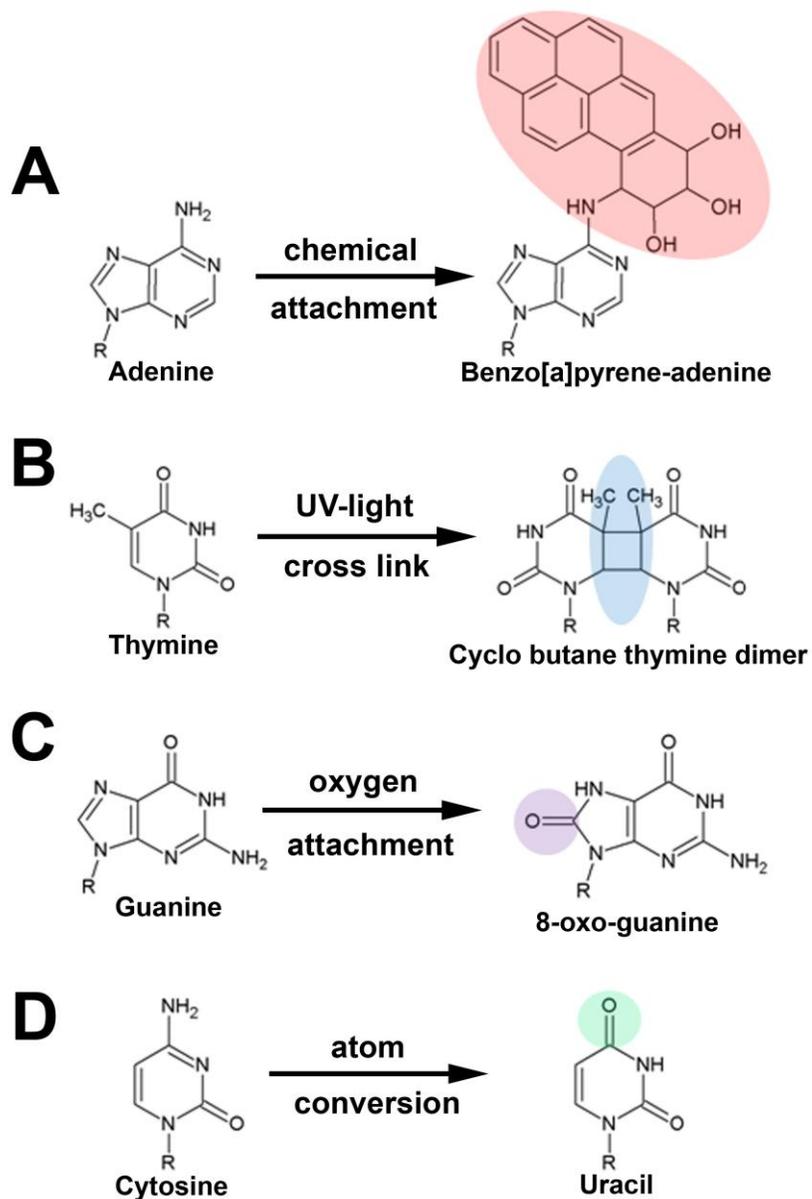


Figure 1.2. Structural comparison of different DNA nucleotide lesions. (A) Attachment of benzo[a]pyrene compound to adenine. (B) UV light induced thymine dimer. (C) Oxidative damage of guanine producing the 8-oxo-guanine lesion. (D) Spontaneous deamination of cytosine into uracil. Sites of nucleotide damage are indicated with coloured shading.

(Uracil) (Figure 1.2D) (Duncan and Miller, 1980). Clearly, DNA is subject to continual damage and insults from a multitude of exogenous and endogenous sources. How this damage translates to sequence variation and mutagenesis is primarily through DNA repair and replication. All organisms contain repair mechanisms in order to correct sites of DNA damage. Two highly conserved and fundamental repair pathways are Base Excision Repair (BER) (Doetsch and Cunningham, 1990) and Nucleotide Excision Repair (NER) (de Laat et al., 1999). These processes remove sites of damage from the DNA strand and then fill in the subsequent gap in a mechanism akin to 'cut and paste' (Friedberg, 2001; Krokan et al., 2000). Unsurprisingly, defects in BER or NER processes induce carcinogenesis and tumour formation (Cheadle et al., 2003; Farrington et al., 2005; Sijbers et al., 1996).

Although BER and NER mechanisms repair sites of DNA damage, many lesions still persist in the genome. These lesions will ultimately be encountered by replicative polymerases during DNA replication. The main task of replicative polymerases is to duplicate the full set of genomic DNA before cell division to ensure that daughter cells receive a complete copy of the genome. Replicative polymerases are highly accurate when copying undamaged DNA molecules, making on average 1 mistake for every 1000,000 bases replicated (McCulloch and Kunkel, 2008; Thomas et al., 1991). However, when they encounter damaged nucleotide bases, their accuracy and activity is dramatically altered. For instance, when C spontaneously deaminates into the non-DNA base U, replicative polymerases misinterpret the lesion as a T base, resulting in an A misincorporation (Wardle et al., 2008). Similarly, when a G nucleotide is oxidized into an 8-oxo-guanine lesion, replicative polymerases incorporate an A nucleotide instead of the

correct G base (Hsu et al., 2004). Thus, high fidelity polymerases induce DNA mutagenesis when replicating through small, chemically altered nucleotide lesions. For bulky or distorting DNA lesions however, different replication mechanisms must be employed.

When replicative polymerases attempt to bypass large DNA adducts, their enzymatic activity is inhibited and the replication fork comes to a halt (Hsu et al., 2005). This effect is due to tight and restrictive active sites which prevent the accommodation of bulky or distorting DNA lesions (Beese et al., 1993; Doublet et al., 1998; Li et al., 1998; Swan et al., 2009). If replication stalling were to persist, the cell may undergo senescence or apoptosis. Thus, the cell requires mechanisms to rescue lesion stalled replication forks. One of these mechanisms involves the recruitment of specialized DNA polymerases that perform DNA lesion replication in a process known as translesion DNA synthesis (TLS).

1.2 Translesion DNA synthesis

When replicative polymerases stall at sites of DNA damage, the stalled replication fork must be rescued in order for DNA duplication to proceed. The rescue operation can occur using two main mechanisms; error-free replication repair via homologous recombination (Berdichevsky et al., 2002) or error-prone bypass via translesion DNA synthesis (Figure 1.3) (Andersen et al., 2008). Translesion DNA synthesis requires specialized DNA polymerases belonging to the Y family, which are recruited to stalled replication forks (Figure 1.3A). The recruitment and regulation of Y family polymerases is coordinated by the mono-ubiquitination of the proliferating cell nuclear antigen

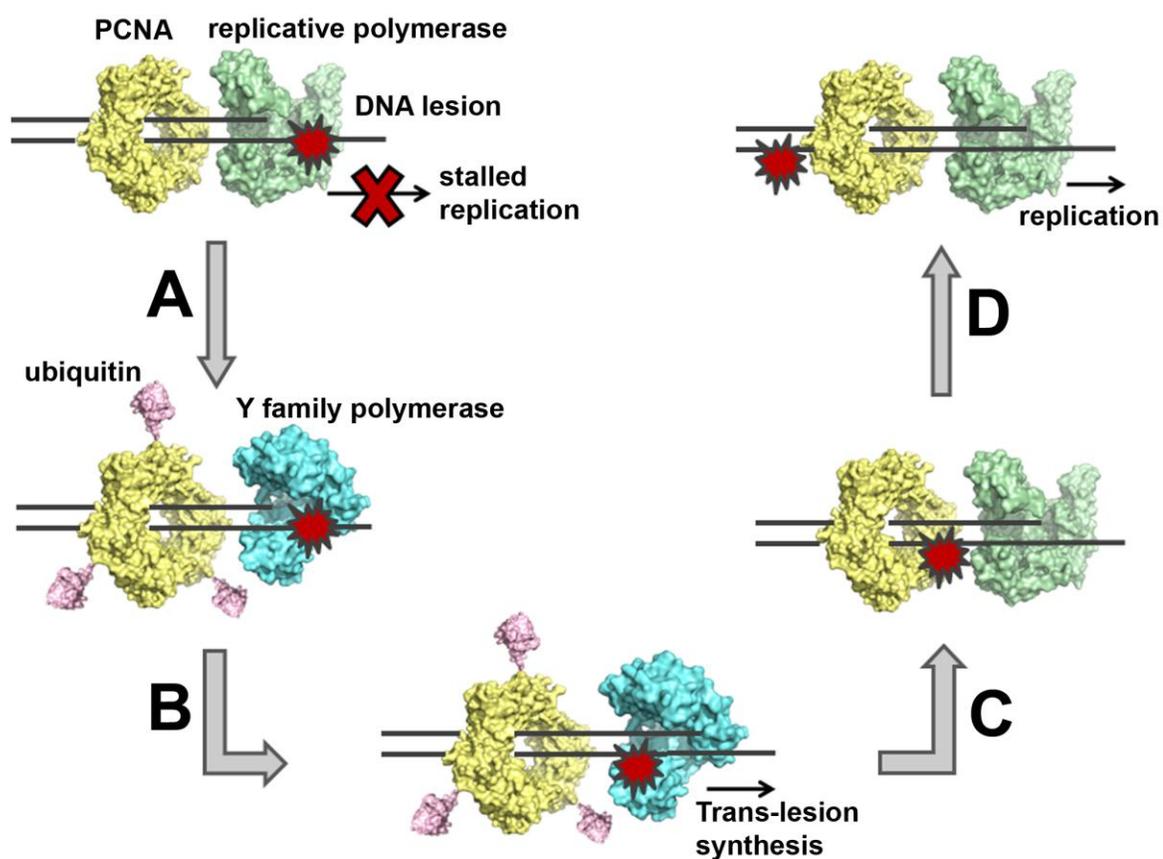


Figure 1.3. The translesion DNA synthesis pathway (A) Polymerase switching between lesion stalled replicative polymerase (green, PDB: 1T7P) and Y family polymerase (blue, PDB: 2ALZ). Mono-ubiquitination (pink, PDB: 1UBQ) of the proliferating cell nuclear antigen (PCNA, yellow, PDB: 1AXC) signals the polymerase switching mechanism. (B) The Y family polymerase bypasses the DNA lesion via translesion DNA synthesis (TLS). (C) The replicative polymerases is recruited back to the replication fork after lesion bypass. Polymerase switching is now coordinated by the de-ubiquitination of PCNA. (D) The replicative polymerase re-initiates DNA replication.

(PCNA) cofactor molecule in Eukaryotes, which specifically signals the TLS pathway (Figure 1.3A) (Haracska et al., 2004; Hoege et al., 2002; Jansen et al., 2007; Plosky and Woodgate, 2004). Once the Y family polymerases have been recruited to the replication fork, they replace the replicative polymerase at the template-primer junction and incorporate nucleotides opposite the damaged nucleotide bases (Figure 1.3B) (Zhuang et al., 2008). Once the lesion has been bypassed by Y family polymerases, the replicative polymerase resumes its position on the DNA strand to proceed with duplication (Figure 1.3C, D). This polymerase switching process is coordinated by the de-ubiquitination of the PCNA molecule (Zhuang et al., 2008). It is important to note that TLS is not a repair process but a bypass mechanism, as the DNA lesion still remains within the genome after replication. The lesion can be repaired at a later time by mechanisms previously mentioned (BER, NER). Although translesion DNA synthesis rescues stalled replication forks, it also contributes to DNA mutagenesis due to the error-prone nature of Y family DNA polymerases (Chabes et al., 2003). Compared to replicative polymerases, members of the Y family have low fidelity with replication errors occurring once in every 100-10,000 bases replicated (Fiala and Suo, 2004; Matsuda et al., 2000; Tissier et al., 2000; Zhang et al., 2000). In fact, the error-prone activity of Y family polymerases is what initially led to their discovery.

1.3 Y family DNA polymerases

1.3.1 History

By the 1980s it was becoming evident that DNA mutagenesis induced by damaging agents was the result of cellular proteins and not the damaging agents themselves. These

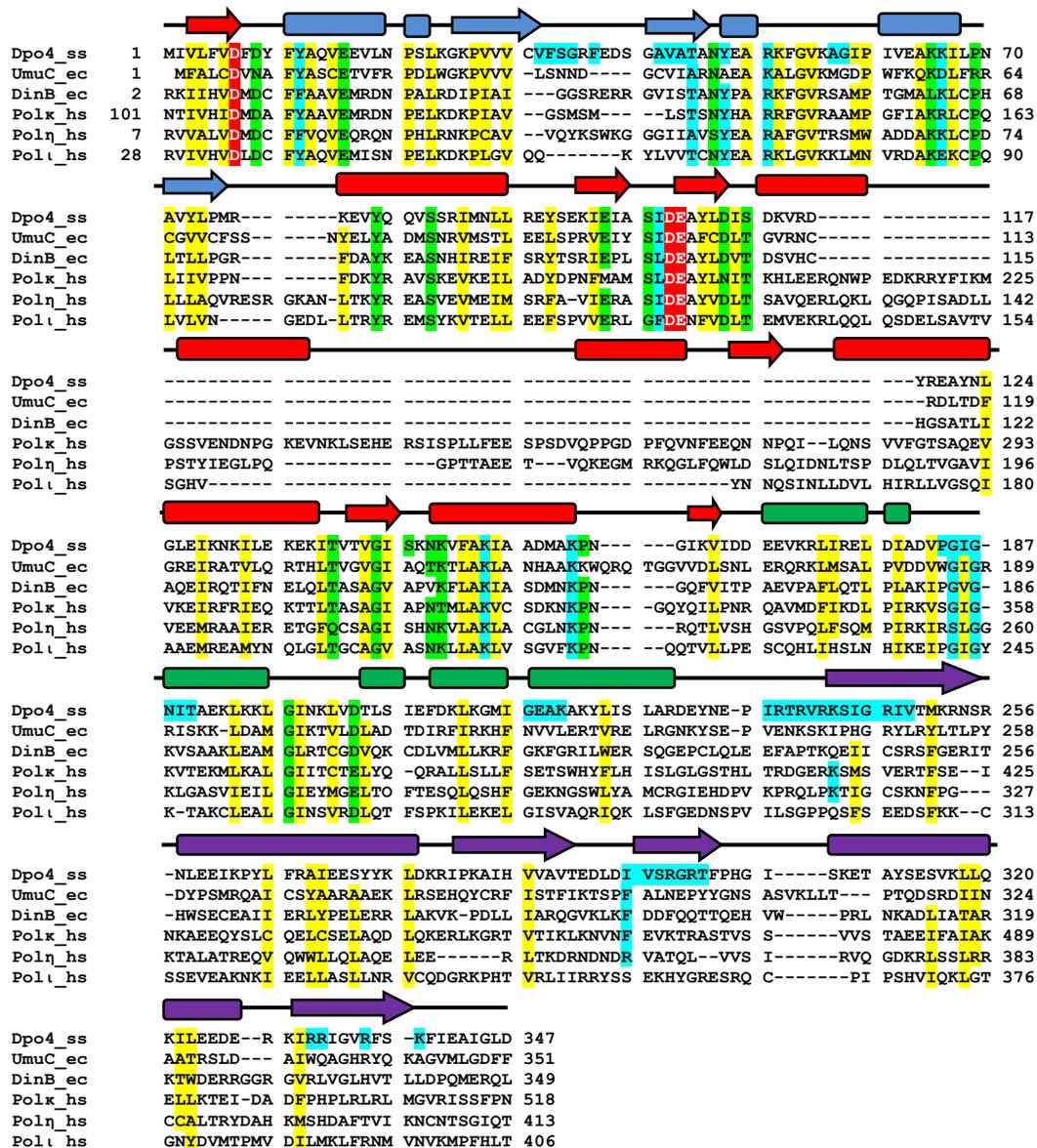


Figure 1.4. Structure based sequence alignment of Y family polymerases. *S. solfataricus* Dpo4, *E. coli* UmuC, *E. coli* DinB, and human polk, polη, and polι are shown. Secondary structure elements for polη are shown above the sequence and are coloured based on domains: palm (red), finger (blue), thumb (thumb), and little finger (purple). Rectangles are α helices and arrows are β sheets. Conserved residues are coloured based on hydrophobic core formation (yellow), structural integrity (green), polymerase activity (red), and DNA binding (blue). Adapted from (Ling et al., 2001).

conclusions were drawn from the observation that mutations in specific genes decreased the mutagenic signature in cells exposed to DNA damaging agents (Kato and Shinoura, 1977; Lawrence and Christensen, 1976; Lemontt, 1971; Steinborn, 1978). Although the mutagenic mechanism of these proteins remained unknown, it was becoming clear through the expanding field of molecular genetics that many organisms contained orthologs of these mutagenic proteins (Kulaeva et al., 1996; Larimer et al., 1989; McDonald et al., 1997; Ohmori et al., 1995) (Figure 1.4). Thus a ‘superfamily’ of mutation-associated enzymes was born. The enzymatic function of these newly discovered mutagenic proteins began to be revealed when the *S. cerevisiae* Rev1 protein was shown to incorporate cytosine nucleotides into a growing DNA primer strand (Nelson et al., 1996). Two landmark studies followed, which demonstrated that *S. cerevisiae* Rad30 protein (pol η) used all four nucleotides for DNA replication and had the ability to replicate through thymine-thymine (TT) dimers similar to undamaged DNA (Johnson et al., 1999; Masutani et al., 1999). Soon after, other members of this mutagenic family were shown to replicate DNA and bypass sites of DNA lesions (Gerlach et al., 1999; McDonald et al., 1999; Reuven et al., 1999; Tang et al., 1999; Tissier et al., 2000; Wagner et al., 1999). These unique enzymes were a newly discovered type of DNA polymerase and were subsequently classified as the Y family. Thus, the field of translesion DNA synthesis began (Ohmori et al., 2001; Woodgate, 1999). These proteins were not initially identified as DNA polymerases due to the absence of sequence conservation with replicative polymerases. It was eventually revealed that Y family DNA polymerases are conserved in all three domains of life (Archaeobacteria, Eubacteria, and Eukaryotes). Thus, the ancient and fundamental

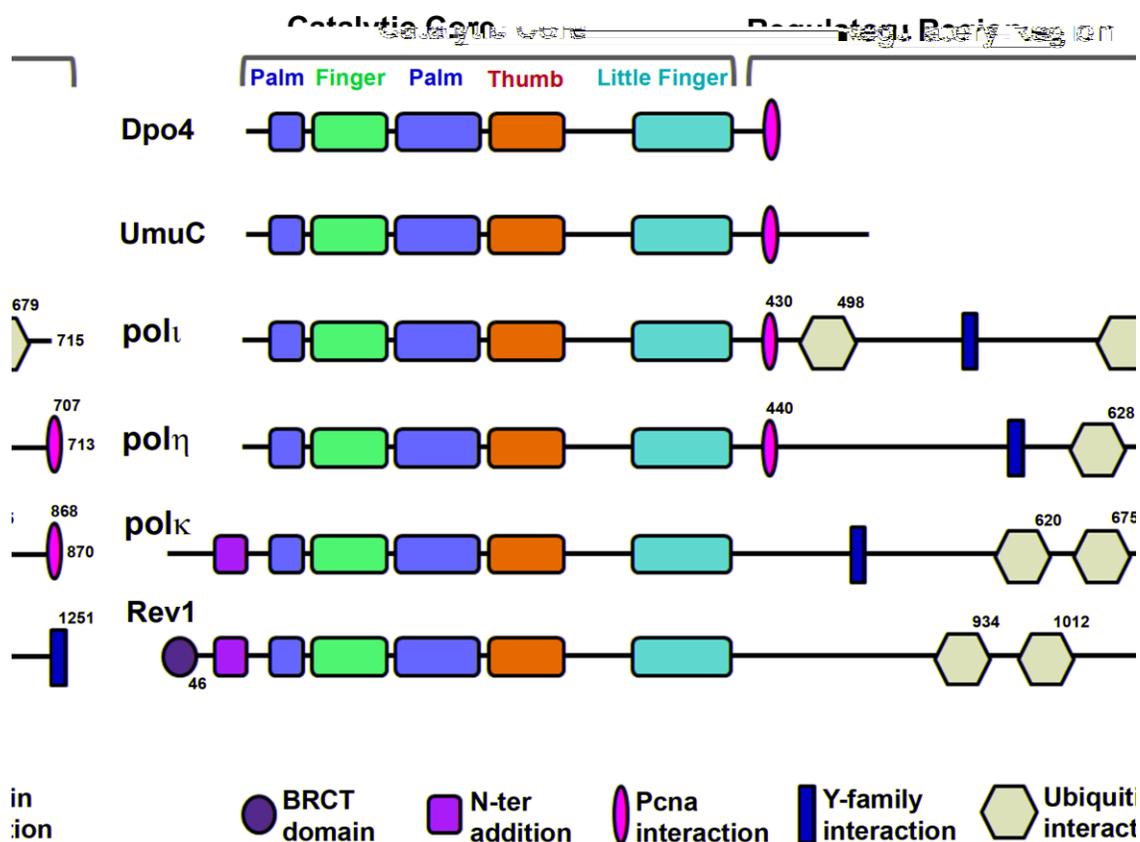


Figure 1.5. Domain organization of Y family DNA polymerases. *S. solfataricus* Dpo4, *E. coli* UmuC, and human pol κ , pol η , and polt are shown. The catalytic core domain organization consisting of palm (red), finger (blue), thumb (green), and little finger (purple) are highly conserved within all Y family polymerases. Pol κ and Rev1 contain additional N-terminal domains (orange) not found in other Y family polymerases. Rev1 also contains an N-terminal BRCT domain (brown) of unknown function. The domain organization of the C-terminal regulatory region is not conserved but has common protein interacting motifs for PCNA (yellow), ubiquitin (cyan), and other Y family polymerases (red). Adapted from Yang and Woodgate (Yang and Woodgate, 2007).

importance of these enzymes truly came to light.

1.3.2 Structure

All Y family DNA polymerases contain a highly conserved N-terminal catalytic core, which is responsible for polymerase activity. In addition, there is a non-conserved C-terminal regulatory region that coordinates protein-protein interactions with binding partners such as PCNA and ubiquitin (Figure 1.5) (Bienko et al., 2005; Haracska et al., 2005; Haracska et al., 2001; Haracska et al., 2001; Haracska et al., 2002; Plosky et al., 2006; Xing et al., 2009). The catalytic core is composed of four domains referred to as the finger, thumb, palm, and little finger (Figure 1.4, 1.5). The finger, thumb, and palm domains of Y family polymerases have similar organization, and functional properties as replicative polymerases (despite the lack of primary sequence conservation between these families) (Yang, 2005). The palm domain contains the catalytic residues which coordinate two active site metal ions (Figure 1.6B). The two-ion-assisted catalytic mechanism appears to be conserved within all DNA polymerases (Yang, 2008; Yang et al., 2006). The B site metal ion stabilizes the incoming nucleotide phosphate groups, while the A site metal ion facilitates phosphodiester bond formation between the incoming nucleotide α phosphate and the 3'-OH group of the primer strand (Vaisman et al., 2005). This reaction induces primer elongation and allows the template DNA strand to be replicated. The finger domain forms a lid on top of the replicating base pair, which is responsible for incoming nucleotide selection (Figure 1.6A) (Glick et al., 2001; Ling et al., 2001). The thumb domain forms part of the DNA binding cleft, which makes contact to the minor groove of the DNA molecule (Figure 1.6) (Ling et al., 2001).

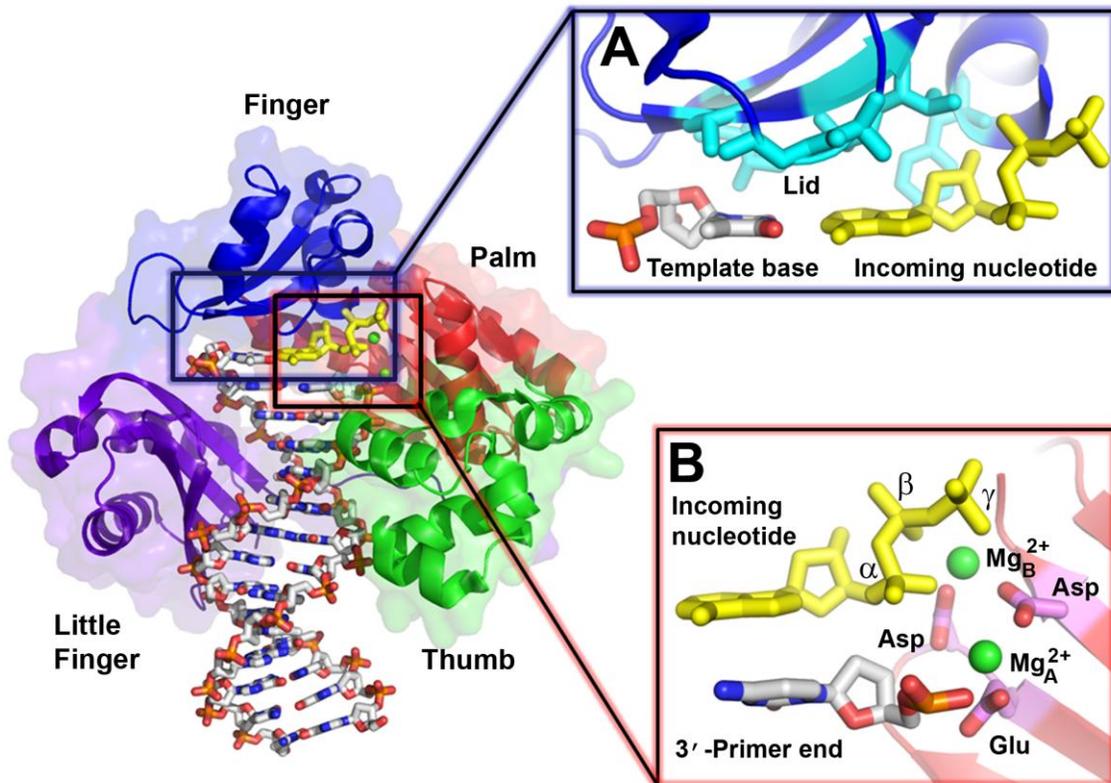


Figure 1.6. The active site of a Y family DNA polymerase (A) Interaction between the lid region (cyan) of the finger domain (blue) and the replicating base pair. (B) Coordination of active site metal ions (green) by conserved catalytic residues in the palm domain (red). The incoming nucleotide is shown in yellow and the primer and template bases are coloured grey. The enzyme is Archaeal Y family polymerase Dpo4 (PDB: 2AGQ).

The little finger domain is unique to the Y family polymerases. It has the lowest sequence conservation among the domains within the catalytic core, and is thought to contribute to the unique lesion bypass properties specific to each Y family polymerase (Boudsocq et al., 2004). The little finger domain contacts the major groove of the DNA molecule, completing the DNA binding cleft with the thumb domain (Figure 1.6) (Ling et al., 2001). The little finger domain is connected to the rest of the catalytic core through a long linker region, which allows large conformational changes upon DNA binding and replication (Silvian et al., 2001; Uljon et al., 2004; Wong et al., 2008). The little finger has been shown to be important in lesion bypass activity, as swapping this domain between two Y family polymerases causes the lesion bypass properties to also be swapped (Boudsocq et al., 2004).

Y family polymerases differ mechanistically from replicative polymerases in several ways. The finger domains of replicative DNA polymerases undergo an induced fit conformation upon nucleotide binding. Specific helices clamp down on the replicating base pair and initiate enzyme catalysis when a proper Watson-Crick base pair is formed (Figure 1.7C) (Franklin et al., 2001; Li et al., 1998; Ollis et al., 1985; Rothwell et al., 2005; Santoso et al.). The induced fit conformation creates a high fidelity enzyme, but also restricts bulky lesion bases from entering the active site. In contrast, Y family polymerase finger domains lack these specific clamping helices and thus are generally in an 'open' conformation (Figure 1.7B, D). In addition, the finger and thumb domains are much smaller than their replicative counterparts, which create an open and solvent accessible active site for DNA lesion accommodation (Figure 1.7A, B). This

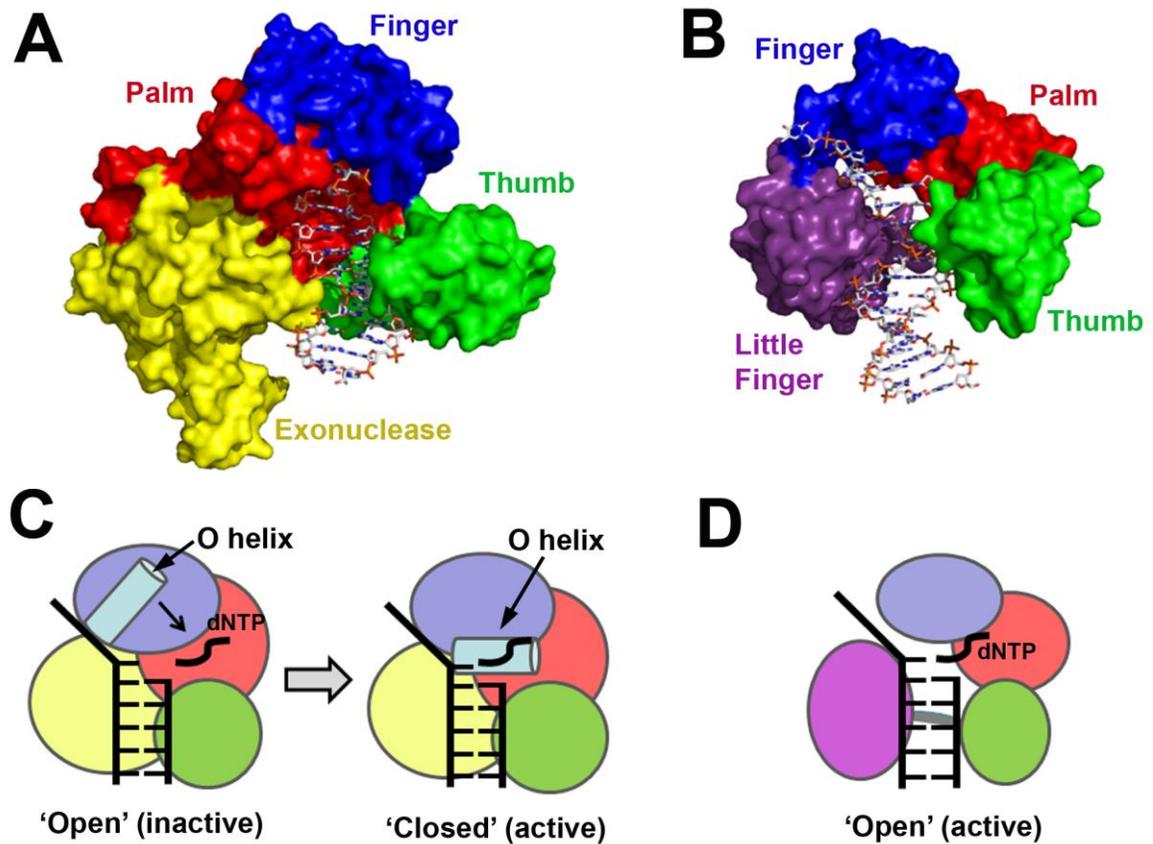


Figure 1.7. Comparison of replicative and Y family polymerases. **(A)** Structure of T7 replicative polymerase (PDB: 1T7P) coloured based on domains: palm (red), finger (blue), thumb (green), and exonuclease (yellow). **(B)** Structure of Archaeal Y family polymerase Dpo4 (PDB: 2AGQ) with same domain colouring as T7 polymerase in panel A. The unique little finger domain is coloured purple. **(C)** Schematic of induced fit conformation mechanism in replicative polymerases. **(D)** Schematic of the open and active conformation mechanism of Y family polymerases.

structural feature allows the accommodation of bulky and distorting DNA lesions, but also increases error-prone replication (Friedberg et al., 2001). In addition, Y family DNA polymerases lack an exonuclease domain and thus have no proofreading ability, which further contributes to their error-prone nature.

The open active site of Y family polymerases allows permissive base pairing to exist, which is a crucial requirement for the bypass of many DNA lesions. Compared to replicative polymerases, the finger domains of Y family polymerases have minimal contacts to the DNA and replicating base pair. Thus, how these enzymes preferentially select incoming nucleotides is not fully understood. Interestingly, the finger domain contains highly variable loop regions which contact the replicating base pair and are thought to account for differences in nucleotide specificity between Y family polymerases (Figures 1.4, 1.6).

1.3.3 Members and functions

Eukaryotes have several Y family DNA polymerases, each with characteristic enzyme activities. Humans contain four Y family DNA polymerases (polymerase eta; pol η , polymerase iota; pol ι , polymerase kappa; pol κ , and Rev1) that likely coordinate the bypass of hundreds of different DNA lesions created in the genome. Eubacteria and Archaeobacteria contain no more than two Y family DNA polymerases. Thus, Eubacteria and Archaeobacteria Y family polymerases require more broad range activity than their Eukaryotic counterparts as they solely perform all the TLS requirements of the cell. It is remarkable that so few enzymes are able to bypass the vast repertoire of damage

created in the genome. However, this broad range activity also foreshadows the error-prone and mutagenic nature of these polymerases.

Rev1 was the first Y family polymerase identified and is only found in Eukaryotic organisms (Lemontt, 1971) with no apparent orthologs in Eubacteria or Archaeobacteria. This polymerase has a unique enzymatic property of incorporating only incoming C nucleotides (dCTP) opposite undamaged and damaged DNA substrates (Haracska et al., 2002; Zhang et al., 2002). Thus, Rev1 is known as a template independent deoxycytidyl transferase. Structural studies have revealed that the protein molecule ejects the template base out of the active site and replaces it with an arginine residue (Nair et al., 2005; Swan et al., 2009) from a unique N-terminal domain (Figure 1.8A). The arginine residue preferentially hydrogen bonds with dCTP, allowing its subsequent incorporation into the DNA primer strand (Figure 1.8A). Although the biological function of this unique dCTP incorporation by Rev1 is unknown, it has been suggested that this activity may allow for the error-free bypass of minor groove adducted guanine lesions (Washington et al., 2004).

Polk is closely related to the most ancient Y family polymerases, with orthologs in Eubacteria (DinB, *E.coli*,) and Archaeobacteria (Dpo4, *S.solfataricus*) (Boudsocq et al., 2001; Jarosz et al., 2007; Lee et al., 2005). Interestingly, polk contains an extra N-terminal domain similar to Rev1, which is not found in the other Y family polymerases. The extra N-terminal domain of polk has been shown through X-ray crystallography to form clamping helices, which enclose the solvent exposed major groove and encircle the DNA helix (Lone et al., 2007) (Figure 1.8B). Polk has one of the highest fidelities of any Y family DNA polymerase (Ohashi et al., 2000). In addition, polk is highly capable of

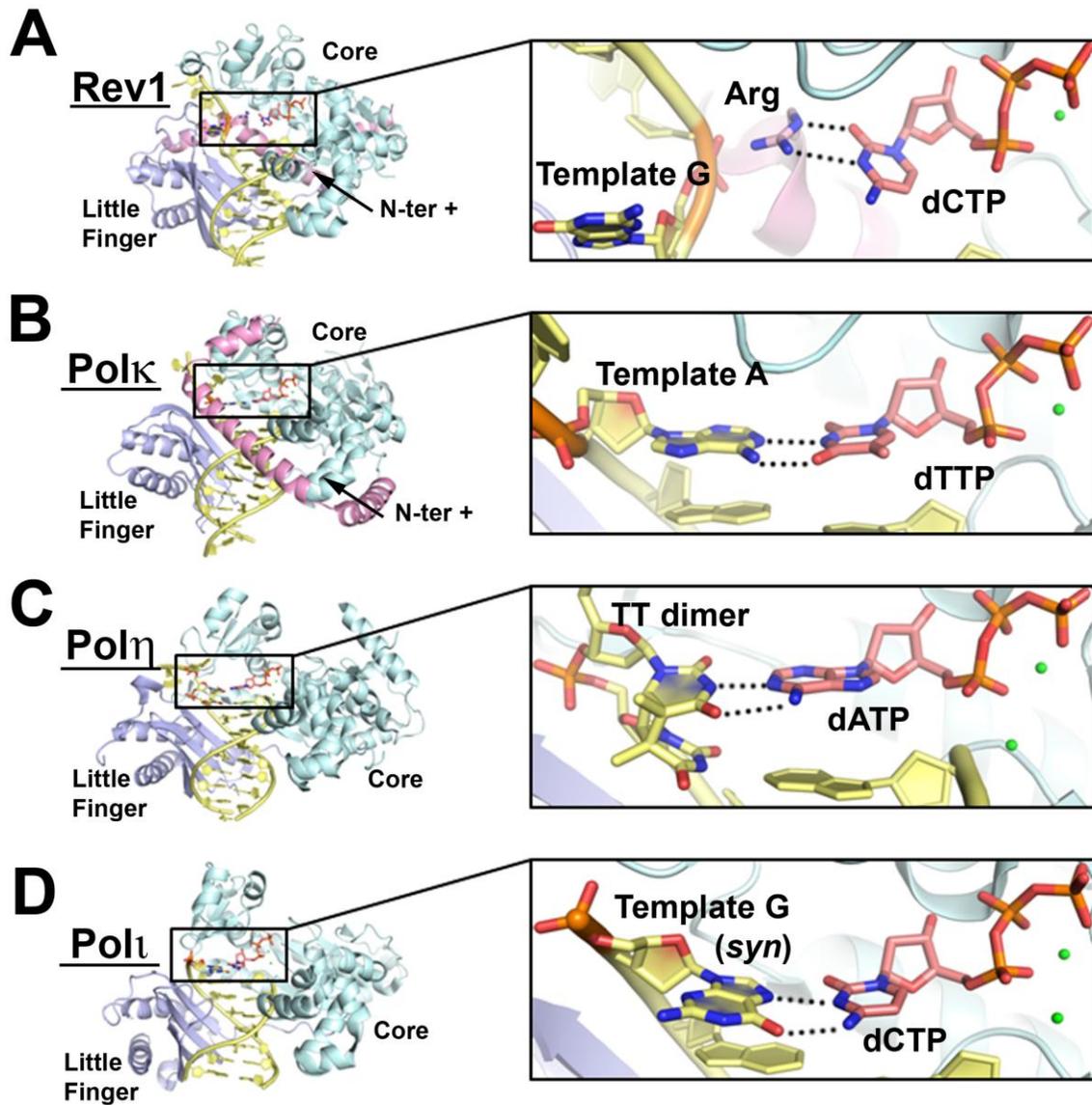


Figure 1.8. Structural comparison of human Y family polymerases. (A) Rev1 template independent dCTP incorporation (PDB: 3GQC). (B) Polk replicating base pair with open minor groove (PDB: 2OH2). (C) Pol η error-free TT dimer replication (PDB: 3MR4). (D) Polt Hoogsteen base pair replication (PDB: 2ALZ). The finger, palm, and thumb domains are coloured cyan and labelled ‘core’. The little finger domain is coloured light blue and the additional N terminal domains (N-ter +) in polk and Rev1 are coloured pink.

extending primers beyond mis-matched bases (Washington et al., 2002). Thus, it has been suggested that polk may function as a DNA lesion extension polymerase (Haracska et al., 2002). Interestingly, cells depleted in polk display hypersensitivity to the carcinogenic compound benzo[a]pyrene, indicating a possible role in bypassing or extending bulky hydrophobic lesions (Ogi et al., 2002). Structural studies support this hypothesis based on the observed N terminal helices, which may protect hydrophobic lesions from solvent, and a large minor groove cavity in the active site, which may accommodate bulky DNA adducts (Lone et al., 2007) (Figure 1.8B).

Pol η is found in all Eukaryotic organisms with orthologs in Eubacteria (*E.coli*, PolV) (Lee et al., 2005) and is able to bypass many different types of DNA damage (Masutani et al., 2000; Zhang et al., 2000). However, Pol η appears to be specialized for replication opposite UV light induced T-T dimers, incorporating the correct A nucleotides with the same efficiency as undamaged DNA (Johnson et al., 1999; Masutani et al., 1999). This functional activity was confirmed *in vivo* with the discovery that defects in pol η lead to the disease Xeroderma Pigmentosum Variant Syndrome (XP-V) (Masutani et al., 1999). Patients with XP-V display hypersensitivity to UV light with skin tumour formation after sunlight exposure due to an inability to accurately replicate TT dimers. Structural studies have revealed that pol η is able to accommodate two template bases in its active site, which facilitates the bypass of cross linked thymine dimers (Biertumpfel et al., 2010; Silverstein et al.) (Figure 1.8C). The ability of pol η to replicate cross linked nucleotides also enables the bypass of replication arresting chemotherapeutic agents such as cisplatin (Alt et al., 2007). Thus, pol η is involved in cancer cell resistance to chemotherapeutic agents (Albertella et al., 2005). Interestingly, pol η also appears to function on

undamaged DNA, utilizing its error-prone replication for somatic hypermutation of immunoglobulin genes (Zeng et al., 2001).

Pol η is a paralog of Eukaryotic pol η and is the most recently evolved Y family DNA polymerase (McDonald et al., 2001). Pol η is only found in Metazoans, from *Drosophila melanogaster* to *Homo sapiens* (Ishikawa et al., 2001), indicating a highly specialized function, which still remains to be elucidated. Pol η is the only known DNA polymerase whose primary mode of DNA replication is via Hoogsteen base pairing (Johnson et al., 2006; Nair et al., 2006; Nair et al., 2004) (Figure 1.8D). In addition, pol η has the highest error-rate of any DNA polymerase on thymine bases, preferring to incorporate mismatched G over the correct A nucleotide (Zhang et al., 2000). Interestingly, this unique error-prone replication opposite thymine bases by pol η is responsible for the high rates of DNA mutagenesis observed in XP-V patients (Wang et al., 2007). These unique activities make pol η one of the most mechanistically interesting Y family polymerase. Although the function of pol η is unknown, it has been suggested that Hoogsteen base pairing may facilitate the bypass of minor groove adducted lesions by projecting the lesion attachment into the solvent exposed major groove (Nair et al., 2006). In addition, it has been shown that cells depleted in pol η become sensitive to oxidizing agents, suggesting a role for pol η in oxidative DNA damage bypass (Petta et al., 2008).

Clearly, the *in vivo* function and lesion specificity of Eukaryotic Y family DNA polymerases remains virtually unknown. One of the difficulties in classifying Y family polymerase function is that the Eukaryotic TLS pathway is performed by a combination of different Y family DNA polymerases (Prakash and Prakash, 2002; Shachar et al., 2009). These enzymes are shown to functionally interact, forming multi-polymerase

complexes that co-localize at damaged replication foci (Kannouche et al., 2003; Ohashi et al., 2004; Tissier et al., 2004). In addition, it has been shown that multiple Y family polymerases are necessary for the efficient bypass of single DNA lesions (Johnson et al., 2001; Johnson et al., 2000). Thus, these enzymes functionally overlap, making the classification of lesion specificity difficult.

1.4 Implications

The ability of Y family DNA polymerases to replicate through sites of DNA damage is fundamental to a cell's ability to tolerate DNA damaging agents. However, these enzymes also induce mutations within the genome and thus likely contribute to evolution, aging, carcinogenesis and disease. In addition, these specialized polymerases also contribute to cancer cell resistance against chemotherapeutic agents such as cisplatin by bypassing the drug modified sites on DNA (Albertella et al., 2005). Understanding how these enzymes mechanistically bypass DNA damage and how they select incoming nucleotides for incorporation will give great insight into the process of mutagenesis leading to cancer formation and evolution. Furthermore, these results may aid in the design of better chemotherapeutic agents that block the activity of these enzymes, as well as identifying mechanisms to alter or regulate Y family polymerase activity through drug-protein interactions.

1.5 Scope of thesis

The focus of this thesis is to mechanistically explain how Y family polymerases select incoming nucleotides opposite a variety of different DNA templates. Four main questions regarding incoming nucleotide selection are investigated. 1) How do Y family

polymerases differentiate between ribo and deoxyribonucleotides during DNA replication? 2) What is the basis of incoming nucleotide selection opposite undamaged DNA? 3) How does nucleotide selection change when replicating small non-distorting lesions? 4) How do Y family polymerases incorporate nucleotides opposite large and bulky hydrophobic lesions?

The model Y family polymerase Dpo4 from *Sulfolobus solfataricus* was chosen to study the first question of deoxyribonucleotide selection due to its generalizability to all members of the Y family. Human polt was chosen to study the three remaining questions regarding Y family polymerase DNA replication due to its extremely high error-rates on undamaged thymine bases, its ability to replicate small oxidative lesions in an error-free manner, and its unique mutagenic replication opposite bulky lesions created from environmental pollutants. Particular emphasis is placed on the finger domain, which directly contacts the replicating base pair and is likely to influence nucleotide specificity during DNA replication.

1.5.1 Hypothesis

The incoming nucleotide specificity of Y family DNA polymerases across different DNA templates is determined by the finger domain, particularly, the lid region, which forms the replicating base pair contact interface.

1.5.2 Objectives

- 1) To determine the structural basis of ribonucleotide discrimination during DNA replication by the model Y family polymerase Dpo4 (Chapter 2).

- 2) To elucidate the structural mechanism of mis-matched nucleotide incorporation opposite an undamaged thymine base by polt (Chapter 3).
- 3) To investigate the structural basis of error-free replication opposite a small, highly mutagenic 8-oxo-guanine lesion by polt (Chapter 4).
- 4) To investigate the structural basis of error-prone replication opposite a large, hydrophobic aminopyrene DNA lesion (Chapter 5).

1.6 References

- Albertella, M.R., Green, C.M., Lehmann, A.R. and O'Connor, M.J. (2005) A role for polymerase eta in the cellular tolerance to cisplatin-induced damage. *Cancer Res*, **65**, 9799-9806.
- Alt, A., Lammens, K., Chiocchini, C., Lammens, A., Pieck, J.C., Kuch, D., Hopfner, K.P. and Carell, T. (2007) Bypass of DNA lesions generated during anticancer treatment with cisplatin by DNA polymerase eta. *Science*, **318**, 967-970.
- Andersen, P.L., Xu, F. and Xiao, W. (2008) Eukaryotic DNA damage tolerance and translesion synthesis through covalent modifications of PCNA. *Cell Res*, **18**, 162-173.
- Beese, L.S., Derbyshire, V. and Steitz, T.A. (1993) Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science*, **260**, 352-355.
- Berdichevsky, A., Izhar, L. and Livneh, Z. (2002) Error-free recombinational repair predominates over mutagenic translesion replication in *E. coli*. *Mol Cell*, **10**, 917-924.
- Bienko, M., Green, C.M., Crosetto, N., Rudolf, F., Zapart, G., Coull, B., Kannouche, P., Wider, G., Peter, M., Lehmann, A.R., Hofmann, K. and Dikic, I. (2005) Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. *Science*, **310**, 1821-1824.
- Biertumpfel, C., Zhao, Y., Kondo, Y., Ramon-Maiques, S., Gregory, M., Lee, J.Y., Masutani, C., Lehmann, A.R., Hanaoka, F. and Yang, W. (2010) Structure and mechanism of human DNA polymerase eta. *Nature*, **465**, 1044-1048.
- Boudsocq, F., Iwai, S., Hanaoka, F. and Woodgate, R. (2001) *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4): an archaeal DinB-like DNA polymerase with

- lesion-bypass properties akin to eukaryotic poleta. *Nucleic Acids Res*, **29**, 4607-4616.
- Boudsocq, F., Kokoska, R.J., Plosky, B.S., Vaisman, A., Ling, H., Kunkel, T.A., Yang, W. and Woodgate, R. (2004) Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication. *J Biol Chem*, **279**, 32932-32940.
- Braithwaite, E., Wu, X. and Wang, Z. (1998) Repair of DNA lesions induced by polycyclic aromatic hydrocarbons in human cell-free extracts: involvement of two excision repair mechanisms in vitro. *Carcinogenesis*, **19**, 1239-1246.
- Chabes, A., Georgieva, B., Domkin, V., Zhao, X., Rothstein, R. and Thelander, L. (2003) Survival of DNA damage in yeast directly depends on increased dNTP levels allowed by relaxed feedback inhibition of ribonucleotide reductase. *Cell*, **112**, 391-401.
- Cheadle, J.P., Dolwani, S. and Sampson, J.R. (2003) Inherited defects in the DNA glycosylase MYH cause multiple colorectal adenoma and carcinoma. *Carcinogenesis*, **24**, 1281-1282; author reply 1283.
- Collins, A.R. (1999) Oxidative DNA damage, antioxidants, and cancer. *Bioessays*, **21**, 238-246.
- de Boer, J., Andressoo, J.O., de Wit, J., Huijmans, J., Beems, R.B., van Steeg, H., Weeda, G., van der Horst, G.T., van Leeuwen, W., Themmen, A.P., Meradji, M. and Hoeijmakers, J.H. (2002) Premature aging in mice deficient in DNA repair and transcription. *Science*, **296**, 1276-1279.
- De Bont, R. and van Larebeke, N. (2004) Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis*, **19**, 169-185.
- de Laat, W.L., Jaspers, N.G. and Hoeijmakers, J.H. (1999) Molecular mechanism of nucleotide excision repair. *Genes Dev*, **13**, 768-785.
- Doetsch, P.W. and Cunningham, R.P. (1990) The enzymology of apurinic/apyrimidinic endonucleases. *Mutat Res*, **236**, 173-201.
- Doublet, S., Tabor, S., Long, A.M., Richardson, C.C. and Ellenberger, T. (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature*, **391**, 251-258.
- Duncan, B.K. and Miller, J.H. (1980) Mutagenic deamination of cytosine residues in DNA. *Nature*, **287**, 560-561.
- Farrington, S.M., Tenesa, A., Barnetson, R., Wiltshire, A., Prendergast, J., Porteous, M., Campbell, H. and Dunlop, M.G. (2005) Germline susceptibility to colorectal cancer due to base-excision repair gene defects. *Am J Hum Genet*, **77**, 112-119.

- Feldman, G., Remsen, J., Shinohara, K. and Cerutti, P. (1978) Excisability and persistence of benzo(a)pyrene DNA adducts in epithelioid human lung cells. *Nature*, **274**, 796-798.
- Fiala, K.A. and Suo, Z. (2004) Pre-Steady-State Kinetic Studies of the Fidelity of *Sulfolobus solfataricus* P2 DNA Polymerase IV. *Biochemistry*, **43**, 2106-2115.
- Franklin, M.C., Wang, J. and Steitz, T.A. (2001) Structure of the replicating complex of a pol alpha family DNA polymerase. *Cell*, **105**, 657-667.
- Freeman, S.E., Hacham, H., Gange, R.W., Maytum, D.J., Sutherland, J.C. and Sutherland, B.M. (1989) Wavelength dependence of pyrimidine dimer formation in DNA of human skin irradiated in situ with ultraviolet light. *Proc Natl Acad Sci U S A*, **86**, 5605-5609.
- Friedberg, E.C. (2001) How nucleotide excision repair protects against cancer. *Nat Rev Cancer*, **1**, 22-33.
- Friedberg, E.C. (2003) DNA damage and repair. *Nature*, **421**, 436-440.
- Friedberg, E.C., Fischhaber, P.L. and Kisker, C. (2001) Error-prone DNA polymerases: novel structures and the benefits of infidelity. *Cell*, **107**, 9-12.
- Gerlach, V.L., Aravind, L., Gotway, G., Schultz, R.A., Koonin, E.V. and Friedberg, E.C. (1999) Human and mouse homologs of Escherichia coli DinB (DNA polymerase IV), members of the UmuC/DinB superfamily. *Proc Natl Acad Sci U S A*, **96**, 11922-11927.
- Glick, E., Vigna, K.L. and Loeb, L.A. (2001) Mutations in human DNA polymerase eta motif II alter bypass of DNA lesions. *Embo J*, **20**, 7303-7312.
- Haracska, L., Acharya, N., Unk, I., Johnson, R.E., Hurwitz, J., Prakash, L. and Prakash, S. (2005) A single domain in human DNA polymerase iota mediates interaction with PCNA: implications for translesion DNA synthesis. *Mol Cell Biol*, **25**, 1183-1190.
- Haracska, L., Johnson, R.E., Unk, I., Phillips, B.B., Hurwitz, J., Prakash, L. and Prakash, S. (2001) Targeting of human DNA polymerase iota to the replication machinery via interaction with PCNA. *Proc Natl Acad Sci U S A*, **98**, 14256-14261.
- Haracska, L., Kondratyck, C.M., Unk, I., Prakash, S. and Prakash, L. (2001) Interaction with PCNA is essential for yeast DNA polymerase eta function. *Mol Cell*, **8**, 407-415.
- Haracska, L., Prakash, L. and Prakash, S. (2002) Role of human DNA polymerase kappa as an extender in translesion synthesis. *Proc Natl Acad Sci U S A*, **99**, 16000-16005.

- Haracska, L., Prakash, S. and Prakash, L. (2002) Yeast Rev1 protein is a G template-specific DNA polymerase. *J Biol Chem*, **277**, 15546-15551.
- Haracska, L., Torres-Ramos, C.A., Johnson, R.E., Prakash, S. and Prakash, L. (2004) Opposing effects of ubiquitin conjugation and SUMO modification of PCNA on replicational bypass of DNA lesions in *Saccharomyces cerevisiae*. *Mol Cell Biol*, **24**, 4267-4274.
- Haracska, L., Unk, I., Johnson, R.E., Phillips, B.B., Hurwitz, J., Prakash, L. and Prakash, S. (2002) Stimulation of DNA synthesis activity of human DNA polymerase kappa by PCNA. *Mol Cell Biol*, **22**, 784-791.
- Hoege, C., Pfander, B., Moldovan, G.L., Pyrowolakis, G. and Jentsch, S. (2002) RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature*, **419**, 135-141.
- Hoeijmakers, J.H. (2009) DNA damage, aging, and cancer. *N Engl J Med*, **361**, 1475-1485.
- Hsu, G.W., Huang, X., Luneva, N.P., Geacintov, N.E. and Beese, L.S. (2005) Structure of a High Fidelity DNA Polymerase Bound to a Benzo[a]pyrene Adduct That Blocks Replication. *J. Biol. Chem.*, **280**, 3764-3770.
- Hsu, G.W., Ober, M., Carell, T. and Beese, L.S. (2004) Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase. *Nature*, **431**, 217-221.
- Ishikawa, T., Uematsu, N., Mizukoshi, T., Iwai, S., Iwasaki, H., Masutani, C., Hanaoka, F., Ueda, R., Ohmori, H. and Todo, T. (2001) Mutagenic and nonmutagenic bypass of DNA lesions by *Drosophila* DNA polymerases dpoleta and dpoliota. *J Biol Chem*, **276**, 15155-15163.
- Jansen, J.G., Foustieri, M.I. and de Wind, N. (2007) Send in the clamps: control of DNA translesion synthesis in eukaryotes. *Mol Cell*, **28**, 522-529.
- Jarosz, D.F., Godoy, V.G. and Walker, G.C. (2007) Proficient and accurate bypass of persistent DNA lesions by DinB DNA polymerases. *Cell Cycle*, **6**, 817-822.
- Johnson, R.E., Haracska, L., Prakash, L. and Prakash, S. (2006) Role of Hoogsteen edge hydrogen bonding at template purines in nucleotide incorporation by human DNA polymerase iota. *Mol Cell Biol*, **26**, 6435-6441.
- Johnson, R.E., Haracska, L., Prakash, S. and Prakash, L. (2001) Role of DNA polymerase zeta in the bypass of a (6-4) TT photoproduct. *Mol Cell Biol*, **21**, 3558-3563.
- Johnson, R.E., Prakash, S. and Prakash, L. (1999) Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Poleta. *Science*, **283**, 1001-1004.

- Johnson, R.E., Washington, M.T., Haracska, L., Prakash, S. and Prakash, L. (2000) Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. *Nature*, **406**, 1015-1019.
- Kannouche, P., Fernandez de Henestrosa, A.R., Coull, B., Vidal, A.E., Gray, C., Zicha, D., Woodgate, R. and Lehmann, A.R. (2003) Localization of DNA polymerases eta and iota to the replication machinery is tightly co-ordinated in human cells. *Embo J*, **22**, 1223-1233.
- Kato, T. and Shinoura, Y. (1977) Isolation and characterization of mutants of *Escherichia coli* deficient in induction of mutations by ultraviolet light. *Mol Gen Genet*, **156**, 121-131.
- Krokan, H.E., Nilsen, H., Skorpen, F., Otterlei, M. and Slupphaug, G. (2000) Base excision repair of DNA in mammalian cells. *FEBS Lett*, **476**, 73-77.
- Kulaeva, O.I., Koonin, E.V., McDonald, J.P., Randall, S.K., Rabinovich, N., Connaughton, J.F., Levine, A.S. and Woodgate, R. (1996) Identification of a DinB/UmuC homolog in the archeon *Sulfolobus solfataricus*. *Mutat Res*, **357**, 245-253.
- Larimer, F.W., Perry, J.R. and Hardigree, A.A. (1989) The REV1 gene of *Saccharomyces cerevisiae*: isolation, sequence, and functional analysis. *J Bacteriol*, **171**, 230-237.
- Lawrence, C.W. and Christensen, R. (1976) UV mutagenesis in radiation-sensitive strains of yeast. *Genetics*, **82**, 207-232.
- Lee, C.H., Chandani, S. and Loechler, E.L. (2005) Homology modeling of four Y-family, lesion-bypass DNA polymerases: The case that *E. coli* Pol IV and human Pol κ are orthologs, and *E. coli* Pol V and human Pol η are orthologs. *J. Mol. Graph. Model.*, **25**, 87-102.
- Lemontt, J.F. (1971) Mutants of yeast defective in mutation induced by ultraviolet light. *Genetics*, **68**, 21-33.
- Li, Y., Korolev, S. and Waksman, G. (1998) Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: structural basis for nucleotide incorporation. *Embo J*, **17**, 7514-7525.
- Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709-715.
- Ling, H., Boudsocq, F., Woodgate, R. and Yang, W. (2001) Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. *Cell*, **107**, 91-102.

- Lone, S., Townson, S.A., Uljon, S.N., Johnson, R.E., Brahma, A., Nair, D.T., Prakash, S., Prakash, L. and Aggarwal, A.K. (2007) Human DNA polymerase kappa encircles DNA: implications for mismatch extension and lesion bypass. *Mol Cell*, **25**, 601-614.
- Lu, T., Pan, Y., Kao, S.Y., Li, C., Kohane, I., Chan, J. and Yankner, B.A. (2004) Gene regulation and DNA damage in the ageing human brain. *Nature*, **429**, 883-891.
- Masutani, C., Araki, M., Yamada, A., Kusumoto, R., Nogimori, T., Maekawa, T., Iwai, S. and Hanaoka, F. (1999) Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. *Embo J*, **18**, 3491-3501.
- Masutani, C., Kusumoto, R., Iwai, S. and Hanaoka, F. (2000) Mechanisms of accurate translesion synthesis by human DNA polymerase eta. *Embo J*, **19**, 3100-3109.
- Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M., Araki, M., Iwai, S., Takio, K. and Hanaoka, F. (1999) The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta. *Nature*, **399**, 700-704.
- Matsuda, T., Bebenek, K., Masutani, C., Hanaoka, F. and Kunkel, T.A. (2000) Low fidelity DNA synthesis by human DNA polymerase-eta. *Nature*, **404**, 1011-1013.
- McCulloch, S.D. and Kunkel, T.A. (2008) The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res*, **18**, 148-161.
- McDonald, J.P., Levine, A.S. and Woodgate, R. (1997) The *Saccharomyces cerevisiae* RAD30 gene, a homologue of *Escherichia coli* dinB and umuC, is DNA damage inducible and functions in a novel error-free postreplication repair mechanism. *Genetics*, **147**, 1557-1568.
- McDonald, J.P., Ropic-Otrin, V., Epstein, J.A., Broughton, B.C., Wang, X., Lehmann, A.R., Wolgemuth, D.J. and Woodgate, R. (1999) Novel human and mouse homologs of *Saccharomyces cerevisiae* DNA polymerase eta. *Genomics*, **60**, 20-30.
- McDonald, J.P., Tissier, A., Frank, E.G., Iwai, S., Hanaoka, F. and Woodgate, R. (2001) DNA polymerase iota and related rad30-like enzymes. *Philos Trans R Soc Lond B Biol Sci*, **356**, 53-60.
- Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2005) Rev1 employs a novel mechanism of DNA synthesis using a protein template. *Science*, **309**, 2219-2222.
- Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2006) Hoogsteen base pair formation promotes synthesis opposite the 1,N6-ethenodeoxyadenosine lesion by human DNA polymerase iota. *Nat Struct Mol Biol*, **13**, 619-625.

- Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2006) An incoming nucleotide imposes an anti to syn conformational change on the templating purine in the human DNA polymerase- ι active site. *Structure*, **14**, 749-755.
- Nair, D.T., Johnson, R.E., Prakash, S., Prakash, L. and Aggarwal, A.K. (2004) Replication by human DNA polymerase- ι occurs by Hoogsteen base-pairing. *Nature*, **430**, 377-380.
- Nelson, J.R., Lawrence, C.W. and Hinkle, D.C. (1996) Deoxycytidyl transferase activity of yeast REV1 protein. *Nature*, **382**, 729-731.
- Ogi, T., Shinkai, Y., Tanaka, K. and Ohmori, H. (2002) Polk protects mammalian cells against the lethal and mutagenic effects of benzo[*a*]pyrene. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 15548-15553.
- Ohashi, E., Bebenek, K., Matsuda, T., Feaver, W.J., Gerlach, V.L., Friedberg, E.C., Ohmori, H. and Kunkel, T.A. (2000) Fidelity and processivity of DNA synthesis by DNA polymerase kappa, the product of the human DINB1 gene. *J Biol Chem*, **275**, 39678-39684.
- Ohashi, E., Murakumo, Y., Kanjo, N., Akagi, J., Masutani, C., Hanaoka, F. and Ohmori, H. (2004) Interaction of hREV1 with three human Y-family DNA polymerases. *Genes Cells*, **9**, 523-531.
- Ohmori, H., Friedberg, E.C., Fuchs, R.P., Goodman, M.F., Hanaoka, F., Hinkle, D., Kunkel, T.A., Lawrence, C.W., Livneh, Z., Nohmi, T., Prakash, L., Prakash, S., Todo, T., Walker, G.C., Wang, Z. and Woodgate, R. (2001) The Y-family of DNA polymerases. *Mol Cell*, **8**, 7-8.
- Ohmori, H., Hatada, E., Qiao, Y., Tsuji, M. and Fukuda, R. (1995) dinP, a new gene in *Escherichia coli*, whose product shows similarities to UmuC and its homologues. *Mutat Res*, **347**, 1-7.
- Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G. and Steitz, T.A. (1985) Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature*, **313**, 762-766.
- Petta, T.B., Nakajima, S., Zlatanou, A., Despras, E., Couve-Privat, S., Ishchenko, A., Sarasin, A., Yasui, A. and Kannouche, P. (2008) Human DNA polymerase ι protects cells against oxidative stress. *Embo J*, **27**, 2883-2895.
- Plosky, B.S., Vidal, A.E., Fernandez de Henestrosa, A.R., McLenigan, M.P., McDonald, J.P., Mead, S. and Woodgate, R. (2006) Controlling the subcellular localization of DNA polymerases ι and η via interactions with ubiquitin. *Embo J*, **25**, 2847-2855.

- Plosky, B.S. and Woodgate, R. (2004) Switching from high-fidelity replicases to low-fidelity lesion-bypass polymerases. *Curr Opin Genet Dev*, **14**, 113-119.
- Prakash, S. and Prakash, L. (2002) Translesion DNA synthesis in eukaryotes: a one- or two-polymerase affair. *Genes Dev*, **16**, 1872-1883.
- Reuven, N.B., Arad, G., Maor-Shoshani, A. and Livneh, Z. (1999) The mutagenesis protein UmuC is a DNA polymerase activated by UmuD', RecA, and SSB and is specialized for translesion replication. *J Biol Chem*, **274**, 31763-31766.
- Rothwell, P.J., Mitaksov, V. and Waksman, G. (2005) Motions of the fingers subdomain of klenTaq1 are fast and not rate limiting: implications for the molecular basis of fidelity in DNA polymerases. *Mol Cell*, **19**, 345-355.
- Santoso, Y., Joyce, C.M., Potapova, O., Le Reste, L., Hohlbein, J., Torella, J.P., Grindley, N.D. and Kapanidis, A.N. (2010) Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *Proc Natl Acad Sci U S A*, **107**, 715-720.
- Setlow, R.B. (1966) Cyclobutane-type pyrimidine dimers in polynucleotides. *Science*, **153**, 379-386.
- Shachar, S., Ziv, O., Avkin, S., Adar, S., Wittschieben, J., Reissner, T., Chaney, S., Friedberg, E.C., Wang, Z., Carell, T., Geacintov, N. and Livneh, Z. (2009) Two-polymerase mechanisms dictate error-free and error-prone translesion DNA synthesis in mammals. *Embo J*, **28**, 383-393.
- Sijbers, A.M., de Laat, W.L., Ariza, R.R., Biggerstaff, M., Wei, Y.F., Moggs, J.G., Carter, K.C., Shell, B.K., Evans, E., de Jong, M.C., Rademakers, S., de Rooij, J., Jaspers, N.G., Hoeijmakers, J.H. and Wood, R.D. (1996) Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell*, **86**, 811-822.
- Silverstein, T.D., Johnson, R.E., Jain, R., Prakash, L., Prakash, S. and Aggarwal, A.K. (2010) Structural basis for the suppression of skin cancers by DNA polymerase eta. *Nature*, **465**, 1039-1043.
- Silvian, L.F., Toth, E.A., Pham, P., Goodman, M.F. and Ellenberger, T. (2001) Crystal structure of a DinB family error-prone DNA polymerase from *Sulfolobus solfataricus*. *Nat Struct Biol*, **8**, 984-989.
- Steinborn, G. (1978) Uvm mutants of *Escherichia coli* K12 deficient in UV mutagenesis. I. Isolation of uvm mutants and their phenotypical characterization in DNA repair and mutagenesis. *Mol Gen Genet*, **165**, 87-93.
- Swan, M.K., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2009) Structural basis of high-fidelity DNA synthesis by yeast DNA polymerase delta. *Nat Struct Mol Biol*, **16**, 979-986.

- Swan, M.K., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2009) Structure of the human Rev1-DNA-dNTP ternary complex. *J Mol Biol*, **390**, 699-709.
- Tang, M., Shen, X., Frank, E.G., O'Donnell, M., Woodgate, R. and Goodman, M.F. (1999) UmuD'(2)C is an error-prone DNA polymerase, Escherichia coli pol V. *Proc Natl Acad Sci U S A*, **96**, 8919-8924.
- Thomas, D.C., Roberts, J.D., Sabatino, R.D., Myers, T.W., Tan, C.K., Downey, K.M., So, A.G., Bambara, R.A. and Kunkel, T.A. (1991) Fidelity of mammalian DNA replication and replicative DNA polymerases. *Biochemistry*, **30**, 11751-11759.
- Tissier, A., Kannouche, P., Reck, M.P., Lehmann, A.R., Fuchs, R.P. and Cordonnier, A. (2004) Co-localization in replication foci and interaction of human Y-family members, DNA polymerase pol eta and REVI protein. *DNA Repair (Amst)*, **3**, 1503-1514.
- Tissier, A., McDonald, J.P., Frank, E.G. and Woodgate, R. (2000) poliota, a remarkably error-prone human DNA polymerase. *Genes Dev*, **14**, 1642-1650.
- Uljon, S.N., Johnson, R.E., Edwards, T.A., Prakash, S., Prakash, L. and Aggarwal, A.K. (2004) Crystal structure of the catalytic core of human DNA polymerase kappa. *Structure*, **12**, 1395-1404.
- Vaisman, A., Ling, H., Woodgate, R. and Yang, W. (2005) Fidelity of Dpo4: effect of metal ions, nucleotide selection and pyrophosphorolysis. *EMBO J.*, **24**, 2957-2967.
- Wagner, J., Gruz, P., Kim, S.R., Yamada, M., Matsui, K., Fuchs, R.P. and Nohmi, T. (1999) The dinB gene encodes a novel E. coli DNA polymerase, DNA pol IV, involved in mutagenesis. *Mol Cell*, **4**, 281-286.
- Wang, Y., Woodgate, R., McManus, T.P., Mead, S., McCormick, J.J. and Maher, V.M. (2007) Evidence that in xeroderma pigmentosum variant cells, which lack DNA polymerase eta, DNA polymerase iota causes the very high frequency and unique spectrum of UV-induced mutations. *Cancer Res*, **67**, 3018-3026.
- Wardle, J., Burgers, P.M., Cann, I.K., Darley, K., Heslop, P., Johansson, E., Lin, L.J., McGlynn, P., Sanvoisin, J., Stith, C.M. and Connolly, B.A. (2008) Uracil recognition by replicative DNA polymerases is limited to the archaea, not occurring with bacteria and eukarya. *Nucleic Acids Res*, **36**, 705-711.
- Washington, M.T., Johnson, R.E., Prakash, L. and Prakash, S. (2002) Human DINB1-encoded DNA polymerase kappa is a promiscuous extender of mispaired primer termini. *Proc Natl Acad Sci U S A*, **99**, 1910-1914.
- Washington, M.T., Minko, I.G., Johnson, R.E., Haracska, L., Harris, T.M., Lloyd, R.S., Prakash, S. and Prakash, L. (2004) Efficient and error-free replication past a

- minor-groove N2-guanine adduct by the sequential action of yeast Rev1 and DNA polymerase zeta. *Mol Cell Biol*, **24**, 6900-6906.
- Wong, J.H., Fiala, K.A., Suo, Z. and Ling, H. (2008) Snapshots of a Y-family DNA polymerase in replication: substrate-induced conformational transitions and implications for fidelity of Dpo4. *J Mol Biol*, **379**, 317-330.
- Woodgate, R. (1999) A plethora of lesion-replicating DNA polymerases. *Genes Dev*, **13**, 2191-2195.
- Xing, G., Kirouac, K., Shin, Y.J., Bell, S.D. and Ling, H. (2009) Structural insight into recruitment of translesion DNA polymerase Dpo4 to sliding clamp PCNA. *Mol Microbiol*, **71**, 678-691.
- Yang, W. (2005) Portraits of a Y-family DNA polymerase. *FEBS Lett*, **579**, 868-872.
- Yang, W. (2008) An equivalent metal ion in one- and two-metal-ion catalysis. *Nat Struct Mol Biol*, **15**, 1228-1231.
- Yang, W., Lee, J.Y. and Nowotny, M. (2006) Making and breaking nucleic acids: two-Mg²⁺-ion catalysis and substrate specificity. *Mol Cell*, **22**, 5-13.
- Yang, W. and Woodgate, R. (2007) What a difference a decade makes: insights into translesion DNA synthesis. *Proc Natl Acad Sci U S A*, **104**, 15591-15598.
- Zeng, X., Winter, D.B., Kasmer, C., Kraemer, K.H., Lehmann, A.R. and Gearhart, P.J. (2001) DNA polymerase eta is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. *Nat Immunol*, **2**, 537-541.
- Zhang, Y., Wu, X., Rechkoblit, O., Geacintov, N.E., Taylor, J.S. and Wang, Z. (2002) Response of human REV1 to different DNA damage: preferential dCMP insertion opposite the lesion. *Nucleic Acids Res*, **30**, 1630-1638.
- Zhang, Y., Yuan, F., Wu, X., Rechkoblit, O., Taylor, J.S., Geacintov, N.E. and Wang, Z. (2000) Error-prone lesion bypass by human DNA polymerase eta. *Nucleic Acids Res*, **28**, 4717-4724.
- Zhang, Y., Yuan, F., Wu, X. and Wang, Z. (2000) Preferential incorporation of G opposite template T by the low-fidelity human DNA polymerase iota. *Mol Cell Biol*, **20**, 7099-7108.
- Zhang, Y., Yuan, F., Xin, H., Wu, X., Rajpal, D.K., Yang, D. and Wang, Z. (2000) Human DNA polymerase kappa synthesizes DNA with extraordinarily low fidelity. *Nucleic Acids Res*, **28**, 4147-4156.
- Zhuang, Z., Johnson, R.E., Haracska, L., Prakash, L., Prakash, S. and Benkovic, S.J. (2008) Regulation of polymerase exchange between Poleta and Poldelta by

monoubiquitination of PCNA and the movement of DNA polymerase holoenzyme. *Proc Natl Acad Sci U S A*, **105**, 5361-5366.

Chapter 2¹

2 Structural basis of ribonucleotide discrimination by a Y family DNA polymerase

2.1 Introduction

Loss of ribonucleotide discrimination by DNA polymerases leads to replicative stress and genome instability (Nick McElhinny et al.). The ability to differentiate between deoxyribonucleotides (dNTPs) and ribonucleotides (NTPs) is an essential function of all DNA polymerases that must be highly specific, in part, due to the excess of NTP concentrations in the cell (Nick McElhinny et al.; Traut, 1994). Most DNA polymerases utilize bulky side chain residues in their active sites to discriminate against NTPs (Joyce, 1997). The identity of this residue is typically glutamine for A family polymerases (Astatke et al., 1998; Patel and Loeb, 2000), and tyrosine or phenylalanine for Y and B family polymerases (Brown et al., 2009; DeLucia et al., 2003; Niimi et al., 2009; Yang et al., 2002). X-ray crystallographic studies have revealed that these residues create a stacking interaction with the deoxyribose sugar of incoming dNTP's, and form a hydrogen bond between their backbone amine groups and the 3'-OH group of the deoxyribose sugar moiety (Biertumpfel et al., 2010; Doublet et al., 1998; Ling et al., 2001; Lone et al., 2007; Nair et al., 2004). Mutating these residues to smaller amino acids generally results in a loss of NTP discrimination (Astatke et al., 1998; Bonnin et al., 1999; DeLucia et al., 2003; Gao et al., 1997; Yang et al., 2002). It has been hypothesized that these bulky active site residues prevent NTP incorporation by clashing with the 2'-

¹ The contents of this chapter have been submitted for publication.

OH group of ribose sugars and have thus been coined ‘steric gates’ (DeLucia et al., 2003; Jarosz et al., 2006). Although the ‘steric gate’ residue is clearly involved in NTP discrimination, no structural studies have been performed on steric gate mutant DNA polymerases. Thus, exactly how dNTP/NTP selectivity is achieved still remains elusive.

Y family polymerases have evolved open, solvent accessible active sites, which accommodate bulky and distorting DNA lesions. Consequently, these polymerases can accommodate an incoming nucleotide in different conformations, which allows permissive base pairing to facilitate DNA lesion bypass (Bauer et al., 2007; Kirouac and Ling, 2009; Ling et al., 2003). Remarkably, the solvent accessible active sites of Y family polymerases, which have minimal contacts to incoming nucleotides, are still highly discriminatory against NTPs (DeLucia et al., 2003; Jarosz et al., 2006; Niimi et al., 2009). Although the steric gate residue is likely responsible for this NTP discrimination, structural investigations are required to fully understand how a single residue can induce such high selectivity in this permissive environment.

In order to elucidate the structural role of the steric gate residue in the active site of a Y family DNA polymerase, we generated a single point mutation (Y12A) in Dpo4, a model Y family polymerase from the archaeon *Sulfolobus solfataricus*. Here we report that mutating this conserved residue in Dpo4, creates a defect in ribonucleotide discrimination and produces a pseudo DNA/RNA polymerase. In addition, the Y12A mutant has low nucleotide incorporation efficiency for both dNTPs and NTPs. Crystal structures were obtained for the Dpo4 Y12A ternary complex, incorporating either dATP or ATP opposite template dT, demonstrating the first structure of a DNA polymerase incorporating a ribonucleotide. These results reveal for the first time, how a Y family

polymerase structurally discriminates against ribonucleotides and furthers our understanding of the general enzymatic mechanism of Y family DNA polymerases.

2.2 Results

2.2.1 Impaired ribonucleotide discrimination by Dpo4 Y12A

To test the ability of the Dpo4 Y12A mutant to discriminate against ribonucleotides, primer extension assays were performed with the wild type and mutant Dpo4 proteins. On undamaged DNA, Dpo4 wild type incorporates dNTPs with high efficiency, extending the primer strand to the end of the DNA template (Figure 2.1A). In contrast to dNTP incorporation, Dpo4 wild type incorporates virtually no NTPs and is thus highly discriminatory against incoming ribonucleotides (Figure 2.1A). In contrast to Dpo4 wild type, the Y12A mutant is able to incorporate NTPs and has therefore lost the ability to discriminate against ribonucleotides (Figure 2.1B). The Dpo4 Y12A mutant also incorporates dNTPs, however, the efficiency of incorporation is reduced compared to wild type (Figure 2.1B). Although Dpo4 Y12A is able to incorporate NTPs, the efficiency of incorporation is further reduced compared to dNTPs. The fidelity of Dpo4 Y12A was also tested and compared to wild type. The Y12A mutant prefers to incorporate the correct dATP or ATP opposite template dT with minimal mis-incorporations of the other nucleotides, similar to wild type Dpo4 (Figure 2.1C, D).

2.2.2 Crystal structures of Dpo4 Y12A mutant

In order to elucidate the acquired function of ribonucleotide incorporation, we crystallized the Dpo4 Y12A mutant in ternary complex with DNA and either incoming dATP or ATP nucleotides. The DNA substrate used for crystallization consisted of a

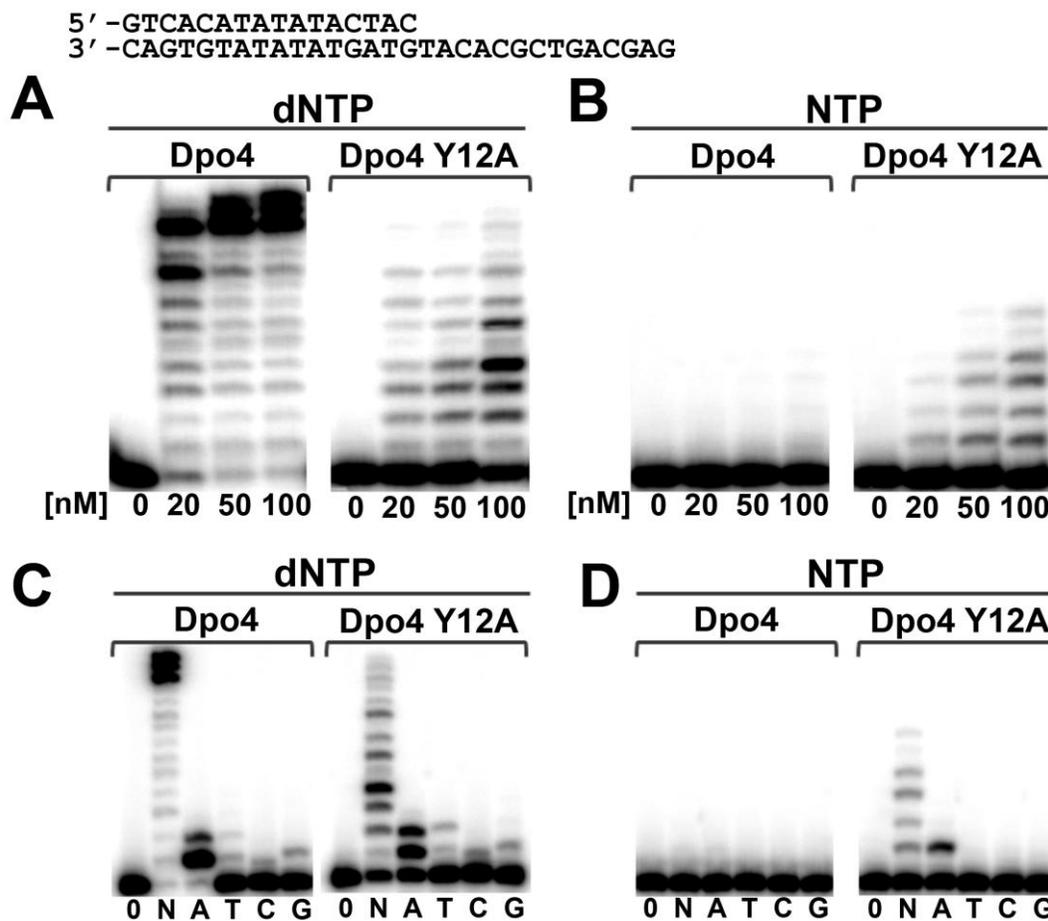


Figure 2.1 Incorporation of dNTP and NTP by Dpo4 and Dpo4 Y12A mutant. **(A)** Incorporation of dNTP by Dpo4 (Left) and Dpo4 Y12A (Right) indicating a decrease in efficiency for Dpo4 Y12A. **(B)** Incorporation of NTP by Dpo4 (Left) and Dpo4 Y12A (Right) indicating a loss of ribonucleotide discrimination by Dpo4 Y12A. Enzymes were incubated with DNA and all four nucleotides concentrations ranging from 0-100 nM. **(C)** Fidelity of dNTP incorporation by Dpo4 (Left) and Dpo4 Y12A (Right). **(D)** Fidelity of NTP incorporation by Dpo4 (Left) and Dpo4 Y12A (Right). Enzymes were incubated with DNA and either no nucleotides, all four nucleotide (N), or individual nucleotides (A, T/U, C, G).

13/18-nt duplex containing a dideoxy 3' primer end and a template thymine base 5' to the template-primer junction. Incoming dATP or ATP were incubated with the DNA substrate and co-crystallized with the Dpo4 Y12A mutant. Both nucleotides were hydrolyzed to dADP and ADP similar to dideoxy nucleotides in the wild type Dpo4 structures (Ling et al., 2001). The resulting two structures are referred to as Y12A-dADP and Y12A-ADP according to the identity of the incoming nucleotide in the active site. The Y12A-dADP and Y12A-ADP structures have space group $P2_12_12$ and diffracted X-rays to 2.65 and 2.4 Å resolutions respectively (Table 2.1). Both structures have been refined with Ca^{2+} as the active site metal ions, similar to previous Dpo4 structures, due to the presence of 100mM $Ca(Ac)_2$ in the crystallization buffer (Wong et al., 2008). There are two active site metal ions in the Y12A-dADP structure, which are found in the same A and B positions as wild type Dpo4 structures. There is only one active site metal ion present in the Y12A-ADP structure, which is located in the B site similar to previous Dpo4 structures (Kirouac and Ling, 2009; Ling et al., 2001; Vaisman et al., 2005). The absence of a metal ion in site A is likely due to the high mobility of ions at this position, which has been previously observed (Kirouac and Ling, 2009; Nair et al., 2005; Vaisman et al., 2005).

The Dpo4 Y12A structures are virtually identical to previously solved wild type Dpo4 structures (Figure 2.3A) (Ling et al., 2001; Vaisman et al., 2005). Comparing all Ca atoms between both Y12A structures and wild type Dpo4 (PDB 2AGQ) gave root-mean-square deviations (r.m.s.d.) between 0.30 and 0.36 Å. In addition, the active site residues contacting the replicating base pair are positioned identically between both Y12A structures and wild type Dpo4. Therefore, the Y12A mutation does not induce any

Table 2.1 Summary of crystallographic data for Dpo4 Y12A structures

Crystal	Y12A- ADP	Y12A- dADP
Space group	P21212	P21212
Complexes per AU ^a	1	1
Unit Cell		
<i>a, b, c</i> (Å)	95.6, 102.0, 52.3	96.4, 102.0, 52.2
Resolution range (Å) ^b	30.0-2.40 (2.46-2.40)	30.0-2.65 (2.79-2.65)
<i>R</i> _{merge} ^b	4.7 (53.1)	7.7 (51.3)
<i>I</i> / σ <i>I</i>	20.5 (2.0)	13.1 (2.1)
Completeness (%) ^b	95.0 (81.9)	99.5(95.0)
Redundancy ^b	3.6 (2.8)	3.1 (2.8)
No. Refelctions	19896	15456
<i>R</i> _{work} / <i>R</i> _{free}	22.8 / 26.5	19.7 / 25.5
No. Atoms		
Protein	2737	2655
DNA	570	570
dNTP	27	26
Ions	2	3
Waters	79	84
B factors		
Protein	61.2	53.7
DNA	68.6	61.7
dNTP	79.5	45.8
Ions	66.1	62.9
Waters	72.3	49.0
R.m.s. deviations		
Bonds (Å)	0.009	0.009
Angles (°)	1.32	1.29

^a AU means asymmetric unit

^b Data in the highest resolution shell are in parentheses

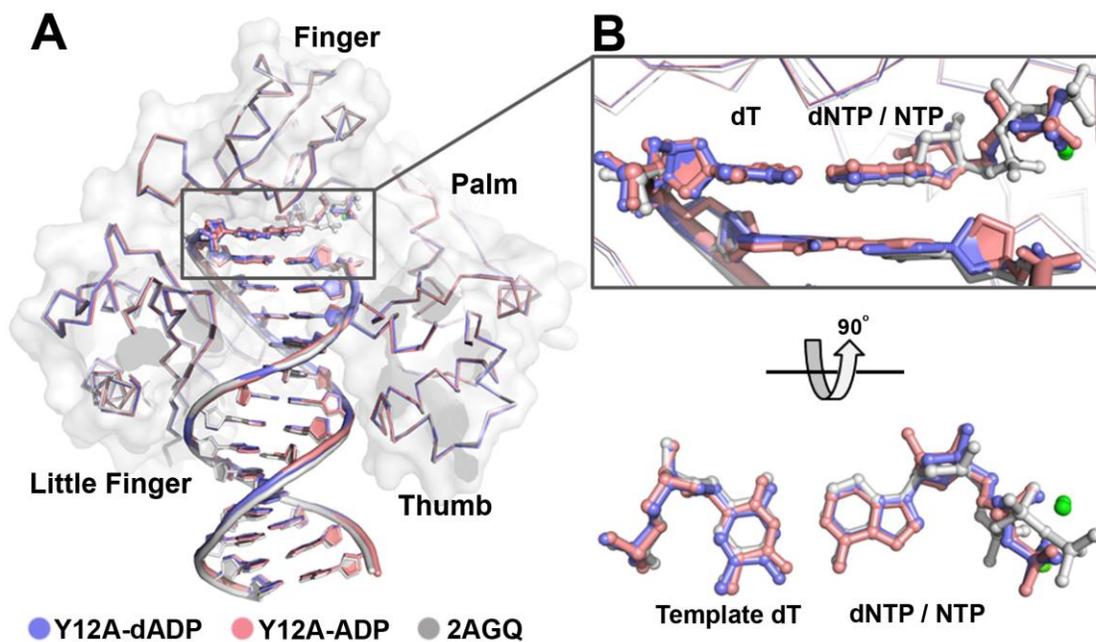


Figure 2.2 Superposition of Dpo4 wild type and Dpo4 Y12A ternary complexes. (A) Superposition of wild type Dpo4 (2AGQ, grey), Y12A-dADP (Blue), and Y12A-ADP (Red). (B) Close up view of the incoming nucleotides and template base from the side and top view. The backbone atoms are shown in ribbon diagram, the incoming nucleotides are shown in ball and stick, and surface representation of 2AGQ is shown in grey. Active site metal ions are shown as green spheres.

disturbance in the protein structure except the side chain difference in the active site. The major difference is observed in the positioning of the incoming nucleotides (Figure 2.3B). The bases of incoming dADP and ADP are positioned nearly identical to that of dATP in the Dpo4 wild type structure (Vaisman et al., 2005) (Figure 2.3B). However, the sugar groups of both dADP and ADP are slightly shifted towards the minor groove of the DNA. In addition, the α and β phosphates are positioned similar to ddADP in wild type Dpo4 (PDB: 1JX4) and differently from dATP in wild type Dpo4 (PDB: 2AGQ) due to the hydrolysis of the γ phosphate (Ling et al., 2001; Vaisman et al., 2005).

2.2.3 Positions of incoming dADP and ADP in the Y12A mutant

The Dpo4 Y12A mutant, which has lost the ability to discriminate against ribonucleotides, induces subtle but distinct conformational changes on both incoming dADP and ADP nucleotides. Both Y12A-dADP and Y12A-ADP structures have the template thymine base in a standard anti conformation, positioned identically to wild type Dpo4 (Figure 2.2). The incoming dADP of Y12A forms a Watson-Crick base pair with the template T base similar to wild type Dpo4 with incoming dATP (Figure 2.3A) (Vaisman et al., 2005). In addition, the 3'-OH of the deoxy ribose sugar makes a hydrogen bond with the amide group of the A12 residue, similar to wild type Dpo4 (Figure 2.3A) (Vaisman et al., 2005). However, since the tyrosine 12 residue has been mutated to alanine, the stacking interaction with the deoxy ribose is lost causing the sugar group to move towards the minor groove (Figure 2.4B). The 3'-OH of deoxyribose shifts ~ 1 Å compared to dATP in the wild type Dpo4 structure (Figure 2.4B). The phosphate groups are coordinated by two metal ions similar to wild type Dpo4.

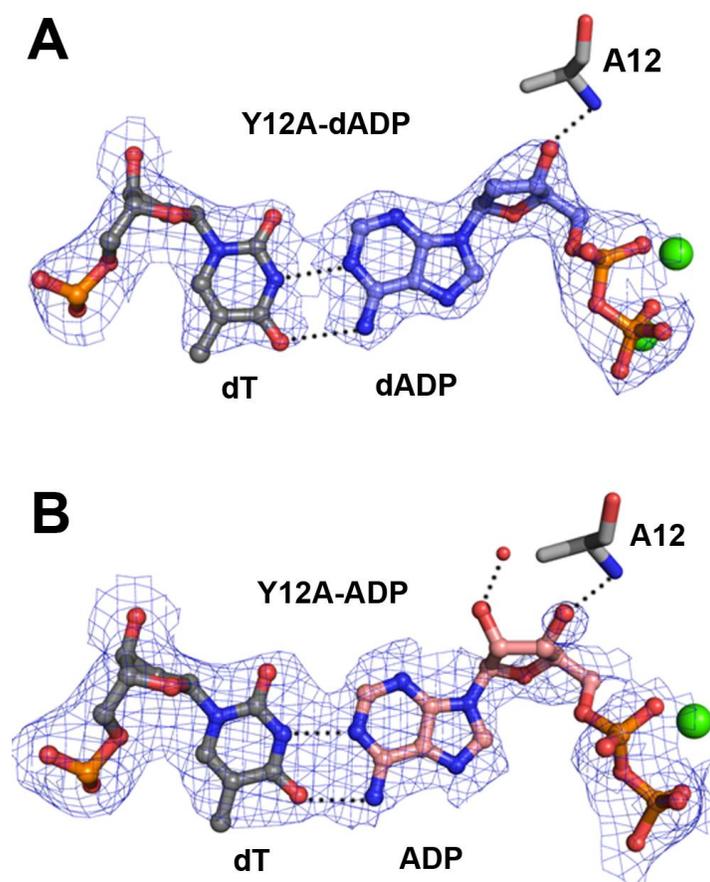


Figure 2.3 Conformation of replicating base pairs in the Y12A ternary structures. **(A)** The Y12A-dADP structure showing the template dT in grey and the incoming dADP in blue. **(B)** The 'active' Y12A-ADP structure showing the template dT in grey and the incoming ATP in red. Hydrogen bonding is shown as dashed black lines, metal ions are shown as green spheres, water molecules are shown as red spheres, and the A12 residue is shown in stick model coloured grey. The 2Fo-Fc electron density map is shown in blue, contoured at 1 σ level.

The incoming ADP is similar to dADP in Y12A-dADP (Figure 2.2). Watson-Crick base pairing occurs between the template T base and incoming dADP (Figure 2.3B). The 3' -OH of the ribose sugar makes a weak hydrogen bond with the amine group of the A12 residue, while the 2' -OH is projected into the open pocket created by the Y12A mutant with a hydrogen bond to a water molecule (Figure 2.3B). The electron density is poor for the ribose sugar of the ADP, indicating a high degree of mobility and destabilization of the nucleotide.

2.3 Discussion

2.3.1 The steric gate and ribonucleotide discrimination

Our structural observations have finally verified the long standing hypothesis that the steric gate residue blocks NTP incorporation by clashing with the 2' -OH group (Astatke et al., 1998; Gao et al., 1997; Joyce, 1997). When the Dpo4 steric gate residue is mutated to an Ala, the 2' -OH group of the ribose sugar is able to be accommodated in the polymerase active site, which allows a productive nucleotide conformation to exist. In the presence of the wild type tyrosine residue, the 2' -OH group would clash with the steric gate (within 1.5 Å), thus preventing NTP incorporation (Figure 2.4A). These results are consistent with pre-steady state kinetic analysis demonstrating the wild type Dpo4 incorporates a dNTP with 9,200 fold greater efficiency than a corresponding NTP, while the efficiency difference is only 5 fold with the Dpo4 Y12A mutant (Sherrer et al., 2010). A similar mechanism likely occurs in the closely related Archaeal Y family polymerase Dbh from *Sulfolobus acidocaldarius*, which also displays a loss of ribonucleotide discrimination upon mutating the conserved steric gate residue (DeLucia et al., 2003).

2.3.2 The steric gate and nucleotide incorporation

Interestingly, the steric gate residue of Dpo4 appears to influence nucleotide incorporation efficiency. The Dpo4 Y12A mutation causes significant reduction in activity for both dNTP and NTP incorporation. Pre-steady state kinetics reveal a 3-19 fold decrease in dNTP incorporation efficiency between Dpo4 wild type and the Y12A mutant (Sherrer et al., 2010). Similarly, NTP incorporation efficiency is reduced by 3-30 fold compared to dNTP's for the Y12A mutant (Sherrer et al., 2010). This decreased activity can be rationalized by comparing the structures of wild type Dpo4 with Y12A-dADP and Y12A-ADP. The Y12A mutation creates a more open and permissive active site, which likely destabilizes dNTP and NTP nucleotides by abolishing the sugar group stacking interaction and allowing greater conformational heterogeneity. In addition, the altered puckering of NTP ribose sugars likely induces higher mobility and greater destabilization compared to dNTP nucleotides. This greater conformational heterogeneity is demonstrated by the poor density of the ADP sugar moiety in Y12A-ADP and results in further reduced incorporation efficiency. Thus, it appears the sugar group must be stabilized and positioned adjacent to the steric gate residue for efficient catalysis to occur.

Interestingly, the dADP and ADP nucleotides in Y12A-dADP and Y12A-ADP are identical to ddADP in a previous wild type Dpo4 structure (PDB: 1JX4) (Figure 2.4C, D) (Ling et al., 2001). The ddADP nucleotide also shifts its sugar group towards the minor groove of the DNA due to the absence of the 3' -OH group, which normally hydrogen bonds to the Tyr12 backbone amine (Figure 2.4C, D). Thus similar to dADP and ADP nucleotides within the Dpo4 Y12A structures, ddADP nucleotides in wild type Dpo4 are

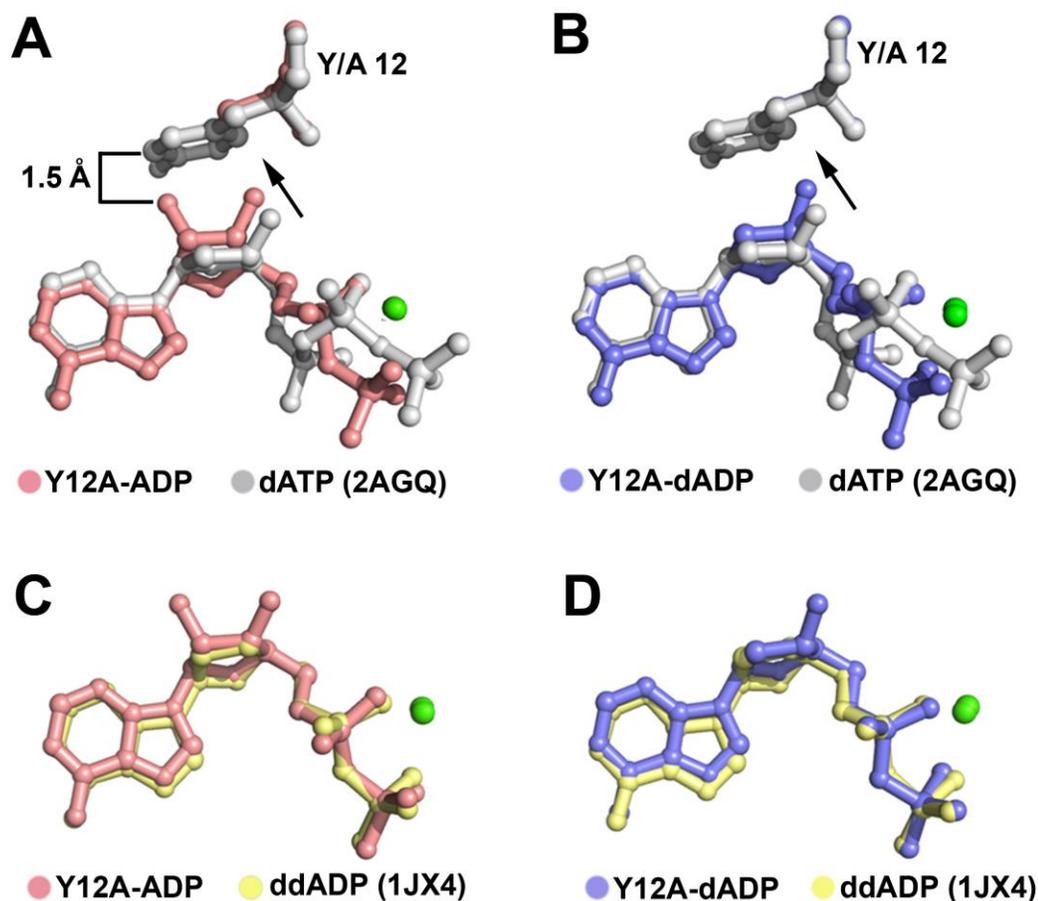


Figure 2.4 Comparison of incoming nucleotides between Dpo4 and Dpo4 Y12A. (A) Superposition of dATP, ADP and amino acid 12 between wild type Dpo4 (2AGQ, grey) and Y12A-ADP (red). (B) Superposition of dATP, dADP and amino acid 12 between wild type Dpo4 (2AGQ, grey) and Y12A-dADP (blue). Black arrows indicate movement of the sugar groups and clashing distance is highlighted with a bracket between the 2' -OH of ADP and the Y12 residue of Dpo4 wt (2AGQ). (C) Superposition of ddADP and ADP between wild type Dpo4 (1JX4, yellow) and Y12A-ADP (red). (D) Superposition of ddADP, dADP between wild type Dpo4 (1JX4, yellow) and Y12A-dADP (blue). Metal ions are shown as green spheres.

destabilized. Consequently, the incorporation efficiency of ddATP nucleotides is significantly reduced compared to dNTPs for both Y family polymerases and high fidelity polymerases (Ling et al., 2003; Zhou et al., 2005). Thus, the stability of the nucleotide sugar group appears critical for nucleotide incorporation efficiency.

Interestingly, the nucleotide incorporation fidelity is maintained for the Dpo4 Y12A mutant. In addition, nucleotide incorporation specificity is maintained between dNTPs and NTPs for Dpo4 Y12A. Previous work has shown the NTP incorporation fidelity of Dpo4 Y12A to be in the range of 10^{-3} - 10^{-4} , which is identical to dNTP incorporation fidelity of wild type Dpo4 (Sherrer et al., 2010). Thus, the steric gate residue of Dpo4 has little influence on incoming nucleotide selection in spite of its critical function in positioning and stabilizing incoming nucleotides. This result is consistent with the structural observations that incoming nucleotide selection for Y family DNA polymerases is determined by hydrogen bonding and base stacking potentials of the incoming nucleotide base and not the sugar group (Kirouac and Ling, 2009). In addition, these results indicate that dNTP and NTP incorporation by Dpo4 Y12A are catalyzed by a similar reaction mechanism.

2.3.3 Sugar positioning and gamma phosphate hydrolysis

The gamma phosphates of both dADP and ADP nucleotides in the Y12A structures have been hydrolyzed. Interestingly, this phenomenon has only been observed for incoming ddNTP nucleotides in Y family polymerase structures (Kirouac and Ling, 2009; Ling et al., 2001). Thus, the Dpo4 Y12A mutant has generated the first observed dNTP gamma phosphate hydrolysis in a DNA polymerase structure. It appears that sugar positioning of the incoming nucleotide is involved in phosphate hydrolysis as both ddADP in Dpo4 and

dADP in Dpo4 Y12A have an identical sugar position, which is different from dATP in Dpo4. As the sugar group shifts to the minor groove, the alpha phosphate moves away from the 3' end of primer strand, and the phosphate groups of the incoming nucleotides are more likely exposed to the catalytic site of Dpo4 for hydrolysis. Therefore, the sugar position likely contributes to gamma phosphate hydrolysis of ddNTP in wild type Dpo4 and dNTP/NTP in the Dpo4 Y12A mutant by re-positioning the phosphate groups in the active site. The gamma phosphate hydrolysis likely contributes to decreased nucleotide incorporation efficiency for ddNTP in wild type Dpo4 and dNTP in Dpo4 Y12A. Mutation of the catalytic residues (D105A/ E106A) abolishes all enzymatic activity and indicates that Dpo4 uses the same active site for both nucleotide incorporation and γ phosphate hydrolysis (Ling et al., 2001). The phosphate hydrolysis side reaction would compete at the active site with the phosphodiester bond reaction, thus reducing the efficiency of the nucleotide incorporation.

2.4 Conclusions

Here we have described the first structure of a Y family DNA polymerase steric gate mutant. We have validated the hypothesis that the steric gate residue prevents NTP incorporation by clashing with the 2' -OH group of ribose sugars. In addition, we have shown how the steric gate of Dpo4 is critical for stabilizing incoming nucleotides for optimal enzymatic efficiency. Lastly, we propose that altered sugar positioning may induce gamma phosphate hydrolysis of incoming nucleotides, causing decreased incorporation efficiency.

2.5 Methods

2.5.1 DNA and protein preparation

Oligonucleotides for crystallization were purchased from Keck Oligo Inc. and were purified by ion exchange chromatography, desalted, and lyophilized before use. The 13-nt primer (5'-GGGGGAAGGACTC^{dd}-3') containing a dideoxy 3' end (C^{dd}) was annealed to an 18-nt template (5'-TTCATGAGTCCTTCCCCC-3'), resulting in a T base at the first replicating position. Oligonucleotides used for primer extension assays were purchased from Sigma-Aldrich and gel purified. The 30nt template (5'-GAGCAGTCGCACATGTAGTATCTCTGTGAC-3') was annealed to the 16nt primer (5'-GTCACAGAGATACTAC-3') resulting in a T base at the first position beyond the primer-template junction. The primer was 5'-end labelled using [γ -³²P]ATP and T4 polynucleotide kinase. The 5'-labelled primer was annealed to the template at a 1.5:1 molar ratio. The Y12A mutation was generated by QuickChange site-directed mutagenesis (Stratagene). The Dpo4 wild type and Dpo4 Y12A mutant proteins were expressed in *E.coli* and purified as previously described (Ling et al., 2001).

2.5.2 Primer extension assays

The DNA substrate (10 nM) was incubated with either Dpo4 or Dpo4 Y12A (10 nM) and either all four dNTPs or rNTPs (20-100 μ M) or individual dNTPs or rNTPs (100 μ M) and reacted at 37°C for 2 min in buffer containing 40mM Tris (pH 8.0), 5mM MgCl₂, 250 μ g/ml BSA, 10mM DTT and 2.5% glycerol. Reactions were terminated with loading buffer containing 95% formamide, 20mM EDTA, and 0.025% bromophenol blue. Reactions were resolved on a 20% polyacrylamide gel containing 7 M urea. The gels were visualized using a PhosphorImager.

2.5.3 Crystallization and structure determination

Ternary complexes were formed for Y12A-dATP and Y12A-ATP by incubating protein (0.2 mM) and DNA in a 1:1.2 molar ratio with either dATP or ATP (2 mM) in the presence of 5 mM MgCl₂. Crystals were obtained in 12.5% PEG 3350 + 100 mM Ca(Ac)₂ + 100 mM Hepes pH 7.0 + 2.5% glycerol. Crystals were flash frozen in liquid nitrogen using a solution of 20% PEG 3350 + 100 mM Ca(Ac)₂ + 100 mM Hepes pH 7.0 + 20% ethylene glycol as a cryoprotectant. The data was collected using a Rigaku-MSR RU-200 generator and a mar345 image plate detector. The Y12A-dADP data was processed with MOSFLM (Leslie, 1992) and the Y12A-ADP data was processed with DENZO (Otwinowski and Minor, 1997) and SCALEPACK (Otwinowski and Minor, 1997). Both structures were solved by molecular replacement using PHASER (McCoy et al., 2005) with Dpo4 wild type (1JX4) as the search model. Refinement was performed using PHENIX (Adams et al., 2010). Model building was performed with COOT (Emsley and Cowtan, 2004) and figures were created with PYMOL (DeLano, 2002). Both structures have good stereochemistry, with over 95% of the residues in the allowed regions of the Ramachandran plot. The atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB) with accession codes 3PR4 and 3PR5 for the Y12A-dATP and Y12A-ATP structures respectively.

2.6 References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C. and Zwart, P.H. (2010) PHENIX: a comprehensive Python-based system for

- macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*, **66**, 213-221.
- Astatke, M., Ng, K., Grindley, N.D. and Joyce, C.M. (1998) A single side chain prevents *Escherichia coli* DNA polymerase I (Klenow fragment) from incorporating ribonucleotides. *Proc Natl Acad Sci U S A*, **95**, 3402-3407.
- Bauer, J., Xing, G., Yagi, H., Sayer, J.M., Jerina, D.M. and Ling, H. (2007) A structural gap in Dpo4 supports mutagenic bypass of a major benzo[a]pyrene dG adduct in DNA through template misalignment. *Proc Natl Acad Sci U S A*, **104**, 14905-14910.
- Biertumpfel, C., Zhao, Y., Kondo, Y., Ramon-Maiques, S., Gregory, M., Lee, J.Y., Masutani, C., Lehmann, A.R., Hanaoka, F. and Yang, W. (2010) Structure and mechanism of human DNA polymerase ϵ . *Nature*, **465**, 1044-1048.
- Bonnin, A., Lazaro, J.M., Blanco, L. and Salas, M. (1999) A single tyrosine prevents insertion of ribonucleotides in the eukaryotic-type phi29 DNA polymerase. *J Mol Biol*, **290**, 241-251.
- Brown, J.A., Fiala, K.A., Fowler, J.D., Sherrer, S.M., Newmister, S.A., Duym, W.W. and Suo, Z. (2009) A Novel Mechanism of Sugar Selection Utilized by a Human X-Family DNA Polymerase. *J Mol Biol*.
- DeLano, W.L. (2002) *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA, USA.
- DeLucia, A.M., Grindley, N.D. and Joyce, C.M. (2003) An error-prone family Y DNA polymerase (DinB homolog from *Sulfolobus solfataricus*) uses a 'steric gate' residue for discrimination against ribonucleotides. *Nucleic Acids Res*, **31**, 4129-4137.
- Doublet, S., Tabor, S., Long, A.M., Richardson, C.C. and Ellenberger, T. (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature*, **391**, 251-258.
- Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60**, 2126-2132.
- Gao, G., Orlova, M., Georgiadis, M.M., Hendrickson, W.A. and Goff, S.P. (1997) Conferring RNA polymerase activity to a DNA polymerase: a single residue in reverse transcriptase controls substrate selection. *Proc Natl Acad Sci U S A*, **94**, 407-411.
- Jarosz, D.F., Godoy, V.G., Delaney, J.C., Essigmann, J.M. and Walker, G.C. (2006) A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates. *Nature*, **439**, 225-228.

- Joyce, C.M. (1997) Choosing the right sugar: how polymerases select a nucleotide substrate. *Proc Natl Acad Sci U S A*, **94**, 1619-1622.
- Kirouac, K.N. and Ling, H. (2009) Structural basis of error-prone replication and stalling at a thymine base by human DNA polymerase iota. *Embo J*, **28**, 1644-1654.
- Leslie, A.G.W. (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography*, **26**.
- Ling, H., Boudsocq, F., Plosky, B.S., Woodgate, R. and Yang, W. (2003) Replication of a *cys--syn* thymine dimer at atomic resolution. *Nature*, **424**, 1083-1087.
- Ling, H., Boudsocq, F., Woodgate, R. and Yang, W. (2001) Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. *Cell*, **107**, 91-102.
- Lone, S., Townson, S.A., Uljon, S.N., Johnson, R.E., Brahma, A., Nair, D.T., Prakash, S., Prakash, L. and Aggarwal, A.K. (2007) Human DNA polymerase kappa encircles DNA: implications for mismatch extension and lesion bypass. *Mol Cell*, **25**, 601-614.
- McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C. and Read, R.J. (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr*, **61**, 458-464.
- Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2005) Human DNA polymerase iota incorporates dCTP opposite template G via a G.C + Hoogsteen base pair. *Structure*, **13**, 1569-1577.
- Nair, D.T., Johnson, R.E., Prakash, S., Prakash, L. and Aggarwal, A.K. (2004) Replication by human DNA polymerase-iota occurs by Hoogsteen base-pairing. *Nature*, **430**, 377-380.
- Nick McElhinny, S.A., Kumar, D., Clark, A.B., Watt, D.L., Watts, B.E., Lundstrom, E.B., Johansson, E., Chabes, A. and Kunkel, T.A. Genome instability due to ribonucleotide incorporation into DNA. *Nat Chem Biol*, **6**, 774-781.
- Nick McElhinny, S.A., Watts, B.E., Kumar, D., Watt, D.L., Lundstrom, E.B., Burgers, P.M., Johansson, E., Chabes, A. and Kunkel, T.A. Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc Natl Acad Sci U S A*, **107**, 4949-4954.
- Niimi, N., Sassa, A., Katafuchi, A., Gruz, P., Fujimoto, H., Bonala, R.R., Johnson, F., Ohta, T. and Nohmi, T. (2009) The steric gate amino acid tyrosine 112 is required for efficient mismatched-primer extension by human DNA polymerase kappa. *Biochemistry*, **48**, 4239-4246.

- Otwinowski, Z. and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol*, **276**, 307-326.
- Patel, P.H. and Loeb, L.A. (2000) Multiple amino acid substitutions allow DNA polymerases to synthesize RNA. *J Biol Chem*, **275**, 40266-40272.
- Sherrer, S.M., Beyer, D.C., Xia, C.X., Fowler, J.D. and Suo, Z. (2010) Kinetic Basis of Sugar Selection by a Y-Family DNA Polymerase from *Sulfolobus solfataricus* P2. *Biochemistry*.
- Traut, T.W. (1994) Physiological concentrations of purines and pyrimidines. *Mol Cell Biochem*, **140**, 1-22.
- Vaisman, A., Ling, H., Woodgate, R. and Yang, W. (2005) Fidelity of Dpo4: effect of metal ions, nucleotide selection and pyrophosphorolysis. *EMBO J.*, **24**, 2957-2967.
- Wong, J.H., Fiala, K.A., Suo, Z. and Ling, H. (2008) Snapshots of a Y-family DNA polymerase in replication: substrate-induced conformational transitions and implications for fidelity of Dpo4. *J Mol Biol*, **379**, 317-330.
- Yang, G., Franklin, M., Li, J., Lin, T.C. and Konigsberg, W. (2002) A conserved Tyr residue is required for sugar selectivity in a Pol alpha DNA polymerase. *Biochemistry*, **41**, 10256-10261.
- Zhou, G.H., Gotou, M., Kajiyama, T. and Kambara, H. (2005) Multiplex SNP typing by bioluminometric assay coupled with terminator incorporation (BATI). *Nucleic Acids Res*, **33**, e133.

Chapter 3¹

3 Structural basis of error-prone replication and stalling opposite a thymine base by human DNA polymerase ι

3.1 Introduction

Human Y family DNA polymerase ι ($\text{pol}\iota$) is a specialized polymerase that does not utilize Watson-Crick base pairing for DNA replication. Instead, this enzyme functions by inducing a *syn* conformation on template purines, which results in Hoogsteen base pairing with the correctly matched incoming nucleotide (Johnson et al., 2005; Nair et al., 2005; Nair et al., 2006). The ability to induce a nucleotide *syn* conformation by $\text{pol}\iota$ appears to serve as the mechanism for replication opposite damaged template purines. Structural evidence demonstrates that the 1, N⁶ ethenodeoxyadenosine and N2 ethyl guanine lesions are presented in the *syn* conformation protruding into the solvent accessible major groove of the DNA helix (Nair et al., 2006; Pence et al., 2008), which allows base pairing with the correct incoming nucleotide.

DNA replication by $\text{pol}\iota$ on template pyrimidines displays extremely high error rates, while incorporation opposite template purines is more accurate (Kunkel et al., 2003; Tissier et al., 2001; Tissier et al., 2000; Zhang et al., 2000). Opposite a template thymine (T), $\text{pol}\iota$ prefers to incorporate a guanine (G) up to 2.5 fold over the correctly paired adenine (A) in a metal-dependent manner (Frank and Woodgate, 2007). In addition, $\text{pol}\iota$ has inefficient replication past a template T base causing a signature T template stall

¹ The contents of this chapter have been published: Kirouac KN, Ling H (2009) Structural basis of error-prone replication and stalling at a thymine base by human DNA polymerase ι *EMBO J* 11: 1644-54.

(Zhang et al., 2000). A similar pattern of mis-incorporation and replication stalling by polt is observed opposite template uracil (U) (Vaisman and Woodgate, 2001). It has been proposed that G mis-incorporation opposite template U could restore the genomic sequence of cytosines (C) that have undergone deamination. The biological role of this preferred mis-insertion of G opposite template T or U within cells remains unknown. However, such an unusual and highly specific property could serve a unique function in DNA maintenance.

The error-prone replication on template T by polt has been implicated in the high rates of DNA mutagenesis presented in patients with the UV-sensitive disorder; *Xeroderma Pigmentosum Variant* (XP-V) syndrome (Wang et al., 2007). When the human Y family DNA polymerase η (pol η) is inactivated by mutations, polt takes over its specialized role of bypassing UV-induced thymine-thymine (T-T) dimers. However, the preference of mis-insertion opposite template T by polt results in an increase in mutagenesis and the presentation of the disease. Although polt is responsible for increased DNA mutagenesis when functioning out of context, this enzyme does appear to play a role in tumour suppression. Mice deficient in both pol η and polt have an earlier onset of UV-light induced tumours than pol η deficiency alone (Dumstorf et al., 2006; Ohkumo et al., 2006), indicating a role for polt in UV-induced lesion bypass. The unique T template mis-incorporation by polt has been extensively reported, but has remained mechanistically unexplained.

Here, we report three crystal structures of polt in complex with DNA containing a template T base in the active site, which is paired with either correct (A) or incorrect (T

or G) incoming nucleotides. Our results reveal, for the first time, the structural basis of preferred G mis-incorporation and stalling on a template T base by polt.

3.2 Results

3.2.1 Polt-DNA-dNTP complexes with template DNA bending at the thymine base

In order to position the thymine base at the polt active site, DNA substrates for crystallization were designed containing a thymine 5' to the template-primer junction. The first substrate (substrate 1) is a 15/9-nt duplex DNA with two thymines 5' to the template-primer junction. The second substrate (substrate 2) is an 18-nt self-complementary duplex containing a dideoxy 3' primer end for trapping ternary complexes. Incoming 2', 3'-dideoxy ATP (ddATP) was incubated with DNA substrate 1 and co-crystallized with polt, while dGTP and dTTP were incubated with DNA substrate 2 and co-crystallized with polt. Interestingly, the polt-DNA-ddATP complex was trapped at the first T from the template-primer junction without the expected one incorporation, probably due to the replication stalling at T. The resulting three crystal structures are denoted as T:ddADP, T:dGTP and T:dTTP according to the replicating base pair in the active site. The ternary complex crystals are in two different space groups (C2 for T:ddADP; P6₅22 for T:dGTP and T:dTTP) and diffract to 2.0 Å, 2.0 Å, and 2.2 Å resolutions, respectively (Table 3.1). In T:ddADP, incoming ddATP was hydrolyzed to ddADP. Hydrolysis in this manner has been observed in Dpo4 from *Sulfolobus solfataricus*, the model enzyme of the Y-family DNA polymerase, due to a weak phosphatase activity of the polymerase

Table 3.1 Summary of crystallographic data for polt template thymine structures

Crystal	T:ddADP	T:dTTP	T:dGTP
Space group	C2	P6 ₅ 22	P6 ₅ 22
Complexes per AU ^a	2	1	1
Unit cell			
<i>a, b, c</i> (Å)	140.2, 71.8, 127.4	97.2, 97.2, 201.9	98.1, 98.1, 203.7
β (°)	112.5		
Resolution range (Å) ^b	52.0-2.0 (2.04-2.00)	27.0-2.2 (2.26-2.20)	24.0-2.0 (2.07-2.00)
R_{merge} ^b	7.2 (61.7)	8.44 (83.8)	12.1(57.7)
$I / \sigma I$	24.6 (2.1)	36.8 (2.3)	57.9 (3.0)
Completeness (%) ^b	99.0 (96.3)	99.9 (100)	99.7 (100)
Redundancy ^b	3.6 (3.1)	11.6 (11.7)	13.3 (8.9)
No. reflections (test)	73972 (2%, 1551)	28762 (2%, 604)	38810 (2%, 826)
$R_{\text{work}} / R_{\text{free}}$	20.6 / 25.3	21.2 / 25.8	21.8 / 24.9
No. atoms			
Protein	6028	2889	2903
DNA	846	326	323
dNTP	50	28	30
Mg ²⁺ ions	-	2	2
Ca ²⁺ ions ^c	9	-	-
Waters	735	279	238
B-factors			
Protein	62.5	43.4	43.5
DNA	52.7	40.1	44.5
dNTP	29.0	32.1	48.8
Metal ions	35.0	24.5	50.7
Water	39.2	45.1	46.8
R.m.s.d. bond lengths (Å)	0.016	0.016	0.019
R.m.s.d. bond Angles (°)	1.96	1.72	1.82

^a AU means asymmetric unit

^b Data in the highest resolution shell are in parentheses

^c Only two Ca²⁺ ions are coordinated at each active site

(Ling et al., 2001). Within the T:ddADP structure, the active site metal ions have been refined as Ca^{2+} , due to the presence of 150 mM CaCl_2 within the crystallization buffer and the high electron density. In addition, anomalous signal peaks were observed at the active site metal ion positions after generating an anomalous map at a wavelength of 0.98322 Å. This is analogous to Dpo4 structures crystallized with 100 mM CaAc_2 (Ling et al., 2001). Anomalous signal peaks were not observed for the T:dTTP or T:dGTP structures, which were both crystallized in the absence of Ca^{2+} ions. Thus, the T:dTTP and T:dGTP structures were refined with two active site Mg^{2+} ions. To ensure the polt structure containing Ca^{2+} ions was accurately representing polt fidelity opposite template T, primer extension assays were performed with polt in the presence of 150 mM CaCl_2 (Figure 3.1). These assays demonstrate that Ca^{2+} ions do not change the nucleotide incorporation specificity of polt and that the T:ddADP structure can be confidently used for structural interpretation (Figure 3.1). The divalent cation in the B site is positioned identically within all three T template structures, as well as previous polt structures containing template purines (Nair et al., 2006) and Dpo4 ternary structure (Vaisman et al., 2005). The divalent cation in the A site however, is mobile with variable positions in all three structures. Divalent ion mobility within the A site has also been reported previously for polt (Nair et al., 2005) and Dpo4 (Vaisman et al., 2005).

Polt in all three ternary structures is essentially identical to that of the previously solved, purine-template polt structures (Nair et al., 2005; Nair et al., 2006; Nair et al., 2004) (Figure 3.2A). The pair-wise comparisons on all C α atoms produced root mean square deviations (rmsd) within 0.7 Å among our three complex structures. In addition, the C α rmsd is ~0.8 Å between T:ddADP and a previously solved polt ternary complex (PDB:

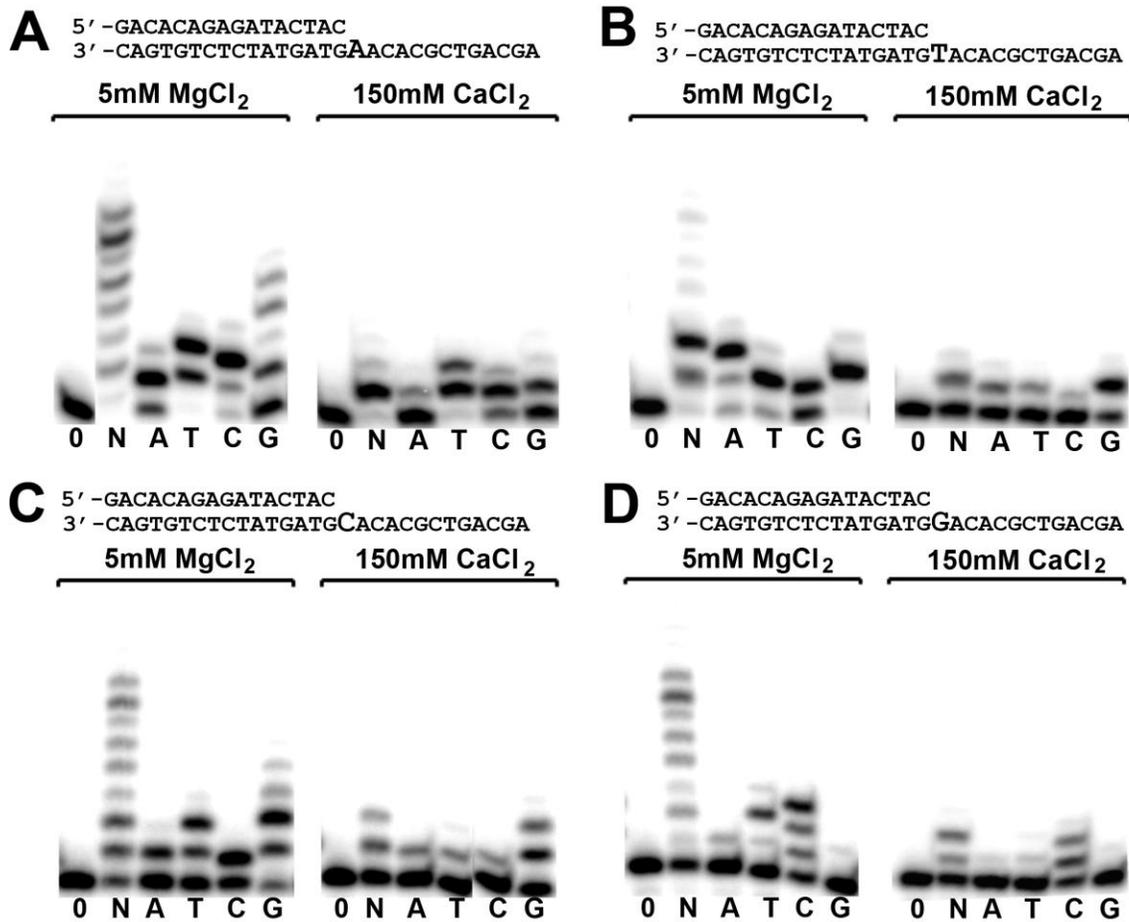


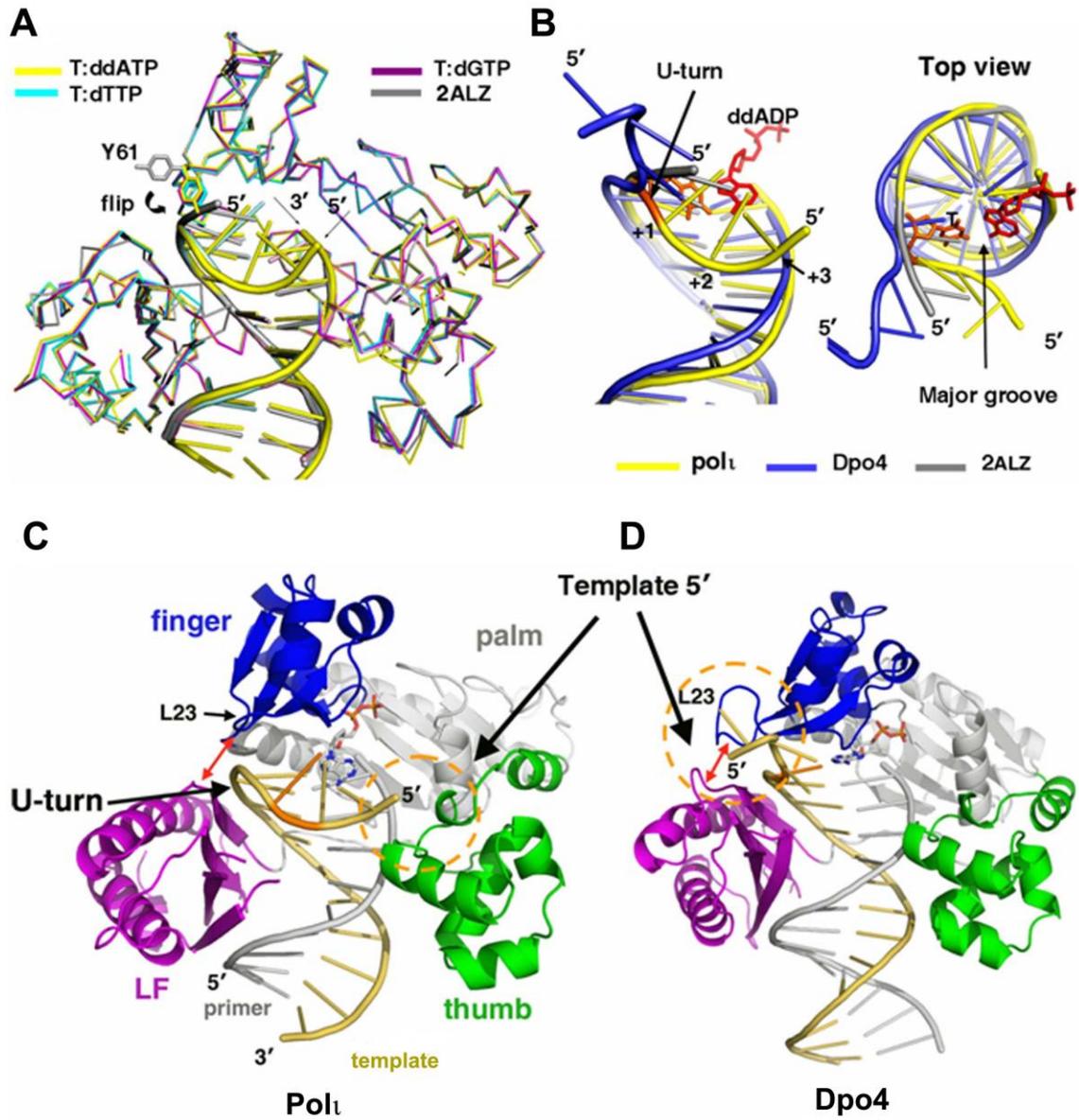
Figure 3.1 The effect of calcium on poliovirus replication. Primer extension assays were carried out in the presence of either 5mM MgCl₂ or 150mM CaCl₂ opposite template thymine (A), adenine (B), cytosine (C), or guanine (D). Poliovirus was incubated with DNA substrates in the presence of either no incoming nucleotides (0), all four incoming nucleotides (N), or individual nucleotides (A, T, C, G) as labelled below each lane.

2ALZ) containing a purine base at the template position. The close agreement between all of these polt structures indicates that polt, like other Y-family polymerases, does not undergo significant conformational change when replicating through different DNA substrates (Bauer et al., 2007; Ling et al., 2001; Nair et al., 2006).

We use the first complex structure (T:ddADP) to describe the general features of the three ternary complexes (Figure 3.2), since these three structures are identical, except for the replicating base pair. All residues of polt have the same side-chain conformations in the current three structures, which are identical to those of the previously reported purine-template polt structures, except for Tyr 61. Tyr 61 flips its side chain conformation 100° from that seen in the purine-template polt structures, and moves its aromatic ring 9 Å closer to the template DNA (Figure 3.2A). The unique Tyr61 orientation is observed in all three of our template thymine structures reported here, which are in two different crystal forms. Thus, this conformation is independent of crystal packing and is likely induced by the DNA substrate that contains a template T base.

A striking structural change is observed in the single-stranded DNA template when comparing our T template base structures with previous polt structures containing template purines and with Y-family polymerase Dpo4 (Figure 3.2). The previous polt and Dpo4 structures project the single stranded template DNA away from the active site in extended conformations (Figure 3.2). In our T template base structures, the single stranded DNA is flipped back upon itself in a 'U' shaped conformation, enclosing the replicating base pair from the DNA's major groove and approaching the polymerase thumb domain across the major groove (Figure 3.2B,C). The DNA backbone is bent ~90° after the template T base towards the major groove, and the +1 nucleotide (5' to T)

Figure 3.2 Comparison of polt-DNA-nucleotide ternary structures. The colour schemes are shown either as the colour of the appropriate labels or the colour bars in the panels. **(A)** Superposition of T:ddADP (yellow), T:dTTP (cyan), T:dGTP (magenta), and a previously solved polt ternary complex (PDB: 2ALZ, grey). Proteins are shown as α traces and Tyr 61 is shown as sticks to highlight its conformational change as a result of the DNA U-turn. Incoming nucleotides were omitted for clarity. **(B)** DNA superposition of T:ddADP (yellow), Dpo4 ternary structure (PDB: 2AGQ, blue), and a previously solved polt ternary complex (PDB: 2ALZ, grey). Top view is also shown with the template T base in orange and the incoming ddADP in red. **(C)(D)** Polt (T:ddADP) and Dpo4(PDB: 2ALZ) ternary complexes. DNA template strands are shown in yellow, T bases in orange, and primer strands in grey. The U-turn DNA and position of the 5' template end are indicated by the appropriate arrows. LF represents the little finger domain.

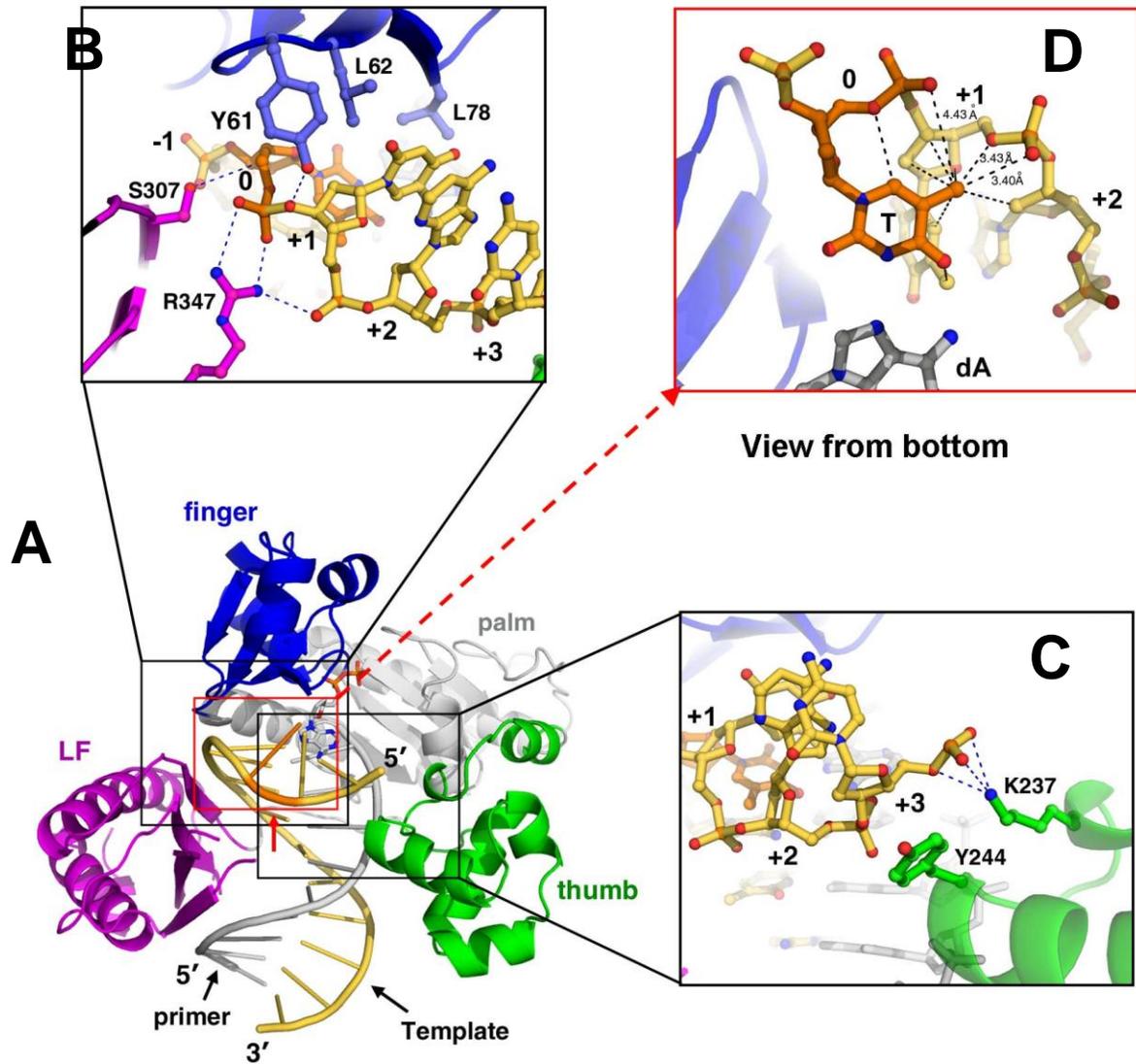


is oriented 90° to the template T base. This latter characteristic is completely different from the template/+1 base relationships in the extension template strands in the other Y-family polymerase structures (Figure 3.2B). In our structures, the +1 base lies perpendicular to the template T base (position 0) due to DNA strand binding. All three of our polt structures display this unique ‘U-turn’ DNA conformation after the T base, irrespective of the incoming nucleotide or the identity of the bases flanking the template T.

3.2.2 Contributions of polt domains and T template base to template stabilization

The unique ‘U-turn’ DNA conformation in our polt structures is stabilized by both polt and the unique T template base within the active site, which likely induces replication stalling. The single-stranded template DNA downstream of the T base is held in position by the finger domain, the little finger (LF) domain and the thumb domain of polt (Figure 3.3A). Tyr 61, Leu 62 and Leu 78 from the finger domain contact the +1 nucleotide and provide strong hydrophobic interactions to the backbone sugar via the aromatic side chain of Tyr 61 and to the +1 nucleotide base from interactions with the two leucine residues (Figure 3.3B). The positively charged Arg 347 and polar Ser 307 from the LF domain contact the phosphate backbone of the template at nucleotide positions 0 and +1 (Figure 3.3B). The 5' end of the template is in contact with the thumb domain at Tyr 244 and Lys 237 (Figure 3.3C). Tyr 244 stacks with the sugar of the +3 nucleotide, and the positively charged Lys 237 interacts with the negatively charged phosphate to fix the free 5' end of the template strand in front of the replicating base pair (Figure 3.3C). Furthermore, the $\sim 90^\circ$ bend at the T base is stabilized by the interactions between the

Figure 3.3 Structure of T:ddADP showing template DNA ‘U-turn’ stabilization. Numbers indicate template nucleotide positions relative to T at position 0. Hydrogen bonding is shown as blue dashed lines. **(A)** Overall T:ddADP structure is shown with DNA template strand in yellow, T base in orange, and the primer strand in grey. The finger domain is shown in blue, little finger (LF) domain in purple, thumb domain in green, and palm domain in grey. **(B)** Zoom-in view of the ‘U-turn’ stabilization by the polt finger (blue) and LF (magenta) domains. **(C)** Zoom-in view of the ‘U-turn’ stabilization by the polt thumb domain (green). **(D)** Zoom-in view of the ‘U-turn’ stabilization by the template T base (orange). Hydrophobic interactions are shown as black dashed lines. View is seen from underneath and looking up through the red square in panel A.

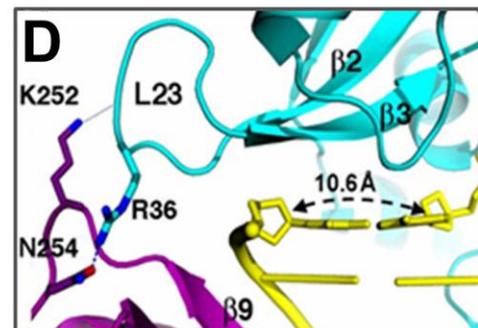
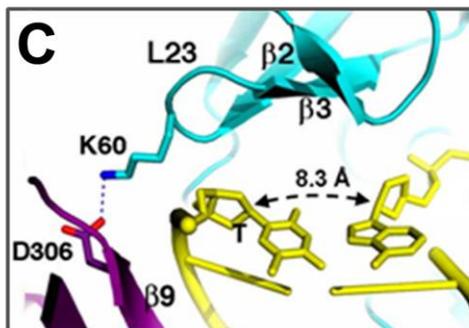
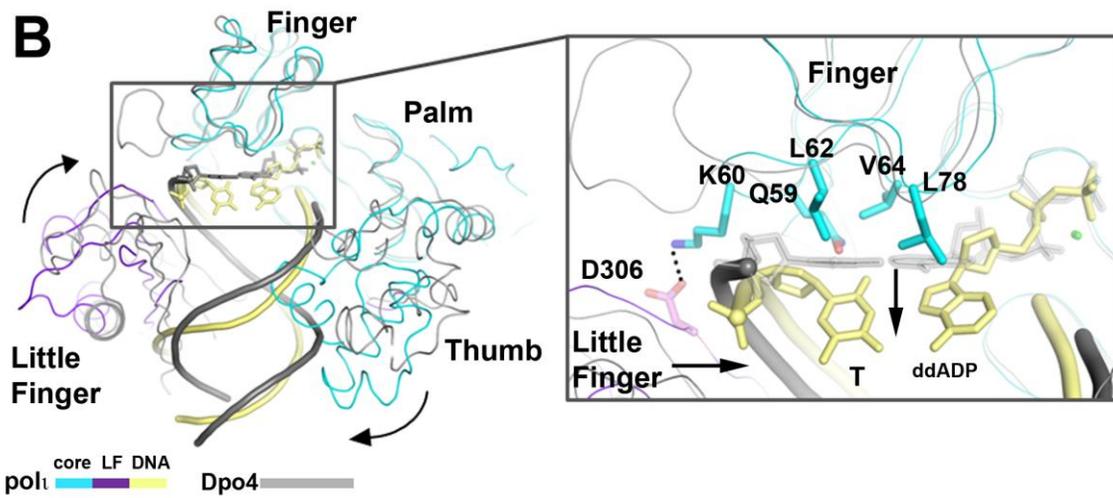
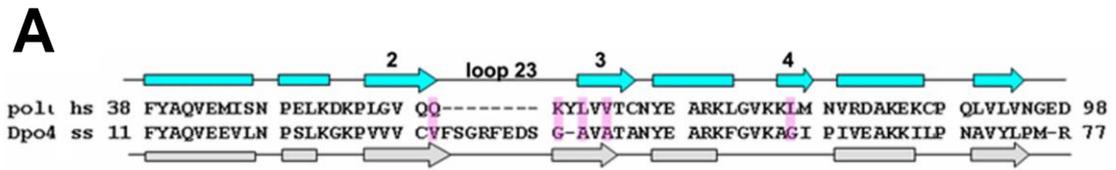


methyl group of the template T base and the bent single-stranded template DNA (Figure 3.3D). Although the extensive contacts between this unique methyl group and the +1 nucleotide reinforce the unusual bending template, the methyl group may not be an absolute requirement, due to a similar stalling effect opposite template U (Vaisman and Woodgate, 2001). The unique ‘U-turn’ conformation is not observed in the presence of template purine bases (Nair et al., 2005; Nair et al., 2006). Two of our template thymine structures (T:dTTP and T:dGTP) have the same sized DNA substrate, single stranded DNA overhang, and crystal form as the previous purine template structures. However, these template thymine structures adopted the ‘U-turn’ conformation, similar to the T:ddADP structure, which has a different DNA substrate and crystal form (Figure 3.2A). This excludes any structural variation caused by differences in the single stranded DNA and packing environments of the complexes. It appears that the purine bases A and G are too large to be accommodated in the U-turn conformation observed in our structures, which would disrupt bending. However, there are likely other unidentified factors also involved in preventing this conformation. Although template C has a similar size to T and U, which may lend itself to a ‘U-turn’ structure, such a conformation may not be stable when template C is in the active site, as no significant replication stalling has been observed with this template base. The U-turn interactions are not involved with the template bases in the double strand DNA except the template T and are mainly involved with backbone atoms on the downstream single strand DNA. This is consistent with the observations that the stalling only depends on the T base and not the bases flanking it (Zhang et al., 2000). The current T template structures clearly show the structural basis for the signature T template stalling by pol λ .

3.2.3 Role of the finger domain and the polt active site

Structural comparison of polt and Dpo4 ternary complexes indicate that the Y-family polymerases are highly structurally conserved in the core area that forms the DNA-binding cleft, except for the lid region of the finger domains (Figure 3.4A). The finger domains are structurally conserved with most secondary structural elements aligned between polt and Dpo4 (Figure 3.4A). However, there are two striking differences in these domains, which affect the shape of the active sites and their interactions with DNA substrate. These differences are concentrated in the fragment between the β -strands $\beta 2$ and $\beta 3$, which is in the non-conserved substrate recognition site that contacts the template DNA and the replicating base pair in the active site (Yang, 2003). First, polt has a much shorter loop between $\beta 2$ and $\beta 3$ than that of Dpo4 (Figure 3.4A). This loop (L23) forms a structural interface for the finger domains to interact with the LF domains, as well as the single stranded template DNA in Dpo4 (Figure 3.4B-D). The shorter loop of polt causes the LF domain to rotate inward to the finger domain and the $\beta 9$ strand moves ~ 2 Å towards the template strand, creating a narrowed active site in polt (Figure 3.4B,C). The narrowed active site limits the C1'-C1' distance of the replicating base pair to within 9 Å in polt (Figure 3.4C). The C1'-C1' distances are 8.3 Å, 8.6 Å, and 8.9 Å in the structures of T:ddADP, T:dTTP, and T:dGTP, respectively. In contrast, the replicating base pair in the active site of Dpo4 has a C1'-C1' distance of ~ 10.6 Å (Figure 3.4D), which is a common strand width for B form DNA in all other Y-family polymerase structures (Alt et al., 2007; Lone et al., 2007; Nair et al., 2005). The second difference between polt and Dpo4 finger domains lies in the fragment that contacts the replicating base pair within the active site. Polt has relatively large amino-acid side chains (Gln 59, Lys 60, Leu 62, Val

Figure 3.4 Polt and Dpo4 active site comparison. **(A)** Structure-based sequence alignment of amino acids for the finger domains of polt (cyan) and Dpo4 (grey). Numbers 2, 3, and 4 indicate the second, third, and fourth β -sheets. Secondary structure is indicated as rectangles for α -helices and arrows for β -sheets. Residues interacting with the replicating base pair are highlighted in magenta. **(B)** Superposition of T:ddADP (cyan, purple) and ternary Dpo4 (type I)-DNA-nucleotide (1JX4, grey). The incoming nucleotides are shown as sticks for Dpo4 (grey) and T:ddADP (yellow). Curved arrows indicate domain movement relative to Dpo4. Zoom in view of active sites show the little finger and bulky amino acids in polt finger domain pushing the template base in towards the primer and downwards (black arrows) **(C)(D)** Close-up views of active sites showing finger-LF domain interactions in polt and Dpo4. Finger domains are cyan, LF domains are purple, and DNA is yellow.



64, Leu 78) contacting the replicating base pair at the active site (Figure 3.4B) when compared to Dpo4, which has relatively small amino acids (Val 32, Gly 41, Ala 42, Ala 44, Gly 58) for the contacts. The larger side chains of Gln 59, Leu 62, Val 64, and Leu 78 in polI push the replicating base pair towards the major groove, effectively tilting it off plane relative to that of Dpo4 (Figure 3.4B). Lys 60 and the residues from strand β 9 of the polI LF domain squeeze the template base towards the incoming nucleotide and make the C1'-C1' distance shorter than 9 Å (Figure 3.4B,C). These structures reveal that the finger domain is not only important for contacting the replicating base pair but is also an essential factor for restricting the C1'-C1' distance in polI. Therefore, the polI finger domain is most likely the functional domain responsible for the nucleotide specificity during replication.

3.2.4 Conformation and stability of replicating base pairs in the polI active site

The unique polI active site and the T template base make the replicating base pairs in the ternary complexes different in conformation from those in other Y-family polymerase ternary structures. All three of our polI structures have the template T base in a normal *anti* conformation when it is paired with an incoming nucleotide in the active site (Figure 3.5). Nucleotide binding does not induce a conformational change in the T template base as observed in the purine-template structures (Nair et al., 2005; Nair et al., 2006). Instead, the incoming dNTPs of the replicating base pairs in our polI structures adopt different conformations, depending on their fit in the enzyme active site. Due to large residues from finger domain, the template T base is pushed out of the stacking area with underneath bases and tilted off plane by the finger domain in all three structures (Figure

3.5). The tilt (τ) and roll (ρ) angles of the off plane T from the underlying base are around 6° and 16° , respectively.

In the T:ddADP structure, the incoming ddADP adopts a *syn* conformation and forms a Hoogsteen base pair with the template T (Figure 3.5A). The Hoogsteen base pair in the T:ddADP structure fits the narrowed active site with a C1'-C1' distance of 8.4 Å, which is similar to other reported polt structures (Nair et al., 2005; Nair et al., 2006; Nair et al., 2006) and is smaller than the required C1'-C1' distance of ~ 10.6 Å for proper Watson-Crick base pairing (Ling et al., 2003; Ling et al., 2001; Ling et al., 2004; Ling et al., 2004; Wong et al., 2008). In addition, incoming ddADP is flipped out of the stacking area of the underlying base pair towards the major groove due to its *syn* conformation. The ddADP is tilted $\sim 20^\circ$ off plane with the underlying base pair and has an elongated stacking distance of about 4 Å, which weakens the stability of the replicating base pair further (Figure 3.5A).

In the T:dTTP structure, the mismatched incoming dTTP is in an *anti* conformation (Figure 3.5B). The narrowed active site holds the pyrimidine-pyrimidine base pair well, due to the pair being smaller than the common pyrimidine-purine base pair in contacting distance. The narrowed active site thus stabilizes the small pyrimidine-pyrimidine mismatched base pair. The C1'-C1' distance of the T:T base pair is 8.5 Å, which would not be stable in an active site that accommodates a standard Watson-Crick base pair with C1'-C1' distance of ~ 10.6 Å. Interestingly, the incoming dGTP is also in the *anti* conformation, which has not been observed in other polt structures that contain template purine bases in the active site (Nair et al., 2005; Nair et al., 2006) (Figure 3.5C). The

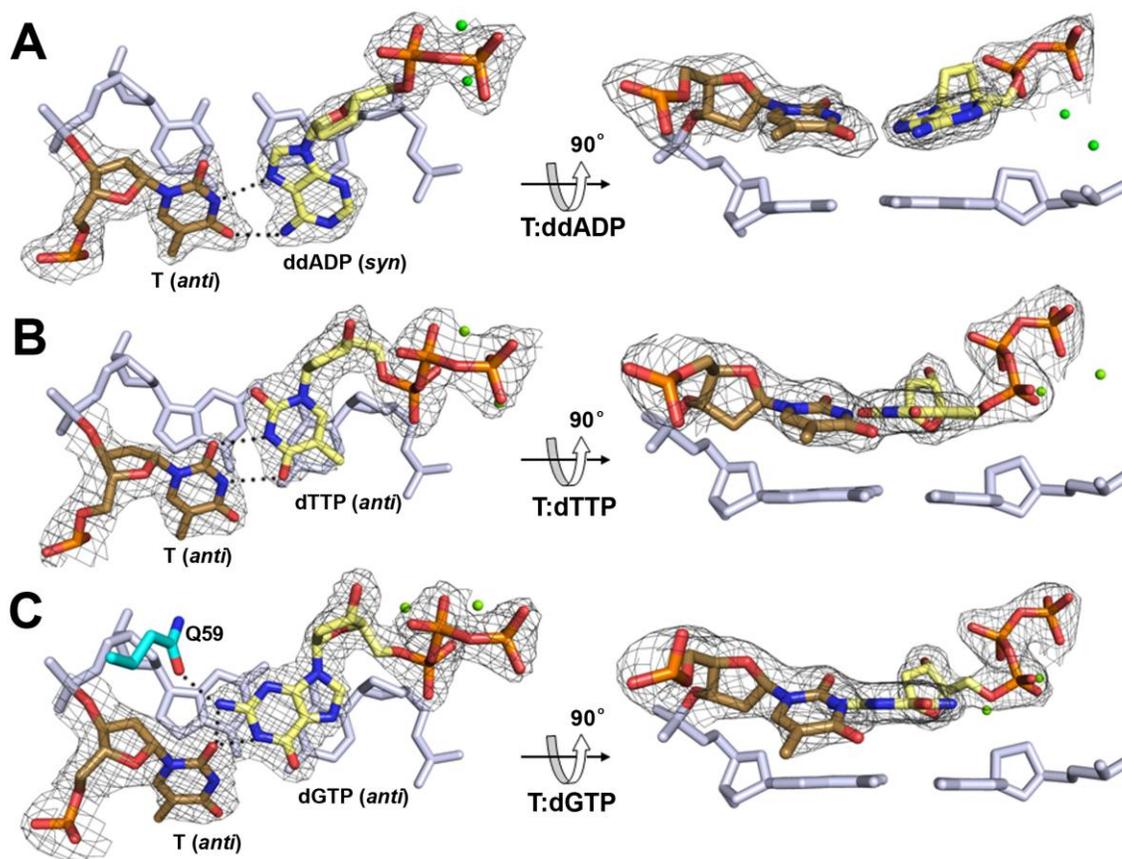


Figure 3.5 Base stacking and hydrogen bonding of replicating base pairs. (A) T:ddADP structure (B) T:dTTP structure (C) T:dGTP structure. The template T base is shown in brown, the incoming nucleotide is shown in yellow, and the underlying base pair is shown in grey. Hydrogen bonds are represented as dashed lines. Protein side chains involved in hydrogen bonding are shown in cyan. Green spheres represent divalent cations. The $2F_o - F_c$ electron density map is contoured at 1σ level.

C1'-C1' distance is restricted to 8.9 Å, which causes the template T to tilt an extra 15° off plane in order to accommodate the *anti* conformation of the dGTP nucleotide. Our structural observation shows that a *syn* conformation on purine nucleotides can occur in the template or incoming nucleotide position and is the result of a narrowed active site that constrains the C1'-C1' distance of the replicating base pair. In contrast to ddADP, the bases of dTTP and dGTP in *anti* conformations remain within the active site, parallel to the underlying base pair (Figure 3.5B,C) with stacking distances in the normal range of 3.2 - 3.6 Å. Compared to incoming A and T bases, the G base of dGTP has the largest stacking surface due to its purine base and *anti* conformation. Base stacking between the incoming nucleotide and the underlying base pair is critical to the stability and preference of nucleotide incorporation (Yang, 2006). Therefore, G is most favourable and A is the least favourable opposite template T in terms of its base-stacking properties.

In addition to base stacking, incoming bases are also stabilized by hydrogen bonding with template bases. There are two hydrogen bonds between incoming A (N⁶ and N7) and the template T (N3 and O⁴) in the Hoogsteen base pair in T:ddADP (Figure 3.5A). In the T:dTTP structure, there are also two hydrogen bonds formed between the template T (O² and N3) and the incoming dTTP (N3 and O⁴) (Figure 3.5B). Accordingly, the hydrogen-bonding forces of the replicating base pairs are comparable in these two complexes. Thus, the loss of base stacking on ddADP makes the mismatched dTTP more favourable for incorporation than the A base. Incoming dGTP also forms two hydrogen bonds to the template T, as well as a unique third hydrogen bond between its N² atom and OE1 of Gln59 from the finger domain (Figure 3.5C). This special hydrogen bonding with Gln59 is the first to be identified in a polt structure and reveals a unique stabilizing force that

favours, over other bases, the mis-incorporation of dGTP opposite template T by polt. The structural observation that incoming dGTP is the most stable incoming nucleotide is supported by the observation that dGTP binding affinity opposite template T by polt is greater than dTTP or dATP and is the same for dTTP binding affinity opposite template A (Washington et al., 2004).

3.2.5 Role of polt finger domain in base incorporation specificity and replication stalling

The finger domain of polt contacts the replicating base pair and pulls the LF domain towards the active site, which contributes to the replication specificity. In order to confirm that the polt finger domain determines the replication specificity, we generated two Dpo4-t chimeric proteins with functional domains switched between Dpo4 and polt. The LF domain has been implicated in contributing to replication specificity (Boudsocq et al., 2004), thus the finger or LF domains of Dpo4 were replaced with the corresponding counterparts of polt in the chimeras. Four DNA substrates containing either T, A, C, or G at the first replicating position and a T base at the eighth replicating position were used for the functional assays (Figure 3.6). The chimeric proteins were tested by primer extension assays with wild-type Dpo4 and polt as controls.

Opposite the T template base, polt has a high mis-incorporation rate of G and T (vertical arrows in Figure 3.6A) opposite template T, as the primers (bottom bands) are almost fully reacted for dGTP (lane G) and dTTP (lane T). Multiple bands are observed due to the low processivity of these enzymes. In contrast, Dpo4 incorporates the correct A nucleotide preferentially (lane A) with dramatically reduced reactions with dTTP (lane T), dGTP (lane G), and dCTP (lane C) compared to dATP (Figure 3.6A). Opposite the A

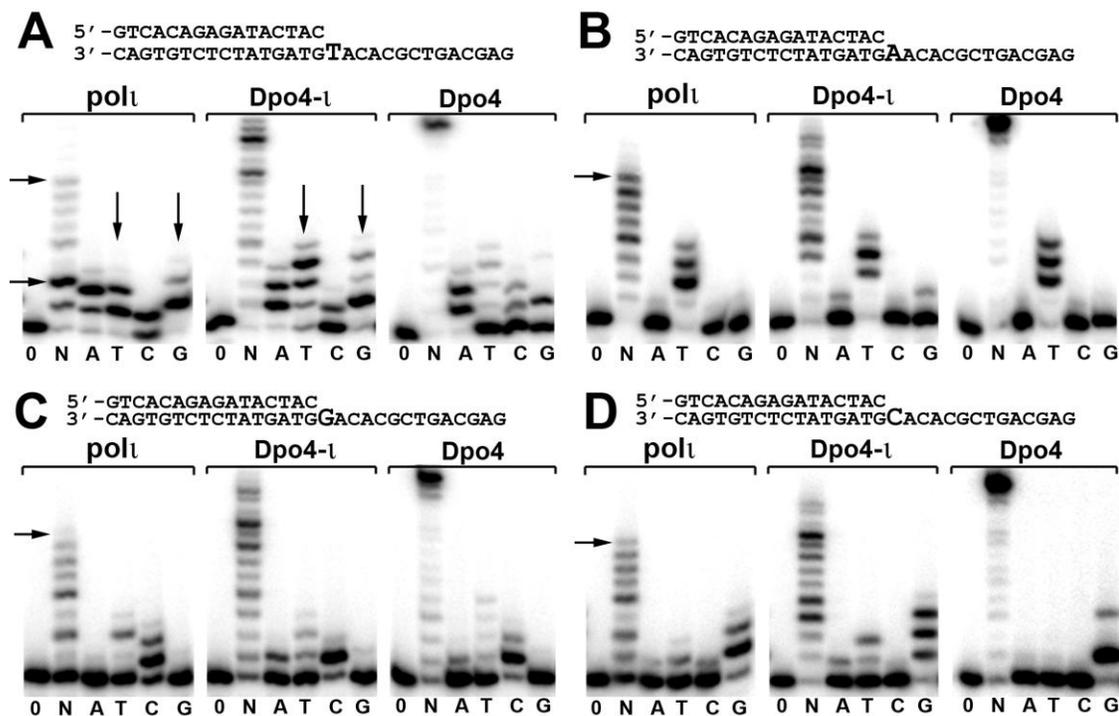


Figure 3.6 The role of the finger domain in nucleotide incorporation specificity. Primer extension analysis was used to examine nucleotide incorporation opposite (A) template T base, (B) template A base, (C) template G base, and (D) template C base by polI, Dpo4-l (finger domain) chimera, and Dpo4. The first replicating template base and the T template base at the 8th position are bolded. Horizontal arrows indicate replication stalling, while vertical arrows indicate mis-incorporation. Enzymes were incubated with DNA and either no nucleotides (0), all four dNTPs (N), or individual dNTPs (A,T,C,G).

template (Figure 3.6B), both Dpo4 and pol ι have quite accurate incorporation with preference for the correct incoming nucleotide dTTP (lanes T in Figure 3.6B). The polymerases against the G template in Figure 2.6C show similar patterns, with C preferentially inserted. Interestingly, Dpo4 replicates the C template (Figure 3.6D) accurately, while pol ι preferentially inserts G with significant mis-incorporations of T (vertical arrows in Figure 3.6D). Remarkably, Dpo4- ι finger domain chimera (Dpo4- ι) adopted a high mis-incorporation rate of G and T opposite template T similar to pol ι , as the primer bands in the lanes T and G are almost fully reacted (vertical arrows in Figure 3.6A), which is similar to pol ι and different from Dpo4 (Figure 3.6A). Accordingly, opposite template A, G, and C, the Dpo4- ι resembles pol ι and differs from wild type Dpo4 (Figure 3.6B,C). The primer extension assays indicate that replication specificity is dominated by the finger domain of the Y-family polymerases as the finger swapping converts Dpo4 into a pol ι -like protein in terms of nucleotide incorporation. Interestingly, the other Dpo4- ι chimera with the LF domain swapped into Dpo4 (Dpo4- ι -LF) does not show any base incorporation pattern changes from Dpo4 to pol ι (data not shown), in contrast to what we observed in the Dpo4- ι mutant. This LF-replacement chimera is very different from the Dpo4-Dbh chimeric proteins in which the enzymatic properties of the mutants are mainly influenced by their LF domains (Boudsocq et al., 2004). In the latter case, the LF domain is swapped between two very similar Y-family polymerases Dpo4 and Dbh, which have almost identical substrate recognition sites. The Dpo4-Dbh chimeras show functional differences of the LF domains between homologues sharing very similar finger domains (Boudsocq et al., 2004). In our case, the dramatic difference between the finger domains of Dpo4 and pol ι masks the influence from the LF domains.

Overall, the mutagenesis data clearly support the structural observations that the finger domain plays an important role in nucleotide incorporation specificity, particularly for G and T mis-insertion opposite template T and in determining the replication specificity.

Replication stalling is observed for polt at the T template bases (horizontal arrows in Figure 3.6) but not for Dpo4 and Dpo4- τ . When all four nucleotides (lanes N) are present in the assay, polt has poor extension beyond the first T template base (Figure 3.6A) and stops at the downstream 8th T base (labelled with horizontal arrows in Figure 3.6), while Dpo4 extends the primer to the end of the template DNA (top bands in Figure 3.6) with better processivity than polt. The finger domain alone does not appear to control the stalling property of the enzymes, as the chimeric protein Dpo4- τ extends replication beyond the 8th T base. This finding is consistent with our structural observation that three domains, instead of the finger domain alone, contribute to the stabilization of the ‘U-turn’ DNA that leads to replication stalling at the T base.

3.3 Discussion

3.3.1 The polt finger domain creates a unique active site that induces low fidelity opposite pyrimidine bases

Polt displays a wide diversity on its fidelity between template purines versus template pyrimidines (Kunkel et al., 2003; Tissier et al., 2001). Although this enzyme has low error rates opposite template purines, it has the highest error rate of any known polymerase opposite template thymines (Johnson et al., 2000; Tissier et al., 2000; Zhang et al., 2000). The diversified substrate-recognition site in the finger domain changes in size and residue identity across the Y-family members and is expected to be responsible for the specificity of nucleotide incorporation (Ling et al., 2001; Yang, 2003). The β 2-

turn- β 3 loop (L23) is the only part of the polymerase core that contacts the LF domain in the Y-family ternary complex structures. Our structural analyses indicate that the shorter polt L23 of the finger domain induces a movement of the LF domain towards the template strand, which results in a narrowed active site. It is conceivable that the finger domain causes the LF shifting towards the template in the active site, as the LF is the most flexible domain in the Y-family structures (Wong et al., 2008).

Polt promotes a T:G mismatch by its unique narrowed active site and specific interactions with the replicating base pair. As a good structural fit, the narrowed active site supports pyrimidine:pyrimidine mismatches when the template base is a pyrimidine. Higher mis-incorporations of T:T, T:C, C:T and C:C were observed in our primer extension assays for both polt and Dpo4-t relative to mis-incorporations opposite template purines. When the template base is a purine, the smaller, incoming pyrimidine maintains its base stacking in *anti* conformations (Nair et al., 2006), while a larger, incoming purine nucleotide would be difficult to fit in the narrowed active site for a purine:purine mismatch. This mechanism prevents mis-incorporation against purine template bases and allows for accurate replication. Our structural and biochemical analyses are consistent with the well documented, high error rates against template pyrimidine bases and relatively high fidelity against purine bases by polt (Johnson et al., 2000; Tissier et al., 2000; Zhang et al., 2000).

Interestingly, in the presence of Mn^{2+} , polt has increased fidelity opposite template thymine with a preference of incorporating the correct A nucleotide instead of incorrect G (Frank and Woodgate, 2007). The Mn^{2+} ion has a more relaxed and mobile coordination

than Mg^{2+} within the active site of polymerases. This effect likely allows the incoming nucleotide to adopt a variety of conformations that would not be possible with Mg^{2+} . In this manner, Mn^{2+} ion coordination by polt may render a favourable interaction that selects A over G opposite template T.

3.3.2 Replication stalling is stabilized by conserved residues over three domains

Another unique feature of polt is a pronounced stalling of replication in extending a primer strand opposite a T base (Zhang et al., 2000). Our structures of polt reveal a unique template DNA ‘U-turn’ conformation at the DNA’s single-stranded side that may effectively stall replication. The back-bending ‘U-turn’ conformation is stabilized by specific interactions from the unique methyl group of the template T base and a collection of interactions from three domains of polt. Three domains are involved in interacting with the “U-turn”: the finger domain (Tyr 61, Leu 62, and Leu 78), little finger domain (Ser 307 and Arg 347), and thumb domain (Lys 237 and Tyr 244). These combined interactions stabilize the bent single-stranded DNA and appear to hinder its translation into the active site for primer elongation. In addition, the highly bent DNA may also reduce the catalytic efficiency of polt, due to the observation that dNTP incorporation opposite template T is much slower than opposite template A (Washington et al., 2004). Replication stalling by polt may be involved in recruiting another polymerase for primer extension after insertion opposite template T or U.

3.3.3 Conclusions

Polt uniquely replicates DNA with a constrained active site, creating shorter C1'-C1' strand distances, with the finger domain projecting the template base out towards the

solvent-accessible major groove and stabilizing a mismatched G base via H-bonding. The finger domain of polt is responsible for the unique active site, and in turn, its replication specificity. This feature allows polt to maintain a relatively high fidelity on template purines, yet induce high rates of mis-incorporation on template pyrimidines. The high fidelity on template purines by polt appears to play a role in translesion synthesis by allowing accurate replication through adducted purine bases. The biological role of polt's low fidelity is still unclear; however, it is apparent that when functioning out of context, this unique replication specificity induces high rates of DNA mutagenesis.

3.4 Methods

3.4.1 Dpo4- ι chimeric proteins

To construct the Dpo4- ι finger domain chimera (Dpo4- ι) and the Dpo4- ι little finger domain chimera (Dpo4- ι -LF), plasmid vector pET-22b containing the Dpo4 gene and plasmid vector pHis-parrallel1 containing the polt gene were used as templates for PCR. For Dpo4- ι , the N-terminus of Dpo4 was cloned to the beginning of the finger domain using primers A: (5'-CGTTACTGCCATGGTTGTTCTTTTCGTTG-3') and B: (5'-TTCTACTTGTGCATAAAAGCAGTCAAAATCAACGAAAAGAACAAT-3'). The result was an N-terminus Dpo4 PCR product containing an NcoI cutting site at the N terminus and a C terminal overhang, which was complementary to the beginning of the polt finger domain. The C-terminus of Dpo4 was cloned past the end of the finger domain using primers C: (5'-GTTGGTATTAGTTAATGGAGAAGACAAGGAAGTATATCAGCAAGTTTC-3') and D: (5'-GCTAGTTATTGCTCAGC-3'). The result was a C-terminal Dpo4 PCR product with an N terminal overhang, complementary to the end of the polt finger domain. The finger domain of polt was cloned using primers E: (5'-TGC

TTTTATGCACAAGTAGAAATG-3') and F: (5'-GTCTTCTCCATTA ACTAATACC AAC-3'). The N-terminal Dpo4 product was joined with the polI finger domain product using primers A and F. The resulting N-terminal Dpo4-polI finger domain product was joined with the C-terminal Dpo4 product using primers A and D to produce the final product of a Dpo4 gene containing the finger domain of polI (Dpo4-t).

For Dpo4-t-LF, the N-terminus of Dpo4 was cloned up to the beginning of the little finger domain using primers A: (5'- C GTT ACT GCC ATG GTT GTT CTT TTC GTT GAT TTT GAC TAC TTT TAC GCT C -3') and B: (5'- AGT TCT TAT AGG CTC GTT ATA CTC GTC TCT AGC TAG A -3'). The result was an N-terminus Dpo4 PCR product containing an NcoI cutting site at the N terminus. The little finger domain of polI was cloned using primers C: (5'- GCC GTT ACT GCC ATG GTT GTT CTT TTC GTT GAT TTT GAC TAC TTT TAC GCT C -3') and D: (5'- CAT CCT CGA GAC CTA CTT AGC AGT ATT TAG TGC TTT AAG GTT GCA GAA GC-3'). The result was a polI little finger domain with an N-terminal overhang, complimentary to the end of the Dpo4 product. The N-terminal Dpo4 product was joined with the polI little finger domain product using primers A and D to produce the final product of a Dpo4 gene containing the little finger domain of polI (Dpo4-t-LF). Both Dpo4-t and Dpo4-t-LF genes were cloned into the pHis-parallel1 vector and confirmed by sequencing.

3.4.2 Primer Extension Assays

DNA substrate (10 nM) was incubated with either Dpo4, Dpo4-t, Dpo4-t-LF or hpolI (10 nM) and 100 μ M of either all four dNTPs or individual dNTPs at 37°C for 2min in reaction buffer containing 40 mM Tris (pH 8.0), 5 mM MgCl₂, 250 μ g/ml BSA, 10 mM DTT, and 2.5% glycerol. Reactions carried out in the presence of 150 mM CaCl₂ were

incubated at 37°C for 60min. Reactions were terminated with loading buffer (95% formamide, 20 mM EDTA, 0.025% xylene, 0.025% bromophenol blue) and resolved on a 20% polyacrylamide gel containing 7 M urea. Gels were visualized using a PhosphorImager.

3.4.3 Protein Preparation

Human DNA pol τ (amino acid 1-420) was cloned into pGST-parrallell vector and the subsequent glutathione S-transferase-tagged pol τ was over expressed in Escherichia coli strain DE3. The pol τ -GST fusion protein was purified using affinity chromatography and cleaved using a histidine-tagged tobacco etch virus (TEV) protease, which was subsequently removed using nickel affinity chromatography. The cleaved pol τ containing 2 extra N-terminal residues was further purified using an SP column. Dpo4 used for functional assays was purified as previously described (Ling et al., 2001). The His tagged Dpo4- τ chimeric proteins used for functional assays were overexpressed in Escherichia coli strain DE3 and purified using nickel affinity chromatography followed by an SP column

3.4.4 DNA preparation

Oligonucleotides for crystallization were purchased from Keck Oligo Inc. and gel purified. The 9-nt primer (5'-GTGGATGAG-3') was annealed to a 15-nt template (5'-CTCATTCTCATCCAC-3'), and the self-annealing 18-nt oligonucleotide (5'-TCATGGGTCCTAGGACCC^{dd}-3') was annealed with itself to give a DNA substrate with two replicative ends. Oligonucleotides used for primer extension assays were purchased from Sigma Aldrich and gel purified. A 30-nt template: (5'-GAGCAGTCGCACATGTAGTATCTCTGTGAC-3') was annealed to a 16-nt primer

(5'-GTCACAGAGATACTAC-3') resulting in a template T base at the first and eighth position beyond the primer-template junction. The primer was 5'-end labelled using [γ -³²P]ATP and T4 polynucleotide kinase. The 5'-labelled primer was mixed with template DNA at a 1.5:1 molar ratio and heated at 95°C, followed by slow cooling to form the annealed DNA substrate.

3.4.5 Crystallization and Structure Determination

Ternary complexes were formed for T:ddADP, T:dGTP, and T:dTTP by incubating protein (0.2 mM) and DNA in a 1:1.2 ratio with either ddNTP or dNTP (1mM), and MgCl₂ (5 mM). Crystals of the T:dGTP and T:dTTP complex were obtained in 12% PEG 5000 MME + 0.2M NH₄SO₄ + 5% glycerol + 0.1 M MES, pH 6.5, while crystals of the T:ddADP complex were obtained in 12% PEG 3350 + 0.15 M CaCl₂ + 0.01 M DTT + 5% glycerol. All crystals were flash frozen in liquid nitrogen using paratone-N as a cryo protectant. X-ray diffraction data were collected on the beamline 24-ID-C at the Advanced Photon Source in Argonne National Laboratory. The data were processed and scaled using HKL (Otwinowski and Minor, 1997).

All three structures were solved using molecular replacement with a previously solved ternary complex (PDB:2ALZ) as a search model. Rigid body refinement was performed using REFMAC (Murshudov et al., 1997), followed by restrained refinement and then TLS refinement. Electron density was well defined for all structures except for the first 27 residues of the N-terminus, loop regions 332-337, 350-356, and 371-378 and the last 6 residues of the C-terminus. Additionally, the +4 and +5 nucleotide within T:ddADP and the +3, and +2 nucleotides within T:dGTP and T:dTTP were disordered. All structures have good stereochemistry with over 95% of the residues in the most favoured region of

the Ramachandran plot. The atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 3GV5, 3GV7 and 3GV8 for structures T:ddADP, T:dTTP, and T:dGTP, respectively.

3.5 References

- Alt, A., Lammens, K., Chiocchini, C., Lammens, A., Pieck, J.C., Kuch, D., Hopfner, K.P. and Carell, T. (2007) Bypass of DNA lesions generated during anticancer treatment with cisplatin by DNA polymerase ϵ . *Science*, **318**, 967-970.
- Bauer, J., Xing, G., Yagi, H., Sayer, J.M., Jerina, D.M. and Ling, H. (2007) A structural gap in Dpo4 supports mutagenic bypass of a major benzo[a]pyrene dG adduct in DNA through template misalignment. *Proc Natl Acad Sci U S A*, **104**, 14905-14910.
- Boudsocq, F., Kokoska, R.J., Plosky, B.S., Vaisman, A., Ling, H., Kunkel, T.A., Yang, W. and Woodgate, R. (2004) Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication. *J Biol Chem*, **279**, 32932-32940.
- Dumstorf, C.A., Clark, A.B., Lin, Q., Kissling, G.E., Yuan, T., Kucherlapati, R., McGregor, W.G. and Kunkel, T.A. (2006) Participation of mouse DNA polymerase ι in strand-biased mutagenic bypass of UV photoproducts and suppression of skin cancer. *Proc Natl Acad Sci U S A*, **103**, 18083-18088.
- Frank, E.G. and Woodgate, R. (2007) Increased catalytic activity and altered fidelity of human DNA polymerase ι in the presence of manganese. *J Biol Chem*, **282**, 24689-24696.
- Johnson, R.E., Prakash, L. and Prakash, S. (2005) Biochemical evidence for the requirement of Hoogsteen base pairing for replication by human DNA polymerase ι . *Proc Natl Acad Sci U S A*, **102**, 10466-10471.
- Johnson, R.E., Washington, M.T., Haracska, L., Prakash, S. and Prakash, L. (2000) Eukaryotic polymerases ι and ζ act sequentially to bypass DNA lesions. *Nature*, **406**, 1015-1019.
- Kunkel, T.A., Pavlov, Y.I. and Bebenek, K. (2003) Functions of human DNA polymerases ϵ , κ and ι suggested by their properties, including fidelity with undamaged DNA templates. *DNA Repair (Amst)*, **2**, 135-149.
- Ling, H., Boudsocq, F., Plosky, B.S., Woodgate, R. and Yang, W. (2003) Replication of a *cys--syn* thymine dimer at atomic resolution. *Nature*, **424**, 1083-1087.

- Ling, H., Boudsocq, F., Woodgate, R. and Yang, W. (2001) Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. *Cell*, **107**, 91-102.
- Ling, H., Boudsocq, F., Woodgate, R. and Yang, W. (2004) Snapshots of replication through an abasic lesion; structural basis for base substitutions and frameshifts. *Mol Cell*, **13**, 751-762.
- Ling, H., Sayer, J.M., Plosky, B.S., Yagi, H., Boudsocq, F., Woodgate, R., Jerina, D.M. and Yang, W. (2004) Crystal structure of a benzo[a]pyrene diol epoxide adduct in a ternary complex with a DNA polymerase. *Proc Natl Acad Sci U S A*, **101**, 2265-2269.
- Lone, S., Townson, S.A., Uljon, S.N., Johnson, R.E., Brahma, A., Nair, D.T., Prakash, S., Prakash, L. and Aggarwal, A.K. (2007) Human DNA polymerase kappa encircles DNA: implications for mismatch extension and lesion bypass. *Mol Cell*, **25**, 601-614.
- Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*, **53**, 240-255.
- Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2005) Human DNA polymerase iota incorporates dCTP opposite template G via a G.C + Hoogsteen base pair. *Structure*, **13**, 1569-1577.
- Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2005) Rev1 employs a novel mechanism of DNA synthesis using a protein template. *Science*, **309**, 2219-2222.
- Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2006) Hoogsteen base pair formation promotes synthesis opposite the 1,N⁶-ethenodeoxyadenosine lesion by human DNA polymerase iota. *Nat Struct Mol Biol*, **13**, 619-625.
- Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2006) An incoming nucleotide imposes an anti to syn conformational change on the templating purine in the human DNA polymerase-iota active site. *Structure*, **14**, 749-755.
- Nair, D.T., Johnson, R.E., Prakash, S., Prakash, L. and Aggarwal, A.K. (2004) Replication by human DNA polymerase-iota occurs by Hoogsteen base-pairing. *Nature*, **430**, 377-380.
- Ohkumo, T., Kondo, Y., Yokoi, M., Tsukamoto, T., Yamada, A., Sugimoto, T., Kanao, R., Higashi, Y., Kondoh, H., Tatematsu, M., Masutani, C. and Hanaoka, F. (2006) UV-B radiation induces epithelial tumors in mice lacking DNA

- polymerase eta and mesenchymal tumors in mice deficient for DNA polymerase iota. *Mol Cell Biol*, **26**, 7696-7706.
- Otwinowski, Z. and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol*, **276**, 307-326.
- Pence, M.G., Blans, P., Zink, C.N., Hollis, T., Fishbein, J.C. and Perrino, F.W. (2008) Lesion bypass of N2-ethylguanine by human DNA polymerase Iota. *J Biol Chem*.
- Tissier, A., Frank, E.G., McDonald, J.P., Vaisman, A., Fernandez de Henestrosa, A.R., Boudsocq, F., McLenigan, M.P. and Woodgate, R. (2001) Biochemical characterization of human DNA polymerase iota provides clues to its biological function. *Biochem Soc Trans*, **29**, 183-187.
- Tissier, A., McDonald, J.P., Frank, E.G. and Woodgate, R. (2000) poliota, a remarkably error-prone human DNA polymerase. *Genes Dev*, **14**, 1642-1650.
- Vaisman, A., Ling, H., Woodgate, R. and Yang, W. (2005) Fidelity of Dpo4: effect of metal ions, nucleotide selection and pyrophosphorolysis. *EMBO J*, **24**, 2957-2967.
- Vaisman, A. and Woodgate, R. (2001) Unique misinsertion specificity of poliota may decrease the mutagenic potential of deaminated cytosines. *Embo J*, **20**, 6520-6529.
- Wang, Y., Woodgate, R., McManus, T.P., Mead, S., McCormick, J.J. and Maher, V.M. (2007) Evidence that in xeroderma pigmentosum variant cells, which lack DNA polymerase eta, DNA polymerase iota causes the very high frequency and unique spectrum of UV-induced mutations. *Cancer Res*, **67**, 3018-3026.
- Washington, M.T., Johnson, R.E., Prakash, L. and Prakash, S. (2004) Human DNA polymerase iota utilizes different nucleotide incorporation mechanisms dependent upon the template base. *Mol Cell Biol*, **24**, 936-943.
- Wong, J.H., Fiala, K.A., Suo, Z. and Ling, H. (2008) Snapshots of a Y-family DNA polymerase in replication: substrate-induced conformational transitions and implications for fidelity of Dpo4. *J Mol Biol*, **379**, 317-330.
- Yang, W. (2003) Damage repair DNA polymerases Y. *Curr Opin Struct Biol*, **13**, 23-30.
- Yang, W. (2006) Poor base stacking at DNA lesions may initiate recognition by many repair proteins. *DNA Repair (Amst)*, **5**, 654-666.
- Zhang, Y., Yuan, F., Wu, X. and Wang, Z. (2000) Preferential incorporation of G opposite template T by the low-fidelity human DNA polymerase iota. *Mol Cell Biol*, **20**, 7099-7108.

Chapter 4¹

4 A unique active site promotes error-free replication opposite an 8-oxo-guanine lesion by human DNA polymerase ι

4.1 Introduction

The existence of all animals on our planet depends on oxygen. However, this essential molecule also represents a toxic precursor and a potent mutagen due to its conversion to oxygen radicals by aerobic respiration. Abundant oxygen radicals threaten the well being of organisms, due to their high reactivity with many biological compounds, in particular, DNA. One of the most abundant and mutagenic oxidative DNA lesions, 7,8-dihydro-8-oxo-guanine (8-oxo-G), forms when an oxygen radical covalently attaches to the C8 atom of a guanine base (Figure 4.1A). Approximately 1000 of these oxidative DNA lesions arise daily in every cell of our body (Collins, 1999).

The 8-oxo-G lesion is highly mutagenic due to its dual coding properties and the ability of high fidelity replicative polymerases to replicate through the lesion. Structural studies demonstrate that the 8-oxo-G lesion can adopt two alternate conformations (*anti* or *syn*) in the active site of DNA polymerases (Briebe et al., 2004; Hsu et al., 2004; Rechkoblit et al., 2009; Vasquez-Del Carpio et al., 2009). The *anti* conformation allows correct base pairing with an incoming cytosine (C) nucleotide, while the *syn* conformation forms a stable mis-pairing with an incoming adenine (A) nucleotide in a normal *anti* conformation (Figure 4.1A). These alternate conformations of 8-oxo-G promote high A mis-incorporation rates for most DNA polymerases (Shibutani et al., 1991).

¹ The contents of this chapter have been submitted for publication.

Consequently, the 8-oxo-G lesion can induce a high frequency of guanine (G) to thymine (T) transversions. Although the mutagenic potential of 8-oxo-G has been well characterized at the structural level, the structural basis of preferential correct C incorporation by eukaryotic DNA polymerases still remains elusive. Given the sheer volume and the dual coding nature of 8-oxo-G, this lesion represents a major hurdle for cells to overcome in maintaining genome integrity.

Organisms evolved repair mechanisms to deal with the abundant 8-oxo-G:A mis-pairs produced in the genome. In the Base Excision Repair (BER) pathway, an adenine glycosylase (MUTY) removes the mis-matched A base, and the subsequent gap can be filled by a specialized polymerase that strongly prefers the correct C incorporation opposite 8-oxo-G (Takao et al., 1999). Although the identity of this DNA polymerase remains unknown, Y family DNA polymerases, which specialize in replicating through DNA lesions, likely play a role.

The human Y family DNA polymerase iota (*pol* ι) appears to play an important role in cellular protection against oxidative stress. *Pol* ι is recruited to chromatin after cells are exposed to oxidative DNA damage, and down regulating *pol* ι greatly increases the sensitivity of cells to oxidizing agents (Petta et al., 2008). In addition *pol* ι displays BER activity *in vitro* and *in vivo* and functionally interacts with the BER scaffold protein XRCC1 (Bebenek et al., 2001; Petta et al., 2008; Prasad et al., 2003). Furthermore, *pol* ι displays 5' deoxyribose phosphate (dRP) lyase activity, a characteristic of BER polymerases (Bebenek et al., 2001). Lastly, *pol* ι is one of only a few DNA polymerases that preferentially incorporate the correct dC nucleotide opposite 8-oxo-G (Zhang et al., 2001). Other human Y family polymerases, such as *pol* η and *pol* κ , have dA mis-

incorporation rates of around 45% to 60% respectively (McCulloch et al., 2009; Zhang et al., 2000).

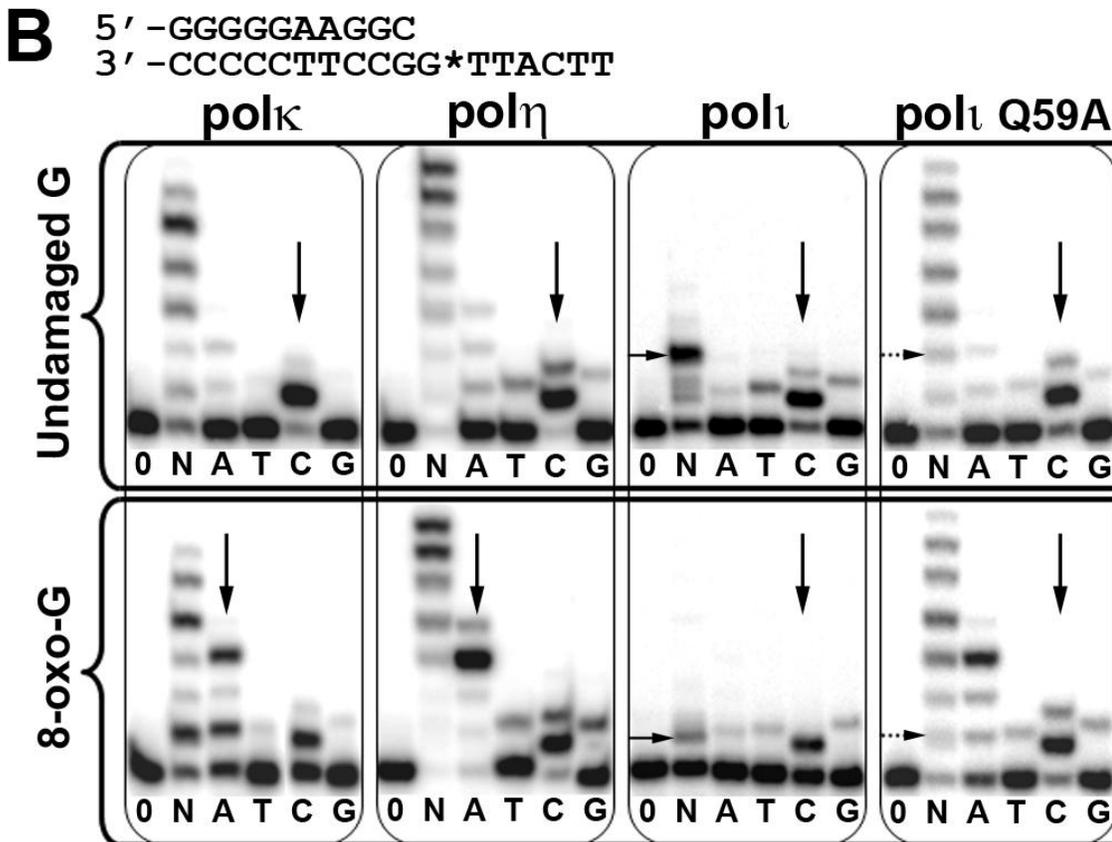
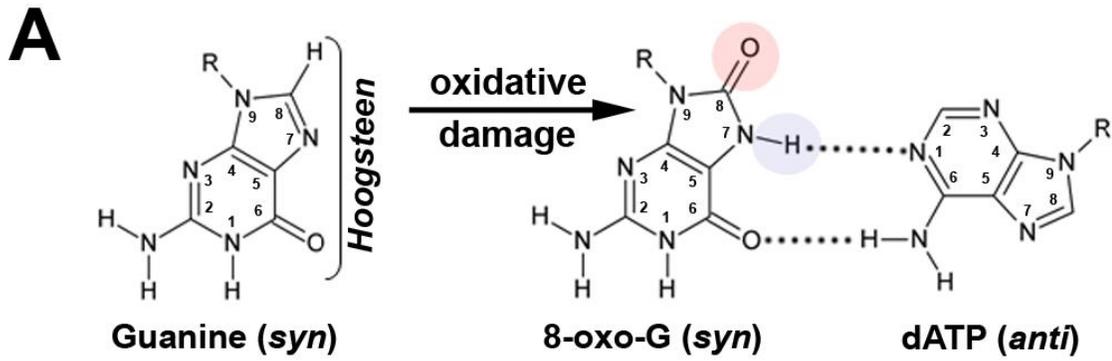
In the present study, we report four crystal structures of polt in complex with DNA containing an 8-oxo-G lesion at the template position, paired with either correct dCTP nucleotide, or incorrect dATP, dTTP and dGTP nucleotides. Our results demonstrate, for the first time, how a specialized eukaryotic DNA polymerase, polt, preferentially incorporates the correct C nucleotide opposite an 8-oxo-G lesion, with high specificity over the mis-matched A, T, and G nucleotides.

4.2 Results and discussion

4.2.1 8-oxo-guanine replication fidelity

Primer extension assays were carried out with three different human Y-family polymerases (polk, polη, and polt) to assess their 8-oxo-G bypass activity and fidelity (Figure 4.1B). Opposite undamaged G, all three polymerases insert the correct C nucleotide with high specificity (Figure 4.1B, vertical arrows). In the presence of 8-oxo-G, only polt maintains a C incorporation preference over the mis-matched nucleotides. Both polk and polη demonstrate equal or greater A mis-incorporation rates over C opposite 8-oxo-G with little to no decrease in replication efficiency compared to undamaged DNA (Figure 4.1B). This fits well with the previously reported 8-oxo-G polymerase structures that have the O⁸ atom residing in a non clashing environment, which has minimal effects on reaction efficiency (Briebe et al., 2004; Hsu et al., 2004; Rechkoblit et al., 2009; Rechkoblit et al., 2006; Vasquez-Del Carpio et al., 2009; Zang et al., 2006). In contrast, polt displays dramatically reduced replication efficiency opposite

Figure 4.1 8-oxo-guanine and nucleotide specificity of Y family DNA polymerases. **(A)** Structural changes of guanine to 8-oxo-guanine (8-oxo-G). Bracket emphasizes the Hoogsteen edge and colored circles represent sites of modification on 8-oxo-guanine. The mismatched OG:dATP base pair is shown on the right. **(B)** Primer extension assays show differences in nucleotide incorporation by pol κ , pol η , pol ι , and pol ι Q59A mutant for undamaged G (top) and 8-oxo-G (bottom) bases. The DNA substrate used for replication assays is shown at the top with G* representing the site of modification. Vertical black arrows indicate nucleotide insertion preference, while horizontal black arrows indicate replication stalling. Horizontal dashed arrows indicate the loss of stalling. Enzymes were incubated with DNA in the absence of nucleotides (0), presence of all four nucleotides (N), or individual nucleotides (A, T, C, G).



8-oxo-G in spite of preferential dC incorporation (Figure 4.1B). The replication assay results are consistent with previously published 8-oxo-G fidelities for polt, polk and polη (McCulloch et al., 2009; Zhang et al., 2001; Zhang et al., 2000) and indicate that polt is a specialized human Y family polymerase that replicates the 8-oxo-G lesion in an error-free manner.

4.2.2 Polt/8-oxo-G/dNTP ternary complexes

In order to elucidate how polt selects the correct C nucleotide opposite 8-oxo-G, we sought to crystallize polt in ternary complex with the 8-oxo-G lesion. The DNA substrate used for crystallization was a self-annealing oligonucleotide, forming a double stranded DNA substrate with two primer-template junctions. Both junctions contain the 8-oxo-G lesion at the first template base position ready for dNTP incorporation. The DNA substrate also contains dideoxy 3' ends in order to inhibit DNA polymerization and thus trap ternary complexes for crystallization. The DNA oligo sequence and modification were confirmed by mass spectrometry. The DNA substrate was incubated with polt and co-crystallized with each incoming nucleotide (dCTP, dATP, dTTP, dGTP) separately. The resulting four crystal structures are denoted as OG:dCTP, OG:dATP, OG:dTTP, and OG:dGTP, according to the incoming nucleotide against the lesion in the active site. All four structures have the same crystal form (space group, P6₅22) and diffract to 1.95 Å, 2.03 Å, 2.09 Å and 2.45 Å resolutions, respectively (Table 4.1).

Polt in all four 8-oxo-G complexes is virtually identical, with root-mean-square deviations (r.m.s.d) within 0.2 Å between all C α atoms. In addition, the C α r.m.s.d. between the current 8-oxo-G structures and a previously solved polt ternary structure with undamaged G (PDB: 2ALZ) (Nair et al., 2005) is within 0.3 Å. This result indicates

Table 4.1 Summary of crystallographic data for polt 8-oxo-guanine ternary structures

Data collection	OG:dCTP	OG:dATP	OG:dGTP	OG:dTTP
Space group	P6 ₅ 22	P6 ₅ 22	P6 ₅ 22	P6 ₅ 22
Mol/AU ^a	1	1	1	1
Unit Cell				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	99.2, 99.2, 202.3	98.7, 98.7, 202.3	98.3, 98.3, 202.6	97.5, 97.5, 201.7
<i>α</i> , <i>β</i> , <i>γ</i> (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Reflections				
Total	367072	430405	124779	193066
Unique	43185	38429	21891	34476
Resolution (Å) ^b	50.0-1.95 (1.98-1.95)	50.0-2.03 (2.07-2.03)	50.0-2.45 (2.49-2.45)	50.0-2.09 (2.13-2.09)
Completeness (%)	99.1 (95.6)	99.8 (100)	99.0 (100)	99.7 (99.7)
<i>I</i> / <i>σ</i> <i>I</i>	25.5 (2.3)	37.9 (2.9)	27.6 (2.0)	21.9 (2.0)
<i>R</i> _{merge} ^b	7.5 (50.4)	7.4 (59.3)	6.1 (57.8)	6.9 (58.0)
Refinement statistics				
<i>R</i> _{work} / <i>R</i> _{free}	22.5/25.7	22.7/25.3	22.5/28.0	22.7/26.6
No. atoms				
Protein	2983	2979	2857	2962
DNA	332	332	332	332
dNTP	28	30	31	29
Ions ^c	2	2	2	2
Water	398	297	151	256
Avg B factor				
Protein	51.4	57.1	63.0	54.5
DNA	35.6	41.1	52.7	36.8
dNTP	35.7	47.6	56.7	33.2
Ions	24.9	35.7	52.6	24.8
Water	46.8	50.2	52.5	52.5
R.m.s. deviations				
Bonds (Å)	0.006	0.008	0.006	0.006
Angles (°)	1.27	1.34	1.26	1.19
^a Mol/AU represents the number of molecules per asymmetric unit.				
^b Data in the highest resolution shell are in parentheses				
^c One non-catalytic Mg ²⁺ ion exists in all structures				

that polt does not undergo significant conformational changes when replicating through the 8-oxo-G lesion. The significant difference between the current four structures lies within the replicating base pairs. The incoming nucleotides in all four ternary structures are in an active position with the α phosphates within 4 Å to the putative 3'-end of the primer strands. One active site Mg^{2+} ion is present in the standard B site positions of the active sites, while the ion density is not well defined for the A sites. Lack of ion coordination at the A site has been previously observed for polt due to ion mobility at this location (Ling et al., 2001; Nair et al., 2005).

4.2.3 8-oxo-G template base positioning

The 8-oxo-G template base is oriented in a *syn* conformation opposite the incoming nucleotides in all four structures (Figure 4.2A). The *syn* conformation of 8-oxo-G presents the Hoogsteen edge exclusively for hydrogen bonding. Polt induces *syn* conformations on template purines due to an exceptionally narrow active site, which restricts the C1'-C1' distance to under 9Å (Kirouac and Ling, 2009; Nair et al., 2004). Accordingly, the C1'-C1' distance of all four replicating base pairs in the current 8-oxo-G structures is less than 9Å (Figure 4.2A). Compared to undamaged dG in the polt active site, 8-oxo-G is pushed out towards the solvent exposed major groove by ~1Å and tilted ~30° off the base stacking plane due to its O⁸ atom clashing with the OE1 atom of Gln59 from the finger domain (2.9 - 3.1 Å) (Figure 4.2B). Consequently, the 8-oxo-G base is shifted away from its optimal hydrogen bonding position with the incoming nucleotides and further crowds the already narrow active site (Figure 4.2B). Interestingly, the position of the 8-oxo-G base is almost identical to that of undamaged T within the polt active site (PDB: 3GV7). Undamaged T also shifts into the major groove with tilting due

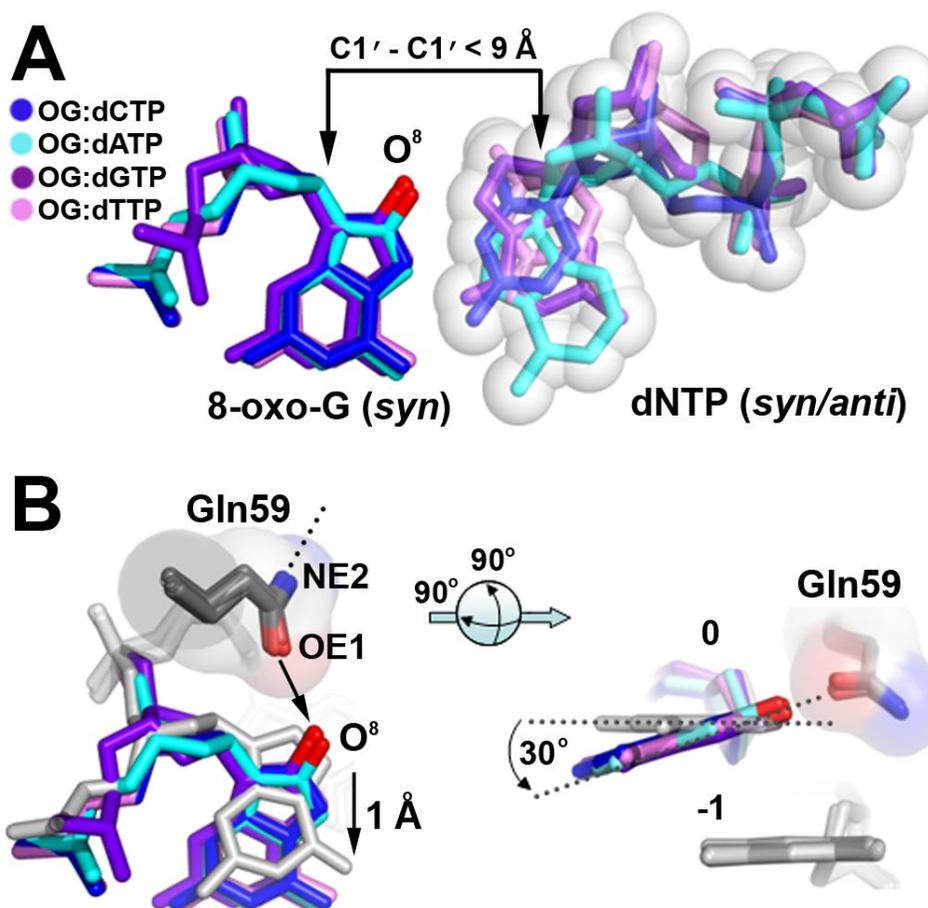


Figure 4.2 8-oxo-G:dNTP positioning in the PolI/DNA/nucleotide structures. **(A)** Superposition of replicating base pairs from all four polI/8-oxo-G/nucleotide structures: OG:dCTP (blue), OG:dATP (cyan), OG:dGTP (purple), and OG:dTTP (pink). The oxygen modification on 8-oxo-guanine is coloured red and the C1'-C1' distances are indicated with black arrows. Sphere representation is shown for the incoming nucleotides. **(B)** Positioning (left) and tilting (right) of the 8-oxo-G bases by Gln59. The 8-oxo-G bases are superimposed with a previous template G polI structure (PDB:2ALZ, grey) with a black arrow indicating repulsion between the OE1 and O⁸ atoms and a black dashed line on the NE2 atom indicating H-bonding. The template bases are numbered 0 and the underlying bases are numbered -1. Dashed lines indicate the planes of the guanine bases.

to its exocyclic oxygen group (O^2) approaching Gln59 in a similar manner to that of the O^8 atom in 8-oxo-G (Kirouac and Ling, 2009). Thus, Gln59 appears to sterically position template bases containing exocyclic groups approaching its OE1 atom. Template G or A bases are not tilted in polt (Nair et al., 2006), as these bases lack exocyclic oxygen atoms that approach Gln59. The specific positioning of the 8-oxo-G base by Gln59 appears to contribute to incoming nucleotide selection by influencing hydrogen bonding potential. Indeed, Q59A mutation increases polt's dATP mis-incorporation opposite 8-oxo-G.

Noticeably, the electron density does not cover the 8-oxo-G base as well as observed in other polymerase structures (Briebe et al., 2004; Hsu et al., 2004; Rechkoblit et al., 2009; Rechkoblit et al., 2006; Vasquez-Del Carpio et al., 2009) (Figure 4.3). The poor density likely results from the crowded active site in which the *syn* lesion base is squeezed by Gln59 and the incoming dNTP (Figure 4.2). Deformation of the density indicates certain structural mobility of the lesion base which has higher B-factors than neighbouring bases. The density appears slightly shrunken along the latitudinal dimension of the purine base and elongated through the longitudinal dimension, according to the *syn* orientation (Figure 4.3, left panels), which connects the electron density between the lesion base and the incoming nucleotides. Alternate conformations of *anti/syn* 8-oxo-G would account for this deformation of electron density. To reflect this observation, we modeled the majority of the lesion bases (~80%) in the *syn* conformation and ~20% in an *anti* conformation, which satisfies the electron density better than *syn* conformations alone (Figure 4.3). However, the *anti* conformation excludes stable binding of incoming dNTPs (Figure 4.3, right panels) and thus forms non-productive complexes. This structural observation is consistent with the reduced efficiency observed for polt opposite

Figure 4.3 Replicating base pairs in pol1/8-oxo-G structures. (A) OG:dCTP (B) OG:dATP (C) OG:dGTP (D) OG:dTTP. For the *syn* 8-oxo-G structures (left hand side), double sided arrows indicate steric repulsion from Gln59 and between crowded bases, and dashed lines indicate hydrogen bonding with bonding and repulsion distances indicated. Divalent metal ions (green) and water molecules (red) are shown as spheres. For the *anti* conformation structures (right hand side), the replicating base pairs of the *syn* forms are shown in light grey as reference. Black arrows indicate potential steric clashing between the incoming nucleotides and 8-oxo-guanine in an *anti* conformation. The replicating base pairs are shown with $2F_o - F_c$ electron density maps contoured at 1.0σ level.

8-oxo-G compared to undamaged G (Zhang et al., 2001) (Figure 4.1B). The Gln59 clashing with the extra O8 atom of 8-oxo-G makes the *syn* conformation less stable than that of the undamaged dG base in polI and induces the non-productive 8-oxo-G *anti* conformation. Accordingly, it has been predicted by molecular simulation work that large major-groove lesions prefer the *anti* conformation within the polI active site (Donny-Clark et al., 2009). However, the 8-oxo-G lesion is the smallest major-groove adduct with only one atom attachment (O⁸). Thus, the 8-oxo-G *syn* conformation can still be accommodated within the polI active site with only minor disturbances in stability, which explains why only a small portion of *anti* conformations are observed in our structures.

4.2.4 8-oxo-G:dNTP replicating base pairs

The nucleotide bases approach the *syn* 8-oxo-G base in different conformations (Figure 4.3). The smaller pyrimidine dCTP and dTTP nucleotides are in *anti* conformations, while the larger purine based dATP and dGTP are in *syn* conformations in order to fit in the narrow active site of polI. In the OG:dCTP structure, the *anti* dCTP base forms the strongest hydrogen bonding network with the 8-oxo-G Hoogsteen edge among all four replication base pairs (Figure 4.3A). Two bonds occur between the N3 and N⁴ atoms of dCTP and the N7 and O⁶ atoms of the 8-oxo-G base. In addition to hydrogen bonding, the small pyrimidine base of dCTP opposite 8-oxo-G conforms to the narrow polI active site. In contrast, the *syn* dATP base forms only one hydrogen bond with the 8-oxo-G Hoogsteen edge in the OG:dATP structure (Figure 4.3B). Thus the dATP nucleotide would be less favorable than dCTP in terms of H-bonding potential. Furthermore, *syn*

purine bases are energetically less favored than the *anti* conformations in solution (Salter et al., 2003; Stolarski et al., 1984). Thus, the narrow polI active site forces dATP to adopt an energetically unfavorable conformation, which also increases the entropy cost of incorporation. These effects together enable polI to discriminate against purine-purine mis-matched base pairs and favors a less strained purine-pyrimidine pair.

In the OG:dTTP structure, dTTP has loose H-bonding to the lesion template with one weak potential H-bond (3.2 Å) between its N3 atom and the 8-oxo-G O⁸ atom (Figure 4.3C). The dGTP base is in unfavourable *syn* conformation and forms one hydrogen bond between its O⁶ atom and the N7 atom of the 8-oxo-G in the OG:dGTP structure (Figure 4.3D). Thus, both dTTP and dGTP nucleotides would be less favourable than dCTP for replication opposite 8-oxo-G by polI. Overall, dCTP is the most favourable incoming nucleotide opposite 8-oxo-G in terms of nucleotide conformation and H-bonding to the *syn* lesion base. The structural analysis fits the biochemical observations that dCTP is preferentially incorporated opposite 8-oxo-G by polI.

4.2.5 Glutamine 59 and polI 8-oxo-G activity

From our structural observations we postulated that Gln59 from the polI finger domain would influence base pair selection and bypass efficiency opposite 8-oxo-G by clashing with the O8 atom. To test this hypothesis, we mutated polI Gln59 to Ala (Q59A mutant) and assayed the enzyme for incorporation specificity and activity. The Q59A mutant increases dATP mis-incorporation against 8-oxo-G compared to wild type polI (Figure 4.1B). The increased dA insertion by the Q59A mutation shifts the 8-oxo-G fidelity of

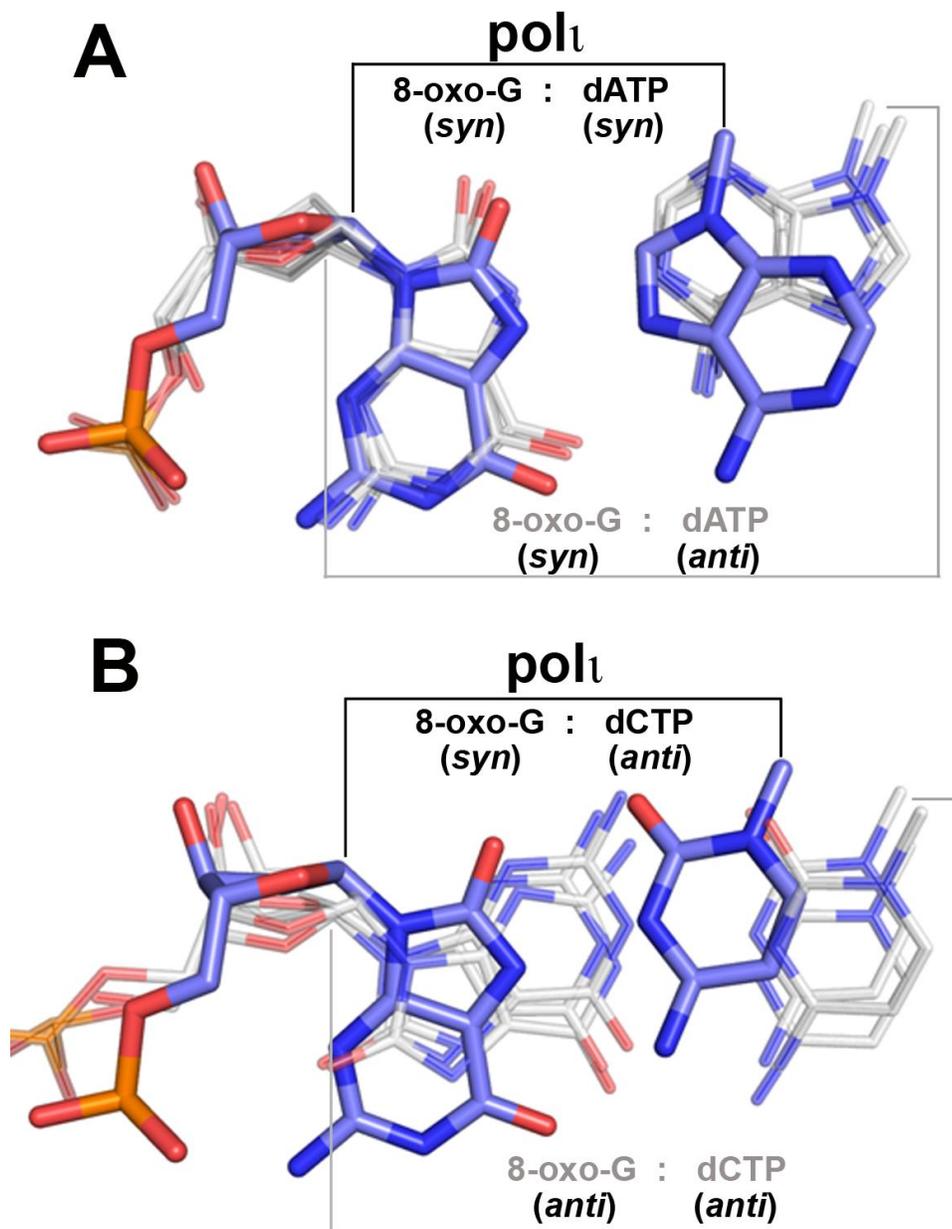
polI towards that of pol η and pol κ . In addition, the reaction efficiency opposite 8-oxo-G has dramatically increased for the mutant. This result is consistent with our structural observations, which demonstrate the Q59 residue in polI reduces productive 8-oxo-G ternary complexes. It is likely that the Gln59 mutation creates a less restrictive polI active site, which alleviates the 8-oxo-G steric clash in the active site. This would reduce non-productive 8-oxo-G *anti* conformations and possibly allow the accommodation of a mis-matched *anti* dATP nucleotide.

Opposite the undamaged and lesion templates, polI displays replication inhibition as shown by stalling bands in the presence of all four nucleotides (horizontal black arrows) (Figure 4.1B). The undamaged DNA stalling is the result of a T base in the template strand which inhibits polI replication (Kirouac and Ling, 2009). Interestingly, the stalling bands disappear in the Q59A mutant assays for both undamaged and damaged DNA substrates (horizontal dashed arrows) (Figure 4.1B), indicating Gln59 contributes to stalling opposite 8-oxo-G similar to T bases. The stalling at 8-oxo-G lesions implies that polI may require other DNA polymerases to extend replication beyond the lesion site in a process known as multiple polymerase translesion replication (Livneh et al.).

4.2.6 Structural comparison of 8-oxo-G replication

It has been demonstrated that active site flexibility around the template base allows conformational alteration of 8-oxo-G in DNA polymerases with standard active site dimensions (Beard et al., 2010). This template flexibility allows the *syn* 8-oxo-G base to form a potent base pair with mis-matched *anti* dATP, causing A mis-incorporations by

Figure 4.4 Structural comparison of 8-oxo-guanine replication. **(A)** Superposition of the 8-oxo-G:dA base pair within pol ι (OG:dATP, blue) and other DNA polymerases with *syn* 8-oxo-G (grey). For clarity, only the base atoms were shown for the dA nucleotides. From left to right, the grey base pairs correspond to pol kappa (PDB: 3IN5), BF pol I (PDB: 1U49), and T7 pol (PDB: 1TK8). The narrow pol ι C1'-C1' distance is indicated with black lines on top, while the standard C1'-C1' distance of the other polymerases is indicated with grey lines on bottom. **(B)** Superposition of the 8-oxo-G:dC base pair within pol ι (OG:dCTP, blue) and other DNA polymerases with *anti* 8-oxo-G (grey). From left to right, the grey base pairs correspond to Dpo4 (PDB: 2C2E), BF pol I (PDB: 1U47), and T7 pol (PDB: 1TKD). The narrow pol ι C1'-C1' distance is indicated with black lines on top, while the standard C1'-C1' distance of the other polymerases is indicated with grey lines on bottom.



virtually all classes of DNA polymerases (Figure 4.4A) (Hsu et al., 2004; Krahn et al., 2003; Vasquez-Del Carpio et al., 2009). Such mutagenic base pairing has been observed for *Bacillus stearothermophilus* fragment (BF polI) and T7 polymerase (T7) from the high fidelity A family (Briebe et al., 2004; Hsu et al., 2004), pol β from the X-family (Batra et al., 2010; Krahn et al., 2003), and pol κ and Dpo4 from the Y-family (Vasquez-Del Carpio et al., 2009; Zang et al., 2006) (Figure 4.4A). However, for polt this mutagenic *syn*-OG:anti-dATP mis-match is too wide to be accommodated in its narrow active site (Figure 4.4A). Consequently, polt induces a dATP *syn* conformation, which dramatically reduces the stability of the OG:dATP mis-matched base pair (figure 4.4A). Thus, polt appears to be the only known DNA polymerase that can replicate the 8-oxo-G base in a *syn* conformation without inducing A mis-incorporations.

Interestingly, it has been shown that the 8-oxo-G *syn* conformation is more favourable than the *anti* conformation in high fidelity polymerases due to the O8 atom of *anti* 8-oxo-G clashing with the DNA backbone phosphates (Briebe et al., 2004; Hsu et al., 2004). This results in a decreased stability for the correct *anti*-OG:*anti*-dCTP Watson-Crick base pair. In the polt active site, the 8-oxo-G base is also induced into a *syn* conformation from clashing interactions. However, due to the crowded nature of the polt active site, the *anti* 8-oxo-G clashes with the incoming nucleotides instead of the DNA backbone. Although the *syn* 8-oxo-G conformation appears more favourable than *anti*, most DNA polymerases can replicate the lesion in both *anti* and *syn* orientations due to standard C1'-C1' distances (Figure 4.4B). In addition, some polymerases are able to alter the *syn/anti* equilibrium by hydrogen bonding with O8 atom of *anti* 8-oxo-G (Eoff et al., 2007). In contrast, polt replicates the 8-oxo-G lesion exclusively using the mutagenic

Hoogsteen edge (Figure 4.4). Remarkably, polt is able to achieve high fidelity 8-oxo-G replication opposite the Hoogsteen edge by creating a unique base pairing environment, which destabilizes incoming A nucleotides.

4.2.7 Implications for polt in the repair of oxidative DNA damage

Polt has a strong preference for inserting C against the 8-oxo-G lesion compared to other polymerases, which have A mis-incorporation preference (Hsu et al., 2004). However, polt may also be specialized for functioning over a broad range of oxidative damage rather than being solely specific for 8-oxo-G. Indeed, other BER DNA polymerases such as pol λ have also been implicated in accurate 8-oxo-G replication in vivo (van Loon and Hubscher, 2009). Interestingly, polt is also associated with another highly abundant oxidative lesion resulting from the oxidative deamination of cytosine: 5-hydroxy-uracil (5-OH-U). Unlike most DNA polymerases that incorporate A opposite the 5-OHU base, polt incorporates a G nucleotide (Vaisman and Woodgate, 2001). Polt also preferentially inserts G opposite template T, which structurally resembles the 5-OH-U lesion (Kirouac and Ling, 2009; Tissier et al., 2000; Zhang et al., 2000). Thus, similar to 8-oxo-G replication, polt could restore the correct incorporation opposite cytosine lesions during DNA replication or oxidative DNA damage repair.

During DNA replication a high frequency of 8-oxo-G bases will be mis-matched with A nucleotides by high fidelity replicative polymerases. These mutagenic base pairs must be repaired using a specialized form of BER, where excision of the dA mis-match occurs, followed by error-free TLS opposite the 8-oxo-G base. This process allows restoration of

the original genomic sequence. Polt likely contributes to this specialized TLS/BER pathway (Petta et al., 2008), which greatly reduces the mutational burden of 8-oxo-G:A mis-matches. Once the correct sequence has been restored, the 8-oxo-G base can be repaired using the standard BER pathway. In this manner, polt could greatly reduce mutagenesis induced by oxidative damage, which suggests possible roles for this enzyme in cellular protection against oxidizing agents.

4.2.8 Conclusions

This work describes the first structural mechanism of preferential C incorporation opposite the highly mutagenic 8-oxo-G lesion by a eukaryotic DNA polymerase. There are two unique features of the polt active site which allow high fidelity 8-oxo-G replication: 1) the narrow size fixes purine bases in *syn* conformations, which destabilizes mis-matched A and favors pyrimidine C; 2) the conserved Gln59 residue from the finger domain positions the *syn*8-oxo-G base, which optimizes H-bonding for the dCTP nucleotide. The unique polt active site may be specialized for accommodating a variety of mutagenic oxidative lesions, and is likely involved in restoring the genomic sequence of mis-matched oxidative base pairs via a specialized BER/TLS pathway.

4.3 Methods

4.3.1 Protein Preparation

Polt used for crystallization (amino acids 1-420) was purified as previously described (Kirouac and Ling, 2009). Proteins used for replication assays (polt 1-430, pol η 1-445, polk 19-523, and poltQ59A 1-430) were cloned from full length cDNA into the “His-

parallel" vector (Sheffield et al., 1999), overexpressed in BL21(DE3) cells and purified by nickel affinity and ion exchange chromatography. The Q59A point mutation within pol ι (1-430) was created by Quickchange mutagenesis using primers A (5'-GACAAACCTTTAGGGGTTCAAGCTAAATATTTGGTGGTTACCTGTAAC-3') and B (5'-GTTGCAGGTAACCACCAAATATTTAGCTTGAACCCCTAAAGGTTTGTC-3'). The resulting vector was sequenced to confirm the mutation.

4.3.2 DNA Preparation

The 8-oxo-G substrate used for crystallization was purchased from Keck Oligo Inc. The purified oligonucleotide containing a dideoxy 3' cytosine (C^{dd}) and an 8-oxo-G (G*) lesion: (5'-TCAG*GGGTCCTAGGACCC^{dd}-3') was confirmed by Mass Spectrometry and annealed with itself to give a DNA substrate with two replicative ends. Oligonucleotides used for primer extension assays were purchased from Sigma Aldrich. A 12-nt primer (5'-GGGGGAAGGAC-3') was annealed to either an 18-nt 8-oxo-G (G*) template (5'-TTCATTG*GTCCTTCCCC-3') or an undamaged 18-nt G template (5'-TTCATTGGTCCTTCCCC-3'). The primer was 5'-end labelled using [γ -³²P]ATP and T4 polynucleotide kinase.

4.3.3 Primer Extension Assays

DNA substrates (10 nM) were incubated with pol ι , pol η , pol κ , or pol ι Q59A (10 nM) and 100 μ M of all four dNTPs or individual dNTPs at 37°C for 2 min. The pol ι reactions were carried out with 80 μ M dNTP. The reaction buffer contained 40 mM Tris (pH 8.0),

5 mM MgCl₂, 250 ug/ml BSA, 10 mM DTT, and 2.5% glycerol. Reactions were terminated with loading buffer (95% formamide, 20 mM EDTA, 0.025% xylene, 0.025% bromophenol blue) and resolved on a 20% polyacrylamide gel containing 7 M urea. Gels were visualized using a PhosphorImager.

4.3.4 Crystallization and Structure Determination

Ternary complexes were formed for OG:dATP, OG:dCTP, OG:dTTP, and OG:dGTP by incubating protein (0.2 mM) and DNA in a 1:1.2 ratio with dNTP (2mM), and MgCl₂ (5 mM). Crystals of all four complexes were obtained in 15% PEG 5000 MME, 0.2M NH₄SO₄, 2.5% glycerol, and 0.1 M MES, pH 6.5. Extensive seeding was performed to increase the diffraction quality of the crystals. Crystals were flash frozen in liquid nitrogen directly from dehydrated crystallization drops to prevent crystal cracking. X-ray diffraction data were collected on beamlines 24-ID-C and 24-ID-E at the Advanced Photon Source in Argonne National Laboratory. The data were processed and scaled using HKL2000 (Otwinowski and Minor, 1997). All four structures were solved by molecular replacement using PHASER (McCoy et al., 2005) with a previously solved ternary complex (PDB: 3GV5) as a search model. Structural refinement was performed using PHENIX (Adams et al., 2010), starting with rigid body refinement, followed by positional and B-factor refinement, and lastly TLS refinement (Painter and Merritt, 2006). Model building and inspection were done using the program COOT (Emsley and Cowtan, 2004). All structures have good stereochemistry with over 95% of the residues in the most favoured region of the Ramachandran plot. Figures were created using PYMOL (DeLano, 2002).

4.4 References

- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C. and Zwart, P.H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*, **66**, 213-221.
- Alt, A., Lammens, K., Chiocchini, C., Lammens, A., Pieck, J.C., Kuch, D., Hopfner, K.P. and Carell, T. (2007) Bypass of DNA lesions generated during anticancer treatment with cisplatin by DNA polymerase ϵ . *Science*, **318**, 967-970.
- Batra, V.K., Beard, W.A., Hou, E.W., Pedersen, L.C., Prasad, R. and Wilson, S.H. (2010) Mutagenic conformation of 8-oxo-7,8-dihydro-2'-dGTP in the confines of a DNA polymerase active site. *Nat Struct Mol Biol*, **17**, 889-890.
- Beard, W.A., Batra, V.K. and Wilson, S.H. (2010) DNA polymerase structure-based insight on the mutagenic properties of 8-oxoguanine. *Mutat Res*.
- Bebenek, K., Tissier, A., Frank, E.G., McDonald, J.P., Prasad, R., Wilson, S.H., Woodgate, R. and Kunkel, T.A. (2001) 5'-Deoxyribose phosphate lyase activity of human DNA polymerase ι in vitro. *Science*, **291**, 2156-2159.
- Briebe, L.G., Eichman, B.F., Kokoska, R.J., Doublié, S., Kunkel, T.A. and Ellenberger, T. (2004) Structural basis for the dual coding potential of 8-oxoguanosine by a high-fidelity DNA polymerase. *Embo J*, **23**, 3452-3461.
- Collins, A.R. (1999) Oxidative DNA damage, antioxidants, and cancer. *Bioessays*, **21**, 238-246.
- DeLano, W.L. (2002) *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA, USA.
- Donny-Clark, K., Shapiro, R. and Broyde, S. (2009) Accommodation of an N-(deoxyguanosin-8-yl)-2-acetylaminofluorene adduct in the active site of human DNA polymerase ι : Hoogsteen or Watson-Crick base pairing? *Biochemistry*, **48**, 7-18.
- Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60**, 2126-2132.
- Eoff, R.L., Irimia, A., Angel, K.C., Egli, M. and Guengerich, F.P. (2007) Hydrogen bonding of 7,8-dihydro-8-oxodeoxyguanosine with a charged residue in the little finger domain determines miscoding events in *Sulfolobus solfataricus* DNA polymerase Dpo4. *J Biol Chem*, **282**, 19831-19843.

- Hsu, G.W., Ober, M., Carell, T. and Beese, L.S. (2004) Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase. *Nature*, **431**, 217-221.
- Kirouac, K.N. and Ling, H. (2009) Structural basis of error-prone replication and stalling at a thymine base by human DNA polymerase iota. *Embo J*, **28**, 1644-1654.
- Krahn, J.M., Beard, W.A., Miller, H., Grollman, A.P. and Wilson, S.H. (2003) Structure of DNA polymerase beta with the mutagenic DNA lesion 8-oxodeoxyguanine reveals structural insights into its coding potential. *Structure*, **11**, 121-127.
- Ling, H., Boudsocq, F., Woodgate, R. and Yang, W. (2001) Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. *Cell*, **107**, 91-102.
- Livneh, Z., Ziv, O. and Shachar, S. Multiple two-polymerase mechanisms in mammalian translesion DNA synthesis. *Cell Cycle*, **9**, 729-735.
- McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C. and Read, R.J. (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr*, **61**, 458-464.
- McCulloch, S.D., Kokoska, R.J., Garg, P., Burgers, P.M. and Kunkel, T.A. (2009) The efficiency and fidelity of 8-oxo-guanine bypass by DNA polymerases delta and eta. *Nucleic Acids Res*, **37**, 2830-2840.
- Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2005) Human DNA polymerase iota incorporates dCTP opposite template G via a G.C + Hoogsteen base pair. *Structure*, **13**, 1569-1577.
- Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2006) An incoming nucleotide imposes an anti to syn conformational change on the templating purine in the human DNA polymerase-iota active site. *Structure*, **14**, 749-755.
- Nair, D.T., Johnson, R.E., Prakash, S., Prakash, L. and Aggarwal, A.K. (2004) Replication by human DNA polymerase-iota occurs by Hoogsteen base-pairing. *Nature*, **430**, 377-380.
- Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307-326.
- Painter, J. and Merritt, E. (2006) TLSMD web server for the generation of multi-group TLS models. *J. Appl. Crystallogr.*, **39**, 109-111.
- Petta, T.B., Nakajima, S., Zlatanou, A., Despras, E., Couve-Privat, S., Ishchenko, A., Sarasin, A., Yasui, A. and Kannouche, P. (2008) Human DNA polymerase iota protects cells against oxidative stress. *Embo J*, **27**, 2883-2895.

- Prasad, R., Bebenek, K., Hou, E., Shock, D.D., Beard, W.A., Woodgate, R., Kunkel, T.A. and Wilson, S.H. (2003) Localization of the deoxyribose phosphate lyase active site in human DNA polymerase ι by controlled proteolysis. *J Biol Chem*, **278**, 29649-29654.
- Rechkoblit, O., Malinina, L., Cheng, Y., Geacintov, N.E., Broyde, S. and Patel, D.J. (2009) Impact of conformational heterogeneity of OxoG lesions and their pairing partners on bypass fidelity by Y family polymerases. *Structure*, **17**, 725-736.
- Rechkoblit, O., Malinina, L., Cheng, Y., Kuryavyi, V., Broyde, S., Geacintov, N.E. and Patel, D.J. (2006) Stepwise translocation of Dpo4 polymerase during error-free bypass of an oxoG lesion. *PLoS Biol*, **4**, e11.
- Salter, E., Wierzbicki, A., Sperl, G. and Thompson, W. (2003) Quantum mechanical study of the Syn and Anti conformations of solvated cyclic GMP. *Struc. Chem.*, **14**, 527-533.
- Sheffield, P., Garrard, S. and Derewenda, Z. (1999) Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. *Protein Expr Purif*, **15**, 34-39.
- Shibutani, S., Takeshita, M. and Grollman, A.P. (1991) Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature*, **349**, 431-434.
- Stolarski, R., Hagberg, C.E. and Shugar, D. (1984) Studies on the dynamic syn-anti equilibrium in purine nucleosides and nucleotides with the aid of ^1H and ^{13}C NMR spectroscopy. *Eur J Biochem*, **138**, 187-192.
- Takao, M., Zhang, Q.M., Yonei, S. and Yasui, A. (1999) Differential subcellular localization of human MutY homolog (hMYH) and the functional activity of adenine:8-oxoguanine DNA glycosylase. *Nucleic Acids Res*, **27**, 3638-3644.
- Tissier, A., Frank, E.G., McDonald, J.P., Iwai, S., Hanaoka, F. and Woodgate, R. (2000) Misinsertion and bypass of thymine-thymine dimers by human DNA polymerase ι . *Embo J*, **19**, 5259-5266.
- Vaisman, A. and Woodgate, R. (2001) Unique misinsertion specificity of polioota may decrease the mutagenic potential of deaminated cytosines. *Embo J*, **20**, 6520-6529.
- van Loon, B. and Hubscher, U. (2009) An 8-oxo-guanine repair pathway coordinated by MUTYH glycosylase and DNA polymerase λ . *Proc Natl Acad Sci U S A*, **106**, 18201-18206.
- Vasquez-Del Carpio, R., Silverstein, T.D., Lone, S., Swan, M.K., Choudhury, J.R., Johnson, R.E., Prakash, S., Prakash, L. and Aggarwal, A.K. (2009) Structure of human DNA polymerase κ inserting dATP opposite an 8-OxoG DNA lesion. *PLoS One*, **4**, e5766.

- Zhang, H., Irimia, A., Choi, J.Y., Angel, K.C., Loukachevitch, L.V., Egli, M. and Guengerich, F.P. (2006) Efficient and high fidelity incorporation of dCTP opposite 7,8-dihydro-8-oxodeoxyguanosine by *Sulfolobus solfataricus* DNA polymerase Dpo4. *J Biol Chem*, **281**, 2358-2372.
- Zhang, Y., Yuan, F., Wu, X., Taylor, J.S. and Wang, Z. (2001) Response of human DNA polymerase iota to DNA lesions. *Nucleic Acids Res*, **29**, 928-935.
- Zhang, Y., Yuan, F., Wu, X., Wang, M., Rechkoblit, O., Taylor, J.S., Geacintov, N.E. and Wang, Z. (2000) Error-free and error-prone lesion bypass by human DNA polymerase kappa in vitro. *Nucleic Acids Res*, **28**, 4138-4146.
- Zhang, Y., Yuan, F., Wu, X. and Wang, Z. (2000) Preferential incorporation of G opposite template T by the low-fidelity human DNA polymerase iota. *Mol Cell Biol*, **20**, 7099-7108.

Chapter 5¹

5 DNA replication through a nitrated polycyclic aromatic hydrocarbon lesion by human DNA polymerase ι

5.1 Introduction

Polycyclic aromatic hydrocarbons (PAH) are the most abundant and widespread organic pollutants on earth, arising from the combustion of carbon containing materials. 1-nitropyrene (1-NP) is the most prevalent nitrated PAH (NPAH) in the environment, and is the major NPAH produced from the combustion of diesel fuel (Scheepers et al., 1994). The 1-NP compound is abundant in urban air particulate and has been shown to induce DNA mutagenesis (Andersson et al., 2009; Watt et al., 2007), apoptosis (Asare et al., 2008), and increased levels of reactive oxygen species (Andersson et al., 2009) in mammalian cells. In addition, 1-NP has been shown to elicit mammary gland tumour formation in animal models (Elbayoumy et al., 1988; Hirose et al., 1984). Thus, 1-NP and related NPAH compounds likely have a major impact on human health, especially in populations living in urban or industrialized areas. The toxicity of the 1-NP compound arises from its metabolic nitro-reduction in the human body, creating a highly reactive species that attacks genomic DNA (Figure 5.1A) (Djuric et al., 1988; Sherrer et al., 2009). Nitro-reduction of 1-NP mainly occurs by bacterial flora in the gastro-intestinal tract, but has also been reported in human endothelial cells (Andersson et al., 2009), which may represent a link between air pollution and decreased cardiovascular health. The reduced 1-NP compound predominantly reacts with the C8 atom of guanine nucleotides,

¹ The contents of this chapter have been submitted for publication.

producing the bulky N-[deoxyguanosine-8-yl]-1-aminopyrene (aminopyrene-guanine, APG) adduct(Howard et al., 1983) (Figure 5.1A). The mutagenic signature of the APG lesion is the induction of G to T transversions(Watt et al., 2007), which is likely responsible for the carcinogenic effects of this highly abundant pollutant.

Bulky DNA lesions, such as APG, inhibit DNA replication of high fidelity DNA polymerases due to an inability to be accommodated in the restrictive active site of these enzymes(Hsu et al., 2005). In order to rescue adduct stalled replication forks, cells must recruit specialized DNA polymerases belonging to the Y family, which can replicate through bulky or distorting DNA lesions due to a spacious and solvent accessible active site(Broyde et al., 2008; Yang and Woodgate, 2007). Although these polymerases rescue stalled replication forks, they are also highly error-prone and thus have the potential to induce somatic mutations in the genome during lesion bypass(Matsuda et al., 2000; Rogozin et al., 2001; Tissier et al., 2000; Wang et al., 2007; Zhang et al., 2000). Therefore, it is likely that Y family DNA polymerases are contributing to the mutagenic signature of cells exposed to the 1-NP compound.

To investigate how Y family DNA polymerases replicate and induce mutations opposite the APG lesion, functional assays combined with X-ray crystallography was performed with different Y family DNA polymerases. Human DNA polymerase iota (*poli*) was crystallized in ternary complex with DNA containing the APG lesion, incorporating either correct or incorrect incoming nucleotides. These structures represent the first glimpse of a DNA polymerase replicating directly opposite a PAH lesion and reveal the first structural mechanism of NPAH induced mutagenesis by a human DNA polymerase. In addition, we crystallized the Archaeal Y family polymerase Dpo4 in an extension

complex with APG, incorporating one nucleotide upstream from the lesion. This structure reveals how the APG lesion can induce replication stalling even in highly specialized Y family DNA polymerases. Taken together, this work demonstrates how PAH compounds induce genomic mutagenesis and cellular toxicity and provides a mechanism by which specific environmental pollutants can directly impact human health.

5.2 Results

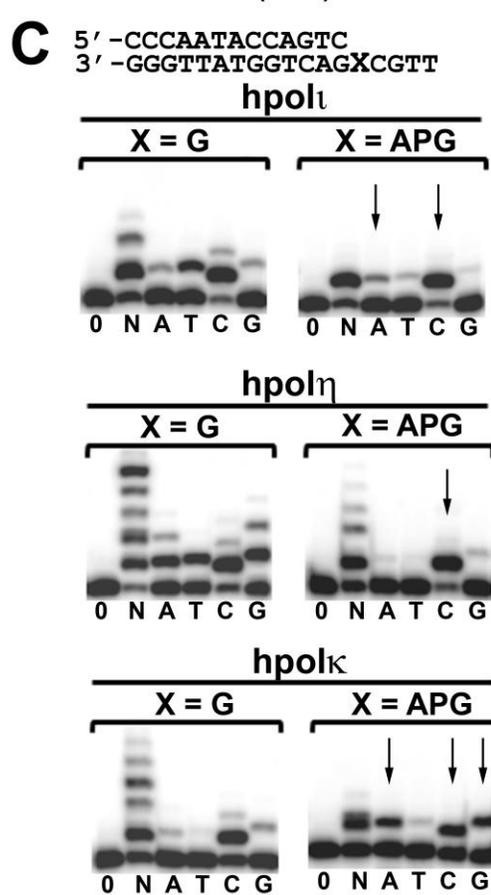
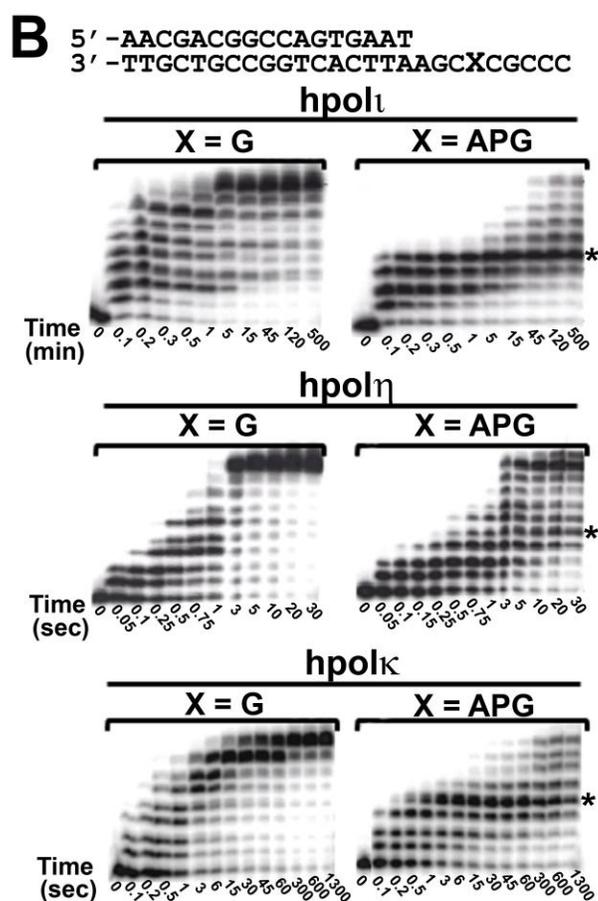
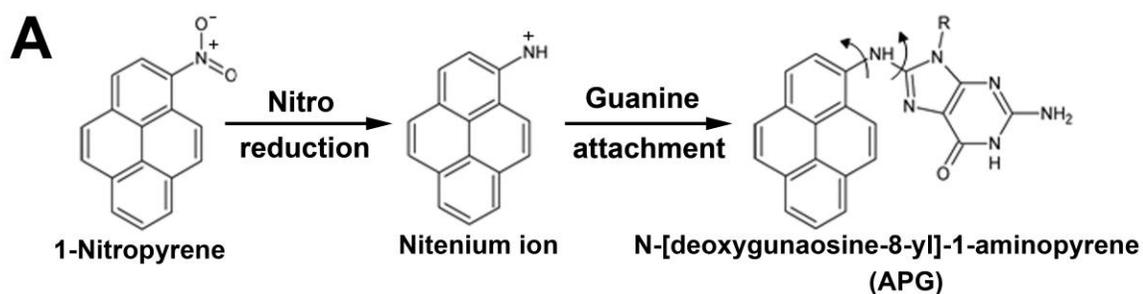
5.2.1 Y family polymerase bypass and fidelity opposite APG

To test and compare the activity of different Y family DNA polymerases during APG replication, we performed primer extension assays using human pol iota (pol ι), human pol eta (pol η), and human pol kappa (pol κ). To test the APG bypass activity of these three human enzymes, running start primer extension assays were performed with either undamaged G or the APG lesion located 4 nucleotides upstream from the primer-template junction (Figure 5.1B). In the presence of undamaged G, all three enzymes efficiently extended the primer to the end of the template DNA (Figure 5.1B). However, pol ι is the least processive of the three enzymes and thus requires a longer time point for full primer elongation. In the presence of the APG lesion the three Y family polymerases have different APG bypass abilities. Pol ι has the least efficient APG bypass ability, displaying significant replication stalling at the site of DNA damage (Figure 5.1B, asterisk). The strong stalling bands at the site of the lesion indicate an inability to extend past the APG adduct after single nucleotide incorporation, which persists after 8 hours of reaction. In contrast, pol η is able to bypass the APG lesion similar to undamaged G with no significant replication stalling. Pol κ APG bypass is similar to pol ι , with significant replication stalling at the site of DNA damage (Figure 5.1B, asterisk). However, pol κ is

able to resolve the APG stalling after 20 minutes of reaction, indicating a higher APG bypass activity than pol ι . Thus, different Y family DNA polymerases have large differences in APG bypass abilities and activity.

To test the fidelity of the three human Y family polymerases, standing start primer extension assays were performed with undamaged G or the APG lesion directly upstream of the primer-template junction (Figure 5.1C). Opposite undamaged G, all three Y family polymerases preferentially incorporate the correct C nucleotide with varying degrees of fidelity (Figure 5.1C). Mis-incorporation bands are observed for all three enzymes, which is a general feature of Y family DNA polymerases due to open and solvent accessible active sites. Opposite the APG lesion, all three polymerases maintain preferential C incorporation (Figure 5.1C). However, the fidelity of each enzyme has been dramatically altered from undamaged G. Surprisingly, for pol ι , and pol η , the enzymatic fidelity increases substantially compared to undamaged G with significant reductions in mis-incorporation bands. In contrast, pol κ 's fidelity is remarkably reduced, with A and G mis-incorporations almost equal to that of C (Figure 5.1C). Although pol ι prefers the correct C nucleotide opposite the APG lesion, there is significant incorporation of the mis-matched A nucleotide, which is the mutagenic signature of the APG adduct. Similar to the running start assays, pol ι and pol κ display replication stalling opposite the APG lesion in the presence of all four nucleotides. Interestingly, pol η also displays some replication stalling when the lesion is located directly upstream from the primer-template junction. These results indicate that different Y family polymerases have different fidelities opposite the APG lesion, with pol η having the highest fidelity, followed by pol ι , and then pol κ .

Figure 5.1 1-nitropyrene guanine attachment and Y family DNA polymerase activity. (A) 1-nitropyrene is reduced in the human body creating a reactive nitrenium ion which attaches to the C8 atom of guanine nucleotides producing the APG lesion. (B) Running start primer extension assays with three different human Y family polymerases (pol η , pol κ , pol ι) in the presence of undamaged G or the APG lesion. Enzymes were incubated with DNA substrates and reacted in the presence of all four nucleotides at various times points, indicated under each lane. Replication stalling at the site of the APG lesion is denoted with asterisks on the right hand side of the gels. (C) Standing start primer extension assays with undamaged G or the APG lesion at the first replication position were performed with pol η , pol κ , and pol ι . Enzymes were incubated with DNA substrates and reacted with all four nucleotides (N) or individual nucleotides (A, T, C, and G) for either 0.5 min for undamaged G or 30 min for the APG lesion. Vertical arrows indicate nucleotide preferences opposite the APG lesion. DNA substrates are shown above the gels.



5.2.2 Polt-APG-dNTP ternary structures

In order to elucidate how A nucleotides are mis-incorporated opposite APG and how this bulky adduct induces replication stalling in Y DNA family polymerases, we sought to crystallize polt in ternary complex with the APG lesion. DNA substrate for crystallization was designed containing the APG lesion directly upstream from the primer-template junction, ready for dNTP incorporation (Figure 5.2A). The DNA substrate was incubated with polt and co-crystallized with either incoming dCTP or dATP nucleotides. The subsequent structures are denoted as APG:dCTP and APG:dATP according to the identity of the replicating base pairs in the active site. Both polt-APG ternary crystals diffracted to 2.9 Å resolutions (Table 5.1) and represent the first structures of a DNA polymerase replicating directly opposite a PAH lesion. Polt in APG:dCTP and APG:dATP structures is very similar, with C α backbone root mean square deviations (r.m.s.d.) within 0.5 Å (Figure 5.2B). Only slight differences are observed in domain positioning, which is likely due to differences in nucleotide conformations in the active site. However, comparing both APG polt structures with a previously solved polt ternary complex with undamaged G (Nair et al., 2005), reveals that small conformational changes occur in the little finger domain and the DNA substrate upon APG binding. The little finger domains have moved downwards by ~5° compared with the undamaged G structure (Figure 5.2B), which is in response to the 5' end of the template DNA moving out towards the solvent exposed major groove. In addition, the bottom of the DNA substrate has rotated towards the thumb domain by ~10° in order to accommodate the shift in the little finger. These small domain and DNA re-orientations are necessary to accommodate the bulky APG lesion in the polt active site. The polt-

Table 5.1 Summary of crystallographic data for polt aminopyrene ternary structures

Data collection	APG:dCTP	APG:dATP	APG-1:dC
Space group	P6 ₃ 22	P6 ₃ 22	P6 ₁
Mol/AU ^a	1	1	1
Unit Cell			
<i>a, b, c</i> (Å)	98.0, 98.0, 194.6	98.9, 98.9, 194.2	56.2, 56.2, 290.6
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å) ^b	50.0-2.90 (2.95-2.90)	50.0-2.90 (2.90-2.90)	50.0-2.00 (2.12-2.00)
Unique reflections	13253	13665	34904
Completeness (%) ^b	97.2 (98.2)	98.5 (98.8)	99.7 (95.6)
Redundancy ^b	7.0 (7.2)	4.0 (4.2)	4.1 (2.3)
<i>I</i> / σ <i>I</i> ^b	29.4 (2.7)	27.5 (2.7)	33.8 (5.15)
<i>R</i> _{merge} ^b	7.8 (63.3)	6.2 (55.2)	4.9 (22.3)
Refinement statistics			
<i>R</i> _{work} / <i>R</i> _{free}	23.6 / 28.1	22.7 / 27.9	19.1 / 22.6
No. atoms			
Protein	2863	2855	2755
DNA	326	326	609
dNTP	28	30	31
Ions	2	2	3
Water	27	19	482
Avg B factor			
Protein	86.1	88.0	26.4
DNA	78.3	84.4	30.2
dNTP	78.0	85.6	38.1
Ions	60.8	66.5	29.3
Waters	70.8	76.1	37.7
R.m.s. deviations			
Bonds (Å)	0.005	0.010	0.008
Angles (°)	1.04	1.53	1.32
^a Mol/AU represents the number of molecules per asymmetric unit.			
^b Data in the highest resolution shell are in parentheses			
^c One non-catalytic Mg ²⁺ ion exists in all structures			

APG structures have the same space group as previously solved polt structures (P6₅22) and the largest surface area interaction is between the protein and DNA substrate, indicating that the observed domain and DNA movements are not due to the packing environment of the crystal lattice.

5.2.3 APG positioning in the polt active site

The orientation of the APG lesion base in the polt active site is dramatically altered compared to undamaged G(Nair et al., 2005). Traditionally, undamaged template purines adopt a *syn* conformation in the polt active site, orienting the C8 atom towards the protein occluded minor groove and forming a Hoogsteen base pair with the incoming nucleotide(Nair et al., 2006). Polt induces *syn* conformations due to a remarkably narrow active site that restricts the C1' - C1' distance to under 9 Å. A Watson-Crick base pair requires a C1' - C1' distance of ~10.6 Å, and thus, this mode of base pairing is highly unfavourable in the narrow polt active site(Kirouac and Ling, 2009). Since the APG base has an aminopyrene ring system attached to the C8 atom, a *syn* conformation would be impossible in the polt active site. Thus, it was of great interest to see how polt could accommodate and bypass the APG lesion using an alternate mechanism to that of previous template purine structures. In the APG:dCTP structure, the APG base is in an *anti* conformation, due to an inability to adopt a *syn* orientation (Figure 5.2C). Since purine *anti* conformations are unfavourable in the polt active site, the APG base and DNA backbone are pushed slightly out of the active site towards the solvent exposed major groove with the aminopyrene attachment ejected away from the DNA helix (Figure 5.2C). This alteration in the template base and DNA backbone accommodates an increased C1' - C1' distance, which allows the APG lesion to adopt an *anti* conformation

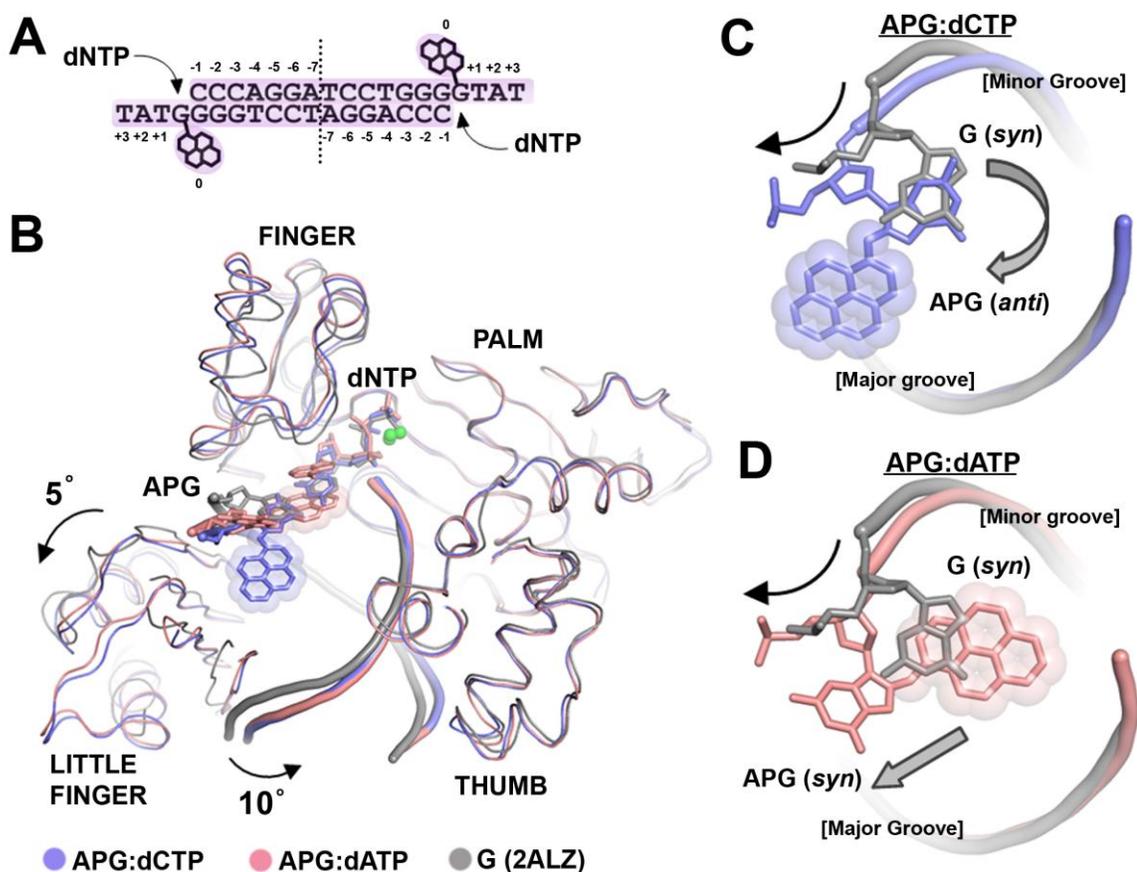


Figure 5.2 Superposition of polt-APG ternary structures. (A) DNA substrate used for crystallization. Numbering is relative to the APG template at position 0. Vertical dashed line indicates axis of symmetry. (B) Superposition of APG:dCTP (blue), APG:dATP (pink), and a previous polt ternary complex with undamaged G (PDB: 2ALZ, grey). Domains are labelled and arrows indicate positioning relative to the undamaged G structure. The aminopyrene lesion is shown with incoming nucleotides and metal ions (green spheres). (C) Positioning of APG in APG:dCTP (blue) and (D) APG:dATP (red) relative to undamaged G. View is looking from top, down through the DNA helix. Black arrows indicate backbone DNA movement and grey block arrows indicate APG base movement. Major and minor groove sides of the DNA helices are labelled.

for base pairing. In the APG:dATP structure, the APG base maintains a *syn* conformation, however, the APG base and DNA backbone are fully projected into the solvent exposed major groove with the aminopyrene attachment occupying the position of the replicating base pair in the active site (Figure 5.2D). The base of the incoming dATP nucleotide is tilted up and stacking on the aminopyrene ring, while the phosphate moiety keeps a regular position as in other polt structures (Kirouac and Ling, 2009; Nair et al., 2006; Nair et al., 2004).

5.2.4 Replicating base pair conformations

Interestingly, the positioning of the APG lesion base is directly influenced by the identity of the incoming nucleotide. Both incoming dCTP and dATP nucleotides are coordinated by one active site Mg^{2+} ion located at the A site and the α phosphate is in reaction distance to the 3' -OH group of the primer strand. Lack of ion coordination at the B site has previously been observed for polt and Dpo4 due to ion mobility at this location (Nair et al., 2005; Vaisman et al., 2005). Incoming dCTP forms a Watson-Crick base pair with the APG base in an *anti* conformation (Figure 5.3A). In this orientation the hydrophobic aminopyrene moiety is positioned in the solvated major groove with no stabilizing interactions. Thus, the aminopyrene ring is highly mobile, resulting in poor electron density for the hydrophobic pyrene attachment. For incoming dATP, the nucleotide base is too large to pair with the APG lesion base. Thus, the dATP base forms a stacking interaction above the hydrophobic lesion, allowing the aminopyrene attachment to be accommodated in the polt active site (Figure 5.3B). In this orientation the hydrophobic aminopyrene ring is highly stabilized through stacking interactions, which is reflected by the well defined electron density for the lesion. Similar intrahelical conformations of

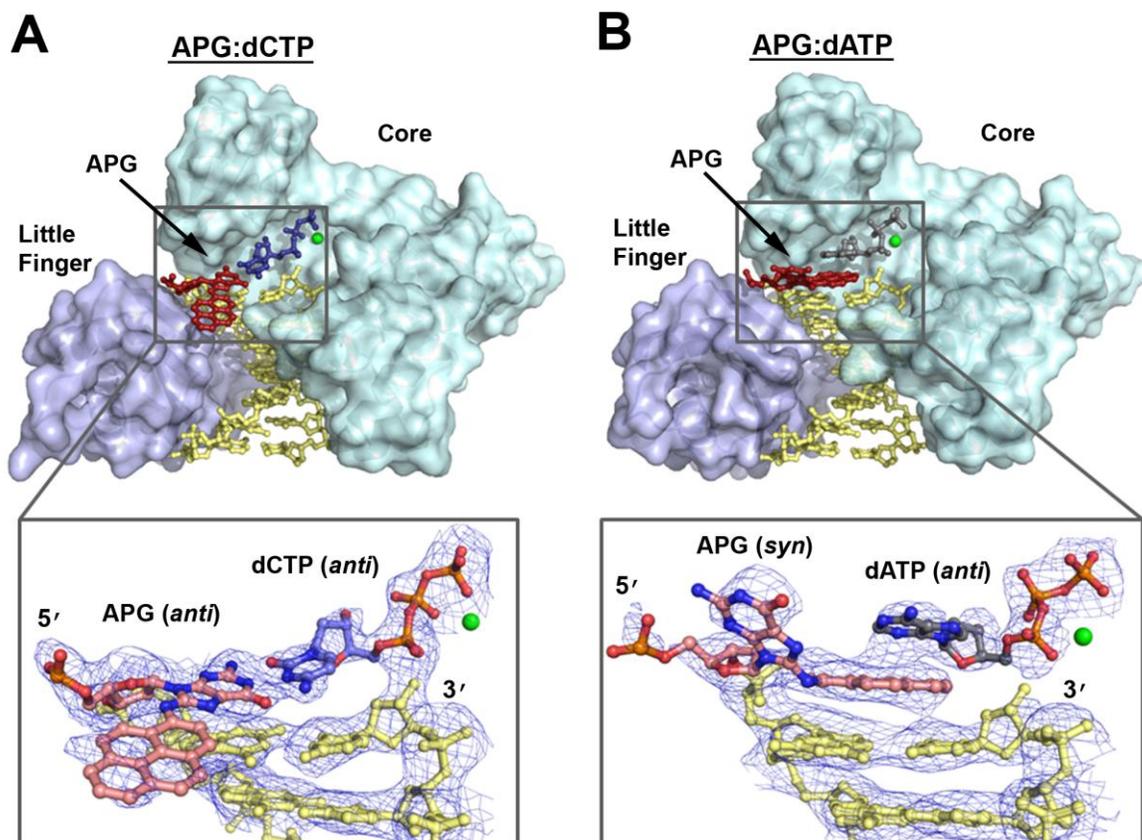


Figure 5.3 Replicating base pair conformations in the polt-APG ternary structures. (A) APG:dCTP and (B) APG:dATP ternary structures are shown. The finger, thumb, and plam domains are coloured cyan and labelled ‘core’, while the little finger domains are coloured light blue. The DNA is coloured in yellow, the APG lesion in red, dCTP in blue, and dATP in grey. Zoom in views of both active sites are shown below the structures with the $2F_o - F_c$ electron density map contoured at 1σ . Active site metal ions are shown as green spheres.

PAH lesions have been reported (Bauer et al., 2007; Ling et al., 2004), suggesting this is a common mechanism to stabilize PAH ring systems. The APG:dATP structure reveals the first glimpse of a PAH lesion inducing dA mis-incorporations in a growing DNA primer strand. No hydrogen bonds are formed between the dATP base and APG lesion, thus, dATP mis-incorporation is dictated purely by base stacking interactions. This indicates that the dATP nucleotide is less stabilized than correct dCTP opposite APG and would be incorporated less efficiently. The polt-APG insertion structures match well with our in vitro polt replication assays, where dCTP is most preferred opposite APG, and dATP mis-insertion is second preferred.

The replicating base pairs in the polt-APG structures not only reveal a structural basis for polt-APG fidelity, but also give insight into polt-APG replication stalling. Looking from the top of the replicating base pairs down through the DNA helix, reveals how the APG base pairs would translate through the polymerase core (Figure 5.4A,B). When the APG lesion is paired with the correct dC nucleotide, the hydrophobic pyrene attachment is projected out of the DNA helix (Figure 5.4A). This conformation would likely inhibit DNA translocation through the polymerase core due to the pyrene attachment clashing with the little finger domain, which lies directly underneath the lesion in this conformation. The hydrophobic aminopyrene ring would likely interact with the protein molecule in order to shield itself from the solvent, which would inhibit DNA replication. In contrast, when the APG base is paired with the mis-matched dA nucleotide, the hydrophobic aminopyrene attachment is stacking directly in the middle of the DNA helix, mimicking a base pair (Figure 5.4B). In this conformation the DNA helix would easily translate through the polymerase core due to the aminopyrene ring system being shielded

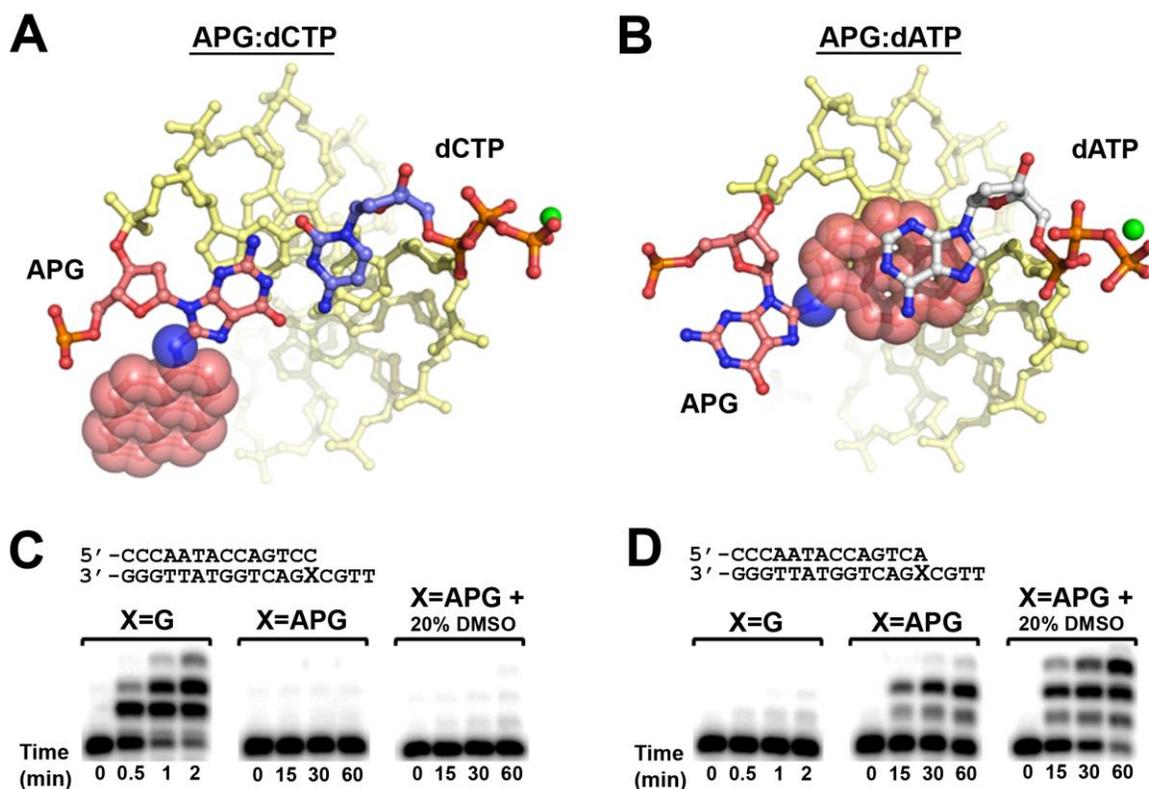


Figure 5.4 APG lesion conformation and replication stalling. (A) Top view of the APG:dCTP and (B) APG:dATP replicating base pairs. The DNA helix is in yellow, the APG lesion in red with sphere representation for the hydrophobic ring system. The dCTP is coloured in blue, dATP in grey, and metal ions are shown as green spheres. (C) Primer extension assays were performed with undamaged G or the APG lesion paired with correct C or (D) mis-matched A from the primer strand. PolI was incubated with DNA substrates and all four nucleotides and reacted for various time points indicated under each lane in the presence or absence of 20% DMSO. DNA substrates are shown above the gels.

from protein interactions. To test this structural hypothesis of polt-APG replication stalling, primer extension assays were performed with undamaged G or the APG lesion paired with correct C or mis-matched A at the primer-template junction (Figure 5.4C,D). In the presence of an undamaged G:C base pair, polt can efficiently extend the primer strand (Figure 5.4C). However, in the presence of an APG:C base pair, primer extension by polt is completely inhibited. The addition of 20% DMSO to the reaction buffer in order to make the solvent more favourable for the hydrophobic lesion, does little to alleviate the APG:C base pair stalling. This indicates that the aminopyrene attachment is likely completely buried into the protein side chains of the little finger domain when polt is extending past the correctly matched dC nucleotide. Interestingly, when an A nucleotide is paired instead of the correct C, a complete opposite effect is observed. In the presence of an undamaged G:A base pair, the primer extension by polt is nearly inhibited (Figure 5.4D). This is likely due to distorted geometry and altered base stacking of a mis-matched G:A primer template junction. In contrast, when an APG:A base pair is present, polt is able to efficiently extend the primer. This is consistent with our structural observations which demonstrate the APG ring system intercalates into the DNA helix and mimics a DNA base pair when matched with the incorrect A nucleotide. Although the hydrophobic APG ring system is intercalated into the DNA helix when paired with mis-matched A, the edge of the aminopyrene ring is still exposed to solvent due to the accessibility of the DNA major groove when in complex with Y family polymerases. Thus, the addition of 20% DMSO increases the extension efficiency of APG:dA by likely further stabilizing the aminopyrene lesion into the DNA helix (Figure 5.4D). A similar effect of greater extension past APG:dA compared to APG:dC is observed for pol η , pol κ ,

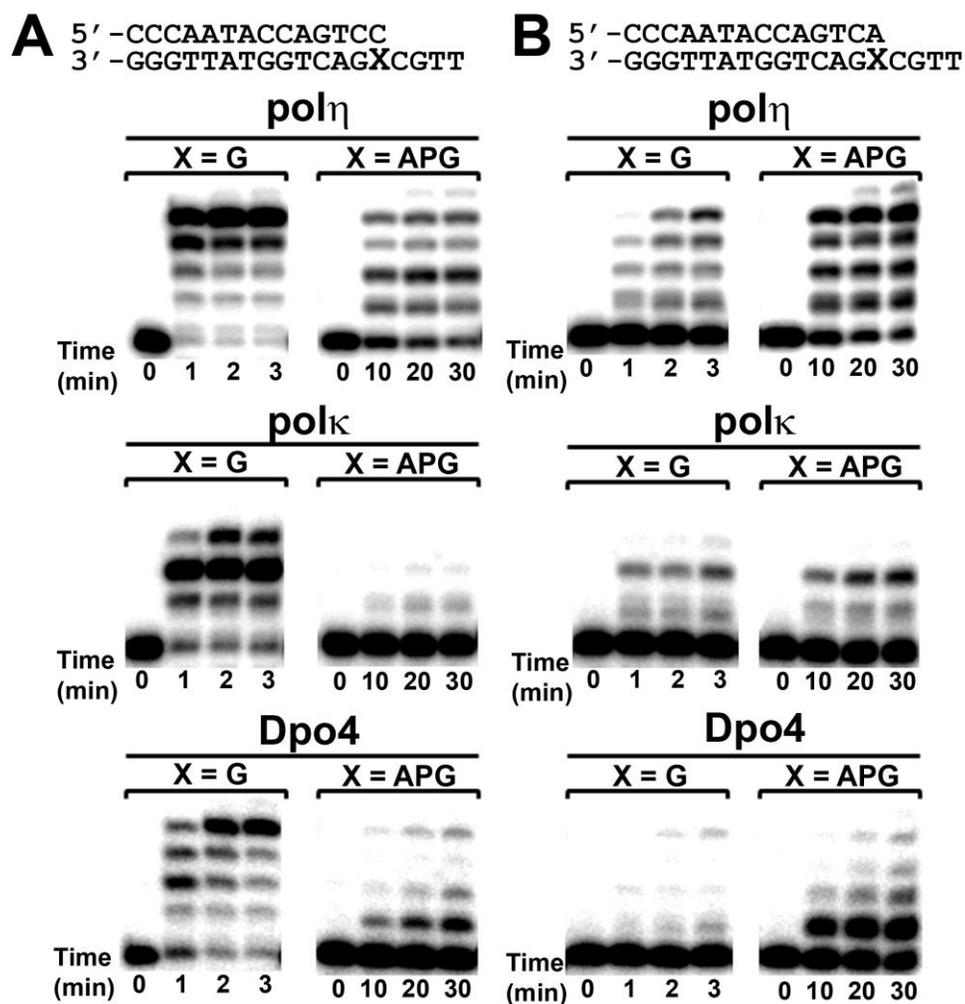


Figure 5.5 APG stalling characteristics of different Y family DNA polymerases. Primer extension assays were performed with pol η , pol κ , and Dpo4. Enzymes were incubated with undamaged G or the APG lesion paired with (A) correct C or (B) mis-matched A bases from the primer strand. Reactions were carried out with all four nucleotides at various time points indicated under each lane. DNA substrates are shown above the gels.

and Dpo4, a model Y family DNA polymerase from Archaea (Figure 5.5). Taken together, these results confirm our structural hypothesis of APG:C base pair induced stalling and validates the nucleotide conformations in the polt-APG ternary structures.

5.2.5 APG:dC induced replication stalling

To further uncover the molecular basis of replication stalling past an APG:C base pair, we sought to determine an APG:C extension structure. Although we were unable to obtain polt extension crystals, we managed to crystallize Dpo4 in an extension complex with the APG:dC base pair. Dpo4 is a model Y family polymerase from archaea, and displays almost identical replication stalling and fidelity as polt opposite the APG lesion (Sherrer et al., 2009). Thus, the Dpo4 APG:dC extension structure serves as a good molecular model to explain polt APG:C induced stalling. The DNA substrate for crystallization was designed containing the template APG lesion paired with correct dC at the primer-template junction (Figure 5.6A). The DNA substrate was incubated with Dpo4 and crystallized with incoming dGTP, which pairs with a template C base upstream from the APG primer-template junction. This structure is referred to as APG-1:dC according to the identity and position of the APG base pair and diffracted to 2.0 Å resolution (Table 5.1).

Similar to polt, Dpo4 has individual domain movements and rotation of the DNA substrate, in order to accommodate the APG lesion in the polymerase core (Figure 5.6B). Interestingly, one dGMP molecule is observed stacking under the replicating base pair due to a large structural gap. No dGMP was included during crystallization, thus this compound likely arises from degradation of the 3' end of the template DNA. The 3' -5' exonuclease activity has been previously observed for Dpo4

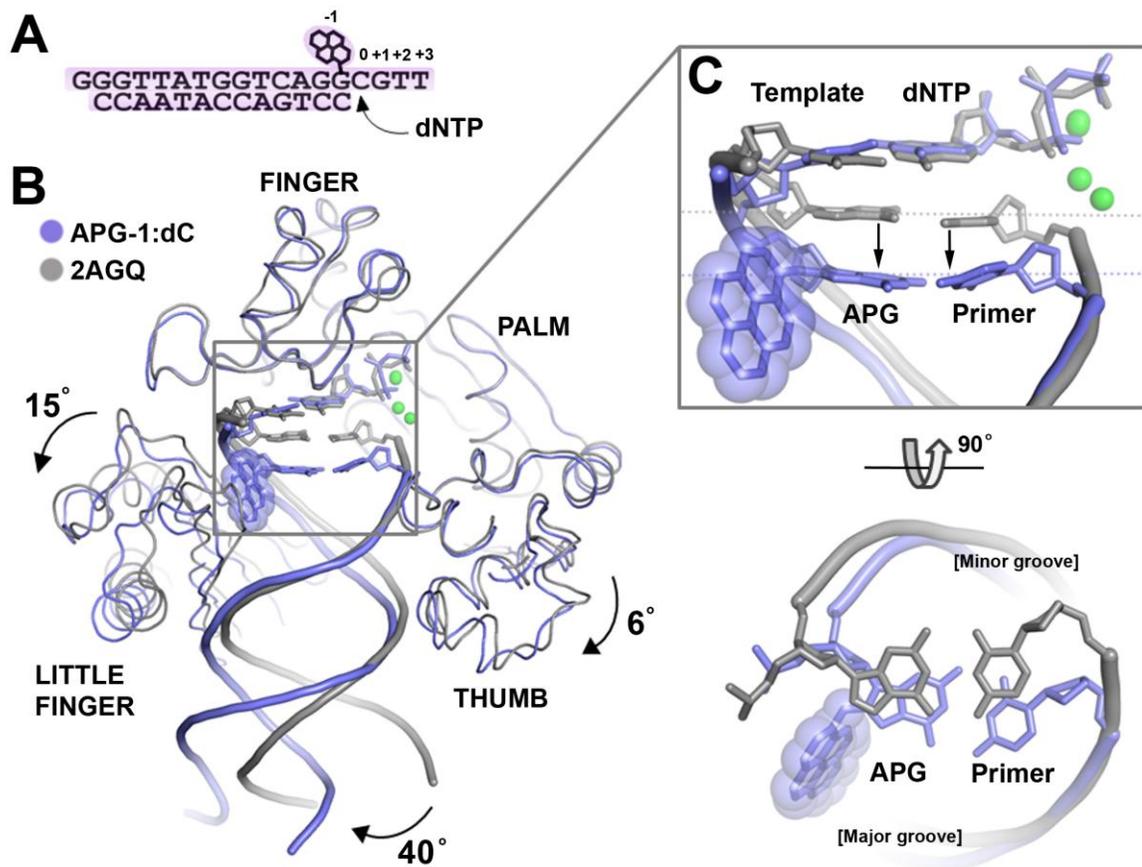


Figure 5.6 Superposition of Dpo4 APG extension structure. (A) DNA substrate used for crystallization. Numbering is relative to the template C base at the position 0. (B) Superposition of the APG-1:dC structure (blue) with a previously solved Dpo4 ternary complex with undamaged DNA (PDB: 2AGQ, grey). The dGMP in APG-1:dC has been removed for clarity. Individual domains are labelled and arrows indicate movement relative to Dpo4 in complex with undamaged DNA. The APG lesion is shown with transparent spheres and incoming nucleotides are shown in stick model with coordinated metal ions as solid green spheres. (C) Zoom in view of the superimposed active sites showing how the APG:dC base pair moves down into a non-standard position. Top view shows the major and minor groove sides of the DNA helix.

(Vaisman et al., 2005; Wong et al., 2008). Although the replicating base pair is in a standard orientation, the underlying base pair containing the APG lesion has moved down the polymerase core to an inactive position, where the primer strand is too far (~ 6 Å) from the α phosphate of the incoming nucleotide for phosphodiester bond formation (Figure 5.6C, 5.7A). This non-productive shift in the APG:dC base pair is stabilized by the hydrophobic aminopyrene lesion interacting with the little finger domain and the dGMP nucleotide filling the gap in the DNA helix. Such an interaction would likely exist in polt primer extension after the APG:dCTP insertion. The aminopyrene ring system is sandwiched through stacking interactions with Arg332 of the little finger and a base from the single stranded template DNA (dG +1) (Figure 5.7B). In addition, Val289 and Ile295 from the Little Finger form van der Waals interactions with the aminopyrene ring system, which further shield the APG lesion from the solvent (Figure 5.7B). Thus, the need to stabilize the projected hydrophobic aminopyrene lesion through protein interactions inhibits DNA replication. This result is consistent with our primer extension assays, which indicate APG-protein interactions induce replication stalling in the presence of an APG:dC base pair.

5.3 Discussion

Exposure to urban air pollution increases morbidity and mortality rates in human populations (Samet et al., 2000). One of the main contributors to the detrimental health effects of air pollution is exposure to PAHs. Although there are hundreds of different carcinogenic PAH compounds found in atmospheric pollution, the metabolites of these compounds all share three common characteristics: 1) hydrophobicity, 2) preferential attachment to guanine nucleotides, and 3) induction of G to T transversions at the site of

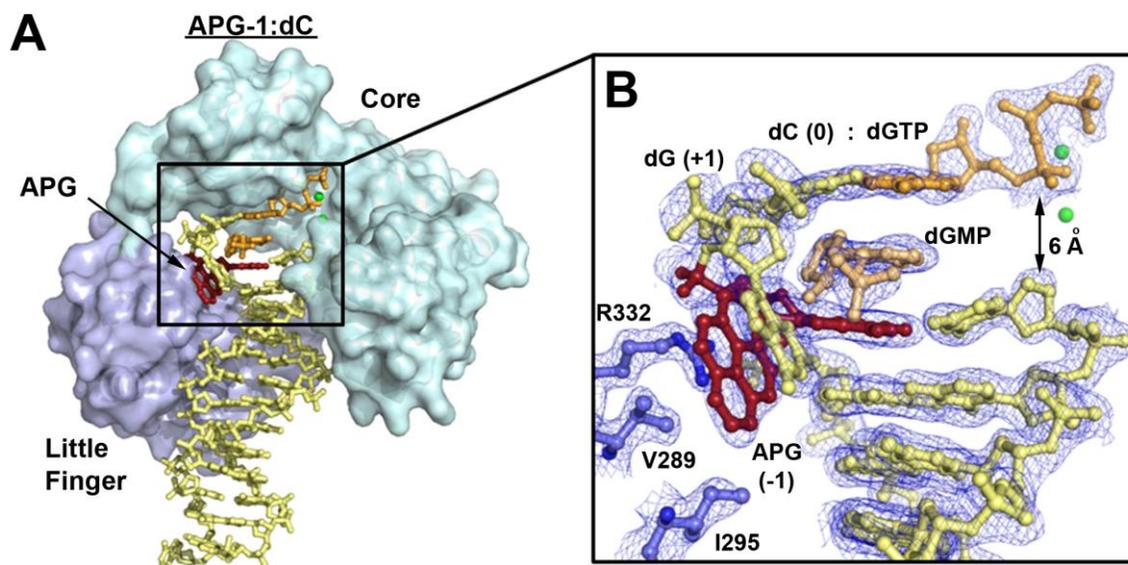


Figure 5.7 APG positioning in a stalled Dpo4 extension complex. **(A)** Structure of APG-1:dC is shown with finger, thumb, and palm coloured cyan and labelled ‘core’. The little finger domain is coloured in light blue, the DNA helix in yellow, the APG lesion in red and the incoming nucleotide in orange. Active site metal ions are shown as green spheres. **(B)** Zoom in view of the APG-1:dC active site showing the aminopyrene contacts with residues from the little finger domain (R332, V289, I295) and single stranded DNA nucleotide (dG +1). Large distance between the primer strand and incoming nucleotide is shown with black double arrow. A dGMP nucleotide is shown stacking in the structural gap between the replicating base pair (0) and the underlying APG:dC pair (-1).

the guanine adduct (Chiapperino et al., 2002; Denissenko et al., 1996; Vogel et al., 2001; Yoon et al., 2001; Zhao et al., 2006). The APG:dATP pol β structure reveals the first mechanism of how A nucleotides are mis-incorporated opposite PAH-guanine lesions. The bulky hydrophobic ring system of the PAH-lesion stabilizes itself away from the solvent by stacking above the primer-template junction. This conformation allows the mis-matched A nucleotide to stack on top of the ring system and be incorporated into the growing primer strand. After one more round of DNA replication, the A mis-match base will result in the signature G to T transversion.

The A nucleotide base has the highest base stacking potential of all four nucleotides and thus is the most favoured mis-matched nucleotide for stacking above the PAH ring system. A similar mechanism is observed for Y family polymerases on blunt end DNA, where the A nucleotide is preferentially incorporated by stacking interactions alone (Fiala et al., 2007). Thus the intercalated PAH ring system can mimic a replicating base pair in the DNA helix and allow template-independent dA mis-incorporations, resulting in G to T transversions. Unsurprisingly, G to T transversions are highly abundant in lung cancers from smokers, who are exposed to high concentrations of PAH compounds through cigarette smoke inhalation (Hainaut and Pfeifer, 2001; Pfeifer et al., 2002). Although the size of the PAH ring system and different chemical attachments could alter the base stacking potential in the DNA helix, the stacking mechanism of A mis-incorporation is likely common to all genotoxic PAH DNA lesions. Furthermore, a common mechanism of A mis-incorporation may explain the widespread mutagenic potential of PAH compounds, and their propensity to elicit carcinogenesis and disease.

Interestingly, A mis-incorporations opposite PAH-guanine lesions may provide cells with a survival advantage over correct C incorporation. Our structural and biochemical studies have revealed that the mis-match A nucleotide promotes primer elongation by stabilizing the PAH ring system in the DNA helix, while the correctly matched C nucleotide induces replication stalling by projecting the PAH ring system into the protein molecule. Thus, correct C incorporation may induce excessive replication fork stalling leading to cellular toxicity and apoptosis, which is characteristic of many PAH DNA lesions(Asare et al., 2008; Raychoudhury and Kubinski, 2003; Tithof et al., 2002). In contrast, the mis-matched A nucleotide would likely promote replication fork progression, which may induce a selective pressure for incorporating mis-matched A nucleotides. Although many factors must contribute to PAH induced cellular toxicity, the ability of PAH DNA lesions to induce replication stalling in specialized Y family polymerases is likely to play a role.

One key question that remains to be answered is which Y family polymerase is the predominant enzyme for PAH DNA lesion bypass. For the APG lesion, it would appear from *in vitro* assays that pol η is the likely candidate, due to its high fidelity and efficient bypass of the NPAH adduct. However, *in vivo* experiments measuring 1-NP induced mutagenesis and cellular toxicity(Asare et al., 2008; Watt et al., 2007) more closely match the APG lesion bypass activity and fidelity of pol ι or pol κ . A more likely scenario is that multiple polymerases are required for PAH lesion bypass, with each having distinct functionalities, such as, nucleotide insertion or primer extension past the lesion. Indeed, it has been previously suggested that multiple polymerases are involved in benzo[a]pyrene lesion bypass as well as many other forms of DNA damage(Livneh et al., 2010; Shachar et al., 2009; Zhao et al., 2006). In addition, it is likely that different Y

family polymerases are specialized at replicating opposite different PAH lesions, depending on the site of nucleotide attachment (major vs minor groove), the size of the PAH ring system, and the presence of chemical side chain ring attachments. Supporting this idea is the observation that opposite a benzo[a]pyrene lesion pol η is error-prone (Zhang et al., 2000) and pol κ is error-free (Zhang et al., 2000), while opposite the APG lesion pol η is error-free and pol κ is error-prone. Thus, Y family polymerase activity and fidelity varies greatly opposite different types of PAH lesions.

In conclusion we have shown how an abundant NPAH organic pollutant (1-Nitropyrene) produced in automobile exhaust alters the coding sequence of genomic DNA, which induces carcinogenesis and disease. We have also shown how this pollutant may induce cellular toxicity by stalling replication fork progression. These results bring awareness to the health dangers imposed by urban air pollution and highlight the importance of developing fuel efficient and clean energy technologies.

5.4 Methods

5.4.1 Protein Preparation

Human DNA pol τ and Dpo4 protein used for crystallization were expressed and purified as previously described (Kirouac and Ling, 2009). Human proteins used for replication assays (pol τ 1-430, pol η 1-445, pol κ 19-523) contained N-ter histadine tags and were expressed in *E.coli* and purified by nickel affinity followed by ion exchange chromatography.

5.4.2 DNA preparation

The APG DNA substrates used for crystallization and activity assays were purified using ion exchange chromatography on a SourceQ column. For polI, the self-annealing 18-nt oligo containing an APG (G*) lesion (5'-TCAG*GGGTCCTAGGACCC-3') was annealed with itself to give a DNA substrate with two replicative ends. For Dpo4, the 18-nt oligo containing an APG (G*) lesion (5'-TTGCG*GACTGGTATTGGG-3') was annealed with a 13-nt primer containing a dideoxy cytosine (C^{dd}) at the 3' end (5'-CCAATACCAGTCC^{dd}-3'). Undamaged oligonucleotides used for primer extension assays were purchased from Keck Oligo Inc. and purified by ion exchange. For the running start activity assays a 17-nt primer was annealed to either an undamaged 26-nt template or an APG containing 26-nt template (Shen et al., 2002). For the standing start fidelity assays a 12-nt primer (5'-CCAATACCAGTC-3') was annealed to either the 18-nt APG (G*) template used for Dpo4 crystallization or an 18-nt undamaged G template (5'-TTGCGGACTGGTATTGGG-3'). For the extension assays, a 13-nt primer containing C at the 3' end (5'-CCAATACCAGTCC-3') or A at the 3' end (5'-CCAATACCAGTCA-3') were annealed to either the 18-nt undamaged template or the 18-nt APG template. The primers were 5'-end labelled using [γ -³²P]ATP and T4 polynucleotide kinase. The 5'-labelled primers were mixed with the template DNA at a 1.5:1 molar ratio and heated to 60°C, followed by slow cooling to form the annealed DNA substrates.

5.4.3 Primer Extension Assays

DNA substrates (10 nM) were incubated with either polI, pol η , pol κ , or Dpo4 (10 nM) and 100 μ M of either all four dNTPs or individual dNTPs at 37°C in reaction buffer

containing 40 mM Tris (pH 8.0), 5 mM MgCl₂, 250 ug/ml BSA, 10 mM DTT, and 2.5% glycerol. For the standing start fidelity assays, reactions were carried out for 30 seconds with undamaged DNA and 30 minutes for the APG DNA. Reaction times for all other experiments are indicated below the gel. Reactions were terminated with loading buffer (95% formamide, 20 mM EDTA, 0.025% xylene, 0.025% bromophenol blue) and resolved on a 20% polyacrylamide gel containing 7 M urea. Gels were visualized using a PhosphorImager.

5.4.4 Crystallization and Structure Determination

Ternary complexes were formed for NPG:dCTP and NPG:dATP by incubating polt protein (0.2 mM) and DNA in a 1:1.2 ratio with dNTP (5mM), and MgCl₂ (5 mM). Crystals of both complexes were obtained in 15% PEG 5000 MME + 0.2M NH₄SO₄ + 2.5% glycerol + 0.1 M MES, pH 6.5. Mico-crystals appeared overnight after optimized streak seeding and grew for approximately one month before diffraction quality crystals were obtained. Crystals were flash frozen in liquid Nitrogen directly from dehydrated crystallization drops to prevent crystal cracking. X-ray diffraction data were collected using the micro focus beamline at 24-ID-E at the Advanced Photon Source in Argonne National Laboratory. The ternary complex for NPG-1 was formed by incubating Dpo4 protein (0.2 mM) and DNA in a 1:1.2 ratio with dGTP (2mM), and MgCl₂ (5 mM). Crystals were obtained in 20% PEG 3350 + 0.2M CaAc₂ + 2.5% glycerol + 0.1 M Hepes, pH 7.0 and were cryo protected with the addition of 20% Ethylene glycol to the mother liquor. X-ray diffraction data were collected on beamline 24-ID-C at the Advanced Photon Source in Argonne National Laboratory

All data were processed and scaled using HKL (Otwinowski and Minor, 1997). All three structures were solved by molecular replacement using PHASER(McCoy et al., 2005) with a previously solved ternary complex (PDB: 3GV5 (iota) and PDB: 1JX4 (Dpo4)) as initial search models. Structural refinement was performed using PHENIX (Adams et al., 2010), starting with rigid body refinement followed by restrained and B factor refinement and lastly TLS refinement (Painter and Merritt, 2006). All structures have good stereochemistry with over 95% of the residues in the most favoured region of the Ramachandran plot. Model building was performed using COOT(Emsley and Cowtan, 2004) and figures were created with PYMOL(DeLano, 2002).

5.5 References

- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C. and Zwart, P.H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*, **66**, 213-221.
- Andersson, H., Piras, E., Demma, J., Hellman, B. and Brittebo, E. (2009) Low levels of the air pollutant 1-nitropyrene induce DNA damage, increased levels of reactive oxygen species and endoplasmic reticulum stress in human endothelial cells. *Toxicology*, **262**, 57-64.
- Asare, N., Landvik, N., Lagadic-Gossmann, D., Rissel, M., Tekpli, X., Ask, K., Lag, M. and Holme, J. (2008) 1-Nitropyrene (1-NP) induces apoptosis and apparently a non-apoptotic programmed cell death (paraptosis) in Hepa1c1c7 cells. *Toxicol. Appl. Pharm.*, **230**, 175-186.
- Bauer, J., Xing, G., Yagi, H., Sayer, J.M., Jerina, D.M. and Ling, H. (2007) A structural gap in Dpo4 supports mutagenic bypass of a major benzo[a]pyrene dG adduct in DNA through template misalignment. *Proc Natl Acad Sci U S A*, **104**, 14905-14910.

- Broyde, S., Wang, L., Rechkoblit, O., Geacintov, N.E. and Patel, D.J. (2008) Lesion processing: high-fidelity versus lesion-bypass DNA polymerases. *Trends Biochem Sci*, **33**, 209-219.
- Chiapperino, D., Kroth, H., Kramarczuk, I.H., Sayer, J.M., Masutani, C., Hanaoka, F., Jerina, D.M. and Cheh, A.M. (2002) Preferential Misincorporation of Purine Nucleotides by Human DNA Polymerase η Opposite Benzo[a]pyrene 7,8-Diol 9,10-Epoxide Deoxyguanosine Adducts. *J. Biol. Chem.*, **277**, 11765-11771.
- DeLano, W.L. (2002) *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA, USA.
- Denissenko, M.F., Pao, A., Tang, M. and Pfeifer, G.P. (1996) Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. *Science*, **274**, 430-432.
- Djuric, Z., Fifer, E., Yamazoe, Y. and Beland, F. (1988) DNA-binding by 1-nitropyrene and 1,6-dinitropyrene in vitro and in vivo effects of nitroreductase induction. *Carcinogenesis*, **9**, 357-364.
- Elbayoumy, K., Rivenson, A., Johnson, B., Dibello, J., Little, P. and Hecht, S. (1988) A study of chemical carcinogenesis. 114. Comparative tumorigenicity of 1-nitropyrene, 1-nitrosopyrene, and 1-aminopyrene administered by gavage to sprague-dawley rats. *Cancer Res.*, **48**, 4256-4260.
- Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60**, 2126-2132.
- Fiala, K.A., Brown, J.A., Ling, H., Kshetry, A.K., Zhang, J., Taylor, J.S., Yang, W. and Suo, Z. (2007) Mechanism of template-independent nucleotide incorporation catalyzed by a template-dependent DNA polymerase. *J Mol Biol*, **365**, 590-602.
- Hainaut, P. and Pfeifer, G.P. (2001) Patterns of p53 G \rightarrow T transversions in lung cancers reflect the primary mutagenic signature of DNA-damage by tobacco smoke. *Carcinogenesis*, **22**, 367-374.
- Hirose, M., Lee, M., Wang, C. and King, C. (1984) Induction of rat mammary-gland tumors by 1-nitropyrene, a recently recognized environmental mutagen. *Cancer Res.*, **44**, 1158-1162.
- Howard, P.C., Heflich, R.H., Evans, F.E. and Beland, F.A. (1983) Formation of DNA adducts in vitro and in *Salmonella typhimurium* upon metabolic reduction of the environmental mutagen 1-nitropyrene. *Cancer Res*, **43**, 2052-2058.
- Hsu, G.W., Huang, X., Luneva, N.P., Geacintov, N.E. and Beese, L.S. (2005) Structure of a High Fidelity DNA Polymerase Bound to a Benzo[a]pyrene Adduct That Blocks Replication. *J. Biol. Chem.*, **280**, 3764-3770.

- Kirouac, K.N. and Ling, H. (2009) Structural basis of error-prone replication and stalling at a thymine base by human DNA polymerase iota. *Embo J*, **28**, 1644-1654.
- Ling, H., Sayer, J.M., Plosky, B.S., Yagi, H., Boudsocq, F., Woodgate, R., Jerina, D.M. and Yang, W. (2004) Crystal structure of a benzo[a]pyrene diol epoxide adduct in a ternary complex with a DNA polymerase. *Proc Natl Acad Sci U S A*, **101**, 2265-2269.
- Livneh, Z., Ziv, O. and Shachar, S. (2010) Multiple two-polymerase mechanisms in mammalian translesion DNA synthesis. *Cell Cycle*, **9**, 729-735.
- Matsuda, T., Bebenek, K., Masutani, C., Hanaoka, F. and Kunkel, T.A. (2000) Low fidelity DNA synthesis by human DNA polymerase-eta. *Nature*, **404**, 1011-1013.
- McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C. and Read, R.J. (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr*, **61**, 458-464.
- Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2005) Human DNA polymerase iota incorporates dCTP opposite template G via a G.C + Hoogsteen base pair. *Structure*, **13**, 1569-1577.
- Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2006) Hoogsteen base pair formation promotes synthesis opposite the 1,N6-ethenodeoxyadenosine lesion by human DNA polymerase iota. *Nat Struct Mol Biol*, **13**, 619-625.
- Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2006) An incoming nucleotide imposes an anti to syn conformational change on the templating purine in the human DNA polymerase-iota active site. *Structure*, **14**, 749-755.
- Nair, D.T., Johnson, R.E., Prakash, S., Prakash, L. and Aggarwal, A.K. (2004) Replication by human DNA polymerase-iota occurs by Hoogsteen base-pairing. *Nature*, **430**, 377-380.
- Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307-326.
- Painter, J. and Merritt, E. (2006) TLSMD web server for the generation of multi-group TLS models. *J. Appl. Crystallogr.*, **39**, 109-111.
- Pfeifer, G.P., Denissenko, M.F., Olivier, M., Tretyakova, N., Hecht, S.S. and Hainaut, P. (2002) Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene*, **21**, 7435-7451.
- Raychoudhury, S.S. and Kubinski, D. (2003) Polycyclic aromatic hydrocarbon-induced cytotoxicity in cultured rat Sertoli cells involves differential apoptotic response. *Environ Health Perspect*, **111**, 33-38.

- Rogozin, I.B., Pavlov, Y.I., Bebenek, K., Matsuda, T. and Kunkel, T.A. (2001) Somatic mutation hotspots correlate with DNA polymerase eta error spectrum. *Nat Immunol*, **2**, 530-536.
- Samet, J.M., Dominici, F., Curriero, F.C., Coursac, I. and Zeger, S.L. (2000) Fine particulate air pollution and mortality in 20 U.S. cities, 1987-1994. *N Engl J Med*, **343**, 1742-1749.
- Scheepers, P., Velders, D., Martens, M., Noordhoek, J. and Bos, R. (1994) Gas-chromatographic mass-spectrometric determination of nitro polycyclic aromatic hydrocarbons in airborne particulate matter from workplace atmospheres contaminated with diesel exhaust. *J. Chromatogr. A*, **677**, 107-121.
- Shachar, S., Ziv, O., Avkin, S., Adar, S., Wittschieben, J., Reissner, T., Chaney, S., Friedberg, E.C., Wang, Z., Carell, T., Geacintov, N. and Livneh, Z. (2009) Two-polymerase mechanisms dictate error-free and error-prone translesion DNA synthesis in mammals. *Embo J*, **28**, 383-393.
- Shen, X., Sayer, J.M., Kroth, H., PotÈn, I., O'Donnell, M., Woodgate, R., Jerina, D.M. and Goodman, M.F. (2002) Efficiency and Accuracy of SOS-induced DNA Polymerases Replicating Benzo[a]pyrene-7,8-diol 9,10-Epoxy A and G Adducts. *J. Biol. Chem.*, **277**, 5265-5274.
- Sherrer, S., Brown, J., Pack, L., Jasti, V., Fowler, J., Basu, A. and Suo, Z. (2009) Mechanistic Studies of the Bypass of a Bulky Single-base Lesion Catalyzed by a Y-family DNA Polymerase. *J. Biol. Chem.*, **284**, 6379-6388.
- Tissier, A., McDonald, J.P., Frank, E.G. and Woodgate, R. (2000) poliota, a remarkably error-prone human DNA polymerase. *Genes Dev*, **14**, 1642-1650.
- Tithof, P.K., Elgayyar, M., Cho, Y., Guan, W., Fisher, A.B. and Peters-Golden, M. (2002) Polycyclic aromatic hydrocarbons present in cigarette smoke cause endothelial cell apoptosis by a phospholipase A2-dependent mechanism. *Faseb J*, **16**, 1463-1464.
- Vaisman, A., Ling, H., Woodgate, R. and Yang, W. (2005) Fidelity of Dpo4: effect of metal ions, nucleotide selection and pyrophosphorolysis. *EMBO J.*, **24**, 2957-2967.
- Vogel, U., Thein, N., Moller, P. and Wallin, H. (2001) Pharmacological coal tar induces G:C to T:A transversion mutations in the skin of muta mouse. *Pharmacol Toxicol*, **89**, 30-34.
- Wang, Y., Woodgate, R., McManus, T.P., Mead, S., McCormick, J.J. and Maher, V.M. (2007) Evidence that in xeroderma pigmentosum variant cells, which lack DNA polymerase eta, DNA polymerase iota causes the very high frequency and unique spectrum of UV-induced mutations. *Cancer Res*, **67**, 3018-3026.

- Watt, D.L., Utzat, C.D., Hilario, P. and Basu, A.K. (2007) Mutagenicity of the 1-nitropyrene-DNA adduct N-(deoxyguanosin-8-yl)-1-aminopyrene in mammalian cells. *Chem Res Toxicol*, **20**, 1658-1664.
- Wong, J.H., Fiala, K.A., Suo, Z. and Ling, H. (2008) Snapshots of a Y-family DNA polymerase in replication: substrate-induced conformational transitions and implications for fidelity of Dpo4. *J Mol Biol*, **379**, 317-330.
- Yang, W. and Woodgate, R. (2007) What a difference a decade makes: insights into translesion DNA synthesis. *Proc Natl Acad Sci U S A*, **104**, 15591-15598.
- Yoon, J.H., Smith, L.E., Feng, Z., Tang, M., Lee, C.S. and Pfeifer, G.P. (2001) Methylated CpG dinucleotides are the preferential targets for G-to-T transversion mutations induced by benzo[a]pyrene diol epoxide in mammalian cells: similarities with the p53 mutation spectrum in smoking-associated lung cancers. *Cancer Res*, **61**, 7110-7117.
- Zhang, Y., Yuan, F., Wu, X., Rechkoblit, O., Taylor, J.S., Geacintov, N.E. and Wang, Z. (2000) Error-prone lesion bypass by human DNA polymerase eta. *Nucleic Acids Res*, **28**, 4717-4724.
- Zhang, Y., Yuan, F., Wu, X., Wang, M., Rechkoblit, O., Taylor, J.S., Geacintov, N.E. and Wang, Z. (2000) Error-free and error-prone lesion bypass by human DNA polymerase kappa in vitro. *Nucleic Acids Res*, **28**, 4138-4146.
- Zhang, Y., Yuan, F., Xin, H., Wu, X., Rajpal, D.K., Yang, D. and Wang, Z. (2000) Human DNA polymerase kappa synthesizes DNA with extraordinarily low fidelity. *Nucleic Acids Res*, **28**, 4147-4156.
- Zhao, B., Wang, J., Geacintov, N.E. and Wang, Z. (2006) Poleta, Polzeta and Rev1 together are required for G to T transversion mutations induced by the (+)- and (-)-trans-anti-BPDE-N2-dG DNA adducts in yeast cells. *Nucleic Acids Res*, **34**, 417-425.

Chapter 6

6 Discussion

6.1 Summary

Our initial studies focused on the model Y family DNA polymerase Dpo4 and how this enzyme discriminates against ribonucleotides (NTP's). Using site-directed mutagenesis and X-ray crystallography, the structural basis of NTP discrimination was revealed. A single conserved aromatic residue in the finger domain was shown to act as a steric gate by clashing with the 2'-OH group of NTP ribose sugars. In addition, this residue was shown to be critical for optimal enzymatic efficiency.

Our next study centred on the error-prone activity of polt opposite undamaged template thymine bases. Through X-ray crystallography and domain swapping experiments, we elucidated the structural mechanism of preferential G mis-incorporation opposite a thymine base and thymine induced replication stalling. These experiments revealed that the finger domain of polt is directly responsible for the unique nucleotide selection and enzymatic activity opposite template thymine. The finger domain of polt induces error-prone thymine replication by creating a unique active site that governs the hydrogen bonding and base stacking potential of incoming nucleotides.

It was then investigated how polt promotes error-free replication opposite the highly mutagenic 8-oxo-guanine lesion. Using X-ray crystallography and site-directed mutagenesis, the finger domain of polt was again implicated in nucleotide incorporation specificity. However, unlike error-prone thymine replication, the finger domain promotes

error-free replication opposite 8-oxo-guanine. This study gives insight into how polt may contribute to cellular protection against oxidizing agents.

The last study demonstrated how nucleotide preference is determined by polt opposite large and bulky DNA lesions. It was revealed that polt can utilize alternate replication mechanisms depending on the nature of the DNA lesion. The bulky and hydrophobic aminopyrene lesion induced A mis-incorporations due to strong base stacking potential with the A base and inhibited DNA replication through protein-DNA lesion interactions. Thus, the finger domain of polt influences nucleotide selection opposite bulky lesions by dictating the template lesion conformation.

6.2 Mechanistic insights

6.2.1 Incoming nucleotide selection

The structural studies of polt in ternary complex with various template bases indicate that incoming nucleotide selection in Y family DNA polymerases is determined by two main forces: 1) hydrogen bonding potential of the incoming nucleotide with the template base, and 2) base stacking potential of the incoming nucleotide with the underlying base pair. The nucleotide with the strongest hydrogen bonding and base stacking potential is incorporated with the highest efficiency, regardless of whether this nucleotide is the correct match. The finger domain influences these two potentials by forming the protein-replicating base pair contact interface, which determines nucleotide positioning. For polt, the finger domain creates a narrow active site and contains large side chain residues which project the template base out towards the solvent exposed major groove. These

two factors were shown to have profound effects on incoming nucleotide specificity. In addition, the finger domain residue glutamine 59 is critical for polt enzymatic activity by creating unique protein-template base interactions. For other Y family polymerases, such as yeast pol η , the finger domain has also been shown to play a role in incoming nucleotide incorporation(Alt et al., 2007).

The insights into nucleotide selection gleaned from the polt structures reveal why Y family DNA polymerases have such high error-rates of replication compared to replicative polymerases. While Y family polymerases incorporate nucleotides based on hydrogen bonding and base stacking interactions alone, replicative polymerases introduce another level of selection by analyzing the replicating base pair geometry. The finger domains of replicative polymerases interact with the minor groove side of the replicating base pair(Hsu et al., 2004). If the minor groove geometry is planar, this indicates Watson-Crick base pairing and results in nucleotide incorporation. If the minor groove is non-planar, the base pair is non-Watson-Crick, and the nucleotide is rejected. Since Y family polymerases cannot select nucleotides based on pairing geometry, mis-matched nucleotides can be efficiently incorporated. In addition, the large and solvent accessible active site of Y family polymerases allow multiple nucleotide conformations to exist, which further enhances mis-matched base pairing potential. This error-prone replication likely introduces somatic mutagenesis within the genome and contributes to cancer formation and aging.

6.2.2 Undamaged DNA replication

Interestingly, our work on polt demonstrates that undamaged DNA replication is more mutagenic than lesion bypass. For example, undamaged thymine replication is error-

prone while 8-oxo-guanine replication is error-free. Similar observations of error-prone undamaged replication have also been noted for pol η (Matsuda et al., 2000). Thus, undamaged DNA replication by Y family polymerases may threaten the integrity of an organism's genome. The extent to which Y family polymerases replicate undamaged DNA *in vivo* is currently unknown. Although it is clear that pol η is involved in undamaged error-prone replication causing somatic hypermutation of immunoglobulin genes(Zeng et al., 2001), whether other Y family polymerases also contribute to this process remains controversial(Faili et al., 2002; Shimizu et al., 2005). Interestingly, during the process of lesion bypass Y family polymerases likely encounter undamaged DNA nucleotides upstream and downstream from the lesion site(Northam et al.). Thus, error-prone undamaged DNA replication may occur during TLS even if the lesion is replicated in an error-free manner. Indeed, it has been shown that nucleotides flanking DNA lesions are hot spots for mutation(McCulloch et al., 2004), indicating the mutagenicity of the TLS pathway despite the accuracy of lesion bypass.

6.2.3 Lesion bypass

Although lesion bypass mechanisms vary between Y family polymerases and are highly dependent on the size and chemical nature of the lesion, common themes in bypass mechanisms are prevalent. For small DNA lesions, it appears Y family polymerases replicate through the lesion site in a similar manner to undamaged DNA. For pol ι , the small 8-oxo-guanine lesion was replicated similar to an undamaged G, with the Hoogsteen edge involved in base pairing. Other Y family polymerases have been shown to replicate small lesions using mechanisms identical to that of undamaged DNA replication(Rechkoblit et al., 2009; Rechkoblit et al., 2006). These results indicate that

most small DNA lesions cannot be recognized by Y family polymerases. This may be true for all DNA polymerases as even replicative polymerases can bypass small DNA lesions(Sabouri et al., 2008). Although Y family polymerases are not required for the bypass of most small DNA lesions, they are likely needed for error-free replication under certain situations, such as polt during oxidative lesion BER(Petta et al., 2008; Prasad et al., 2003).

The structural study of polt in complex with the bulky and hydrophobic aminopyrene lesion indicates that the solvent exposed major groove is critical for accommodating and replicating large DNA adducts. Even though the polt active site is relatively restrictive for a Y family polymerase, the major groove is able to accommodate multiple conformations of the bulky adduct. This permissive nature of the polt active site is fundamental to the bypass of the bulky aminopyrene lesion. Other Y family polymerases have also been shown to sample multiple conformations of bulky adducts in order to identify the most productive orientation to facilitate bypass(Bauer et al., 2007; Ling et al., 2004).

Although polt is able to incorporate nucleotides opposite DNA lesions, replication is stalled after a single nucleotide is incorporated. This indicates that polt must require additional polymerases for the extension reaction during TLS. In addition, this stalling mechanism may recruit other DNA modifying enzymes such as DNA ligase(Tomkinson et al., 2001) during 8-oxo-guanine BER. The stalling signature of polt may also serve as a form of auto-regulation in order to prevent high rates of error-prone replication past the DNA lesion. Other Y family polymerases such as polk also display lesion stalling opposite particular DNA adducts(Zhang et al., 2000), however the effect is lesion specific

and less dramatic than polt. Thus, it appears that polt is likely involved in a unique or specialized form of TLS.

6.2.4 Rational drug design

The ability of Y family polymerases to replicate through cross-linked nucleotides is a major contributor to cancer cell resistance against replication-arresting chemotherapeutic agents such as Cisplatin(Albertella et al., 2005). These drugs impart their anti-cancer effects by stalling replication in rapidly dividing cells, which is a common characteristic of cancer cells. Our structural and biochemical studies of aminopyrene lesion bypass indicate that hydrophobic ring systems are highly inhibitory to Y family polymerase replication. Thus, adding hydrophobic ring attachments to DNA adduct inducing chemotherapeutics may greatly increase their efficacy by preventing Y family polymerase bypass. An alternate route to improve replication arresting chemotherapeutic agents is to combine the drug with a Y family polymerase specific inhibitor.

The little finger domain appears to be a promising target for Y family polymerase inhibitors for a number of reasons. First, the little finger domain is only found in Y family polymerases, thus, drugs which bind this unique domain may have little cross-reactivity with other polymerases and enzymes. Second, the little finger domain is critical for Y family polymerase activity and lesion bypass(Boudsocq et al., 2004; Ling et al., 2001). Inhibitors of this domain could target two main areas: The linker region and the DNA binding interface. The linker region connects the little finger to the thumb domain of the polymerase and is required for large conformational movements which mediate DNA substrate binding and lesion bypass(Wong et al., 2008). Thus, inhibiting the movement of this linker region would inhibit Y family polymerase activity. Another

potential site to target is the interface between the little finger and the major groove of the DNA substrate. Designing drugs to interact with this region of the little finger would inhibit DNA substrate binding and prevent lesion bypass.

In addition to inhibiting Y family polymerase activity, altering the function or nucleotide specificity of these enzymes would be highly desirable. For instance, a drug that would increase the fidelity of polt replication opposite TT dimers could be a viable treatment option for XP-V syndrome by reducing DNA mutagenesis induced by polt. The narrow C1'-C1' distance in the polt active site influences error-prone replication opposite thymine bases by causing incoming A nucleotides to adopt an unfavourable *syn* conformation. This unique active site geometry is created through an interaction between the finger and little finger domains. Thus, a drug that could inhibit this interaction could increase the active site dimensions of polt and possibly promote a higher TT dimer fidelity and thus reduce mutagenesis in XP-V patients. However, altering polt fidelity opposite thymine bases could have detrimental effects, as the *in vivo* function of this unique specificity is still unknown.

6.3 Future directions

6.3.1 Function

Further research is required to elucidate the biological role of each Eukaryotic Y family DNA polymerase within the cell. Particularly, we need to understand which enzymes are responsible for different classes of DNA lesion bypass (ie. oxidative lesions, UV light lesions, and environmental pollutant lesions). Although functional overlap may make this type of classification difficult, utilizing a variety of experimental approaches may

help. One approach could be immunoprecipitation of Y family polymerases followed by tryptic digestion and liquid chromatography-mass spectrometry (LC MS/MS) to identify novel binding partners (Vasilescu and Figeys, 2006). For instance pol η has been implicated in BER processes based on its interaction with the DNA repair scaffold protein XRCC1 (Petta et al., 2008). Similar experiments could be performed after cellular exposure to different DNA damaging agents in order to identify replication or repair machinery complexes involved in TLS. In addition, such analysis may identify specific polymerase partners that function concertedly during DNA lesion bypass of specific adducts.

Continued structural studies of Eukaryotic Y family polymerases are invaluable to functional characterization. Structurally comparing Y family polymerases opposite many different forms of DNA damage will give insight into which polymerase has an optimal bypass mechanism for specific lesions. Combining this structural information with *in vitro* activity and fidelity assays will facilitate functional classification by matching these results with cellular TLS activity and mutagenic profiles. For instance, cellular exposure to aminopyrene adducts results in replication stalling and G to T transversions (Watt et al., 2007), which coincides with the structure and functional activity of pol η .

6.3.2 Regulation

Considering the error-prone nature of Y family polymerases and their contribution to DNA mutagenesis and cancer, it is not surprising that Y family polymerase activity must be tightly regulated. Although it is accepted that PCNA mono-ubiquitination coordinates the TLS pathway (Bienko et al.; Terai et al.), exactly how this modification recruits Y family polymerases to the replication fork and enables polymerase switching is unknown.

Future structural studies will need to focus on either full length Y family polymerases or domains of the regulatory region in complex with mono-ubiquitinated PCNA. Such complexes would give great insight into how mono-ubiquitination can regulate Y family polymerase activity and lesion bypass. Another important structural study is to investigate how the regulatory region of different Y family polymerases interacts and how this enables multi-polymerase complex formation and co-localization to replication foci.

Given that multiple Y family polymerases can be present at a single DNA lesion, it remains a mystery how the cell determines which polymerase to use opposite a given lesion. Perhaps the cell resolves this dilemma by substrate competition, whereby the polymerase with the highest binding and bypass activity is selected for lesion bypass. In addition, it has been shown that Y family polymerases themselves are modified by ubiquitin(Bienko et al.) and the small ubiquitin-like modifier (SUMO)(Kim and Michael, 2008), which influences lesion bypass. Thus, post-translational modification appears to be a key regulator of the TLS pathway and could be a promising area in future research in Y family DNA polymerases.

6.4 Conclusions

This work describes the first detailed structural mechanism of Y family DNA polymerase nucleotide selection during different DNA replication contexts. The finger domain is directly involved in nucleotide selection by rejecting incoming NTP's and influencing hydrogen bonding and base stacking potentials of incoming dNTP's. In addition to nucleotide selection, the finger domain allows the accommodation of bulky DNA adducts

via the solvent exposed major groove, facilitating lesion bypass. These results shed light on how Y family polymerase activity promotes cellular tolerance to DNA damaging agents and its contribution to carcinogenicity, aging, and evolution. Furthermore, our structural insights into Y family polymerase activity will aid in the design of novel drugs to inhibit or functionally alter these potentially therapeutic targets.

6.5 References

- Albertella, M.R., Green, C.M., Lehmann, A.R. and O'Connor, M.J. (2005) A role for polymerase eta in the cellular tolerance to cisplatin-induced damage. *Cancer Res*, **65**, 9799-9806.
- Alt, A., Lammens, K., Chiocchini, C., Lammens, A., Pieck, J.C., Kuch, D., Hopfner, K.P. and Carell, T. (2007) Bypass of DNA lesions generated during anticancer treatment with cisplatin by DNA polymerase eta. *Science*, **318**, 967-970.
- Bauer, J., Xing, G., Yagi, H., Sayer, J.M., Jerina, D.M. and Ling, H. (2007) A structural gap in Dpo4 supports mutagenic bypass of a major benzo[a]pyrene dG adduct in DNA through template misalignment. *Proc Natl Acad Sci U S A*, **104**, 14905-14910.
- Bienko, M., Green, C.M., Sabbioneda, S., Crosetto, N., Matic, I., Hibbert, R.G., Begovic, T., Niimi, A., Mann, M., Lehmann, A.R. and Dikic, I. (2010) Regulation of translesion synthesis DNA polymerase eta by monoubiquitination. *Mol Cell*, **37**, 396-407.
- Boudsocq, F., Kokoska, R.J., Plosky, B.S., Vaisman, A., Ling, H., Kunkel, T.A., Yang, W. and Woodgate, R. (2004) Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication. *J Biol Chem*, **279**, 32932-32940.
- Faili, A., Aoufouchi, S., Flatter, E., Gueranger, Q., Reynaud, C.A. and Weill, J.C. (2002) Induction of somatic hypermutation in immunoglobulin genes is dependent on DNA polymerase iota. *Nature*, **419**, 944-947.
- Hsu, G.W., Ober, M., Carell, T. and Beese, L.S. (2004) Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase. *Nature*, **431**, 217-221.
- Kim, S.H. and Michael, W.M. (2008) Regulated proteolysis of DNA polymerase eta during the DNA-damage response in *C. elegans*. *Mol Cell*, **32**, 757-766.

- Ling, H., Boudsocq, F., Woodgate, R. and Yang, W. (2001) Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. *Cell*, **107**, 91-102.
- Ling, H., Sayer, J.M., Plosky, B.S., Yagi, H., Boudsocq, F., Woodgate, R., Jerina, D.M. and Yang, W. (2004) Crystal structure of a benzo[a]pyrene diol epoxide adduct in a ternary complex with a DNA polymerase. *Proc Natl Acad Sci U S A*, **101**, 2265-2269.
- Matsuda, T., Bebenek, K., Masutani, C., Hanaoka, F. and Kunkel, T.A. (2000) Low fidelity DNA synthesis by human DNA polymerase-eta. *Nature*, **404**, 1011-1013.
- McCulloch, S.D., Kokoska, R.J., Masutani, C., Iwai, S., Hanaoka, F. and Kunkel, T.A. (2004) Preferential cis-syn thymine dimer bypass by DNA polymerase eta occurs with biased fidelity. *Nature*, **428**, 97-100.
- Northam, M.R., Robinson, H.A., Kochenova, O.V. and Shcherbakova, P.V. (2009) Participation of DNA polymerase zeta in replication of undamaged DNA in *Saccharomyces cerevisiae*. *Genetics*, **184**, 27-42.
- Petta, T.B., Nakajima, S., Zlatanou, A., Despras, E., Couve-Privat, S., Ishchenko, A., Sarasin, A., Yasui, A. and Kannouche, P. (2008) Human DNA polymerase iota protects cells against oxidative stress. *Embo J*, **27**, 2883-2895.
- Prasad, R., Bebenek, K., Hou, E., Shock, D.D., Beard, W.A., Woodgate, R., Kunkel, T.A. and Wilson, S.H. (2003) Localization of the deoxyribose phosphate lyase active site in human DNA polymerase iota by controlled proteolysis. *J Biol Chem*, **278**, 29649-29654.
- Rechkoblit, O., Malinina, L., Cheng, Y., Geacintov, N.E., Broyde, S. and Patel, D.J. (2009) Impact of conformational heterogeneity of OxoG lesions and their pairing partners on bypass fidelity by Y family polymerases. *Structure*, **17**, 725-736.
- Rechkoblit, O., Malinina, L., Cheng, Y., Kuryavyi, V., Broyde, S., Geacintov, N.E. and Patel, D.J. (2006) Stepwise translocation of Dpo4 polymerase during error-free bypass of an oxoG lesion. *PLoS Biol*, **4**, e11.
- Sabouri, N., Viberg, J., Goyal, D.K., Johansson, E. and Chabes, A. (2008) Evidence for lesion bypass by yeast replicative DNA polymerases during DNA damage. *Nucleic Acids Res*, **36**, 5660-5667.
- Shimizu, T., Azuma, T., Ishiguro, M., Kanjo, N., Yamada, S. and Ohmori, H. (2005) Normal immunoglobulin gene somatic hypermutation in Pol kappa-Pol iota double-deficient mice. *Immunol Lett*, **98**, 259-264.
- Terai, K., Abbas, T., Jazaeri, A.A. and Dutta, A. (2010) CRL4(Cdt2) E3 ubiquitin ligase monoubiquitinates PCNA to promote translesion DNA synthesis. *Mol Cell*, **37**, 143-149.

- Tomkinson, A.E., Chen, L., Dong, Z., Leppard, J.B., Levin, D.S., Mackey, Z.B. and Motycka, T.A. (2001) Completion of base excision repair by mammalian DNA ligases. *Prog Nucleic Acid Res Mol Biol*, **68**, 151-164.
- Vasilescu, J. and Figeys, D. (2006) Mapping protein-protein interactions by mass spectrometry. *Curr Opin Biotechnol*, **17**, 394-399.
- Watt, D.L., Utzat, C.D., Hilario, P. and Basu, A.K. (2007) Mutagenicity of the 1-nitropyrene-DNA adduct N-(deoxyguanosin-8-yl)-1-aminopyrene in mammalian cells. *Chem Res Toxicol*, **20**, 1658-1664.
- Wong, J.H., Fiala, K.A., Suo, Z. and Ling, H. (2008) Snapshots of a Y-family DNA polymerase in replication: substrate-induced conformational transitions and implications for fidelity of Dpo4. *J Mol Biol*, **379**, 317-330.
- Zeng, X., Winter, D.B., Kasmer, C., Kraemer, K.H., Lehmann, A.R. and Gearhart, P.J. (2001) DNA polymerase eta is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. *Nat Immunol*, **2**, 537-541.
- Zhang, Y., Yuan, F., Wu, X., Wang, M., Rechkoblit, O., Taylor, J.S., Geacintov, N.E. and Wang, Z. (2000) Error-free and error-prone lesion bypass by human DNA polymerase kappa in vitro. *Nucleic Acids Res*, **28**, 4138-4146.

Appendices

Appendix 1: Copyright approval for previously published work (Chapter 3)

Rightslink® by Copyright Clearance Center

Page 1 of 1



Title: Structural basis of error-prone replication and stalling at a thymine base by human DNA polymerase I

Author: Kevin N Kirouac, Hong Ling

Publication: The EMBO Journal

Publisher: Nature Publishing Group

Date: May 14, 2009

Copyright © 2009, Nature Publishing Group

Logged in as:
Kevin Kirouac

LOGOUT

Author Request

If you are the author of this content (or his/her designated agent) please read the following. If you are not the author of this content, please click the Back button and select an alternative [Requestor Type](#) to obtain a quick price or to place an order.

Ownership of copyright in the article remains with the Authors, and provided that, when reproducing the Contribution or extracts from it, the Authors acknowledge first and reference publication in the Journal, the Authors retain the following non-exclusive rights:

- a) To reproduce the Contribution in whole or in part in any printed volume (book or thesis) of which they are the author(s).
- b) They and any academic institution where they work at the time may reproduce the Contribution for the purpose of course teaching.
- c) To reuse figures or tables created by them and contained in the Contribution in other works created by them.
- d) To post a copy of the Contribution as accepted for publication after peer review (in Word or Tex format) on the Author's own web site, or the Author's institutional repository, or the Author's funding body's archive, six months after publication of the printed or online edition of the Journal, provided that they also link to the Journal article on NPG's web site (eg through the DOI).

NPG encourages the self-archiving of the accepted version of your manuscript in your funding agency's or institution's repository, six months after publication. This policy complements the recently announced policies of the US National Institutes of Health, Wellcome Trust and other research funding bodies around the world. NPG recognises the efforts of funding bodies to increase access to the research they fund, and we strongly encourage authors to participate in such efforts.

Authors wishing to use the published version of their article for promotional use or on a web site must request in the normal way.

If you require further assistance please read NPG's online [author reuse guidelines](#).

BACK

Copyright © 2010 Copyright Clearance Center, Inc. All Rights Reserved. [Privacy statement](#).
Comments? We would like to hear from you. E-mail us at customer@copyright.com

Curriculum Vitae

Name: Kevin Kirouac

Post-secondary Education and Degrees: University of Western Ontario
London, Ontario, Canada
2001-2005 B.M.Sc Honors Biochemistry

The University of Western Ontario
London, Ontario, Canada
2005-2010 Ph.D. Biochemistry

Honours and Awards: Margaret P. Moffat Research Award
University of Western Ontario
2009

Schulich Graduate Scholarship
University of Western Ontario
2007-2010

Graduate Research Scholarship
University of Western Ontario
2005-2007

Publications:

- Kirouac K**, Ling H. (2009) Structural basis of error-prone replication and stalling at a thymine base by Human DNA polymerase ι . *EMBO J.* 28: 1644-54
- Xing G, **Kirouac K**, Shin YJ, Bell SD, Ling H. (2009) Structural insight into recruitment of translesion DNA polymerase Dpo4 to sliding clamp PCNA. *Mol Microbiol.* 71: 678-91

Selected conference presentations:

- “Error-free replication opposite an oxidative DNA lesion by a human Y family DNA polymerase.” American Crystallographic Association, July 25-30 2010, Chicago, Illinois, USA. *Oral presentation*
- “Structural insights into error-prone replication and stalling opposite a template thymine base by a human DNA polymerase.” American Crystallographic Association, July 25-30 2009, Toronto, Canada. *Oral presentation*
- “Structural insights into T template stalling and nucleotide mis-incorporation by human DNA polymerase ι .” 12th Annual DNA Replication and Repair Symposium, June 28 2008, Buffalo, Michigan, U.S.A. *Oral presentation.*