ABSTRACT

SUAREZ, SAMUEL CHARLES. Molecular Determinants of Human DNA Polymerase η Fidelity. (Under the direction of Dr. Scott D. McCulloch).

DNA damage is a ubiquitous challenge to replication in cells, as damage causes replicative polymerase stalling. However, once DNA has been unwound at the replication fork, replication must proceed in the presence of damage to prevent more deleterious and almost assuredly mutagenic consequences. Alleviation of replicative polymerase stalling is accomplished by specialized DNA polymerases that can synthesize across from DNA lesions using the damage as a template, a process termed translesion synthesis (TLS). DNA polymerase η (pol η) is the best understood of these polymerases, and lack of pol η synthesis activity in the human cancer prone syndrome Xeroderma pigmentosum variant (XPV) leads to cancer susceptibility upon sunlight exposure. XPV cells display higher mutation rates when exposed to UV light. This prevention of mutagenesis occurs despite pol η having fidelity that is thousands of fold lower than replicative polymerases when copying both damaged and undamaged DNA. Pol η has been implicated in replication past the UV induced *cis-syn* cyclobutane pyrimidine dimers (CPD) and the oxidative lesion 7,8-dihydro-8-oxo-guanine (8-oxoG) in cells. We sought to better understand the molecular basis of efficient but moderate to low fidelity bypass by pol η. We

have examined polymerase properties as well as replication fidelity opposite these 2 lesions and with undamaged DNA.

To this end, we have created and purified a set of single amino acid substitution mutants in and surrounding the active site of the protein, utilizing the truncated catalytic core of the protein as a model. We assessed these mutants for overall synthesis activity as well as bypass fidelity opposite both T-T CPD and 8-oxoG. Our results show that several residues are critical for polymerase function, and altering these amino acids have multiple effects on polymerase properties. The R55A mutant abolishes polymerase activity while the Q38A, Y52E, and R61A mutants display altered fidelities. Y52E increased fidelity at both lesions and undamaged DNA, while R61A increased fidelity when copying T-T CPD. Also notable, Q38A increased fidelity opposite 8-oxoG, while it decreased fidelity opposite a T-T CPD.

One proposed means of increasing pol η fidelity is interaction with replication accessory proteins that assist in replication at the replication fork. We purified the full-length form of pol η, containing known protein:protein interaction domains in the C-terminus, and examined the effect of adding RPA to the bypass reaction. We saw no change in fidelity when examining fidelity opposite T-T CPD or 8-oxoG. We sought to confirm our results by also expressing two previously identified mutants with specific fidelity signatures, Q38A and Y52E. These full-

length mutants recapitulated the fidelity effects seen in the truncated mutants when copying damaged DNA, and these fidelity signatures were unchanged with the addition of RPA.

Taken together, these results indicate that the major determinant of pol η fidelity is the active site structure of the protein. The active site sequence is robust and certain amino acids play a critical role in the molecular mechanism of synthesis by the enzyme. Further clues as to the effects of altered polymerase function could be addressed by experiments expressing these mutant proteins in cells lacking pol η. Additional investigation is necessary to recapitulate a more complete set of proteins with known functions in TLS, as interaction with other proteins could possibly alter fidelity. This work emphasizes that TLS is a damage tolerance process that could potentially cause mutations if perturbed. In order to avoid the certainly mutagenic consequences of strand breaks, cells utilize damage tolerance at the cost of potential mutagenesis. This balance between tolerance and mutagenesis has implications for multiple disease processes and human health.

Molecular Determinants of Human DNA Polymerase η Fidelity

by
Samuel Charles Suarez

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Toxicology

Raleigh, North Carolina

2014

APPROVED BY:

_____	_____
Scott D. McCulloch	Robert C. Smart
Committee Chair	Vice-Chair
_____	_____
Yoshiaki Tsuji	David W. Threadgill

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

DEDICATION

For my parents, Joe and Marty, and my wife, Danielle.

BIOGRAPHY

Sam Suarez was born on March 5, 1984 in Long Branch, New Jersey. After spending his first year of life exploring Manhattan, his parents relocated to central Florida, finally settling in the thriving metropolis of DeBary. After graduating Deltona High School in 2002, he received a scholarship to study at Tulane University in New Orleans, LA. At Tulane, Sam chose to major in biological chemistry, while also pursuing a minor in economics. In the early hours of August 28, 2005, Sam was forced to evacuate New Orleans due to the approach of Hurricane Katrina. In the days that followed, it became increasingly clear that no one would be attending school in New Orleans for the foreseeable future. Sam journeyed back to central Florida, working for a time until returning to New Orleans triumphantly on New Year's Eve after 4 months of exile. A research project during his last year as an undergraduate served as his first exposure to independent laboratory science. This project led him to apply to graduate school, eventually leading to his enrollment in Ph.D. program in the department of Toxicology at North Carolina State University.

Sam enjoys spending his time away from work in the outdoors, either riding one of his multiple bikes, or running the trails around the Triangle area. He has completed multiple triathlons and running races while attending graduate school,

including multiple half-marathons, multiple half-ironmans, and the Myrtle Beach marathon.

Sam met his wife, Danielle, in 2009 through mutual friends during a stint participating in an adult kickball league. They married in July of 2013. They reside in Cary, with their sheltie, Bella.

ACKNOWLEDGMENTS

There are so many people to thank, without each of whom this work would not have been possible.

I would like to thank Dr. Scott McCulloch for being an excellent mentor and teacher. Your patience has been greatly appreciated. I was woefully inexperienced in lab science when I began at NC State, and your guidance has helped mold me into a (decently) competent scientist. I had no idea this subject existed when I entered graduate school and exploring it has been an immensely rewarding experience. If he had kept a tab of equipment I have broken in lab, I would have had to work one of these years for free. I don't think he holds it against me (too much).

I would like to thank the members of my committee: Dr. Rob Smart, Dr. Yoshi Tsuji, and Dr. David Threadgill. Each of your perspectives has been invaluable in moving this project forward. I appreciate the time you have taken out of your busy schedules to advise me, and I'm a better scientist today because of your guidance.

I would like to thank each of my compatriots in the McCulloch lab: Shannon Toffton, Renee Beardslee, and Kim Herman. Working alongside each of you has been a privilege, and each of your contributions has been immeasurable. Discussions with each of you helped crystallize my thinking on my project, and my perspective

has broadened because of each of you. Without Shannon, I would have spent an extra year ordering and searching for supplies. Without Renee, I would have lost touch with reality long ago. Without Kim, I would have been discouraged to the point of despair multiple times.

The front office and support staff of the Program in Environmental and Molecular Toxicology have helped me immensely. Thank you to Janet Roe (Go Saints!), Jackie Broughton, and Jeanne Burr.

My parents, Joe and Marty, have been incredibly supportive throughout this process, and I can't thank them enough for all the opportunities they've provided me. Your love, support, and encouragement mean the world to me, and I'm incredibly lucky to have you both as parents. I'm sure they're happier than I am that I've finally finished college.

To my wife Danielle, I think we make a pretty good team. I'm amazed that we had time to fall in love and get married while in I was in graduate school, but somehow we did it. Your support sustains me everyday, and I look forward to moving on to the next step in our life.

TABLE OF CONTENTS

LIST OF TABLES ix

LIST OF FIGURES x

LIST OF SYMBOLS AND ABBREVIATIONS xii

GENERAL INTRODUCTION 1

DNA Replication 1
DNA Damage and Repair 5
Translesion Synthesis 6
Pol η 11
Structure of Pol η 16
Fidelity of Pol η 19
Role of Pol η in processes other than TLS 22
Pol η in vitro versus in vivo 23

RATIONALE 35

CHAPTER 1 – Biochemical analysis of active site mutations of human polymerase η 36

Abstract 37
Introduction 38
Materials and Methods 42
Reagents and materials 42
Expression vector and protein purification 43
Polymerase and lesion bypass assays 46
Forward mutation assay 47
Results 48
Active site mutations of pol η 48
Polymerase activity 49
Lesion bypass efficiency 51

Lesion bypass fidelity 52
Forward Gap Fidelity 55
Discussion 57
Conflict of interest statement 65
Acknowledgements 65
Funding information 65

CHAPTER 2 – Biochemical analysis of DNA polymerase η fidelity in the presence of replication protein A 81

Abstract 82
Introduction 83
Materials and Methods 88
DNA Oligomer Sequences 88
Recombinant Proteins 89
Pol η 89
RPA 92
Lesion bypass fidelity assay 94
DNA binding 95
Results and Discussion 96
Conflict of Interest Statement 107
Acknowledgements 107
Funding information 107

GENERAL DISCUSSION 127

Effects of Single Amino Acid Substitution Mutants 129
Effects of Accessory Proteins 136
Other Factors Possibly Influencing TLS Fidelity 139
Reflections on This Work 141
TLS as a Damage Tolerance Process 144

REFERENCES 147

LIST OF TABLES

CHAPTER 1

Table 1.1 Activity and lesion bypass fidelity for single amino acid mutants of human pol η 66

Table 1.2 Base substitution and insertion/deletion error rates of pol η mutants copying undamaged DNA in a gapped plasmid 77

Supplementary Table 1.1 Primer sequences used in site-directed mutagenesis reactions to generate single amino acid mutants in the catalytic core of human DNA polymerase η (amino acids 1-511) 78

Supplementary Table 1.2 Base substitution error spectrum of truncated pol η mutants as calculated from forward gap filling assay 79

CHAPTER 2

Table 2.1 The effect of RPA on the lesion bypass fidelity of truncated pol η 108

Supplementary Table 2.1 Truncated and full-length pol η comparison 109

LIST OF FIGURES

GENERAL INTRODUCTION

Figure I.1 Simplified model of a eukaryotic replication fork 25

Figure I.2 Components of DNA Replication Fidelity 27

Figure I.3 Selected chemical structures of DNA damage that can be bypassed by pol η 28

Figure I.4 Polymerase switching model of TLS 29

Figure I.5 Models of translesion synthesis 31

Figure I.6 Domain structure of pol η 34

CHAPTER 1

Figure 1.1 Location of amino acids altered in this study within the active site of human pol η 68

Figure 1.2 Denaturing polyacrylamide gel electrophoresis based separation of lesion bypass efficiency assay reaction products 70

Figure 1.3 Lesion bypass efficiency by pol η mutants 71

Figure 1.4 Error rate when copying a *cis-syn* cyclobutane thymine-thymine dimer by pol η mutants 73

Figure 1.5 Error rate when copying an 8-oxoG lesion by pol η mutants .. 75

CHAPTER 2

Figure 2.1 RPA binding to substrates used in the lesion bypass fidelity assay 110

Figure 2.2 Substrate Diagrams 112

Figure 2.3 Wild type, full-length pol η damaged error rates 113

Figure 2.4 Wild type, Q38A, and Y52E damaged error rates 115

Figure 2.5 Wild Type, Q38A, and Y52E undamaged error rates 117

Supplementary Figure 2.1 Sypro-Red stained protein gels of purified proteins 119

Supplementary Figure 2.2 Colorless plaque frequency 121

Supplementary Figure 2.3 Complex error rate (10^{-4}) 123

Supplementary Figure 2.4 Pol η primer extension in the presence or absence of RPA 125

LIST OF SYMBOLS AND ABBREVIATIONS

α – alpha

δ – delta

ε – epsilon

η – eta

ζ – zeta

ι – iota

κ – kappa

λ – lambda

θ – theta

ν – nu

6-4 PP – pyrimidine (6-4) pyrimidone photoproducts

8-oxoG – 7,8-dihydro-8-oxo-guanine

ATR – ATM and Rad3 related protein kinase

ATRIP – ATR interacting protein

CPD – *cis-syn* cyclobutane pyrimidine dimers

dNTPs – deoxynucleotide triphophosphates

indel – insertion/deletion

MYH – Mut Y homolog

NER – nucleotide excision repair

PCNA – proliferating cell nuclear antigen

PIP – PCNA interacting peptide

pol – polymerase

pol η – DNA polymerase η

RFC – replication factor C

RIR – Rev1 interacting region

RPA – replication protein A

SNP – single nucleotide polymorphisms

SSB – single stranded binding protein

ssDNA – single-stranded DNA

T-T CPD – *cis-syn* thymine-thymine cyclobutane pyrimidine dimer

TLS – translesion synthesis

UBM – ubiquitin binding motif

UBZ – ubiquitin binding zinc finger

UV – ultraviolet light

XP – Xeroderma pigmentosum

XPV – Xeroderma pigmentosum variant

GENERAL INTRODUCTION

Duplication of genetic material is essential to all organisms, from prokaryotes to complex multicellular eukaryotes. All known organisms on Earth (excluding some viruses) utilize DNA as their genetic material, and the task of replication is accomplished at the molecular level by enzymes classified as DNA polymerases. Humans are known to have 16 different DNA polymerases, each responsible for a varying portion of the goal of accurate copying of the genome.[1] More than 60 years has passed since Watson and Crick first described the molecular structure of DNA; and in their paper they stated "It has not escaped our notice that the specific pairing we have postulated immediately suggests a copying mechanism for the genetic material."[2] In the years since this postulation, we have come to understand the complexity inherent in the processes necessary to preserve the sequence of the genome, while also allowing for the genetic variation necessary for life to adapt.[1,3]

DNA Replication

Replication of the majority of eukaryotic nuclear DNA is accomplished by two B-family replicative DNA polymerases: pol δ, and pol ε.[1] The polymerases catalyze chain elongation by linking the 5' α-phosphate of an incoming nucleotide precursor onto the 3' hydroxyl group of the previous nucleotide in the existing

chain. Thus, DNA is always synthesized in cells in a 5'→3' fashion. The two DNA strands pair in an anti-parallel manner (i.e. one is 5'→3', the other 3'→5') and chromosomal DNA replication takes place on both strands simultaneously in the same direction. This requires two modes of replication. The two strands of the fork are copied with one strand, the leading strand, occurring continuously for stretches of thousands of bases and replication of the opposite strand must occur in a discontinuous fashion, in shorter stretches termed Okazaki fragments. [4] Despite this knowledge, only recently have experiments begun to elucidate which of the replicative polymerases perform each of these functions. [5] In the current model, Pol ε performs continuous, leading strand synthesis,[6,7] while pol δ performs discontinuous, lagging strand synthesis.[7-11] These polymerases are assisted by a multitude of replication accessory proteins, each of which contributes to the task of complete genome duplication.[12]

Copying of the DNA requires more than polymerization of nucleotide chains. The DNA is first unwound by the MCM2-7 helicase complex.[13] The heterotrimeric single stranded binding protein replication protein A (RPA) binds this unwound DNA to prevent re-annealing.[14] Synthesis of RNA-DNA hybrid primers, from which pol δ and ε can initiate copying is accomplished by pol α.[1] The five subunit clamp

loading complex replication factor C (RFC) loads the homotrimeric sliding clamp proliferating cell nuclear antigen (PCNA) onto DNA[15]. Once PCNA is loaded onto primer termini, synthesis proceeds and PCNA:polymerase interaction functions to increase the processivity of both pol δ and pol ε (See Waga and Stillman[15] and references therein). The primase, helicase, sliding clamp, and single stranded binding proteins perform their functions at and surrounding the replication fork to assist the replicative polymerases in the monumental task of copying chromosomal DNA.[16] A schematic model of the proposed replication fork is shown in Figure I.1.

Estimates of the fidelity of DNA replication indicate that there is 1 error per 10^9 to 10^{10} bases copied, which equates to approximately 1 error per human genome duplication.[17] This high fidelity is a combination of multiple factors, including (but not limited to): free energy differences between correct Watson-Crick base pairing and incorrect base pairing, replicative polymerase selectivity, proofreading exonuclease activity of the replicative polymerases, and the mismatch repair system (see Loeb[18] and McCulloch[3] and references therein). These factors are represented in Figure I.2. These factors work in concert to ensure that the genome is copied accurately, and genetic information is passed on.

Free energy differences between correct and incorrect base pairs can only explain a small amount of the overall fidelity of replication.[19] The active sites of polymerases account for a large portion of discrimination between insertion of the correct base opposite the template and exclusion of incorrect bases. The active sites of replicative polymerases are structurally very well defined and can accommodate only the geometries of correct Watson-Crick base pairing.[20,21] Base mismatches fall outside this defined geometry and are disfavored for both incorporation and extension with continued synthesis.[22,23] Despite the high selectivity provided by replicative polymerase active sites, incorporation of mispaired nucleotides does occur occasionally, and it can be dealt with by the intrinsic proofreading exonuclease activities of pols δ and ε.[3] There is also evidence that errors made by the exonuclease deficient pol α can be proofread by the proofreading activity of pol δ.[24] This mechanism could also be a possibility for other polymerases that lack intrinsic proofreading.[25] These factors, combined with overall surveillance of newly replicated DNA by mismatch repair that detects and corrects mismatches that do become stably incorporated into DNA, contribute to the overall estimates of high replication fidelity.[3,18,26] However, these estimates represent situations of unperturbed DNA replication, a condition that is not always possible.

DNA Damage and Repair

DNA damage is a ubiquitous and constant threat from multiple sources, a fact that is evidenced by the multitude of mechanisms present to ameliorate DNA damage. DNA damage encompasses a wide range of alterations; from the breaking one or both strands of the phosphate backbone of the DNA helix, to loss of base information by hydrolysis of the glycosidic (base-sugar) bond, to the adduction of one or more atoms to any part of the nucleic acid chain (base, sugar or phosphate backbone). Each type of damage has a repair mechanism to process the damaged DNA strand, and many of these mechanisms have overlapping specificities.[27] DNA damage occurs thousands of times per cell per day. Exogenous chemical and energetic exposures can cause DNA alterations. Endogenous sources can also damage the DNA as a byproduct of normal cellular processes.[28] Exogenously caused damage occurs in multiple forms, due to exposures both natural and human caused. Ultraviolet light is a component wavelength of sunlight, and its energy can create multiple types of DNA damage, including *cis-syn* cyclobutane pyrimidine dimers (CPD) and pyrimidine (6-4) pyrimidone photoproducts (6-4 PP).[29] The chemotherapeutic drug cisplatin adducts DNA in multiple ways, including, but not limited to intrastrand crosslinking of purines.[30] Metabolic activation of polycyclic

aromatic hydrocarbons adds reactive oxygen groups to the hydrocarbon rings which can attach to bases in DNA.[31] Endogenous and spontaneous mechanisms also cause DNA damage. Reactive oxygen species are a byproduct of cellular metabolism and can create multiple DNA adducts, including 7,8-dihydro-8-oxo-guanine (8-oxoG) and thymine glycol.[32] Spontaneous hydrolysis of nucleotides causes abasic sites where the base coding information is lost.[28] Selected chemical structures of these lesions are shown in Figure I.3; and with respect to DNA replication, most if not all DNA alterations function as a block to normal DNA synthesis. This is due to the well-defined and constrained active sites of pol δ and pol ε.[33]

Translesion Synthesis

Once the helicase has segregated DNA into separate strands at the replication fork, excision of damage becomes undesirable, because cutting of the single strand would lead to more deleterious consequences such as a double strand break.[34] Blockage of the replicative polymerases retards fork progression, and this slowing of synthesis at or behind the replication fork is the definition of replication stress.[35] One method to cope with this damage-induced replication stress is the continued synthesis of DNA past these lesions. This process is termed translesion synthesis (TLS), as it utilizes specialized polymerases of the Y-family (pols η, κ, ι Rev1) and

one of the B-family (pol ζ) to copy past DNA damage, inserting bases across from the altered DNA template that would normally serve as a block to replication.[36] This process is classified as damage tolerance since it does not remove the damage. This contrasts with DNA repair processes that remove adducts or join separated DNA strands back together.[27]

Numerous signaling events and post-translational modifications have been implicated the recruitment of TLS polymerases to the replication fork. The best recognized of these post-translational modifications is the monoubiquitination of PCNA, and this modification is theorized to be the major signal for the switch between synthesis by replicative polymerases and TLS polymerases.[37] Pol η and the bypass of UV-induced CPD is the most well understood of these TLS events and is presented here as a model. UV-induced DNA lesions cause stalling of replicative polymerases, leading to replication stress and RPA coating the single stranded DNA (ssDNA) unwound by continued action of the helicase.[35] RPA-coated ssDNA induces activation of ATM and Rad-3 related protein kinase (ATR) through its binding partner ATR interacting protein (ATRIP).[38] ATR activates both CHK1 and p53, leading to their actions as effectors of checkpoint response, as well as other targets important to prevent later origins of replication from firing.[39] ATR also can

phosphorylate pol η directly on Ser601, and this modification requires interaction with the E3 ubiquitin ligase Rad18.[40] Evidence exists that ATR phosphorylates the p32 subunit of RPA, playing a further role in replication stress response signaling.[41,42] Downstream of ATR, p53 can induce pol η protein expression through increased mRNA transcription and the XPV gene (The XPV gene product is pol η) contains a p53 response element.[43] p53 related ubiquitin ligases MDM2 and Pirh2 post-translationally modify pol η with ubiquitin. MDM2 polyubiquitinates pol η and targets the polymerase for proteasomal degradation.[44] Pirh2 can monoubiquitinate pol η at 4 lysines in its final 31 amino acids, preventing its association with PCNA and limiting its ability to perform TLS.[45,46] This implies that some as yet unknown protein must deubiquitinate pol η for it to perform TLS.[45] p53 stabilization and activity are influenced multiple autoregulatory feedback loops. Both MDM2 and Pirh2 are induced by p53 in this autoregulatory manner, and negatively regulate p53.[47] This induction of pol η by p53 and negative regulation by p53 related targets could serve to balance the need for pol η in the response to DNA damage while also preventing its activity in inappropriate situations. Another negative regulator of TLS is p21 interaction with PCNA, and this contact prevents PCNA interaction with Y-family polymerases.[48-51] p21 degradation after UVC irradiation is dependent on its

interaction with PCNA and this degradation promotes Y-family polymerase foci formation.[50,51] This degradation of PCNA-coupled p21 could play a role in allowing pol η and other Y-family polymerase access to DNA to facilitate TLS.[51] These factors demonstrate polymerase access to DNA is tightly controlled and also ensure pol η can be recruited to stalled replication forks or post-replication gaps to perform TLS at the appropriate time.

The E3 ubiquitin ligase Rad18 plays a related role in monoubiquitination of PCNA.[52] The initial event of Rad18 recruitment is the same as checkpoint signaling, the covering of ssDNA with RPA.[53,54] Long regions of RPA-coated ssDNA recruits Rad18 to sites of stalled replication, and pol η is recruited along with it.[55-57] Evidence has shown that pol η also plays a role in targeting Rad18 to PCNA, as no PCNA-interacting domain for Rad18 has yet been identified, and the latter 300 amino acids of pol η have been shown to enhance Rad18-dependent monoubiquitination of PNCA.[58] Evidence suggests that ubiquitinated PCNA promotes maximal bypass of damage, and promotes the switch from replicative polymerase synthesis to TLS. The most well understood pathway leading to TLS is this monoubiquitination event at lysine 164 of the PCNA monomer by Rad18, but evidence exists that a separate, Rev1 dependent, monoubiquitin-PCNA independent pathway also transpires.[59]

Once PCNA monoubiquitination occurs and pol η is at the site of damage, bypass of a T-T CPD or one of the polymerase's other cognate lesions can occur. The polymerase switching model of TLS past a T-T CPD and eventual resumption of replicative polymerase synthesis is depicted in Figure I.4. This figure is also applicable to TLS by other polymerases, as shown when comparing Figure I.5 (panels A and B) with Figure I.4. Once the damage is copied past, replicative polymerase synthesis resumes.

How the polymerase that accomplishes TLS is chosen is as yet unknown, but evidence suggests this choice is lesion specific.[60] There are two complementary, non-exclusive models of how differing polymerases can accomplish bypass. The first is complete, efficient bypass of the lesion by one polymerase that is able to do so. The polymerase then is able to "sense" it has bypassed the lesion and dissociates, allowing replicative synthesis to continue. This model is shown for the example of pol η bypassing a T-T CPD in Figure I.5 panel A. The second model requires one insertion by a first polymerase, then extension of the paired, damaged primer-terminus by a second polymerase. This model is demonstrated in the example of pol ι/pol ζ bypassing a 6-4 PP in Figure I.5 panel B.[3] The timing of lesion bypass also has 2 differing models. Examples are shown in Figure I.5, panel C and D. The

polymerase-switching model (Figure I.5 C) of TLS most likely comes into play during active replication, while the gap filling model of TLS is utilized when repriming of DNA replication occurs downstream of the lesion, allowing replication to proceed and leaving unreplicated gaps behind.[60] These models are non-exclusive as well, and evidence suggest both bypass at the replication fork and post-replicative bypass happen in cells.[34,61,62] Depending on the timing of the damaging event in the cell cycle and the type of damage encountered, TLS could occur by a combination of these models.

Pol η

The discovery of eukaryotic translesion synthesis polymerases occurred as a result of studies of the autosomal recessive genetic disorder Xeroderma pigmentosum (XP).[63] XP patients are prone to sunlight-induced skin cancers, and approximately 80% of the patients are deficient in one of 7 proteins involved in nucleotide excision repair (NER).[64] The remaining 20% have intact NER but have a defect in DNA synthesis after exposure to certain genotoxic agents, including ultraviolet light (UV). These 20% are classified as Xeroderma pigmentosum variant (XPV).[65] XPV patients have normal excision repair (in contrast to classical XP), but still show sensitivity to sunlight, freckled skin and 100% tumor incidence at times

much earlier (20-30 years typically) than the median age of incidence in normal populations.[64] This defect in XPV patients is due to mutations within the XPV gene, which encodes DNA polymerase η (pol η).[66] Ectopic expression of either the human or mouse form of pol η in XPV cells corrects the defect in DNA synthesis.[67] This synthesis defect in XPV is due to the inability of these cells to efficiently copy past the UV-induced lesions present in their DNA.[68] Pol η is widely conserved in eukaryotes, with examples occurring in organisms as diverse as yeast, plants, worms, flies, mice, humans and even thermophilic deep-sea worms.[69,70] Identification of the XPV gene as a bona fide DNA polymerase led to the realization that multiple DNA polymerases exist with the ability to perform TLS opposite a variety of structurally diverse lesions. In eukaryotes, these TLS polymerases are the members of the Y-family including pol η, as well as the B-family polymerase, pol ζ. Y-family polymerases were classified as a separate family due to their lack of primary sequence homology with replicative polymerases of the B-family.[60]

It has been proposed that each TLS polymerase has a set of cognate lesions, and this allows individual polymerases to efficiently bypass lesions that are accommodated by the polymerase active site.[71,72] The first evidence as to the identity of pol η cognate lesions comes from studies published in 1976 of XPV cells. XPV

variant cells display higher mutation rates than normal cells when exposed to ultraviolet irradiation in the UVC spectrum, and this exposure reduced survival.[73,74] Subsequent studies show that absence of pol η increases mutation rates when exposed to DNA damaging agents creating both UV damage and 8-oxoG, leading to the idea that pol η participates in a pathway that is "error-free". [66,75-79] The identity of the molecular defect in XPV was not discovered until 1999, but once nonsense, missense, frameshift mutations and large deletions in the XPV gene were identified as the cause of this defect,[64,66,80-86] investigations into pol η began in earnest. Mouse models of XPV provide further evidence that pol η helps organisms cope with UV induced DNA lesions. Pol η knockout mice are viable, fertile, and show no abnormalities when visually observed for one year.[87] When exposed to UV light, the mice recapitulate the phenotype of XPV of skin freckling, and 100% of pol η knockout mice develop skin tumors following a regiment of UV exposure. Mouse embryonic fibroblasts isolated from these mice were sensitive killing when exposed to UV.[87,88] These all provide evidence of pol η's participation in the tolerance of UV-induced damage.

Identification of the XPV gene as encoding human pol η and the Rad30 gene encoding the *S. cerevisiae* homolog of pol η allowed for creation of expression vectors

capable of producing purified proteins.[66,69,75,89,90] This led to in vitro biochemical experiments using the isolated proteins and chemically synthesized lesions to determine the lesion specificity of bypass by pol η. In addition to CPD, evidence exists that pol η is able to bypass a number of other DNA modifications and adducts, including 8-oxoG, cisplatin adducts, and thymine glycol.[76,90-94] Interestingly, pol η is unable to efficiently bypass 6-4 PP, another lesion created by UV.[68] This also is true for other lesions, such as an abasic site and BPDE adducts.[68,91,95] This argues that bypass by pol η is lesion specific. These preliminary studies were qualitative in nature and on the whole used conditions that were less comparable to the putative mechanism of bypass in vivo. Generally, these studies involved longer time points and large polymerase excess over substrate, or assessed the ability of the polymerase to insert single nucleotides opposite the lesions tested. The definition of a bypass event has evolved to describe insertion across from the lesion and extension of at least 1 nucleotide past the lesion,[95] as polymerases other than TLS polymerases do not extend damaged or mismatched primer termini efficiently.[23,60] Studies that examine the products of a single round of synthesis and require multiple insertion and extension events utilizing damaged templates can better approximate the function of the polymerase in vivo.[96]

The transition to quantitative studies provides further insight into the molecular mechanism of pol η bypass. Pol η is specialized for its primary purpose of lesion bypass, as pol η preferentially copies T-T CPD and 8-oxoG containing DNA more readily than the corresponding undamaged DNA sequence.[97,98] Pol η is not nearly as processive as replicative polymerases, copying less that 10 nucleotides per synthesis event.[90]. Pol η is also able to sense when it has bypassed the lesion and switches to less processive synthesis after bypassing T-T CPD, providing another mechanism to prevent unnecessary synthesis by pol η.[97] This differs for 8-oxoG, as the human enzyme does not display this switch to less processive synthesis after efficiently bypassing the lesion. This is not the case when examining *Saccharomyces cerevisiae* pol η, as yeast pol η displays a similar switch when copying 8-oxoG and T-T CPD as human pol η does when copying T-T CPD.[98] This is the first of many examples of differing properties between these two pol η homologs. These differences may arise because *S. cerevisiae* possesses only 3 TLS polymerases (pols η, ζ, and Rev1) while humans possess at least 5 (pols η, ζ, and Rev1, but also ι, κ, λ, and possibly θ and ν as well).

Structure of pol η

Y-family polymerases retain the classical "right hand" structure conserved in other polymerases.[99] They also share similar domain structure to other polymerases, specifically the fingers, palm and thumb domains.[21] Y-family members contain an additional fourth domain, referred to as the little finger, or polymerase-associated domain; which sits alongside the fingers of the polymerase.[100,101] The lack of amino acid sequence homology between Y-family and other polymerases translates into differences in the details of the active site structure. Early crystal structures of various Y-family polymerases revealed a more open active site, suggesting a mechanism by which adducted DNA could be accommodated by the polymerases.[100-104] (For comparison of Y-Family polymerases and replicative polymerases active site structures, please see figures described in previous reviews.[3,36,59,105]) These active sites are still differentiated between members of the Y-family, possibly providing a basis for their differing properties.

Early random mutation screens investigating the role of amino acids in the polymerase active site were conducted with other Y-family homologs as a basis for selection of the region of investigation,[106,107] as high quality crystal structures of human pol η in complex with DNA were unavailable until more recently [85] These

random mutation studies identified specific amino acids that were conserved in pol η and were important for bypass of T-T CPD. Sequence alignments also show conserved amino acids throughout the active sites of Y-family polymerases.[81] Once crystal structures were solved for human pol η with DNA within the active site, it was shown that pol η actually fits 2 bases simultaneously into its active site cleft.[85] This provides a molecular basis for the bypass of the structure of a T-T CPD, and is reminiscent of the same phenomenon seen in the Y-family polymerase Dpo4.[100] T-T CPDs are created by the linking of adjacent thymines with a cyclobutane ring, as shown in Figure I.3. This rigidly connects the two bases together and kinks the backbone of the DNA strand.[108] The ability to accommodate 2 bases while linked allows the polymerase to copy past the damage and bypass the lesion with greater efficiency than copying undamaged DNA.[85,97] It also provides a mechanism for the preferential dissociation after bypass of a T-T CPD, as the rigidity imparted by the damage causes steric conflicts with the polymerase and promotes pol η dissociation after bypass.[85] This open active site crystal also provides a structural basis for the ability of pol η to bypass the chemically similar cisplatin crosslink (Fig I.3),[109] and has provided understanding of the mechanism of somatic hypermutation at mutation hotspots in immunoglobulin genes.[110] The structure of 8-oxoG bypass by

human pol η is an unanswered question. Crystal structures of pol β with 8-oxoG in the active site indicates the lesion is able to form a base pair with adenine with geometry nearly identical to the correct Watson-Crick base pair after rotation of 8-oxoG around its glycosidic bond.[111] As of yet, no crystal structure has been produced of human pol η with 8-oxoG in its active site, but the open nature of the crystal structure would be expected to accommodate the ambiguous coding potential 8-oxoG represents, based on the measured efficiency of bypass.

Crystal structures of human pol η have focused on the structure of the active site, which exists in the first 432 of the 713 amino acids of the protein.[85,109,110] The C-terminal end of the protein contains domains necessary for regulation of polymerase function and protein:protein interaction.[59] A diagram of pol η active site domains and regulatory regions is displayed in Figure I.6. In addition to the structures of the four catalytic domains, 4 separate types of regulatory motifs have been described. The PCNA interacting domain, termed the PCNA interacting peptide (PIP) has been identified.[112,113] This PIP motif occurs twice in the protein, at the C-terminus of the little finger domain and at the extreme C-terminus of the protein,[71,105] and one can functionally substitute for the other in the interaction between PCNA and pol η.[113] Another motif for interaction with PCNA is the ubiquitin binding zinc finger

domain (UBZ).[114] This motif mediates interactions with monoubiquitinated PNCA, as well as monoubiquitinated pol ι.[114,115]. Pol η also possesses two regions that mediate interactions with another Y-family member, Rev1.[116,117] These Rev1 interacting regions (RIR) are seen in all other polymerases in the Y-family, and Rev1 has a conserved region shown to interact with the other members of the Y-family.[71] Finally, pol η's C-terminal region contains its nuclear localization sequence. All of these motifs likely play a role in the regulation and interactions between members of the Y-family and other proteins necessary for TLS.

Fidelity of Pol η

Y-family polymerases are suited to performing TLS, and as stated above, early crystal structures showed open, solvent accessible actives site is a mechanism used to accommodate the distortions to the structure of DNA caused by DNA damage.[100-104] As a consequence of their open active sites, their fidelity on undamaged DNA is thousands of fold lower than replicative polymerases.[3] (Figure I.2) For example, on undamaged DNA, Pol η makes 1 error for every ~30 nucleotides it copies.[118,119] By comparison, pol δ and pol ε make 1 error per 100,000 and 500,000 bases copied, respectively.[3] Pol η can also make tandem bases substitution (insertion of 2 consecutive incorrect bases and extension of the resulting mismatched primer

termini) errors at a much higher rate than replicative polymerases.[118,119] Pol κ has a somewhat higher fidelity than pol η, but still much lower than replicative polymerases, making 1 error per 175 bases copied when copying undamaged DNA.[120] Pol ι is unusual in that it prefers to insert guanine opposite template thymine, rather than the adenine of the correct Watson-Crick base pair.[121] Along with open active sites, Y-family polymerases lack intrinsic proofreading exonuclease activities present in the replicative polymerases δ and ε. These error rates imply that Y-family polymerase access to undamaged DNA must be tightly controlled in order to prevent mutagenesis.[59]

The use of an assay that allow for the quantification of individual bypass and extension events allowed for further insights into the fidelity of TLS by pol η. This assay is an adaptation of the earlier LacZ complementation color screening system for undamaged DNA which relies on gapped M13 phage DNA to screen for base substitutions and nucleotide insertion/deletion (indel) events. This system allows for screening of thousands of bacteriophage plaques, each plaque representing an individual bypass event. These plaques can be further analyzed by sequencing of the mutant plaque genome to allow determination of base substitution and indel spectra. Use of this system requires the presence of all 4 deoxynucleotide

triphosphates (dNTPs) in competition, which allows for further information unable to be gained when measuring insertion of single nucleotides. This assay is adaptable to the study of multiple lesions,[122] and has been used to study the bypass fidelity of pol η when copying past T-T CPD, 8-oxoG and abasic sites.[95,96,98]

Since pol η has been implicated in the in vivo bypass of CPD and 8-oxoG, [73,74,79,123,124] the fidelity of bypass by pol η has been studied quantitatively for these two lesions. The fidelity of bypass of a T-T CPD by human pol η is roughly equal to fidelity when copying undamaged DNA (i.e. 3-5% errors).[97] This fidelity is similar to yeast pol η.[75] Pol η also has interesting pattern of making errors opposite a T-T CPD, making errors opposite the 3'-T of the dimer (first insertion) much more often than opposite the 5'-T (second insertion) of the dimer.[97]

When analyzing the in vitro bypass of 8-oxoG, human pol η makes errors ~50% of the time, emphasizing the mutagenic potential of the lesion.[91,98,125] Yeast pol η shows much higher fidelity, making errors opposite 8-oxoG in only 2-3% of bypass events, an over 15-fold reduction from human pol η.[76,98] The similarity between the error rates for 8-oxoG and T-T CPD suggest that in yeast this fidelity is sufficient to prevent mutagenesis.[98] Multiple factors have been suggested as mechanisms to raise the fidelity of TLS by pol η, including extrinsic proofreading and modification of

polymerase fidelity by interaction with accessory proteins.[3] Extrinsic proofreading has been shown to improve the fidelity of bypass of T-T CPD by yeast pol η.[126] Replication accessory proteins such as RPA, RFC and PCNA do not increase the fidelity of T-T CPD or 8-oxoG bypass by yeast pol η,[98,127] but one report has shown PCNA and RPA together can improve human pol η fidelity greater than 21-fold in single-nucleotide insertion experiments.[125] Despite this evidence, the complete mechanism that raises pol η fidelity to the level that suppresses mutagenesis is less than clear.

Role of Pol η in processes other than TLS

The primary characterization of pol η's function in replication is as a TLS polymerase. However, pol η has been implicated in other cellular replication processes as well. Targeting of pol η's reduced fidelity opposite undamaged DNA to immunoglobulin genes is useful in generation of antibody diversity and somatic hypermutation.[128-130] Pol η is also involved in A/T modification of immunoglobulin genes in mice, as pol η knockout mice show reduced levels of this mutation in these genes.[131,132] Pol η can also play a role in homologous recombination.[133] Pol η has been shown to contribute to replication of DNA structures such as G-quadruplexes and common fragile sites.[134,135] It can also assist in copying non-B form DNA.[136] Pol η has

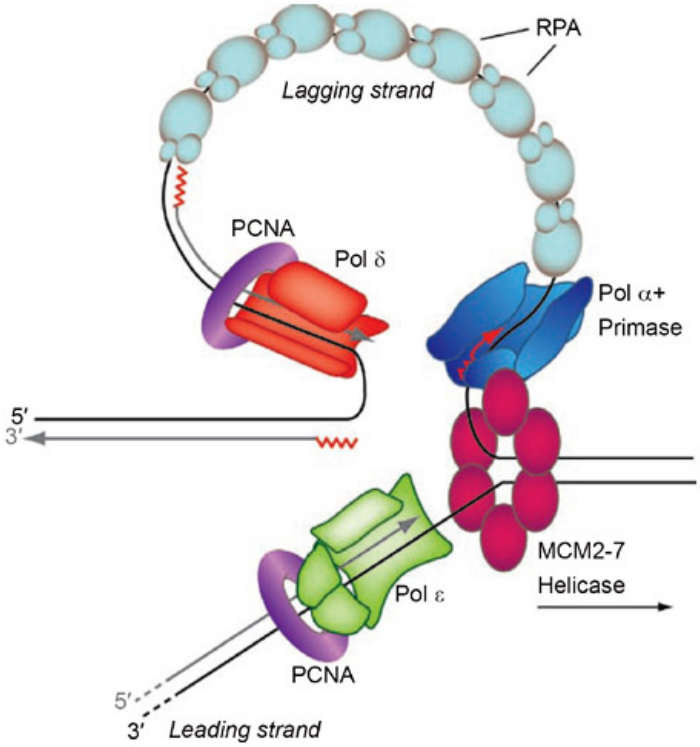
also been shown to interact with Werner syndrome protein, a helicase and exonuclease whose absence causes the premature aging syndrome of the same name.[137] Participation of pol η in these processes in which damage is not necessarily involved provides relevance to studying the fidelity and bypass of pol η not only on damaged templates, but also the fidelity of synthesis on undamaged DNA.

Pol η in vitro versus in vivo

Early characterizations of pol η bypass identified it as participating in an "error-free" pathway of bypass.[66,75,76,89,138] However, evidence has clearly shown that pol η bypass of lesions and copying of undamaged DNA in vitro is not high fidelity in any sense.[91,96,98,125] Multiple possibilities have been proposed explain how the in vitro fidelity of bypass could be altered, including extrinsic proofreading, interaction with replication accessory proteins, and interaction with mismatch or other repair processes.[3,139-141] As the active site of the polymerase forms the basis of the fidelity and efficiency of bypass by pol η, we wished to examine the role of selected amino acids implicated in bypass of the two quantitatively studied lesions bypassed by pol η, T-T CPD and 8-oxoG. Using the biochemical tools and assays at our disposal, we selectively mutated amino acids implicated in the function of the polymerase in order to better understand the molecular mechanism of pol η. These amino acids

were selected using the crystal structure, information from previous random mutation screens, homology with other polymerases, and SNPs present in our region of interest.[85,106,107,142-144] We also have the tools at our disposal to produce purified replication accessory proteins present at the replication fork. Using these same assays, we wished to test the ability of these proteins to affect the fidelity of bypass by pol η.

Figure I.1 – Simplified model of a eukaryotic replication fork. Protein depictions are based on currently accepted subunit composition of *S. cerevisiae* proteins but are not meant to be accurate structure-based models. The assignment of pol ε to the leading strand is based on multiple reports,[6,7] but has not been definitively established for all replication. Pol δ is consequently assigned to the lagging strand, consistent with earlier reports.[8-11] MCM 2-7 helicase(magenta); RPA (light blue ovals); PCNA (purple torus); pol α-primase complex (blue); RNA-DNA hybrid primer (red zig-zag and arrow); pol δ (red); pol ε (green); template strand DNA (black lines); newly synthesized DNA (gray lines). Figure reprinted and legend adapted from McCulloch and Kunkel.[3]


RPA
Lagging strand
PCNA
Pol δ
Pol α+
Primase
5′
3′
MCM2-7
Helicase
Pol ε
PCNA
5′
3′
Leading strand

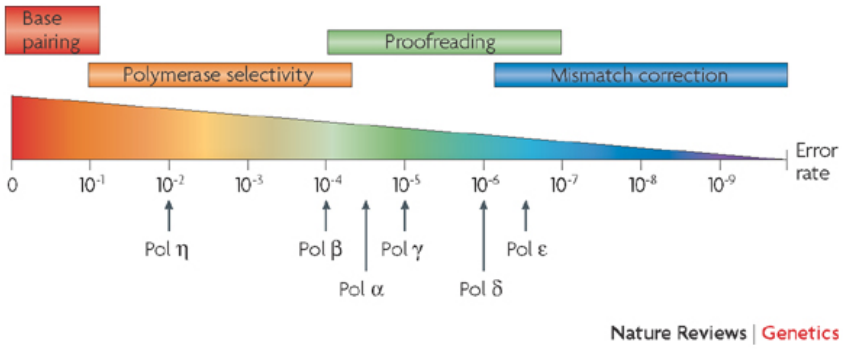


Figure I.2 – Components of DNA Replication Fidelity. Summarized above the line are the contributions of the different steps that enhance the fidelity of DNA replication. Below the line are representative values for the rates of single-base substitutions by different DNA polymerases when copying undamaged DNA. Figure is reprinted from Loeb and Monnat.[18]

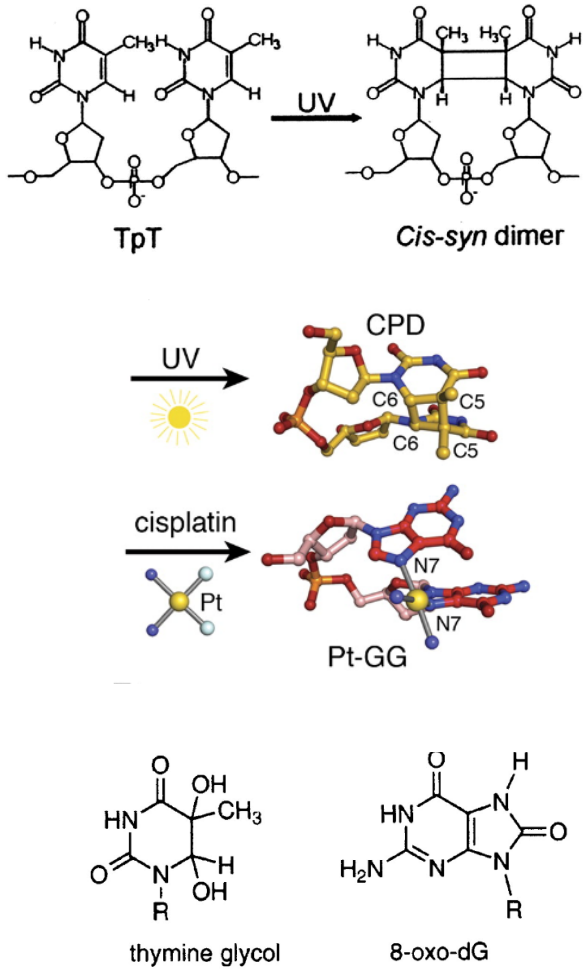
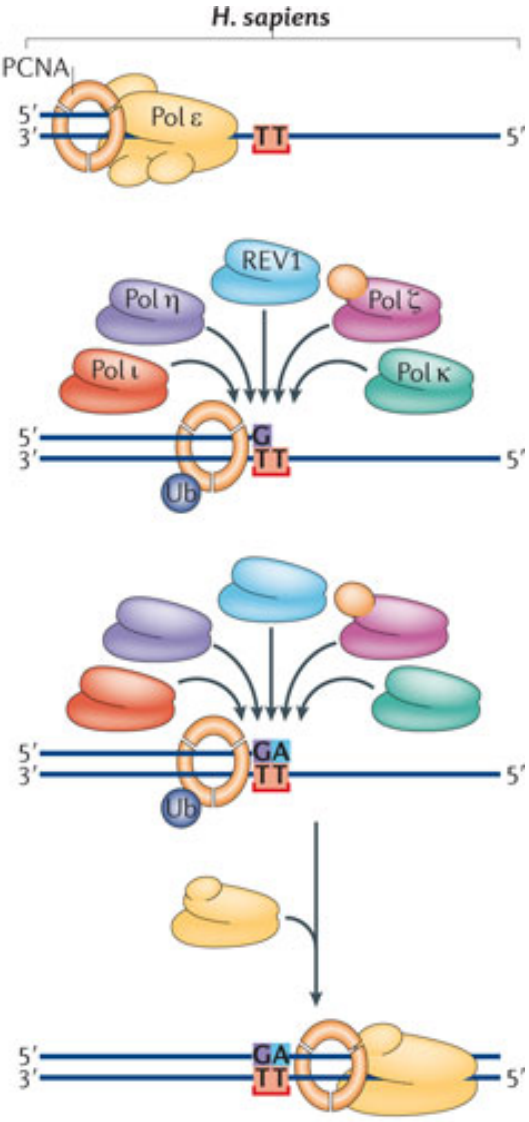


Figure I.3 – Selected chemical structures of DNA damage that can be bypassed by pol η. Structures represent the UV-induced T-T CPD, adducts created by the chemotherapeutic cisplatin, and the oxidative lesions 8-oxoG and thymine glycol. Figures adapted from previously published reports.[32,108,109]

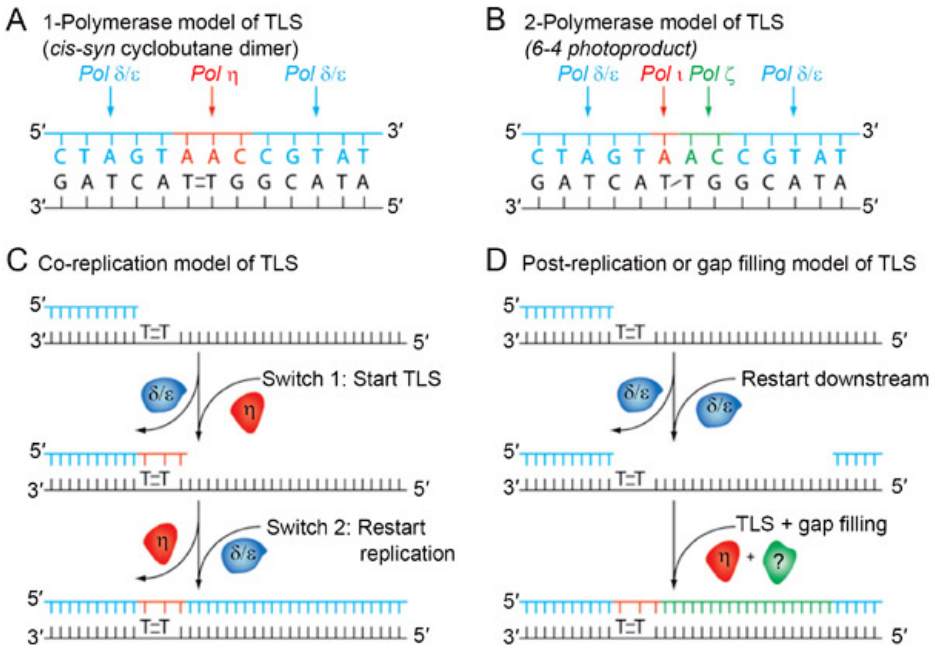
Figure I.4 – Polymerase switching model of TLS. Replicative polymerases like pol ε stall at the site of a T-T CPD, represented by a TT in the figure. TLS polymerases are recruited to the site at the same time as PCNA is monoubiquitinated. In humans, at least five TLS polymerases can be recruited to sites of arrested replication, including Pol η, Pol ι, Pol κ, REV1 and Pol ζ. The likelihood that errors will be made when bypassing the lesion will depend on the nature of the DNA lesion and the polymerase utilized. The extension step may be facilitated by the same enzyme that performed the (mis)insertion or by a completely different polymerase. It is thought that pol η catalyzes the complete bypass reaction past the T-T CPD pictured. Once the nascent DNA chain has been extended beyond the lesion, the replicative DNA polymerase replaces the TLS polymerase. Figure adapted from Sale et al.[59]



Nature Reviews | Molecular Cell Biology

Figure I.5 – Models of translesion synthesis. A. The 1-polymerase model of TLS, shown here for a T-T CPD, states that a single polymerase is responsible for the complete bypass of a lesion, including insertion opposite all lesion bases and extension from the primer terminus opposite a damaged template base. **B.** The 2-polymerase model of TLS, shown here for a thymine-thymine 6-4 photoproduct, states that different polymerases are responsible for the insertion steps at the various lesion positions. In the example given, note that while pol ζ is responsible for extension from the template-3' T primer terminus, it also carries out an insertion at the 5' T position of the lesion. For a single base lesion, the insertion step would be opposite undamaged DNA. Note that for both examples given, the actual TLS reaction is flanked relatively closely both upstream (1-2 bases) and downstream (1-5 bases) of the lesion by replicative polymerase synthesis. **C.** Model for TLS that occurs at a replication fork during the process of ongoing synthesis. **D.** Model for TLS that takes place as a "gap-filling" reaction, away from the main replication machinery. Note that both of these models are consistent with either the 1- or 2-polymerase model of TLS given in panels **A.** and **B**. In both cases, post-translational modification of PCNA and possible other proteins is critical for the polymerase switch.

Figure I.5 continued – Note that panels A and B are models of the actual TLS process while panels **C** and **D** depict models for the timing of TLS. As such, there is overlap between the panels. Figure reprinted and legend adapted from McCulloch and Kunkel.[3]


A 1-Polymerase model of TLS
(cis-syn cyclobutane dimer)
Pol δ/ε
Pol η
Pol δ/ε
5′
3′
C T A G T A A C C G T A T
G A T C A T=T G G C A T A
3′
5′
B 2-Polymerase model of TLS
(6-4 photoproduct)
Pol δ/ε
Pol ι
Pol ζ
Pol δ/ε
5′
3′
C T A G T A A C C G T A T
G A T C A T–T G G C A T A
3′
5′
C Co-replication model of TLS
5′
3′
T=T
5′
δ/ε
η
Switch 1: Start TLS
5′
3′
T=T
5′
η
δ/ε
Switch 2: Restart
replication
5′
3′
T=T
5′
D Post-replication or gap filling model of TLS
5′
3′
T=T
5′
δ/ε
δ/ε
Restart downstream
5′
3′
T=T
5′
TLS + gap filling
η
+
?
5′
3′
T=T
5′

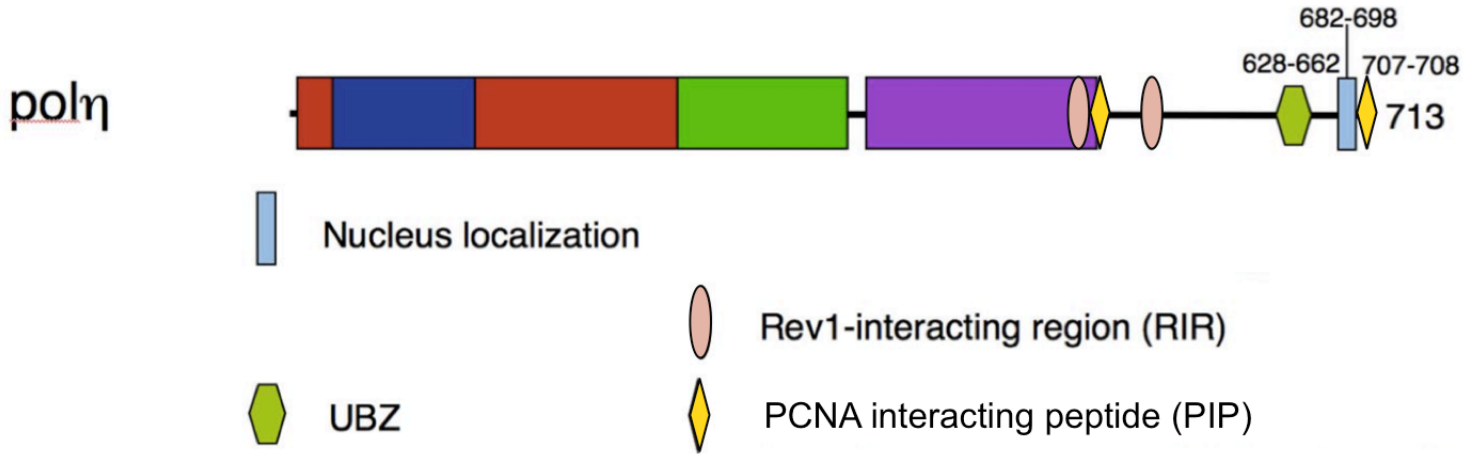


Figure I.6 – Domain structure of pol η. The 4 domains of the catalytic (active) site of pol η are indicated as colored blocks and are identified as follows: palm (red), fingers (blue), thumb (green), and little finger (purple). Regulatory regions are marked as colored shapes as indicated by the legend. UBZ stands for ubiquitin binding zinc finger. Figure is a combination of information from 4 reviews, [36,59,71,105] and design as well as style of the figure is adapted from Yang and Woodgate.[105]

RATIONALE

DNA polymerase η synthesizes across from damaged DNA templates in order to prevent deleterious consequences caused by the inability of replicative polymerases to efficiently copy past these templates.[60] While imperfect, this bypass is overall positive for a cell, as cells deficient in pol η display higher mutation rates when exposed to DNA damaging agents,[73,74,79,123,124] despite the fact that the fidelity of translesion synthesis is much lower than replicative polymerase synthesis.[3,97,98] This lowered fidelity can also have undesired consequences and implications for genome stability. This dissertation research focuses on determining the biochemical underpinnings of the efficient but moderate-to-low fidelity bypass of T-T CPD and 8-oxoG DNA lesions by pol η.

This work has focused on 2 questions:

1. What effect do single amino acid substitution mutants in and surrounding the active site of pol η have on fidelity and bypass of the polymerase?
2. Can replication accessory proteins present at the replication fork alter the fidelity of replication of damaged and undamaged DNA by pol η in the presence of all 4 dNTPs?

CHAPTER 1

Biochemical analysis of active site mutations of human polymerase η[145]

Samuel C. Suarez[§], Renee A. Beardslee[§], Shannon M. Toffton[§], and Scott D. McCulloch[§]

[§] *Department of Environmental and Molecular Toxicology, North Carolina State University, Campus Box 7633, Raleigh, NC 27695*

Article available as originally published at:

http://www.sciencedirect.com/science/article/pii/S0027510713000195

Running title: Active site mutants of pol η

Keywords: Mutagenesis, DNA damage, translesion synthesis, DNA polymerase eta, 8-oxoG, thymine-thymine CPD

Abstract

DNA polymerase η (pol η) plays a critical role in suppressing mutations caused by the bypass of *cis-syn* cyclobutane pyrimidine dimers (CPD) that escape repair. There is evidence this is also the case for the oxidative lesion 7,8-dihydro-8-oxo-guanine (8-oxoG). Both of these lesions cause moderate to severe blockage of synthesis when encountered by replicative polymerases, while pol η displays little no to pausing during translesion synthesis. However, since lesion bypass does not remove damaged DNA from the genome and can possibly be accompanied by errors in synthesis during bypass, the process is often called 'damage tolerance' to delineate it from classical DNA repair pathways. The fidelity of lesion bypass is therefore of importance when determining how pol η suppresses mutations after DNA damage. As pol η has been implicated in numerous in vivo pathways other than lesion bypass, we wanted to better understand the molecular mechanisms involved in the relatively low-fidelity synthesis displayed by pol η. To that end, we have created a set of mutant pol η proteins each containing a single amino acid substitution in the active site and closely surrounding regions. We determined overall DNA synthesis ability as well as the efficiency and fidelity of bypass of thymine-thymine CPD (T-T CPD) and 8-oxoG containing DNA templates. Our

results show that several amino acids are critical for normal polymerase function, with changes in overall activity and fidelity being observed. Of the mutants that retain polymerase activity, we demonstrate that amino acids Q38, R61, and Y52 play key roles in determining polymerase fidelity, with substitution of alanine(Q38A and R61A) and glutamic acid (Y52E) causing both increases and decreases in fidelity. Remarkably, the Q38A mutant displays increased fidelity during synthesis opposite 8-oxoG but decreased fidelity during synthesis opposite a T-T CPD.

Introduction

DNA polymerase η, a member of the Y-family, plays a critical role in bypassing DNA lesions, most notably *cis-syn* cyclobutane pyrimidine dimers (CPD) created by exposure of DNA to ultraviolet light. It synthesizes past lesions that block replicative polymerases and would otherwise halt replication fork progression. This action serves to lower the mutagenic potential they represent, and cells lacking functional pol η display markedly higher mutation rates after UV light exposure.[59,71,73,74] Despite this observation, the fidelity of pol η during the bypass event is far from perfect or even high fidelity in the classic sense of the word.[91,97,98] Replicative polymerases are capable of copying 10^5-10^6 nucleotides without making an error,[3,19,146] while the fidelity of translesion synthesis (TLS) by polymerase η can

be 3-4 orders of magnitude lower. Despite early characterizations of TLS by pol η as being "error free",[75,147-149] during thymine-thymine CPD (T-T CPD) bypass the enzyme produces errors in the range of 1 in 30,[97] which is also the average error rate when copying large stretches of undamaged DNA.[118,119] The fidelity of human pol η when bypassing 8-oxoG is even lower, with both steady state kinetic assays[91,125] and those requiring complete bypass in the presence of all four deoxynucleotides[98] showing that dATP and dCTP are inserted at roughly equal frequencies. On an absolute scale this is very poor fidelity, although there is evidence that errors generated during bypass can be detected and removed by the exonuclease activities of the replicative polymerases.[126] It is often stated that 8-oxoG is not a blocking lesion, although there are multiple reports that demonstrate it can impede the normal synthesis of replicative polymerases.[98,150-152] Additionally, while multiple polymerases have been shown to be able to bypass 8-oxoG under some conditions,[125,153,154] pol η does so with greater than 100% efficiency compared to an undamaged G in the same context.[98] We hypothesize that the ability of pol η to efficiently synthesize past the lesion makes it one of the preferred polymerase to do so in vivo, acknowledging other polymerases may also play a role.[125,155] In this scenario, the low-fidelity of bypass is a required trade-off to prevent the

accumulation of single stranded regions of unreplicated DNA that are even more problematic. The combination of high-efficiency but low-fidelity lesion bypass by pol η therefore represents a balance that is assumed to lead to an overall positive outcome for the cell.

The TLS properties of pol η are conserved across a wide range of species. For instance, deep sea worms have pol η homologs that allow T-T CPD bypass, despite presumably never being exposed to UV light.[70] While the properties of pol η from different species are generally in agreement, there are a few differences. For instance, the human, mouse and budding yeast (*S. cerevisiae*) forms of pol η display similar properties when it comes to T-T CPD bypass and their fidelity when copying undamaged DNA.[75,97,118,119,127,147,149,156] With respect to 8-oxoG bypass, however, there is a marked difference. DNA pol η from yeast displays the same high efficiency of bypass with bypass fidelity in the same range as a T-T CPD.[76,98,157] As noted above, while human and mouse pol η are able to efficiently bypass 8-oxoG, their fidelity when doing so is remarkably low, in the range of 1 in 2, or 50%.[91,98,125] Despite these in vitro findings there is evidence that in vivo pol η suppresses mutagenesis by this lesion.[79] Some possible explanations for these seemingly disparate results are extrinsic proofreading of errors, as has been shown for T-T CPD bypass, and

modulation of the fidelity by association with replication accessory proteins.[125,126] Interestingly, while the latter mechanism has been demonstrated by one assay for human pol η, it has been ruled out by a different assay for yeast pol η.[98]

Early crystal structures of Y-family polymerases suggested that a very open, solvent accessible active site was one of the means by which the unique properties of the enzymes were achieved.[100-104] More recent crystal structures have shown for both human and yeast pol η that the active site can in fact fit two template bases.[85,158] This provides an explanation for how the chemically linked CPD can be so readily bypassed, as can lesions created by the crosslinking chemotherapy agent cisplatin.[90,92,109,159-161] Interestingly, the structures of DNA pol η in complex with undamaged DNA show two bases in the active site as well. This is in line with the observed low fidelity when copying undamaged DNA,[97,118,119,148,156] and may help to explain how pol η is able to bypass other DNA lesions such as 8-oxoG and thymine glycol.[90,91,157] While the crystal structure of yeast pol η bound to 8-oxoG containing DNA is informative for the preferred dCTP incorporation by that enzyme,[162] there is currently no structure available of human pol η in complex with 8-oxoG to shed light on the remarkably low fidelity bypass it exhibits.[91,98] However, the available crystal structures of the human enzyme show which amino acid residues interact

with the DNA and incoming nucleotide triphosphate and have allowed us to select several candidate residues for detailed biochemical analysis.

As pol η has been implicated in processes other than TLS,[128,129,133-135,137] we sought to understand the molecular determinants that govern human pol η function. To investigate the link between lesion bypass fidelity and lesion bypass ability, we have generated several mutant forms of pol η, each containing a single amino acid substitution in the active site and surrounding regions of the catalytic core of human pol η (Figure 1.1). These include three single nucleotide polymorphisms (SNP's) identified in the NCBI dbSNP database[144] as well as several picked from the crystal structures and alignments of pol η sequences.[81] Each mutant was purified and tested for overall DNA polymerase activity as well as lesion bypass ability and fidelity.

Materials and Methods

Reagents and materials

Oligonucleotides were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). and Midland Certified Reagent Co. (Midland, TX). Phosphoramidite precursors for the damaged nucleotides were purchased (by Midland Certified Reagent Co.) from Glen Research (Sterling, VA). Nucleotides and restriction enzymes were obtained from New England Biolabs (Ipswich, MA). All polymerase

and exonuclease reactions were performed in 40 mM Tris pH 8.0, 250 μg/ml BSA, 10 mM DTT, 10 mM $MgCl_2$, 60 mM KCl, and 1.25% glycerol. Polymerase reactions were supplemented with 1 mM or 0.1 mM final concentration of each dNTP for the forward gap filling or oligonucleotide-based assays, respectively. All cell lines, bacteriophage and reagents for lesion bypass and forward mutation assays have been previously described.[122,163]

Expression vector and protein purification

The pET21b-XPV vector was obtained through the generosity of the laboratory of Dr. Thomas Kunkel (NIEHS, Research Triangle Park, NC). This vector codes for the first 511 amino acids of human pol η containing a C-terminal 6X histidine tag after a 2 amino acid linker (519 amino acids total; ~56 kD; vector sequence available on request). Amino acid changes were introduced using the Stratagene QuikChange II XL site-directed mutagenesis kit using conditions recommended by the manufacturer. Primers used for mutagenesis reactions are given in Supplementary Table 1.1. Changes were verified and no additional mutations in the ORF were confirmed by sequencing performed by Genewiz (South Plainfield, NJ). The expression vector (coding for either wild type or the single amino acid changes) was introduced into BL21 (DE3) cells by electroporation, grown

overnight at 37 °C on LB media containing 100 μg/ml ampicillin, and then a single colony was used to inoculate 10 ml LB broth with 100 μg/ml ampicillin and grown to saturation at 30 °C/250 RPM. Saturated 10 ml culture was used to inoculate 1 L LB broth (no ampicillin) that was incubated at 30 °C/250 RPM until the OD_{595} reached 0.4-0.6. IPTG was added to 0.5 mM final concentration and the culture was incubated at 15°C/250 RPM for an additional 18 hours. Cells were harvested by centrifugation, the pellet washed twice with PBS, and the cell pellet was resuspended in a small volume (~5 ml) of PBS that was dripped into liquid N_2, creating ~3-5 mm diameter drops that were stored at -80 °C. Cells were lysed using a SPEX Sample Prep 6870 Freezer/Mill. Cell drops were cooled for 10 minutes in liquid N_2 and then subjected to 6 cycles of: 1 minute grinding (10 impacts per second) and 1 minute cooling. The resulting lysed cell powder was thawed and resuspended in 20-40 ml ice cold Buffer A (25 mM sodium phosphate pH 7.4, 10% glycerol, 300 mM NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol) supplemented with 0.2 mM PMSF and Roche Complete Protease inhibitor tablets (1 tablet per 50 ml total volume). The crude cell lysate was sonicated (Branson 250 sonifier; output-2, duty cycle 50%) 6-10 times for 30 seconds with 30 seconds on ice in between cycles. Cell lysate was centrifuged at 20,000g for 30 minutes at 4 °C. The clarified extract

was loaded onto a 5 ml HiTrap Chelating HP column ({GE Life Sciences; charged with $NiSO_4$ and then equilibrated in Buffer A). The column was washed with 10 column volumes (CV) Buffer A, then with 5 CV Buffer B (25 mM sodium phosphate pH 7.4, 10% glycerol, 100 mM NaCl, 5 mM β-mercaptoethanol) supplemented with 10 mM imidazole. Bound protein was eluted using a 20 CV linear gradient of Buffer B containing from 10 to 600 mM imidazole. Fractions containing protein (as judged from A_{280} readings) were analyzed by SDS-PAGE and fractions with the expected protein size were pooled and concentrated using an Amicon Ultra -15 (30K MWCO) centrifugal filter (Millipore), then diluted 1:10 with Buffer C (25mM Tris-Cl pH 7.4, 10% glycerol, 1 mM DTT) containing 100 mM NaCl. The resulting solution was filtered through a 0.2 μm filter and loaded onto Mono S 5/50 GL column (1ml bed volume) equilibrated with Buffer C (100 mM NaCl). The column was washed with 10 CV Buffer C containing 100 mM NaCl, then eluted with a 30 CV linear gradient of Buffer C containing from 100 to 600 mM NaCl. Protein fractions identified by A_{280} were analyzed by SDS-PAGE. Selected fractions containing highly pure polymerase were aliquoted and frozen in liquid N_2 and stored at -80 °C until further use.

Polymerase and lesion bypass assays

Determination of polymerase activity and bypass efficiency were performed essentially as described previously[95,97] using the template 5'-TCGGTACCGGGTTAxCCTTTGGAGTCGACCTGC-3' (where x represents either undamaged G or 8-oxo-G and the underlined TT are either undamaged or a *cis-syn* cyclobutane thymine-thymine dimer), and primers 5'-Cy5-GCAGGTCGACTCCAAAG-3' (G/8-oxoG reactions) or 5'-Cy5-GCAGGTCGACTCCAAAGGC-3' (T-T CPD reactions). Substrate was created by mixing primer (5 μM final concentration) with 1.1X molar equivalent template in 25 mM Tris pH 8.0 and 100 mM NaCl, heating to 95 °C for 5 minutes and then cooling to room temperature over 3 hours, protected from light. Assays using these "running start" substrates were performed using substrate to enzyme ratios and time points (indicated in figure legends and text) empirically determined sufficient to keep termination probabilities constant over time. Reaction products were resolved by 10% dPAGE and imaged with a Storm 865 imager. Analysis of band intensity was done using Image Quant TL software (GE Life Sciences) and the termination probabilities, bypass amounts, and primer utilization were quantitated as described previously.[95,97,98] The overall activity on undamaged DNA was

calculated from the primer utilization amounts and normalized for amount of polymerase in the reaction and time. Lesion bypass fidelity assays were performed as previously described as was calculation of mutant frequencies and error rates after sequencing.[95,98,122] Substrates for the lesion bypass fidelity assays were created as above using the template sequence 5'-CCAGCTCGGTACCGGGTTAxCCTTTGGAGTCGACCTGCAGAAATT-3' for 8-oxoG reactions (x is either G or 8-oxoG) or 5'-AGGAAACAGCTATGACCATGATTACGAATTCCAGCTCGGTACCGGGxxAGCCTTTGGAGTCGACCTGCAGAAATT-3' for T-T CPD experiments (x represents either undamaged TT or a *cis-syn* T-T CPD). Regions paired with primer are underlined.

Forward mutation assay

The fidelity of the purified polymerases when copying undamaged DNA was measured utilizing the well-characterized 407 base gap filling forward mutation assay.[163] Reactions contained between 50 and 500-fold excess polymerase over substrate and were incubated at 37 °C for 1 hour. Between 1200 and 4800 plaques resulting from each gap filling reaction were counted for plaque color phenotype and from an unselected subset of the mutants, DNA was amplified using TempliPhi

(GE Life Sciences; following manufacturers recommended procedures) and the filled region analyzed for changes after sequencing. Mutation frequencies and error rates were calculated as described previously.[163]

Results

Active site mutations of pol η

Candidate amino acids for change were determined from previously published reports and sequence alignments of Y-family polymerases.[85,106,107,142,143] Single nucleotide polymorphisms in the candidate region were determined from the NCBI SNP database.[144] The group of mutants created and expressed is presented in Table 1.1 and their location within the structure of the protein is shown in Figure 1.1. Predicted function of SNP's is based on the published structure of human pol η.[144] Other amino acids in these regions play roles in coordination of a magnesium ion required for catalytic activity, contact with the template bases both at the site of nucleotide insertion and slightly upstream, and alignment of the incoming nucleotide.[85] Purification of *E. coli* expressed protein was achieved for wild type and 10 individual amino acid substitution mutants and was consistently greater than 95% pure, as judged from SYPRO Red stained gels. Two attempts at purification of the F17L mutant resulted in inadequate purified protein yields, presumably from

incorrectly folded protein as the expected overexpressed bands was observed in crude cell extracts. As would be expected from preparations of the naturally exonuclease deficient pol η,[118] all preparations were confirmed to be free of detectable mismatched primer:terminus exonuclease activity (A:G mispair; data not shown).

Polymerase activity

To evaluate whether the amino acid substitutions had any effect on the ability to synthesize DNA, we first performed polymerase assays under conditions of substrate excess and calculated the amount of substrate used per minute per μg protein (Table 1.1). In these assays, we empirically determined the substrate to enzyme ratio needed to utilize no more than ~30% of the total substrate within 8 minutes. These values were chosen as they consistently give 'single hit' conditions, as defined by the termination probability of the enzyme at any given position remaining constant over time.[95] The overall polymerase activity of the various preparations (picomole substrate used/min/μg protein) was calculated from the actual primer utilization values. Preps with activity less than 10% compared to wild type were purified a second time to ensure it was not an artifact of the purification process. The various mutants tested ranged from complete lack of detectable

polymerase activity to having several fold more activity. Interestingly, all three of the naturally occurring isoforms of pol η tested (M14V, R81C, and E82D) displayed increased overall activity while only one of the other mutants (Q38A) tested did. Three mutants (Y52E, R61A, and S62G) displayed no change or slightly reduced activity, while Q38V and S62A displayed drastically reduced but still detectable activity. It is interesting that the Q38V and S62A forms had reduced activity while changes to different amino acids at these same residues had near normal or increased polymerase activity (Q38A and S62G; 80% and 210% of wild type, respectively), demonstrating that it is not just changing the wild type sequence but to which residue that needs to be considered. The R55A form, which coordinates the incoming dNTP phosphate group, was devoid of any detectable polymerase activity despite twice being purified in a soluble form. Only forms of pol η that were able to generate completely copied duplex substrates (lesion bypass fidelity assay) and able to completely fill the 407-base gapped plasmid DNA (forward mutation assay) were evaluated further. Therefore, the R55A and F17L forms are not included in subsequent analyses. While we did not perform pre-steady state kinetic analysis to determine the active fraction of each preparation, all proteins were expressed and purified under similar conditions and we assume that the differences in observed

activity are intrinsic to the different amino acid sequence. Repeat preparations of both wild type and selected mutants show similar activities (within 10%; data not shown).

Lesion bypass efficiency

We analyzed all mutants that had requisite levels of polymerase activity in order to determine their ability to bypass either 8-oxo-G or a T-T CPD relative to an otherwise identical undamaged sequence. The goal was to determine if the changes in activity observed on undamaged DNA extended to damaged DNA or if there was a differential effect caused by the lesion containing substrates. Short times and high substrate to enzyme ratios (20:1 to 400:1) were used to ensure single substrate-polymerase interactions.[95,122] All preparations showed the expected low processivity, with none of the mutants being able to insert more than 8-10 nucleotides during a single round of synthesis (Figure 1.2). This is consistent with several published reports of pol η and other Y-family polymerase activity.[97,118,120,147,164] Only Q38A (71% efficiency) and Y52E (100% efficiency) display somewhat reduced ability to bypass 8-oxoG compared to wild type (150% efficiency). All other mutants tested displayed the preferred copying of damaged DNA compared to undamaged DNA that wild type polymerase shows (i.e. bypass efficiency values of greater than 100%).[97,98,127] We

note that even the lowest observed 8-oxoG bypass efficiency seen here is still 2-3 fold higher than *S. cerevisiae* replicative polymerase δ[98] and pol ε.[150] When measuring T-T CPD bypass efficiency, we also observed more efficient bypass of damaged DNA compared to undamaged (i.e. values greater than 100%) for most mutant forms of the protein, although the magnitude of the effect does seem to be diminished somewhat compared to wild type (Figure 1.3B). The only exceptions were Q38A and Y52E (91% and 46%, respectively). Again, these values are still much higher than the values for replicative polymerases.[97,165]

Lesion bypass fidelity

We next investigated the bypass fidelity of these mutant forms of human pol η during synthesis past two common lesions: a *cis-syn* cyclobutane thymine-thymine dimer and 7,8-dihydro-8-oxo-guanine. The assay used employs a partially duplex oligonucleotide with sequence that matches a part of the *LacZα* open reading frame but containing an amber stop codon, within which the lesion is located. After recovery of the synthesized strand, the newly copied DNA is annealed to gapped M13mp18 DNA and transfected into *E. coli*. Inaccurate bypass of either lesion generates a sequence that gives a dark blue M13 plaque phenotype, while accurate bypass causes a readily distinguishable light blue plaque.[95,122] The frequency of dark

blue plaques is therefore an indication of lesion bypass fidelity and detailed spectrum information is obtained from sequencing of DNA from mutant plaques. As can be seen in Table 1.1, of the mutants tested, there were few drastic differences in the overall fidelity of either undamaged or damaged DNA. With the exceptions of Y52E for all three templates and Q38A for the T-T CPD containing template, differences in dark blue plaque frequency were less than 3-fold compared to wild type. This indicates that overall none of these changes cause a major change in the fidelity of pol η. The Y52E mutant has previously been shown to have moderately higher fidelity than the wild type enzyme during synthesis when copying undamaged DNA.[107] but here we show this extends to two different damaged templates as well. This is the first report of this mutant for bypass fidelity opposite 8-oxoG, where we observed an almost 10-fold drop in the dark blue plaque frequency compared to wild type protein (29% for wild type and 3.2% for Y52E). This frequency is very similar to that given by the yeast pol η enzyme in this same assay.[98]

A more thorough analysis of the bypass fidelity is achieved by sequencing of mutant plaque DNA. This allows for a detailed analysis of the spectrum of changes that gives insight into exactly what bases are being inserted during bypass. From this

analysis (Figures 1.4 and 1.5), we see that the increase (2.2% to 7.0% dark blue plaque frequency) by the Q38A form is caused almost solely by an increase in T misinsertions at the 3′T, causing the T→A error rate to increase ~20-fold compared to wild type (17 versus 350 x 10^{-4}, WT and Q38A, respectively) (Figure 1.4B). The T→C error rates are the same as wild type (320 versus 520 x 10^{-4} for WT and Q38A, respectively) (Figure 1.4A). This same error occurs with increased frequency when copying undamaged DNA as well. None of the other mutants gave such a drastic increase in error rate during T-T CPD bypass. Both the Y52E and R61A mutants show more modest decreases in error rate during T-T CPD bypass for both T→A and T→C changes.

When copying template G bases, misinsertion of dATP causes a G→T mutation. Even though this specific error is much less frequent than the most common pol η error, T→C (71 vs 650 x 10^{-4}, G→T and T→C, respectively for WT protein), the Q38A (9 x 10^{-4}), Y52E (4 x 10^{-4}), R61A (20 x 10^{-4}), and S62G (6 x 10^{-4}) mutants all show a decrease in the error rate of at least 3-fold compared to wild type (Figure 1.5). However, of these only the Y52E (490 x 10^{-4}) and Q38A (750 x 10^{-4}) proteins also show a decrease in error rate when copying 8-oxoG (wild type rate of 3500 x 10^{-4}; 7.1 and 4.6-fold decreases, respectively). This supports the idea that Y52E

is in general a higher fidelity enzyme[107] while suggesting that the R61A and S62G changes have more limited effects. It is remarkable that Q38A displays better fidelity than wild type when copying both undamaged G and 8-oxoG, because of the lower fidelity it displays when copying either undamaged or damaged template T. To our knowledge, this is the first report of a single amino acid change in pol η (or any polymerase) that causes a lower fidelity phenotype when copying one lesion (T-T CPD) and a higher fidelity phenotype when copying another lesion (8-oxoG). It is also interesting that the Q38V protein displays neither of these attributes, being essentially like WT for both T-T CPD and 8-oxoG bypass. It should be noted here that none of the SNP's tested (M14V, R81C and E82D) displayed any difference compared to wild type for lesion bypass fidelity of either lesion.

Forward Gap Fidelity

To further investigate the changes in fidelity we observed, a more robust forward fidelity assay was performed. In this assay, a gap of 407 bases is copied by the purified polymerase, allowing determination of a diverse group of changes including all 12 base:base mismatches, insertions, deletions, and other complex mutations.[163] These experiments confirmed and extended the results of the lesion bypass assay (Table 1.2). The overall single base substitution error rate, the most

common of all errors detected, was the lowest for Y52E and R61A forms (~3 -fold reduction compared to WT), similar to the lesion bypass assay. We use a 3-fold change as the minimum required to be considered different. Most samples were less drastically affected and displayed fold changes between 0.5 and 1.1 compared to wild type. That is, not significantly different when all possible errors are grouped together. The spectrum of errors from the various forms did, however, generate some interesting results.

The Q38A protein was more likely to give single base insertions (5.3-fold increase) and complex errors, including tandem base substitutions (3.3-fold increase). This did not extend to single base deletions however. In fact, all of the mutant proteins were less likely to generate single base deletions compared to wild type (0.3 to 0.9-fold decrease), while all but the Y52E protein seemed more likely to generate single base insertions (up to 5.3-fold increase). Again, these changes are relatively small. Interestingly, we never observed a single base insertion in mutant plaques from the Y52E form (28 samples, 11,396 bases analyzed). This is consistent with the larger decrease in complex mutations observed by this form (0.1-fold compared to WT). Further breaking down the single base substitution error rates (Supplementary Table 1.2) shows that each of the different mutant forms of pol η

has a distinct means of achieving the overall fidelity it does. For example, the higher fidelity Y52E form displays reduced error rates at template purines and only modest changes (if any) at template pyrimidines. This is also the case for the R61A protein. Interestingly, the Q38A protein displayed an increase in error rate for two specific changes: C→A and T→A, both of which are caused by dTTP misinsertion (10.4 and 8.2-fold increases compared to WT, respectively). The latter is the same error the Q38A protein made in the lesion bypass reversion assay.

Discussion

We have generated multiple single amino acid substitution mutants of human DNA polymerase η in and around the active site in an attempt to better understand the molecular mechanism by which the unique properties of this polymerase are achieved. Specifically, we are interested in the high efficiency but moderate-to-low fidelity bypass of two common lesions, 8-oxoG and a T-T CPD. By studying both properties with these mutant forms of the polymerase we were able to identify several amino acid residues that affect the overall function of the polymerase and some that have very specifics effects on fidelity. We used a truncated version of pol η (encoding the N-terminal 511 amino acids of the enzyme) as a model for the polymerase catalytic activity. This form of the protein can be

generated in highly pure form and in relatively large quantities using an *E. coli* expression system and is amenable to creation of site-specific mutants. While lacking the C-terminal regions of the polymerase required *in vivo* for ubiquitination and other critical protein-protein interactions,[59,71] an even shorter form of pol η has been used to describe the crystal structure of the catalytic residues and in that report they argue the truncated form is a suitable model of polymerase catalytic activity.[85] We also provide evidence for this, as the wild type of our truncated protein displays very similar biochemical properties to the full-length enzyme for both lesion bypass efficiency and fidelity. To our knowledge, the regulatory and protein interaction domains of the full-length protein have not been shown to alter the nucleotide incorporation properties of the enzyme. Indeed, using yeast proteins we have previously shown that the fidelity and efficiency of both T-T CPD and 8-oxoG are unchanged by the presence of replication accessory proteins.[98,127] One hallmark of pol η is the ability to preferentially copy damaged DNA over undamaged DNA. We feel this property is of critical importance and at times overlooked in its role in suppressing mutagenesis during TLS. Of the amino acid residues investigated, there was not a correlation between lesion bypass efficiency and fidelity. One change (Y52E) caused an increase in overall polymerase fidelity that was accompanied by a

decrease in lesion bypass efficiency opposite both 8-oxoG and a T-T CPD, although the observed efficiency is still greater than that of the replicative polymerases.[98,126] One potential consequence of this observation would be an increased chance of polymerase stalling opposite a lesion during bypass, leading to an increased chance of polymerase switching or extrinsic proofreading. But another form (Q38A) also displayed decreased bypass efficiency of both lesions that was accompanied by a decrease in fidelity opposite one of them (T-T CPD) and an increase in fidelity opposite the other (8-oxoG). Other changes (R61A and S62A) caused varying degrees of fidelity changes when copying either undamaged or damaged DNA but displayed little to no changes in bypass efficiency. From these results we conclude that the bypass efficiency of pol η opposite these two lesions is not entirely dependent on what nucleotide is inserted during the synthesis step. There also did not seem to be a correlation between the overall activity of the enzyme and either fidelity or bypass efficiency. Of note is the fact that none of the three SNP's tested (M14V, R81C and E82D) had any major effect on the polymerase properties. While this is a negative finding, it is important to keep in mind this applies only the outcomes described here. It is possible these SNPs affect the protein in other ways that are undetectable in these assays. Analyzing the in vivo effects of these SNPs as

well as selected other mutants might allow for a more thorough understanding of the connection between bypass efficiency, fidelity and cell fate after DNA damage has occurred.

We found that one of our mutants, R55A, was completely devoid of DNA polymerase activity. This arginine is conserved in Y-family enzymes in archaebacteria, eubacteria, yeast, plants, flies, worms, mice and in multiple human Y-family enzymes.[81,106] This supports the notion it is a critical residue for polymerase function. Based on the recent crystal structures of human pol η[85] and modeling of earlier yeast pol η structures,[107] it appears the positively charged arginine interacts with the negatively charged phosphate of the incoming nucleotide, stabilizing it in the active site. Replacing this residue with an uncharged alanine would be expected therefore to affect polymerization. Evidence for this is contained in a previously reported random mutation screen of human pol η that selected for still active protein.[106] In that experiment, only a single mutant was recovered that contained a mutation at this residue (from arginine to lysine). Presumably this mutant was still active because the change did not remove the positive charge. It will be interesting to assay the properties of a R55L form of pol η as well as other Y-family proteins to test this idea.

The finding that Y52E displays better than wild type fidelity confirms and expands previous work.[107] This residue is also involved in coordinating the incoming nucleotide, so changing from tyrosine to glutamic acid is predicted to destabilize the incoming nucleotide. A mismatched nucleotide requires even more stabilization so it is not surprising that Y52E displays higher fidelity. Our data is the first to demonstrate this quantitatively for lesion bypass as well (Figures 1.4 and 1.5) and supports the idea that pol η copies damaged and undamaged DNA in the same way. These results suggest that the Y52E form of pol η is a general antimutator, copying both damaged and undamaged substrates with higher fidelity. The R61A change has recently been shown (using steady-state kinetics of single nucleotide insertion) to reduce overall polymerase activity as well as misinsertion of dGTP opposite undamaged template T.[85] We see the same effect for both undamaged and the 3′T of a T-T CPD, and also show that it has decreased (relative to wild type) misinsertion of dCTP opposite both undamaged and damaged T's (Figure 1.4). In fact, we see relatively large (at least 5-fold) decreases in the misinsertion of at least one dNTP opposite all four template bases when the R61A protein copies large stretches of undamaged DNA (Supplementary Table 1.2). This arginine residue is conserved in pol η from multiple species.[107] We do not see a similar increase in

fidelity when this mutant copies 8-oxoG (Figure 1.5). Human pol η has yet to be crystallized in complex with 8-oxoG, but it is suspected that the misincorporation of dATP opposite this lesion occurs via Hoogsteen pairing as has been observed for T7 DNA polymerase and pol β.[111,152] We note here that since R61 normally serves to prevent Hoogsteen pairing,[85] our results suggest that perhaps the mechanism of pol η copying past 8-oxoG is in fact different than other higher fidelity polymerases. This idea has also been put forth for yeast pol η, which has been crystallized bound to 8-oxoG templates[162] Of interest here is the observation that in yeast pol η, when a dATP residue is modeled opposite 8-oxoG, there are no steric conflicts, but that in order to be in ideal position for catalysis, the dATP residue would need to "move slightly 'inward'".[162] In yeast pol η, R73 (equivalent to R61 in human) sits on top of the incoming dCTP residue during 8-oxoG bypass. The adjacent M74 (yeast) amino acid is larger than the adjacent S62 (human). It's possible that the smaller residue allows more room for the dATP to move inward, explaining the large increase in dATP insertion by human pol η opposite 8-oxoG compared to the yeast enzyme.[76,91,98,157] Changing R61 to alanine or S62 to alanine or glycine, as we have done here, would allow even more room and hence explain why no increase in fidelity was observed with these mutants.

Perhaps the most interesting finding was the fact that the Q38A mutant displayed both increased and decreased fidelity, dependent on the template being copied. While it displayed an increase in fidelity compared to wild type polymerase opposite both G and 8-oxoG (8 and 5-fold, respectively), its fidelity when copying the undamaged or T-T CPD sequence was the same as wild type for dGTP misinsertion and worse for dTTP misinsertion (6 and 20-fold lower fidelity for undamaged TT and T-T CPD, respectively). The homologous Q55 in yeast pol η stabilizes 8-oxoG in the active site, perhaps explaining the observation that of all the mutants we tested, Q38A displayed the lowest bypass efficiency for 8-oxoG. This same form of pol η was recently shown to display increased stalling when bypassing a T-T CPD[85] and we see a similar result here (Figure 1.3B). We suggest that this may be due to the increased misinsertion of dTTP during bypass (Figure 1.4B). The data presented here shows it also has decreased fidelity, at least for a T→A change. These results indicate that the link between bypass efficiency and fidelity is complicated and possibly lesion specific. The Q38A and R61A results suggest that the ability to bypass either 8-oxoG or a T-T CPD may occur by different molecular interactions between the protein and DNA. This could explain the extraordinarily low fidelity of wild type pol η for 8-oxoG bypass and may mean that other factors (e.g. accessory

proteins, extrinsic proofreading) play a larger role in determining the mutagenicity of TLS on this lesion after bypass. An in vivo analysis of selected mutants is needed to determine if similar changes in the mutation rates in cells after genotoxic exposure occurs in a manner consistent with these in vitro experiments.

Conflict of interest statement:

The authors declare no conflict of interest.

Acknowledgements

We thank Dr. Robert Smart and Dr. Thomas Kunkel for helpful conversations and materials. We thank the reviewers for numerous comments and helpful suggestions to increase the quality of this manuscript.

Funding information

This work was supported by the National Institutes of Health [R01 ES016942 to S.M.] and [T32 ES007046]. Partial support of S.T. was by the College of Agriculture and Life Sciences at North Carolina State University. Funding for open access charge: National Institutes of Health.

Table 1.1 – Activity and lesion bypass fidelity for single amino acid mutants of human pol η

Mutant	Function/Role*	Activity**	Dark Blue plaques (%)***		
			Undamaged	T-T CPD	8-oxoG
Wild Type		170 ± 38	7.1%	2.3%	29%
M14V	Active site magnesium coordination****	500 ± 140 (3.0X)	5.3%	2.2%	27%
F17L	Lid of steric gate	*NA*	*NA*	*NA*	*NA*
Q38A	Van der Waals interaction with template base	350 ± 32 (2.1X)	6.4%	7.0%	14%
Q38V	Van der Waals interaction with template base	16 ± 3 (0.1X)	6.8%	1.7%	33%
Y52E	Interaction with incoming nucleotide	120 ± 35 (0.7X)	1.4%	0.5%	3.2%
R55A	Interaction with phosphate of incoming nucleotides	*ND*	*NA*	*NA*	*NA*
R61A	Hydrogen bonds with incoming nucleotide	170 ± 20 (1.0X)	2.7%	0.9%	24%
S62A	Contact with 5′ base upstream of 1st template base	32 ± 14 (0.2X)	7.4%	2.1%	17%
S62G	Contact with 5′ base upstream of 1st template base	130 ± 20 (0.8X)	3.6%	2.8%	23%
R81C	Structure of protein	420 ± 90 (2.5X)	6.1%	4.3%	26%
E82D	Structure of protein	230 ± 87 (1.4X)	4.0%	2.1%	21%

Table 1.1 continued

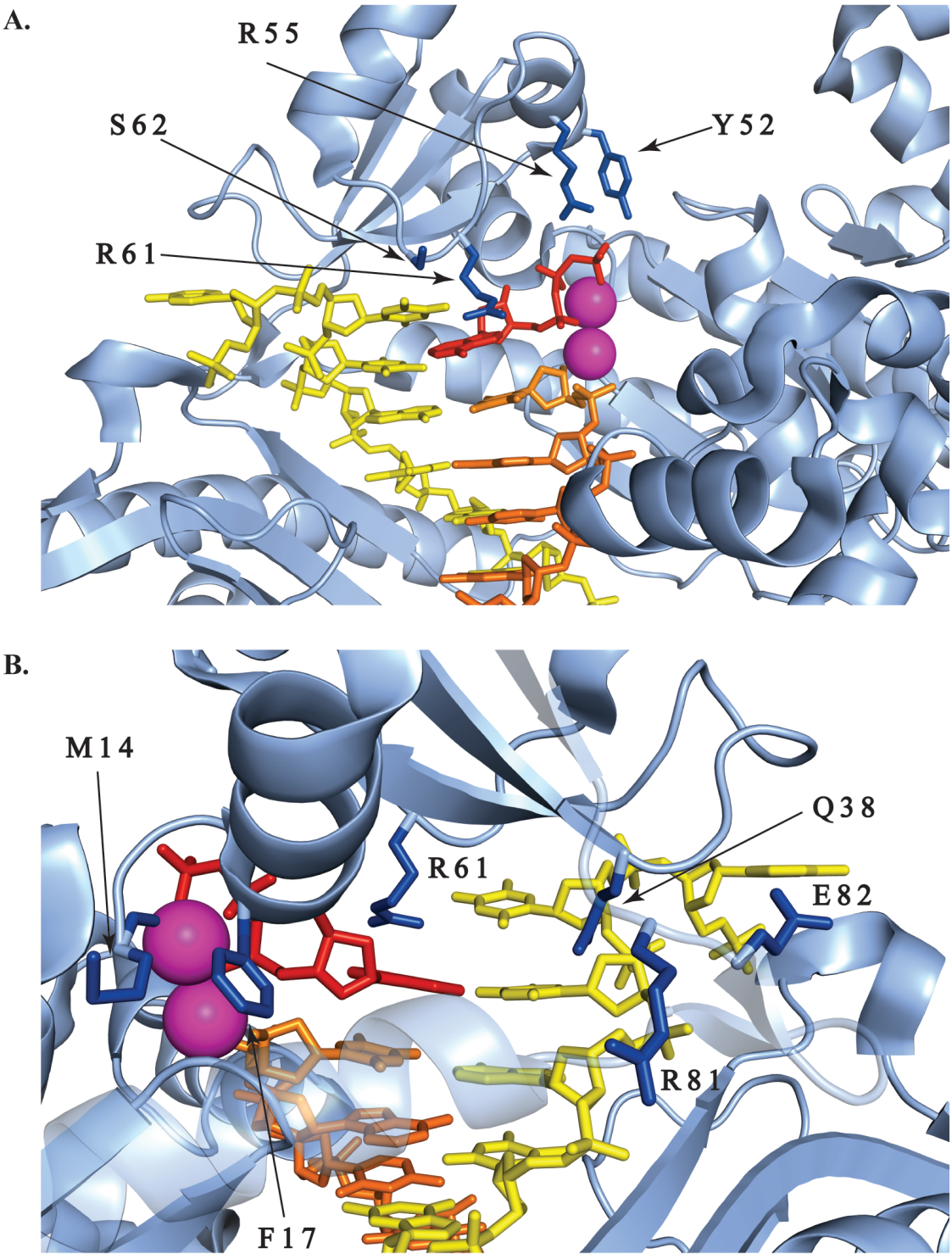
* Predicted function based on published reports and crystal structures

** Picomoles of substrate extended per minute per μg protein. Values in parentheses are relative to WT. NA, not applicable. ND, none detected.

*** Frequency of dark blue plaques observed in lesion bypass assay.

**** This interaction is with the peptide bond, not the side chain.

Figure 1.1 – Location of amino acids altered in this study within the active site of human pol η. DNA template is shown in yellow, primer strand in orange, incoming nucleotide in red and metal ions as magenta spheres. Amino acids of interest are shown in dark blue. **A.** Y52, R55, R61 and S62 residues. **B.** M14, F17, Q38, R61, R81, E82 residues. The light blue transparent helix is 'in front' of the DNA (from the viewpoint of this image) but was lightened so as not to obscure the view. Images were created with the PyMOL Molecular Graphics System[166] using coordinates in PDB: 3MR2.[85]


A.
R55
S62
Y52
R61
B.
M14
Q38
R61
E82
R81
F17

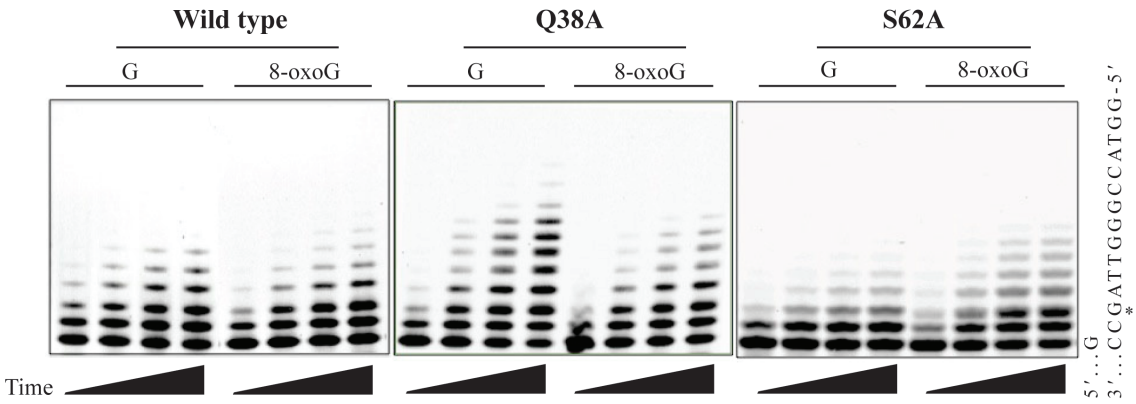
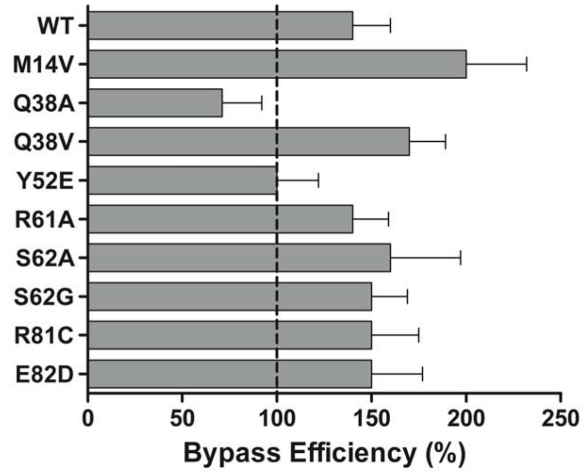


Figure 1.2 – Denaturing polyacrylamide gel electrophoresis based separation of lesion bypass efficiency assay reaction products. Shown are 3 selected forms of the polymerase at time points ranging from 2-8 minutes. After analysis of fluorescently labeled primer strands using ImageQuant TL (GE Life Sciences), reactions were confirmed to be under single interaction conditions, demonstrating constant termination probabilities over time. Primer utilization and bypass amounts were calculated as previously described.[95,122]

Figure 1.3 – Lesion bypass efficiency by pol η mutants. Lesion bypass efficiency assays were performed as described in Materials and Methods. Bypass of either undamaged or damaged template DNA was measured under conditions of single interactions between polymerase and substrate. Substrate:polymerase ratios used ranged from 20:1 (Q38V) to 400:1 (M14V) to account for the differences in overall activity of the various forms. Bypass efficiency is defined as the ratio of damaged to undamaged bypass, i.e. values greater than 100% reflect better bypass of the damaged base. The average bypass efficiency of 4 time points (2, 4, 6, 8 minutes) each from 2 reactions are graphed with the standard deviation shown in the error bars. **A.** Bypass efficiency of 8-oxoG containing templates. **B.** Bypass efficiency of T-T CPD containing templates.

A. 8-oxoG Bypass



B. T-T CPD Bypass

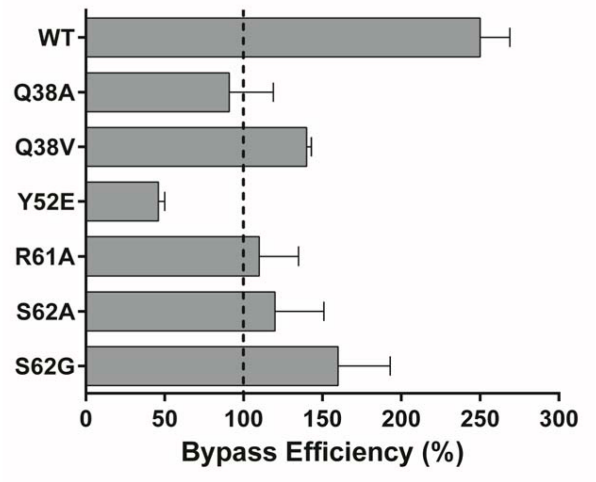
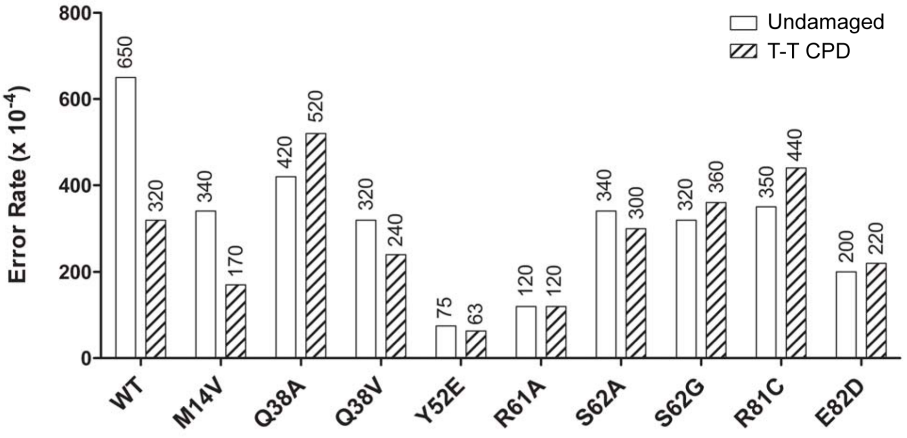


Figure 1.4 – Error rate when copying a *cis-syn* cyclobutane thymine-thymine dimer by pol η mutants. Values given are calculated as described in the Materials and Methods and represent error rates (10^{-4}). Bypass fidelity reactions used a 2:1 substrate:polymerase ratio and were incubated at 37 °C for 30 minutes. Values come from sequencing (for each form) 83-135 dark blue plaques for undamaged DNA and 38-52 dark blue plaques for T-T CPD reactions, with each plaque representing a unique bypass event. **A.** Values for T→C errors (dGTP misinsertion) opposite the 3′T of either undamaged (white bars) or damaged (hatched bars) template DNA. **B.** Values for T→A errors (dTTP misinsertion) opposite the 3′T of either undamaged (white bars) or damaged (hatched bars) template DNA. In both panels the values for wild type pol η are consistent with values generated using full length pol η in the same assay.[97]

A. T to C substitutions



B. T to A substitutions

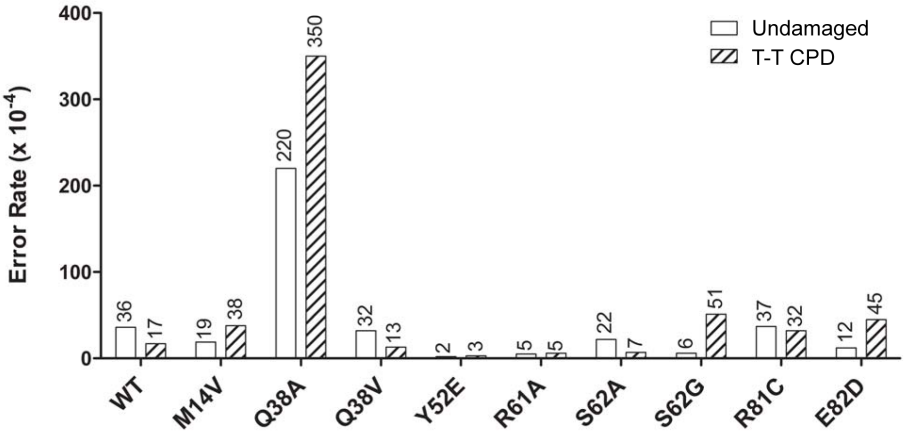


Figure 1.5 – Error rate when copying an 8-oxoG lesion by pol η mutants. Values given are calculated as described in the Materials and Methods and represent error rates (10^{-4}). Bypass fidelity reactions used a 2:1 substrate:polymerase ratio and were incubated at 37 °C for 30 minutes. Values come from sequencing (for each form) 83-135 dark blue plaques for undamaged DNA and 14-48 dark blue plaques for 8-oxoG reactions, with each plaque representing a unique bypass event. The undamaged samples are the same as those described in Figure 4. **A.** Values for G→T errors (dATP misinsertion) opposite the G/8-oxoG base of either undamaged (white bars) or damaged (hatched bars) template DNA. The Y-axis is broken to allow plotting of both undamaged and damaged rates, despite orders of magnitude difference. Only the G→T change is shown, as it is by far the most frequent error when copying 8-oxoG (data not shown and previous work[98]). The values for wild type pol η are consistent with values generated using full length pol η in the same assay.[98]

A. G to T substitutions

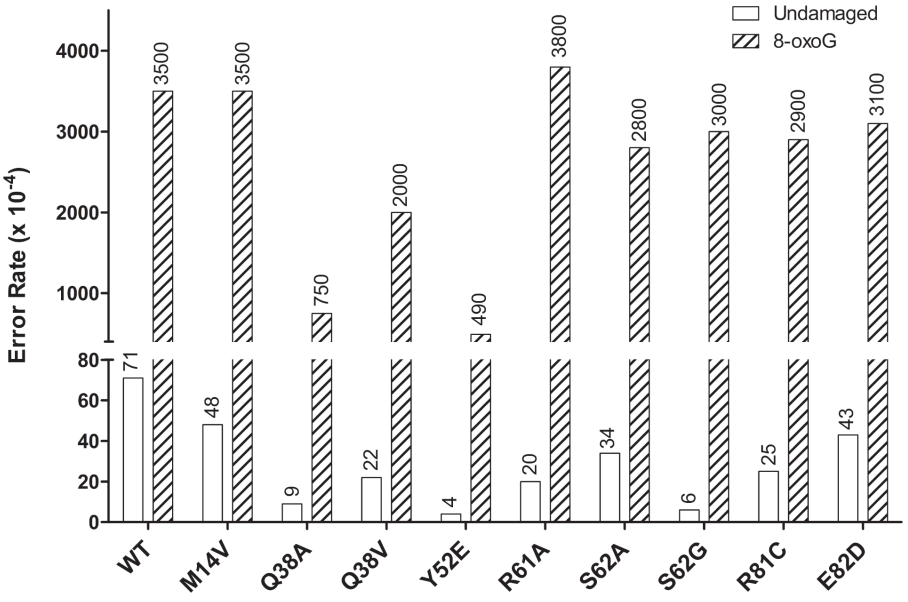


Table 1.2 – Base substitution and insertion/deletion error rates of pol η mutants copying undamaged DNA in a gapped plasmid

	Error rate (x 10^{-4}) *									
Mutant	**Single base substitution**		**Single base deletion**		**Single base insertion**		**Complex****		**Tandem base substitution**	
WT	230		25		3		35		13	
M14V	170	(0.7)	9	(0.4)	5	(1.8)	22	(0.6)	9	(0.7)
Q38A	260	(1.1)	18	(0.7)	16	(5.3)	100	(2.8)	43	(3.3)
Q38V	220	(1.0)	12	(0.5)	9	(3.1)	46	(1.3)	10	(0.8)
Y52E	73	(0.3)	10	(0.4)	≤1	(0.5)	3	(0.1)	1	(0.1)
R61A	74	(0.3)	7	(0.3)	4	(1.4)	7	(0.2)	1	(0.1)
S62A	170	(0.7)	11	(0.4)	5	(1.8)	16	(0.5)	8	(0.6)
S62G	170	(0.7)	14	(0.6)	3	(1.0)	20	(0.6)	8	(0.6)
R81C	180	(0.8)	21	(0.9)	7	(2.3)	36	(1.0)	9	(0.7)
E82D	120	(0.5)	13	(0.5)	8	(2.6)	20	(0.6)	5	(0.4)

* Values given correspond to errors per 10,000 bases copied.

** **Complex errors are defined as multiple changes occurring within 3 bases, such as substitution-deletions and substitution-insertions. The tandem base substitutions error rates shown are a subset of the Complex errors reported. Values in parentheses for mutants are relative to WT rate.**

Supplementary Table 1.1 – Primer sequences used in site-directed mutagenesis reactions to generate single amino acid mutants in the catalytic core of human DNA polymerase η (amino acids 1-511). Sequence differences compared to wild type are underlined.

Mutant	Forward Primer	Reverse Primer
M14V	5′-GTGGTTGCTCTCGTGGACGTGGACTGTTTTTTTGTTC	5′-GAACAAAAAAACAGTCCACGTCCACGAGAGCAACCAC
F17L	5′-GTGGACATGGACTGTTTGTTTGTTCAAGTGGAG	5′-CTCCACTTGAACAAACAAACAGTCCATGTCCAC
Q38A	5′-CCTTGTGCAGTTGTAGCGTACAAATCATGG	5′-CCATGATTTGTAGGCTACAACTGCACAAGG
Q38V	5′-CCTTGTGCAGTTGTAGTGTACAAATCATGG	5′-CCATGATTTGTACACTACAACTGCACAAGG
Y52E	5′-GGAATAATTGCAGTGAGTGAGGAAGCTCGTGCATTTGG	5′-CCAAATGCACGAGCTTCCTCACTCACTGCAATTATTCC
R55A	5′-AGTTATGAAGCTGCTGCATTTGGAGTC	5′-GACTCCAAATGCAGCAGCTTCATAACT
R61A	5′-GCATTTGGAGTCACTGCAAGTATGTGGGC	5′-GCCCACATACTTGCAGTGACTCCAAATGC
S62A	5′-GGAGTCACTAGAGCTATGTGGGCAGAT	5′-ATCTGCCCACATAGCTCTAGTGACTCC
S62G	5′-GGAGTCACTAGAGGTATGTGGGCAGATGGATGC	5′-ATCTGCCCACATACCTCTAGTGACTCC
R81C	5′-TTCTACTGGCACAAGTTTGTGAGTCCCGTGGG	5′-CCCACGGGACTCACAAACTTGTGCCAGTAGAA
E82D	5′-TCTACTGGCACAAGTTCGTGATTCCCGTGGGAA	5′-TTCCCACGGGAATCACGAACTTGTGCCAGTAGA

Supplementary Table 1.2 – Base substitution error spectrum of truncated pol η mutants as calculated from forward gap filling assay. Values given are the calculated frequency of errors per 10,000 bases copied. Rates relative to wild type are given in parentheses. Values in red indicate rates greater than 5 fold above wild type rate. Values in green are rate less than 5 fold the wild type rate. For changes that were not observed, rates were calculated as if a single instance was found and the values presented as less than or equal to (≤) values.

Supplementary Table 1.2 continued

Base	Mutation From → To	Mispair Template·dNTP	Error Rate (x 10^{-4}) WT	M14V	Q38A	Q38V	Y52E	R61A	S62A	S62G	R81C	E82D
A	A → C	A·dGTP	**72**	65 (0.9X)	16 (0.2X)	53 (0.7X)	6 (0.1X)	22 (0.3X)	22 (0.3X)	35 (0.5X)	54 (0.8X)	26 (0.4X)
	A → G	A·dCTP	**60**	27 (0.5X)	40 (0.7X)	21 (0.4X)	6 (0.1X)	11 (0.2X)	38 (0.6X)	35 (0.6X)	38 (0.6X)	16 (0.3X)
	A → T	A·dATP	**150**	65 (0.4X)	88 (0.6X)	126 (0.8X)	6 (0.04X)	22 (0.1X)	22 (0.1X)	52 (0.3X)	87 (0.6X)	58 (0.4X)
	A → N	A·dNTP	**283**	157 (0.6X)	144 (0.5X)	200 (0.7X)	18 (0.1X)	56 (0.2X)	49 (0.2X)	122 (0.4X)	179 (0.6X)	100 (0.4X)
C	C → A	C·dTTP	**5**	18 (3.6X)	52 (10.4X)	13 (2.6X)	≤ 5 (1.0X)	≤ 5 (1.0X)	18 (3.6X)	19 (3.8X)	≤ 4 (0.8X)	9 (1.8X)
	C → G	C·dCTP	**10**	≤ 4 (0.4X)	7 (0.7X)	9 (0.9X)	20 (2.0X)	≤ 5 (0.5X)	≤ 4 (0.4X)	5 (0.5X)	9 (0.9X)	≤ 4 (0.4X)
	C → T	C·dATP	**63**	44 (0.7X)	78 (1.2X)	26 (0.4X)	20 (0.3X)	18 (0.3X)	57 (0.9X)	75 (1.2X)	26 (0.4X)	17 (0.3X)
	C → N	A·dNTP	**78**	62 (0.8X)	130 (1.7X)	47 (0.6X)	39 (0.5X)	18 (0.2X)	75 (1.0X)	99 (1.3X)	40 (0.5X)	26 (0.3X)
G	G → A	G·dTTP	**75**	28 (0.4X)	100 (1.3X)	77 (1.0X)	≤ 6 (0.1X)	6 (0.1X)	11 (0.1X)	54 (0.7X)	28 (0.4X)	22 (0.3X)
	G → C	G·dGTP	**13**	17 (1.3X)	25 (1.9X)	≤ 5 (0.4X)	6 (0.5X)	≤ 6 (0.5X)	23 (1.8X)	18 (1.4X)	17 (1.3X)	11 (0.8X)
	G → T	G·dATP	**13**	17 (1.3X)	42 (3.2X)	11 (0.8X)	6 (0.5X)	6 (0.5X)	6 (0.5X)	24 (1.8X)	11 (0.8X)	16 (1.2X)
	G → N	A·dNTP	**100**	62 (0.6X)	167 (1.7X)	88 (0.9X)	13 (0.1X)	12 (0.1X)	40 (0.4X)	97 (1.0X)	57 (0.6X)	49 (0.5X)
T	T → A	T·dTTP	**20**	22 (1.1X)	163 (8.2X)	38 (1.9X)	10 (0.5X)	23 (1.2X)	40 (2.0X)	24 (1.2X)	18 (0.9X)	21 (1.1X)
	T → C	T·dGTP	**385**	335 (0.9X)	345 (0.9X)	457 (1.2X)	166 (0.4X)	164 (0.4X)	397 (1.0X)	297 (0.8X)	375 (1.0X)	235 (0.6X)
	T → G	T·dCTP	**24**	26 (1.1X)	33 (1.4X)	≤ 4 (0.2X)	24 (1.0X)	5 (0.2X)	9 (0.4X)	8 (0.3X)	31 (1.3X)	38 (1.6X)
	T → N	A·dNTP	**429**	383 (0.9X)	540 (1.3X)	495 (1.2X)	200 (0.5X)	191 (0.4X)	445 (1.0X)	334 (0.8X)	423 (1.0X)	295 (0.7X)
			227	**173 (0.8X)**	**255 (1.1X)**	**215 (0.9X)**	**73 (0.3X)**	**74 (0.3X)**	**172 (0.8X)**	**169 (0.7X)**	**182 (0.8X)**	**122 (0.5X)**

CHAPTER 2

Biochemical analysis of DNA polymerase η fidelity in the presence of replication protein A

Samuel C. Suarez[§], Shannon M. Toffton[§], and Scott D. McCulloch[§,‡]

[§] *Program in Environmental and Molecular Toxicology, Department of Biological Sciences, North Carolina State University, Raleigh, NC 27695*

[‡] *Center for Human Health and the Environment, North Carolina State University, Raleigh, NC 27695*

Abstract

DNA polymerase η (pol η) synthesizes across from damaged DNA templates in order to prevent deleterious consequences like replication fork collapse and double-strand breaks. This process, termed translesion synthesis (TLS), is an overall positive for the cell, as cells deficient in pol η display higher mutation rates. This outcome results despite the fact that the in vitro fidelity of bypass by pol η alone is moderate to low, depending on the lesion being copied. One possible means of increasing the fidelity of pol η is interaction with replication accessory proteins present at the replication fork. We have previously utilized a bacteriophage based screening system to measure the fidelity of bypass using purified proteins. Here we report on the fidelity effects of a single stranded binding protein, replication protein A (RPA), when copying the oxidative lesion 7,8-dihydro-8-oxo-guanine(8-oxoG) and the UV-induced *cis-syn* thymine-thymine cyclobutane pyrimidine dimer (T-T CPD). We observed no change in fidelity dependent on RPA when copying these damaged templates. This result is consistent in multiple position contexts. We previously identified single amino acid substitution mutants of pol η that have specific effects on fidelity when copying both damaged and undamaged templates. In order to confirm our results, we examined the Q38A and Y52E mutants in the same full-

length construct. We again observed no difference when RPA was added to the bypass reaction, with the mutant forms of pol η displaying similar fidelity regardless of RPA status. We do, however, observe some slight effects when copying undamaged DNA, similar to those we have described previously. Our results indicate that RPA by itself does not affect pol η dependent lesion bypass fidelity when copying either 8-oxoG or T-T CPD lesions.

Introduction

DNA replication in the presence of damaged bases requires specialized DNA polymerases in order to prevent more deleterious consequences caused by replicative polymerase stalling.[71] One member of the Y-family, DNA polymerase η (pol η), replicates past UV light induced DNA lesions like *cis-syn* thymine-thymine cyclobutane pyrimidine dimers (T-T CPD) with similar fidelity to that of copying undamaged DNA but with much higher efficiency.[97] Pol η and other Y-family polymerases demonstrate much lower fidelity than replicative polymerases when copying undamaged DNA, and their access to DNA is likely tightly controlled by mechanisms that include (but are not limited to) mono-ubiquitination of the sliding clamp PCNA.[59] In contrast with the similar error rates of 1 error in ~30 insertions when copying T-T CPD and undamaged DNA,[97,118,119] human pol η copies the

ubiquitous oxidative lesion, 7,8-dihydro-8-oxo-guanine (8-oxoG) with an error rate approaching 1 in 2.[91,98,125] Despite this very low fidelity, pol η copies past 8-oxoG more efficiently than it copies undamaged DNA of the same sequence.[98,145,167] This contrasts with the fidelity of *S. cerevisiae* pol η, which copies 8-oxoG with much higher fidelity.[76,98] Despite these error rates, cells deficient in pol η display higher mutation rates when transfected with DNA treated with methylene blue plus visible light, which preferentially creates 8-oxoG lesions in DNA.[79] A similar increase in mutations is seen when XPV cells that are deficient in pol η are exposed to UV light.[73,74] Many possibilities exist to explain this paradox of a polymerase that creates mutations, but whose presence is an overall positive for the cell.

One explanation of how moderate-to-low fidelity bypass by pol η still allows a reduction of mutagenesis is by modification of pol η fidelity by interaction with one or more of the many replication accessory proteins present at a replication fork. However, a long record of DNA replication fidelity studies have shown a less than clear record of interactions with replication accessory proteins that increase polymerase fidelity. When examining the bacteriophage polymerases from RB69, T4 and T7, there is little evidence that replication accessory proteins have large effects on polymerase fidelity.[168-171] Some small changes are observed when examining the

effect of the processivity clamp on *E. coli* Pol III fidelity, but not Pol IV, a lesion bypass polymerase.[172,173] *Thermus thermophilius* single stranded binding protein (SSB) slightly increases the fidelity of the exonuclease deficient *T. thermophilius* polymerase when PCR is performed on the pUC19 plasmid,[174] although an indirect role in protecting the DNA substrate cannot be ruled out in this report. When examining eukaryotic polymerases, the evidence is just as varied. Polymerase α from *S. cerevisiae* exhibits similar mutation frequencies when examining 3 SSBs from yeast, and one of those SSBs resulted in slightly reduced single base deletions when copying 3-5 reiterated nucleotides.[175] Yeast pol α also shows no difference in base substitution fidelity or single base deletion mutant frequency when adding yeast RPA.[176] Pol α from HeLa cell extracts exhibits a ~5-fold reduction in mutation frequencies when copying shuttle vectors in response to addition of human RPA.[177] Calf thymus pol α exhibits reduced terminal misincorporation at pol α pause sites with the addition of RPA,[178] while also reducing misincorporation efficiency about between 5- and 6-fold.[179] Polymerases that accomplish the majority of replication show other effects. PCNA actually increases calf thymus pol δ misincorporations,[180] and PCNA decreases *S. cerevisiae* pol δ fidelity ~2-fold.[181] This contrasts with a report by Fortune et al reporting PCNA and RPA do not decrease base substitutions by *S.*

cerevisiae pol δ, but PCNA and RPA reduce deletions ~10-fold individually and ≥90-fold together.[182] While these reports suggest that replication accessory proteins do have some ability to alter the fidelities of polymerases, they do not speak to the synthesis across from damaged DNA templates, which occurs by polymerases that are significantly different than the replicative polymerases.

Reports examining eukaryotic polymerases involved in TLS have also been varied. The fidelity of the B-family polymerase ζ from *S. cerevisiae*, which plays a role primarily in extending mismatched primer termini, is not affected by the combination of replication factor C (RFC; the 5-subunit PCNA loading complex), PCNA, and RPA.[183] This is similar to results from *S. cerevisiae* pol η when copying both T-T CPD or 8-oxoG in the presence of all three proteins or complexes.[98,127] Both of these reports utilize an assay that requires multiple insertion and extension events past damaged templates. Both reports shows little effect of these proteins on bypass fidelity.[127,183] These reports contrast with work by Maga et al who reported a 6-fold reduction in human pol η misincorporation of dATP when adding just PCNA to a single nucleotide incorporation experiment with 8-oxoG contained within the template. This same misincorporation event is reduced 21-fold when adding both RPA and PCNA.[125] This report utilized polymerase excess over substrate DNA and

single nucleotide insertion kinetics, lacked RFC that could load PCNA onto primer termini, and also lacked any means of blocking PCNA from migrating off the DNA ends. While this is a very interesting result, the results are for PCNA alone as well as PCNA and RPA together, but not RPA alone.

A point that is often overlooked in the discussion of pol η bypass as "error-free" is that *S. cerevisiae* and human pol η differ in 8-oxoG bypass fidelity,[76,91,98] but share similar T-T CPD bypass and undamaged DNA fidelity.[75,97,118,119,147] Taking into account previous reports,[98,125] we wished to determine the contribution of RPA to the fidelity of bypass by human pol η across from 8-oxoG and T-T CPD. We reasoned that the very low fidelity of 8-oxoG bypass we have reported could be mitigated by accessory proteins, despite their apparent lack of ability to do anything to yeast pol η, because the fidelity of the human polymerase by itself is so low. Here we utilized a well-described system that requires both insertion(s) across from the damaged nucleotides as well as extension beyond the lesion by polymerase. Extensive use of this assay and template sequence previously using only polymerase allows us to make a direct comparison of the ability of RPA to modify the fidelity of lesion bypass by human pol η.

Materials and Methods

DNA Oligomer Sequences

Oligonucleotide primers were purchased from Integrated DNA Technologies, Inc (Coralville, IA). Damaged and undamaged templates were purchased from Midland Certified Reagent Co. (Midland, TX). Substrates used for the lesion bypass fidelity assay were as follows. Template sequence is stated. Underlined portions indicate primer annealing. Bold **XX** indicate positions of T-T CPD. Bold **Y** indicates either undamaged G or 8-oxoG. Primers were purchased with cy5 5' end labeling from IDT. 75 mer templates: 5'-biotin-

AGGAAACAGCTATGACCATGATTACGAATTCCAGCTCGGTACCGGGTTA**Y**CC

TTTGGAGTCGACCTGCAGAAATT-biotin.

5'-biotin-

AGGAAACAGCTATGACCATGATTACGAATTCCAGCTCGGTACCGGG**XX**AGC

CTTTGGAGTCGACCTGCAGAAATT-biotin.

75 mer upstream: 5'-biotin-

CCAGCTCGGTACCGGGTTA**Y**CCTTTGGAGTCGACCTGCAGAAATTCACTGGC

CGTCGTTTTACAACGTCGTGACT-biotin.

45 mer: 5'-CCAGCTCGGTACCGGGTTA**Y**CCTTTGGAGTCGACCTGCAGAAATT.

Templates and primers were resuspended in ultrapure H_2O. Substrates were created by mixing primer (5 μM final concentration) with 1.1x molar equivalent template in 25 mM Tris pH 8.0 and 100 mM NaCl, heating to 85 °C for 5 minutes and then cooling to room temperature over 3 hours, protected from light.

Recombinant Proteins

Pol η

The expression vector for n-terminally 6xHis-tagged human pol η (pJM879) and *E. coli* expression strain (BL21 [DE3] derivative RW644) was graciously provided by Dr. Roger Woodgate (NICHHD). Expression and purification was modified from a previously published protocol.[184] RW644 cells were transformed with pJM879 vector using electroporation and plated onto LB media containing 30 μg/mL kanamycin. Isolated colonies were picked and used to inoculate a 12 mL culture of LB with kanamycin. This culture was grown to saturation overnight at 37° C with constant shaking and used to inoculate 1L LB with kanamycin. Cultures were grown for 5 hours at 37° C with constant shaking, after which, cells were centrifuged, washed with PBS, and resuspended in PBS in a volume equal to that of the cell pellet. Resuspended cells were placed drop wise into liquid N_2, freezing them into 3-5mm spheres, which were pooled and stored at -80° C until use. Cell spheres were lysed

using a SPEX Sample Prep 6870 Freezer/Mill. Cells were cooled in liquid N_2 for 10 minutes, and then subjected to 6 cycles of 1 minute grinding (10 impacts per second) followed by 1 minute rest. The resulting lysed cell powder was resuspended in lysis buffer containing 50 mM Tris pH 7.5, 300 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoethanol, and 10% glycerol and was supplemented with 0.2 mM PMSF and Roche Complete Protease inhibitor tablets (1 per 50 mL volume). The crude cell lysate was sonicated (Branson 250 sonifier; output-2, duty cycle 50%) 12-16 times for 30 seconds with one minute on ice in between cycles. Cell lysate was centrifuged for 30 minutes at 40,000g, 4° C and the soluble fraction was passed through a Nalgene filter unit (.45 μM). The filtered lysate was applied to a 5 mL HiTrap Q FF (GE Life Sciences) and 5 mL HiTrap Chelating HP column (GE Life Sciences) charged with $NiSO_4$, connected in sequence, and equilibrated in lysis buffer using an AKTA Purifier (GE Life Sciences). After loading, the HiTrap Q FF was removed and the remaining chelating column was washed with 5 column volumes Wash 1 (50 mM Tris pH 7.5, 10% glycerol, 1 M NaCl, 20 mM imidazole, and 10 mM β-mercaptoethanol). The column was then washed with 5 column volumes Wash 2 (10 mM Na-Pi pH 7.7, 300 mM NaCl, 10% glycerol, 20 mM imidazole, and 10 mM β-mercaptoethanol). The protein was eluted with a step gradient (Wash 2 – Buffer H)

consisting of 4 column volumes at 25% Buffer H followed by 6 column volumes of 100% Buffer H (10 mM Na-Pi pH 7.7, 300 mM NaCl, 10% glycerol, 200 mM imidazole, and 10 mM β-mercaptoethanol). Fractions with highest enrichment for pol η (as determined by SYPRO Red-stained SDS-PAGE) were pooled and loaded directly to a 5 mL BioRad Hydroxyapatite (Bio-Scale Mini CHT Type I) column. The column was washed with 5 column volumes of Buffer H and then an 8 column volume gradient of 10-125 mM sodium phosphate in Buffer H was applied, followed by 2 column volumes of 200 mM sodium phosphate in Buffer H. Fractions with the highest enrichment of pol η were buffer exchanged against Buffer M (20 mM NaPi pH 7.3, 10% glycerol, 100 mM NaCl, and 10 mM β-mercaptoethanol) using an Amicon Ultra-15 centrifugal filter unit (30,000 NMWL) (EMD Millipore). The sample was then applied to a 1mL Mono S 5/50 GL column (GE Life Sciences). The column was washed with 8 column volumes of Buffer M, then with Buffer M containing 200 mM NaCl. Protein was eluted using a gradient of 200-600 mM NaCl in Buffer M. Peak fractions were flash frozen in liquid N_2 in aliquots and stored at -80° C until further use.

Single amino acid substitution mutants were produced by mutating the pJM879 vector using the Agilent Technologies Quikchange II XL site directed

mutagenesis kit according to manufacturer instructions. Primers were purchased from and synthesized by IDT. Primers, with codon changes underlined, were as follows: Q38A-F; 5'-TAAACCGTGCGCGGTTGTCGCATATAAAAGCTGGAAAGGG, Q38A-R; 5'-CCCTTTCCAGCTTTTATATGCGACAACCGCGCACGGTTTA, Y52E-F; 5'-GGTGCCATTATCGCAGTTTCTGAGGAAGCGCGCGCCTT, Y52E-R; 5'-AGGGCGCGCGCTTCCTCAGAAACTGCGATAATGCCACC.

C-terminal 6xHis-tagged truncated human pol η (aa 1-511) was purified as previously described.[145,167]

RPA

The pTYB-RPA vector encoding human RPA was graciously provided by Dr Yue Zou (East Tennessee State University). Purification of the chitin binding protein-intein-RPA70 fusion protein and RPA32, RPA14 subunits was purified as described below, a modification of a previously published report.[185]

BL 21 [DE3] *E. coli* cells were transformed by electroporation and plated onto LB media with 100 μg/mL ampicillin and 50 μg/mL chloramphenicol. An isolated colony was used to inoculate 100 mL 2XYT containing 100 μg/mL ampicillin and 50 μg/mL chloramphenicol. Cultures were grown to saturation overnight with shaking

at 37° C. 20 mL of saturated cultures were used to inoculate each of 4 x 1 L 2XYT media with ampicillin and chloramphenicol and were grown to an OD_{600} of ~0.6. IPTG was added to a final concentration of 1 mM and cultures were grown an additional 3 hours at 25° C, after which cells were harvested by centrifugation, washed with PBS, pooled, and resuspended in a volume equal to that of the cell pellet. Resuspended cells were placed drop wise into liquid N_2 and ground into powder as described above. The resulting lysed cell powder was resuspended in buffer A (20 mM Tris-Cl pH 8.0, 1 M NaCl, .1 mM EDTA, .1% Triton X-100, 1 mM DTT, 10% glycerol, 0.2 mM PMSF and Roche Complete Protease inhibitor tablets [1 per 50 mL volume]). Cell lysate was sonicated as described above and was centrifuged for 30 min at 15,000g, 4° C. 5 mL of chitin beads were equilibrated in buffer A and added to the cleared lysate. This mixture was rotated for 30 minutes at 4° and applied to an empty BioRad Econo-Pac chromatography column (1.5 x12cm, 30mL total volume). Flow was by gravity only. Resin was washed with 100 mL of buffer B (20 mM Tris-Cl pH 8.0, 800 mM NaSCN, 0.25 mM EDTA, 0.01% NP-40, 1 mM DTT, 10% glycerol). The resin was flushed with 25 mL cleave buffer (20 mM Tris-Cl pH 8.0, 500 mM NaCl, 0.1 mM EDTA, 10% glycerol) with 30 mM DTT, sealed, and stored at 4° C overnight. Protein was eluted with 20 mL cleave buffer

containing only 1 mM DTT, and was collected in 1 mL aliquots. Selected peak fractions (as determined by SDS-PAGE) were pooled and buffer exchanged against buffer C (25 mM HEPES-OH pH 7.8, .25 mM EDTA, 1 mM DTT, and 10% glycerol) with 50 mM KCl using an Amicon Ultra-15 centrifugal filter unit (10,000 NMWL) (EMD Millipore). The sample was then applied to a 1 mL Mono Q 5/50 GL column (GE Life Sciences) using an AKTA Purifier (GE Life Sciences), and was washed with 20 mL buffer C with 50 mM KCl. Protein was eluted using a 30 mL gradient from 50 to 500 mM KCl in buffer C. Peak fractions (as determined by SDS-PAGE) were buffer exchanged to buffer C with 150 mM KCl by the same procedure as before. Buffer exchanged sample was flash frozen in liquid N_2 and stored in aliquots at -80° C until use.

Lesion bypass fidelity assay

Cell strains, bacteriophage reagents, and protocol have been previously described.[122] All reactions were performed in 40 mM Tris pH 8.0, 250 μg/mL BSA, 10 mM DTT, 10 mM $MgCl_2$, 60 mM KCl, and 1.25% glycerol. Reactions were supplemented with 100 μM final concentration of each dNTP. Restriction enzymes and nucleotides were purchased from New England Biolabs (Ipswich, MA). Reaction volumes for the lesion bypass assay were 50 μL. Reactions contained 10

pmoles substrate, 10 pmoles polymerase, and 50 pmoles RPA. RPA was added first and pre-incubated at 37° C for 30 minutes and synthesis was initiated by addition of polymerase. Reactions were incubated at 37° C for an additional 30 minutes. Reactions were stopped with the addition of 2 μL of 500 mM EDTA, and were processed according to previously described protocol.[122] After recovery, newly synthesized oligos were annealed to gapped M13mp18 bacteriophage DNA (10-25 fold excess oligo over phage). Annealed gap DNA was transformed into MC1061 cells as previously described and plaque phenotype and numbers were counted. Dark blue plaques were amplified using Templiphi (GE Life Sciences) according to manufacturer protocol and resulting DNA was sequenced by Genewiz (Research Triangle Park, NC). Error rates were calculated as previously described.[122]

DNA binding

DNA binding experiments were performed in 40 mM Tris pH 8.0, 250 μg/mL BSA, 10 mM DTT, 10 mM $MgCl_2$, 60 mM KCl, and 1.25% glycerol. Reactions containing RPA were incubated at 37°C for 30 minutes. Reactions with RPA and polymerase were incubated at 37° for 30 minutes after the addition of RPA, and for an additional 5 minutes at 37° upon the addition of polymerase. Reactions with RPA, polymerase, and dNTPs were incubated at 37° for 30 minutes after the

addition of RPA, 5 minutes at 37° C upon the addition of polymerase, and for an additional 30 minutes at 37° C with the addition of 100 μM final concentration of each dNTP. All reactions were stopped by the addition of 0.5 volumes of ice-cold 40% glycerol. The sample was placed on ice and loaded into a 6% acrylamide (19:1 acrylamide:*bis*-acrylamide) gel. Gels were imaged with a Storm 865 imager (GE Life Sciences).

Results and Discussion

There are conflicting reports on the ability of the replication accessory protein RPA to increase the fidelity of polymerase h during bypass of an 8-oxoG.[98,125] We have previously shown, using an assay that requires multiple incorporations of all four dNTPs, that bypass of 8-oxoG by human pol η occurs with very low fidelity, with dATP being stably misincorporated ~50% of the time.[98] We also demonstrated that the addition of RPA, RFC and PCNA had no effect on the fidelity using the yeast forms of these proteins. However, Maga et al have reported that addition of RPA and PCNA greatly increase the incorporation of dCTP by both human pol η and pol λ.[125] Possible reasons for these conflicting reports include the use of single nucleotide insertion assays versus our assay that utilizes all four nucleotides in

competition, as well as the possibility that yeast and human pol η are different in this respect.

In order to directly test these possible explanations, we used our in vitro, color based screening assay of 8-oxoG bypass fidelity to measure the fidelity of human pol η in the absence and presence of human RPA. We first determined the fidelity of pol η in the presence of RPA when copying 8-oxo-G using a truncated protein that contains the first 511 amino acids, including the catalytic core of the polymerase.[145,167] As shown in Table 1, the frequency of dark blue plaques, the presence of which indicates an error during lesion bypass, is unchanged when RPA is included in the reaction. The observed value of 33% is the same as previously published values.[145] After sequencing mutant plaques, we calculated an error rate of 5000 x 10^{-4} for dATP misinsertion, corresponding to roughly equal incorporation of dATP and dCTP. This is the same rate that has been reported multiple times for pol η alone using this assay as well as steady state single nucleotide insertion kinetic assays.[91,98,125] An error rate of 3400 x 10^{-4} was calculated when a longer substrate was used, and also when adding heat inactivated RPA to the reaction (Supplementary Table 2.1). Changes of at least 3-fold are the threshold that we consider different in

this assay. These data suggest that in this assay, RPA has no detectable effect on the fidelity of 8-oxoG bypass by human pol η.

Since pol η and RPA have not been reported to physically interact, we hypothesized one way that RPA could affect pol η fidelity is the manner in which RPA binds to DNA. We first confirmed RPA binding to a DNA substrate containing either 49 or 51 bases of single stranded template region using non-denaturing PAGE. Representative substrate diagrams are shown in Figure 2. The substrates used contain either a 24 or 26 base region of primer:template duplex on a 75 base long template, giving either 49 or 51 bases of single stranded region in which RPA can bind, well above its largest binding footprint of ~30 bases.[14] As seen in Figure 1A, RPA does indeed bind our substrate DNA, and heat inactivation of RPA (85°C, 15 minutes) abolishes the interaction. Increasing amounts of RPA show increased amount of binding (Figure 1B), and we observe that the presence of damage does not seem to change binding to our substrates (Figure 1C). Importantly, RPA and pol η can also bind to the substrate simultaneously, as seen by the supershift present in Figure 1D. Additionally, synthesis by pol η appears to displace RPA from the DNA, as the addition of deoxynucleotides to the reaction shows a near complete

abolishment of RPA binding the DNA. Presumably the creation of fully duplex DNA is the cause of this lack of binding.

While we cannot definitely say from the gel shift experiments that pol η and RPA physically interact, we recognize the limitations of using a truncated protein that excludes the C-terminal 202 amino acids. Therefore, we then expressed and purified a full-length pol η in an *E. coli* expression system[184] and performed fidelity measurements as described above. As can be seen in Figure 3A, the addition of RPA gives error rates the same as when polymerase only was used (3500×10^{-4} and 4000×10^{-4} for polymerase alone and with RPA, respectively). Heat inactivation of RPA once again shows no difference compared to either polymerase alone or functional RPA (error rate of 3000×10^{-4}). We also examined if adding RPA was able to alter the fidelity when copying a T-T CPD. Using this assay, it has been previously reported that the most common mutations at the 3'-T are changes to C, with changes to A being less frequent, and to G occurring least frequently.[97,127,145,167] Here, we see that the addition of RPA gives error rates the same as those obtained in its absence (Figure 3B). From these data, we conclude that RPA does not change the fidelity of full-length human pol η when copying either 8-oxoG or a T-T CPD.

Given the striking enhancement of fidelity reported by Maga et al,[125] we sought an explanation for how RPA might affect the fidelity of the polymerase when it seems likely that it will have been displaced from the lesion DNA prior to polymerase copying of the lesion. We reasoned that possibly the binding of RPA caused a change in the conformation of the DNA such that the *anti* versus *syn* form was favored for 8-oxoG, allowing more frequent dCTP incorporation. Since we were unable to use the same sequence as that report (it is incompatible with the *LacZ* based screening we use in this assay), we instead created a substrate with a suitable sequence that contained the 8-oxoG lesion further away from the primer terminus, and somewhere within where RPA is expected to bind the substrate (Figure 2). This substrate requires synthesis of many more undamaged bases prior to lesion bypass. In this context, the error rates observed for 8-oxoG bypass were 3300×10^{-4} without RPA, 4000×10^{-4} with RPA, and 3200×10^{-4} with heat inactivated RPA (Figure 3B) These numbers are essentially the same as those seen in the "close" position context and also again show that the addition of RPA does not change the fidelity of the polymerase (Figure 3).

We have recently reported on several single amino acid mutants of pol η that display altered fidelity.[145,167] Given that we failed to see an alteration of fidelity by

the addition of RPA, we propose that the main determinant of pol η dependent 8-oxoG bypass fidelity is the properties of the polymerase itself. To test this idea, we generated full-length pol η that contained 2 of the single amino acid substitutions that we have described as having altered fidelity[145] In the truncated construct, both the Q38A and Y52E mutants show specific changes to fidelity for different errors, depending on the template being copied. If it is only the properties of the polymerase that determine fidelity, we should see the same altered fidelity signature whether RPA is present or not.

Both Q38A and Y52E showed reduced error rates for the most common 8-oxo-G to T errors when compared to wild type polymerase.[145] This effect is recapitulated using the full-length protein (Figure 4A), providing further evidence that the fidelities of full-length and truncated pol η are similar (see also Supplementary Table 2.1). Q38A and Y52E display error rates about 3-fold lower than wild type, and the addition of RPA does not change these values. Additionally, the signature of the Q38A mutant of increased 3'-T to A changes compared to wild type is present and maintained in the presence of RPA (Figure 4B). The Y52E mutant shows a reduction in 3'-T to A changes compared to wild type and is also unaffected by the addition of RPA. The reduction of 3'-T to C changes observed in the Y52E

mutant is also maintained regardless of RPA status (Figure 4C). Overall, the signatures we published for the truncated mutants hold true for the full-length mutants, and the addition of RPA does not alter these signatures when copying damaged DNA templates.

Pol η also plays a role in copying structures other than base lesions,[128,129,133-135,137] so we also examined the ability of RPA to the alter the fidelity when copying undamaged DNA. Interestingly, we see four instances of at least 3-fold change. For changes at undamaged T to A, we see a reduction in error rate from 210 x 10^{-4} to 64 x 10^{-4} when adding RPA to reactions containing the Q38A form of pol η, a difference of just over 3 fold. A similar 3 fold reduction was seen for T to C changes in the Y52E reaction when adding RPA (60 x 10^{-4} to 18 x 10^{-4}). This same reduction level of was seen for wild type when examining undamaged G to T changes when adding RPA (120 x 10^{-4} to 40 x 10^{-4}). For the Q38A mutant, this difference was even more pronounced when examining G to T changes. We observed an over 6 fold reduction, from 130 x 10^{-4} to 21 x 10^{-4} for this mutant. To understand the significance of these changes, we must examine the underlying calculation of these error rates. For the Y52E T to C change, in polymerase only reactions they accounted for 23 of the 36 sequenced plaques (64%), and for 15 out of 30 (50%) when RPA was present.

Therefore, we consider this change real and not the result of low sample number. However, when examining Q38A undamaged error rates for T to A and G to T, the picture is more nuanced. For each of these errors, at least 1 change was observed out of a total of 30-32 sequenced mutant plaques but the sample numbers were not as great. For the T to A change, we found 5 of 32 (16%) for polymerase only, and 3 out of 30 (10%) with RPA. For G to T changes, there were 3 out of 32 (9%) for polymerase alone and 1 out of 30 (3%) with RPA. Since these changes (G to T, T to A) are relatively rare events on undamaged DNA, we would need to sequence significantly more samples to achieve adequate numbers of each type to make more definitive statements on the significance of these changes.

While some of the observed differences in single base substitutions when RPA is added may be a result of low numbers of observed mutants, we also see interesting patterns for this mutant when examining two other types of changes. Due to slight translational read through of the stop codon in the LacZ sequence used in this assay, frameshift errors at the stop codon are detectable as they result in true colorless plaques, rather than light blue (and opposed to the dark blue of true reversion mutations).[95] When examining the colorless plaque frequency for Q38A experiments, the colorless frequency is reduced 2 to 3-fold with the addition of RPA

when examining 8-oxoG and undamaged templates, but remains similar when examining using T-T CPD containing substrates (Supplementary Figure 2.2). Another advantage of utilizing an assay that requires both insertions across from the damaged nucleotides as well as extension beyond the lesion is the ability to detect complex changes, such as tandem base substitutions, which pol η is known to make.[118,119,127,145,167] When looking at complex error rates, we note that Q38A again shows a similar pattern to that of frameshift errors. Complex errors at 8-oxoG are decreased greater than 7 fold with the addition of RPA and are decreased slightly more than 3 fold with the addition of RPA at undamaged bases. This reduction is not seen for complex changes when examining T-T CPD (Supplementary Figure 2.3). The glutamine residue at position 38 in pol η contacts the template base in the active site of the protein,[85] and abolishment of this interaction by substitution with alanine increases the number of complex errors seen when copying undamaged DNA.[145] It is possible that RPA binding to DNA somehow stabilizes the template DNA in the active site for this mutant polymerase, reducing the occurrence of these errors when copying 8-oxoG and undamaged DNA. This stabilization may not be evident with wild type pol η due to intact template contact[85] or with a T-T CPD, which has two

templates bases physically crosslinked.[108] This is reminiscent of effects seen on mutant polymerases of the bacteriophage RB69 [171].

This reduction in error rate, if real, could play a role in reducing the numbers of errors introduced when copying undamaged DNA, for example when pol η is involved in DNA copying during somatic mutation of immunoglobulin genes, homologous recombination, or copying of other complex undamaged DNA structures. We find it interesting that a similar reductions in error rate specifically when copying undamaged DNA has been previously reported for single amino acid substitution mutants in the little finger region of pol η[167] and a reduction in undamaged DNA error rate was also observed in other active site mutants.[145] Combined, these results suggest that the error rate of TLS by wild type pol η, at least for 8-oxoG and a T-T CPD, is controlled largely by the structure of the polymerase active site. Mutations that affect fidelity do so in large part by changing the fidelity on undamaged DNA synthesis while leaving the fidelity of TLS largely untouched. We maintain that the low-to-moderate fidelity of pol η mediated TLS does indeed represent a tradeoff between the risk of introducing base substitution or small frameshift mutations and the risk of having a stalled replication fork and/or un-

replicated stretches of DNA that can lead to aberrant (and much more likely to be mutagenic) processing, like non-homologous end joining.

An unresolved issue we have not yet addressed regarding the ability of replication accessory proteins to affect TLS fidelity is the role of PCNA. In Maga et al, they tested both RPA and PCNA in combination.[125] McCulloch et al also tested the combination of RPA, PCNA, and the clamp loading complex RFC using yeast proteins.[98] While McCulloch et al reported no change in fidelity, it is still possible that in the human system PCNA will play a role in altering fidelity. To these ends, it will be interesting to see if the combination of RPA and PCNA has a combined effect on human pol η, and also whether unmodified and mono-ubiquitylated PCNA have different properties in this regard.

Conflict of interest statement:

The authors declare no conflict of interest.

Acknowledgements

We thank Dr. Roger Woodgate for sharing of the full-length pol η expression vector and cell strains. We thank Dr. Ekaterina Frank for discussion in optimizing the purification. We thank Dr. Yue Zou for the provision of the RPA expression vector.

Funding information

This work was supported by the National Institutes of Health [R01 ES016942 to S.M.] and [T32 ES007046].

Table 2.1 – The effect of RPA on the lesion bypass fidelity of truncated pol η.

Dark blue reversion frequencies and error rates (10^{-4}) from the lesion bypass assay on 45mer templates containing 8-oxoG. Values calculated as previously described.[122] Error rates result from sequencing of 38 and 22 dark blue plaques, respectively.

		Dark Blue Plaque Frequency	Error Rate (10^{-4})
	RPA	**8-oxoG**	**8-oxoG to T**
η-511	-	*29%* [a]	*3500* [a]
	+	33%	5000

[a]Data in italics previously published.[145]

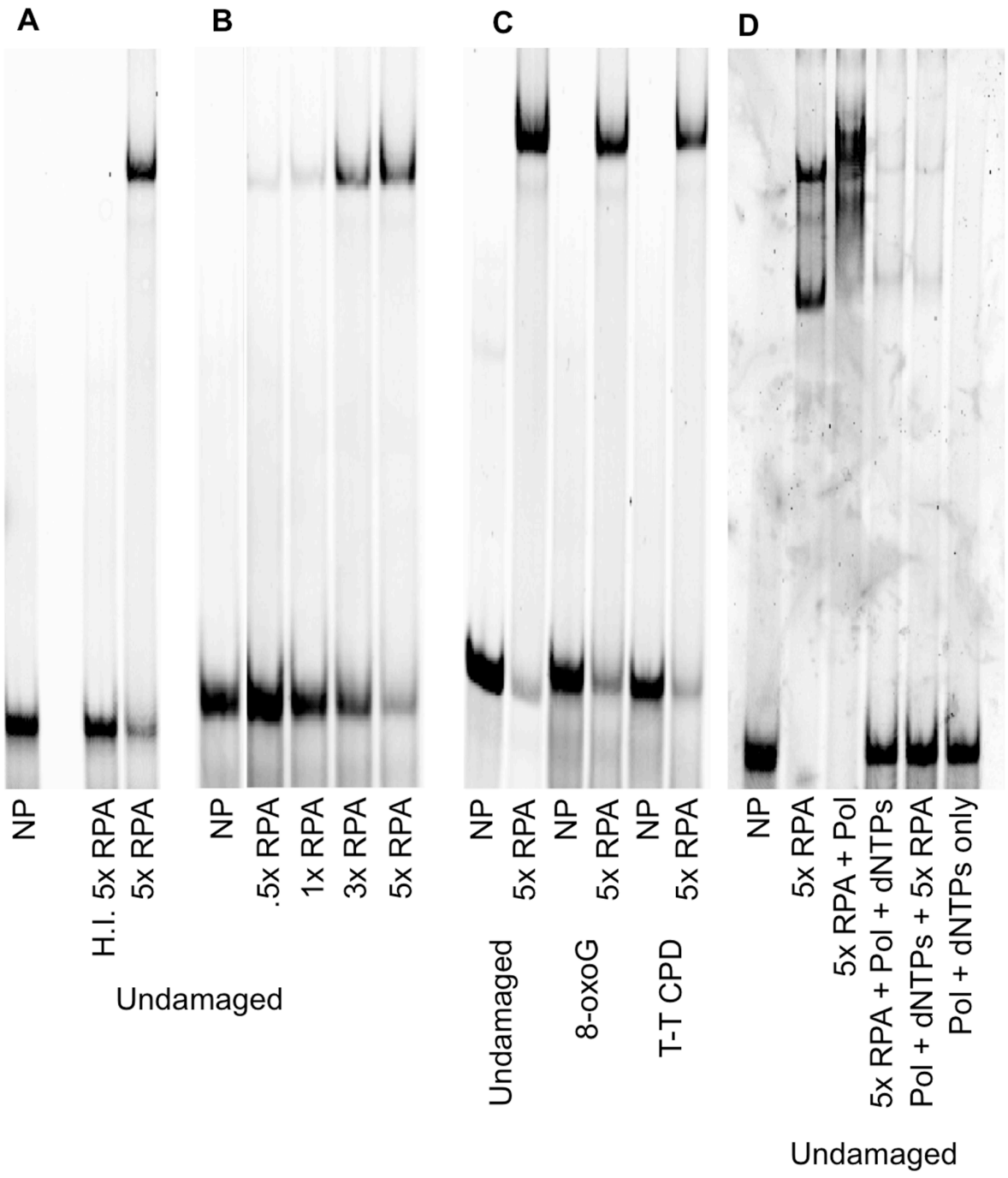
Supplementary Table 2.1 - Truncated and full-length pol η comparison. Dark blue plaque frequencies and error rates (10^{-4}) from lesion bypass fidelity assay on 75mer templates. Templates contained 8-oxoG. H.I. indicates addition of heat inactivated RPA that was created by heating an aliquot of RPA to 85°C for 15 minutes. Values for full-length represent the average of 2 independent experiments. Values calculated as previously described.[122] Error rates result from sequencing between 23 and 47 dark blue plaques.

		Dark Blue Plaque Frequency	Error Rate (10^{-4})
	RPA	**8-oxoG**	**8-oxoG to T**
η-511	-[b]	*29%* [a]	*3500* [a]
	+[b]	33%	5000
	+	25%	3400
	H.I.	23%	3400
η-713	-	28%	3500
	+	32%	4000
	H.I.	28%	3000

[a]Data in italics previously published.[145]

[b]Data collect on a 45mer template

Figure 2.1 – RPA binding to substrates used in the lesion bypass fidelity assay. NP indicates no protein. HI RPA indicates RPA that was heated to 85° for 15 minutes. X RPA indicates fold excess RPA over DNA substrate. Substrates contain 8-oxoG or T-T CPD as indicated. All RPA was bound for 30 minutes at 37° **A.** RPA binding to undamaged DNA and heat inactivation of RPA abolishes binding. **B.** Titration of RPA binding to undamaged DNA. **C.** RPA binding to DNA containing 8-oxoG and T-T CPD. **D.** Consecutive binding of RPA and pol η. RPA is displaced by pol η synthesis. RPA is unable to bind double-stranded DNA after DNA synthesis. All images run on 6% acrylamide:TBE gels. Substrates were 5' end labeled and imaged on a Storm 865 imager (GE Life Sciences).


A
B
C
D
NP
H.I. 5x RPA
5x RPA
Undamaged
NP
.5x RPA
1x RPA
3x RPA
5x RPA
NP
5x RPA
NP
5x RPA
NP
5x RPA
Undamaged
8-oxoG
T-T CPD
NP
5x RPA
5x RPA + Pol
5x RPA + Pol + dNTPs
Pol + dNTPs + 5x RPA
Pol + dNTPs only
Undamaged

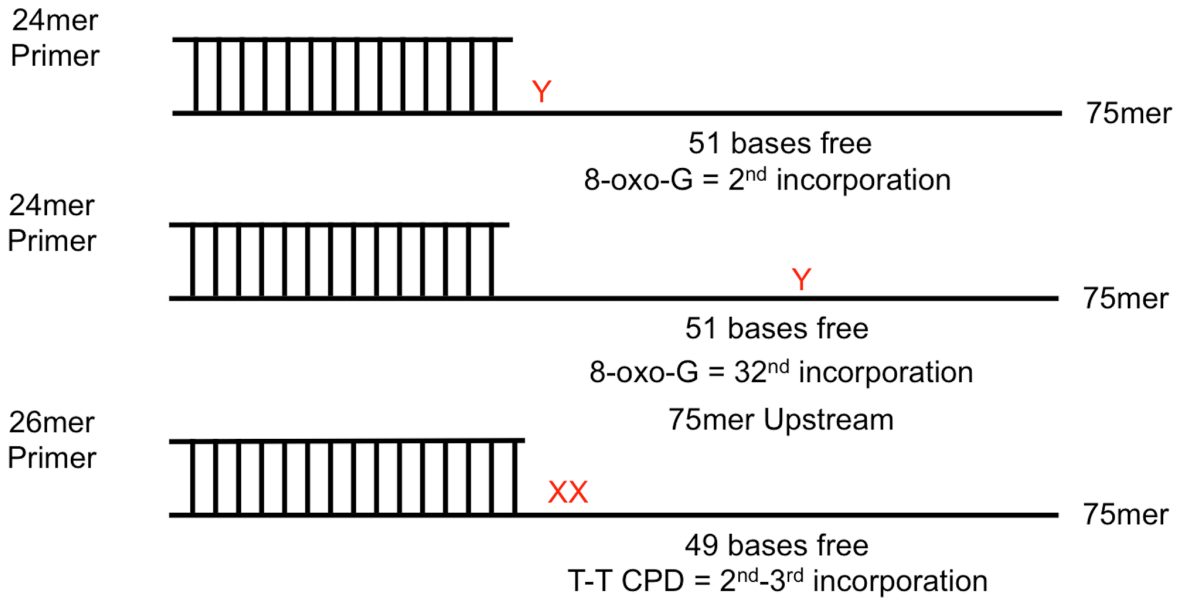


Figure 2.2 – Substrate Diagrams. Schematic diagram of the DNA substrates used in RPA binding and DNA fidelity assays. Y indicates the position of normal guanine or 8-oxoG. 8-oxoG substrates contain a 24-mer primer with the lesion occurring at the 2nd or 32nd incorporation. XX indicates the position of T-T CPD. T-T CPD template contains a 26-mer primer that causes the 3'-T to be the 2nd incorporation after synthesis begins. All templates are 75 bases in length.

Figure 2.3 – Wild type, full-length pol η damaged error rates. Data from the lesion bypass fidelity assay as described in Materials and Methods. Substrate to polymerase ratios were 1:1. Error rates were calculated from sequencing between 23 and 47 dark blue plaques per experiments, with each plaque representing a unique bypass event. White bars indicate the absence of RPA. Dark Grey bars indicate the addition of 5-fold RPA over substrate. Hatched bars indicate the addition of 5-fold heat inactivated RPA over substrate. Heat inactivated RPA created by heating an aliquot of RPA to 85°C for 15 minutes. **A.** 8-oxoG to T (misincorporation of dATP opposite 8-oxoG) error rates(10^{-4}) on 75mer and 75mer upstream templates containing 8-oxoG. Values represent average of 2 independent experiments. **B.** 3'-T to C and 3'-T to A errors (misincorporation of dGTP and dTTP, respectively) on 75mer templates containing T-T CPD.

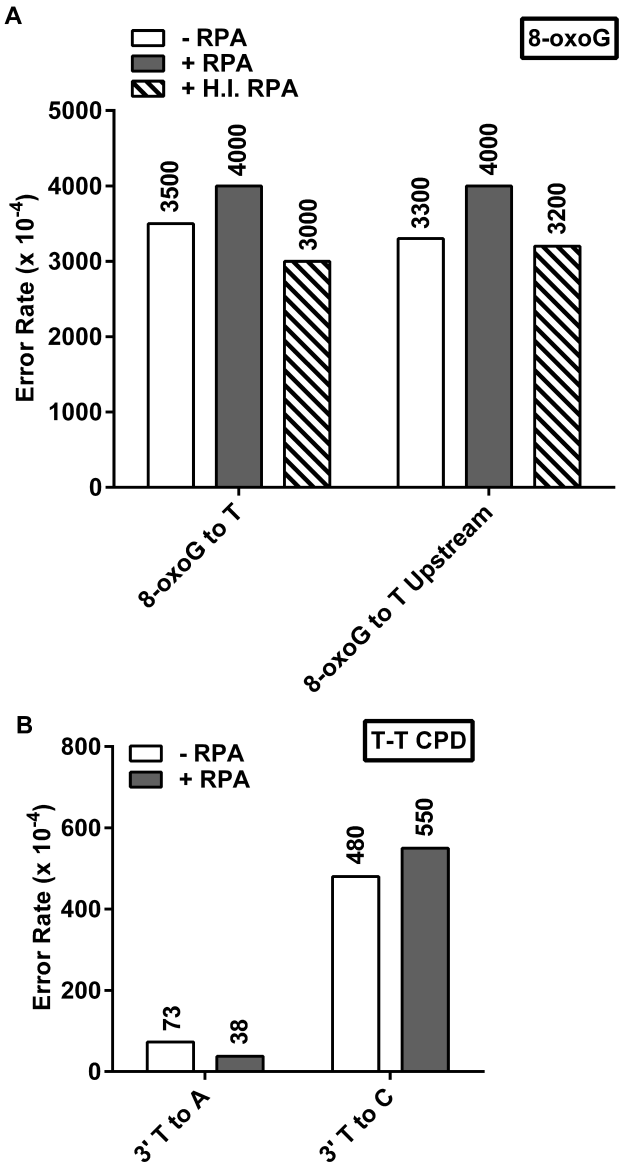

A
- RPA
+ RPA
+ H.I. RPA
8-oxoG
Error Rate (x 10^-4)
5000
4000
3000
2000
1000
0
3500
4000
3000
3300
4000
3200
8-oxoG to T
8-oxoG to T Upstream
B
- RPA
+ RPA
T-T CPD
Error Rate (x 10^-4)
800
600
400
200
0
73
38
480
550
3' T to A
3' T to C

Figure 2.4 – Wild type, Q38A, and Y52E damaged error rates. Data from the lesion bypass fidelity assay as described in Materials and Methods. Error rates were calculated from sequencing between 23 and 47 dark blue plaques per experiment, with each plaque representing a unique bypass event. Substrate to polymerase ratios were 1:1. White bars indicate the absence of RPA. Dark Grey bars indicate the addition of 5-fold RPA over substrate. **A.** 8-oxoG to T error rates (10^{-4}) (misincorporation of dATP opposite 8-oxoG) when copying 75mer templates. **B.** 3'-T to A and **C.** 3'-T to C error rates (10^{-4}) (misincorporation of dGTP and dTTP, respectively) when copying 75mer templates containing T-T CPD.

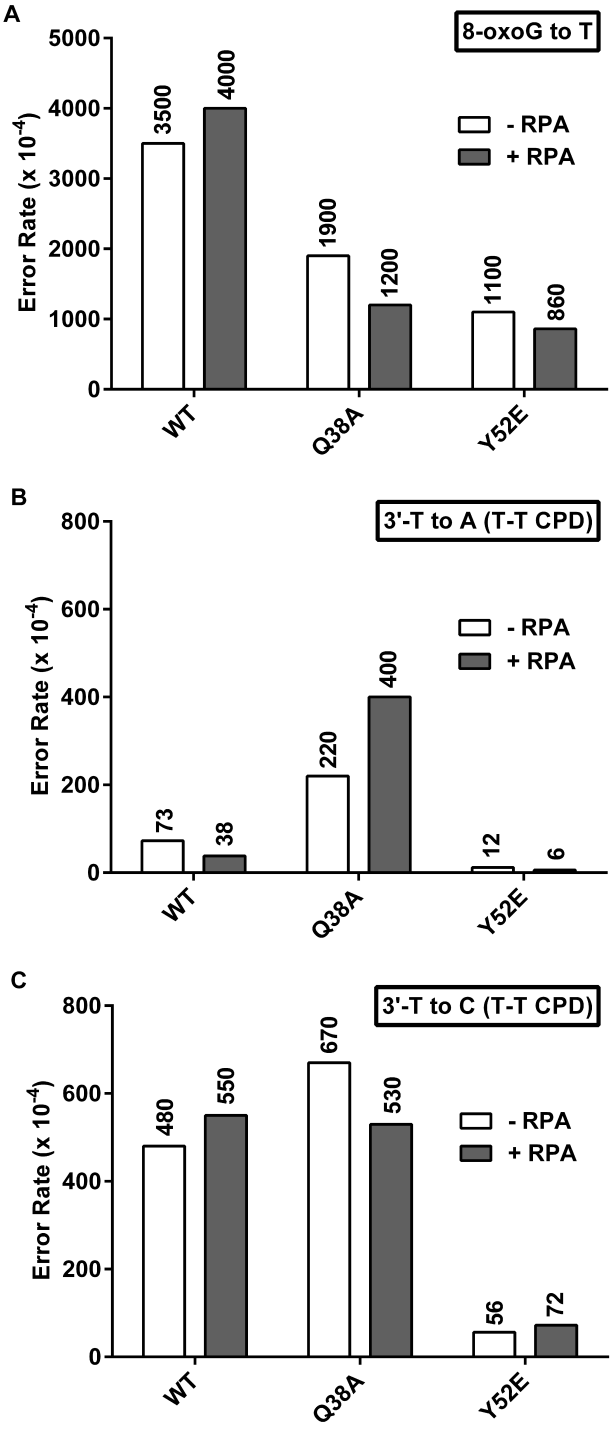
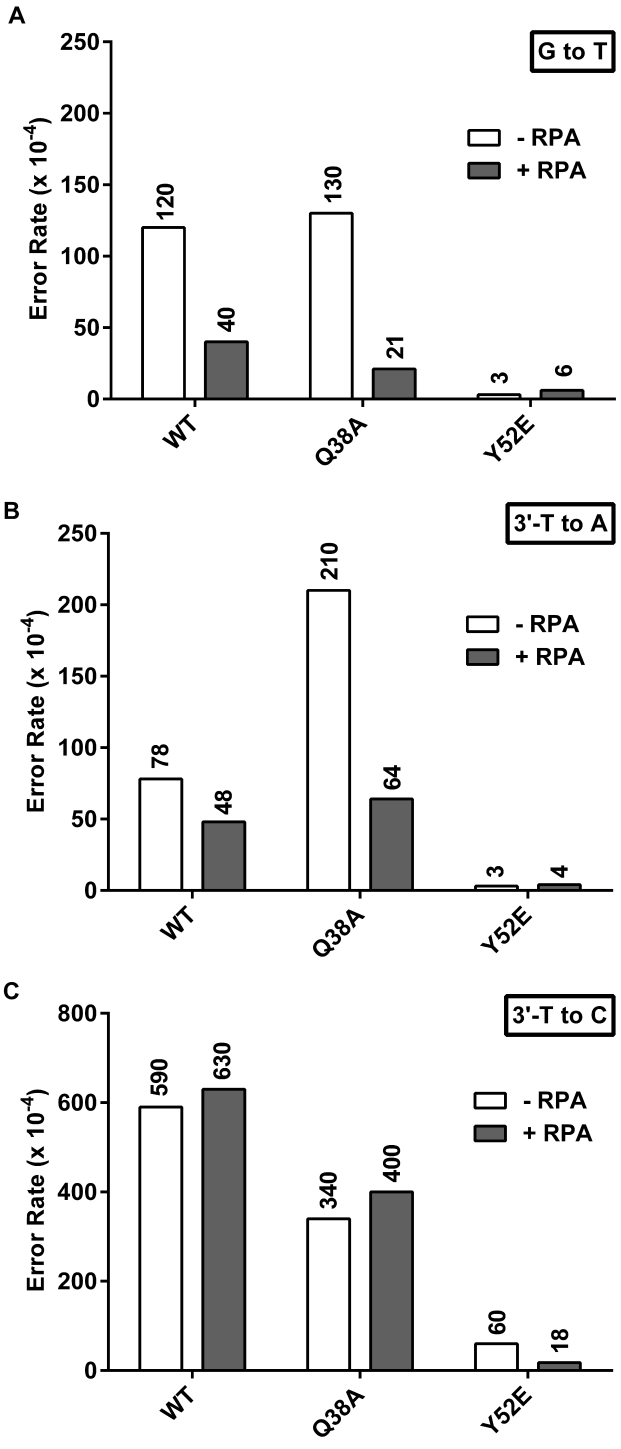
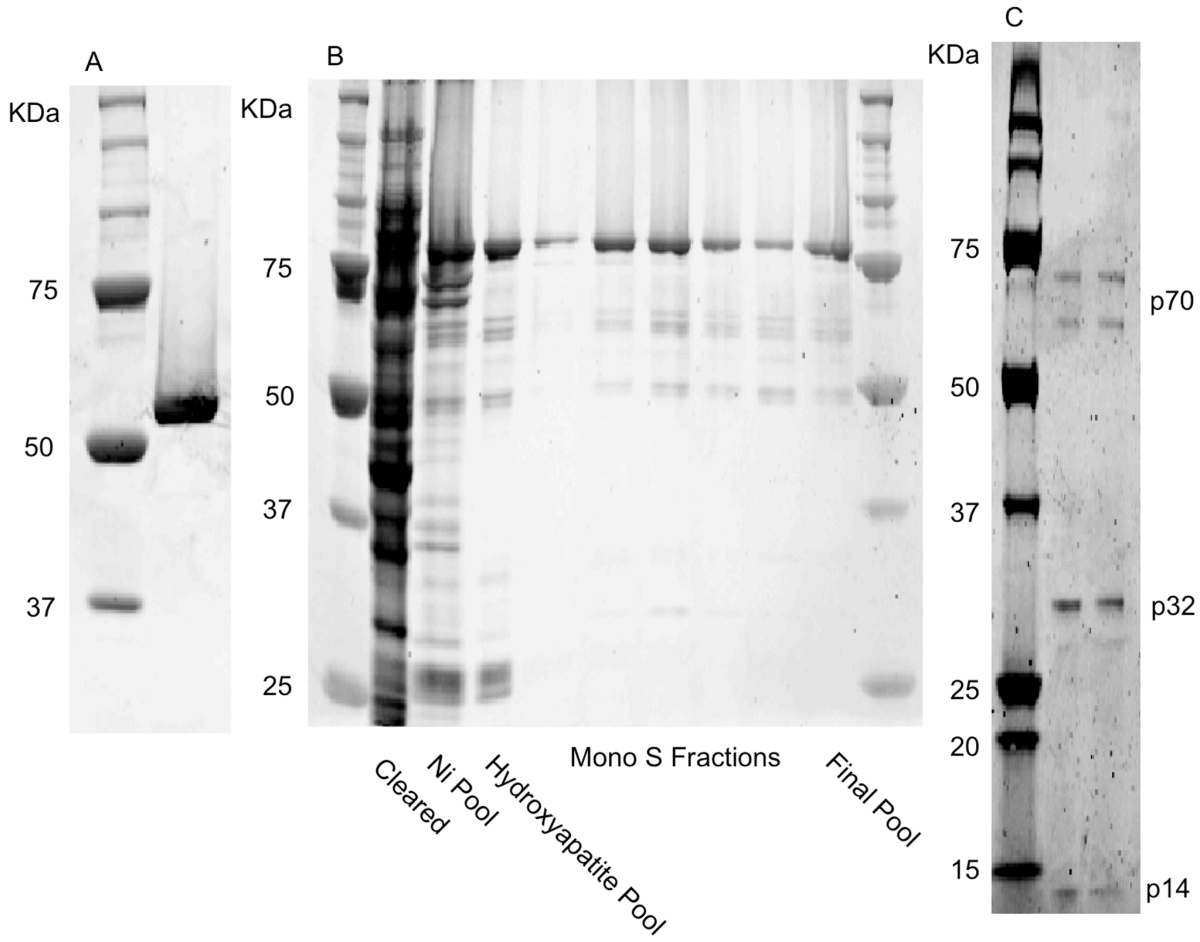

A
8-oxoG to T
- RPA
+ RPA
Error Rate (x 10-4)
5000
4000
3000
2000
1000
0
3500
4000
1900
1200
1100
860
WT
Q38A
Y52E
B
3'-T to A (T-T CPD)
- RPA
+ RPA
Error Rate (x 10-4)
800
600
400
200
0
73
38
220
400
12
6
WT
Q38A
Y52E
C
3'-T to C (T-T CPD)
- RPA
+ RPA
Error Rate (x 10-4)
800
600
400
200
0
480
550
670
530
56
72
WT
Q38A
Y52E

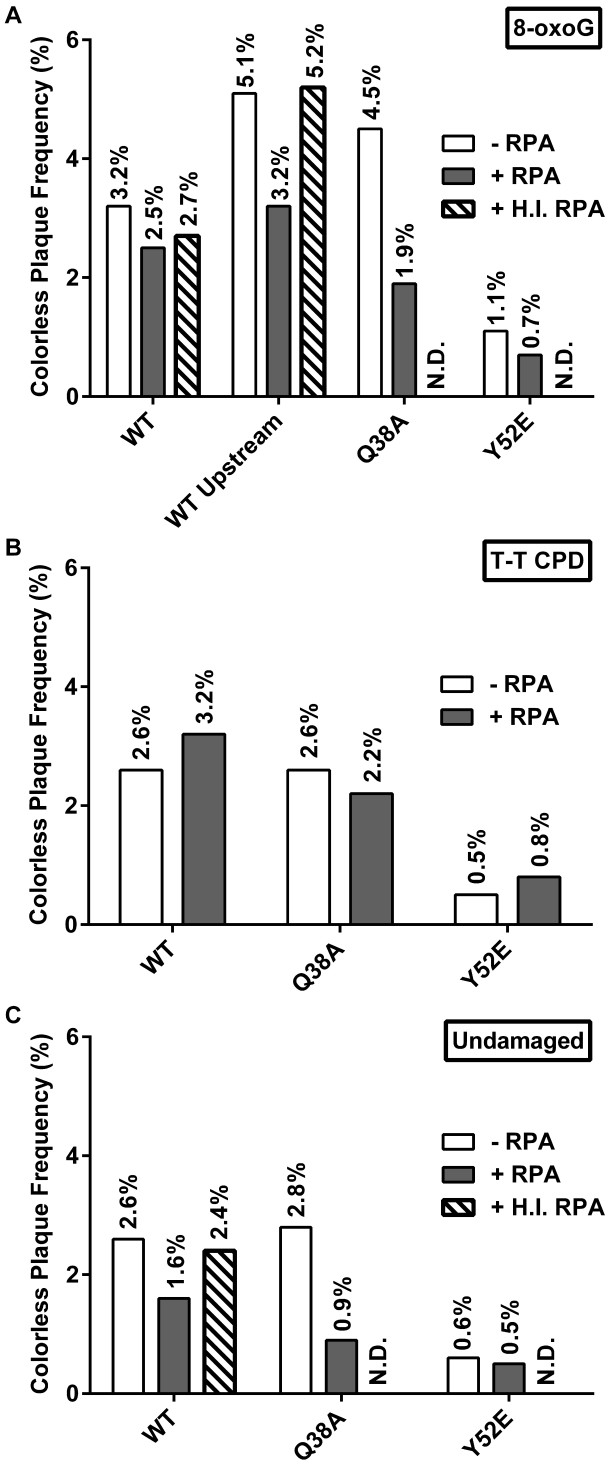
Figure 2.5 – Wild Type, Q38A, and Y52E undamaged error rates. All data from the lesion bypass fidelity assay as described in Materials and Methods. Error rates were calculated from sequencing between 22 and 48 dark blue plaques per experiment, with each plaque representing a unique bypass event. Substrate to polymerase ratios were 1:1. White bars indicate the absence of RPA. Dark Grey bars indicate the addition of 5-fold RPA over substrate. **A.** G to T error rates (10^{-4}) (misincorporation of dATP opposite G) when copying 75mer templates. **B.** 3'-T to A and **C.** 3'-T to C error rates (10^{-4}) (misincorporation of dGTP and dTTP, respectively) when copying 75mer undamaged templates.


A
G to T
- RPA
+ RPA
Error Rate (x 10-4)
250
200
150
100
50
0
120
40
130
21
3
6
WT
Q38A
Y52E
B
3'-T to A
- RPA
+ RPA
Error Rate (x 10-4)
250
200
150
100
50
0
78
48
210
64
3
4
WT
Q38A
Y52E
C
3'-T to C
- RPA
+ RPA
Error Rate (x 10-4)
800
600
400
200
0
590
630
340
400
60
18
WT
Q38A
Y52E

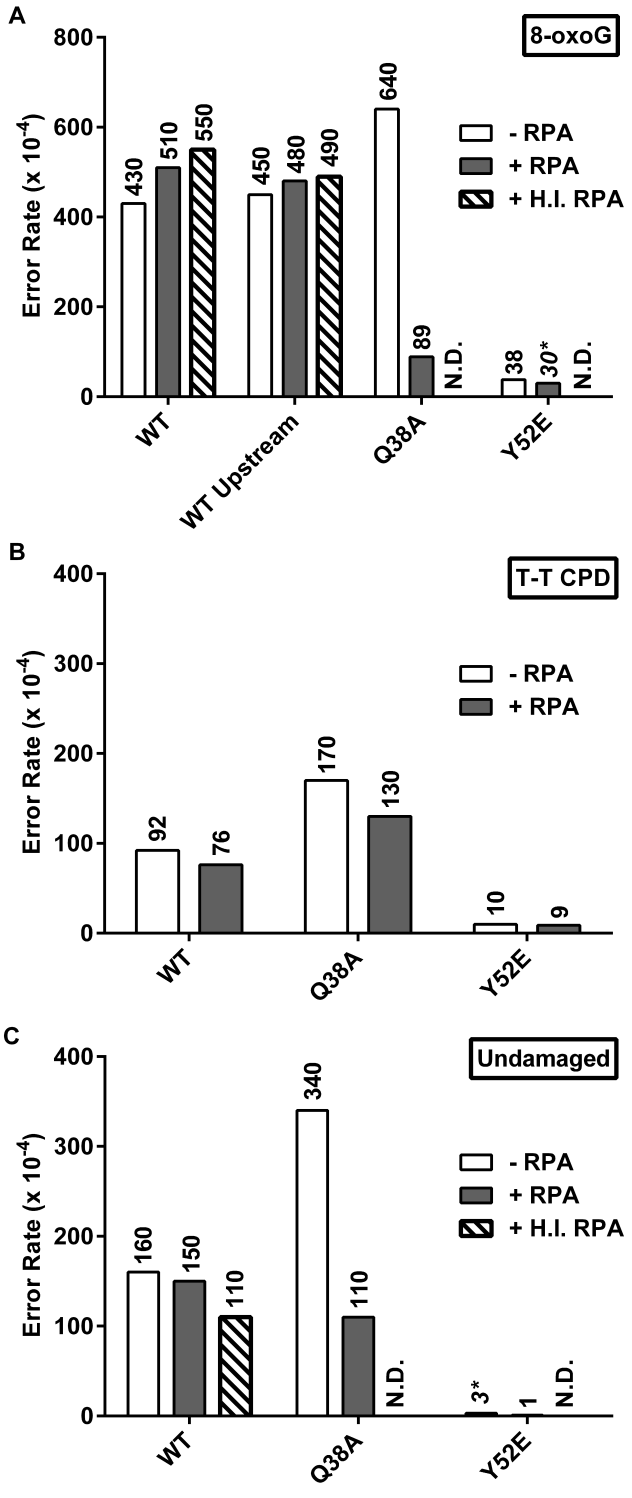
Supplementary Figure 2.1 – Sypro-Red stained protein gels of purified proteins. Sizes are indicated for relevant ladder bands. All gels imaged with a Storm 865 imager (GE Life Sciences). **A.** C-terminal 6x-His tagged truncated pol η (1-511 aa) separated by 10% SDS-PAGE Gel. **B.** Purification overview of N-Terminal 6x-His tagged pol η produced in *E. coli*. Fractions represent samples taken after centrifugation, pool after the nickel column, pool after the hydroxyapatite column, individual fractions from the mono s column, and the pooled result as stored. Samples separated by 10% SDS-PAGE. **C.** Final pools of RPA separated by 4-20% SDS-PAGE. 3 subunits of RPA (p70, p32, p14) marked as indicated.


A
KDa
75
50
37
B
KDa
75
50
37
25
Cleared
Ni Pool
Hydroxyapatite Pool
Mono S Fractions
Final Pool
C
KDa
75
50
37
25
20
15
p70
p32
p14

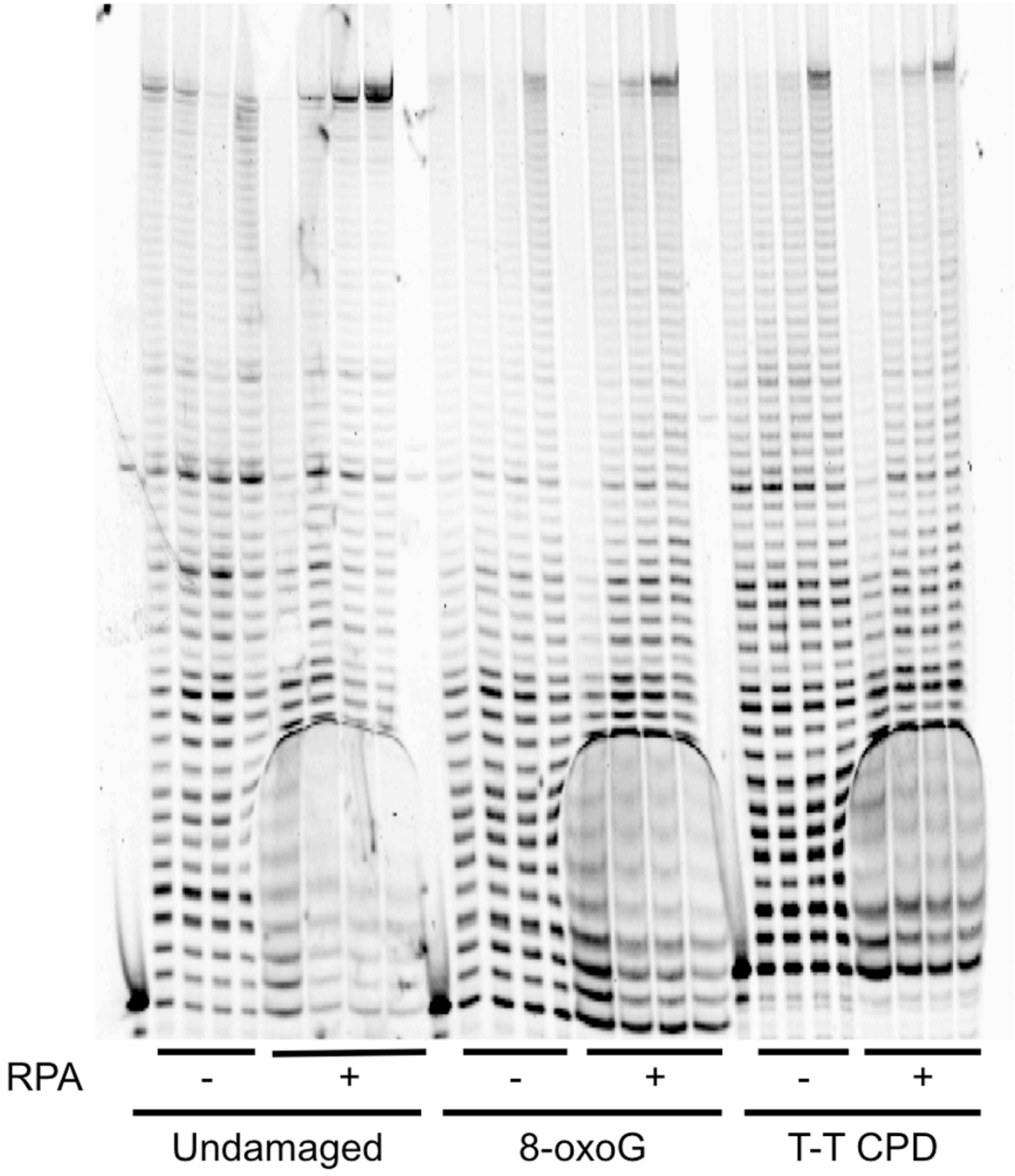
Supplementary Figure 2.2 – Colorless plaque frequency. All data from the lesion bypass fidelity assay as described in Material and Methods. Percentage of colorless plaques of the whole number of plaques counted. Total plaques counted between 1,000 and 20,000 per condition. White bars indicate the absence of RPA. Dark Grey bars indicate the addition of 5-fold RPA over substrate. Hatched bars indicate the addition of 5-fold heat inactivated RPA over substrate.


A
8-oxoG
Colorless Plaque Frequency (%)
6
4
2
0
3.2%
2.5%
2.7%
5.1%
3.2%
5.2%
4.5%
1.9%
N.D.
1.1%
0.7%
N.D.
- RPA
+ RPA
+ H.I. RPA
WT
WT Upstream
Q38A
Y52E
B
T-T CPD
Colorless Plaque Frequency (%)
6
4
2
0
2.6%
3.2%
2.6%
2.2%
0.5%
0.8%
- RPA
+ RPA
WT
Q38A
Y52E
C
Undamaged
Colorless Plaque Frequency (%)
6
4
2
0
2.6%
1.6%
2.4%
2.8%
0.9%
N.D.
0.6%
0.5%
N.D.
- RPA
+ RPA
+ H.I. RPA
WT
Q38A
Y52E

Supplementary Figure 2.3 – Complex error rate (10^{-4}). All data from the lesion bypass fidelity assay as described in Material and Methods. Complex errors defined as any changes at multiple bases with less than 2 correct insertions in between. Examples include 2 sequential base substitutions (majority), Base substitution followed by correct insertion followed by another base substitution, or multiple base deletions. Numbers with an * indicated no mutant plaques sequenced contained a complex change and the error rate is at most the indicated rate (calculated based on the value if there were 1 observed change). White bars indicate the absence of RPA. Dark Grey bars indicate the addition of 5-fold RPA over substrate. Hatched bars indicate the addition of 5-fold heat inactivated RPA over substrate.


A
8-oxoG
Error Rate (x 10-4)
800
600
400
200
0
430
510
550
450
480
490
640
89
N.D.
38
30*
N.D.
- RPA
+ RPA
+ H.I. RPA
WT
WT Upstream
Q38A
Y52E
B
T-T CPD
Error Rate (x 10-4)
400
300
200
100
0
92
76
170
130
10
9
- RPA
+ RPA
WT
Q38A
Y52E
C
Undamaged
Error Rate (x 10-4)
400
300
200
100
0
160
150
110
340
110
N.D.
3*
1
N.D.
- RPA
+ RPA
+ H.I. RPA
WT
Q38A
Y52E

Supplementary Figure 2.4 – Pol η primer extension in the presence or absence of RPA. For each substrate and condition, time points were taken at 3, 6, 9 and 30 minutes. Substrate to polymerase ratio was 10:1. RPA to substrate ratio was 5:1. Samples were incubated without polymerase for 30 minutes at 37° C and synthesis was initiated with the addition of pol η. Samples were separated by 10% Denaturing PAGE and imaged on a Storm 865 imager (GE Life Sciences).


RPA
-
+
-
+
-
+
Undamaged
8-oxoG
T-T CPD

GENERAL DISCUSSION

External exposures form the basis for the study of toxicology. However, a threat is also posed by the exposures humans undergo daily from endogenous sources such as reactive oxygen species, created as a byproduct of metabolism. Both external and endogenous exposures, as well as spontaneous chemical reactions can damage DNA.[28] These exposures can damage other cellular components as well, but DNA is the only biological molecule that must be repaired instead of being resynthesized.[4]. Accurate replication is necessary to prevent possible deleterious consequences of mutations to the DNA sequence. Changes in DNA sequence can cause cancer, aging, and other diseases.[18] Despite mechanisms to prevent replication when damage has occurred, occasions arise where copying occurs in the presence of damage. Replication must continue even when the replicative polymerases pause synthesis and the replication fork stalls forward progression, as excision repair on regions of unwound, single stranded DNA would cause double strand breaks,[34] a situation that is nearly guaranteed to be mutagenic.[186] In order to decrease the likelihood of this possibility, cells possess a separate pathway of damage tolerance that is able to copy past these lesions, alleviating replication stress and allowing the damage to be repaired later by other processes in fully duplexed DNA. This damage

tolerance pathway is termed translesion synthesis, and involves multiple polymerases of the Y-family, and also the B-family polymerase, pol ζ. The general features shared by these polymerases are more open active sites (excluding pol ζ, for which the active site structure is unknown), lack of intrinsic exonuclease proofreading activity, and reduced fidelity when copying undamaged DNA compared to replicative polymerases.[36]

This process of translesion synthesis was initially described as "error-free", as loss of function of the first of these TLS polymerases discovered, pol η, leads to increased mutations and the cancer prone syndrome XPV.[66,80,89] As absence of this proteins leads to carcinogenesis, pol η was the first polymerase to be described as a tumor suppressor.[101] Despite this characterization, the in vitro bypass fidelity of human pol η is moderate to low, with the polymerase making mistakes 3-5% of the time when copying T-T CPD and 35-50% of the time when copying 8-oxoG.[91,97,98,125,145,167] How these in vitro error rates relate to a process that prevents mutagenesis in cells is currently unknown. We have put forth data in this dissertation that furthers our understanding of TLS by pol η at the molecular level. In order to generate this data, we have made use of truncated (1-511 amino acids) and full-length (1-713 amino acids) pol η, purified by sequential affinity and ion-

exchange chromatography. We have produced the wild type sequence of pol η, as well as a set of amino acid substitution mutants encompassing single amino changes in and surrounding the active site.[145] A second, companion study undertaken by our colleague Renee Beardslee examined effects of amino acid substitutions in the β strand (aa 316-324) that functions to align template strand DNA in the catalytic core of pol η.[167] We have also purified the heterotrimeric replication accessory protein A (RPA) and tested its ability to modulate the fidelity of pol η.

Effects of Single Amino Acid Substitution Mutants

The primary amino acid sequence of Y-family polymerases leads to the same right handed structure as replicative polymerases despite little sequence homology between the families.[105] Y-family polymerases generally feature more open active sites and relaxed nucleotide selectivity.[3] The crystal structure of pol η demonstrates its ability to fit two damaged nucleotides into the active site, accommodating damage like the chemically linked T-T CPD and cisplatin crosslink[85,109]. This openness seems well suited to accommodate the smaller, oxidative lesion 8-oxoG. It also provides the means necessary to bypass these lesions more efficiently than the corresponding undamaged DNA.[85,97,98,109] The tertiary protein structure positions individual amino acids in the active site in order to provide the molecular

machinery that accomplishes the task of nucleotide polymerization, template DNA damage bypass, and extension beyond the lesion.[85] Using the crystal structure, previous random mutation screens, and known SNPs in the genome encoding amino acid substitutions to the protein,[85,106,107,142-144] this work sought to determine the individual contribution of amino acids present in and around the active site to the bypass of DNA lesions.[145] In the set of amino acid changes considered in this study and its companion examining a β strand important for positioning of template DNA, only two mutants, R55A, and G320P could reduce polymerase activity to the point of being unusable in our biochemical assays.[145,167] The change of glycine 320 to proline in a conserved beta sheet disrupts the molecular splint region of the polymerase that extensively interacts with template DNA.[85,167] The substitution of arginine 55 with alanine disrupts contact with the phosphate group of the incoming nucleotide.[85,145] These two contacts are apparently critical for activity, as their disruption greatly reduces polymerization. A third change (F17L) was unable to be purified, as changing this amino acid resulted in insufficient proteins yields, despite presence of overexpressed bands of the correct size when examining crude protein extracts by SDS-PAGE. The fact that only 3 of these changes out of 17 mutants studied produced nonfunctional or misfolded protein argues that the amino acid

sequence that form pol η's active site is robust, with only a limited number of conserved amino acids playing a more important role in the function of the polymerase. This is also supported by random mutation experiments that used screening for mutants that complement UV-C sensitivity of pol η deficient yeast. That study demonstrated that single mutations at 20 of 22 active site residues were tolerated, and even mutants with up to 4 different substitutions in this region retain activity comparable to wild type.[106]

Despite plasticity within this region of the protein, certain substitutions are more detrimental and can cause XPV. Of the sequence changes examined by sequencing the XPV gene mutations underlying the disease, approximately 70% of mutations would result in truncations of pol η protein. The remaining 30% contain missense mutations that alter the amino acid sequence of the protein. Some of these mutants cause misfolded proteins, leading to inactivation, while others lead to correctly folded protein that has some alteration in function. The type of mutation is correlated with the severity of the disease phenotype.[86] R361S is one such mutant resulting in a full-length protein that can be purified, and its properties are currently under investigation (R. Beardslee, unpublished data). It will be interesting to purify missense mutants with correctly folded protein resulting in XPV. Some of these

changes could result in altered fidelity or bypass ability, or generally less active polymerases. Characterizing these kind of missense mutations will give us a better idea as to what kinds of changes to the protein are tolerated without causing accelerated disease. This information will complement the study of non-disease causing SNPs, three of which we have shown here to be biochemically similar to wild type for fidelity and bypass ability. Evidence exists that in XPV cells with inactivating truncations, mutations made by low fidelity bypass of CPD by pol ι could lead to the increased mutations seen in XPV.[187] Taken together, this data could provide an idea as to the degree which misinsertions by pol η and other TLS polymerases could raise the mutation rate above background.

To further the goal of determining the consequences of TLS in cells, exogenously expressing pol η mutants in XPV cells could provide insight into the effects of altering synthesis activity. We have characterized the M14V mutant, coded for by a SNP, which exhibits increased bypass efficiency for 8-oxoG compared to wild type. M14V shows the same fidelity profile as wild type, and this is not surprising, as M14V changes contacts to the active site magnesium required for catalysis, not positioning of the nucleotide or template (as determined from PDB file 3MR2[85,166]). Expressing this mutant in a cell could decrease markers of replication

stress upon DNA damage such as phosphorylation of RPA or CHK1.[35] Cells expressing the inactive mutants R55A and G320P could actually increase mutations in XPV cells possessing truncation mutations in their XPV gene. The intact C-terminal region in these pol η mutants could enhance Rad18 mediated monoubiquitination of PCNA. Similar catalytically inactive pol η mutants have been theorized to increase recruitment of other Y-family polymerases due to this monoubiquitination, causing an increase in more error prone synthesis.[58]

We have identified 2 mutants that decreased bypass efficiency for both T-T CPD and 8-oxoG (Q38A and Y52E) and one that has reduced efficiency opposite T-T CPD only (R61A). These mutant polymerases also show effects on fidelity that vary by which change is made. Q38A increases 3'-T to A changes opposite T-T CPD, while decreasing 8-oxoG to T changes. R61A reduces both 3'-T to C and 3'-T to A changes opposite T-T CPD. Y52E reduces errors for all 3 changes considered, 8-oxoG to T changes and 3'-T to C and 3'-T to A changes opposite T-T CPD. Exogenously expressing these mutants with altered fidelity and bypass in XPV cells could provide insights as to molecular properties necessary to rescue the cell phenotype. XPV cells display a phenotype of altered DNA synthesis and increased mutation.[63,73,74] Expressing mutants with increased fidelity but lowered bypass

ability could provide clues as to the level of bypass needed to cope with DNA damage and prevent this phenotype. It is also possible that decreased bypass efficiency could provide opportunities for low fidelity bypass by pol ι. Expression of pol η with reduced bypass and increased fidelity could actually increase mutation rates in these cells due to increased synthesis opportunities for pol ι or other TLS polymerases. The altered bypass efficiencies are still 2-3 fold greater than yeast replicative polymerases,[98,145,150] and this bypass efficiency may be sufficient to rescue the XPV cell phenotype and prevent synthesis by other, more error prone polymerases. If bypass ability is adequate but reduced, these cells would be expected to show decreases in mutation rate compared to cells expressing wild type pol η, but could also display increase markers of replication stress such as phosphorylation of RPA or CHK1.[35] Selection of an XPV cell line for this experiment expressing a catalytically inactive variant of pol η instead of a truncated protein would be useful to control for perturbations in ubiquitin signaling caused by alteration of the C-terminus of the protein.[58] These results could provide further insight into the mechanism and result of bypass by pol η.

The molecular mechanism by which human pol η copies past 8-oxoG is as yet unknown. Crystal structures of the polymerase in ternary complex with 8-oxoG and

the correct C and incorrect A in the active site could provide insight into the high error rates seen when bypassing 8-oxoG. Human, mouse, and yeast pol η are the only Y-family polymerases whose fidelity of 8-oxoG bypass has been published using an assay that involve insertion across from the lesion as well as extension the resulting damaged primer termini in the presence of all 4 dNTPs, and all 3 homologs of pol η copy 8-oxoG better than they copy the corresponding undamaged DNA sequence.[98] Two other members of the Y-family, human pols ι and κ, have been examined using single nucleotide experiments to measure insertion opposite 8-oxoG. Insertion efficiency of pol ι is reduced when inserting C opposite 8-oxoG compared to undamaged G, and pol ι prefers to insert the correct dCTP.[188] Pol κ insertion efficiency is reduced when examining both the correct C and incorrect A insertion events opposite 8-oxoG, but pol κ preferentially incorporates the incorrect dATP.[189] Yeast pol η, human pol ι, and human pol κ, have been crystalized with the correct dCTP paired with 8-oxoG in the active site.[162,190,191] Pol ι and pol κ have both crystallized with the incorrect dATP paired with 8-oxoG as well.[190-192] Yeast and human pol η have been crystallized with a T-T CPD in the active site.[85,158] Each of these crystal structures provides insight into the molecular mechanisms of bypass of these lesions by each polymerase, but as of yet no crystal structure has been

published to explain the low fidelity and high efficiency bypass of 8-oxoG by human pol η.

Effects of Accessory Proteins

One of the factors proposed to modulate pol η fidelity is interaction with replication accessory proteins present at the replication fork. Examining the history of the replication fidelity field shows limited examples of fidelity differences when examining the effect of replication accessory proteins, as discussed in chapter 2. Investigations with yeast pol η along with yeast RFC, PCNA and RPA have shown little to no effect of fidelity when bypassing a T-T CPD or 8-oxoG using the quantitative assays also used in this work.[98,127] However, yeast and human pol η show different fidelity with respect to copying 8-oxoG,[98] and one report has shown an enhancement of insertion fidelity opposite 8-oxoG by human pol η with the addition of human RPA and PCNA. This report utilized single nucleotide insertion experiments, lacked RFC to load PCNA onto primer termini, lacked any means of preventing PCNA diffusion off of substrates and did not report results of RPA alone.[125] We sought to clarify this result using our color based screening system that requires insertion across from the damage and extension of the damaged primer terminus with all 4 dNTPs in competition. Addition of RPA to the bypass reaction

showed no differences in the fidelity of wild type pol η when copying past either T-T CPD or 8-oxoG. To provide confirmation, we tested two mutants, Q38A and Y52E, that have unique fidelity signatures compared to wild type. We first generated these same mutants in the full-length protein, and verified the signature was still present. When adding RPA, these mutants displayed fidelities for lesion bypass that did not differ from conditions when RPA was absent. The only effects seen when RPA was present were instances of increased fidelity on undamaged DNA in certain changes for wild type, Q38A, and Y52E.

The advantage of using an assay that requires insertion across from and extension past the lesion to measure fidelity is we can gain information about sequence changes other than base substitutions. Translational read through of the stop codon in the LacZ sequence results in light blue plaques where the sequence has not been changed, or the template strand of the gapped phage DNA is expressed. Indel events creating frameshifts in the gene display a true colorless plaque phenotype.[95] Calculating the portion of colorless plaques as a percentage of the whole gives insight into the rate of these indel mutations the same way dark blue plaque frequency informs us about base substitution mutations.[122] Another advantage of this assay is the ability to detect complex mutations that require

multiple misinsertion events in close proximity, such as tandem base substitutions, which pol η makes at higher rates than other polymerases.[118,119] When examining these types of changes in the Q38A mutant, some effects on fidelity were observed upon addition of RPA. The colorless plaque frequency and frequency of complex errors when copying both 8-oxoG and undamaged DNA were reduced between 3-7 fold when RPA was present. These effects on colorless plaque frequency and complex error rates were not observed when copying T-T CPD in this mutant. We speculate this is due to the additional rigidity provided by the cyclobutane ring present in the dimer. Presumably, this effect cannot be seen in the wild type or Y52E mutant due to intact protein-DNA template contacts. The Q38A change abolishes this conserved contact with the template base in the active site.[85] The change in DNA conformation imparted by the binding of RPA for these two changes in this mutant possibly cannot be produced when the rigid T-T CPD is in the active site. This data indicates, when copying damaged DNA, the structure of the polymerase active site is the main determinant of fidelity.

We saw no difference when adding RPA to the bypass reaction of pol η copying T-T CPD or 8-oxoG, suggesting the main determinant of pol η fidelity opposite these lesions is the active site of the polymerase itself. However, this does

not take into account interactions with other polymerases and proteins present at the replication fork. Pol η is known to interact with both unmodified and monoubiquitinated PCNA, as well as Rev1 and pol ι.[37,52,115,193-195] Rev1 is also targeted to regions of single stranded DNA and primer termini, possibly playing an additional role in targeting pol η and other TLS polymerases to regions of stalled replication.[196] Phosphorylation and ubiquitination also play a role in regulation of pol η function, and these sites of post-translational modification are located in the C-terminus of the protein.[40,45,197] The regions of pol η known to be involved in protein:protein interactions also occur in the C-terminal region of the protein (Figure I.4). It is possible that these interactions in the C-terminus could affect functions in the active site located in the N-terminus of the protein, but no crystal structure of the complete pol η protein is yet available to inform our understanding of the contacts between the catalytic core and the interacting domains located in the C-terminus. It will be interesting to see if any of these modifications or interactions modulates the fidelity or efficiency of pol η bypass.

Other Factors Possibly Influencing TLS Fidelity

Another possible mechanism to prevent TLS fidelity from creating mutagenesis is involvement of mismatch repair or base excision repair. Several

elements of the mismatch repair system have been implicated in TLS processes in response to UV and oxidative damage.[140,141] Physical interactions have been reported for both the MutSα and MutLα complexes with pol η.[139,198] Pol η bypass of 8-oxoG has been shown to be independent of MutSα in *S. cerevisiae*, and perhaps this is due to the higher fidelity of yeast pol η opposite this lesion.[199] MSH2-MSH6 (MutSα) has been shown to play a role in the removal of mispaired 8-oxoG:A in *S. cerevisiae*.[200] No MutY homolog (MYH) has been discovered in budding yeast, and this base excision repair glycosylase processes 8-oxoG:A mispairs in fission yeast and humans.[201,202] These base excision and mismatch repair factors all serve to cope with 8-oxoG when paired with correct or incorrect nucleotides. Despite this difference in processing of mispaired 8-oxoG lesions between species, monoubiquitination of PCNA is observed upon treatment with agents that cause oxidative damage such as hydrogen peroxide and menadione in both budding yeast and humans, implying a role for TLS in tolerance of oxidative damage.[140,203] It is a possible scenario that TLS plays a role in the initial response to oxidative DNA damage, with fidelity corresponding to the species TLS is occurring in (2-3% in yeast, 50% in humans). Base excision repair could then play a role in resolving 8-oxoG:A mispairs and 8-oxoG:C. Correctly paired 8-oxoG:C would be processed to excise 8-oxoG by OGG1 glycosylase, leading

to repair of the damage after resynthesis of the resulting gap by pol β.[204] 8-oxoG:A mispairs would require two or more rounds of base excision repair. First, the incorrect A would have to be removed from the 8-oxoG:A mispair by MYH.[205] Resynthesis would be a more complicated affair, as 8-oxoG would still be present. Pol η could perform this resynthesis with lowered fidelity. Another possibility is resynthesis opposite 8-oxoG by pol λ, which has been implicated in these multiple rounds of oxidative damage dependent base excision repair. These rounds of MYH dependent excision and resynthesis would continue until the correct C was inserted opposite 8-oxoG, leading to excision of 8-oxoG by OGG1 and leaving the correct C as a template for resynthesis by either pol β or λ.[205] These events could serve to correct errors made by pol η or another polymerase that copies 8-oxoG with reduced fidelity. Further exploration of the interaction between TLS, mismatch repair, and base excision repair is needed to determine how these systems could cooperate to reduce the mutagenic potential of TLS.

Reflections on This Work

If we were to begin this work in 2014 instead of 2007, numerous advantages would be available that could have allowed a refined approach to this project. As a newly formed lab, we have had to obtain equipment over time, limiting us initially.

As time has passed, our knowledge and capacity to produce data have increased. Acquisition of multiple pieces of equipment throughout 2009 and 2010, most notably an AKTA Purifier, accelerated our research. We have also taken the opportunity to make use of new techniques available to us, the most notable of which has been the extensive use of fluorescent labels for nucleotide and protein visualization. These techniques have required the procurement of a Storm imager, and access to this instrument has been essential to performing our work.

Other groups have developed new methods of measuring polymerase fidelity using next-generation sequencing equipment since we began our investigations.[206] While the methods utilized in our work are able to provide us with excellent information about the fidelity of our polymerases of interest, these methods are limited to sequence contexts required for the assay. Next-generation sequencing methods are not limited by sequence context, and information as to the effects of surrounding DNA sequence on bypass fidelity by pol η would be interesting to obtain. We attempted to obtain fidelity data utilizing other, more high-throughput, non-sequencing based methods, but technical challenges prevented utilization of those assays.

We have utilized biochemical approaches in order to better understand the molecular mechanisms of synthesis by pol η. If beginning our work in the present day, we would still utilize a similar approach to answer our research questions. The use of purified proteins allows us to examine the properties of the polymerase in an isolated system. Insights into the action of pol η in an in vitro, biochemical context have raised important questions as to the implications of a polymerase that makes errors in synthesis, but is an overall positive for cells. These insights provide complements to experiments performed in cells, and both approaches are valuable. We would have liked expand our experiments into a cell-based system, but were limited by time constraints and the necessity of setting up those experiments from the ground up in our lab.

Refinements in our ability to produce protein could have allowed us to explore more amino acid changes in the active site, including more SNPs present in human populations. The development of a method to produce full-length pol η in *E. coli* has been a boon for our research, as this has allowed us to move away from producing certain proteins in insect cells. Producing proteins through baculovirus infection has provided multiple technical challenges in a new lab, and we are now beginning to overcome these issues. The information provided by the crystal

structure of pol η in complex with DNA has provided invaluable information, and could have provided us context for amino acids in the active site at the beginning of this project. When we began this work, there was limited information on non-synonymous SNPs, and that information has grown immensely. Combining currently available data from the crystal structure as well as recent SNP data could inform us as to which SNPs could possibly alter function, an area of possible future exploration. The ability to express these mutants in a full-length construct in *E. coli* would allow for greater numbers of changes to be explored in a shorter time. Also, information as to SNPs that occur together and exploration of those altered polymerases could provide insights as to possibly altered polymerase function in human populations.

TLS as a Damage Tolerance Process

TLS by Pol η offers a method for tolerance of UV induced CPDs, but can also provide a mechanism for resistance of chemotherapeutic agents such as cisplatin and other platinum containing compounds used in the treatment of cancer.[207,208] The open active site provides a mechanism for bypass of cisplatin crosslinks by pol η.[109] Higher pol η expression also predicts resistance to treatment and shorter survival times in multiple types of cancers treated with cisplatin and related compounds.[209-211]

Understanding the molecular mechanism of pol η and TLS in total could help design adjunct therapies to reduce resistance to already efficacious chemotherapeutics.[36] Indeed, small molecule inhibitors that interact with both monoubiquitinated and unmodified PCNA have been identified and are under investigation as possible adjunct therapies for cancer treatment along with administration with cisplatin.[212,213]

This work emphasizes that TLS is a damage tolerance and not a repair process. This process can be a positive for the cell, as demonstrated when defects in pol η mediating TLS of sunlight associated DNA damage cause the cancer syndrome XPV[66] TLS can also be a negative in the treatment of cancer, as pol η provides a resistance mechanism to the DNA damaging effects of cisplatin.[208] The fact that pol η consistently maintains its fidelity and bypass despite changes to individual amino acids in the active site points to the role of pol η as a polymerase specialized in the efficient bypass of specific lesions. This efficiency comes at the expense of reduced fidelity when copying damaged and undamaged DNA. Despite this reduced fidelity, TLS is an overall positive for the cell. Work must still be done in order to determine how the fidelity of TLS is modified, and more clearly elucidate the regulation of TLS, as misregulation or uncorrected errors in the genome induced by TLS lead to mutagenesis and result in disease.

Genomic instability is a contributor to carcinogenesis, as defects in repair and tolerance contribute to cancer.[214] This could contribute to the mutator phenotype necessary to generate the multiple mutations seen in cancer cells.[18,36] Mutations play a role not only in cancer, but reduced fidelity and genomic instability has been implicated in a host of other diseases and conditions such as Alzheimer's disease, Amytrophic lateral sclerosis, Parkinson's disease, Huntington's disease, Cockayne's syndrome, Trichothiodystrophy, Xeroderma pigmentosum, XPV, and aging.[28,215] Better understanding of the mechanisms of mutagenesis, including the possible contribution of TLS to these mutations, could lead to better understanding of these disease processes, and ultimately improve human health.

REFERENCES

(1) Bebenek, K.; Kunkel, T. A. Functions of DNA Polymerases. *Adv Protein Chem* **2004**, *69*, 137–165.

(2) Watson, J. D.; Crick, F. H. Molecular Structure of Nucleic Acids. *Nature* **1953**, *171*, 737–738.

(3) McCulloch, S. D.; Kunkel, T. A. The Fidelity of DNA Synthesis by Eukaryotic Replicative and Translesion Synthesis Polymerases. *Cell Res* **2008**, *18*, 148–161.

(4) Kornberg, A.; Baker, T. A. DNA Replication; 2nd ed.; W.H. Freeman and Company: New York, NY, 1992.

(5) Kunkel, T. A.; Burgers, P. M. J. Dividing the Workload at a Eukaryotic Replication Fork. *Trends Cell Biol* **2008**, *18*, 521–527.

(6) Pursell, Z. F.; Isoz, I.; Lundström, E.-B.; Johansson, E.; Kunkel, T. A. Yeast DNA Polymerase ε Participates in Leading-Strand DNA Replication. *Science* **2007**, *317*, 127–130.

(7) Miyabe, I.; Kunkel, T. A.; Carr, A. M. The Major Roles of DNA Polymerases Epsilon and Delta at the Eukaryotic Replication Fork Are Evolutionarily Conserved. *PLoS GENETICS* **2011**, *7*, e1002407.

(8) Jin, Y. H.; Obert, R.; Burgers, P. M. J.; Kunkel, T. A.; Resnick, M. A.; Gordenin, D. A. The 3'→5′ Exonuclease of DNA Polymerase δ Can Substitute for the 5′ Flap Endonuclease Rad27/Fen1 in Processing Okazaki Fragments and Preventing Genome Instability. *Proc Natl Acad Sci USA* **2001**, *98*, 5122–5127.

(9) Jin, Y. H.; Ayyagari, R.; Resnick, M. A.; Gordenin, D. A.; Burgers, P. M. J. Okazaki Fragment Maturation in Yeast. *J Biol Chem* **2003**, *278*, 1626–1633.

(10) Garg, P.; Stith, C. M.; Sabouri, N.; Johansson, E.; Burgers, P. M. J. Idling by DNA Polymerase δ Maintains a Ligatable Nick During Lagging-Strand DNA Replication. *Genes Dev* **2004**, *18*, 2764–2773.

(11) Nick-McElhinny, S. A.; Gordenin, D. A.; Stith, C. M.; Burgers, P. M. J.; Kunkel, T. A. Division of Labor at the Eukaryotic Replication Fork. *Mol Cell* **2008**, *30*, 137–144.

(12) Johnson, A.; O'Donnell, M. Cellular DNA Replicases: Components and Dynamics at the Replication Fork. *Annu Rev Biochem* **2005**, *74*, 283–315.

(13) Forsburg, S. L. Eukaryotic MCM Proteins: Beyond Replication Initiation. *Microbiol Mol Biol Rev* **2004**, *68*, 109–131.

(14) Blackwell, L. J.; Borowiec, J. A. Human Replication Protein A Binds Single-Stranded DNA in Two Distinct Complexes. *Mol Cell Bio* **1994**, *14*, 3993–4001.

(15) Waga, S.; Stillman, B. The DNA Replication Fork in Eukaryotic Cells. *Annu Rev Biochem* **1998**, *67*, 721–751.

(16) Leman, A. R.; Noguchi, E. The Replication Fork: Understanding the Eukaryotic Replication Machinery and the Challenges to Genome Duplication. *Genes (Basel)* **2013**, *4*, 1–32.

(17) Loeb, L. A. Mutator Phenotype May Be Required for Multistage Carcinogenesis. *Cancer Res* **1991**, *51*, 3075–3079.

(18) Loeb, L. A.; Monnat, R. J. DNA Polymerases and Human Disease. *Nat Rev Genet* **2008**, *9*, 594–604.

(19) Kunkel, T. A.; Bebenek, K. DNA Replication Fidelity. *Annu Rev Biochem* **2000**, *69*, 497–529.

(20) Kool, E. T. Active Site Tightness and Substrate Fit in DNA Replication. *Annu Rev Biochem* **2002**, *71*, 191–219.

(21) Beard, W. A.; Wilson, S. H. Structural Insights Into the Origins of DNA Polymerase Fidelity. *Structure (Cambridge, MA, U.S.)* **2003**, *11*, 489–496.

(22) Goodman, M. F. Hydrogen Bonding Revisited: Geometric Selection as a Principal Determinant of DNA Replication Fidelity. *Proc Natl Acad Sci USA* **1997**, *94*, 10493–10495.

(23) Johnson, S. J.; Beese, L. S. Structures of Mismatch Replication Errors Observed in a DNA Polymerase. *Cell* **2004**, *116*, 803–816.

(24) Pavlov, Y. I.; Frahm, C.; Nick McElhinny, S. A.; Niimi, A.; Suzuki, M.; Kunkel, T. A. Evidence That Errors Made by DNA Polymerase α Are Corrected by DNA Polymerase δ. *Curr Biol* **2006**, *16*, 202–207.

(25) Nick-McElhinny, S. A.; Pavlov, Y. I.; Kunkel, T. A. Evidence for Extrinsic Exonucleolytic Proofreading. *Cell Cycle* **2006**, *5*, 958.

(26) Li, G.-M. Mechanisms and Functions of DNA Mismatch Repair. *Cell Res* **2008**, *18*, 85–98.

(27) Friedberg, E. C.; Walker, G. C.; Siede, W.; Wood, R. D.; Schultz, R. A.; Ellenberger, T. DNA Repair and Mutagenesis; 2nd ed.; ASM Press: Washington, DC, 2006.

(28) Hoeijmakers, J. H. J. DNA Damage, Aging, and Cancer. *N Engl J Med* **2009**, *361*, 1475–1485.

(29) Pfeifer, G. P.; Besaratinia, A. UV Wavelength-Dependent DNA Damage and Human Non-Melanoma and Melanoma Skin Cancer. *Photochem Photobiol Sci* **2012**, *11*, 90–97.

(30) Siddik, Z. H. Cisplatin: Mode of Cytotoxic Action and Molecular Basis of Resistance. *Oncogene* **2003**, *22*, 7265–7279.

(31) Xue, W.; Warshawsky, D. Metabolic Activation of Polycyclic and Heterocyclic Aromatic Hydrocarbons and DNA Damage: a Review. *Toxicol Appl Pharmacol* **2005**, *206*, 73–93.

(32) Marnett, L. J. Oxyradicals and DNA Damage. *Carcinogenesis* **2000**, *21*, 361–370.

(33) Prindle, M. J.; Loeb, L. A. DNA Polymerase Delta in DNA Replication and Genome Maintenance. *Environ Mol Mutagen* **2012**, *53*, 666–682.

(34) Ulrich, H. D. Timing and Spacing of Ubiquitin-Dependent DNA Damage Bypass. *FEBS Lett* **2011**, *585*, 2861–2867.

(35) Zeman, M. K.; Cimprich, K. A. Causes and Consequences of Replication Stress. *Nat Cell Biol* **2013**, *16*, 2–9.

(36) Sale, J. E. Translesion DNA Synthesis and Mutagenesis in Eukaryotes. *Cold Spring Harb Persp Biol* **2013**, *5*, a012708.

(37) Lehmann, A. R.; Niimi, A.; Ogi, T.; Brown, S.; Sabbioneda, S.; Wing, J. F.; Kannouche, P. L.; Green, C. M. Translesion Synthesis: Y-Family Polymerases and the Polymerase Switch. *DNA Repair* **2007**, *6*, 891–899.

(38) Zou, L.; Elledge, S. J. Sensing DNA Damage Through ATRIP Recognition of RPA-ssDNA Complexes. *Science* **2003**, *300*, 1542–1548.

(39) Sancar, A.; Lindsey-Boltz, L. A.; Ünsal-Kaçmaz, K.; Linn, S. Molecular Mechanisms of Mammalian DNA Repair and the DNA Damage Checkpoints. *Annu Rev Biochem* **2004**, *73*, 39–85.

(40) Gohler, T.; Sabbioneda, S.; Green, C. M.; Lehmann, A. R. ATR-Mediated Phosphorylation of DNA Polymerase η is Needed for Efficient Recovery From UV Damage. *J Cell Biol* **2011**, *192*, 219–227.

(41) Vassin, V. M.; Anantha, R. W.; Sokolova, E.; Kanner, S.; Borowiec, J. A. Human RPA Phosphorylation by ATR Stimulates DNA Synthesis and Prevents ssDNA Accumulation During DNA-Replication Stress. *J Cell Sci* **2009**, *122*, 4070–4080.

(42) Liu, S.; Opiyo, S. O.; Manthey, K.; Glanzer, J. G.; Ashley, A. K.; Amerin, C.; Troksa, K.; Shrivastav, M.; Nickoloff, J. A.; Oakley, G. G. Distinct Roles for DNA-PK, ATM and ATR in RPA Phosphorylation and Checkpoint Activation in Response to Replication Stress. *Nucleic Acids Res* **2012**, *40*, 10780-10794.

(43) Liu, G.; Chen, X. DNA Polymerase η, the Product of the Xeroderma Pigmentosum Variant Gene and a Target of p53, Modulates the DNA Damage Checkpoint and p53 Activation. *Mol Cell Bio* **2006**, *26*, 1398–1413.

(44) Jung, Y.-S.; Qian, Y.; Chen, X. DNA Polymerase Eta is Targeted by Mdm2 for Polyubiquitination and Proteasomal Degradation in Response to Ultraviolet Irradiation. *DNA Repair* **2011**, *11*, 177–184.

(45) Bienko, M.; Green, C. M.; Sabbioneda, S.; Crosetto, N.; Matic, I.; Hibbert, R. G.; Begovic, T.; Niimi, A.; Mann, M.; Lehmann, A. R.; et al. Regulation of Translesion Synthesis DNA Polymerase η by Monoubiquitination. *Mol Cell* **2010**, *37*, 396–407.

(46) Jung, Y.-S.; Hakem, A.; Hakem, R.; Chen, X. Pirh2 E3 Ubiquitin Ligase Monoubiquitinates DNA Polymerase Eta to Suppress Translesion DNA Synthesis. *Mol Cell Bio* **2011**, *31*, 3997–4006.

(47) Harris, S. L.; Levine, A. J. The p53 Pathway: Positive and Negative Feedback Loops. *Oncogene* **2005**, *24*, 2899–2908.

(48) Soria, G.; Speroni, J.; Podhajcer, O. L.; Prives, C.; Gottifredi, V. p21 Differentially Regulates DNA Replication and DNA-Repair-Associated Processes After UV Irradiation. *J Cell Sci* **2008**, *121*, 3271–3282.

(49) Prives, C.; Gottifredi, V. The p21 and PCNA Partnership. *Cell Cycle* **2008**, *7*, 3840–3846.

(50) Soria, G.; Gottifredi, V. PCNA-Coupled p21 Degradation After DNA Damage: the Exception That Confirms the Rule? *DNA Repair* **2010**, *9*, 358–364.

(51) Mansilla, S. F.; Soria, G.; Vallerga, M. B.; Habif, M.; Martínez-López, W.; Prives, C.; Gottifredi, V. UV-Triggered Degradation Facilitates Damaged-DNA Replication and Preserves Genomic Stability. *Nucleic Acids Res* **2013**, *41*, 6942–6951.

(52) Kannouche, P. L.; Wing, J.; Lehmann, A. R. Interaction of Human DNA Polymerase η with Monoubiquitinated PCNA. *Mol Cell* **2004**, *14*, 491–500.

(53) Davies, A. A.; Huttner, D.; Daigaku, Y.; Chen, S.; Ulrich, H. D. Activation of Ubiquitin-Dependent DNA Damage Bypass is Mediated by Replication Protein A. *Mol Cell* **2008**, *29*, 625–636.

(54) Huttner, D.; Ulrich, H. D. Cooperation of Replication Protein A with the Ubiquitin Ligase Rad18 in DNA Damage Bypass. *Cell Cycle* **2008**, *7*, 3629–3633.

(55) Watanabe, K.; Tateishi, S.; Kawasuji, M.; Tsurimoto, T.; Inoue, H.; Yamaizumi, M. Rad18 Guides Polη to Replication Stalling Sites Through Physical Interaction and PCNA Monoubiquitination. *EMBO J* **2004**, *23*, 3886–3896.

(56) Vaziri, C.; Masai, H. Integrating DNA Replication with Trans-Lesion Synthesis via Cdc7. *Cell Cycle* **2010**, *9*, 4818–4823.

(57) Day, T. A.; Palle, K.; Barkley, L. R.; Kakusho, N.; Zou, Y.; Tateishi, S.; Verreault, A.; Masai, H.; Vaziri, C. Phosphorylated Rad18 Directs DNA Polymerase η to Sites of Stalled Replication. *J Cell Biol* **2010**, *191*, 953–966.

(58) Durando, M.; Tateishi, S.; Vaziri, C. A Non-Catalytic Role of DNA Polymerase η in Recruiting Rad18 and Promoting PCNA Monoubiquitination at Stalled Replication Forks. *Nucleic Acids Res* **2013**, *41*, 3079–3093.

(59) Sale, J. E.; Lehmann, A. R.; Woodgate, R. Y-Family DNA Polymerases and Their Role in Tolerance of Cellular DNA Damage. *Nat Rev Mol Cell Biol* **2012**, *13*, 141–152.

(60) Waters, L. S.; Minesinger, B. K.; Wiltrout, M. E.; D'Souza, S.; Woodruff, R. V.; Walker, G. C. Eukaryotic Translesion Polymerases and Their Roles and Regulation in DNA Damage Tolerance. *Microbiol Mol Biol Rev* **2009**, *73*, 134–154.

(61) Temviriyanukul, P.; van Hees-Stuivenberg, S.; Delbos, F.; Jacobs, H.; de Wind, N.; Jansen, J. G. Temporally Distinct Translesion Synthesis Pathways for Ultraviolet Light-Induced Photoproducts in the Mammalian Genome. *DNA Repair* **2012**, *11*, 550–558.

(62) Quinet, A.; Vessoni, A. T.; Rocha, C. R. R.; Gottifredi, V.; Biard, D. S. F.; Sarasin, A.; Menck, C. F. M.; Stary, A. Gap-Filling and Bypass at the Replication Fork Are Both Active Mechanisms for Tolerance of Low-Dose Ultraviolet-Induced DNA Damage in the Human Genome. *DNA Repair* **2014**, *14*, 27–38.

(63) Cordonnier, A. M.; Fuchs, R. P. P. Replication of Damaged DNA: Molecular Defect in Xeroderma Pigmentosum Variant Cells. *Mutat Res-DNA Repair* **1999**, *435*, 111–119.

(64) Gratchev, A.; Strein, P.; Utikal, J.; Goerdt, S. Molecular Genetics of Xeroderma Pigmentosum Variant. *Exp Dermatol* **2003**, *12*, 529–536.

(65) Cordonnier, A. M.; Lehmann, A. R.; Fuchs, R. P. P. Impaired Translesion Synthesis in Xeroderma Pigmentosum Variant Extracts. *Mol Cell Bio* **1999**, *19*, 2206–2211.

(66) Masutani, C.; Kusumoto, R.; Yamada, A.; Dohmae, N.; Yokoi, M.; Yuasa, M.; Araki, M.; Iwai, S.; Takio, K.; Hanaoka, F. The *XPV* (Xeroderma Pigmentosum Variant) Gene Encodes Human DNA Polymerase η. *Nature* **1999**, *399*, 700–704.

(67) Yamada, A.; Masutani, C.; Iwai, S.; Hanaoka, F. Complementation of Defective Translesion Synthesis and UV Light Sensitivity in Xeroderma Pigmentosum Variant Cells by Human and Mouse DNA Polymerase η. *Nucleic Acids Res* **2000**, *28*, 2473–2480.

(68) Masutani, C.; Araki, M.; Yamada, A.; Kusumoto, R.; Nogimori, T.; Maekawa, T.; Iwai, S.; Hanaoka, F. Xeroderma Pigmentosum Variant (XP-v) Correcting Protein From HeLa Cells Has a Thymine Dimer Bypass DNA Polymerase Activity. *EMBO J* **1999**, *18*, 3491–3501.

(69) McDonald, J. P.; Rapic-Otrin, V.; Epstein, J. A.; Broughton, B. C.; Wang, X.; Lehmann, A. R.; Wolgemuth, D. J.; Woodgate, R. Novel Human and Mouse Homologs of *Saccharomyces cerevisiae* DNA Polymerase η. *Genomics* **1999**, *60*, 20–30.

(70) Kashiwagi, S.; Kuraoka, I.; Fujiwara, Y.; Hitomi, K.; Cheng, Q. J.; Fuss, J. O.; Shin, D. S.; Masutani, C.; Tainer, J. A.; Hanaoka, F.; et al. Characterization of a Y-Family DNA Polymerase Eta From the Eukaryotic Thermophile Alvinella Pompejana. *J Nucleic Acids* **2010**, *2010*, 1–13.

(71) Guo, C.; Kosarek-Stancel, J. N.; Tang, T.-S.; Friedberg, E. C. Y-Family DNA Polymerases in Mammalian Cells. *Cell Mol Life Sci* **2009**, *66*, 2363–2381.

(72) Pryor, J. M.; Dieckman, L. M.; Boehm, E. M.; Washington, M. T. Eukaryotic Y-Family Polymerases: a Biochemical and Structural Perspective. In *Nucleic Acid Polymerases*; Murakami, K. S.; Trakselis, M. A., Eds.; Nucleic Acids and Molecular Biology; Springer-Verlag: Berlin, Heidelberg, 2013; Vol. 30, pp. 85–108.

(73) Maher, V. M.; Ouellette, L. M.; Curren, R. D.; McCormick, J. J. Frequency of Ultraviolet Light-Induced Mutations is Higher in Xeroderma Pigmentosum Variant Cells Than in Normal Human Cells. *Nature* **1976**, *261*, 593–595.

(74) Maher, V. M.; Ouellette, L. M.; Curren, R. D.; Justin McCormick, J. Caffeine Enhancement of the Cytotoxic and Mutagenic Effect of Ultraviolet Irradiation in a Xeroderma Pigmentosum Variant Strain of Human Cells. *Biochem Bioph Res Co* **1976**, *71*, 228–234.

(75) Johnson, R. E.; Prakash, S.; Prakash, L. Efficient Bypass of a Thymine-Thymine Dimer by Yeast DNA Polymerase, Polη. *Science* **1999**, *283*, 1001–1004.

(76) Haracska, L.; Yu, S.-L.; Johnson, R. E.; Prakash, L.; Prakash, S. Efficient and Accurate Replication in the Presence of 7,8-dihydro-8-oxoguanine by DNA Polymerase η. *Nat Genet* **2000**, *25*, 458–461.

(77) de Padula, M.; Slezak, G.; van der Kemp, P. A.; Boiteux, S. The Post-Replication Repair RAD18 and RAD6 Genes Are Involved in the Prevention of Spontaneous Mutations Caused by 7,8-dihydro-8-oxoguanine in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **2004**, *32*, 5003–5010.

(78) Sakamoto, A. N.; Stone, J. E.; Kissling, G. E.; McCulloch, S. D.; Pavlov, Y. I.; Kunkel, T. A. Mutator Alleles of Yeast DNA Polymerase ζ. *DNA Repair* **2007**, *6*, 1829–1838.

(79) Lee, D.-H.; Pfeifer, G. P. Translesion Synthesis of 7,8-dihydro-8-oxo-2′-Deoxyguanosine by DNA Polymerase Eta in Vivo. *Mutat Res-Fund Mol M* **2008**, *641*, 19–26.

(80) Johnson, R. E.; Kondratick, C. M.; Prakash, S.; Prakash, L. *hRAD30* Mutations in the Variant Form of Xeroderma Pigmentosum. *Science* **1999**, *285*, 263–265.

(81) Boudsocq, F.; Ling, H.; Yang, W.; Woodgate, R. Structure-Based Interpretation of Missense Mutations in Y-Family DNA Polymerases and Their Implications for Polymerase Function and Lesion Bypass. *DNA Repair* **2002**, *1*, 343–358.

(82) Broughton, B. C.; Cordonnier, A. M.; Kleijer, W. J.; Jaspers, N. G. J.; Fawcett, H.; Raams, A.; Garritsen, V. H.; Stary, A.; Avril, M.-F.; Boudsocq, F.; et al. Molecular Analysis of Mutations in DNA Polymerase η in Xeroderma Pigmentosum-Variant Patients. *Proc Natl Acad Sci USA* **2002**, *99*, 815–820.

(83) Tanioka, M.; Masaki, T.; Ono, R.; Nagano, T.; Otoshi-Honda, E.; Matsumura, Y.; Takigawa, M.; Inui, H.; Miyachi, Y.; Moriwaki, S.; et al. Molecular Analysis of DNA Polymerase Eta Gene in Japanese Patients Diagnosed as Xeroderma Pigmentosum Variant Type. *J Invest Dermatol* **2007**, *127*, 1745–1751.

(84) Inui, H.; Oh, K.-S.; Nadem, C.; Ueda, T.; Khan, S. G.; Metin, A.; Gozukara, E.; Emmert, S.; Slor, H.; Busch, D. B.; et al. Xeroderma Pigmentosum-Variant Patients From America, Europe, and Asia. *J Invest Dermatol* **2008**, *128*, 2055–2068.

(85) Biertümpfel, C.; Zhao, Y.; Kondo, Y.; Ramón-Maiques, S.; Gregory, M.; Lee, J. Y.; Masutani, C.; Lehmann, A. R.; Hanaoka, F.; Yang, W. Structure and Mechanism of Human DNA Polymerase η. *Nature* **2010**, *465*, 1044–1048.

(86) Opletalova, K.; Bourillon, A.; Yang, W.; Pouvelle, C.; Armier, J.; Despras, E.; Martin, L.; Mateus, C.; Robert, C.; Kannouche, P. L.; et al. Correlation of Phenotype/Genotype in a Cohort of 23 Xeroderma Pigmentosum-Variant Patients Reveals 12 New Disease-Causing *POLH* Mutations. *Hum Mutat* **2014**, *35*, 117–128.

(87) Lin, Q.; Clark, A. B.; McCulloch, S. D.; Yuan, T.; Bronson, R. T.; Kunkel, T. A.; Kucherlapati, R. Increased Susceptibility to UV-Induced Skin Carcinogenesis in Polymerase η-Deficient Mice. *Cancer Res* **2006**, *66*, 87–94.

(88) Ohkumo, T.; Kondo, Y.; Yokoi, M.; Tsukamoto, T.; Yamada, A.; Sugimoto, T.; Kanao, R.; Higashi, Y.; Kondoh, H.; Tatematsu, M.; et al. UV-B Radiation Induces Epithelial Tumors in Mice Lacking DNA Polymerase η and Mesenchymal Tumors in Mice Deficient for DNA Polymerase ι. *Mol Cell Bio* **2006**, *26*, 7696–7706.

(89) McDonald, J. P.; Levine, A. S.; Woodgate, R. 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* **1997**, *147*, 1557–1568.

(90) Masutani, C.; Kusumoto, R.; Iwai, S.; Hanaoka, F. Mechanisms of Accurate Translesion Synthesis by Human DNA Polymerase η. *EMBO J* **2000**, *19*, 3100–3109.

(91) Zhang, Y.; Yuan, F.; Wu, X.; Rechkoblit, O.; Taylor, J.-S.; Geacintov, N. E.; Wang, Z. Error-Prone Lesion Bypass by Human DNA Polymerase η. *Nucleic Acids Res* **2000**, *28*, 4717–4724.

(92) Vaisman, A.; Masutani, C.; Hanaoka, F.; Chaney, S. G. Efficient Translesion Replication Past Oxaliplatin and Cisplatin GpG Adducts by Human DNA Polymerase η. *Biochemistry* **2000**, *39*, 4575–4580.

(93) Levine, R. L.; Miller, H.; Grollman, A.; Ohashi, E.; Ohmori, H.; Matsutani, C.; Hanaoka, F.; Moriya, M. Translesion DNA Synthesis Catalyzed by Human Pol η and Pol κ Across 1,N6-Ethenodeoxyadenosine. *J Biol Chem* **2001**, *276*, 18717–18721.

(94) Kusumoto, R.; Masutani, C.; Iwai, S.; Hanaoka, F. Translesion Synthesis by Human DNA Polymerase η Across Thymine Glycol Lesions. *Biochemistry* **2002**, *41*, 6090–6099.

(95) Kokoska, R. J.; McCulloch, S. D.; Kunkel, T. A. The Efficiency and Specificity of Apurinic/Apyrimidinic Site Bypass by Human DNA Polymerase η and Sulfolobus Solfataricus Dpo4. *J Biol Chem* **2003**, *278*, 50537–50545.

(96) McCulloch, S. D.; Kokoska, R. J.; Kunkel, T. A. Efficiency, Fidelity and Enzymatic Switching During Translesion DNA Synthesis. *Cell Cycle* **2004**, *3*, 580–583.

(97) McCulloch, S. D.; Kokoska, R. J.; Masutani, C.; Iwai, S.; Hanaoka, F.; Kunkel, T. A. Preferential *Cis–Syn* Thymine Dimer Bypass by DNA Polymerase η Occurs with Biased Fidelity. *Nature* **2004**, *428*, 97–100.

(98) McCulloch, S. D.; Kokoska, R. J.; Garg, P.; Burgers, P. M. J.; Kunkel, T. A. The Efficiency and Fidelity of 8-oxo-Guanine Bypass by DNA Polymerases δ and η. *Nucleic Acids Res* **2009**, *37*, 2830–2840.

(99) Friedberg, E. C. Suffering in Silence: the Tolerance of DNA Damage. *Nat Rev Mol Cell Biol* **2005**, *6*, 943–953.

(100) Ling, H.; Boudsocq, F.; Woodgate, R.; Yang, W. Crystal Structure of a Y-Family DNA Polymerase in Action: a Mechanism for Error-Prone and Lesion-Bypass Replication. *Cell* **2001**, *107*, 91–102.

(101) Trincao, J.; Johnson, R. E.; Escalante, C. R.; Prakash, S.; Prakash, L.; Aggarwal, A. K. Structure of the Catalytic Core of *S. cerevisiae* DNA Polymerase η: Implications for Translesion DNA Synthesis. *Mol Cell* **2001**, *8*, 417–426.

(102) Uljon, S. N.; Johnson, R. E.; Edwards, T. A.; Prakash, S.; Prakash, L.; Aggarwal, A. K. Crystal Structure of the Catalytic Core of Human DNA Polymerase Kappa. *Structure (Cambridge, MA, U.S.)* **2004**, *12*, 1395–1404.

(103) Ling, H.; Boudsocq, F.; Plosky, B. S.; Woodgate, R.; Yang, W. Replication of a *Cis-Syn* Thymine Dimer at Atomic Resolution. *Nature* **2003**, *424*, 1083–1087.

(104) Ling, H.; Boudsocq, F.; Woodgate, R.; Yang, W. Snapshots of Replication Through an Abasic Lesion: Structural Basis for Base Substitutions and Frameshifts. *Mol Cell* **2004**, *13*, 751–762.

(105) Yang, W.; Woodgate, R. What a Difference a Decade Makes: Insights Into Translesion DNA Synthesis. *Proc Natl Acad Sci USA* **2007**, *104*, 15591–15598.

(106) Glick, E.; Vigna, K. L.; Loeb, L. A. Mutations in Human DNA Polymerase η Motif II Alter Bypass of DNA Lesions. *EMBO J* **2001**, *20*, 7303–7312.

(107) Glick, E.; Chau, J. S.; Vigna, K. L.; McCulloch, S. D.; Adman, E. T.; Kunkel, T. A.; Loeb, L. A. Amino Acid Substitutions at Conserved Tyrosine 52 Alter Fidelity and Bypass Efficiency of Human DNA Polymerase η. *J Biol Chem* **2003**, *278*, 19341–19346.

(108) Park, H.; Zhang, K.; Ren, Y.; Nadji, S.; Sinha, N.; Taylor, J.-S.; Kang, C. Crystal Structure of a DNA Decamer Containing a *Cis-Syn* Thymine Dimer. *Proc Natl Acad Sci USA* **2002**, *99*, 15965–15970.

(109) Zhao, Y.; Biertümpfel, C.; Gregory, M. T.; Hua, Y.-J.; Hanaoka, F.; Yang, W. Structural Basis of Human DNA Polymerase η-Mediated Chemoresistance to Cisplatin. *Proc Natl Acad Sci USA* **2012**, *109*, 7269–7274.

(110) Zhao, Y.; Gregory, M. T.; Biertümpfel, C.; Hua, Y.-J.; Hanaoka, F.; Yang, W. Mechanism of Somatic Hypermutation at the WA Motif by Human DNA Polymerase η. *Proc Natl Acad Sci USA* **2013**, *110*, 8146–8151.

(111) Krahn, J. M.; Beard, W. A.; Miller, H.; Grollman, A. P.; Wilson, S. H. Structure of DNA Polymerase β with the Mutagenic DNA Lesion 8-oxodeoxyguanine Reveals Structural Insights Into Its Coding Potential. *Structure (Cambridge, MA, U.S.)* **2003**, *11*, 121–127.

(112) Haracska, L.; Johnson, R. E.; Unk, I.; Phillips, B.; Hurwitz, J.; Prakash, L.; Prakash, S. Physical and Functional Interactions of Human DNA Polymerase η with PCNA. *Mol Cell Bio* **2001**, *21*, 7199.

(113) Acharya, N.; Yoon, J.-H.; Hurwitz, J.; Prakash, L.; Prakash, S. DNA Polymerase η Lacking the Ubiquitin-Binding Domain Promotes Replicative Lesion Bypass in Humans Cells. *Proc Natl Acad Sci USA* **2010**, *107*, 10401–10405.

(114) Bienko, M.; Green, C. M.; Crosetto, N.; Rudolf, F.; Zapart, G.; Coull, B.; Kannouche, P. L.; Wilder, G.; Peter, M.; Lehmann, A. R.; et al. Ubiquitin-Binding Domains in Y-Family Polymerases Regulate Translesion Synthesis. *Science* **2005**, *310*, 1821–1824.

(115) McIntyre, J.; Vidal, A. E.; McLenigan, M. P.; Bomar, M. G.; Curti, E.; McDonald, J. P.; Plosky, B. S.; Ohashi, E.; Woodgate, R. Ubiquitin Mediates the Physical and Functional Interaction Between Human DNA Polymerases η and ι. *Nucleic Acids Res* **2013**, *41*, 1649-1660

(116) Tissier, A.; Kannouche, P. L.; Reck, M.-P.; Lehmann, A. R.; Fuchs, R. P. P.; Cordonnier, A. M. Co-Localization in Replication Foci and Interaction of Human Y-Family Members, DNA Polymerase Polη and REV1 Protein. *DNA Repair* **2004**, *3*, 1503–1514.

(117) Ohashi, E.; Murakumo, Y.; Kanjo, N.; Akagi, J.-I.; Masutani, C.; Hanaoka, F.; Ohmori, H. Interaction of hREV1 with Three Human Y-Family DNA Polymerases. *Genes Cells* **2004**, *9*, 523-531.

(118) Matsuda, T.; Bebenek, K.; Masutani, C.; Hanaoka, F.; Kunkel, T. A. Low Fidelity DNA Synthesis by Human DNA Polymerase-η. *Nature* **2000**, *404*, 1011–1013.

(119) Matsuda, T.; Bebenek, K.; Masutani, C.; Rogozin, I. B.; Hanaoka, F.; Kunkel, T. A. Error Rate and Specificity of Human and Murine DNA Polymerase η. *J Mol Bio* **2001**, *312*, 335–346.

(120) Ohashi, E.; Bebenek, K.; Matsuda, T.; Feaver, W. J.; Gerlach, V. L.; Friedberg, E. C.; Ohmori, H.; Kunkel, T. A. Fidelity and Processivity of DNA Synthesis by DNA Polymerase κ, the Product of the Human *DINB1* Gene. *J Biol Chem* **2000**, *275*, 39678–39686.

(121) Tissier, A.; McDonald, J. P.; Frank, E. G.; Woodgate, R. Pol ι, a Remarkably Error-Prone Human DNA Polymerase. *Genes Dev* **2000**, *14*, 1642–1650.

(122) McCulloch, S. D.; Kunkel, T. A. Measuring the Fidelity of Translesion DNA Synthesis. *Meth Enzymol* **2006**, *408*, 341–355.

(123) King, N. M.; Nikolaishvili-Feinberg, N.; Bryant, M. F.; Luche, D. D.; Heffernan, T. P.; Simpson, D. A.; Hanaoka, F.; Kaufmann, W. K.; Cordeiro-Stone, M. Overproduction of DNA Polymerase Eta Does Not Raise the Spontaneous Mutation Rate in Diploid Human Fibroblasts. *DNA Repair* **2005**, *4*, 714–724.

(124) Herman, K. N.; Toffton, S.; McCulloch, S. D. Detrimental Effects of UV-B Radiation in a Xeroderma Pigmentosum-Variant Cell Line. *Environ Mol Mutagen* **2014**.

(125) Maga, G.; Villani, G.; Crespan, E.; Wimmer, U.; Ferrari, E.; Bertocci, B.; Hübscher, U. 8-oxo-Guanine Bypass by Human DNA Polymerases in the Presence of Auxiliary Proteins. *Nature* **2007**, *447*, 606–608.

(126) McCulloch, S. D.; Kokoska, R. J.; Chilkova, O.; Welch, C. M.; Johansson, E.; Burgers, P. M. J.; Kunkel, T. A. Enzymatic Switching for Efficient and Accurate Translesion DNA Replication. *Nucleic Acids Res* **2004**, *32*, 4665–4675.

(127) McCulloch, S. D.; Wood, A.; Garg, P.; Burgers, P. M. J.; Kunkel, T. A. Effects of Accessory Proteins on the Bypass of a *Cis-Syn* Thymine-Thymine Dimer by *Saccharomyces cerevisiae* DNA Polymerase η. *Biochemistry* **2007**, *46*, 8888–8896.

(128) Rogozin, I. B.; Pavlov, Y. I.; Bebenek, K.; Matsuda, T.; Kunkel, T. A. Somatic Mutation Hotspots Correlate with DNA Polymerase η Error Spectrum. *Nat Immunol* **2001**, *2*, 530–536.

(129) Zeng, X.; Winter, D. B.; Kasmer, C.; Kraemer, K. H.; Lehmann, A. R.; Gearhart, P. J. DNA Polymerase η is an A-T Mutator in Somatic Hypermutation of Immunoglobulin Variable Genes. *Nat Immunol* **2001**, *2*, 537–541.

(130) Faili, A.; Aoufouchi, S.; Weller, S.; Vuillier, F.; Stary, A.; Sarasin, A.; Reynaud, C.-A.; Weill, J.-C. DNA Polymerase η is Involved in Hypermutation Occurring During Immunoglobulin Class Switch Recombination. *J Exp Med* **2004**, *199*, 265–270.

(131) Delbos, F.; De Smet, A.; Faili, A.; Aoufouchi, S.; Weill, J.-C.; Reynaud, C.-A. Contribution of DNA Polymerase η to Immunoglobulin Gene Hypermutation in the Mouse. *J Exp Med* **2005**, *201*, 1191–1196.

(132) Delbos, F.; Aoufouchi, S.; Faili, A.; Weill, J.-C.; Reynaud, C.-A. DNA Polymerase η is the Sole Contributor of A/T Modifications During Immunoglobulin Gene Hypermutation in the Mouse. *J Exp Med* **2007**, *204*, 17–23.

(133) Kawamoto, T.; Araki, K.; Sonoda, E.; Yamashita, Y. M.; Harada, K.; Kikuchi, K.; Masutani, C.; Hanaoka, F.; Nozaki, K.; Hashimoto, N. Dual Roles for DNA Polymerase η in Homologous DNA Recombination and Translesion DNA Synthesis. *Mol Cell* **2005**, *20*, 793–799.

(134) Bétous, R.; Rey, L.; Wang, G.; Pillaire, M.-J.; Puget, N.; Selves, J.; Biard, D. S. F.; Shin-ya, K.; Vasquez, K. M.; Cazaux, C.; et al. Role of TLS DNA Polymerases Eta and Kappa in Processing Naturally Occurring Structured DNA in Human Cells. *Mol Carcinog* **2009**, *48*, 369–378.

(135) Rey, L.; Sidorova, J. M.; Puget, N.; Boudsocq, F.; Biard, D. S. F.; Monnat, R. J. J.; Cazaux, C.; Hoffmann, J.-S. Human DNA Polymerase η is Required for Common Fragile Site Stability During Unperturbed DNA Replication. *Mol Cell Bio* **2009**, *29*, 3344–3354.

(136) Boyer, A.-S.; Grgurevic, S.; Cazaux, C.; Hoffmann, J.-S. The Human Specialized DNA Polymerases and Non-B DNA: Vital Relationships to Preserve Genome Integrity. *J Mol Bio* **2013**, *425*, 4767–4781.

(137) Kamath-Loeb, A. S.; Lan, L.; Nakajima, S.; Yasui, A.; Loeb, L. A. Werner Syndrome Protein Interacts Functionally with Translesion DNA Polymerases. *Proc Natl Acad Sci USA* **2007**, *104*, 10394–10399.

(138) Cordeiro-Stone, M.; Makhov, A. M.; Zaritskaya, L. S.; Griffith, J. D. Analysis of DNA Replication Forks Encountering a Pyrimidine Dimer in the Template to the Leading Strand. *J Mol Bio* **1999**, *289*, 1207–1218.

(139) Kanao, R.; Hanaoka, F.; Masutani, C. A Novel Interaction Between Human DNA Polymerase η and MutLα. *Biochem Bioph Res Co* **2009**, *389*, 40–45.

(140) Zlatanou, A.; Despras, E.; Braz-Petta, T.; Boubakour-Azzouz, I.; Pouvelle, C.; Stewart, G. S.; Nakajima, S.; Yasui, A.; Ishchenko, A. A.; Kannouche, P. L. The hMsh2-hMsh6 Complex Acts in Concert with Monoubiquitinated PCNA and Pol η in Response to Oxidative DNA Damage in Human Cells. *Mol Cell* **2011**, *43*, 649–662.

(141) Lv, L.; Wang, F.; Ma, X.; Yang, Y.; Wang, Z.; Liu, H.; Li, X.; Liu, Z.; Zhang, T.; Huang, M.; et al. Mismatch Repair Protein MSH2 Regulates Translesion DNA Synthesis Following Exposure of Cells to UV Radiation. *Nucleic Acids Res* **2013**, *41*, 10312–10322.

(142) Glick, E.; Anderson, J. P.; Loeb, L. A. In Vitro Production and Screening of DNA Polymerase η Mutants for Catalytic Diversity. *BioTechniques* **2002**, *33*, 1136–1144.

(143) Niimi, A.; Limsirichaikul, S.; Yoshida, S.; Iwai, S.; Masutani, C.; Hanaoka, F.; Kool, E. T.; Nishiyama, Y.; Suzuki, M. Palm Mutants in DNA Polymerases α and η Alter DNA Replication Fidelity and Translesion Activity. *Mol Cell Bio* **2004**, *24*, 2734–2746

.

(144) Sherry, S. T.; Ward, M. H.; Kholodov, M.; Baker, J.; Phan, L.; Smigielski, E. M.; Sirotkin, K. dbSNP: the NCBI Database of Genetic Variation. *Nucleic Acids Res* **2001**, *29*, 308–311.

(145) Suarez, S. C.; Beardslee, R. A.; Toffton, S. M.; McCulloch, S. D. Biochemical Analysis of Active Site Mutations of Human Polymerase η. *Mutat Res-Fund Mol M* **2013**, *745-746*, 46–54.

(146) Kunkel, T. A. DNA Replication Fidelity. *J Biol Chem* **2004**, *279*, 16895–16898.

(147) Washington, M. T.; Johnson, R. E.; Prakash, S.; Prakash, L. Fidelity and Processivity of *Saccharomyces cerevisiae* DNA Polymerase η. *J Biol Chem* **1999**, *274*, 36835–36838.

(148) Johnson, R. E.; Washington, M. T.; Prakash, S.; Prakash, L. Fidelity of Human DNA Polymerase η. *J Biol Chem* **2000**, *275*, 7447–7450.

(149) Washington, M. T.; Johnson, R. E.; Prakash, S.; Prakash, L. Accuracy of Thymine-Thymine Dimer Bypass by *Saccharomyces cerevisiae* DNA Polymerase η. *Proc Natl Acad Sci USA* **2000**, *97*, 3094–3099.

(150) Sabouri, N.; Viberg, J.; Goyal, D. K.; Johansson, E.; Chabes, A. Evidence for Lesion Bypass by Yeast Replicative DNA Polymerases During DNA Damage. *Nucleic Acids Res* **2008**, *36*, 5660–5667.

(151) Einolf, H. J.; Guengerich, F. P. Fidelity of Nucleotide Insertion at 8-oxo-7,8-dihydroguanine by Mammalian DNA Polymerase δ. *J Biol Chem* **2001**, *276*, 3764–3771.

(152) Brieba, L. G.; Eichman, B. F.; Kokoska, R. J.; Doublié, S.; Kunkel, T. A.; Ellenberger, T. Structural Basis for the Dual Coding Potential of 8-oxoguanosine by a High-Fidelity DNA Polymerase. *EMBO J* **2004**, *23*, 3452–3461.

(153) Shibutani, S.; Takeshita, M.; Grollman, A. P. Insertion of Specific Bases During DNA Synthesis Past the Oxidation-Damaged Base 8-oxodG. *Nature* **1991**, *349*, 431–434.

(154) Vaisman, A.; Woodgate, R. Unique Misinsertion Specificity of Polι May Decrease the Mutagenic Potential of Deaminated Cytosines. *EMBO J* **2001**, *20*, 6520–6529.

(155) Markkanen, E.; Castrec, B.; Villani, G.; Hübscher, U. A Switch Between DNA Polymerases δ and λ Promotes Error-Free Bypass of 8-oxo-G Lesions. *Proc Natl Acad Sci USA* **2012**, *109*, 20401–20406.

(156) Washington, M. T.; Johnson, R. E.; Prakash, L.; Prakash, S. Accuracy of Lesion Bypass by Yeast and Human DNA Polymerase η. *Proc Natl Acad Sci USA* **2001**, *98*, 8355–8360.

(157) Yuan, F.; Zhang, Y.; Rajpal, D. K.; Wu, X.; Guo, D.; Wang, M.; Taylor, J.-S.; Wang, Z. Specificity of DNA Lesion Bypass by the Yeast DNA Polymerase η. *J Biol Chem* **2000**, *275*, 8233–8239.

(158) Silverstein, T. D.; Johnson, R. E.; Jain, R.; Prakash, L.; Prakash, S.; Aggarwal, A. K. Structural Basis for the Suppression of Skin Cancers by DNA Polymerase η. *Nature* **2010**, *465*, 1039–1043.

(159) Bassett, E.; Vaisman, A.; Havener, J. M.; Masutani, C.; Hanaoka, F.; Chaney, S. G. Efficiency of Extension of Mismatched Primer Termini Across From Cisplatin and Oxaliplatin Adducts by Human DNA Polymerases β and η in Vitro. *Biochemistry* **2003**, *42*, 14197–14206.

(160) Bassett, E.; Vaisman, A.; Tropea, K. A.; McCall, C. M.; Masutani, C.; Hanaoka, F.; Chaney, S. G. Frameshifts and Deletions During in Vitro Translesion Synthesis Past Pt-DNA Adducts by DNA Polymerases β and η. *DNA Repair* **2002**, *1*, 1003–1016.

(161) Ummat, A.; Rechkoblit, O.; Jain, R.; Roy Choudhury, J.; Johnson, R. E.; Silverstein, T. D.; Buku, A.; Lone, S.; Prakash, L.; Prakash, S.; et al. Structural Basis for Cisplatin DNA Damage Tolerance by Human Polymerase η During Cancer Chemotherapy. *Nat Struct Mol Biol* **2012**, *19*, 628–632.

(162) Silverstein, T. D.; Jain, R.; Johnson, R. E.; Prakash, L.; Prakash, S.; Aggarwal, A. K. Structural Basis for Error-Free Replication of Oxidatively Damaged DNA by Yeast DNA Polymerase η. *Structure (Cambridge, MA, U.S.)* **2010**, *18*, 1463–1470.

(163) Bebenek, K.; Kunkel, T. A. Analyzing Fidelity of DNA Polymerases. *Meth Enzymol* **1995**, *262*, 217–232.

(164) Haracska, L.; Kondratick, C. M.; Unk, I.; Prakash, S.; Prakash, L. Interaction with PCNA is Essential for Yeast DNA Polymerase η Function. *Mol Cell* **2001**, *8*, 407–415.

(165) McCulloch, S. D.; Kunkel, T. A. Multiple Solutions to Inefficient Lesion Bypass by T7 DNA Polymerase. *DNA Repair* **2006**, *5*, 1373–1383.

(166) The Pymol Molecular Graphics System. Version 1.3. Schroedinger, LLC.

(167) Beardslee, R. A.; Suarez, S. C.; Toffton, S. M.; McCulloch, S. D. Mutation of the Little Finger Domain in Human DNA Polymerase η Alters Fidelity When Copying Undamaged DNA. *Environ Mol Mutagen* **2013**, *54*, 638–651.

(168) Kunkel, T. A.; Patel, S. S.; Johnson, K. A. Error-Prone Replication of Repeated DNA Sequences by T7 DNA Polymerase in the Absence of Its Processivity Subunit. *Proc Natl Acad Sci USA* **1994**, *91*, 6830–6834.

(169) Kroutil, L. C.; West Frey, M.; Kaboord, B. F.; Kunkel, T. A.; Benkovic, S. J. Effect of Accessory Proteins on T4 DNA Polymerase Replication Fidelity. *J Mol Bio* **1998**, *278*, 135–146.

(170) Bebenek, A.; Carver, G. T.; Dressman, H. K.; Kadyrov, F. A.; Haseman, J. K.; Petrov, V.; Konigsberg, W. H.; Karam, J. D.; Drake, J. W. Dissecting the Fidelity of Bacteriophage RB69 DNA Polymerase: Site-Specific Modulation of Fidelity by Polymerase Accessory Proteins. *Genetics* **2002**, *162*, 1003–1018.

(171) Bebenek, A.; Carver, G. T.; Kadyrov, F. A.; Kissling, G. E.; Drake, J. W. Processivity Clamp Gp45 and ssDNA-Binding-Protein Gp32 Modulate the Fidelity of Bacteriophage RB69 DNA Polymerase in a Sequence-Specific Manner, Sometimes Enhancing and Sometimes Compromising Accuracy. *Genetics* **2005**, *169*, 1815–1824.

(172) Bloom, L. B.; Chen, X.; Fygenson, D. K.; Turner, J.; O'Donnell, M.; Goodman, M. F. Fidelity of Escherichia Coli DNA Polymerase III Holoenzyme. *J Biol Chem* **1997**, *272*, 27919–27930.

(173) Kobayashi, S.; Valentine, M. R.; Pham, P.; O'Donnell, M.; Goodman, M. F. Fidelity of Escherichia Coli DNA Polymerase IV. *J Biol Chem* **2002**, *277*, 34198–34207.

(174) Perales, C.; Cava, F.; Meijer, W. J. J.; Berenguer, J. Enhancement of DNA, cDNA Synthesis and Fidelity at High Temperatures by a Dimeric Single-Stranded DNA-Binding Protein. *Nucleic Acids Res* **2003**, *31*, 6473–6480.

(175) Roberts, J. D.; Hamatake, R. K.; Fitzgerald, M. S.; Sugino, A.; Kunkel, T. A. Effect of Accessory Proteins on the Fidelity of DNA Synthesis by Eukaryotic Replicative Polymerases. *Prog Clin Biol Res* **1990**, *340A*, 91–100.

(176) Kunkel, T. A.; Roberts, J. D.; Sugino, A. The Fidelity of DNA Synthesis by the Catalytic Subunit of Yeast DNA Polymerase α Alone and with Accessory Proteins. *Mutat Res* **1991**, *250*, 175–182.

(177) Carty, M. P.; Levine, A. S.; Dixon, K. HeLa Cell Single-Stranded DNA-Binding Protein Increases the Accuracy of DNA Synthesis by DNA Polymerase α in Vitro. *Mutat Res-DNA Repair* **1992**, *274*, 29–43.

(178) Suzuki, M.; Izuta, S.; Yoshida, S. DNA Polymerase α Overcomes an Error-Prone Pause Site in the Presence of Replication Protein-A. *J Biol Chem* **1994**, *269*, 10225–10228.

(179) Maga, G.; Frouin, I.; Spadari, S.; Hubscher, U. Replication Protein A as a "Fidelity Clamp" for DNA Polymerase α. *J Biol Chem* **2001**, *276*, 18235–18242.

(180) Mozzherin, D. J.; McConnell, M.; Jasko, M. V.; Krayevsky, A. A.; Tan, C. K.; Downey, K. M.; Fisher, P. A. Proliferating Cell Nuclear Antigen Promotes Misincorporation Catalyzed by Calf Thymus DNA Polymerase δ. *J Biol Chem* **1996**, *271*, 31711–31717.

(181) Hashimoto, K.; Shimizu, K.; Nakashima, N.; Sugino, A. Fidelity of DNA Polymerase δ Holoenzyme From *Saccharomyces cerevisiae*: The Sliding Clamp Proliferating Cell Nuclear Antigen Decreases Its Fidelity. *Biochemistry* **2003**, *42*, 14207–14213.

(182) Fortune, J. M.; Stith, C. M.; Kissling, G. E.; Burgers, P. M. J.; Kunkel, T. A. RPA and PCNA Suppress Formation of Large Deletion Errors by Yeast DNA Polymerase δ. *Nucleic Acids Res* **2006**, *34*, 4335–4341.

(183) Zhong, X.; Garg, P.; Stith, C. M.; Nick-McElhinny, S. A.; Kissling, G. E.; Burgers, P. M. J.; Kunkel, T. A. The Fidelity of DNA Synthesis by Yeast DNA Polymerase Zeta Alone and with Accessory Proteins. *Nucleic Acids Res* **2006**, *34*, 4731–4742.

(184) Frank, E. G.; McDonald, J. P.; Karata, K.; Huston, D.; Woodgate, R. A Strategy for the Expression of Recombinant Proteins Traditionally Hard to Purify. *Anal Biochem* **2012**, *429*, 132–139.

(185) Yang, Z.-G.; Liu, Y.; Mao, L. Y.; Zhang, J.-T.; Zou, Y. Dimerization of Human XPA and Formation of XPA2-RPA Protein Complex. *Biochemistry* **2002**, *41*, 13012–13020.

(186) Kanaar, R.; Hoeijmakers, J. H. J.; van Gent, D. C. Molecular Mechanisms of DNA Double-Strand Break Repair. *Trends Cell Biol* **1998**, *8*, 483–489.

(187) Wang, Y.; Woodgate, R.; McManus, T. P.; Mead, S.; McCormick, J. J.; Maher, V. M. Evidence That in Xeroderma Pigmentosum Variant Cells, Which Lack DNA Polymerase η, DNA Polymerase ι Causes the Very High Frequency and Unique Spectrum of UV-Induced Mutations. *Cancer Res* **2007**, *67*, 3018–3026.

(188) Zhang, Y.; Yuan, F.; Wu, X.; Taylor, J. S.; Wang, Z. Response of Human DNA Polymerase ι to DNA Lesions. *Nucleic Acids Res* **2001**, *29*, 928–935.

(189) Haracska, L.; Prakash, L.; Prakash, S. Role of Human DNA Polymerase Kappa as an Extender in Translesion Synthesis. *Proc Natl Acad Sci USA* **2002**, *99*, 16000–16005.

(190) Kirouac, K. N.; Ling, H. Unique Active Site Promotes Error-Free Replication Opposite an 8-oxo-Guanine Lesion by Human DNA Polymerase Iota. *Proc Natl Acad Sci USA* **2011**, *108*, 3210–3215.

(191) Irimia, A.; Eoff, R. L.; Guengerich, F. P.; Egli, M. Structural and Functional Elucidation of the Mechanism Promoting Error-Prone Synthesis by Human DNA Polymerase Kappa Opposite the 7,8-dihydro-8-oxo-2'-deoxyguanosine Adduct. *J Biol Chem* **2009**, *284*, 22467–22480.

(192) Carpio, R. V.-D.; Silverstein, T. D.; Lone, S.; Swan, M. K.; Choudhury, J. R.; Johnson, R. E.; Prakash, S.; Prakash, L.; Aggarwal, A. K. Structure of Human DNA Polymerase κ Inserting dATP Opposite an 8-oxoG DNA Lesion. *PLoS ONE* **2009**, *4*, e5766.

(193) Hendel, A.; Krijger, P. H. L.; Diamant, N.; Goren, Z.; Langerak, P.; Kim, J.; Reiβner, T.; Lee, K.-Y.; Geacintov, N. E.; Carell, T.; et al. PCNA Ubiquitination is Important, but Not Essential for Translesion DNA Synthesis in Mammalian Cells. *PLoS GENETICS* **2011**, *7*, e1002262.

(194) Acharya, N.; Yoon, J.-H.; Gali, H.; Unk, I.; Haracska, L.; Johnson, R. E.; Hurwitz, J.; Prakash, L.; Prakash, S. Roles of PCNA-Binding and Ubiquitin-Binding Domains in Human DNA Polymerase η in Translesion DNA Synthesis. *Proc Natl Acad Sci USA* **2008**, *105*, 17724–17729.

(195) Guo, C.; Fischhaber, P. L.; Luk-Paszyc, M. J.; Masuda, Y.; Zhou, J.; Kamiya, K.; Kisker, C.; Friedberg, E. C. Mouse Rev1 Protein Interacts with Multiple DNA Polymerases Involved in Translesion DNA Synthesis. *EMBO J* **2003**, *22*, 6621–6630.

(196) Masuda, Y.; Kamiya, K. Role of Single-Stranded DNA in Targeting REV1 to Primer Termini. *J Biol Chem* **2006**, *281*, 24314–24321.

(197) Chen, Y.-W.; Cleaver, J. E.; Hatahet, Z.; Honkanen, R. E.; Chang, J.-Y.; Yen, Y.; Chou, K.-M. Human DNA Polymerase η Activity and Translocation is Regulated by Phosphorylation. *Proc Natl Acad Sci USA* **2008**, *105*, 16578–16583.

(198) Wilson, T. M.; Vaisman, A.; Martomo, S. A.; Sullivan, P.; Lan, L.; Hanaoka, F.; Yasui, A.; Woodgate, R.; Gearhart, P. J. MSH2-MSH6 Stimulates DNA Polymerase η, Suggesting a Role for A:T Mutations in Antibody Genes. *J Exp Med* **2005**, *201*, 637–645.

(199) Mudrak, S. V.; Welz-Voegele, C.; Jinks-Robertson, S. The Polymerase Eta Translesion Synthesis DNA Polymerase Acts Independently of the Mismatch Repair System to Limit Mutagenesis Caused by 7,8-dihydro-8-oxoguanine in Yeast. *Mol Cell Bio* **2009**, *29*, 5316–5326.

(200) Ni, T. T.; Marsischky, G. T.; Kolodner, R. D. MSH2 and MSH6 Are Required for Removal of Adenine Misincorporated Opposite 8-oxo-Guanine in *S. cerevisiae*. *Mol Cell* **1999**, *4*, 439–444.

(201) Chang, D.-Y.; Lu, A.-L. Functional Interaction of MutY Homolog with Proliferating Cell Nuclear Antigen in Fission Yeast, Schizosaccharomyces Pombe. *J Biol Chem* **2002**, *277*, 11853–11858.

(202) van Loon, B.; Markkanen, E.; Hübscher, U. Oxygen as a Friend and Enemy: How to Combat the Mutational Potential of 8-oxo-Guanine. *DNA Repair* **2010**, *9*, 604–616.

(203) van der Kemp, P. A.; de Padula, M.; Burguiere-Slezak, G.; Ulrich, H. D.; Boiteux, S. PCNA Monoubiquitylation and DNA Polymerase Ubiquitin-Binding Domain Are Required to Prevent 8-oxoguanine-Induced Mutagenesis in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **2009**, *37*, 2549–2559.

(204) Svilar, D.; Goellner, E. M.; Almeida, K. H.; Sobol, R. W. Base Excision Repair and Lesion-Dependent Subpathways for Repair of Oxidative DNA Damage. *Antiox Redox Signaling* **2011**, *14*, 2491–2507.

(205) Sobol, R. W. DNA Repair Polymerases. In *Nucleic Acid Polymerases*; Murakami, K. S.; Trakselis, M. A., Eds.; Springer-Verlag: Berlin, Heidelberg, 2013; Vol. 30, pp. 43–83.

(206) Taggart, D. J.; Camerlengo, T. L.; Harrison, J. K.; Sherrer, S. M.; Kshetry, A. K.; Taylor, J.-S.; Huang, K.; Suo, Z. A High-Throughput and Quantitative Method to Assess the Mutagenic Potential of Translesion DNA Synthesis. *Nucleic Acids Res* **2013,** *41*, e96.

(207) Köberle, B.; Tomicic, M. T.; Usanova, S.; Kaina, B. Cisplatin Resistance: Preclinical Findings and Clinical Implications. *Biochim Biophys Acta* **2010**, *1806*, 172–182.

(208) Galluzzi, L.; Senovilla, L.; Vitale, I.; Michels, J.; Martins, I.; Kepp, O.; Castedo, M.; Kroemer, G. Molecular Mechanisms of Cisplatin Resistance. *Oncogene* **2012**, *31*, 1869–1883.

(209) Ceppi, P.; Novello, S.; Cambieri, A.; Longo, M.; Monica, V.; Iacono, Lo, M.; Giaj-Levra, M.; Saviozzi, S.; Volante, M.; Papotti, M.; et al. Polymerase η mRNA Expression Predicts Survival of Non-Small Cell Lung Cancer Patients Treated with Platinum-Based Chemotherapy. *Clin Cancer Res* **2009**, *15*, 1039–1045

.

(210) Teng, K.-Y.; Qiu, M.-Z.; Li, Z.-H.; Luo, H.-Y.; Zeng, Z.-L.; Luo, R.-Z.; Zhang, H.-Z.; Wang, Z.-Q.; Li, Y.-H.; Xu, R.-H. DNA Polymerase η Protein Expression Predicts Treatment Response and Survival of Metastatic Gastric Adenocarcinoma Patients Treated with Oxaliplatin-Based Chemotherapy. *J Transl Med* **2010**, *8*, 126.

(211) Zhou, W.; Chen, Y.-W.; Liu, X.; Chu, P.; Loria, S.; Wang, Y.; Yen, Y.; Chou, K.-M. Expression of DNA Translesion Synthesis Polymerase η in Head and Neck Squamous Cell Cancer Predicts Resistance to Gemcitabine and Cisplatin-Based Chemotherapy. *PLoS ONE* **2013**, *8*, e83978.

(212) Punchihewa, C.; Inoue, A.; Hishiki, A.; Fujikawa, Y.; Connelly, M.; Evison, B.; Shao, Y.; Heath, R.; Kuraoka, I.; Rodrigues, P.; et al. Identification of Small Molecule Proliferating Cell Nuclear Antigen (PCNA) Inhibitor That Disrupts Interactions with PIP-Box Proteins and Inhibits DNA Replication. *J Biol Chem* **2012**, *287*, 14289–14300.

(213) Inoue, A.; Kikuchi, S.; Hishiki, A.; Shao, Y.; Heath, R.; Evison, B. J.; Actis, M.; Canman, C. E.; Hashimoto, H.; Fujii, N. A Small Molecule Inhibitor of Monoubiquitinated Proliferating Cell Nuclear Antigen (PCNA) Inhibits Repair of Interstrand DNA Crosslink, Enhances DNA Double-Strand Break, and Sensitizes Cancer Cells to Cisplatin. *J Biol Chem* **2014**.

(214) Hanahan, D.; Weinberg, R. A. Hallmarks of Cancer: the Next Generation. *Cell* **2011**, *144*, 646–674.

(215) Ventura, I.; Russo, M. T.; De Luca, G.; Bignami, M. Oxidized Purine Nucleotides, Genome Instability and Neurodegeneration. *Mutat Res-Gen Tox En* **2010**, *703*, 59–65.