An Investigation into the Role, Recruitment, and Regulation of PrimPol in DNA Replication Restart

A thesis submitted to the University of Sussex for the degree of Doctor of Philosophy

Thomas A. Guilliam

March 2017

DECLARATION

This thesis conforms to an 'article format' in which the middle chapters consist of discrete articles which have been published in peer-reviewed journals. The first and final chapters present an overview and discussion of the field and the research undertaken, these chapters contain material already published in review articles.

The second half of Chapter 1 is published in *Nucleic Acids Research* as:

Guilliam, T.A., Keen, B.A., Brissett, N.C., Doherty, A.J., 2015. Primase-polymerases are a functionally diverse superfamily of replication and repair enzymes. Nucleic Acids Res. ***43****, 6651–6664. doi:10.1093/nar/gkv625*

Contributions: T.A.G. researched and wrote the majority of the paper and created four of the six figures presented. B.A.K. was involved in assisting with research and writing. N.C.B. created the crystallographic figures 1 and 5. A.J.D. was responsible for the initial concept of the review, in addition to co-writing the manuscript.

Chapter 2 is published in *Nucleic Acids Research* as:

Guilliam, T.A., Jozwiakowski, S.K., Ehlinger, A., Barnes, R.P., Rudd, S.G., Bailey, L.J., Skehel, J.M., Eckert, K.A., Chazin, W.J., Doherty, A.J., 2015. Human PrimPol is a highly error-prone polymerase regulated by single-stranded DNA binding proteins. Nucleic Acids Res. ***43****, 1056–1068. doi:10.1093/nar/gku1321*

Contributions: T.A.G. purified human PrimPol, RPA, and mtSSB, performed all gel-based biochemical experiments, generated the figures, assisted in pSJ4-*lacZα* polymerase-fidelity assay analysis, developed the model, and wrote the manuscript. S.K.J. purified human Pol δ and PCNA, developed and performed the pSJ4-*lacZα* fidelity assay, and assisted in development of the model. A.E. and W.J.C. performed the nuclear magnetic resonance (NMR) experiments. R.P.B. and K.A.E performed the HSV-tk forward mutation assay. S.G.R. performed the affinity purification of PrimPol for mass spectrometry (MS) and western blot validation. M.J.S. was responsible for MS analysis. A.J.D. was involved in experimental development and co-writing the manuscript.

Chapter 3 is published in *Nucleic Acids Research* as:

Guilliam, T.A., Bailey, L.J., Brissett, N.C., Doherty, A.J., 2016. PolDIP2 interacts with human PrimPol and enhances its DNA polymerase activities. Nucleic Acids Res. ***44****, 3317–3329. doi:10.1093/nar/gkw175*

Contributions: T.A.G. purified all proteins, developed and performed all experiments, analysed the data, generated the figures, and wrote the manuscript. L.J.B. was responsible for generating the PrimPol$^{-/-}$ MRC5 cells. N.C.B. was involved in technical assistance during MS analysis. A.J.D. oversaw the development of experiments and co-wrote the manuscript.

The first half of Chapter 4 is in press at *Methods in Enzymology* as:

Guilliam, T.A., Doherty, A.J., 2017. Current and emerging assays for studying the primer synthesis activities of DNA primases. Methods Enzymol.

Contributions: T.A.G. developed the fluorescence-based primase assay, wrote the manuscript, and produced the figures. A.J.D. developed the initial concept of the paper and co-wrote the manuscript.

The majority of Chapter 5 is in press at *Nature Communications* as:

Guilliam, T.A., Brissett, N.C., Ehlinger, A., Keen, B.A., Kolesar, P., Taylor, E., Bailey, L.J., Lindsay, H., Chazin, W.J., Doherty, A.J. 2017. Molecular basis for PrimPol recruitment to replication forks by RPA. Nat. Commun.

Contributions: T.A.G. purified human PrimPol, PrimPol C-terminal domain (CTD) cancer mutants, and RPA, generated the HEK293 and DT40 stable cell lines, performed the *in vitro* primase assay, gel filtration chromatography (GFC) analysis of CTD cancer mutants, immunoprecipitation experiments, DNA fibre analysis, triton fractionation, generated the figures, developed the model, and wrote the manuscript. N.C.B. was responsible for crystallography and GFC analysis of RBM mutants. A.E. and W.J.C. performed the NMR, isothermal titration calorimetry (ITC), dynamic light scattering (DLS), and size exclusion chromatography with multi-angle light scattering (SEC-MALS) experiments. B.A.K. Assisted in crystallography and gel filtration chromatography studies, and performed circular dichroism (CD). P.K. performed the yeast 2-hybrid experiments. E.T. and H.L. performed the chromatin isolation from *Xenopus* egg extract. L.J.B. generated the wild-type human PrimPol stable DT40 cell line. A.J.D. was involved in experimental design and co-writing of the manuscript.

The majority of Chapter 6 is published in *genes* as:

> *Guilliam, T.A., Doherty, A.J., 2017. PrimPol—Prime Time to Reprime. Genes* ***8****, 20. doi:10.3390/genes8010020*

Contributions: T.A.G. wrote the manuscript and generated the figures. A.J.D. developed the initial concept and co-wrote the manuscript.

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any degree.

Thomas A. Guilliam

University of Sussex

Thomas A. Guilliam

Doctor of Philosophy Genome Stability

An investigation into the role, recruitment, and regulation of PrimPol in DNA replication restart

Summary

DNA replication is one of life's fundamental processes. This delicate task is performed by a complex of molecular machines, known collectively as the replisome. At the heart of the replisome lie the replicative DNA polymerases which catalyse synthesis of daughter DNA strands with astonishing accuracy and efficiency. Nevertheless, these enzymes are prone to stalling upon encountering DNA damage lesions and secondary structures. In order to restart replication, DNA damage tolerance mechanisms are required. This published article-format thesis focusses on a recently discovered primase-polymerase, and member of the archaeo-eukaryotic primase (AEP) superfamily, involved in DNA damage tolerance, known as PrimPol. The work presented here builds on the initial characterisation of the enzyme, which identified potential roles in the bypass of DNA damage through translesion synthesis (TLS) and repriming of replication.

The first presented article consists of a review of the AEP superfamily, functionally repositioning the group under the category of primase-polymerases. In the second paper, evidence is presented to suggest that PrimPol's activity is regulated by single-strand binding proteins, required due to the enzyme's error-prone nature. In the third chapter, a novel PrimPol binding protein, polymerase delta-interacting protein 2 (PolDIP2), is identified and characterised as a stimulatory factor for PrimPol's primer extension activities. Chapter 4 focusses on the development and use of a gel-based fluorescent primase assay to assess PrimPol's ability to reprime downstream of DNA damage lesions and secondary structures. The fifth presented paper identifies the molecular basis for PrimPol's interaction with replication protein A (RPA). Using biophysical, biochemical, and cellular approaches, this paper identifies the mechanism by which PrimPol is recruited to reprime replication. Lastly, in Chapter 6, a review of the presented articles in the context of the wider literature is included. Together, this work supports a role for PrimPol in repriming and restarting DNA replication following stalling at impediments, as well as identifying mechanisms involved in the recruitment and regulation of the enzyme.

Acknowledgements

Firstly, I would like to thank my supervisor, Aidan Doherty, for giving me the opportunity to undertake my doctoral studies in his laboratory and for his guidance and advice over the past few years. Secondly, I would like to thank all members of the Doherty lab, including Stan, Nigel, Laura, Ben, Peter, Przemek, Pierre, and Farimah, who have together provided a supportive and enjoyable environment during my time here. In particular, thanks are due to Stan, not only for teaching me the majority of the biochemical techniques presented in this thesis, but also for his valuable advice on publishing papers and making my PhD productive. Special thanks also go to Laura for teaching me cell biology techniques, and Nigel for lots of technical support and fruitful discussions. I am also grateful to the previous 'PrimPol PhD students', including Ben, Sean, and Julie, for providing a foundation of PrimPol knowledge to build upon.

I would also like to thank Howard Lindsay and David Clancy at Lancaster University for giving me my first opportunity to work in a lab and inspiring me to continue. Little did I know as a lowly undergrad working on PrimPol in Howard's lab, that five years later I would still be working on PrimPol as a lowly PhD student. Thanks are also due to all our collaborators with whom I have shared publications over the past three years including the Chazin, Eckert, Lindsay, Sale, and Hirota laboratories. I'd also like to thank all members of the Genome Damage and Stability Centre for providing an excellent place to work. I am grateful to the University of Sussex for supporting my research.

Finally I would like to thank my family for supporting me from afar and Lauren for joining me 'down south' and keeping me sane during these past few years.

Table of Contents

Abbreviations

3AT	3-Amino-1,2,4-triazole
8-oxo-dG	8-oxo-2'-deoxyguanosine
A	Adenine
A.fu	*Archaeoglobus fulgidus*
aa	Amino acid
AEP	Archaeo-eukaryotic primase
AID	Activation-induced deaminase
Ap	Apurinic/apyrimidinic
APE1	AP endonuclease 1
APS	Ammonium persulphate
APTX	Aprataxin
ARS	Autonomously replicating sequence
AT	Ataxia telangiectasia
ATP	Adenosine triphosphate
ATR	AT and Rad3-related kinase
ATRIP	ATR-interacting protein
Bc	*Bacillus cereus*
BER	Base excision repair
BLM	Bloom syndrome helicase
bp	Base pair
BS^3	Bis(sulfosuccinimidyl)suberate
BSA	Bovine serum albumin
B-W	Binding-washing
C	Cytosine
CD	Circular dichroism
CDC	Cell division cycle
CDK	Cyclin-dependent kinase
CDT1	CDC10 Target 1
Chk1	Checkpoint kinase 1
CldU	5-chloro-2'-deoxyuridine
Cm	Chloramphenicol
CMG	CDC45, MCM, GINS
COC	Cyclic olefin copolymer
CPD	Cyclobutane pyrimidine dimer

CRISPR	Clustered regularly interspaced short palindromic repeat
CS	Cockayne syndrome
Csm3	Chromosome segregation in meiosis 3
CSP	Chemical shift perturbation
CTD	C-terminal domain
C-terminus	Carboxyl-terminus
Ctf4	Chromosome transmission fidelity protein 4
CTNA	Chain-terminating nucleotide analogue
dATP	Deoxyadenosine triphosphate
DBD	DNA-binding domain
dCTP	Deoxycytosine triphosphate
WT	Wild-type
dd	Dideoxynucleotide
ddH_2O	Double-distilled H_2O
DDK	Dbf4-dependent kinase
DDT	DNA damage tolerance
dGTP	Deoxyguanosine triphosphate
D-loop	Displacement-loop
DLS	Dynamic light scattering
DMEM	Dulbecco's modified eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNA-PKcs	DNA-dependent protein kinase catalytic subunit
dNMP	Deoxynucleoside monophosphate
dNTPs	Deoxynucleoside triphosphate
Dox	Doxycycline
Dpb	DNA polymerase binding protein
DSB	Double-strand break
DSBR	Double-strand break repair
dsDNA	Double-strand DNA
DTT	Dithiothreitol
dTTP	Deoxythymidine triphosphate
dUTP	Deoxyuridine triphosphate
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetra-acetic acid
EF	Error frequency
EGTA	Ethyleneglycoltetra-acetic acid

EM	Electron microscopy
EMSA	Electrophoretic mobility shift assay
Exo	Exonuclease
FACT	Facilitates chromatin transcription
FAM	Carboxyfluorescein
FANCJ	Fanconi anemia group J protein
FCS	Fetal calf serum
FEN1	Flap endonuclease 1
FL	Full-length
FPLC	Fast protein liquid chromatography
FUdR	Fluorodeoxyuridine
FWD	Forward
G	Guanine
G1-phase	Growth-phase 1
G2-phase	Growth-phase 2
GFC	Gel-filtration chromatography
GINS	Go-ichi-ni-san (Japanese for 5-1-2-3)
GST	Glutathione S-transferase
HEK293	Human embryonic kidney cell line 293
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HEX	Hexachlorofluorescein
HGT	Horizontal gene transfer
HIV	Human immunodeficiency virus
HLTF	Helicase-like transcription factor
HMG	High mobility group
hNRP	Nucleosome assembly protein 1-like 1
HPLC	High performance liquid chromatography
HR	Homologous recombination
HRP	Horse radish peroxidase
Hs	*Homo sapiens*
HSQC	Heteronuclear single quantum coherence
HSV	Herpes simplex virus
HTS	High-throughput sequencing
HU	Hydroxyurea
IdU	Iododeoxyuridine
Ig	Immunoglobulin
IMAC	Immobilised metal ion affinity chromatography

Indel	Insertion-deletion
IR	Ionising radiation
ITC	Isothermal titration calorimetry
Kb	Kilobase pair
LC-MS/MS	Liquid chromatography tandem mass spectrometry
LFQ	Label-free quantification
LIG	Ligase
LUCA	Last universal common ancestor
MCM	Mini-chromosome maintenance
MEF	Mouse embryonic fibroblast
MGR	Malachite green reagent
MLH	MutL homolog
MMEJ	Microhomology-mediated end joining
MMR	Mismatch repair
MMS	Methylmethanesulfonate
Mrc1	Mediator of replication checkpoint protein 1
mRNA	Messenger RNA
MS	Mass spectrometry
MSH	MutS homolog
Mtb	*Mycobacterium tuberculosis*
mtDNA	Mitochondrial DNA
MTS	Mitochondrial targeting sequence
mtSSB	Mitochondrial single-strand binding protein
NCLDV	Nucleo-cytoplasmic large DNA viruses
NER	Nucleotide excision repair
NFR	Nucleosome free region
NHEJ	Non-homologous end-joining
Ni-NTA	Nickel-nitrolotriacetic acid
NMR	Nuclear magnetic resonance
Nt	Nucleotide
NTD	N-terminal domain
N-terminus	Amino-terminus
OB	Oligonucleotide binding
OBD	Origin-binding domain
O_H	Origin of heavy strand DNA replication
O_L	Origin of light strand DNA replication
ORC	Origin recognition complex

OXPHOS	Oxidative phosphorylation
PAD	Polymerase associated domain
PAGE	Polyacrylamide gel electrophoresis
PARP	Poly(ADP-ribose) polymerase
PBS	Phosphate-buffered saline
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
Pfu	*Pyrococcus furiosus*
Pi	Inorganic phosphate
PIPES	Piperazine-N,N'-bis(2-ethanesulfonic acid)
PMS	Postmeiotic segregation
PMSF	Phenylmethylsulfonyl fluoride (PMSF)
PNKP	Bifunctional polynucleotide phosphatase/kinase
Pol	Polymerase
PolA1	DNA polymerase α catalytic subunit
PolA2	DNA polymerase α subunit B
PolD1	DNA polymerase δ catalytic subunit
PolD2	DNA polymerase δ subunit 2
PolD3	DNA polymerase δ subunit 3
PolD4	DNA polymerase δ subunit 4
PolDIP2	Polymerase delta-interacting protein 2
PolE1	DNA polymerase ε catalytic subunit
PolE2	DNA polymerase ε subunit 2
PolE3	DNA polymerase ε subunit 3
PolE4	DNA polymerase ε subunit 4
PolRMT	Mitochondrial RNA polymerase
PPi	Pyrophosphate
PPL	PrimPol-like protein
Pre-IC	Pre-initiation complex
Pre-RC	Pre-replication complex
PriL	DNA primase large subunit (archaeal)
Prim1	DNA primase small subunit (eukaryotic)
Prim2	DNA primase large subunit (eukaryotic)
PrimPol	Primase-Polymerase
PriS	DNA primase small subunit (archaeal)

PRR	Post-replicative repair
PSI-BLAST	Position-specific iterative basic local alignment search tool
pssDNA	Partial single-stranded DNA
Pu-seq	Polymerase-usage sequencing
PVT-PEI	Polyvinyl toluene-polyethyl-eneimine
Rad	Radiation (gene)
RBD	RPA-binding domain
RBM	RPA-binding motif
RCRE	Rolling circle replication endonuclease
RER	Ribonucleotide excision repair
Rev	Reversionless (gene)
RFC	Replication factor C
RI	Refractive index
RNA	Ribonucleic acid
RNAi	Ribonucleic acid interference
RNAP	RNA polymerase
RNase MRP	RNase mitochondrial RNA processing enzyme
rNTP	Ribonucleotide triphosphate
ROS	Reactive oxygen species
RPA	Replication protein A
RFC	Replication factor C
RPMI	Roswell park memorial institute medium
RRM	RNA recognition motif
SCJ	Sister chromatid junction
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulfate
SEC	Size-exclusion chromatography
SEC-MALS	SEC-Multi-angle light scattering
SHM	Somatic hypermutation
siRNA	Short interfering RNA
Sld	Synthetic lethality with dpb11
SPA	Scintillation proximity assay
S-phase	Synthesis-phase
SSB	Single-strand binding protein
SSBR	Single-strand break repair
ssDNA	Single-strand DNA
STR	Short tandem repeat

T	Thymine
TBE	Tris-borate-EDTA
TCEP	Tris(2-carboxyethyl)phosphine
TDP1	Tyrosyl-DNA phosphodiesterase 1
TdT	Terminal transferase
TEMED	N,N,N',N'-Tetramethylethylenediamine
TFAM	Mitochondrial transcription factor A
Tg	Thymine glycol
Tgo	*Thermococcus gorgonarius*
TLS	Translesion synthesis
T_m	Melting temperature
Tof1	Topoisomerase 1-associated factor 1
TOP	Topoisomerase
TOPRIM	Topoisomerase-primase fold
TS	Template switching
U	Uracil
UBM	Ubiquitin-binding motif
UBZ	Ubiquitin-binding zinc finger
UV	Ultraviolet
VACV	Vaccinia virus
WRN	Werner syndrome helicase
XP	Xeroderma pigmentosum
XPV	Xeroderma pigmentosum variant
XRCC	X-ray cross complementation protein
YPD	Yeast peptone dextrose
ZfKO	Zinc finger knockout
ZnF	Zinc finger
6-4PP	Pyrimidine (6-4) pyrimidone photoproduct

List of Figures

List of Tables

Chapter 1

Introduction

1. Introduction

In any system, replication, variation, and selection, given adequate time, facilitate the spontaneous emergence of organised complexity. In biology, these three pre-requisites form the fundamental basis which underpins the evolution of all life. At its core, biological replication is dependent upon the accurate and efficient duplication of the genome, with errors during this process generating the necessary variation for selection to occur. From a reductionist point of view, therefore, genome duplication, and more specifically DNA replication, is the underlying foundation for the emergence and evolution of all complex life. The introduction to this thesis provides a general background to the process of DNA replication, with a particular focus on the enzymes involved in the direct synthesis of DNA; primases and polymerases. Following a broad review of our current understanding of DNA replication, the strategies employed by the replisome to maintain progression and DNA synthesis in the presence of a range of obstacles will be discussed. Finally, the often under-appreciated role of archaeo-eukaryotic primases (AEPs) in DNA replication, repair, and damage tolerance will be highlighted, before moving on to introduce the focus of this thesis, a recently discovered vertebrate AEP termed PrimPol.

1.1. The Double Helix

In 1953, James Watson and Francis Crick published their now famous model for the structure of deoxyribose nucleic acid (DNA) (Watson and Crick, 1953). Based on unpublished X-ray data from Franklin, Gosling, and Wilkins, coupled with Chargaff's identification of base ratios, Watson and Crick proposed a right handed double-helical structure of DNA (Chargaff et al., 1952; Franklin et al., 1953; Wilkins et al., 1953; Zamenhof et al., 1952). Although previous work had identified DNA as the biological carrier of genetic information, the mechanism for how the molecule could store and transmit such information was not apparent until after the elucidation of its structure (Avery et al., 1944; Griffith, 1928; Hershey and Chase, 1952).

The DNA double helix is composed of two anti-parallel polymers of deoxyribonucleotides. Each deoxyribonucleotide consists of a nitrogenous base bound to the 1' carbon of a deoxyribose sugar, and a phosphate group bound to the 5' carbon of the sugar. The nucleotide monomers bind together via their phosphate groups, which form phosphodiester bonds between the 3' and 5' carbons of the adjacent deoxyribose sugars. Importantly, the asymmetric nature of these bonds gives DNA directionality due to its 5' and 3' termini. Two antiparallel DNA strands bind together in a right-handed, helical fashion by virtue of non-covalent hydrogen bonding between purine (adenine and guanine) and pyrimidine (cytosine and thymine) bases. Here, adenine and thymine pair,

and cytosine and guanine pair, forming two and three complementary hydrogen bonds with each other, respectively (Watson and Crick, 1953). Consequently, each DNA strand is bound to another, which encodes the reverse complementary sequence of its bases. The formation and stabilisation of the double helix through these hydrogen bonds, produces a structure in which the inward-facing nitrogenous bases lie perpendicular to the alternating sugar-phosphate backbone of each DNA strand.

The inherent complementary nature of the double helix structure of DNA, elucidated by Watson and Crick, immediately suggests a possible mechanism of duplication. Indeed the authors remarked that “It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material” (Watson and Crick, 1953). The copying mechanism referred to by Watson and Crick was a semi-conservative DNA replication model. Here, each parental DNA strand would provide a template for the synthesis of a complementary daughter DNA strand. This semi-conservative mode of DNA replication, inferred from the very structure of the molecule itself, was later confirmed, adding further elegance to the structure and function of DNA (Meselson and Stahl, 1958).

Subsequent studies were able to determine the nature of the genetic code stored in DNA (Crick, 1970). It was found that DNA provides the information required for the production of proteins in the sequence of bases which it encodes. DNA is transcribed into mRNA, each three bases of which are known as a ‘codon’; with each codon encoding a specific amino acid (Crick et al., 1961; Nirenberg et al., 1965; Tsugita and Fraenkel-Conrat, 1960). Thus, during the process of translation, each codon is read sequentially to produce a specific amino acid sequence and consequent protein. This linear flow of information in biological systems, from DNA and RNA to proteins, but never from protein to protein, or back to nucleic acid, was termed by Francis Crick “the central dogma of molecular biology” (Crick, 1970).

1.2. DNA Polymerases

Despite the, theoretically, relatively simple “copying mechanism” envisaged by Watson and Crick, the reality of duplicating the genome of any organism is a highly complicated process, requiring a coordinated complex of molecular machines known collectively as the replisome. At the core of the replisome lie the enzymes responsible for template-directed synthesis of the daughter DNA strands, these enzymes are known as DNA polymerases. In humans, DNA polymerases are tasked with accurately copying the ~6 billion base pair diploid genome, packaged into 46 chromosomes and compacted around histones to form chromatin.

1.2.1. A Brief History of DNA Polymerases

In 1956, only three years after the publication of the structure of DNA, the first enzyme capable of performing template-directed DNA synthesis was discovered in *Escherichia coli* (Kornberg et al., 1956). Two years later, this enzyme was purified and designated as “polymerase”, coining the term which would be used to describe the plethora of nucleotidyl transferase enzymes discovered thereafter (Lehman et al., 1958). These studies determined the requirement of deoxynucleotide triphosphates (dNTPs), a primed DNA template, and magnesium (Mg^{2+}) ions as a cofactor, for DNA polymerase activity. Thus highlighting that, although a whole replisome is required for accurate and efficient genome replication *in vivo*, DNA synthesis *in vitro* can be performed by a single enzyme in a relatively simple reaction mixture. In 1959, Arthur Kornberg was awarded The Nobel Prize for his work in discovering the first DNA polymerase. However, the enzyme he discovered, now known as DNA Pol I, is not the polymerase responsible for bulk DNA replication in *E. coli* (De Lucia and Cairns, 1969). Remarkably, the honour of discovering that polymerase was left to Arthur Kornberg’s son, Thomas Kornberg. In 1971, one year after identifying *E. coli* DNA Pol II, Thomas Kornberg and Malcolm Gefter discovered what would become DNA Pol III, the enzyme responsible for genome duplication in *E. coli* (Kornberg and Gefter, 1970, 1971). We now know that *E. coli* possess five DNA polymerases (Pols I, II, III, IV, V), with DNA Pol I playing a role in DNA damage repair and Okazaki fragment maturation (Lehman, 2003).

After the initial identification of DNA polymerase activity by Kornberg *et al.* in 1956, the discovery of polymerase activity in eukaryotes followed closely behind. In 1957, Bollum and Potter identified the activity of DNA Polymerase α (Pol α) in rat liver homogenates (Bollum and Potter, 1957). In the sixty years since these discoveries an array of DNA polymerases have been identified and characterised from all domains of life. These studies have revealed that DNA polymerases are essential not only for genome duplication, but also for the tolerance and repair of DNA damage. We now know that the human genome alone encodes at least 15 DNA-dependent DNA polymerases, excluding PrimPol (Figure 1.1.) (Lange et al., 2011). DNA polymerases can be grouped into seven distinct families based on their phylogenetic relationships with *E. coli* Pol I (A family), Pol II (B family), Pol III (C family), Euryarchaeotic Pol II (D family), human Pol β (X family), and *E. coli* UmuC/DinB and eukaryotic *RAD30*/xeroderma pigmentosum variant (Y family). The 15 human DNA-dependent DNA polymerases can all be grouped into just four of these, the A, B, X, and Y families (Braithwaite and Ito, 1993; Burgers et al., 2001; Ito and Braithwaite, 1991; Zhao and Washington, 2017).

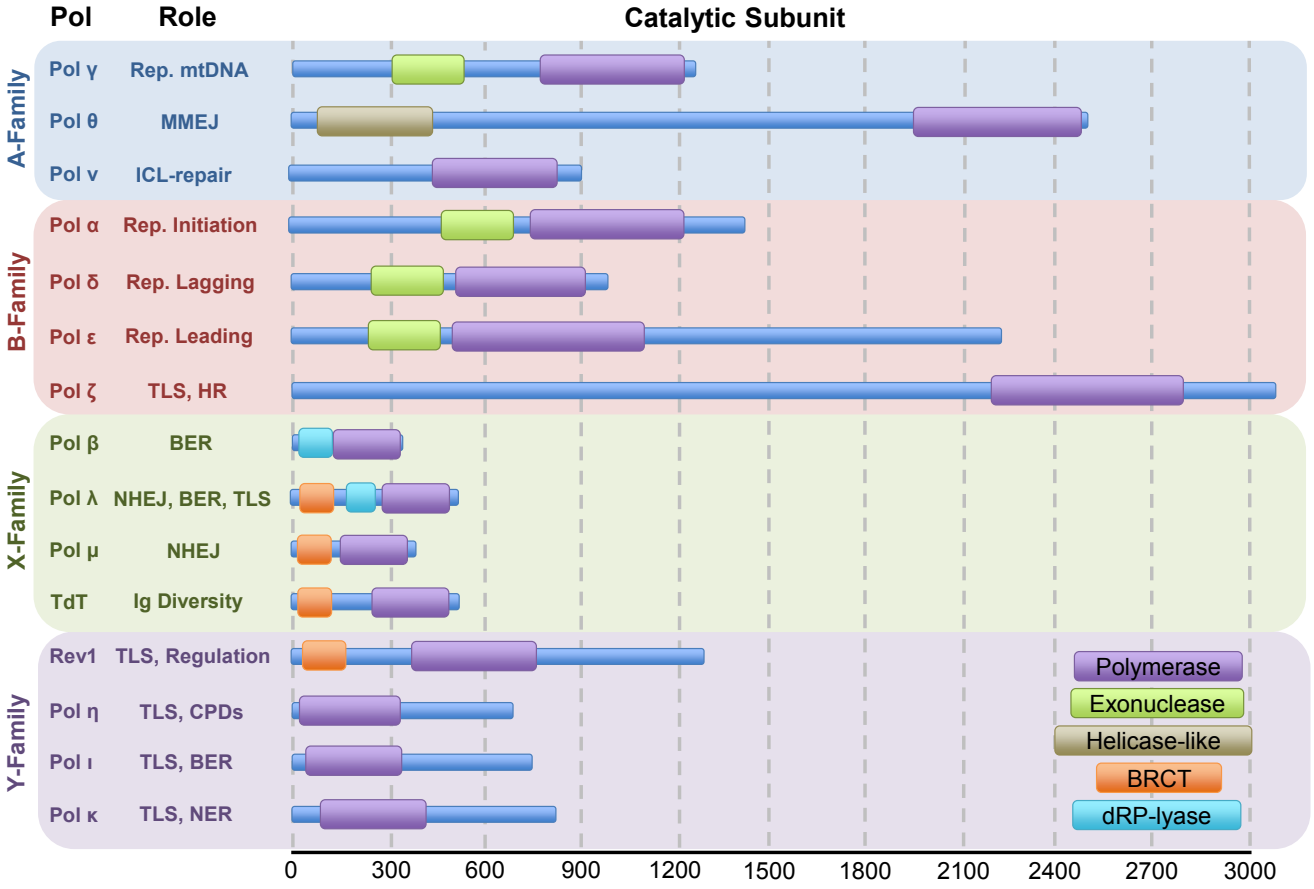


Figure 1.1. The domain architecture of the human DNA-dependent DNA polymerases.
The 15 human DNA-dependent DNA polymerases (excluding PrimPol) are arranged into their respective polymerase families, A-family (blue), B-family (red), X-family (green), and Y-family (purple). The predominant role and domain architecture of the catalytic subunit of each polymerase is indicated to the right, with the length in amino acids displayed below. Note that telomerase is omitted from this list as it is an RNA- not DNA- dependent DNA polymerase. Figure adapted from Lange et al., 2011.

1.2.2. The General Structure and Mechanism of DNA Polymerases

The versatile array of DNA polymerases uncovered since 1956 all conform to a highly conserved general structure and common mechanism of DNA synthesis. Thus, the information gleaned from early mechanistic studies of Pol I has proved applicable to all DNA polymerases discovered thereafter (Kornberg et al., 1956; Lehman et al., 1958; Rothwell and Waksman, 2005). DNA polymerases utilise single-stranded (ss) DNA templates to extend the 3' end of a short DNA or RNA chain (known as a primer) already base-paired to the template. Importantly, the majority of these enzymes cannot initiate DNA synthesise *de novo*, rather they rely on DNA primases (see section 1.6.) to synthesise the initial primer and provide them with the 3' hydroxyl group required for further extension. Following binding to a primer-template substrate, DNA polymerases are able to catalyse the addition of successive incoming dNTPs to the 3' end of the primer in a repetitive fashion (Figure 1.2.A.) (Lujan et al., 2016). Here, an incoming dNTP first base pairs with its complementary partner in the template strand through hydrogen bonding. The polymerase is then able to catalyse nucleophilic attack of the 3' hydroxyl group of the primer terminus on the α-phosphate group of the dNTP to be added. This releases two of the phosphate moieties from the incoming dNTP and provides the energy required to form a phosphodiester bond, thereby linking the incorporated dNTP with phosphate-sugar backbone of the primer strand (Rothwell and Waksman, 2005) (Figure 1.2.A.). In this manner, DNA polymerases always synthesise DNA in a 5' to 3' direction, whilst moving 3' to 5' on the anti-parallel template strand. This is an important feature when considering duplication of the double helix. DNA replication occurs bi-directionally on each strand, producing a forked structure. Since the two template DNA stands are anti-parallel to each other, one will allow continuous synthesis of nascent DNA in a 5' to 3' direction, this is known as the leading strand. However, since the complementary strand is anti-parallel it cannot permit continuous 5' to 3' synthesis and thus requires repeated repriming and discontinuous replication in segments termed Okazaki fragments. These fragments are subsequently processed and ligated together to form the lagging strand (Kainuma-Kuroda and Okazaki, 1975).

As noted in early studies, metal ions, generally Mg^{2+}, are required for DNA synthesis (Lehman et al., 1958). Polymerases utilise a two-metal ion mechanism (metal ion A and B) to catalyse nucleotide incorporation (Figure 1.2.B.). Each metal ion is held in the correct position and orientation by the carboxylate groups of conserved aspartate residues. Metal ion A binds to the 3' hydroxyl group of the primer to facilitate its nucleophilic attack on the incoming dNTP. Whilst the metal ion B binds to the incoming dNTP to facilitate release of the β and γ phosphate groups by stabilising their negative

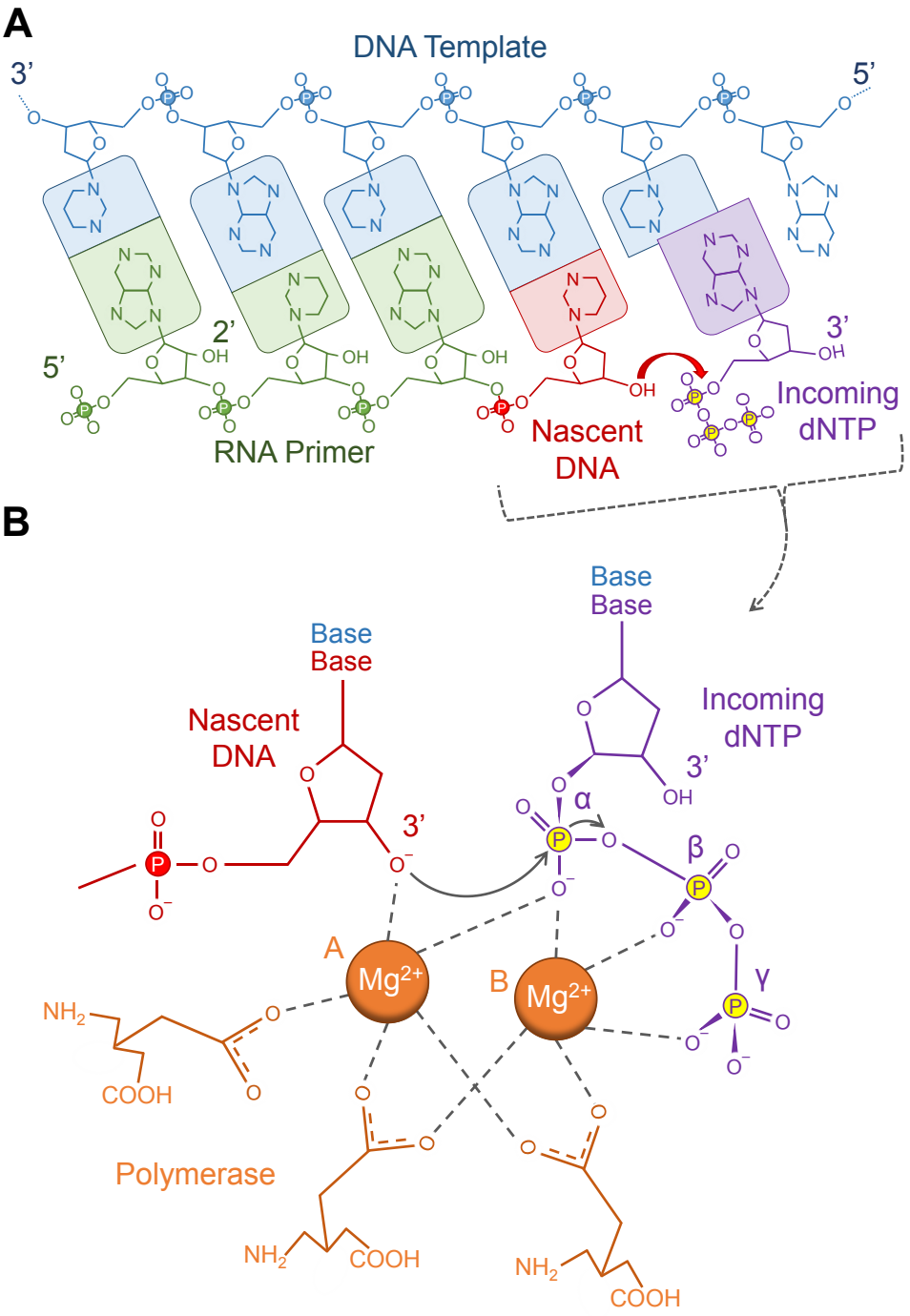


Figure 1.2. The mechanism of DNA synthesis by polymerases.
(A) DNA polymerases catalyse the nucleophilic attack of the 3' hydroxyl group of the growing DNA strand (red) on the 5' phosphate of the incoming dNTP (purple). This reaction is driven by the chemical energy supplied by the triphosphate (yellow) of the incoming dNTP. This mechanism of polymerisation provides directionality to DNA synthesis by only permitting polymerases to extend the 3' end of DNA chains. Due to the antiparallel nature of duplex DNA, polymerases move 3' to 5' on the template DNA strand (blue). Polymerases are unable to initiate DNA synthesis *de novo* and require a DNA or RNA primer. Here an RNA primer is shown (green), identifiable from its 2'-hydroxyl group. Figure adapted from Lujan et al., 2016.
(B) The two-metal ion mechanism of DNA synthesis. The two magnesium ions (orange) are coordinated by catalytic amino acids of the polymerase (orange). Metal ion A activate the 3'-hydroxyl group of the nascent DNA chain (red) for nucleophilic attack of the α-phosphate of the incoming dNTP (purple). Metal ion B interacts wth the β- and γ-phosphates of the dNTP to facilitate their removal. Together these interactions stabilise the pentavalent transition state required for synthesis.

charge. Together, the two metal ions stabilise the pentacovalent transition state to permit catalysis and dNTP incorporation into the growing primer strand (Brautigam and Steitz, 1998).

In order to achieve this synthesis all DNA polymerases conform to a common domain organisation and general overall structure. The first insights into this structure came from crystallographic resolution of the Klenow fragment of *E. coli* Pol I, the polymerase and exonuclease domains of the enzyme (Klenow and Overgaard-Hansen, 1970; Ollis et al., 1985). These studies revealed that the polymerase domain of the enzyme resembles a right hand, consisting of three subdomains termed the “fingers”, “thumb”, and “palm” (Figure 1.3.A. and B.) (Beard and Wilson, 2001; Ollis et al., 1985). Independent of the detailed structural features of their distinct domains, all polymerase structures uncovered to date share this overall architecture (Steitz, 1999). Here, the DNA molecule sits within the palm and is enclosed by the fingers and thumb subdomains, creating a “U” shaped structure (Ollis et al., 1985). The palm subdomain contains the polymerase active site, harbouring the catalytic carboxylate groups required for metal ion coordination, which facilitate phosphoryl transfer. Whilst the fingers subdomain interacts with the incoming dNTP and the thumb subdomain coordinates the DNA template (Figure 1.3.C).

During catalysis, conformational changes in these essential domains facilitate extension of the primer strand. Upon binding the primer-template substrate, a helix-loop-helix motif located in the thumb subdomain of the polymerase undergoes a conformational change to grip the DNA (Li et al., 1998a, 1998b). The fingers subdomain binds the incoming dNTP and subsequently undergoes a large conformational alteration involving an initial 6° rotation in the core of the subdomain, followed by a large 40° inward rotation of the helices in the tip of the fingers towards the palm domain (Li and Waksman, 2001). This change causes a closing of the “hand”, bringing the incoming dNTP into close proximity of the active site and further coordinating the primer-template substrate. In addition to facilitating phosphoryl transfer, this step is also important in ensuring incorporation of the correct incoming nucleotide. The tight closed-conformation of the enzyme around the base-paired incoming dNTP sterically hinders incorrect base pairing and thus aids in preventing misincorporation (Li et al., 1998a; Li and Waksman, 2001). Following phosphodiester bond formation an additional conformational change in the enzyme permits release of the β and γ phosphate groups from the incorporated dNTP as pyrophosphate. This is followed by either dissociation of the enzyme or the addition of another nucleotide.

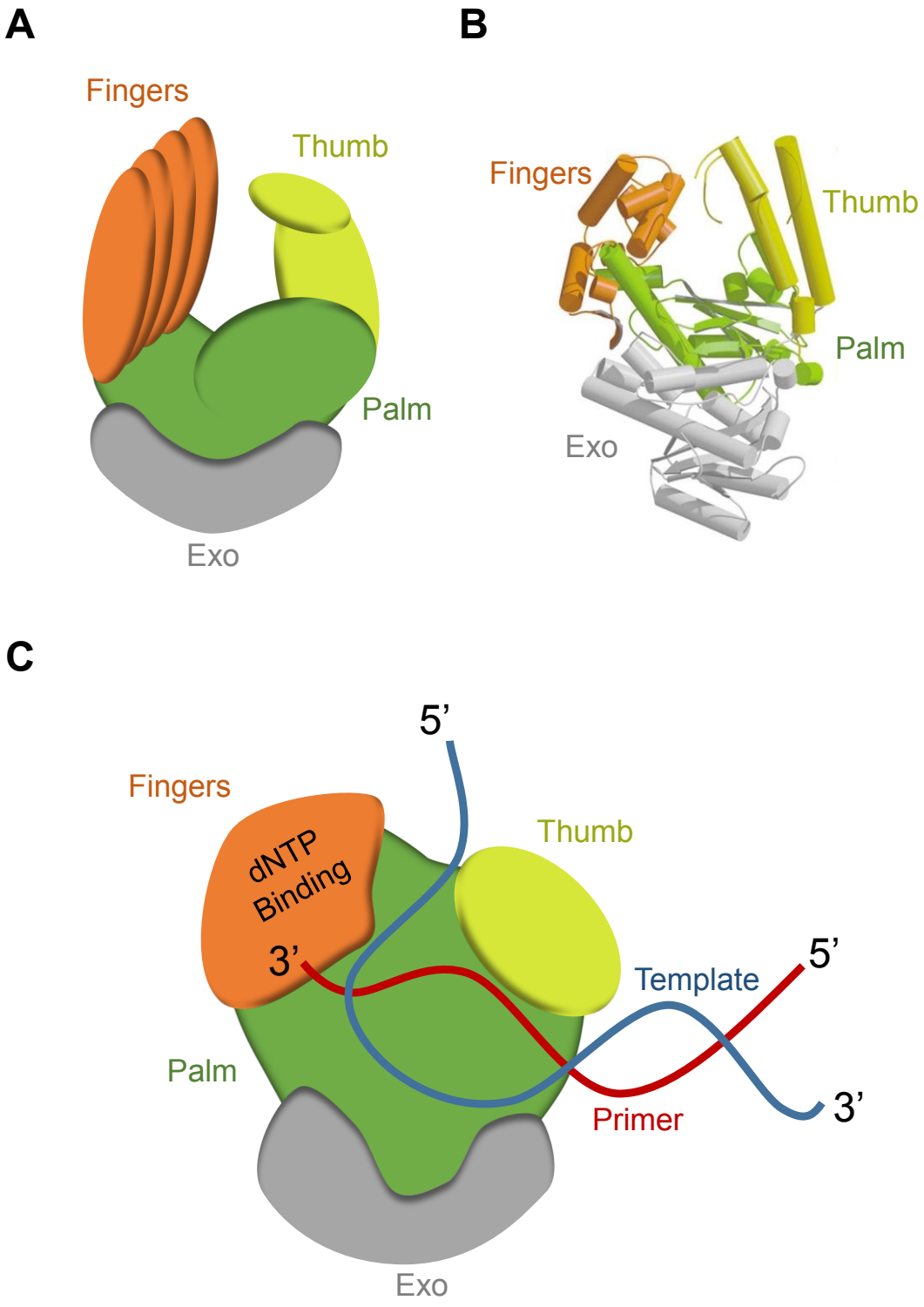


Figure 1.3. The canonical right-hand structure of DNA polymerases.
(A) Cartoon schematic of the canonical DNA right-hand polymerase structure displaying the fingers (orange), thumb (yellow), palm (green), and exonuclease (grey, not present in all polymerases) domains. **(B)** The structure of the first DNA polymerase to be solved, *E. coli* DNA Pol I Klenow fragment (Ollis et al., 1985), showing the resemblance to a right-hand, with the corresponding domains indicated. Figure taken from Beard and Wilson, 2001. **(C)** Schematic of a DNA polymerase in combination with DNA. The DNA sits in the U-shaped cleft of the enzyme, the thumb domain coordinates the DNA template, whilst the fingers domain interacts with the incoming dNTP. The catalytic residues are located in the palm domain.

1.2.3. The Eukaryotic Nuclear Replicative DNA Polymerases

Despite having a plethora of DNA polymerases at their disposal, eukaryotes employ just three of these for the bulk of DNA synthesis during nuclear replication (Table 1.1.) (Hübscher et al., 2002; McCulloch and Kunkel, 2008). These multi-subunit enzymes, Pols α, δ, and ε, all belong to the B family (Lujan et al., 2016). All B family polymerases possess five distinct subdomains, these include an exonuclease domain and an N-terminal domain (NTD), in addition to the familiar fingers, thumb, and palm domains (Franklin et al., 2001; Xia and Konigsberg, 2014, p. 69). Whilst the exonuclease domain displays 3'-5' proofreading activity, used to remove misincorporated nucleotides, the NTD is devoid of catalytic activity, although may play a role in increasing stability and fidelity (ratio of correct over incorrect nucleotide incorporation) (Li et al., 2010; Prindle et al., 2013). The eukaryotic replicative polymerases exhibit a number of distinct features which make them particularly suited to accurate and efficient DNA synthesis (Burgers, 2009; Kunkel and Burgers, 2008). These include large thumb subdomains and tight active sites, in addition to the 3'-5' exonuclease domains, which together act to increase the processivity (the number of nucleotides incorporated in a single binding event) and fidelity of the enzymes.

The first eukaryotic polymerase to be identified, Pol α, is also the first of these polymerases to perform DNA synthesis during replication (Friedberg, 2006; Lujan et al., 2016). The polymerase activity of this enzyme is coupled to the replicative primase, together they comprise a heterotetrameric complex containing: Prim1, Prim2, PolA1, and PolA2 (Muzi-Falconi et al., 2003) (Table 1.1.). During the initiation of replication, Prim1 is responsible for the synthesis of a short RNA primer (7-12 ribonucleotides long), which is subsequently extended by PolA1 to generate a 30-35 nucleotide primer with an RNA 5' end and DNA 3' end (Garg and Burgers, 2005b; Johansson and Dixon, 2013). The two remaining subunits, Prim2 and PolA2, associate with the primase and polymerase subunits, respectively, and are responsible for stabilisation and regulation of catalytic activity. Structural studies of the catalytic PolA1 subunit in three states (apo, binary, and ternary) have permitted insight into the conformational changes of Pol α during synthesis (Perera et al., 2013). These studies reveal that the thumb subdomain makes multiple contacts with the RNA primer through hydrophobic and polar interactions. Interestingly, it was suggested that Pol α may recognise the shape of the RNA-DNA duplex, due to it being in the A-form, rather than typical B-form of DNA usually found *in vivo*. Extension of the RNA primer by Pol α is limited to 10-12 nucleotides, which is equal to one turn of the helix. This has led to speculation that synthesis may be terminated by Pol α upon loss of specific contacts between the thumb subdomain and RNA primer, thereby

Polymerase	Subunit Composition (kDa)	Subunit Function	Polymerase Function	Fidelity (x 10^{-5})	
				Sub.	Indel
Pol α	180 / PolA1	Catalytic subunit (Pol)	Primer synthesis and extension, initiation of leading and lagging strand replication	9.6	3.1
	68 / PolA2	Structural, Interactions			
	55 / Prim2	Accessory (Primase)			
	48 / Prim 2	Catalytic subunit (Primase)			
Pol δ	125 / PolD1	Catalytic subunit	Bulk lagging strand replication, extension of primers on leading strand	1.3 + Exo: < 1.3	5.7 + Exo: 1.3
	66 / PolD3	Multimerisation, PCNA Interaction			
	50 / PolD2	Structural, Interactions			
	12 / PolD4	Regulation, TLS?			
Pol ε	261 / PolE1	Catalytic subunit	Bulk leading strand replication	24 + Exo: < 0.2	5.6 + Exo: 0.05
	59 / PolE2	Multimerisation			
	17 / PolE3	Structural, Interactions			
	12 / PolE4	Structural, Interactions			

Table 1.1. The subunit composition, role, and fidelity of the human replicative DNA polymerases.
The subunit composition and function, in addition to the polymerase function and fidelity of the three replicative nuclear DNA polymerases, Pol α (blue), Pol δ (red), and Pol ε (green), are displayed. For Pol δ and Pol ε the polymerase fidelity in both the absence and presence (+ Exo) of proofreading exonuclease activity is given. Information obtained from Hübscher et al., 2002; McCulloch and Kunkel, 2008.

triggering disassociation of the enzyme and handover to Pol δ (Perera et al., 2013). However, a more recent crystallographic study of Pol α in complex with a DNA:DNA helix shows that contacts previously observed with the RNA:DNA duplex are preserved, suggesting that primer termination is not due to loss of specific contacts upon extension of the RNA primer (Coloma et al., 2016). Rather, this report identifies that the DNA duplex in contact with the enzyme is in the A-B form and the energetic cost of distorting B-DNA to A-B DNA may actually be the factor leading to termination of primer extension.

In addition to well characterised conformational changes to the finger and thumb subdomains, the palm subdomain of Pol α also undergoes structural rearrangements during primer extension (Perera et al., 2013). It is thought that these changes permit the enzyme to translocate at and beyond the RNA/DNA duplex during synthesis. In contrast to Pols δ and ε, Pol α lacks exonuclease activity due to the loss of four critical carboxylate groups and a β-hairpin motif usually found in B-family polymerases (Hogg et al., 2007). Given that Pol α is responsible for the initiation of Okazaki fragment synthesis, and each Okazaki fragment is only ~ 165 nucleotides in length, the enzyme contributes significantly to the total amount of DNA synthesis during replication (Johansson and Dixon, 2013). This significant contribution to replication, coupled with the enzyme's lack of proofreading capabilities, necessitates compensation by Pol δ and the mismatch repair system to remove and correct errors created by the enzyme (McElhinny et al., 2010; Pavlov et al., 2006).

Pol δ was identified as the third mammalian DNA polymerase in 1976, following the discoveries of Pol α in 1957 and Pol β in 1971 (Byrnes et al., 1976; Friedberg, 2006). We now know that the enzyme is responsible for extending the primers synthesised by Pol α, and therefore performs the majority of lagging strand synthesis (Lujan et al., 2016). Intriguing new evidence suggests that Pol δ may also perform this role on the leading strand following initiation of replication, before handing over to Pol ε (see section 1.3.4.) (Daigaku et al., 2015; Yeeles et al., 2017). In humans, Pol δ is composed of four subunits: p125/PolD1, p50/PolD2, p66/PolD3, and p12/PolD4. Here, p125 is the catalytic subunit, whilst p50 serves as a scaffold to mediate interactions with the p66 and p12 subunits (Johansson and Dixon, 2013) (Table 1.1.). Pol δ is adapted to its role as the lagging strand polymerase by displaying extremely high fidelity. The enzyme exhibits an error frequency of 1 in 22,000 incorporated nucleotides, with proofreading activity increasing this fidelity by one to two orders of magnitude (McCulloch and Kunkel, 2008; Prindle et al., 2013; Schmitt et al., 2009). This proofreading activity, enabled by the exonuclease domain, is able to sense mismatched base pairs through the loss of specific hydrogen bond contacts at the N3 and O2 positions of purines and pyrimidines, respectively

(Doublié and Zahn, 2014). These contacts extend up to five base-pairs post-insertion, greatly contributing to the enzyme's fidelity (Swan et al., 2009). Upon loss of the hydrogen bond contacts, a critical β-hairpin segment in the exonuclease domain mediates switching from the polymerase to proofreading mode, which consequently facilitates removal of the incorrectly base-paired nucleotide (Doublié and Zahn, 2014). The importance of this activity for both Pol δ and Pol ε is highlighted by the observation that most mutations of these enzymes involved in cancer development are located in the exonuclease domain (Church et al., 2013; Henninger and Pursell, 2014).

In addition to high fidelity, Pol δ is also suited to bulk DNA replication by its processive nature. In comparison to Pol α, which only extends RNA primers by 10-12 nucleotides, Pol δ from *Saccharomyces cerevisiae* has been shown to incorporate over 7000 nucleotides per binding event *in vitro* (Burgers, 1991). This processivity is partly due to the structural features of the enzyme, such as the large thumb subdomain which mediates contact with the DNA template (Rothwell and Waksman, 2005; Sale et al., 2012). However, processivity is also greatly influenced by interactions with accessory proteins, such as Proliferating Cell Nuclear Antigen (PCNA) and Replication Factor C (RFC). These two proteins, known as the sliding clamp and clamp loader, respectively, form a complex with the Pol δ holoenzyme and greatly increase its processivity (Hashimoto et al., 2003). Here, PCNA encircles the DNA template, after being loaded by RFC, and interacts with Pol δ, thus helping to tether the enzyme to the DNA and prevent dissociation.

In 1989, Pol ε became the fourth nuclear DNA polymerase to be identified in mammalian cells (Friedberg, 2006). Since its discovery, significant evidence has amounted to support a role for the enzyme as the predominant leading strand polymerase (Lujan et al., 2016). Pol ε is a heterotetrameric complex, which in humans consists of: PolE1, PolE2, PolE3, and PolE4, in yeast these are labelled Pol2, Dpb2, Dpb3, and Dpb4. Here, PolE1 or Pol2, is the catalytic centre, possessing both polymerase and exonuclease activity, with the other three subunits performing accessory roles, including interacting with the ssDNA template to increase processivity (Table 1.1.) (Aksenova et al., 2010; Tsubota et al., 2003). Notably, unlike Pol δ, Pol ε does not require interactions with PCNA to achieve high levels of processivity. Structural studies of yeast Pol ε have revealed two key differences between the enzyme and typical B-family polymerases (Hogg et al., 2014). Firstly, Pol ε possesses a new functional domain not previously observed in B-family polymerases. This domain, termed the P domain, is formed from two large insertions in the palm domain of the enzyme, which form an elongated structure that extends out of the palm and towards the dsDNA. This allows the enzyme to grip the

newly synthesised DNA as it leaves the active site and is likely responsible for its high intrinsic processivity (Hogg et al., 2014). Secondly, Pol ε lacks the extended β-hairpin loop found in other B-family polymerases, which is important for switching to the exonuclease active site. Here, it is thought that the novel P domain compensates for this short β-hairpin loop by maintaining close contact between the polymerase and DNA whilst switching active sites (Hogg et al., 2014). Clearly, this compensation is sufficient, given that Pol ε is reported to be the most accurate of all yeast DNA polymerases (Fortune et al., 2005).

1.3. DNA Replication

Accurate and efficient DNA replication, prior to cell division, is dependent not only on the appropriate functioning of DNA polymerases, but of the replisome as a whole which together unpacks, unwinds, and duplicates the genome. Failure to perform this task appropriately risks genomic instability. Consequently, a discrete phase of the cell cycle is dedicated to the process, this is termed the DNA synthesis (S)-phase, which is coordinated to ensure complete genome duplication prior to mitosis and cytokinesis. DNA replication can be broadly partitioned into three main stages: initiation, elongation, and termination. After a discussion of DNA replication origins, each of these stages will be described drawing on recent insights from benchmark *in vitro* reconstitution studies (Devbhandari et al., 2017; Kurat et al., 2017; Yeeles et al., 2017).

1.3.1. DNA Replication Origins and Their Licensing

It has been estimated that a single bi-directional replication fork propagated from a single origin of replication would take more than 40 days to copy just human chromosome 1 (MacAlpine, 2016). In 1966, Huberman and Riggs noted that in order to obviate this issue and complete DNA replication within S-phase, multiple origins of replication must be dispersed across each chromosome (Huberman and Riggs, 1966).

In *S. cerevisiae,* replication origins are denoted by short autonomously replicating sequences (ARSs) which contain a consensus sequence (Leonard and Méchali, 2013). This sequence is bound by the origin recognition complex (ORC), a conserved heterohexameric replication initiation factor (Bell and Kaguni, 2013). However, in higher eukaryotes ORC does not display any sequence specificity, leading to speculation that alternative chromatin features, rather than consensus sequences, are required to specify origins (Remus et al., 2004; Vashee et al., 2003). Indeed, studies in both *Drosophila* and human cells have found that ORC localises in a sequence-independent manner to open chromatin and is frequently associated with promoters and enhancers, in addition to

regions occupied by transcription factors (MacAlpine et al., 2010; Miotto et al., 2016). In conjunction with this, recent reconstitution studies using purified yeast replication proteins and chromatinised templates, found that nucleosomes are able to suppress non-specific ORC binding (Kurat et al., 2017). The authors suggest that this phenomenon may also limit ORC binding to only nucleosome free regions (NFRs) in mammalian cells (Figure 1.4.A.). During S-phase mammalian genomes are divided into temporally distinct early- and late-replicating regions, depending on the density of replication origins and the timing of activation (Rhind and Gilbert, 2013). A recent study identified that ORC density may play a role in determining replication timing (Miotto et al., 2016). Remarkably, the authors were able to accurately model and predict replication timing based solely on the density of ORC binding sites. Interestingly, a lack of ORC binding was observed at common fragile sites, areas of chromosomes prone to breakage, demonstrating the significance of replication origins on genome integrity.

Despite, the lack of a detailed understanding of the precise chromatin features governing origin selection and activation, much is known about events following ORC binding. These events lead to the assembly of the pre-replication complex (pre-RC) in a process known as origin licensing (Siddiqui et al., 2013). Importantly, origin licensing is strictly regulated to ensure that all origins are licensed prior to the initiation of DNA replication and therefore before entering S-phase. In order to achieve this, origin licensing only occurs during late mitosis and G_1-phase of the cell cycle. In *S. cerevisiae*, this regulation is instilled by the action of cyclin dependent kinase (CDK), which phosphorylates a number of initiation proteins to prevent pre-RC assembly (Figure 1.4.A.) (Siddiqui et al., 2013). Upon exit from mitosis, CDK activity drops, removing this inhibition and permitting origin licensing. During G1-phase, ORC bound to replication origins recruits two additional factors, Cell Division Cycle 6 (CDC6) and CDC10 Target 1 (CDT1) (Remus et al., 2009). Together, these three licensing factors direct loading of the mini-chromosome maintenance (MCM) complex, MCM2-7, the replicative helicase, around dsDNA as an inactive double hexamer (Figure 1.4.B.) (Siddiqui et al., 2013). This completes the formation of the pre-RC, which subsequently requires the recruitment of other replication factors to form the pre-initiation complex (pre-IC) before DNA replication can proceed.

1.3.2. Formation of the Pre-Initiation Complex

The formation of the pre-RC during G_1-phase is the first step in preparing origins for firing. However, a second step is required before the initiation of replication can occur. Here, recruitment of additional factors just before and during S-phase forms the pre-IC, leading to helicase activation and the recruitment of the replicative primase and

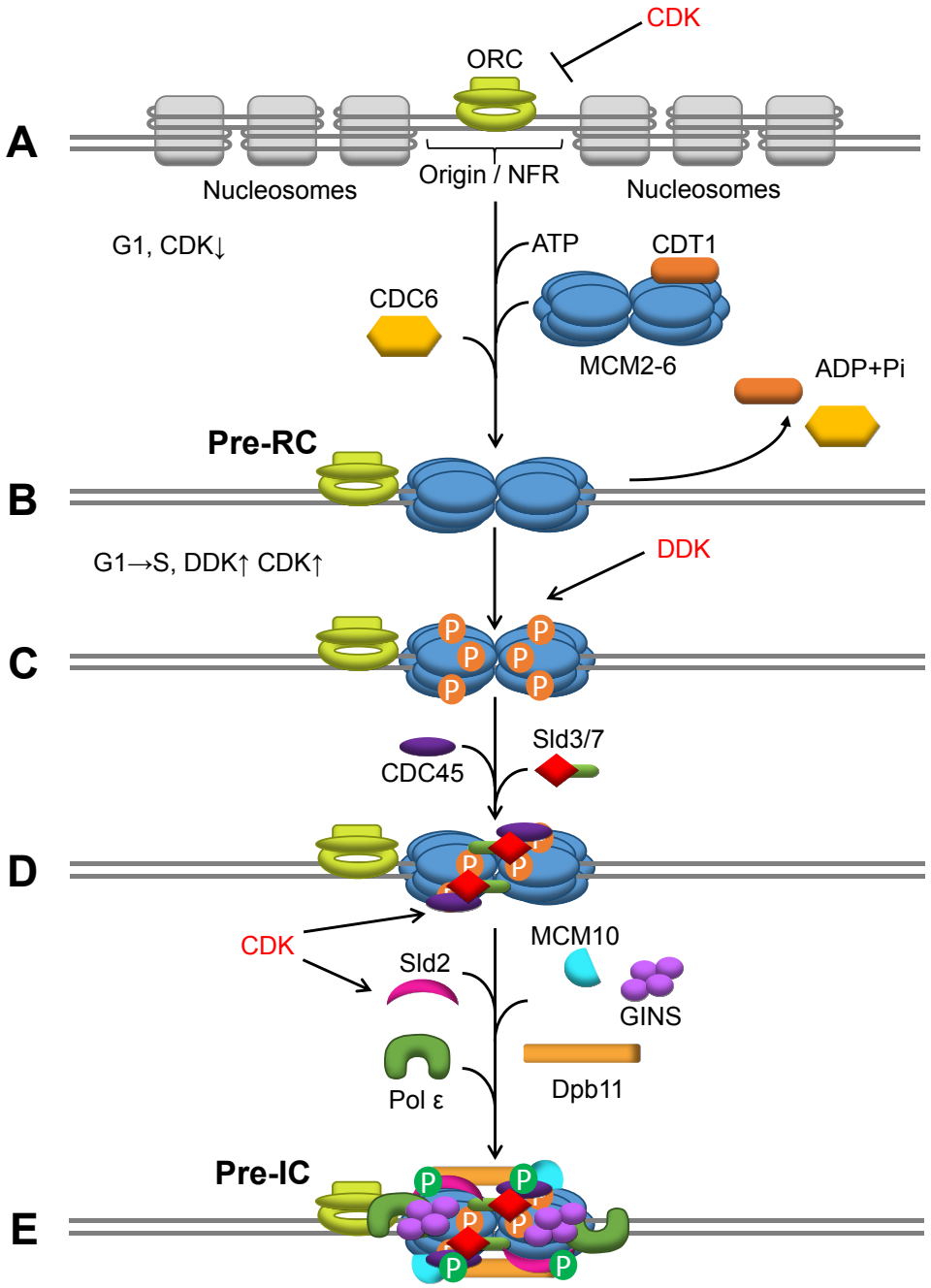


Figure 1.4. Formation of the Pre-RC and Pre-IC at DNA replication origins.
(A) In higher eukaryotes ORC binds to nucleosome free regions (NFRs) in the genome to mark origins of replication. High CDK activity phosphorylates a number of key initiation factors to inhibit Pre-RC assembly during mitosis. **(B)** Upon entry into G1-phase, CDK activity drops, permitting Pre-RC assembly. ORC recruits CDC6 and CDT1 which direct loading of MCM2-6 onto dsDNA as an inactive double hexamer in an ATP-dependent fashion, forming the Pre-RC. **(C)** In late G1 and early S-phase CDK and DDK activity increases. DDK phosphorylates MCM2, 4, and 6. **(D)** Phosphorylation of MCM stimulates recruitment of Sld3/7 and CDC45. **(E)** CDK phosphorylation of Sld2 and 3, further stimulates recruitment of Sld2, Dpb11, GINS, Pol ε, and MCM10. Together, these factors form the Pre-IC, with MCM, GINS, and CDC45 forming the CMG complex. DNA replication is initiated following unwinding of DNA by CMG. Information obtained from Yeeles et al., 2015.

polymerases (Tanaka and Araki, 2013). The model for the events leading up to and during origin firing, including the factors and regulatory mechanisms involved, was recently confirmed following the reconstitution of replication initiation using sixteen purified yeast replication proteins (Yeeles et al., 2015). Here, it was determined that pre-IC formation is regulated by the activity of CDK and Dbf4-Dependent Kinase (DDK). Immediately prior to S-phase, CDK and DDK levels rise, this prevents additional MCM loading and leads to the phosphorylation of a number of key replication factors. Firstly, DDK phosphorylates MCM2, 4, and 6, stimulating recruitment of synthetic lethality with dpb11 (Sld) 3/7 and CDC45 (Figure 1.4.C. and D.) (Deegan et al., 2016). The remaining firing factors, Sld2, DNA polymerase-binding protein 11 (Dpb11), GINS, Pol ε, and MCM10, are then recruited following CDK phosphorylation of Sld3 (Figure 1.4.E.) (Parker et al., 2017). Upon recruitment, CDC45 and GINS form a complex with MCM2-7, known as the CMG complex. Notably, Pol ε also plays a non-catalytic, but critical, role in the assembly of this complex and is recruited to the CMG complex through a direct interaction with GINS (Georgescu et al., 2014). Although not required for stable complex formation, MCM10 is essential for activation of the CMG complex (Yeeles et al., 2015). Thus, this list represents the minimum set of factors required for pre-IC formation and CMG activation. Subsequently, each CMG complex is remodelled, forming a single-stranded 3'-5' DNA translocase responsible for unwinding the double helix, with each of the two MCM hexamers translocating in opposite directions (Figure 1.5.A.) (Fu et al., 2011). Following origin firing in yeast, re-replication is prevented by CDK, which promotes the degradation of CDC6 and nuclear exclusion of CDT1 (Parker et al., 2017). In humans and other higher eukaryotes, Geminin prevents re-replication by binding to CDT1 at origins already licenced (Lutzmann et al., 2006). Importantly, not all licenced origins fire during replication, leaving behind dormant origins which may play a role in DNA damage tolerance (see section 1.5.1.).

1.3.3. The Initiation of DNA Replication

Following the formation and activation of the pre-IC, MCM2-7 unwinds the DNA duplex, forming a bubble of ssDNA. This ssDNA is rapidly bound by Replication Protein A (RPA), a heterotrimeric single-strand binding protein (SSB) which prevents re-annealing, the generation of secondary structures, and degradation of DNA. A number of additional proteins are also recruited, which together form the replication progression complex (RPC). These include Chromosome transmission fidelity protein 4 (Ctf4), Pol α, Chromosome segregation in meiosis 3 (Csm3), Topoisomerase 1-associated factor 1 (Tof1), Mediator of replication checkpoint protein 1 (Mrc1), Facilitates chromatin

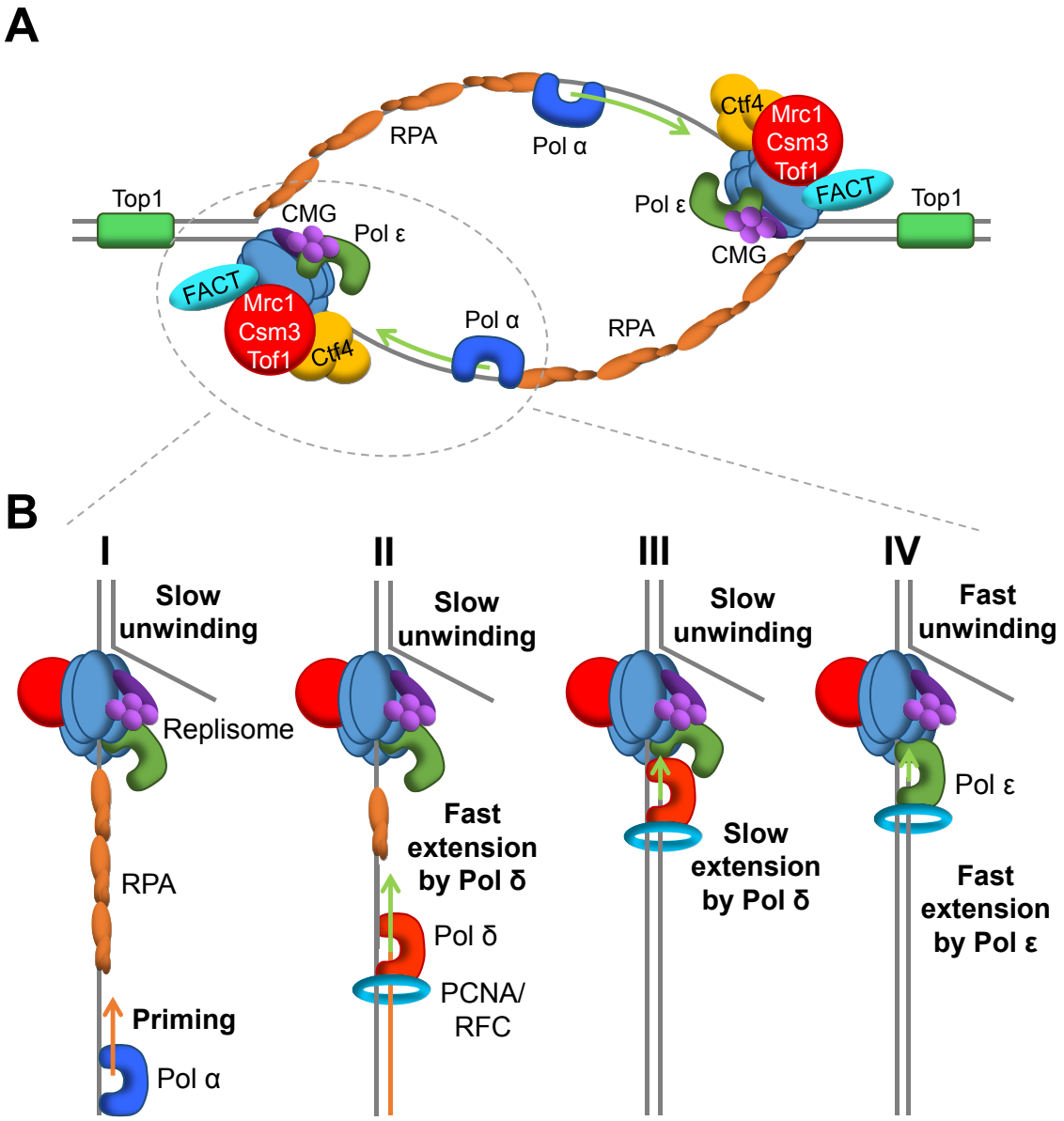


Figure 1.5. The initiation of leading strand replication.
(A) Following Pre-IC formation, the parental DNA duplex is unwound by MCM. This produces ssDNA which is bound by RPA. Pol α, Ctf4, Csm3, Tof1, Mrc1, FACT, and Top1 are recruited to form the replication progression complex. Once recruited, Pol α initiates DNA synthesis on the ssDNA template. **(B)** Initiation of leading strand replication. (I) Slow unwinding of the DNA duplex by MCM permits recruitment and primer synthesis by Pol α on the leading strand template. (II) RFC loads PCNA onto the primer-template, displacing Pol α and stimulating Pol δ recruitment. Pol δ rapidly extends the primer. (III) Rapid synthesis continues until Pol δ reaches the advancing replisome, here synthesis slows as it is limited by the rate of unwinding by MCM. (IV) The nascent DNA chain and PCNA are transferred to Pol ε, which stimulates DNA unwinding and subsequently performs the bulk of leading strand synthesis. Figure B adapted from Yeeles et al., 2017.

transcription (FACT), and DNA topoisomerase I (Top1) (Figure 1.5.A) (Gambus et al., 2006).

Again through reconstitution studies, it was recently demonstrated that Mrc1 (Claspin in metazoans) and Csm3/Tof1 are essential for the establishment and maintenance of *in vivo* replication rates (Yeeles et al., 2017). In conjunction, these studies identified that FACT is critical for replication of chromatinised templates, but dispensable when replicating naked DNA (Kurat et al., 2017). Both MCM10 and Ctf4 have been proposed to link Pol α to CMG, however in reconstitution studies on naked DNA, Pol α functioned distributively, even in the presence of these two factors (Gambus et al., 2009; Ricke and Bielinsky, 2004; Yeeles et al., 2017). Interestingly, experiments using chromatinised templates identified that lagging-strand products became consistent over a range of Pol α concentrations, potentially suggesting that the polymerase is physically tethered to the replisome by interactions with FACT and nucleosomes (Kurat et al., 2017). Alternatively, regular priming may be enforced by pausing of Pol δ at each nucleosome (Devbhandari et al., 2017; Kurat et al., 2017). Therefore, it is also possible that Pol α is simply recruited to the RPA bound ssDNA exposed by MCM unwinding, indeed the enzyme has been previously shown to share an interaction with RPA (Dornreiter et al., 1992).

Regardless of the precise mechanism which localises Pol α to the exposed ssDNA, once in contact the primase subunit is able to facilitate RNA primer synthesis, before extension by the polymerase subunit (Figure 1.5.A.). Consequently, this provides the 3’ hydroxyl utilised for further elongation by the replicative polymerases. In order for these polymerases to take over synthesis a polymerase switch mechanism is required. This occurs in two steps, firstly Pol α is competed from RPA by RFC. RFC is then able to load PCNA, which is subsequently bound by Pol δ (Yuzhakov et al., 1999). This mechanism, which further acts to restrict Pol α activity to the generation of short primers, was recently confirmed in reconstitution experiments (Devbhandari et al., 2017).

1.3.4. Elongation and the Progression of the Replisome

The switch from Pol α to the replicative DNA polymerases δ and ε, marks the transition from the initiation to elongation phase of replication. Notably, however, Pol α activity is still critical to prime Okazaki fragment synthesis throughout replication on the lagging strand. The prevailing model for polymerase usage during the elongation phase of DNA replication, is that Pol ε is responsible for leading strand replication, whilst Pol δ synthesises Okazaki fragments on the lagging strand, which involves extension and displacement of the primers synthesised by Pol α (Figure 1.6.A.) (Kunkel and Burgers, 2008). Indeed, accumulating evidence over the past decade has strongly supported this

model (Kunkel and Burgers, 2008). However, the Prakash laboratory recently challenged this view, presenting evidence to suggest that Pol δ is responsible for synthesis of both the leading and lagging strands, with Pol ε's primary role being the proofreading of errors made by Pol δ on the leading strand (Johnson et al., 2015). This model has now been largely discredited due to concerns over experimental procedures and the observation that Pol ε is unable to proofread errors made by Pol δ (Burgers et al., 2016; Flood et al., 2015).

Whilst recent evidence generally supports the model that Pol ε and Pol δ perform leading and lagging strand synthesis, respectively, it also suggests that this model is not as simple as first thought. Polymerase usage sequencing (Pu-seq) studies in yeast, which use high-throughput sequencing to map ribonucleotides incorporated in the genome by Pol ε and Pol δ mutants during replication, have shed light on the genome wide subtleties of polymerase usage during this process (Clausen et al., 2015; Daigaku et al., 2015; Koh et al., 2015; Reijns et al., 2015). Here, it was found that Pol ε and Pol δ are consistently responsible for leading and lagging strand synthesis, respectively. However, there was a strong bias towards Pol δ synthesis on the leading strand close to replication origins (Daigaku et al., 2015). Recent replisome reconstitution studies by Yeeles *et al.* have now identified the role of Pol δ in leading strand replication following primer synthesis (Figure 1.5.B.). In this report, Pol δ was required for efficient and uniform initiation of leading strand synthesis (Yeeles et al., 2017). Yeeles *et al.* propose a model whereby Pol δ and PCNA take over the 3' end of the primer synthesised by Pol α on the leading strand (Figure 1.5.B. I and II). Initially, as CMG-Pol ε moves away from the primer, synthesis by Pol δ will be rapid, until it catches up with the slow moving CMG (Figure 1.5.B. II). At this point Pol δ will slow down due to an inability to accelerate unwinding by CMG (Figure 1.5.B. III). Here, a polymerase switch occurs in which the 3' end and PCNA are transferred over to Pol ε, which can stimulate CMG unwinding, consequently promoting maximal leading strand replication rates for continued synthesis (Figure 1.5.B. IV). Mrc1 also stimulates unwinding by CMG during replisome progression, with Csm3/Tof1 acting to promote functional association of Mrc1 with the replisome. Interestingly, in conjunction, it was discovered that PCNA plays an important role in promoting leading, as well as lagging, strand replication (Yeeles et al., 2017). This suggests that Pol ε utilises both CMG and PCNA as processivity factors during synthesis. It is likely that CMG tethers the polymerase to the unwinding fork while PCNA promotes continued association with the 3' end of the leading strand (Figure 1.5.B. IV).

On the lagging strand, discontinuous synthesis of Okazaki fragments by the concerted activities of Pol α and Pol δ continues as the replisome progresses (Figure 1.6.A. and

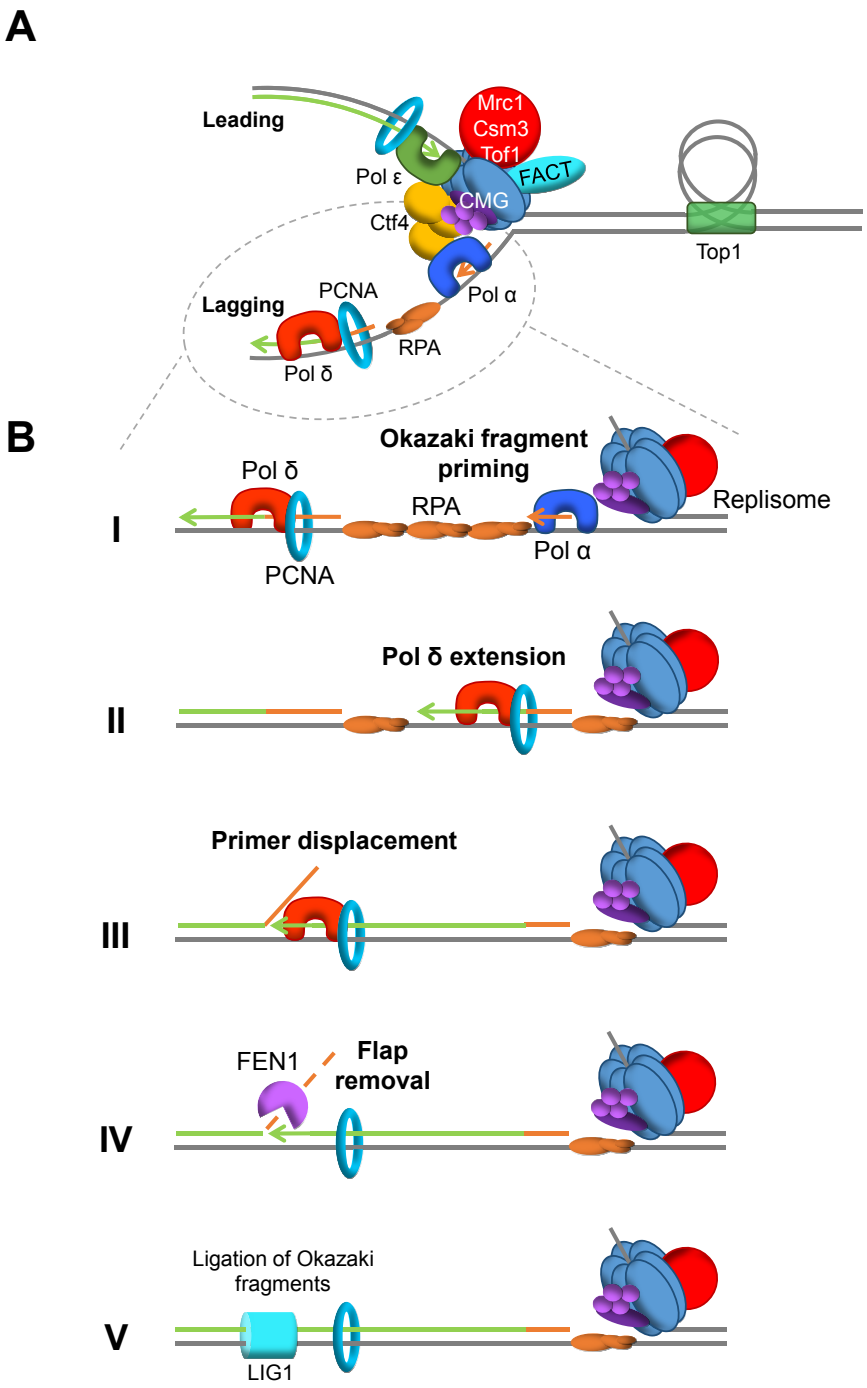


Figure 1.6. DNA replication elongation and lagging strand synthesis.
(A) The elongation phase of DNA replication. Following initiation, the replicative DNA polymerases ε and δ perform the bulk of DNA synthesis on the leading and lagging strands, respectively. Synthesis on the leading strand is predominantly continuous, whereas lagging strand synthesis is discontinuous and requires constant priming by Pol α. Mrc1, Csm3, and Tof1 are required for the maintenance of maximal fork speeds, while FACT is required for nucleosome processing. Top1 works ahead of the replisome to relieve torsional stress caused by unwinding of the DNA duplex. **(B)** Lagging strand DNA synthesis. (I) Discontinuous generation of Okazaki fragments is initiated by Pol α mediated primer synthesis. (II) Pol δ and PCNA replace Pol α and extend the Okazaki fragment primer, displacing RPA from the ssDNA template. (III) Extension continues until Pol δ reaches the primer of the preceding Okazaki fragment downstream, the enzyme displaces the RNA primer generating a flap. (IV) The flap is removed by FEN1, (IV) and the two Okazaki fragments are ligated together by LIG1.

B.), these fragments must be ligated together to produce an intact daughter lagging strand in a process known as Okazaki fragment maturation. Here Pol δ displaces the RNA primer of the previous Okazaki fragment, producing a 5' flap which is subsequently cleaved by flap endonuclease 1 (FEN1) (Maga et al., 2001). The two Okazaki fragments are then ligated together by DNA ligase I (Figure 1.6.B.)(Barnes et al., 1990; Devbhandari et al., 2017; Lehman, 1974; Li et al., 1995; Pascal et al., 2004).

Due to the double helical nature of DNA, unwinding of the two strands during CMG progression causes torsion to build up ahead of the replisome, resulting in positive supercoiling and topological stress (Keszthelyi et al., 2016). In order to release this torsion, specialised enzymes known as topoisomerases are required. During replication in eukaryotes, two such enzymes, Top1 and Top2, are employed to resolve topological stress. Top1 is a type IB topoisomerase which functions by nicking one of the DNA strands, while Top2, a type II topoisomerase, breaks both strands (Keszthelyi et al., 2016; Wang, 1996, 1998). In each case, this activity allows topological stress to be released before re-ligation of the nicked strands. This is critical to relax the positive supercoiling produced ahead of the fork during replication and thus facilitate the continued progression of the replisome (Figure 1.6.A.) (Brill et al., 1987; Hiasa and Marians, 1994a, 1994b; Kim and Wang, 1989). Aside from acting ahead of the replisome, topoisomerases can also function behind the fork in a second pathway. Here the replication fork rotates relative to the unwound template DNA ahead, thereby transferring the DNA intertwines to the region behind the replisome (Keszthelyi et al., 2016). This transfer causes the replicated sister-chromatids to become intertwined, forming pre-catenanes, which can progress to full catenanes upon completion of replication. Pre-catenanes are removed by the action of type II topoisomerases behind the progressing replisome, allowing topological stress to be relieved when template DNA ahead of the fork is not accessible (Baxter, 2015). This pathway is of particular importance upon the convergence of two adjacent replisomes during the termination of replication (Keszthelyi et al., 2016).

1.3.5. The Termination of DNA Replication

DNA replication terminates when adjacent replication forks moving towards each other meet. The location of termination is largely dependent upon the site of initiation and the rate of replication, rather than specific DNA sequences or chromatin features (Greenfeder and Newlon, 1992). There are four main processes required to efficiently terminate replication. These include; the unwinding of the final portion of DNA between

the two replisomes, the filling in and ligation of daughter strands, the removal of dsDNA catenanes, and the disassembly of the replisome (Dewar et al., 2015).

As replisomes converge during termination, topological stress builds up in the unwound DNA in between. However, this DNA is thought to be largely inaccessible to topoisomerases due to steric exclusion by the two replisomes (Keszthelyi et al., 2016). In order to proceed, fork rotation is employed, allowing the topological stress to be transferred behind the fork and unwinding of the final unreplicated DNA to occur (Sundin and Varshavsky, 1981). This rotation generates catenanes which are subsequently removed by Top2. Recent work in *Xenopus* extracts has revealed the mechanism which ensures complete replication at sites of termination in vertebrates (Dewar et al., 2015). Here it was determined that as replisomes converge on opposite strands, due to the CMGs encircling the leading strand, they efficiently pass each other without stalling. This allows the remaining gaps in the daughter strands to be filled in before the CMG contacts the 5' end of the opposing fork's lagging strand. At this point the complex passes over the ssDNA-dsDNA junction and moves onto the dsDNA. Ligation of the daughter DNA strands then occurs and at the same catenanes are resolved. Lastly, the dsDNA bound CMG is ubiquitylated on MCM7, triggering removal by the ATPase p97 (Maric et al., 2014; Moreno et al., 2014).

1.3.6. Mitochondrial DNA Replication

In addition to the nuclear genome, eukaryotes also possess a second "genome" in the mitochondria. In mammals, around 1000 copies of this 16.5 Kb circular DNA molecule are housed in the mitochondria (Wilson et al., 1985). The mitochondrial replisome is distinct from the nuclear DNA replication machinery, and includes Pol γ, Twinkle DNA helicase, mitochondrial single-strand binding protein (mtSSB), mitochondrial RNA polymerase (PolRMT), topoisomerases, mitochondrial transcription factor A (TFAM), DNA ligase III, RNase H1, and RNase mitochondrial RNA processing enzyme (RNase MRP) (Cerritelli et al., 2003; Hance et al., 2005; Simsek et al., 2011; Spelbrink et al., 2001; Tyynismaa et al., 2004; Van Goethem et al., 2001).

The mitochondrial genome has a single origin of replication on each strand. These are termed O_H and O_L on the heavy and light strands respectively, with O_L located two-thirds of the way around the DNA molecular relative to O_H (Falkenberg et al., 2007). Prior to the initiation of DNA synthesis, TFAM, a member of the high-mobility group (HMG) proteins, binds upstream of O_H and unwinds the DNA duplex (Dairaghi et al., 1995; Fisher et al., 1987). This facilitates the recruitment of PolRMT which synthesis an RNA primer at the O_H origin (Chang et al., 1985; Chang and Clayton, 1985). Importantly, this primer

is identical to the one used for the polycistronic transcription of mitochondrial genes and therefore requires processing by the endoribonuclease RNase MRP for replication to occur (Topper and Clayton, 1990). The resulting processed RNA primer can then be utilised by Pol γ for extension (Ropp and Copeland, 1996). During replication, unwinding of the duplex parental DNA is facilitated by the replicative helicase Twinkle which acts as a hexamer without the need for a specialised protein to load it onto the circular mtDNA template (Jemt et al., 2011; Korhonen et al., 2003). Following unwinding, ssDNA is bound by mtSSB, a homotetramer with an analogous function to RPA. Pol γ, Twinkle, and mtSSB form a processive replication machinery and together make up the minimal mitochondrial replisome, capable of generating 16.5 Kb products *in vitro* (Falkenberg et al., 2007).

DNA synthesis at O_L is not initiated until it is exposed by replication of the leading heavy strand. Once in a ssDNA conformation, O_L forms a stem loop structure which facilitates priming by PolRMT (Fusté et al., 2010). Interestingly, this stem loop structure is highly conserved and mutations affecting its structure are under-represented in the mitochondrial genome (Wanrooij et al., 2012). RNase H1 is also thought to play a key role in mtDNA replication, with *RNaseh1*$^{-/-}$ mice displaying significant mitochondrial defects (Cerritelli et al., 2003). Here, RNase H1 is likely involved in the removal of the RNA primers at each origin, likewise DNA ligase III may also be involved in ligation of the resulting nicks.

1.4. Replication Stress

During its progression, the replisome faces numerous obstacles which it must overcome in order to accurately complete genome duplication (summarised in Figure 1.7.). These obstacles arise from a variety of intracellular and extracellular origins, however almost all cause a similar effect; the slowing or stalling of replication fork progression. In this section, the array of sources contributing to replication stress will be described, before focussing more closely on one of these, DNA damage, and the mechanisms employed for its repair.

1.4.1. Sources of Replication Stress

Any factor which leads to the slowing or stalling of the replisome can be considered a source of replication stress (Zeman and Cimprich, 2014). Over recent years, the list of potential sources of replication stress has grown, highlighting the vulnerability of the core replisome to perturbations in its DNA template and the requirement for sufficient resources to maintain progression. In fact, the DNA template itself can act as a source

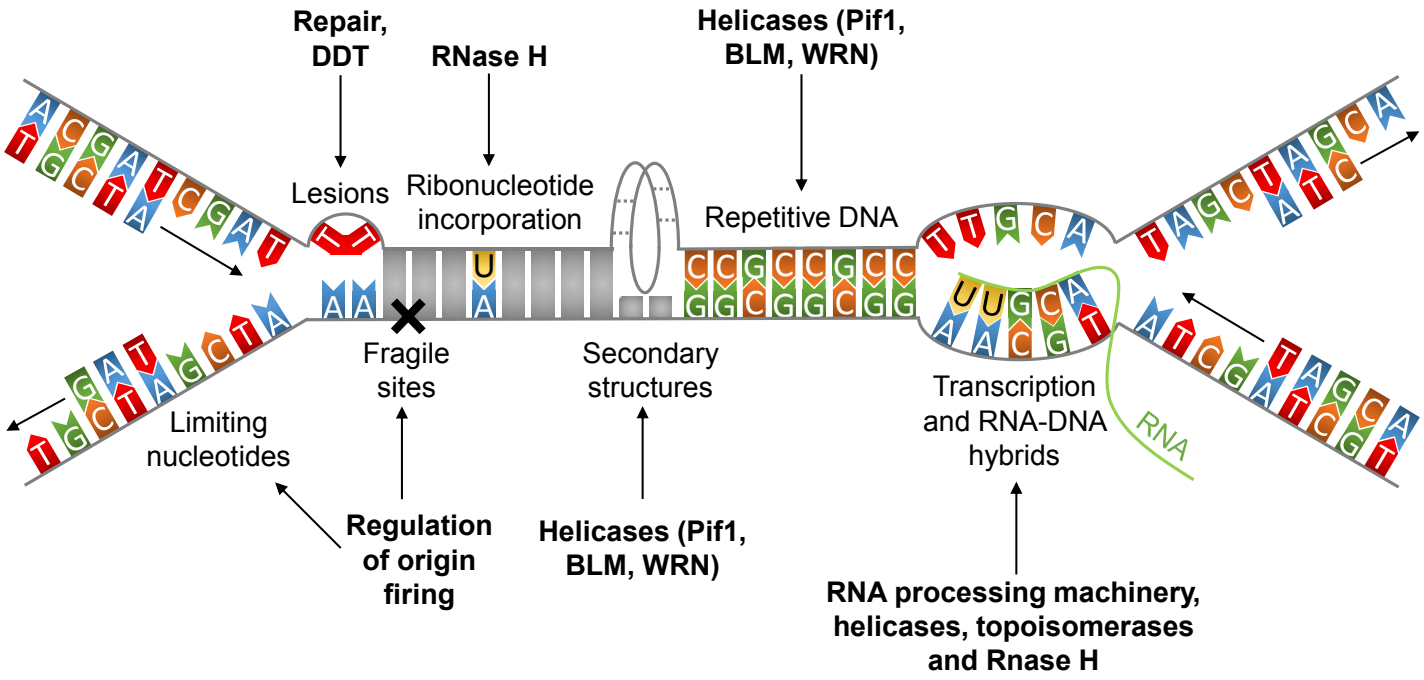


Figure 1.7. The sources of DNA replication stress.
Various sources contribute to replication stress and can impede the completion of genome duplication including, DNA damage, ribonucleotide incorporation, repetitive DNA sequences, limiting pools of nucleotides, fragile sites, DNA secondary structures, and collisions between replication and transcription machinery. Potential pathways and factors involved in remediating each of these sources are highlighted in bold. Figure adapted from Zeman and Cimprich, 2014.

of replication stress. Numerous DNA sequences pose a challenge to the progressing replisome due to their ability to form DNA secondary structures, these include hairpins and triplexes, formed by trinucleotide repeats, which can directly block the replisome or cause slippage (McMurray, 2010). This in turn leads to expansion or contraction of the repeats and can potentially further contribute to replication stress (Kim and Mirkin, 2013). In addition to trinucleotide repeats, G-rich sequences of DNA have the ability to form G-quadruplex structures through hoogsteen base-pairing. These structures can cause the formation of double-strand breaks (DSBs) and deletions when incorrectly replicated and thus require specialised helicases for their unwinding (Bochman et al., 2012; Paeschke et al., 2013). Aside from DNA secondary structures, nicks and gaps present in the template can additionally stall the replisome or contribute to DSB formation. These nicks and gaps are often caused by DNA repair pathways or topoisomerases during processing (Zeman and Cimprich, 2014).

Although extremely accurate when forming correct base-pairs, the replicative polymerases can themselves contribute to replication stress through misincorporation of ribonucleotides (rNTPs). Both Pol δ and Pol ε incorporate rNTPs at a high rate during replication, which must be removed through ribonucleotide excision repair (Dalgaard, 2012). Loss of this pathway is lethal in mammals and causes damage sensitivity in yeast, highlighting the detrimental affect misincorporated rNTPs can have on the cell (Lazzaro et al., 2012; Reijns et al., 2012). The replicative polymerases are stalled upon encountering rNTPs in the template strand and aberrant processing of rNTPs by Top1 can produce non-ligatable nicks, both contributing to replication stress (Kim et al., 2011; Nick McElhinny et al., 2010; Williams et al., 2013).

Protein complexes bound to the DNA template during replication can also hinder fork progression. Perhaps the most significant of these is transcription machinery, due to replication and transcription sharing the same DNA template (Bermejo et al., 2012; Helmrich et al., 2013). Collisions between replication and transcription complexes can cause DSBs, this is a particular problem at early replicating fragile sites. These are highly transcribed regions of the genome which are replicated early in S-phase, thus increasing the chance of collision (Barlow et al., 2013). As well as direct collision, replication and transcription on the same DNA template can generate topological stress. Indeed, this is supported by studies in yeast which found that replication stress is produced by converging replication and transcription machinery even before they collide (Bermejo et al., 2011, 2012). Loss of RNA processing elements can further increase the risk of replication-transcription collision and topological stress by slowing the rate of transcription (Huertas and Aguilera, 2003; Li and Manley, 2005; Paulsen et al., 2009;

Stirling et al., 2012; Wahba et al., 2011). Additionally, this can promote the formation of unresolved R-loop structures, RNA-DNA duplexes with a displaced ssDNA strand, which can also stall progressing replisomes (Aguilera and García-Muse, 2012). RNA processing elements help to prevent this scenario by preventing rehybridisation of the RNA transcript, with specialised helicases assisting in unwinding RNA-DNA duplexes which do occur (Alzu et al., 2012; Yüce and West, 2013). Therefore, loss of these components contributes to replication stress.

In addition to early replicating fragile sites, common fragile sites are prone to causing DSBs even at low levels of replication stress (Debatisse et al., 2012). Interestingly, the replication fork progresses through these sites at a normal rate, suggesting there is no direct block produced by secondary structures or specific template elements (Debatisse et al., 2012). Instead, it is thought that these regions are vulnerable to stress due to the lack of origins, which may impair the ability to overcome 'normal' replication obstacles at these sites (Zeman and Cimprich, 2014). Improper control of replication initiation can also contribute to replication stress. Firing of too many origins can result in depletion of nucleotide pools and thereby slowed fork rates (Beck et al., 2012; Sørensen and Syljuåsen, 2012). Whereas firing of too few origins can cause under-replication and the loss of genetic material (Debatisse et al., 2012; Shima et al., 2007). Likewise, chromatin compaction may contribute to inhibition of replisome progression. In support of this, increased replication-dependent DSBs have been observed in yeast heterochromatin regions and relaxation of chromatin at fragile sites reduces the occurrence of breakage (Jiang et al., 2009; Lambert and Carr, 2013a).

In summary, it is clear that almost all the elements required for efficient replisome progression, such as a pristine DNA template and sufficient dNTPs, can cause replication stress when impaired. Many of these impairments emerge from endogenous processes in the cell, such as transcription, or inherent elements of the genome, for example secondary-structure forming sequences. However, perhaps the most significant and well-characterised source of replication stress arises from DNA damage, which can be generated from both endogenous and exogenous sources.

1.4.2. DNA Damage

Surprisingly, given that it's the carrier of genetic material, DNA is a rather unstable molecule which is liable to damage and decay (Lindahl, 1993). The scale of this problem for the preservation of genomic integrity, is highlighted by the estimation that there are up to 10^5 spontaneous DNA lesions generated per cell every day (Hoeijmakers, 2009). This DNA damage arises from three main sources, these are, spontaneous damage due

to the inherent instability of the DNA molecule, by-products of normal endogenous cellular metabolism, and external physical and chemical mutagens.

Spontaneous damage to DNA occurs in two main ways. The first of these is depurination which is estimated to occur around 9000 times per cell each day (Nakamura et al., 1998). This is caused by hydrolysis of the base-sugar bond, resulting in the loss of the base while the phosphodiester backbone remains intact. Consequently, an apurinic/apyrimidinic (Ap) site is generated which can stall replisome progression and lead to base substitution mutations following replication (Loeb and Preston, 1986). Secondly, spontaneous deamination can cause interconversion between bases, the most common example of this is deamination of cytosine to uracil which can lead to G→A transition mutations upon replication (Duncan and Miller, 1980).

Normal cellular metabolism contributes significantly to DNA damage through the generation of reactive oxygen species (ROS). These include hydrogen peroxide, superoxide anions, and hydroxyl radicals, which are primarily a product of oxidative phosphorylation (OXPHOS) in the mitochondria (Cadet et al., 2003). ROS attack the DNA, generating lesions, the most common of which are 8-oxo-2'-deoxyguanosine (8-oxo-dG) lesions, caused by the oxidation of guanine (Halliwell and Aruoma, 1991). Although 8-oxo-dG lesions can be bypassed by the replisome, they direct misincorporation of adenine in the daughter strand, producing GC→TA transversion mutations. Meanwhile, oxidation of thymine moieties can produce thymine glycol (Tg) lesions which act to stall the replisome (Ghosh et al., 2008).

Lastly, there are numerous exogenous physical and chemical mutagens which damage DNA and threaten genomic integrity. Ionising radiation (IR) causes both direct damage to DNA through the creation of DNA strand breaks and also indirectly by generating ROS in the cell (vanAnkeren et al., 1988). Ultraviolet radiation, primarily from sunlight, causes covalent linkages to form between adjacent pyrimidines, forming bulky DNA lesions including cyclobutane pyrimdine dimers (CPDs) and pyrimidine (6-4) pyrimidone photoproducts (6-4PPs). Both of these lesions significantly distort the double helix and pose as potent obstacles to the replicative polymerases (Rastogi et al., 2010). Chemical agents, many of which are used in cancer chemotherapy, produce a variety of DNA lesions. These include alkylating agents, such as methyl methanesulfonate (MMS), which attach alkyl groups to bases, and crosslinking agents, including cisplatin, which can produce covalent linkages between bases in both the same strand (intrastrand crosslinks) and different strands (interstrand crosslinks) of DNA (Ciccia and Elledge, 2010). Intercalating agents, such as the anthracycline class of chemotherapeutic drugs,

interfere with base stacking rather than causing covalent modifications, consequently affecting replication and generating mutations (Minotti et al., 2004).

1.4.3. An Overview of DNA Repair

Given the vulnerability of DNA to damage and decay, cells have evolved numerous DNA repair mechanisms to counteract this damage prior to replication, thereby reducing the chance of mutagenesis (summarised in Figure 1.8.). These mechanisms include, mismatch repair (MMR), base excision repair (BER), nucleotide excision repair (NER), single-strand break repair (SSBR), non-homologous end joining (NHEJ), and homologous recombination (HR) (Iyama and Wilson III, 2013).

MMR increases the fidelity of DNA replication 1000-fold by removing and correcting bases which have been misincorporated by the replicative polymerases (Hsieh and Yamane, 2008). Humans possess seven different MMR proteins; mutL homolog 1 (MLH1), MLH3, mutS homolog 2 (MSH2), MSH3, MSH6, postmeiotic segregation increased 1 (PMS1), and PMS2 (Pal et al., 2008). Together these proteins are able to recognise misincorporated bases and form a complex to facilitate excision of the affected DNA before resynthesis by Pol δ and subsequent ligation.

BER works in a similar manner to MMR to repair base lesions, such as those caused by oxidative damage and spontaneous deamination (Almeida and Sobol, 2007; Hitomi et al., 2007). In this mechanism, DNA glycosylases hydrolyse the base-sugar bond before strand incision through apurinic/apyrimidinic lyase (AP lyase) endonuclease activity (Hitomi et al., 2007). Pol β can subsequently insert a base into the gap left from incision before DNA ligase III and X-ray repair cross complementing protein-1 (XRCC1) seals the DNA, this is termed short patch BER (Dianova et al., 2004). Alternatively, Pol β, δ, or ε can work in conjunction with PCNA and poly(ADP-ribose) polymerase 1 (PARP) in long patch BER where 2 to 13 bases are displaced during gap-filling synthesis, producing a flap which is processed by FEN1 and DNA ligase I (Iyama and Wilson III, 2013).

More bulky lesions, which distort the DNA backbone, are removed through NER (Iyama and Wilson III, 2013). Dysfunction in any of the main NER components causes Xeroderma pigmentosum (XP), an autosomal recessive disorder with eight different genetic complementation groups, representing the main NER proteins (XPA to XPG). Together these proteins, along with additional repair factors, bind to the distorted DNA, unwind the duplex, and remove a section of ssDNA 24-32 nucleotides in length (Lehmann, 2011). The resulting ssDNA is then filled in by the replicative DNA polymerases and ligated. NER can also be coupled to transcription, here stalling of RNA polymerase (RNAP) triggers removal of the lesion and restart of transcription by the

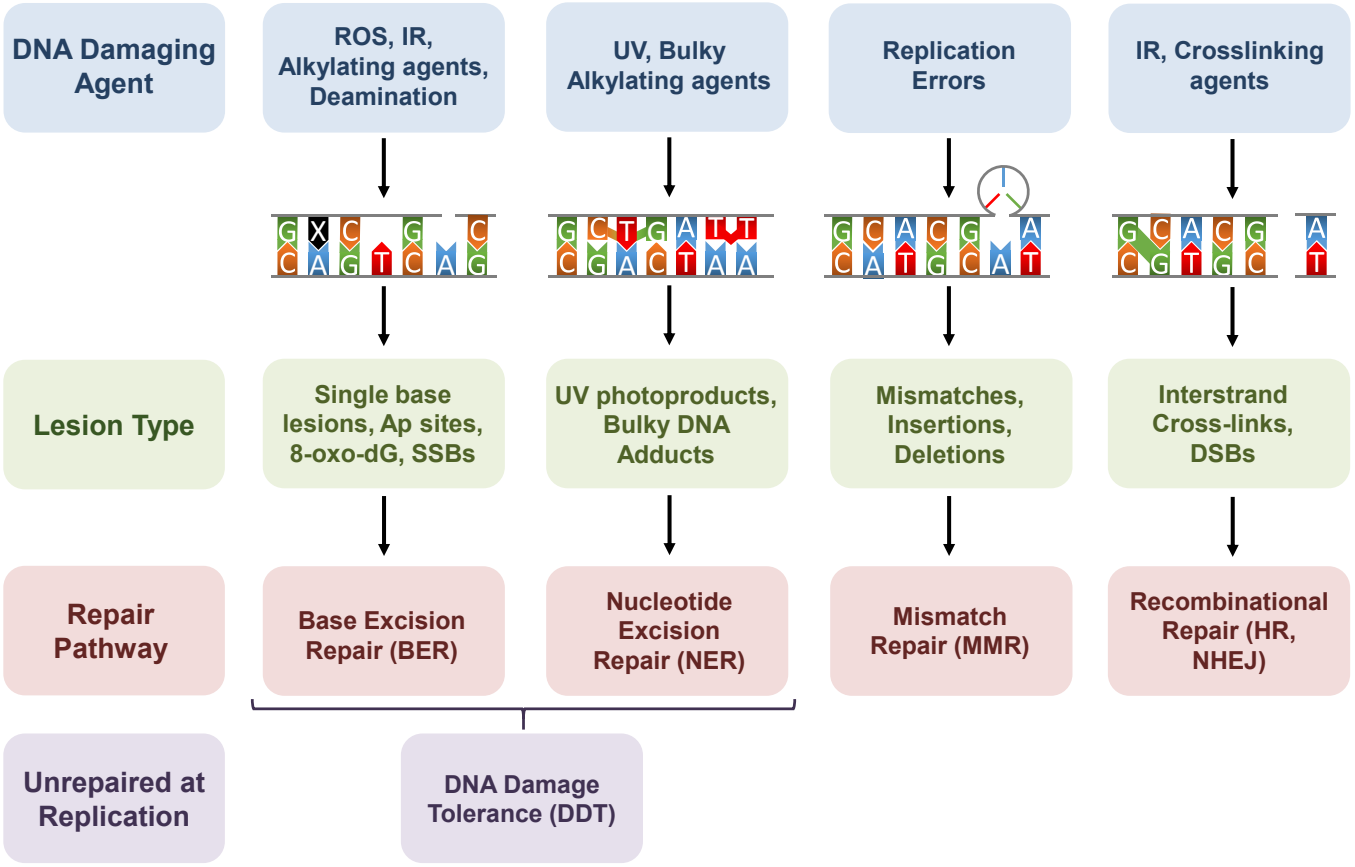


Figure 1.8. Summary of DNA damaging agents and repair mechanisms.
Numerous endogenous and exogenous agents contribute to DNA damage. This damage must be accurately repaired to prevent genome instability. Some of the agents contributing to DNA damage (blue), in addition to the types of damage produced (green) and the repair pathways employed to remove this damage (red) are displayed. DNA damage lesions which are unrepaired by BER or NER prior to replication can be bypassed by the replisome through DDT mechanisms (purple).

Cockayne syndrome (CS) proteins CSA and CSB, named after the autosomal recessive disease caused by their mutation (Iyama and Wilson III, 2013).

Single-strand breaks occur at very high rates, estimated to be in the order of tens of thousands per cell per day (Lindahl, 1993). These breaks are formed by canonical DNA damage, in addition to enzymatic intermediates of repair pathways and catalytic intermediates of topoisomerases (Iyama and Wilson III, 2013). Due to the range of potential causes, single-strand breaks can generate a diverse range of damaged 3' and 5' termini. Consequently, specialised enzymes including, DNA ligases, tyrosyl-DNA phosphodiesterase 1 (TDP1), aprataxin (APTX), polynucleotide kinase 3'-phosphatase (PNKP), and AP endonuclease 1 (APE1), are employed to process and repair the DNA termini before sealing the break. Intriguingly, mutation of many of these factors, such as TDP1, APTX, and PNKP, are linked to neurological disorders, rather than cancer, suggesting that these mechanisms may be in non-dividing cells (Gueven et al., 2004; Shen et al., 2010; Takashima et al., 2002).

Although occurring at a much lower rate, DSBs are extremely deleterious to the cell, with the potential to cause significant genome instability and ultimately cell death if left unrepaired (Bohgaki et al., 2010; Jackson and Bartek, 2009). Two major pathways are involved in the repair DSBs: HR and NHEJ. To operate HR requires a homologous sister chromatid and therefore is only active during the S and G2-phases of the cell cycle. Here, it primarily functions to repair DSBs caused by replication fork collapse as a result, for example, of polymerase stalling lesions (Iyama and Wilson III, 2013). The pathway involves the sensing of the DSB, 5'-3' resection to generate 3' ssDNA overhangs, stand invasion of the intact homologous region of the sister chromatid, polymerase extension and D-loop formation, and finally Holliday junction formation and resolution or synthesis-dependent strand annealing. In the absence of a sister chromatid NHEJ is used for DSB repair. The process involves three main steps which direct ligation of the two ends of the break, these include: the sensing and recognition of the break, processing of termini to remove damage and reveal microhomology, and joining and ligation of the two ends. Despite being the major DSB repair pathway in higher eukaryotes, NHEJ can often be error prone due to loss of genetic information at the break (Lieber, 2010).

1.5. DNA damage tolerance

The structural features of the replicative DNA polymerases, which confer inherent high fidelity, also make the enzymes intolerant to distortions in the template DNA. Consequently, unrepaired DNA damage lesions and secondary structures, which persist into S-phase, block replication and cause fork stalling. The stalling of replicases at these

sites can elicit uncoupling of leading and lagging strand synthesis, due to the ability of the replicative helicase to bypass the lesion and facilitate continued unwinding of duplex DNA (Lopes et al., 2006). The effect of this on the lagging strand is limited due the discontinuous nature of synthesis there. However, uncoupling of leading strand replication can generate stretches of RPA-bound ssDNA on one side of the fork, producing a substrate for ataxia-telangiectasia mutated- and Rad3-related (ATR) –ATR-interacting protein (ATRIP) binding and checkpoint kinase 1 (Chk1) activation (Zou and Elledge, 2003). This RPA-bound ssDNA also triggers DNA damage tolerance (DDT) mechanisms which allow replication to be completed in the presence of fork stalling lesions. Eukaryotes utilise two such mechanisms: recombination-dependent template switching (TS) which uses a homologous template, usually the sister chromatid, to bypass the lesion in a generally “error-free” manner, and translesion synthesis (TLS) where specialised DNA polymerases directly synthesise past the lesion in what is considered “error-prone” bypass. Additionally, the firing of dormant origins can help to ensure complete genome duplication in the presence of replication stress (summarised in Figure 1.9.). In this section, each of these mechanisms will be outlined and discussed before focussing on their timing during replication and relationship with repriming.

1.5.1. Firing of Dormant Origins

During origin licensing in G1-phase a ~ 20-fold excess of MCM2-7 is loaded onto DNA over the level required for the number of origins used in normal DNA replication (Edwards et al., 2002). This excess MCM is not located at the same site as ORC and does not affect normal replication rates when depleted to a level of two per origin in *Xenopus* egg extract (Edwards et al., 2002; Mahbubani et al., 1997; Oehlmann et al., 2004; Ritzi et al., 1998). Additionally, cells with a reduced amount of chromatin bound MCM display normal rates of S-phase progression (Tsao et al., 2004). However, loss of excess MCM sensitises DNA synthesis, cellular proliferation, and cell survival, to replication inhibitors, suggesting a role in the maintenance of DNA replication following perturbation (Ge et al., 2007; Woodward et al., 2006). Indeed, studies of human U2OS cells identified that DNA replication inhibitors induced firing of dormant origins not used in normal DNA replication (Ge et al., 2007). The firing of these origins was dependent on the excess chromatin bound MCM. Coupled with this, replication or checkpoint inhibition causes a decrease in replication origin spacing, demonstrating that additional origins are fired under these conditions (Gilbert, 2007; Maya-Mendoza et al., 2007; Woodward et al., 2006).

It is thought the firing of dormant origins is regulated by a simple passive mechanism (Ge et al., 2007). Following activation of an origin cluster, the dormant origins in that

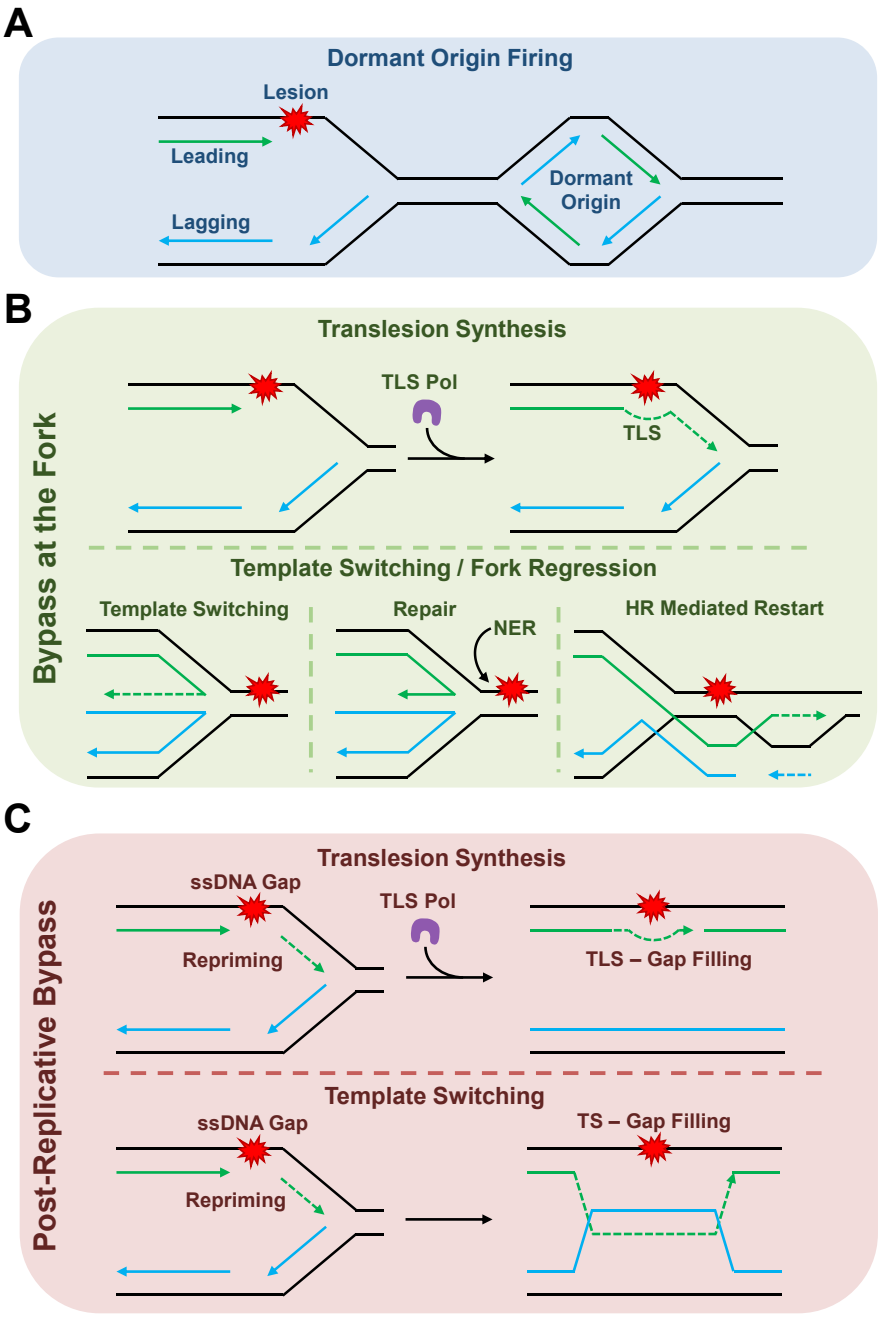


Figure 1.9. DNA damage tolerance mechanisms.
(A) DNA replication can proceed in presence of replisome-stalling obstacles though the activation of dormant origins of replication downstream of the damage. This allows unreplicated regions in between adjacent replisomes to be filled in. **(B)** DDT mechanisms at the replication fork. DNA damage lesions can be directly overcome by the action of TLS polymerases which are able to replicate over the lesion. Alternatively, fork regression can occur, generating a four way Holliday junction. Here, the nascent lagging strand can be utilised as a template for extension of the nascent leading strand, before reverse branch-migration places the nascent leading strand in a position downstream of the lesion. Fork regression also places the stalling lesion in a dsDNA context which may permit repair and removal by NER to allow replisome progression. Finally, homologous recombination (HR) can be used, whereby the nascent leading strand is extruded and undergoes template exchange with homologous sequences which are used as a template for further extension. **(C)** Post-replicative DDT mechanisms. Repriming of replication downstream of the lesion/obstacle leaves behind a ssDNA gap which can be subsequently filled in. This gap-filling may be performed by a TLS polymerase or through template switching where the sister chromatid is utilised as a template.

cluster have only a short time to fire before they are replicated by converging forks from neighbouring origins. Thus, if fork progression is slowed or stalled there is more time, and therefore an increased probability, of the origin firing before it is replicated. In support of this, dormant origin firing is not dependent on checkpoint kinase activity (Ge et al., 2007). However, it has also been speculated that phosphorylation of MCM by ATR may play a role in activating dormant origins during replication stress (Blow et al., 2011). Regardless of the mechanism controlling MCM activation, by firing dormant origins unreplicated DNA in between two stalled or slowed converging forks can be filled in, allowing bulk DNA synthesis to be completed (Figure 1.9.A.). The requirement of dormant origins in times of replication stress explains the excess MCM loaded onto DNA during origin licensing in G1-phase. Importantly, new origins cannot be licensed during S-phase, as this would occur on both replicated and unreplicated DNA, consequently leading to rereplication of nascent strands.

1.5.2. Translesion Synthesis

Although the firing of dormant origins allows unreplicated DNA between stalled forks to be filled in, it does not permit synthesis opposite the stalling lesion and therefore cannot ensure complete genome duplication alone. In addition to ATRIP/ATR, the generation of RPA-bound ssDNA, following leading strand replication uncoupling, provides a substrate for binding of E3 ubiquitin ligase Rad18 and E2 ubiquitin-conjugating enzyme Rad6, which together mediate monoubiquitylation of PCNA (Davies et al., 2008; Hoege et al., 2002). This enhances the affinity of TLS polymerases for PCNA through their ubiquitin-binding zinc finger (UBZ) and ubiquitin-binding motif (UBM) domains (Bienko et al., 2005). It is thought that this increased affinity for PCNA serves to recruit TLS polymerases to the damage site, here the polymerase can facilitate extension past the lesion, allowing replication to proceed and placing the damage in a dsDNA context to permit repair (Figure 1.9.B. and C.). Despite being firmly established in yeast, PCNA ubiquitylation as a mechanism of TLS polymerase recruitment in higher eukaryotes has not yet been confirmed and remains a subject of much debate (Despras et al., 2012; Durando et al., 2013; Göhler et al., 2011; Hedglin et al., 2016; Hendel et al., 2011; Sabbioneda et al., 2008, 2009; Zhao and Washington, 2017). One recent study identified that switching from human Pol δ to TLS polymerases occurs independently of PCNA monoubiquitylation and is instead a passive process based on the intrinsic DNA binding properties of the different polymerases (Hedglin et al., 2016). This report also further suggests that PCNA monoubiquitylation may rather play a role in altering the local chromatin structure surrounding the lesion to facilitate access of TLS polymerases.

The best characterised human TLS polymerases include Pols η, ι, κ, and Rev1 of the Y-family, and Pol ζ of the B-family (Table 1.2.). Overall, these enzymes conform to the typical polymerase right-hand topology. However, they also possess a number of modified features which make them suitable for the bypass of DNA lesions. In comparison to the replicative DNA polymerases, TLS polymerases have much more spacious catalytic sites, stubbier finger and thumb subdomains, and lack exonuclease proofreading domains (Sale et al., 2012). Y-family TLS polymerases also possess an additional polymerase-associated domain (PAD) or "little-finger", which is involved in binding the DNA template and is thought to confer lesion specificity to each polymerase (Boudsocq et al., 2004). These features are exemplified by Pol η which can incorporate A-A opposite a T-T CPD with similar accuracy to unmodified T-T (Johnson et al., 1999). Importantly, Pol η can accommodate both of the thymine bases of the CPD in its particularly large active site, these bases are then further stabilised to allow accurate and efficient bypass (Biertümpfel et al., 2010; Silverstein et al., 2010). Following incorporation opposite the lesion, a specialised β-strand in the enzyme's little-finger subdomain provides a positively charged surface to counter the distortion created by the CPD (Silverstein et al., 2010). This prevents frameshifts and slippage being caused by the distorted DNA duplex. Together, these specialised features allow the TLS polymerases to directly bypass bulky and distorting lesions, but also result in low fidelity and poor processivity on undamaged DNA. As a consequence, the contribution of these enzymes to DNA synthesis is kept to a minimum. Indeed, steric clashes cause Pol η to dissociate from the DNA template after only three post-lesion bases have been incorporated (Silverstein et al., 2010).

Interestingly, vertebrate Rev1 is able to interact with the other Y-family TLS polymerases and Pol ζ (Guo et al., 2003; Murakumo et al., 2001; Ohashi et al., 2004, 2009; Tissier et al., 2004). These interactions appear to play an important role in the recruitment of TLS polymerases to the replication fork during lesion bypass. Indeed, a number of studies have demonstrated that PCNA ubiquitylation is dispensable for lesion bypass at the replication fork, with polymerases potentially being recruited through their interaction with Rev1 instead (Bienko et al., 2005; Despras et al., 2012; Edmunds et al., 2008; Nelson et al., 2000). Here, it is likely that Rev1 simply serves as a scaffold to link TLS polymerases to PCNA, due to its shared interaction with both (Sale et al., 2012). Additionally, it has been suggested that interactions between multiple TLS polymerases and PCNA can form higher order structures termed "PCNA tool belts" (Boehm et al., 2016). Intriguingly, these "PCNA tool belts", and TLS polymerases linked to PCNA through Rev1, have been identified in single-molecule studies and appear to be able to interconvert without

Pol	Family	Features
Rev1	Y	• Incorporates dCMP opposite dG and Ap sites • Acts as a scaffold by interacting with other Y-family Pols and Pol ζ • Generates mutations at G-C base pairs during immunoglobulin gene somatic hypermutation • Plays a role in G-quadruplex replication
Pol η	Y	• Accurately bypasses CPDs • Defects lead to XPV • Generates mutations at A-T base pairs during immunoglobulin gene somatic hypermutation • Accumulates in replication foci
Pol ι	Y	• Has a unique replication fidelity – accurately replicates dA, but is error-prone when replicating dT • Accumulates in replication foci, but for a shorter time than Pol η • May play a role in BER
Pol κ	Y	• Prone to making -1 frameshift mutations, but accurately bypasses a number of N^2-dG lesions • Has additional roles in the repair synthesis step of NER • Most accurate of the Y-family Pols • Efficient extender of mispaired termini
Pol ζ	B	• Similar fidelity to Pol α • Plays an 'extender' role in TLS • The gene is embryonic lethal in mice • Required for post-replicative damage tolerance of incorporated rNTPs

Table 1.2. The canonical eukaryotic TLS polymerases.
Displayed are the canonical and best characterised of the eukaryotic TLS polymerases, including the Y-family polymerases η (red), ι (green), κ (purple), and Rev1 (blue), in addition to Pol ζ of the B-family (orange). Adapted from Sale et al., 2012.

dissociation (Boehm et al., 2016). Thus, this complex formation may permit rapid testing of multiple TLS polymerases to identify the most suitable one for bypassing a specific lesion *in vivo*.

1.5.3. Template Switching and Recombination-Mediated Restart

Following fork stalling and Rad6/Rad18 binding, the additional recruitment of Rad5 and the heterodimeric Mms2-Ubc13 complex facilitates polyubiquitylation of PCNA (Hoege et al., 2002; Ulrich and Jentsch, 2000). This in turn inhibits TLS and promotes TS (Branzei et al., 2004; Haracska et al., 2004; Pfander et al., 2005; Zhang and Lawrence, 2005). In addition to its E3-ubiquitin ligase activity, Rad5 also exhibits DNA-dependent ATPase activity which is linked to a DNA helicase function required for DDT. This helicase activity is able to promote replication fork regression *in vitro* and similar results have been observed with one of the human Rad5 homologues, helicase-like transcription factor (HLTF) (Achar et al., 2011; Blastyák et al., 2007). Fork regression or reversal can potentially promote restart in a number of ways (Figure 1.9.B.). Firstly, fork reversal causes rewinding of the nascent strands, which base-pair together, forming a four-way Holliday junction. This returns the stalling lesion to duplex DNA and may facilitate repair and removal, with nucleolytic degradation of the branched structure subsequently allowing replication restart (Courcelle et al., 2003). Alternatively, the displaced nascent leading strand can utilise the nascent lagging strand as a template through TS, before reverse branch migration places the 3' end of the lagging strand beyond the lesion to allow continued extension (Branzei and Szakal, 2016; Higgins et al., 1976). A third scenario is also possible whereby restart occurs through a HR-mediated mechanism (Lambert et al., 2010). Here, the nascent leading strand is extruded and undergoes template exchange with homologous sequences, this leads to displacement loop (D-loop) formation to allow continued elongation and produces Holliday junctions which subsequently require resolution (Lambert et al., 2010; Minca and Kowalski, 2011). This mechanism can, however, lead to strand invasion events which use only limited homology, potentially causing genomic rearrangements (Branzei and Szakal, 2016; Lambert and Carr, 2013b). As such, this pathway appears to function as a last-resort or "salvage" pathway due its error-prone habit of using templates other than the sister chromatid (Branzei and Szakal, 2016).

The identification of the fork remodelling activities of Rad5 and HLTF, in addition to DNA translocases with strand annealing capabilities, such as SWI/SNF-related matrix-associated actin-dependent regulator of chromatin (SMARCAL1), legitimise the possibility that fork regression and TS function at the replication fork to mediate lesion

bypass (Bansbach et al., 2009). However, accumulating evidence suggests that TS occurs in a predominantly post-replicative fashion to fill in gaps left opposite lesions after repriming events, here error-free DDT occurs through strand exchange with the homologous sister chromatid (Figure 1.9.C.) (Branzei and Szakal, 2016).

1.5.4. Post-Replicative Repair and the Timing of DDT

In 1968, Rupp and Howard-Flanders observed that NER deficient *E. coli* could perform DNA replication in the presence of UV damage with only a slight delay (Rupp and Howard-Flanders, 1968). Moreover, they identified the presence of ssDNA gaps opposite UV lesions in the nascent daughter strands, which were subsequently filled and sealed in a process termed “post-replicative repair” (PRR). This seminal work, the first to describe DDT, also suggested that it functions behind the fork to fill in gaps left opposite lesions as a result of repriming. This is the obvious case for lagging strand synthesis where initiation of a new Okazaki fragment would easily allow resumption of replication after fork stalling. However, the long-standing view of leading strand synthesis as a continuous process, coupled with the discovery of TLS polymerases, led to the prevailing view for many years that DDT functions at the replication fork to promote continued progression. However, more recent studies are provoking a return to the original model for DDT, where both TS and TLS predominantly function as PRR mechanisms (Figure 1.9.C.).

Importantly, for PRR to operate on the leading strand, reinitiation of replication by repriming downstream of the blockage is required. In *E. coli*, origin independent leading strand repriming by the replicative primase, DnaG, has been demonstrated and is now a well-established method of replication restart (Heller and Marians, 2006; Yeeles and Marians, 2011). Further studies have confirmed that this repriming mechanism operates without dissociation of the replication machinery, consequently instilling innate damage tolerance on the core replisome and allowing lesions to be skipped in an efficient manner (Yeeles and Marians, 2013). Although leading strand repriming has not yet been demonstrated using a eukaryotic system *in vitro*, given recent advances in reconstituting eukaryotic replication, it seems likely the question of its existence will soon be addressed. Indeed, accumulating *in* vivo evidence points to a leading strand repriming mechanism in eukaryotes. Notably, in both yeast and mammalian systems, ssDNA gaps in daughter strands have been observed following DNA replication in the presence of UV damage (Lehmann, 1972; Lopes et al., 2006). Additionally, both TLS and TS mechanisms can be functionally separated from chromosomal replication without adverse effects (Daigaku et al., 2010; Karras and Jentsch, 2010). Post-replicative TLS in yeast is further supported

by the observation that Rev1 expression is cell-cycle regulated and peaks in G2-phase (Waters and Walker, 2006). In human fibroblasts, high UV exposure during DNA replication was found to only produce a marginal reduction in fork speed which was independent of Pol η (Elvers et al., 2011). Here, discontinuous elongation of replication forks was also observed, leading to speculation that repriming downstream of UV lesions and not TLS at the fork, was maintaining replisome progression. Likewise, in *Xenopus* egg extracts, continued primer synthesis at replication forks stalled by aphidicolin has been demonstrated (Van et al., 2010).

Recent studies provide further evidence for leading strand repriming in yeast and suggest that TS DDT mechanisms predominantly operate behind the fork in a gap-filling manner (Figure 1.9.C.). Firstly, impairment of repriming pathways and replicative helicase-primase coupling did not affect bulk DNA replication, but did cause a defect in error-free bypass by TS (Fumasoni et al., 2015). This decrease in TS was accompanied with increased mutagenic DDT, fork reversal, and genome rearrangements. Here, the authors suggest that in WT cells, repriming allows replication restart, leaving a ssDNA gap which is filled by TS. In the absence of repriming, leading/lagging replication uncoupling occurs, and forks are restarted by mutagenic reversal and HR mechanisms. Secondly, electron microscopy (EM) studies of TS switching intermediate structures known as sister-chromatid junctions (SCJs), support the post-replicative gap-filling model (Giannattasio et al., 2014). Factors important for gap-processing during Okazaki fragment processing, including the 9-1-1 complex and exonuclease 1 (Exo1) also appear to play an important role in TS, lending further support to this model (Karras et al., 2013; Vanoli et al., 2010).

Thus, the weight of evidence is clearly in support of post-replicative DDT as a conserved mechanism in eukaryotes as well as bacteria. Nevertheless, it remains likely that DDT at the fork also occurs *in vivo*. In particular, fork HR-mediated DDT is likely to be required to overcome obstacles which cannot be bypassed by the replicative helicase, and are therefore not amenable to repriming, such as interstrand crosslinks. At these regions fork stabilisation through reversal will also be required to prevent error-prone restart and genomic instability (Branzei and Szakal, 2016). Regardless of the contribution of both mechanisms, the requirement of leading strand repriming for PRR suggests that primases play an important role in maintaining fork progression during replication stress and demonstrates that these enzymes are not only required for the initiation of replication at origins. Indeed, in recent years primases have emerged as a diverse class of enzymes with a range of roles in DNA metabolism.

1.6. Primase-Polymerases are a Functionally Diverse Superfamily of Replication and Repair Enzymes

Until relatively recently, DNA primases were viewed simply as a class of proteins that synthesise short RNA primers requisite for the initiation of DNA replication. However, recent studies have shown that this perception of the limited activities associated with these diverse enzymes can no longer be justified. Numerous examples can now be cited demonstrating how the term 'DNA primase' only describes a very narrow subset of these nucleotidyltransferases, with the vast majority fulfilling multifunctional roles from DNA replication to damage tolerance and repair. This section of the introduction focuses on the archaeo-eukaryotic primase (AEP) superfamily, drawing on recently characterised examples from all domains of life to highlight the functionally diverse pathways in which these enzymes are employed. The broad origins, functionalities and enzymatic capabilities of AEPs emphasises their previous functional misannotation and supports the necessity for a reclassification of these enzymes under a category called Primase-Polymerases within the broader functional grouping of polymerases. Importantly, the repositioning of AEPs in this way better recognises their broader roles in DNA metabolism and encourages the discovery of additional functions for these enzymes, aside from those highlighted here.

1.6.1. Two Distinct Primases: DnaG and AEP Primase Superfamilies

Early studies identified that the *E. coli* DnaG protein is responsible for the initiation of Okazaki fragment synthesis by the generation of short RNA primers (Bouché et al., 1978). All bacteria possess DNA primases belonging to the DnaG superfamily, which fulfil the canonical primer synthesis role during DNA replication. Typically, these monomeric DnaG-like replicative primases are helicase-associated, permitting the synthesis of RNA primers of between 10 to 60 nucleotides in length on most ssDNAs (Frick and Richardson, 2001). These enzymes contain a characteristic catalytic domain of the topoisomerase-primase (TOPRIM fold, composed of an α/β core with four conserved strands and three helices, in addition to two conserved catalytic motifs (Keck et al., 2000). The first of these motifs, a conserved glutamate, is thought to act as a general base during nucleotide polymerisation. The second motif contains two conserved aspartates (DxD), which coordinate Mg^{2+} ions required for catalytic activity (Aravind et al., 1998a).

Although functionally related, the AEP superfamily is evolutionarily and structurally distinct from the bacterial DnaG primases (Aravind et al., 1998a; Aravind and Koonin,

2001; Augustin et al., 2001; Frick and Richardson, 2001; Keck et al., 2000). In common with prokaryotic DnaG primases, AEPs are absolutely required for the initiation of DNA replication in archaea and eukaryotes (Frick and Richardson, 2001). Despite this, DnaG-like primases have been identified in archaeal genomes and, similarly, AEPs are also found to be distributed across all domains of life. Replicative primases of the AEP superfamily typically form a heterodimeric complex containing both a small catalytic subunit (PriS/Prim1) and a large accessory subunit (PriL/Prim2). As previously mentioned, in eukaryotes this heterodimer forms a complex with the DNA Pol α subunits (PolA1 and PolA2) that together initiate DNA replication (Frick and Richardson, 2001). The AEP superfamily is distinguished by a characteristic catalytic core composed of two modules; an N-terminal $(\alpha\beta)_2$ unit that has no equivalent structural homology to other proteins in the structural database (PDB) and a C-terminal unit, which like the A-, B-, and Y-family DNA polymerases, is a highly derived RNA recognition motif (RRM) (Figure 1.10.). This catalytic core harbours three conserved motifs (motifs I, II and III), an hhhDhD/E motif (where 'h' is a hydrophobic residue), an sxH motif (where 's' is a small residue and 'x' is any residue) and an hD/E motif (Iyer et al., 2005). The first and third of these motifs are involved in divalent metal ion coordination for catalysis, whilst the sxH motif is required for nucleotide binding (Augustin et al., 2001; S.-H. Lao-Sirieix et al., 2005; Lipps et al., 2004). Multiple mutagenesis studies have shown these motifs to be essential for catalysis (Augustin et al., 2001; Bianchi et al., 2013; Copeland and Tan, 1995; Della et al., 2004; Galal et al., 2012; Klinedinst and Challberg, 1994; Lao-Sirieix and Bell, 2004; Lipps et al., 2003; Mikhailov and Rohrmann, 2002). In addition to these motifs, some AEPs also possess additional associated domains including zinc-binding and helicase domains (Figure 1.11.).

Despite the apparent uniqueness of the AEP catalytic fold, the highly conserved catalytic aspartate residues of these enzymes are superposable with the catalytic core of the X-family DNA polymerases, including Pol β (Augustin et al., 2001; Kirk and Kuchta, 1999a). However, this apparent similarity is thought to be a result of convergent evolution as the secondary structural contexts in which these aspartate residues are located differs (S. Lao-Sirieix et al., 2005; S.-H. Lao-Sirieix et al., 2005). Nevertheless, this similarity, coupled with the requirement of divalent metal ions for catalysis, allows inference of a two-metal ion mechanism of catalysis, similar to that employed by DNA polymerases (Kirk and Kuchta, 1999a; Steitz et al., 1994). This mechanism is now supported by the structure of a pre-ternary complex of a mycobacterial AEP bound to DNA (Brissett et al., 2011).

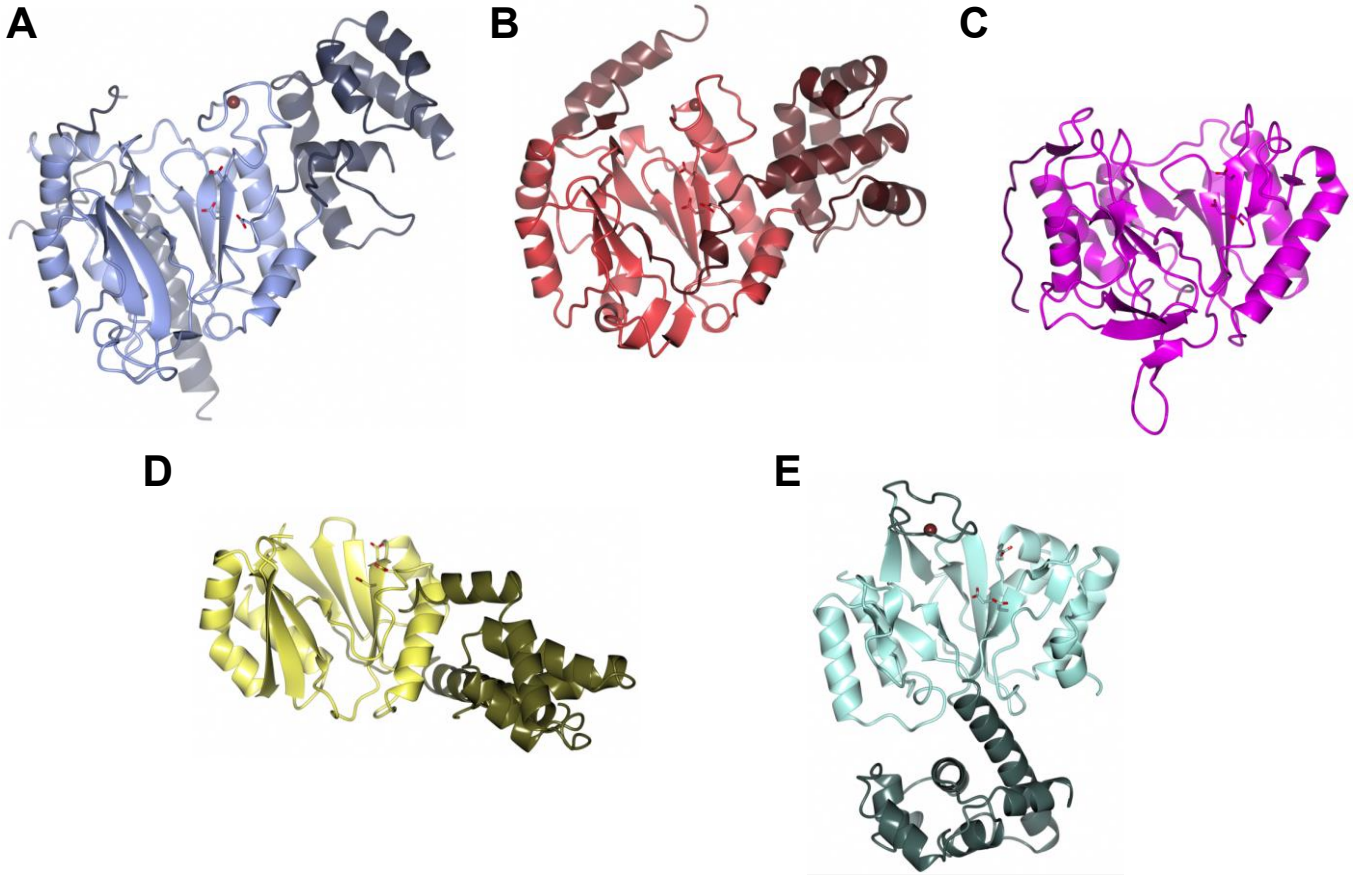


Figure 1.10. Architecture of AEP catalytic subunits from the major domains of life.
Representative examples of the crystal structures of AEPs that have been elucidated. **(A)** Human primase subunit (Prim1) (PDBID: 4RR2), sky blue. **(B)** The primase small catalytic subunit (PriS/Prim1) from the archaeal species *P. horikoshi* (PDBID: 1V33), pale crimson**. (C)** The NHEJ repair polymerase (PolDom/LigD-Pol) from *M. tuberculosis* (PDBID: 2IRU), magenta. **(D)** The AEP domain of RepB' encoded by the *E.coli* plasmid RSF1010 (PDBID: 3H20), gold. **(E)** The AEP domain of ORF904 encoded by the *S. islandicus* plasmid pRN1 (PDBID: 3M1M), sea green. The conserved catalytic core of these enzymes is shown in a lighter hue and catalytic triads are rendered as sticks with the acidic oxygens coloured red. Where present, the coordinated zinc atoms in the zinc finger domains are coloured tan.

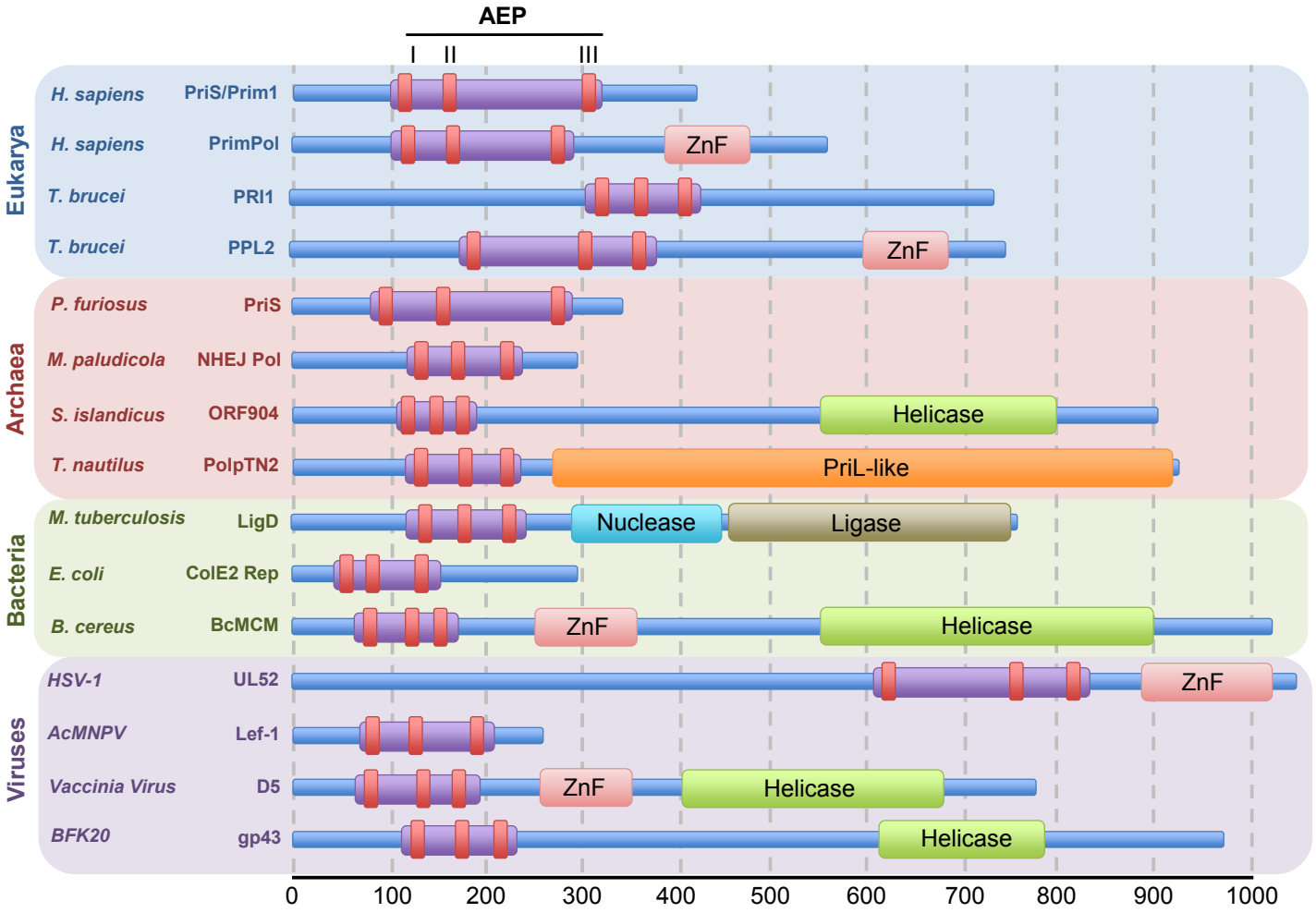


Figure 1.11. Domain organisation of members of the AEP superfamily.
The AEP superfamily is formed of a number of divergent enzymes with varying domain organisations. Some representative examples of members of the AEP superfamily are displayed here with the three signature catalytic motifs of the AEPs are depicted in red, in addition to any accessory domains associated with this domain. The blue, red, green and purple backgrounds correspond to the different domains of life in which these primase family members belong.

1.6.2. Evolutionary History of AEPs

The lack of homology between the bacterial and archaeal/eukaryotic primase superfamilies also extends to other replicative proteins of these domains, including DNA polymerases and helicases. This clear distinction between the replicative machinery employed by bacteria and archaea/eukaryotes has generated debate as to how these two groups of replicative enzymes arose. Notably, in contrast to the differences in DNA replication machinery, the core components of transcription and translation are conserved across domains (Leipe et al., 1999; Sweetser et al., 1987). This observation led to a hypothesis that the two replicative systems evolved twice independently from a common ancestor which utilised reverse transcription to replicate an RNA/DNA genome (Leipe et al., 1999) (Figure 1.12.A.). The evolution of DNA replication-competent cells then subsequently led to the elimination of the reverse transcription pathway. In support of this model, a number of primases and polymerases possess, or can be engineered to exhibit, reverse transcriptase activity (Gill et al., 2014; Jozwiakowski et al., 2015; Jozwiakowski and Connolly, 2011; Myers and Gelfand, 1991; Ong et al., 2006). A second model proposes that the last universal common ancestor (LUCA) possessed both an AEP and TOPRIM primase (Hu et al., 2012) (Figure 1.12.B.). In bacteria, selective pressure resulted in the loss of AEPs as replicative primases and, similarly, in archaea TOPRIM primases lost their role in priming replication. Importantly, many bacteria and archaea still retained their respective AEP and TOPRIM primases, however the roles of these enzymes changed with the AEPs employed in DNA repair processes, e.g. non homologous end-joining (NHEJ) in bacteria (Della et al., 2004) and the TOPRIM primases potentially utilised for RNA degradation in archaea (Le Breton et al., 2007; Walter et al., 2006). This model also predicts that in eukarya the DnaG primase was lost and other proteins were acquired to replace its roles (Hu et al., 2012). An alternative scenario to these models is that LUCA possessed either a TOPRIM primase or an AEP and subsequent selective pressure led to the emergence of the second primase superfamily in either the bacterial or archaeal/eukaryotic lineages (Figure 1.12.C.). In this case, AEPs could have been acquired later by bacteria and viruses through horizontal gene transfer to fulfil alternative roles in DNA replication, repair and damage tolerance. Which of these models is likely to be correct remains to be established.

Despite the lack of homology between the primase superfamilies, the evolutionary history of the AEP superfamily displays an interesting parallel with that of the DnaG TOPRIM primases. Iyer *et al.* reported an extensive *in silico* analyses of the AEP superfamily and identified that the closest relatives of the AEP-fold, amongst the RRM-like proteins, are the rolling circle replication endonucleases (RCRE) and origin-binding domain proteins

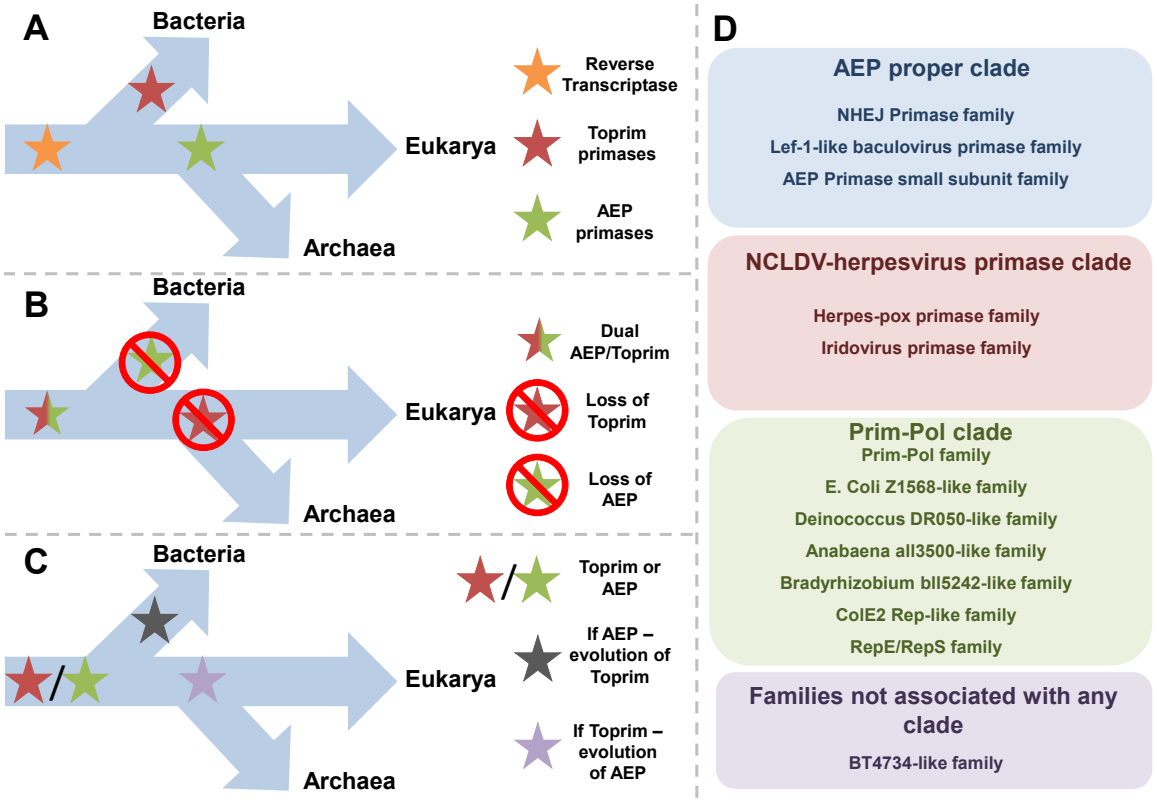


Figure 1.12. Alternative hypotheses for the evolution of AEP and Toprim primases.
(A) The first model of primase evolution suggests that primases evolved independently twice from a last universal common ancestor (LUCA). Bacterial ancestors evolved toprim primases and archaeal and eukaryotic ancestors evolved AEP primases. Subsequent horizontal gene transfer occurred between the two lineages to account for AEP primase's role in NHEJ in bacteria and toprim-type primase's role in archaeal RNA degradation. (**B)** The second model of primase evolution suggests that LUCA had a dual primase replication mechanism, consisting of both AEP and toprim primases. During the evolution of bacteria, they lost the replicative function of the AEP primases but retained them for the auxiliary function for NHEJ-mediated DNA repair. During the evolution of the archaeal/eukaryotic lineage, the replicative function of toprim primases was lost but their auxiliary role in archaeal RNA degradation was retained. (**C)** The third model of primase evolution suggests that LUCA had either an AEP or toprim-like primase. Significant, evolutionary pressures could then have driven the evolution or acquisition of a second class of primase. (**D)** 12 of the 13 major AEP families can be arranged into three higher order clades, the AEP Proper Clade, the NCLDV-Herpesvirus primase clade and the Prim-Pol clade.

(OBDs) of the papovaviruses. This close evolutionary relationship between the AEPs and RCRE echoes that of the DnaG TOPRIM primases with topoisomerases. Intriguingly, this reveals that both primase superfamilies share close evolutionary ties with nucleases, which offer an alternative solution to the DNA replication initiation problem (Iyer et al., 2005). Specifically, transfer of the 5' end of a nicked DNA strand to a tyrosine on the nuclease allows the elongation of the free 3' OH group by a DNA polymerase for synthesis of the new strand. In several families of DNA viruses and phage, this method of initiating rolling circle replication is employed. Iyer *et al*. suggest that RCRE and OBDs share a common ancestor with the AEPs that possessed polymerase activity. The RCRE subsequently evolved from this enzyme by acquiring nuclease and losing polymerase activities, meanwhile OBDs lost all catalytic activity. However, the authors also accept the possibility that the AEP-RCRE-OBD common ancestor was simply a nucleic acid binding protein, which utilised its divalent cation coordinating acidic residue to aid in DNA binding. This ancestral protein may then have acquired nuclease activity, whilst polymerase activity could have been independently acquired in numerous descendent lineages (Iyer et al., 2005).

To date, the AEP superfamily can be classified into 13 major families, 12 of these can be further organised into three higher-order clades; the AEP proper clade, the NCLDV-herpesvirus clade and the PrimPol clade (Figure 1.12.D.) (Iyer et al., 2005). Regardless of the somewhat murky evolutionary origins of the AEP superfamily, studies in recent years have illustrated that these enzymes have diversified to fulfil a range of specialist roles in DNA replication, repair and damage tolerance, as will be described in the following sections.

1.6.3. Archaeal Primases can act as Replicative Polymerases

As previously mentioned, in eukaryotes the replicative heterodimeric primase (Prim1/2) complexes with Pol α (PolA1/2) to form a tetrameric complex. The resolution of the crystal structure of the human heterodimeric primase identified that the small subunit (Prim1) utilises the same set of functional residues for primer initiation and elongation (Figure 1.10.), in addition, this study also identified the mode of association between the primase and Pol α (Kilkenny et al., 2013; Vaithiyalingam et al., 2014). This close partnership allows the polymerase subunit to easily access and extend the RNA primer synthesised by Prim1, before hand-off to the more processive replicases. However, this complexity does not exist in archaea, which lack Pol α subunits or any apparent interaction between the PriS/L complex and the archaeal B-family replicases (Le Breton et al., 2007; Liu et al., 2001). Remarkably, the replicative primase is able to synthesise

and elongate its own primers. Evidence for archaeal primases as DNA-dependent DNA polymerases was first noted in the archaeal Prim1 homologue, PriS, from the hyperthermophile *Pyrococcus furiosus* (*Pfu*) (Bocquier et al., 2001). In addition to displaying DNA polymerase activities (Figure 1.13.), *Pfu* PriS can also synthesise primers using dNTPs, as well as rNTPs. This contrasts with eukaryotic Prim1, which can only synthesise RNA primers (Liu et al., 2001). This ability to synthesise DNA primers and perform DNA-dependent DNA polymerase activity has also been observed in the PriS/L complexes from other archaea, including *Pyrococcus horikoshii* (Matsui et al., 2003), *Sulfolobus solfactaricus* (Lao-Sirieix and Bell, 2004), *Thermococcus kodakaraensis* (Galal et al., 2012), and *Archaeoglobus fulgidus* (Jozwiakowski et al., 2015). In fact, for each of these species, except *Thermococcus kodakaraensis,* the replicative primase actually shows a preference to prime using dNTPs over rNTPs. Strikingly, the primer elongation capacity of these enzymes ranges from less than 500 bases in length to >7 kilobases (Bocquier et al., 2001; Galal et al., 2012; Lao-Sirieix and Bell, 2004; Le Breton et al., 2007; Liu et al., 2001). Together, these studies provided the first evidence that archaeal replicative primases can also be utilised in a role similar to that of Pol α in eukaryotes, thus establishing that these enzymes can act and be classified as primase-polymerases or Prim-Pols.

1.6.4. Primases Acting as Extra-Chromosomal Plasmid Replicases

In addition to PriS, some archaea possess additional AEPs encoded by extra-chromosomal plasmids, which are thought to partake in both the initiation and replication of these small molecules of DNA. The first of these to be identified was ORF904 encoded by the pRN1 plasmid (~5 kb) of *Sulfolobus islandicus*. ORF904 belongs to a novel family of primases present sporadically in crenarchaeal plasmids and Gram-positive bacterial plasmids, the Prim-Pol family (Iyer et al., 2005). The enzyme is composed of an N-terminal AEP domain and a C-terminal helicase / translocase domain. This AEP domain displays both DNA-dependent RNA/DNA primase and DNA polymerase activity (Figure 1.13.), with the helicase domain exhibiting DNA-dependent ATPase activity (Lipps et al., 2003). Notably, ORF904 shows a preference to generate DNA primers and, in the presence of dNTPs, it can extend these primers by several kilobases (Beck and Lipps, 2007). The crystal structure of its AEP domain revealed that this enzyme shares strong structural similarities with the *Pyrococcus* archaeal primase (Lipps et al., 2003), particularly in the arrangement of the metal coordinating acidic residues, which display strict conservation within the β-sheet region (Figure 1.10.) (S.-H. Lao-Sirieix et al., 2005). Interestingly, both enzymes possess zinc-binding motifs adjacent to the catalytic centre however, somewhat surprisingly, these motifs are located in unrelated positions in each

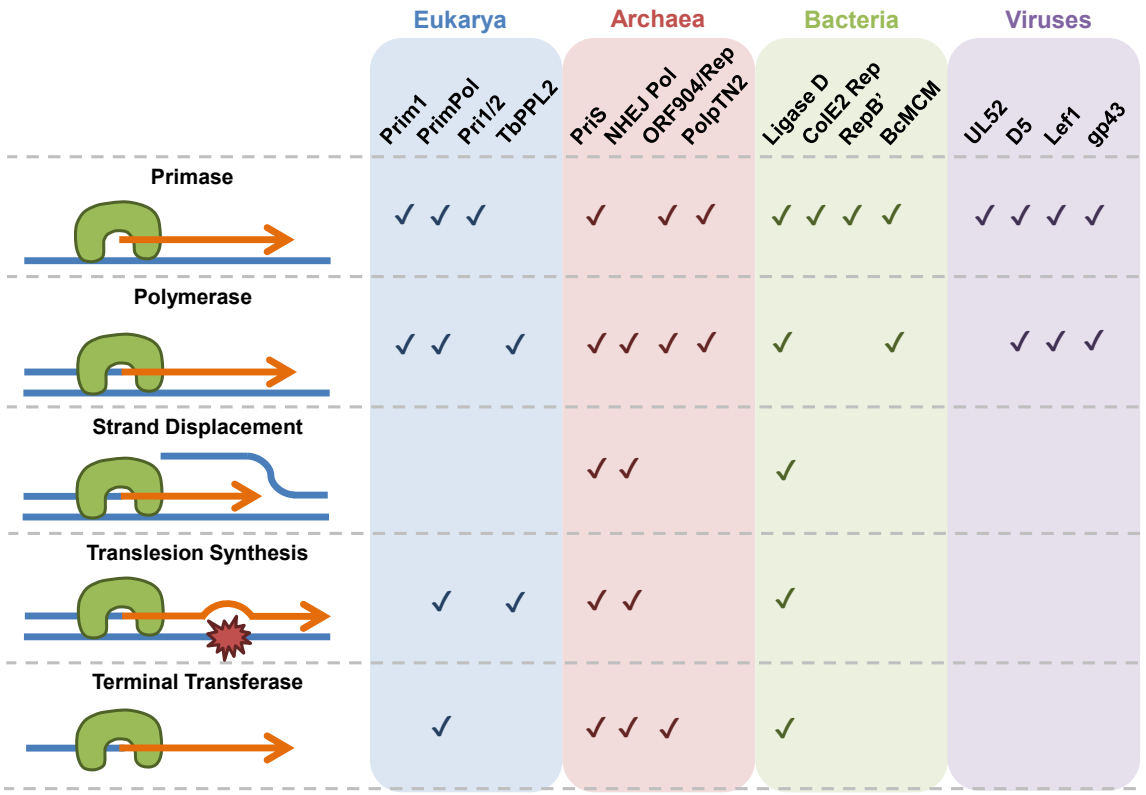


Figure 1.13. Nucleotidyltransferase activities associated with AEP members.
AEP-type primase family members possess many more activities, in addition to catalysing primer synthesis DNA for replication. The reported additional nucleotidyltransferase activities for each of the different AEPs are depicted, including polymerase activity (either DNA-dependent DNA polymerase or DNA dependent RNA polymerase), lesion bypass, terminal transferase and strand-displacement. The observed ability of each enzyme to perform the indicated activity is noted by a tick. The blue, red, green and purple backgrounds correspond to the domain of life in which the primase family is found. References describing the associated activities described here: (Bartlett et al., 2013; Beck and Lipps, 2007; Bianchi et al., 2013; Brooks and Dumas, 1989; Conaway and Lehman, 1982; Copeland and Wang, 1993; Crute and Lehman, 1991; De Silva et al., 2009; Della et al., 2004; Desogus et al., 1999; Falco et al., 2004; García-Gómez et al., 2013; Gill et al., 2014; Gong et al., 2005; Halgasova et al., 2012; Hines and Ray, 2011, 2010; Hu et al., 2012; Jozwiakowski et al., 2015; Keen et al., 2014; Klinedinst and Challberg, 1994; Le Breton et al., 2007; Lipps et al., 2004; Liu et al., 2001; Mikhailov and Rohrmann, 2002; Pitcher et al., 2007, 2005; Prato et al., 2008; Samuels et al., 2009; Sanchez-Berrondo et al., 2012; Takechi et al., 1995; Takechi and Itoh, 1995).

case. This observation has led to the suggestion that the common ancestor of both enzymes did not contain a zinc-binding domain and that two independent insertion events occurred to produce this domain in the evolution of each family (S.-H. Lao-Sirieix et al., 2005). In addition to ORF904, a highly related protein has been identified on the pIT3 plasmid of *Sulfolobus solfataricus* called Rep that also comprises an AEP domain fused to a putative helicase (Prato et al., 2008). The replicative N-terminal domain of this protein, termed Rep245, also possesses dNTP / rNTP-dependent primer synthesis and DNA polymerase activities (Figure 1.13.).

A recent report described the intriguing enzymatic activities of an AEP called PolpTN2 encoded by the pTN2 plasmid of *Thermococcus nautilus.* PolpTN2 is uniquely a fusion of an N-terminal domain homologous to PriS and a C-terminal domain related to PriL (Figure 1.11.) (Gill et al., 2014). This domain conformation is at odds with other plasmid-encoded primases, which are typically fused to helicases. Nevertheless, similar to other archaeal plasmid-encoded primases, PolpTN2 exhibits primase and DNA polymerase activities. The primase activity of PolpTN2 is exclusively limited to using dNTPs. In addition, the enzyme also has terminal transferase activity, which is greatly enhanced by the removal of the PriL-like region of the protein. This removal also confers reverse transcriptase activity to the primase (Gill et al., 2014). Interestingly, PolpTN2 and Rep(pIT3) lack a zinc-binding motif present in most other AEPs. The observation that the zinc-binding motifs of each AEP family may have arisen independently suggests that these plasmid-encoded AEPs may represent evolutionarily ancestral AEPs.

Bacteria, like archaea, also harbour extra-chromosomal plasmid DNA. Two decades ago, a Rep protein from the colicin E2 (ColE2) plasmid was found to have DNA primase activity (Figure 1.13.) (Takechi et al., 1995; Takechi and Itoh, 1995). A decade later, it was shown that this primase was in fact a member of the AEP family, distantly related to the archaeal AEPs ORF904 (pRN1) and Rep (pIT3) (Iyer et al., 2005). However, it seems that, unlike the archaeal plasmid AEPs, Rep (ColE2) functions solely as an RNA primase, rather than as a DNA primase-polymerase. This enzyme, in addition to DNA polymerase I, is required for ColE2 DNA replication *in vitro*. Rep (ColE2) binds specifically to the plasmid's origin of replication where it initiates synthesis through the generation of a short RNA primer, allowing DNA polymerase I to subsequently proceed with DNA replication (Beck and Lipps, 2007). Thus, it appears that Rep (ColE2) functions as a plasmid-specific bacterial primase.

Another bacterial plasmid, RSF1010, found in a broad host range of over Gram-negative and some Gram-positive bacteria, encodes three Rep proteins; RepA, a helicase, RepB',

an AEP primase and RepC, a replication initiator protein (Scherzinger et al., 1984, 1991). RSF1010 contains two primase recognition sites *ssiA* and *ssiB*, each of which are recognised by RepB' allowing the independent synthesis of two primers that can then be extended by the host DNA polymerase III. The crystal structure of RepB' revealed the presence of two distinct domains; a large N-terminal domain containing two anti-parallel β-sheets flanked by six α-helices, and a smaller C-terminal region with a bundle of five α-helices. Notably, the enzyme lacks a zinc-binding motif (Figure 1.10.) (Geibel et al., 2009). This structure reveals a strong similarity between the N-terminal domain of RepB' and the catalytic domain of *P. furiosus* PriS. However, these enzymes share limited sequence homology, in addition to differences in ssDNA template recognition and in their requirements for priming. The structure of the catalytic core of RepB' bound to a *ssiA* recognition site has provided significant insights into DNA recognition by these primases (Figure 1.14.C.) and suggested a mechanism for initiation of plasmid DNA replication (Hines and Ray, 2011). Interestingly, RepB' displays a high degree of thermostability, presumably a result of its structural similarity to primases of the thermophilic archaea, raising interesting questions about the evolutionary origins of the RSF1010 plasmid (Geibel et al., 2009).

The two bacterial plasmid AEPs discussed here, therefore, stand in contrast with those of archaea. The bacterial Rep (ColE2) and RepB' enzymes represent prototypical AEPs, employed purely in initiating replication through synthesis of a short RNA primer. In contrast, the archaeal plasmid AEPs are proficient primase-polymerases, able to initiate and proceed with bulk replication of their host plasmid DNA. The conservative primase ability and lack of polymerase activity exhibited by these bacterial primases should not, however, be thought typical of all bacterial AEPs. The *Bacillus cereus* genome encodes *Bc*MCM (mini-chromosome maintenance), an AEP/MCM primase/helicase from an integrated prophage. *Bc*MCM was originally identified through BLAST analysis as an MCM homologue, with an N-terminal region of weak homology to AEPs (Figure 1.11.) (McGeoch and Bell, 2005). Initial biochemical studies identified 3'-5' helicase and ssDNA-stimulated ATPase activity, but also noted the absence of any primase activity (Samuels et al., 2009). However, a more recent structure/function study was able to detect not only helicase activity, but also primase and DNA-dependent DNA polymerase activities (Figure 1.13.) (Sanchez-Berrondo et al., 2012). Interestingly, like many archaeal AEPs, *Bc*MCM displays a strong preference for dNTPs during primer synthesis and extension. Together, these findings suggest that BcMCM may act as an important multi-functional enzyme, potentially being deployed in special circumstances during *B. cereus* DNA replication, such as the re-initiation of leading strand replication following

fork stalling. Importantly, *Bc*MCM is not the only bacterial AEP with an unconventional cellular role, as discussed in section 1.6.6. multifunctional AEPs are also required for DNA DSB, and probably other, repair processes in most bacterial species (Della et al., 2004).

1.6.5. Viral AEPs Involved in DNA Replication

Many of the AEPs distributed across the bacterial, archaeal, and eukaryotic genomes appear to have viral origins. Indeed, many viruses encode their own AEPs including, UL52-like primases from herpes simplex viruses, D5-like primases from NCLDVs and Lef-1 primases from phage and baculoviruses (Iyer et al., 2005). As is the case for cellular AEPs, viral AEPs also fulfil a number of key roles in DNA metabolism, particularly during replication.

Perhaps the most well studied of the viral AEPs is the UL5-UL8-UL52 heterotrimeric primase-helicase complex found in the herpes simplex virus family (Crute et al., 1989). Originally identified in herpes simplex virus type 1 (HSV-1), a large double-stranded DNA virus, the UL5-UL8-UL52 complex is encoded by three of the seven genes essential for replication of the HSV-1 chromosome (Crute and Lehman, 1991). Of these three proteins, UL52 was identified as the AEP responsible for priming DNA replication (Crute and Lehman, 1991; Klinedinst and Challberg, 1994), UL5 has helicase activity (Gorbalenya et al., 1989) and UL8 appears to be required for utilisation of primers by the UL30/UL42 polymerase. However, UL8 is dispensable for helicase and primase activity of UL5/UL52 (Marsden et al., 1997; Sherman et al., 1992). Where most primases have a zinc-binding motif in their catalytic domains (Augustin et al., 2001; S.-H. Lao-Sirieix et al., 2005; Lipps et al., 2004), UL52 has a strand-rich zinc finger domain that is located separately at its C-terminus (Figure 1.11.) and is absolutely required for primase activity *in vitro* (Biswas and Weller, 1999). The UL52 primase is capable of producing ribonucleotide primers of ~8-12 nucleotides in length, which are critical for initiating replication of the 153 kilobase viral genome (Crute and Lehman, 1991).

Another group of large viruses encoding AEPs are the poxviruses, which includes smallpox, that undertake DNA replication in the cytoplasm of infected cells (Moss, 2007). Most studies of the poxviruses have focussed on vaccinia virus (VACV), which possesses D5, an AEP/helicase fusion protein (Figure 1.11.) (Silva et al., 2007). This enzyme has a C-terminal domain belonging to the helicase superfamily III and an N-terminal region with sequence and structural features similar to AEPs (Silva et al., 2007). The N-terminal AEP domain of D5 is essential for viral replication in VACV-infected cells (Silva et al., 2007). In addition, this enzyme exhibits primase activity *in vitro* and stringent

template specificity, strongly suggesting a priming role for this enzyme in VACV DNA replication (Figure 1.13.) (De Silva et al., 2009; Silva et al., 2007). Based on extensive *in silico* analysis, Iyer *et al*. grouped the D5-like primases of poxviruses, irdoviruses, mimivirus, and African swine fever virus with the herpes simplex virus primases and their eukaryotic homologues, including eukaryotic PrimPol (Iyer et al., 2005). These enzymes, in addition to the A468R-like proteins of phycodnaviruses, make up the NCLDV-herpesvirus clade of AEPs. However, it should be noted that not all viral AEPs belong to this primase clade.

Unlike the UL52 herpesvirus and D5-like poxvirus AEPs, Lef-1-like primases of baculoviruses represent a family of AEPs that are more closely related to the replicative and NHEJ AEPs (see section 1.6.6.) that collectively form the AEP-proper clade (Iyer et al., 2005). Strikingly, the Lef-1-like primases of baculoviruses have the capacity to synthesise RNA primers that are extended by up to several kilobases in length (Mikhailov and Rohrmann, 2002) (Figure 1.13.). This ability is in line with the extension activities of the archaeal replicative primase PriS from *Pyrococcus*, supporting the fact that these enzymes belong to the same AEP clade. However, it has been suggested that the extension capabilities of Lef-1-like primases may be limited by other replication factors *in vivo* (Mikhailov and Rohrmann, 2002). Nevertheless, this ability raises the possibility that these enzymes may play additional roles in primer extension.

In contrast to the RNA primase activities of the viral AEPs discussed above, the gp43-like proteins of corynephage BFK20, do not share this rNTP incorporation preference. Instead, the gp43-like proteins, part of the Prim-Pol clade of AEPs that includes ORF904 and Rep(pIT3), can only incorporate dNTPs (Halgasova et al., 2012). In addition, the gp43-like proteins, similar to the archaeal AEPs, display both DNA primase and polymerase activities (Halgasova et al., 2012). Thus, showing that even within viruses, AEPs form a diverse group of enzymes with varying catalytic capabilities and potentially divergent roles.

1.6.6. Primases Involved in DNA Double-Strand Break Repair

Around the time that archaeal primases were first reported to also be template-dependent polymerases (Bocquier et al., 2001), AEP orthologues were unexpectedly identified in prokaryotic genomes (Aravind and Koonin, 2001; Koonin et al., 2000; Weller and Doherty, 2001). These AEP genes were frequently found to be co-operonic with Ku (Aravind and Koonin, 2001; Doherty et al., 2001; Weller and Doherty, 2001), a protein responsible for binding the ends of DNA DSBs during NHEJ in eukaryotes. These findings provided early clues that a conserved NHEJ pathway may exist in prokaryotes

and that AEPs may be intrinsically involved in this DSB repair process. Subsequent studies identified that a *bona fide* NHEJ DSB repair apparatus exists in bacteria (Della et al., 2004; Gong et al., 2005; Pitcher et al., 2007; Weller et al., 2002) and that these AEPs form part of a larger multi-protein repair complex known as ligase D (LigD). More recently, a closely related NHEJ apparatus has also been identified in some archaeal species (Bartlett et al., 2013). In mycobacteria, LigD is a fusion protein composed of AEP, nuclease and ligase domains (Della et al., 2004; Pitcher et al., 2005). However, in many species these "domains" exist as individual co-operonically expressed proteins that form a functional NHEJ complex (Bartlett et al., 2013). Prokaryotic NHEJ is therefore thought to be essentially facilitated by a Ku-LigD complex that possesses all of the activities required to bind to the break termini and catalyse re-joining of DSBs (Bartlett et al., 2013; Della et al., 2004; Gong et al., 2005; Pitcher et al., 2007). NHEJ AEPs are capable of performing an astonishing range of nucleotidyltransferase activities, presumably to accommodate the myriad of end configurations produced during formation of DSBs. Specifically, these enzymes can catalyse template-dependent DNA/RNA polymerase, terminal transferase, strand-displacement and gap-filling synthesis, with a notable preference to incorporate ribonucleotides (Bartlett et al., 2013; Della et al., 2004; Pitcher et al., 2007). In addition, these AEPs can readily extend mismatched primer-template termini and perform TLS bypass of 8-oxo-dG lesions and Ap sites (Pitcher et al., 2007) (Figure 1.13.).

Since the unexpected discovery that AEPs function as components of the DSB repair machinery in bacteria and archaea, there has been much conjecture about why members of the primase family evolved to become the primary NHEJ polymerases. It is likely that these bespoke repair enzymes, which belong to the AEP proper clade that also includes the replicative primases, evolved from a primordial AEP with an innate capacity to make short RNA primers into a novel class of adaptable end-joining polymerases capable of processing DNA ends during break repair. A comparison of the sequences and structures of the NHEJ AEPs with the replicative enzymes (PriS), reveals that whilst both share a common catalytic architecture (Figure 1.10.), there are several distinctive adaptations. NHEJ AEP polymerases possess a number of distinctive DNA binding modes that distinguishes them from related enzymes, enabling them to operate even at the extreme ends of DNA. These enzymes possess a positively charged surface pocket that enables them to bind specifically to a 5' phosphate, either close to or at the terminus of a DSB, thus stably tethering the enzyme to the break to permit end-processing (Figure 1.14.A.). In addition, they have also evolved prominent surface loops (Loops 1 and 2) that facilitate a remarkable ability to promote break synapsis (Figure 1.14.B.), a process

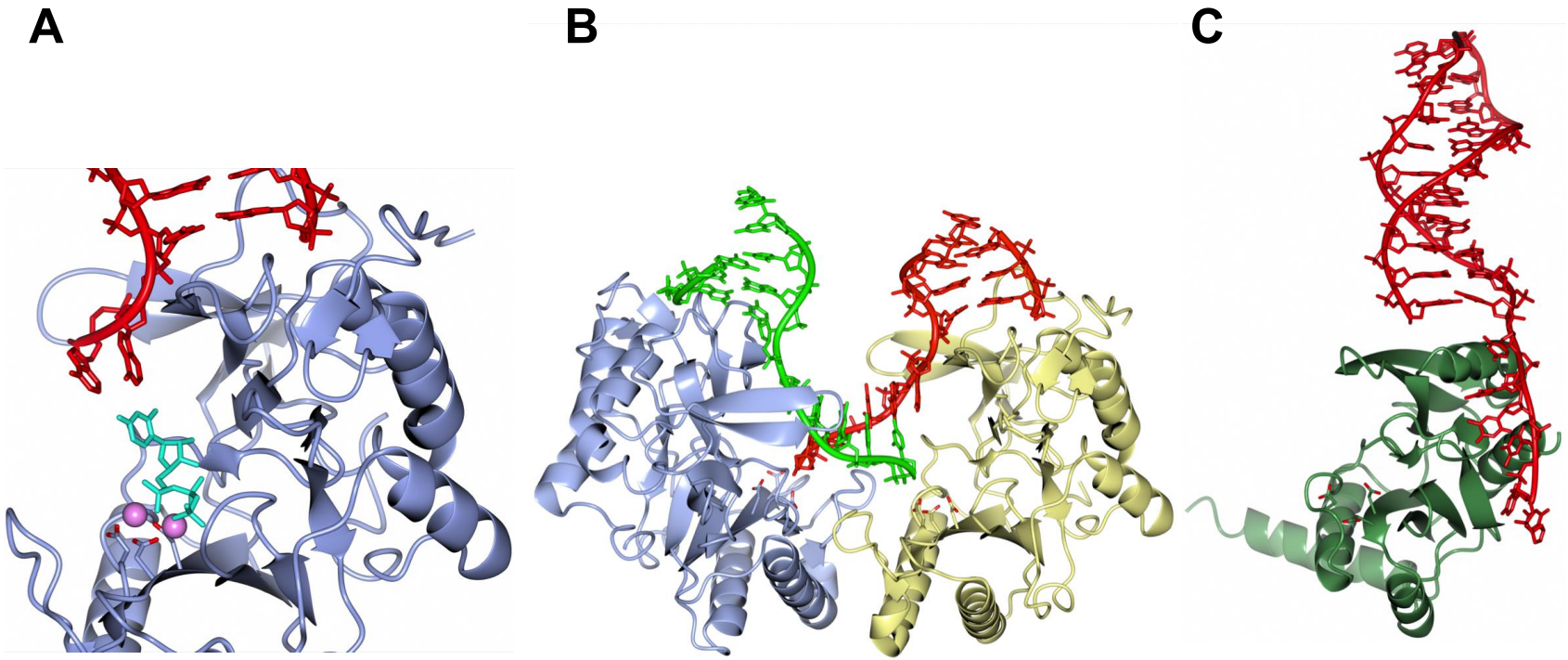


Figure 1.14. Structures of AEPs bound to DNA substrates.
Structural examples of AEP members bound DNA intermediates. **(A)** Structure a pre-ternary catalytic conformation of a NHEJ repair polymerase (PolDom / LigD-Pol) from *M. tuberculosis* bound to a dsDNA break with a 3' overhanging terminus (PDBID: 4PKY, ice blue) and UTP and manganese cofactors, coloured cyan and pink, respectively. **(B)** Crystal structure of a micro-homology mediated end-joining (MMEJ) intermediate showing an NHEJ repair polymerase-mediated synapsis of a DSB (PDBID: 4MKY, ice blue and lemon). **(C)** Structure of the AEP domain of RepB' bound to a *ssiA* DNA replication initiation site (PDBID: 3H25, lawn green). The catalytic residues are rendered as sticks with the acidic oxygens coloured red. DNA strands are coloured red or green.

that permits breaks to be annealed back together by a mechanism known as microhomology-mediated end joining (MMEJ). During this process, each side of the break is first bound by a polymerase, forming a pre-ternary complex in anticipation of receiving the other end of the break. Subsequently, the surface loops, conserved only in these AEPs, act in concert to “present” one end of the DSB to the other side in order to promote and accelerate break annealing. In the case of the 3’ overhangs, this process configures the break to allow productive gap-filling synthesis to occur *in trans* (Brissett et al., 2007, 2011, 2013). This mechanism also provides a molecular basis for the template-dependent terminal transferase synthesis catalysed by these enzymes at the 3’ ends of DNA. Although these unprecedented MMEJ processes were initially considered by some to be specific to these polymerases, an analogous polymerase-mediated MMEJ mechanism has since been reported for archaeal PriS, mammalian Pol θ, and terminal transferase (TdT), suggesting that this is a functionally conserved mechanism (Gouge et al., 2015; Hu et al., 2012; Kent et al., 2015).

Although the biological roles of the NHEJ and replicative AEPs are clearly distinct, these enzymes are closely related, belonging to the same clade, and therefore are likely to share common features despite their divergent evolution. The crystal structures of a number of catalytic intermediates of mycobacterial NHEJ AEPs bound to DNA have provided some important insights into the shared common catalytic mechanisms of AEPs and also explained why these enzymes may be suited to the task of break repair (Brissett et al., 2007, 2011, 2013). The structure of a preternary NHEJ AEP-DNA complex has revealed that these enzymes, in common with polymerases, employ a two metal mechanism of catalysis, with binding of the second metal dependent on engagement of the incoming nucleotide with both the active site and template strand (Figure 1.14.A.). As discussed, these AEPs have the ability to “accept” an incoming primer strand provided *in trans* by an adjacent AEP pre-ternary complex (Figure 1.14.B.). Significant in this regard, these enzymes can bind to and extend an incoming primer as short as a dinucleotide (Brissett et al., 2011). This mechanism is highly reminiscent of the initiation step performed by replicative primases. Here, a binary complex between the enzyme and ssDNA is formed first, followed by binding of the 3´nucleotide to form a pre-ternary complex. This is followed by recruitment of a 5´-nucleotide, which acts as the “primer”, to form a ternary complex. Notably, both NHEJ and replicative AEPs can catalyse an unconventional addition of a ribonucleotide in the 3´-5´ direction, followed by a more conventional 5´-3´elongation step. This innate ability of AEPs of this clade to accept short primers may explain why they were the most appropriate enzymes to evolve further, by

the acquisition of additional surface loops and phosphate binding residues, into highly effective NHEJ repair polymerases.

1.6.7. Primases Involved in DNA Damage Tolerance

In addition to the apparent absence of Pol α homologues, many archaeal species also lack canonical TLS polymerases, with only a subset of species possessing Y-family polymerases. Furthermore, many archaea do not encode NER or photolyase pathways to remove potential replication fork stalling UV-light induced DNA damage (Kelman and White, 2005). This raises the question as to how archaeal species lacking these pathways tolerate DNA damage, which is of paramount importance given the extreme environments in which many of these species reside. Recently, it was reported the replicative primase, PriS, from the Y-family deficient archaeal species (*A. fulgidus* and *P. furiosus)* is inherently damage tolerant (Figure 1.13.) (Jozwiakowski et al., 2015). Strikingly, it was identified that PriS from these organisms is capable of faithfully bypassing highly DNA-distorting CPDs, in addition to 8-oxo-dG. The extreme environments inhabited by thermophilic archaeal species generate significant amounts of cytosine deamination, generating uracil base adducts that induce profound fork stalling when encountered by the archaeal replicases (B- and D-family polymerases). Notably, PriS also replicates past templating uracil bases, even when stalled replicative polymerases are bound (Jozwiakowski et al., 2015), suggesting that PriS assists the replisome in maintaining active fork progression during genome duplication. These findings further corroborate that archaeal replicative primases, in addition to primer synthesis, play additional roles in replication. Likewise, PrimPol, the focus of this thesis, is also tolerant to duplex-distorting DNA lesions, as will be discussed later in this introduction (see section 1.7.).

1.6.8. Essential Roles for Multiple AEP Orthologues in Trypanosomes

Kinetoplastids are a group of single-celled protozoa characterised by the presence of kinetoplasts, networks of circular DNA found inside a large single mitochondrion and composed of both maxi-circles (20-40 kb) and mini-circles (0.5-1 kb) (Shapiro and Englund, 1995). Within each organism, this network of kinetoplast DNA must be duplicated prior to division. One particularly well studied kinetoplastid protozoan is *Trypanosoma brucei*, the causative agent of human African trypanosomiasis (Brun et al., 2010). Until recently, the kinetoplastid replication machinery of *T. brucei* could not be reconstituted as the primase involved in kinetoplast replication had not yet been identified. Two such primases have since been characterised, PRI1 and PRI2, responsible for maxi-circle and mini-circle replication initiation, respectively (Hines and

Ray, 2010, 2011). These primases belong to the NCLDV herpesvirus clade of AEPs, and each contains an RNA recognition motif and a PriCT-2 motif. In each case, RNAi depletion of these enzymes in *T. brucei* causes inhibition of cell growth and the depletion of kinetoplast DNA, clearly suggesting that these AEPs fulfil vital roles in priming kinetoplast DNA replication (Hines and Ray, 2010, 2011).

Recently, two PrimPol-like orthologues (PPLs), referred to as PPL1 and PPL2, have been identified in *T. brucei* (Rudd et al., 2013). PPL1 is capable of synthesising RNA primers up to 50 nucleotides in length on poly(dT) templates. In contrast, PPL2 does not appear to exhibit any primase activity, representing another example where an AEP has ceased to function as a primase (Figure 1.13.). However, both PPL1 and PPL2 possess damage tolerance synthesis activities, specifically an ability to bypass 6-4PPs and 8-oxo-dG lesions. Perhaps most surprising was the finding that PPL2 is essential for cell survival. Knockdown of PPL2 results in cell cycle arrest following bulk genome duplication. These cells accumulate a lot of DNA damage and die in a pre-mitotic phase. It has been proposed that PPL2 functions as a TLS polymerase that assists in restarting replication downstream of damage or DNA structures during the completion of genome duplication in G_2 (Rudd et al., 2013). This inability to bypass damage likely leads to the generation of double-strand breaks observed when PPL2 is knocked down. The existence of PPL2 is probably a result of the duplication and subsequent diversification of PPL1 to remedy DNA replication issues specific to trypanosomes and other protists, such as the replication of repetitive sequence elements. These examples again demonstrate how AEPs have diversified to fill a range of roles in both nuclear and mitochondrial DNA metabolism.

1.6.9. Why do Primases Prefer to Incorporate RNA into DNA?

A common feature of many AEP enzymes, including replicative primases, is their marked preference to incorporate NTPs, rather than dNTPs, into the synthesised strand. A pertinent question here is why these specialized polymerases have maintained this preference to prime replication or repair damaged DNA by synthesising RNA, which is much less stable due to the presence of a 2’ OH moiety that makes the sugar much more prone to hydrolysis. This is particularly surprising in the case of mammalian DNA replication, where the replicative primases incorporate lots of RNA into the newly synthesised DNA that must then be excised and replaced before genome duplication is completed.

A shared structural feature of all AEP-related primases is their open and malleable active sites that, unlike canonical polymerases, enables them to accommodate a wide variety

of DNA configurations including: ss / ds DNA, mismatches, lesions and even termini of DSBs (Figure 1.10. and 1.14.). However, the price to be paid for this catalytic flexibility is low fidelity. To illustrate this point in more detail let us examine the NHEJ AEPs. These enzymes are highly adaptive polymerases that have effectively lost their primase activity and evolved to accommodate a wide range of DNA configurations in their active sites and perform a extensive variety of extension activities to ensure that DSB are repaired, irrespective of the nature of the break. However, they also preferentially incorporate NTPs to fill in any gaps with RNA, which are then preferentially ligated to seal the breaks (Bartlett et al., 2013). Why do these enzymes prefer to fill the gaps with DNA instead of RNA? The likely explanation for this preference is because of the very low fidelity exhibited by these polymerases. By incorporating RNA into the repaired breaks, the enzyme is "flagging up" the bases that it has incorporated. Once DSB repair has been completed by NHEJ, ribonucleases (e.g. RNase H2) can then excise the RNA, which can then be replaced with DNA by more accurate patch repair polymerases. In the case of DNA replication, primer synthesis is also a highly inaccurate process and given the large number of regions where RNA is incorporated into the genome, particularly on the lagging strand, this would result in the introduction of a very high mutagenic load during every round of replication. To prevent this from occurring in cells it is likely that, as during NHEJ, RNA is preferred for primer synthesis to demarcate regions of low fidelity synthesis that are subsequently excised and replaced with DNA in a more faithful synthesis process that occurs before the completion of genome synthesis.

1.6.10. Primase-Polymerases – Initiating Replication is Only the Beginning

Since the first discovery of replicative DNA polymerases, well over half a century ago, it has been fully appreciated that there are also a more diverse range of families within this general classification, giving rise to many additional enzymes that have distinct roles in a wide variety of nucleic acid metabolism processes in cells, from replication to transcription. In contrast, the possibility that DNA primases may also have additional members and roles in cells has largely been overlooked until relatively recently. The reasons for this are partially down to their name, which categorically assigns a sole function to the bespoke nucleotidyltransferase activity associated with the first identified replicative primases, thus deterring significant further investigation of additional activities and functions associated with this grouping of enzymes. Whilst this primase label appropriately describes the *de novo* synthesis of RNA primers associated with toprim-related DnaG primases and a small subset of the AEP-proper clade (Prim1), in most cases it is a misnomer that does not adequately describe the function of the vast majority of members that make up this superfamily (Figure 1.15.). AEP primases and canonical

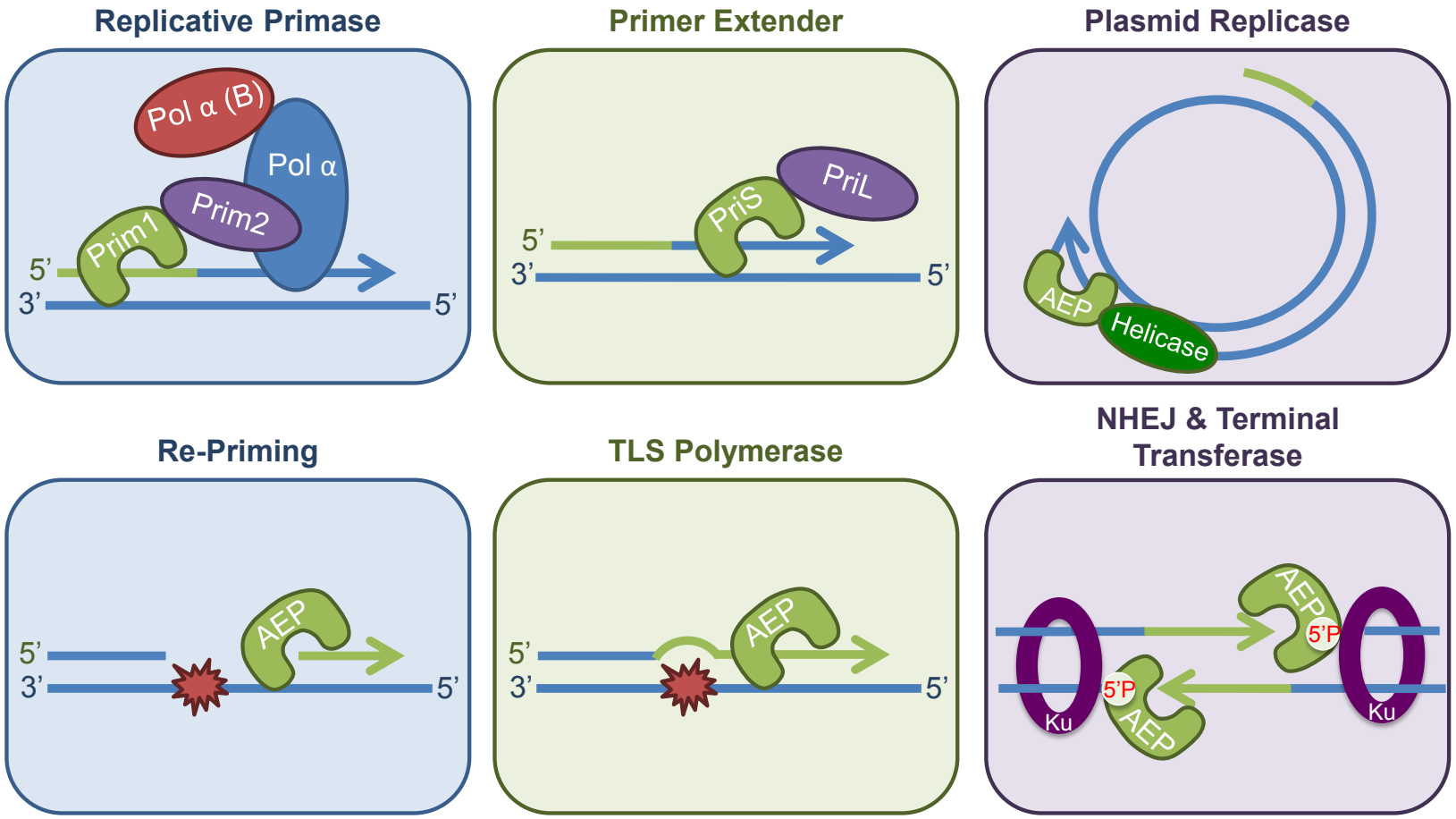


Figure 1.15. Diversity of functional roles fulfilled by AEPs.
AEP superfamily members are employed in many different biological roles, in addition to replicative primases. These enzymes are also utilised as primer extenders, plasmid replicases, damage-tolerance re-priming enzymes, TLS polymerases, NHEJ DNA break repair and terminal transferase polymerases.

DNA polymerases most likely evolved from a common ancestral nucleic acid recognition domain (Iyer et al., 2005), the RRM, and share a number of common catalytic features, including divalent metal-dependent catalysed extension of nucleic acids in a 5'-3' direction. Therefore, it is more appropriate that all members of the AEP primase superfamily should be considered as belonging to the broad general grouping of enzymes called polymerases and, within this umbrella term, be further sub-classified as belonging to a sub-group of enzymes called Primase-Polymerases (Prim-Pols) to reflect their dual origins and, in most cases, their capacity to perform both synthesis functions (Figure 1.13.). This term has already gained acceptance to describe a number of different microbial and eukaryotic AEPs involved in a diverse range of functions, including plasmid replication, lesion bypass, and repriming.

Although many additional roles have been identified for Prim-Pols in all domains of life over the last decade or so (Figure 1.15.), much more remains to be discovered about the diverse functions and pathways in which these highly adaptable enzymes operate. For example, given that they can perform both priming and template-dependent synthesis events, it is likely that Prim-Pols undertake roles in other key cellular pathways, such as restart and bypass mechanisms associated with replication fork stalling at structural impediments. Given the propensity of many Prim-Pols to incorporate RNA during synthesis, it is also likely that they may also be involved in a range of transcriptional-related processes. AEP members are associated with some CRISPR operons suggesting potential roles in other processes, such as viral "immunity" in microbes. These are just some potential examples for novel AEP functions and many more await to be discovered.

In addition to understanding the myriad of different functions associated with Prim-Pols, the identification of novel AEPs also provides an opportunity to explore potential associations between these enzymes and human disease. Critically, many Prim-Pols are also important for the life cycles of pathogenic bacteria, viruses and protozoa that infect mammals and therefore these enzymes are potentially attractive targets for the development of novel anti-microbial agents that specifically target essential DNA replication and repair pathways in these organisms.

1.7. PrimPol

Until recently, the only AEP member identified in higher eukaryotes was the Pol α-associated, PriS homologue, Prim1. However, a second eukaryotic AEP has now been described and characterised, this enzyme is named PrimPol (alternative names CCDC111, FLJ33167, EukPrim2 or hPrimPol1). PrimPol was originally identified as a

novel uncharacterised member of the NCLDV-herpes virus clade of AEPs (Iyer et al., 2005). PrimPol orthologues are present across a diverse range of unicellular and multicellular eukaryotes including species of animals, plants, and primitive early eukaryotes, such as fungi, protists, and algae. However, the enzyme is notably absent from a number of eukaryotes including *Caenorhabditis elegans* and *Drosophila melanogaster*. Importantly, this interrupted distribution suggests that PrimPol was acquired through horizontal gene transfer from a viral source and subsequently lost on a number of independent occasions (Iyer et al., 2005). PrimPol, similar to *A. fulgidus* and *P. furiosus* PriS*,* possesses both primase and DNA damage tolerance polymerase activities (Figure 1.13.) (Bianchi et al., 2013; García-Gómez et al., 2013; Keen et al., 2014b; Mourón et al., 2013; Rudd et al., 2013; Wan et al., 2013). Specifically, PrimPol is capable of bypassing UV-induced 6-4PPs, in addition to 8-oxo-dG lesions (Bianchi et al., 2013; García-Gómez et al., 2013). The enzyme is composed of two domains, an N-terminal AEP domain, consisting of three canonical catalytic motifs, and a C-terminal UL52-like zinc finger, required for primase but not polymerase activity and conserved across the NCLDV-herpesvirus clade of AEPs (Figure 1.10.) (Iyer et al., 2005).

PrimPol is required for the maintenance of replication fork progression, with PrimPol knockout cells displaying decreased forks rates and increased sensitivity to DNA damaging agents (Bianchi et al., 2013). Furthermore, PrimPol has been implicated in the restart of stalled replication forks (Keen et al., 2014b; Mourón et al., 2013; Wan et al., 2013), involving either TLS or repriming downstream of damage. As well as its involvement in nuclear DNA synthesis, a substantial proportion of PrimPol also localises to the mitochondria where it is thought to aid in the replication of the small circular mitochondrial genome, potentially by assisting 8-oxo-dG bypass or through repriming (Bianchi, 2013; García-Gómez et al., 2013). In line with this, PrimPol-knockdown cells have been reported to exhibit mitochondrial DNA defects (Bianchi, 2013; García-Gómez et al., 2013).

Coupled with the increasing evidence pointing towards a conserved repriming mechanism across all domains of life, there are some suggestions that PrimPol may fulfil this role in higher eukaryotes. Importantly, the dispensability of the zinc finger (ZnF) domain of PrimPol for polymerase and TLS activities, coupled with its strict requirement for primase activity, has allowed separation of function studies to be performed *in vivo* (Keen et al., 2014b; Mourón et al., 2013). These studies revealed that an intact ZnF domain is required for PrimPol to maintain normal replication fork rates following UV damage, suggesting that the primase activity of the enzyme is necessary for replication restart. Importantly, the potential involvement of PrimPol in repriming DNA replication

begs the question as to why the replicative primase Prim1 cannot also be employed in this role? It seems likely that, at least for lagging strand synthesis, this will be the case. Here, repriming would simply require the initiation of a new Okazaki fragment downstream of the lesion, facilitated by Prim1-dependent repriming. However, the requirement for PrimPol to reprime synthesis on the leading strand may be of more importance due to its preference to utilise dNTPs during primer synthesis, thus preventing incorporation of RNA. This offers a distinct advantage over Prim1 as it eliminates the possibility of RNA processing or hydroysis that would lead to formation of breaks on the leading strand that would eventually result in potentially lethal DSBs.

The overall aim of the articles presented in this thesis was to build upon the initial characterisation of PrimPol, with a specific focus on the mechanisms controlling recruitment and regulation. To this end, PrimPol's interaction with recruitment and regulatory partners was investigated using both cellular and biochemical approaches. The first of the articles focusses on the regulation of PrimPol's enzymatic activities by both RPA and mtSSB, in addition to the characterisation of the enzyme's fidelity as a DNA polymerase (Chapter 2). The second article investigates the effect of a novel PrimPol interacting partner, Polymerase δ-interacting protein 2 (PolDIP2), on the enzyme's primer extension activities using *in vitro* and *in vivo* approaches (Chapter 3). Chapter 4 describes the development of a fluorescent primase assay and demonstrates how this technique was used to investigate PrimPol's repriming activity in two further reports. The final results chapter provides an in depth characterisation of PrimPol's interaction with RPA and proposes a mechanism for the recruitment and regulation of the enzyme at stalled replication forks (Chapter 5). These articles are presented chronologically and thus in the context of the wider PrimPol literature at the time they were published. In the final chapter, a review and discussion of the current PrimPol literature is provided, giving an overview and placing the work presented here in the wider context of this growing body of knowledge (Chapter 6). Together, the articles presented here build upon the initial characterisation of the enzyme and strongly support a role for PrimPol as the facilitator of leading strand repriming in higher eukaryotes, following recruitment and regulation by both RPA and PolDIP2.

Chapter 2

Human PrimPol is a Highly Error-Prone Polymerase Regulated by Single-Stranded DNA Binding Proteins

2.1. Abstract

PrimPol is a recently identified polymerase involved in eukaryotic DNA damage tolerance, employed in both repriming and TLS mechanisms to bypass nuclear and mitochondrial DNA lesions. In this report, we investigate how the enzymatic activities of human PrimPol are regulated. We show that, unlike other TLS polymerases, PrimPol is not stimulated by PCNA and does not interact with it *in vivo*. We identify that PrimPol interacts with both of the major single-strand binding proteins, RPA and mtSSB *in vivo.* Using NMR spectroscopy, we characterize the domains responsible for the PrimPol-RPA interaction, revealing that PrimPol binds directly to the N-terminal domain of RPA70. In contrast to the established role of SSBs in stimulating replicative polymerases, we find that SSBs significantly limit the primase and polymerase activities of PrimPol. To identify the requirement for this regulation, we employed two forward mutation assays to characterise PrimPol's replication fidelity. We find that PrimPol is a mutagenic polymerase, with a unique error specificity that is highly biased towards insertion-deletion errors. Given the error-prone disposition of PrimPol, we propose a mechanism whereby SSBs greatly restrict the contribution of this enzyme to DNA replication at stalled forks, thus reducing the mutagenic potential of PrimPol during genome replication.

2.2. Introduction

Accurate and efficient DNA replication is essential for the maintenance of genomic integrity. The replication machinery is a highly specialised multi-protein complex employed for this purpose, with the replicative DNA polymerases, Pol α, Pol δ, and Pol ε, tasked with the majority of bulk DNA synthesis in the eukaryotic nucleus. In mitochondria, this task is undertaken by Pol γ. These enzymes are superbly adapted to maximise faithful DNA synthesis, however this high degree of specialisation comes at a cost. Helix-distorting DNA lesions and structures, which persist into the S-phase of the cell cycle, present an obstacle to the replicative polymerases, causing stalling of the replication fork (Aguilera and Gómez-González, 2008). In response, cells employ a variety of DDT mechanisms to facilitate lesion/structure bypass and permit continued replication (Li and Heyer, 2008; Sale et al., 2012).

Mechanisms of replication restart include HR, in which an alternative sister template permits extension of the stalled primer terminus (Li and Heyer, 2008; Sale et al., 2012). Firing of dormant origins, discontinuous generation of Okazaki fragments on the lagging strand, or repriming of the replication fork downstream of the lesion on the leading strand, can also restart stalled forks (Heller and Marians, 2006). An alternative mechanism is TLS. Here, specialised DNA polymerases, predominantly of the Y-family, rescue stalled replication forks by directly synthesising across the damaged nucleotides. In contrast to replicative DNA polymerases, TLS polymerases display low fidelity during replication of undamaged DNA templates, thus requiring strict regulation (Sale et al., 2012). The primary level of regulation for TLS polymerases comes with their inherent distributive character. Additional regulation of access to the replisome is proposed to occur, in part, through post-translational modification of the PCNA (Friedberg et al., 2005). Collision of the replication fork with DNA lesions, and consequent stalling, stimulates mono-ubiquitylation of PCNA, increasing its affinity for TLS polymerases, thus promoting recruitment of these enzymes to the fork. Following bypass of the lesion, the TLS polymerase dissociates and the high fidelity replicative polymerases proceed with DNA synthesis (Friedberg et al., 2005). The polymerase switch, therefore, acts to limit DNA replication by the low fidelity TLS polymerases, permitting access only when lesion bypass is required. Recently, a novel primase-polymerase called PrimPol has been reported to be involved in DDT and TLS during both nuclear and mitochondrial DNA replication (Rudd et al., 2014).

PrimPol is a eukaryotic DNA primase-polymerase, belonging to the AEP superfamily, that undertakes lesion bypass roles in both nuclear and mitochondrial DNA replication

(Bianchi et al., 2013; García-Gómez et al., 2013; Rudd et al., 2013, 2014; Wan et al., 2013). This enzyme is capable of synthesising primers using either rNTPs or dNTPs, conferring the ability to reprime and restart replication downstream of DNA lesions. PrimPol also possesses robust template-dependent DNA polymerase activity, which it can utilise to bypass highly distorting 6-4 PPs and oxidative lesions, including the common 8-oxo-dG lesion, thus establishing PrimPol as a competent TLS polymerase (Bianchi et al., 2013; García-Gómez et al., 2013; Rudd et al., 2013; Wan et al., 2013). PrimPol possesses two distinct domains; an enzymatic AEP polymerase domain required for the catalytic activities of the enzyme and a UL52-like ZnF domain necessary for primase activity and modulating the processivity and fidelity of the enzyme (Keen et al., 2014b). PrimPol knockout ($PrimPol^{-/-}$) cells display increased sensitivity to DNA damaging agents and decreased replication fork rates (Bianchi et al., 2013), in addition to defects in mtDNA replication (García-Gómez et al., 2013). Furthermore, $PrimPol^{-/-}$ mouse embryonic fibroblasts (MEFs) have increased metaphase aberrations and chromatid breaks, increasing substantially following low-dose aphidicolin treatment (Bianchi et al., 2013). In trypanosomes, deletion of a PrimPol orthologue leads to growth arrest in G_2 followed by cell death (Rudd et al., 2013). Recent studies have established the involvement of PrimPol in DDT through at least two mechanisms, repriming and TLS. However, the regulation of PrimPol's contribution to DNA replication has not previously been explored.

In this report, we describe how the enzymatic activities of PrimPol are regulated. We observe that, in contrast to other TLS polymerases, PrimPol does not interact with PCNA and is not stimulated by its presence *in vitro.* Pull-down studies identify that human PrimPol interacts with RPA, the nuclear SSB, and its mitochondrial equivalent, mtSSB. We find that PrimPol interacts with the N-terminal domain of the RPA70 subunit (RPA70N), RPA has previously been shown to stimulate the activity of Pol α and Pol δ (Braun et al., 1997; Tsurimoto and Stillman, 1989), mtSSB also stimulates its respective polymerase partner, Pol γ (Oliveira and Kaguni, 2010). However, in stark contrast we demonstrate that both of these proteins act to significantly limit both the primase and polymerase activities of PrimPol. We demonstrate that PrimPol is an error-prone DNA polymerase, with a strong preference to generate base insertions and deletions, thus necessitating strict regulation during its involvement in DNA synthesis. We propose that SSBs potentially act to limit the contribution of PrimPol to DNA replication in order to limit error-prone synthesis during the bypass of lesions and other genetic obstacles.

2.3. Materials and Methods

2.3.1. Affinity Purification of PrimPol Complexes for Mass Spectrometry Analysis

For the large scale affinity purification of soluble human PrimPol for MS analysis, thirty 175 cm^2 flasks of confluent Flp-In T-REx-293 cells engineered for PrimPol expression were used, 1 day before harvesting, PrimPol expression was induced in 15 of these flasks by addition of 10 ng/ml doxycycline. Following harvesting and collection, cell pellets (~1 g each) were lysed in 15 ml lysis buffer (150mM NaCl, 30mM Tris pH 7.4, 0.5% NP-40, with Roche protease inhibitor cocktail) and incubated at 4°C on a rocker for 20 minutes. Input was retained (500 µl) and 1 ml of *Strep*-tactin resin (packed volume) added to the lysate and placed on a rocker for 2 hours at 4°C. Washes were also performed in batch mode (with lysis buffer containing 0.1% NP40) and then the *Strep*-tactin resin was transferred to a gravity flow column and washed further. Five successive 500 µl elutions with lysis buffer containing 2 mM desthiobiotin were performed, and each snap frozen with 10% glycerol. Following Western blot analysis to determine which affinity purifications were successful, the chosen elutions were concentrated using a VivaSpin 10,000 kDa molecular filter before resolving on a Bis-Tris 4-20% gel and colloidal Coomassie staining (Invitrogen). Whole lane gel extraction was performed with each lane being divided into 1-2 mm bands, which were placed in a 96-well plate before trypsin digestion and MS analysis.

2.3.2. Mass spectrometry

Polyacrylamide gel slices (1-2 mm) containing the purified proteins were prepared for MS analysis using the Janus liquid handling system (PerkinElmer, UK). Briefly, the excised protein gel pieces were placed in a well of a 96-well microtitre plate and destained with 50% v/v acetonitrile and 50 mM ammonium bicarbonate, reduced with 10 mM DTT, and alkylated with 55 mM iodoacetamide. After alkylation, proteins were digested with 6 ng/µL Trypsin (Promega, UK) overnight at 37°C. The resulting peptides were extracted in 2% v/v formic acid, 2% v/v acetonitrile. The digest was analysed by nano-scale capillary LC-MS/MS using a Ultimate U3000 HPLC (ThermoScientific, San Jose, USA) to deliver a flow of approximately 300 nL/min. A C18 Acclaim PepMap100 5 µm, 100 µm x 20 mm nanoViper (Thermo Scientific, San Jose, USA), trapped the peptides prior to separation on a C18 Acclaim PepMap100 3 µm, 75 µm x 150 mm nanoViper (Thermo Scientific Dionex, San Jose, USA). Peptides were eluted with a gradient of acetonitrile. The analytical column outlet was directly interfaced via a modified

nano-flow electrospray ionisation source, with a hybrid linear quadrupolefourier transform mass spectrometer (LTQ Orbitrap XL/ETD, Thermo Scientific, San Jose, USA). Data dependent analysis was carried out, using a resolution of 30,000 for the full MS spectrum, followed by eight MS/MS spectra in the linear ion trap. MS spectra were collected with an automatic target gain control of $5x10^5$ and a maximum injection fill time of 100 ms over a m/z range of 300–2000. MS/MS scans were collected using an automatic gain control value of $4x10^4$ and a threshold energy of 35 m/z for collision induced dissociation. LC-MS/MS data were then searched against a protein database (UniProt KB) using the Mascot search engine programme (Matrix Science, UK) (Perkins et al., 1999). Database search parameters were set with a precursor tolerance of 5 ppm and a fragment ion mass tolerance of 0.8 Da. One missed enzyme cleavage was allowed and variable modifications for oxidized methionine, carbamidomethyl cysteine, pyroglutamic acid, phosphorylated serine, threonine and tyrosine were included. MS/MS data were validated using the Scaffold programme (Proteome Software Inc., USA) (Keller et al., 2002). All data were additionally interrogated manually.

2.3.3. Construction of Human PrimPol Mutants

Human PrimPol and $PrimPol^{1-487}$ were cloned as previously described (Keen et al., 2014b). $PrimPol^{24-354}$ was constructed by PCR using the following forward and reverse primers; FWD: GTTTCTTCATATGCGGTTGTCCTCAGTGATAGACC, REV: 5'-GTTTCTTGCGGCCGCGATACTGTTAAAATATCCAACC-3'.

2.3.4. Expression and Purification of Recombinant Proteins

Wild-type PrimPol and $PrimPol^{24-354}$ were expressed in *E. coli* SHuffle Express cells overnight at 16°C, proteins were then purified as previously described (Keen et al., 2014b). Human recombinant PCNA, Pol δ, RPA, and mtSSB proteins were expressed and purified as reported previously (Longley et al., 2009; Masuda et al., 2007). In addition, kTaq-Pol A and Tgo-Pol B (exo-) were purified as previously described (Engelke et al., 1990; Evans et al., 2000). Protein concentrations were determined based on absorbance at 280 nm corrected with the protein specific extinction coefficient. Extinction coefficient values for each of the recombinant proteins were calculated using the ProtParam tool (ExPASy). Phage T4 SSB and T4 polymerase were purchased from New England Biolabs. Pol γ^{exo-} was a kind gift from Dr Whitney Yin (University of Texas, USA).

2.3.5. Electrophoretic Mobility Shift Assays

Electrophoretic mobility shift assays (EMSAs) were performed for 60 min at room temperature in 20 µL volumes containing 50 mM Potassium acetate, 20 mM Tris-acetate pH 7.9, 10 mM Magnesium acetate, 1 mM DTT, 60 nM single stranded fluorescein labelled DNA (sequence 16, Table 2.1.), and varying concentrations of mtSSB or RPA (as indicated on Figure 2.1.). Reactions were supplemented with 2 µL 25% (w/v) Ficoll and resolved on a 5% (v/v) native polyacrylamide gel at 75 V for 60min in 0.5x TBE buffer. Gels were scanned using a FujiFilm FLA-5100 image reader.

2.3.6. NMR Methods

RPA70N and RPA32C were expressed and purified as described previously (Arunkumar et al., 2005; Frank et al., 2014). ^{15}N-^{1}H HSQC experiments were performed at 25°C on a Bruker Avance III 800 MHz NMR spectrometer with a cryogenically cooled probe. Spectra were acquired for samples of ^{15}N-enriched RPA32C or RPA70N alone and in the presence of full-length PrimPol or PrimPol$^{1\text{-}487}$. All samples were equilibrated in a buffer containing 20 mM HEPES, 80 mM NaCl, 2 mM DTT, and 5% deuterium oxide.

2.3.7. DNA Primase Assays

DNA primase assays were performed using the previously described protocol (Keen et al., 2014b), in buffer containing 50 mM Potassium acetate, 20 mM Tris-acetate pH 7.9, 10 mM Magnesium acetate and 1 mM DTT. The templating oligonucleotide sequence can be found as sequence 10 in Table 2.1. In assays containing SSBs, reactions were supplemented with 4 µM mtSSB, 8 µM T4 SSB, or 8 µM RPA, before the addition of PrimPol. Note that twice as much RPA than mtSSB was used due to the increased level of RPA required to completely shift the DNA probe in EMSA reactions (Figure 2.1.). An excess of SSBs over ssDNA was used to ensure that the template was fully coated, taking into account the binding site size of the protein and the length of the ssDNA binding interface. Reaction products were resolved on a 15% (v/v) polyacrylamide gel containing 7M urea and 1x TBE buffer at 850 V for 2.5 hours in 1x TBE buffer. Gels were scanned with a FujiFilm FLA-5100 image reader.

2.3.8. DNA Primer Extension Assays

Primer extension assays were performed using 5’ Hexachlorofluorescein labelled DNA primers (ATDbio) (sequences 1-4 in Table 2.1.) annealed to complementary unlabelled DNA templates (sequences 5-15 in Table 2.1.). Extensions were carried out at 37°C in 20 µL volumes containing; 100 nM of the assayed polymerase (or 3U/mL of T4 Pol), 20

#	Oligonucleotide	Modification	Sequence (5'-3')
1	HP-16 Primer	5'-HEX	CACTGACTGTATGATG
2	HP-20 Primer	5'-HEX	TGTCGTCTGTTCGGTCGTTC
3	HP-27 Primer	5'-HEX	TGTCGTCTGTTCGGTCGTTCGGTCTTC
4	HP-28 Primer	5'-HEX	TGTCGTCTGTTCGGTCGTTCGGTCTTCA
5	ND-50 Template$_{TT}$	N/A	CGCGCAGGGCGCACAACAGCCTTGAAGACCGAACGACCGAACAGACGACA
6	ND-50 Template$_{AA}$	N/A	CGCGCAGGGCGCACAACAGCCAAGAAGACCGAACGACCGAACAGACGACA
7	ND-50 Template$_{CC}$	N/A	CGCGCAGGGCGCACAACAGAGCCGAAGACCGAACGACCGAACAGACGACA
8	ND-50 Template$_{GG}$	N/A	CGCGCAGGGCGCACAACAGCCGGGAAGACCGAACGACCGAACAGACGACA
9	ND-97 Template	3'-Biotin	ACCGCGAACTTGAATTCTAGTTCAGTCTAAATGCTCTCAAGCACTGAGCAATTCACAACATAT GGCTTTCGATTACCGAACGACCGAACAGACGACA
10	Poly(dT)-60	5'-Biotin	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
11	6-4PP Template	N/A	CTCGTCAGCATCT^TCATCATACAGTCAGTG
12	CPD Template	N/A	CGCGCAGGGCGCACAACAGCCT=TGAAGACCGAACGACCGAACAGACGACA
13	8-oxo-dG Template	N/A	CGCGCAGGGCGCACAACAGCC8-oxo-dGTGAAGACCGAACGACCGAACAGACGACA
14	dU Template	N/A	CGCGCAGGGCGCACAACAGCCUTGAAGACCGAACGACCGAACAGACGACA
15	Ap Template	N/A	CGCGCAGGGCGCACAACAGCCApTGAAGACCGAACGACCGAACAGACGACA
16	ND-50 ssDNA Template	5'-FAM	CGCGCAGGGCGCACAACAGCCTTGAAGACCGAACGACCGAACAGACGACA
17	Primase Template	5'-Biotin	GTCTTCTATCTCGTCTATATTCTATTGTCTCTATGAATACCTTCATCAGTCTCACATAGATGC ATC

Table 2.1. Primer/template sequences.
Sequences of the DNA oligonucleotides used in primase and primer extension assays. Lesions within the sequences are denoted in red.

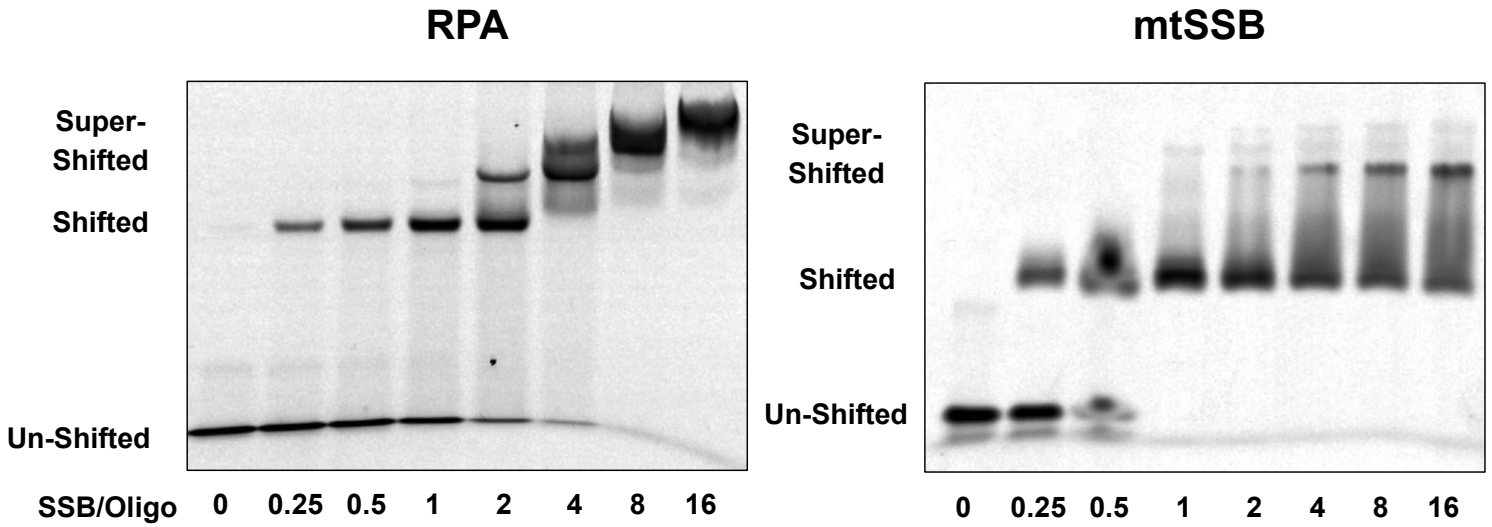


Figure 2.1. EMSA confirmation of RPA and mtSSB ssDNA binding.
'SSB/Oligo' below each gel indicates the molar ratio of SSB to the ssDNA template. Note that more than twice as much RPA, compared to mtSSB, was required to fully shift the ssDNA template. Super-shifted bands indicate multiple SSBs bound to the ssDNA template, demonstrating that a large excess of protein over DNA is required to fully coat the template.

nM primer-template substrate, 100 µM dNTPs (NEB), 50 mM Potassium acetate, 20 mM Tris-acetate pH 7.9, 10 mM Magnesium acetate, 1 mM DTT, and 2 µg BSA (NEB). In the case of single nucleotide incorporation assays, 100 µM of the individual dNTP to be assayed was added in place of all dNTPs. For assays using SSBs, DNA templates were pre-incubated on ice with 200 nM mtSSB, 400 nM RPA, or 400 nM T4 SSB, before the addition of enzymes. Again, twice as much RPA than mtSSB was used due to the apparent lower affinity of RPA for DNA as perceived in EMSA experiments (Figure 2.1.). Extension reactions were monitored over varying time courses, typically 1min, 3min, 5min, and 10min (except where indicated), and quenched with 20 µL stop buffer (95% (v/v) formamide, 0.05% (v/v) bromophenol blue, 0.09% (v/v) xylene cyanol and 200 nM competitor oligonucleotide). Products were boiled at 95°C for 5min before resolution on a 15% (v/v) polyacrylamide/7 M urea gel. Gels were scanned using an FLA-5100 image reader (Fujifilm). Primer extension products were quantified using ImageQuant TL software (GE Life Sciences).

2.3.9. The pSJ4 Plasmid-Based *lacZα* Fidelity Assay

The fidelity of human PrimPol was determined using the pSJ4 plasmid-based *lacZ*α reporter gene assay. The pSJ4 plasmid is a customised version of the previously described pSJ3 plasmid (Keith et al., 2013). The differentiating feature of the pSJ4 plasmid is a short 64 nt long gap (versus the 163 nt long gap of pSJ3), which is more efficiently filled up by distributive DNA polymerases *in vitro*. This specific feature was necessary to implement in the pSJ4 plasmid to make it suitable for measurements of the fidelity of human PrimPol. Typical pSJ4 gap filling reactions were carried out in a total volume of 10 µl comprising: reaction buffer (50 mM Potassium acetate, 20 mM Tris-acetate pH 7.9, 10 mM Magnesium acetate, 1 mM DTT, and 2 µg BSA), 20 fmol of gapped pSJ4 plasmid, 100 µM of each dNTP and 50 nM PrimPol. The gap filling reaction was carried out at 37°C for 30 minutes and completion was confirmed using an analytical digestion with EcoRI (Fermentas) restriction endonuclease followed by 1% agarose electrophoresis. As a control the pSJ4 *lacZ*α reporter gene assay was used to measure the fidelity of two well characterised thermostable DNA polymerases, the Klenow fragment kTaq-Pol A and Tgo-Pol B (exo$^-$). The fidelity of both of the thermostable polymerases was measured as described previously (Keith et al., 2013).

2.3.10. *In vitro* HSV-*tk* Mutagenesis Assay

Primed ssDNA and gapped-duplex substrates were prepared as previously described (Eckert et al., 2002; Hile and Eckert, 2008). Primer extension reactions were initiated with 1.6 µM PrimPol in buffer containing 10 mM Bis Tris Propane-HCl pH 7.0, 10 mM

$MgCl_2$, 1 mM DTT, 500 µM dNTPs, and 200 nM primed ssDNA substrate in 50 µl total volume. Reactions were terminated after 15 minutes. The enzyme was used in excess conditions due to PrimPol's distributive synthesis pattern and ssDNA binding capacity. The 81 nt HSV-*tk* target sequence was isolated by MluI and StuI digestion and hybridized to a gapped heteroduplex DNA molecule. After confirming hybridization by agarose gel electrophoresis, FT334 *E. coli* (*upp, tdk)* were transformed with hybridized DNA. Transformed bacteria were plated on VBA plates containing 50 µg/mL chloramphenicol (Cm) in the absence or presence of 40 µM 5-fluoro-2′-deoxyuridine (FUdR) to determine HSV-*tk* mutation frequencies, as described (Eckert et al., 2002). The observed HSV-*tk* frequency is calculated as the ratio of $FUdR^R/Cm^R$ to Cm^R colonies, and was determined for three independent reactions. DNA sequence analysis was conducted on independent mutants isolated from two independent PrimPol reactions per template. The polymerase error frequency (Pol EF) was calculated by subtracting the ssDNA background mutation frequency from the HSV-TK frequency. To correct the Pol EF for HSV-*tk* mutants with multiple polymerase errors, the Pol EF_{est} was derived as described (Opresko et al., 2000), using the following formula:

$$Pol\ EF_{est} = Pol\ EF / \sum_{n=1}^{3} \left(\frac{1}{n}\right)\left(\frac{mutants\ with\ n\ errors}{total\ mutants\ analyzed}\right)$$

where n is the total number of detectable errors that were > 10 nucleotides apart within the same sequence.

2.4. Results

2.4.1. PrimPol is not Stimulated by PCNA

Mono-ubiquitylation of PCNA in response to DNA damage increases its affinity for TLS polymerases such as Pol η, Pol κ, Pol ι, and Rev1, promoting their recruitment to the replication fork in order to facilitate lesion bypass (Lehmann et al., 2007). *In vitro* studies have shown that the ability of Pol η and Rev1 to bypass an abasic site is stimulated by the presence of mono-ubiquitylated PCNA over un-modified PCNA. However, on an undamaged template, PCNA stimulates the polymerase activities of these enzymes to a similar degree, regardless of its ubiquitylation status (Garg and Burgers, 2005a). In order to investigate whether PrimPol is stimulated by PCNA, we assessed the impact of both un-modified and mono-ubiquitylated PCNA on the polymerase activity of the enzyme *in vitro*. To do this, we employed primer extension assays on a 97-mer DNA template (sequence 9, Table 2.1.) annealed to a 20 nucleotide primer (sequence 2, Table 2.1.). Unlike the stimulating effect of PCNA on the polymerase activity of Pol η and Rev1, we

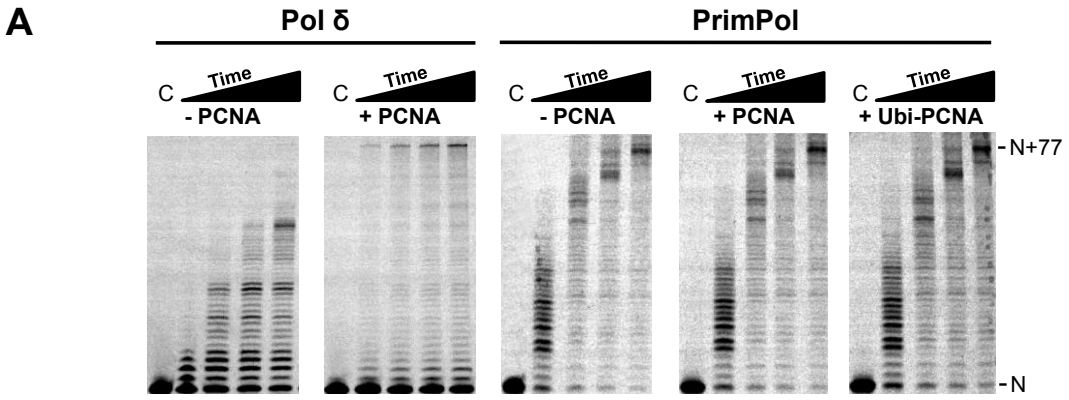
find that both un-modified and mono-ubiquitylated PCNA have no stimulatory effect on the polymerase activity of PrimPol (Figure 2.2.A.). In contrast, the presence of un-modified PCNA increased the processivity of Pol δ in the same conditions (Figure 2.2.A.). These results demonstrate that PrimPol, unlike other TLS polymerases, is not stimulated by either un-modified or mono-ubiquitylated PCNA, suggesting that the enzyme is regulated by another distinct mechanism.

2.4.2. PrimPol Interacts with RPA and mtSSB *in vivo*

To identify cellular factors that associate with PrimPol, and thus may be involved in regulating this polymerase *in vivo*, we purified PrimPol from cultured human cells and identified co-purifying proteins using MS. To facilitate affinity purification of PrimPol, we fused it to a *Strep*-tag, which exploits the high affinity and specific binding between streptavidin and its natural ligand biotin (Schmidt et al., 1996). Specifically, the eight amino acid long *Strep*-Tag II (WSHPQFEK) was used, which allows affinity purification with the streptavidin derivative *Strep*-Tactin and specific elution with desthiobiotin (Schmidt et al., 1996; Voss and Skerra, 1997). The Flp-In T-REx system was used to generate a stable cell line with doxycycline-inducible expression of *Strep*-tagged PrimPol. Affinity purified *Strep*-tagged PrimPol and co-purifying proteins were resolved by SDS-PAGE and excised gel bands analysed by MS.

Proteins identified in the MS analysis were ranked according to percentage of total spectra in the induced sample, and the fold enrichment calculated for each. A large set of proteins (1249) were identified, of these ~550 were present only in the induced sample and a further 65 showed a >3-fold enrichment, with PrimPol enriched 20-fold. Input of these proteins into the Database of Annotation, Visualisation, and Integrated Discovery (DAVID) clustered these proteins into a number of functionally related groups (Dennis et al., 2003; Huang et al., 2008). Consistent with the dual localisation of PrimPol, two of the predominant groups were nuclear and mitochondrial proteins. A large proportion of DNA and nucleotide binding proteins were also present, and more specifically proteins involved in DNA replication and repair, such as RPA (Figure 2.2.B.). In contrast, no mitochondrial replication enzymes were present, although mtSSB was enriched.

To validate the potential PrimPol interacting proteins from the preliminary MS analysis, small-scale affinity purifications of *Strep*-tagged PrimPol from whole cell lysate were performed and analysed by Western blot. Following addition of doxycycline a predominant species of ~69 kDa was detected by Western blot analysis with an anti-PrimPol antibody (Figure 2.2.C.), and furthermore, endogenous PrimPol was also detected, with little *Strep*-tagged PrimPol visible in the non-induced lysate (Figure



B

Protein identified	Unique peptides		Sequence coverage %		Total number of spectra	
	E2	E3	E2	E3	E2	E3
RPA70	1	9	2	13	1	14
RPA32	2	2	7	7	2	3
mtSSB	5	6	31	39	5	14

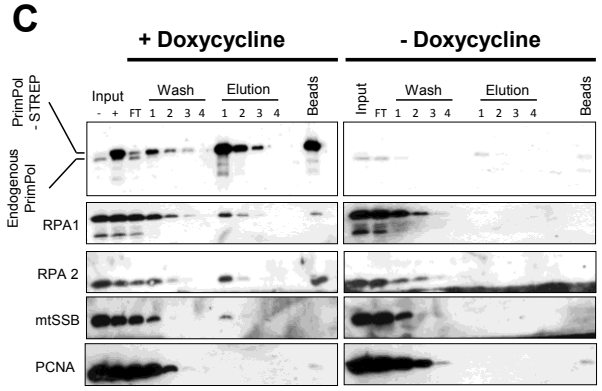


Figure 2.2. PrimPol is not regulated by PCNA but does interact with SSBs.
(A) PrimPol or Pol δ (100 nM) were incubated with primer template substrates (20 nM) and dNTPs (100 μM) at 37ºC for increasing times (1, 3, 5, 10 mins) in the absence or presence of PCNA (200 nM). PrimPol's primer extension activity was not stimulated in the presence of either un-modified or mono-ubiquitylated PCNA (Ubi-PCNA). In contrast, Pol δ shows increased processivity when PCNA is present. 'C' indicates the no enzyme control. **(B)** Identification of binding partners of PrimPol as analysed through mass spectrometry analysis, showing enrichment of RPA subunits 1 and 2, and mtSSB. The fold enrichment of each protein is indicated on the right of the table. **(C)** Western blot validation of PrimPol interacting proteins identified in the mass-spectrometry analysis. PrimPol co-purifies with both mtSSB and RPA, but not with PCNA. Flp-In T-Rex-293 cells engineered for inducible expression of PrimPol^{FlagStrep} were grown in the presence or absence of doxycycline (10 ng/ml, 24 hours) and PrimPol^{FlagStrep} was affinity purified from the soluble lysate using *Strep*-Tactin resin. Fractions from the affinity purification were analysed by Western blot with the indicated antibodies. Input '+' and '-' represent the clarified lysate of cells grown in the presence or absence of doxycycline, respectively.

2.2.C.). Analysis of the affinity purification with RPA70 and RPA32 antibodies, and the mitochondrial equivalent, mtSSB, after stringent washing, all gave specific bands in the elutions (Figure 2.2.C), indicating that PrimPol associates with these proteins. The RPA findings agree with recent studies by Wan *et al*. (Wan et al., 2013), which reported that PrimPol interacts with RPA and that this interaction may be required for recruitment of PrimPol to stalled replication forks (Wan et al., 2013). ATAD3, a mitochondrial membrane associated ATPase and core nucleoid component, was also detected in the elutions, however ATAD3 did appear to bind to the *Strep*-tactin resin in the non-induced sample (data not shown). Notably, analysis of the affinity purification using an anti-PCNA antibody did not give detectable bands in the elutions (Figure 2.2.C.), suggesting that, unlike other TLS polymerases, PrimPol does not interact with PCNA. This result agrees with the inability of PCNA to stimulate PrimPol and further suggests that PrimPol is not regulated by PCNA *in vivo*. Although many potential interactions were identified by MS, we have validated that both major classes of cellular SSBs (RPA and mtSSB) co-purify with PrimPol, whilst PCNA does not.

2.4.3. RPA70N Protein Recruitment Domain of RPA Mediates the Interaction with PrimPol

In order to cross-validate the interaction between RPA and PrimPol, as well as characterise the domains responsible, we screened the two primary RPA protein recruitment domains DBD-N of RPA1 (RPA70N) and the winged-helix domain of RPA2 (RPA32C) using NMR spectroscopy. To this end, ^{15}N-^{1}H Heteronuclear Single Quantum Coherence (HSQC) NMR spectra of ^{15}N-enriched RPA70N$^{1\text{-}120}$ and RPA32C$^{172\text{-}270}$ were acquired in the absence and presence of two-fold molar excess of unlabelled PrimPol. These spectra monitor amide chemical shifts, which are sensitive to their local chemical environment. Thus, binding of a ligand is expected to perturb the location and/or intensity of the peaks from residues at the binding site. We note that additional chemical shift perturbations can occur within globular protein interaction domains as result of structural changes induced by ligand binding.

The NMR analysis of the interaction of RPA32C with PrimPol revealed no significant chemical shift perturbations (Figure 2.3.). In contrast, addition of PrimPol to RPA70N generated substantial perturbations (Figure 2.4.A). The primary effect observed is loss of signal intensity for the majority of peaks (Figure 2.4.A. red), which we attribute to the large increase in mass as the ~13 kDa protein tumbles much more slowly when part of the ~78 kDa complex. These observations indicate that RPA70N serves as the primary PrimPol interaction site on RPA.

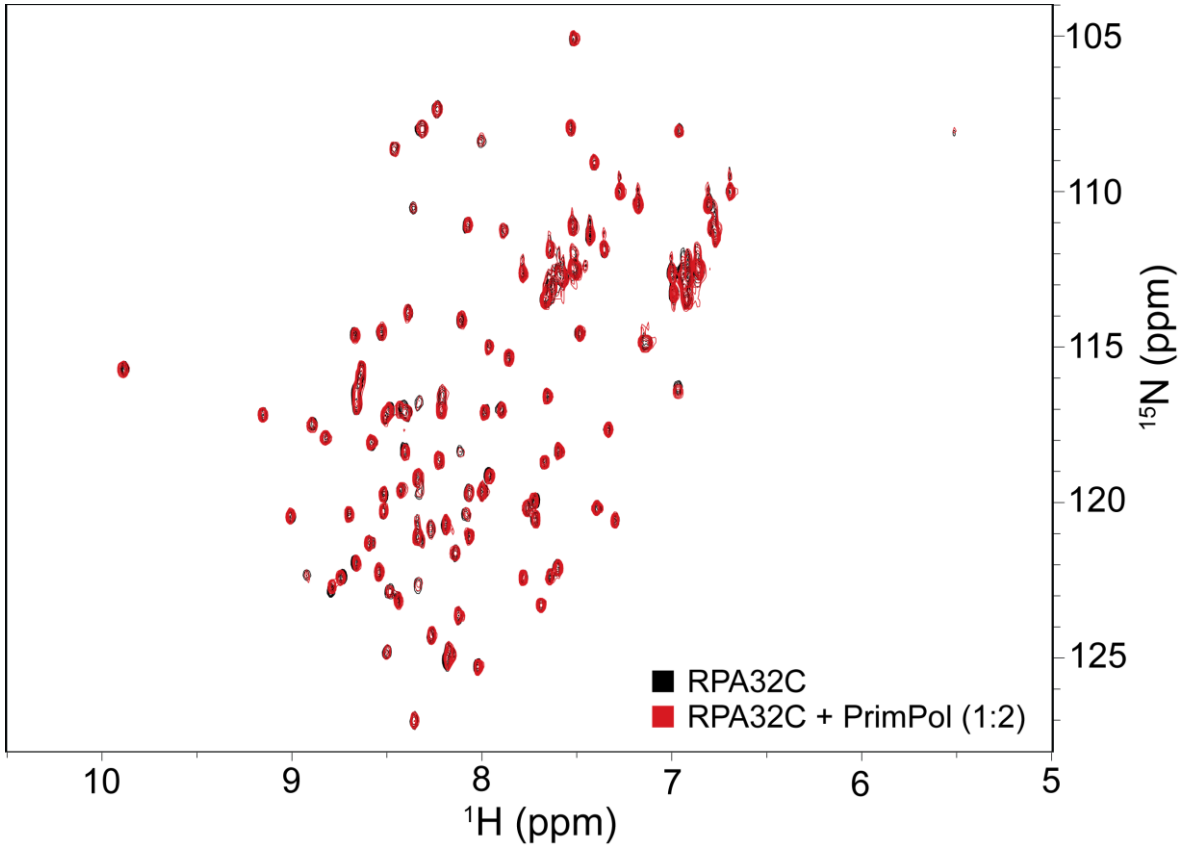


Figure 2.3. NMR titration of RPA32C with PrimPol.
^{15}N-^{1}H HSQC overlay of ^{15}N-enriched RPA32C alone (black) or in the presence of two-fold molar excess of PrimPol (red).

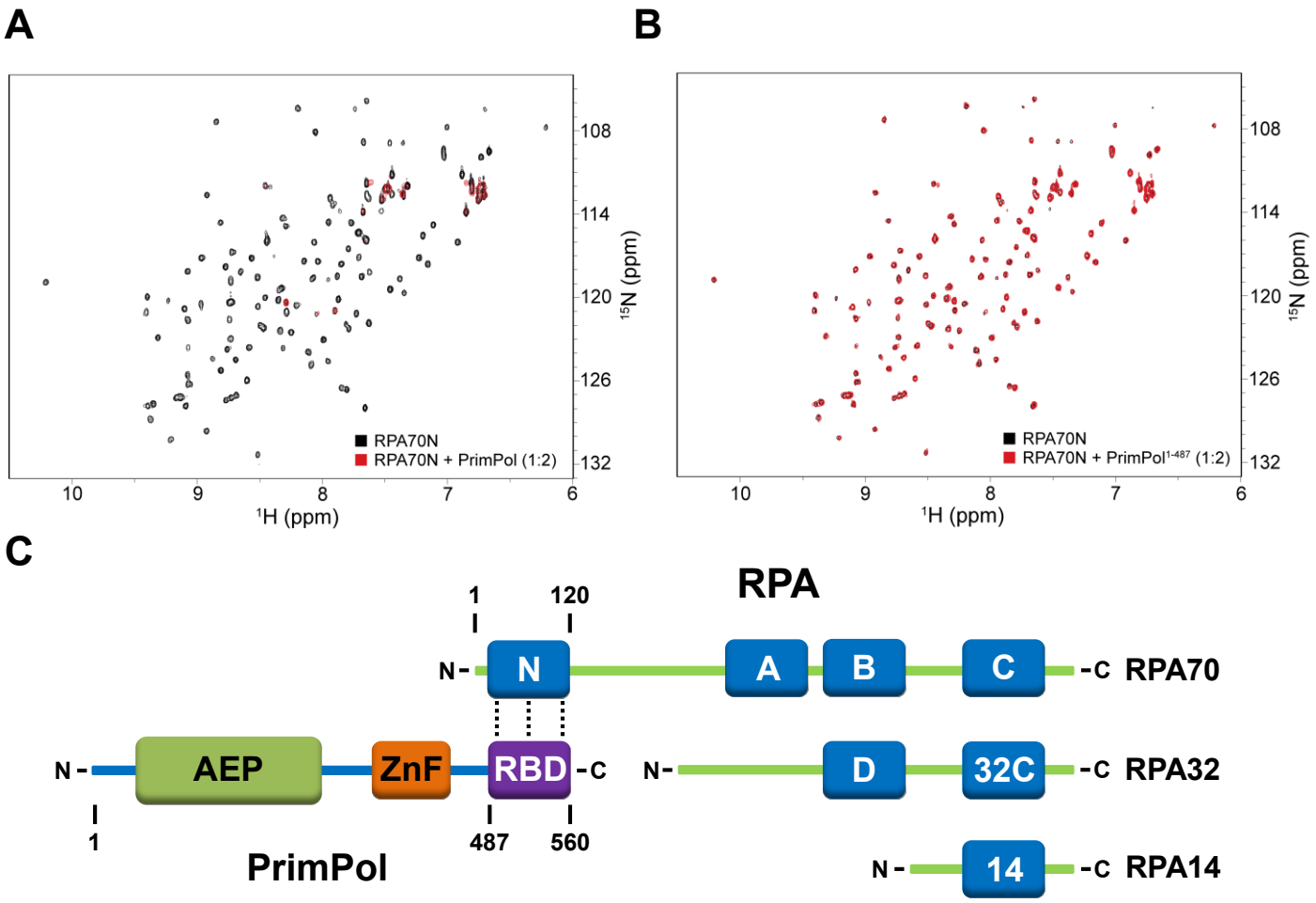


Figure 2.4. Characterisation of the RPA-PrimPol domain interactions.
^{15}N-^{1}H HSQC overlays of 50 µM ^{15}N-enriched RPA70N alone (**A-B,** black) and in the presence of 100 µM of PrimPol (**A**, red) or PrimPol$^{1\text{-}487}$ (**B**, red). All spectra were obtained as 25 °C in a buffer containing, 20 mM HEPES, 80 mM NaCl, 2 mM DTT, and 5% deuterium oxide at pH 7.1. **(C)** Schematic showing the RPA binding domain (RBD) of PrimPol and the RPA70N domain of RPA where it binds.

Having mapped the primary interaction domain of RPA, we next used NMR to better define the interaction region of PrimPol. Previous immunoprecipitation data suggested that the C-terminal region of PrimPol is responsible for its interaction with RPA (Wan et al., 2013). We therefore examined the binding to RPA70N of a PrimPol deletion construct lacking the C-terminal 73 residues (PrimPol$^{1-487}$). Figure 2.4.B. shows that the HSQC spectrum of RPA70N with PrimPol$^{1-487}$ closely resembled that of the free protein. Thus, loss of the C-terminal region causes PrimPol to lose its ability to bind to RPA70N. Together, these results support a model in which the primary interaction between PrimPol and RPA is mediated by the RPA70N and PrimPol$^{488-560}$ domains (Figure 2.2.C.).

2.4.4. RPA and mtSSB Suppress *de novo* Primer Synthesis by PrimPol

Recent studies identifying that PrimPol's C-terminal RPA-interacting domain is required for foci formation and appropriate functioning of the enzyme *in vivo*, suggest that RPA may act to recruit PrimPol to the replication fork (Wan et al., 2013). However, the effect of RPA on the enzymatic activities of PrimPol has not previously been studied. To determine the effect of both RPA and mtSSB on the primase activity of PrimPol, we analysed the enzyme's ability to synthesise primers on a 60-mer poly-dT ssDNA template (sequence 10, Table 2.1.), coated with either RPA or mtSSB. T4 SSB coated ssDNA was also included as a non PrimPol-interacting control. As observed previously (Bianchi et al., 2013; Keen et al., 2014b), PrimPol facilitated primer synthesis on the ssDNA template in the absence of SSBs. However, RPA and mtSSB strongly inhibited the ability of PrimPol to synthesise primers, both with dNTPs and rNTPs (Figure 2.5.). PrimPol also failed to synthesise primers on ssDNA coated with T4 SSB. This suggests that SSBs supress the primer synthesis ability of PrimPol by blocking potential DNA binding sites for the enzyme. These findings echo previous studies of the Pol α complex, which demonstrated that primer synthesis was suppressed on ssDNA templates coated with RPA (Collins and Kelly, 1991).

2.4.5. RPA and mtSSB Impede Primer Extension by PrimPol

In contrast to the suppression RPA imposes on Pol α during primer synthesis, RPA stimulates the polymerase and processivity of the enzyme when elongating primers (Braun et al., 1997). This implies that RPA acts to prevent Pol α binding to ssDNA, hence negating primer synthesis but actively promotes primer extension. Furthermore, RPA and mtSSB have been shown to stimulate the polymerase activities of Pol δ and Pol γ, respectively (Oliveira and Kaguni, 2010; Tsurimoto and Stillman, 1989). Therefore, we next investigated the effect of RPA and mtSSB on the polymerase extension activity of PrimPol. T4 SSB was again included as a non-interacting control. In order to investigate

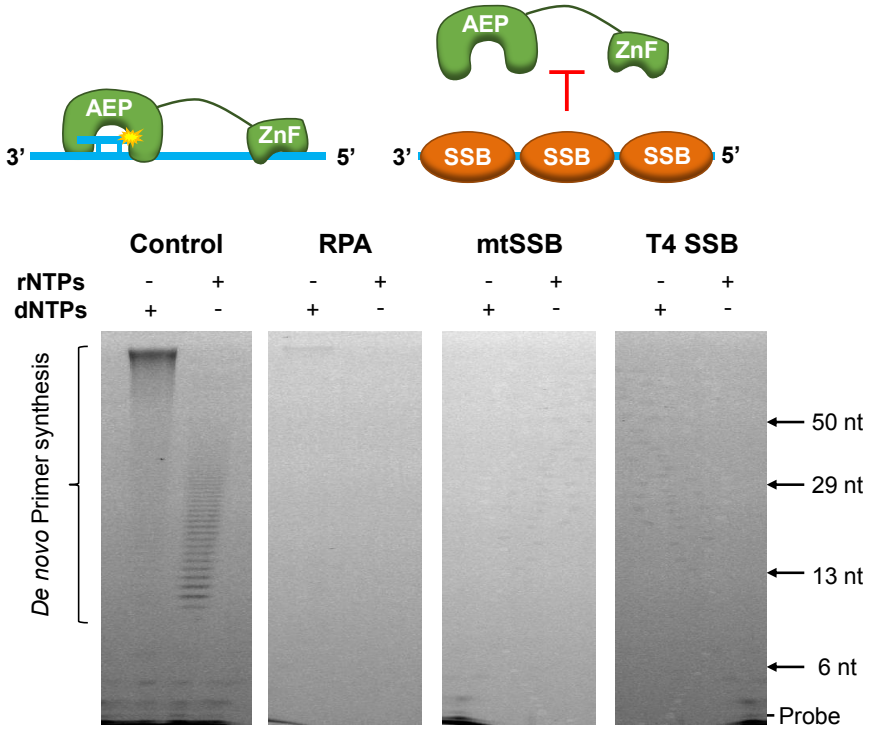


Figure 2.5. The effect of SSBs on the primase activity of human PrimPol.
Single-stranded poly-dT templates (500 nM) were incubated with dNTPs or rNTPs (500 μM) and human PrimPol (1 μM), either alone or in the presence of RPA (8 μM), mtSSB (4 μM), or T4 SSB (8 μM) for 1hr at 37ºC (see 'materials and methods' for details). In the absence of SSBs PrimPol is capable of *de novo* primer synthesis using either dNTPs or rNTPs. However, when templates are coated with SSBs PrimPol is unable to synthesise primers. The schematic above represents how this inhibition is likely a result of RPA and mtSSB blocking PrimPol binding sites on the ssDNA template.

this, we employed a standard primer extension assay in the presence of RPA, mtSSB, or T4 SSB. This represents a physiologically relevant situation in which the replication fork has stalled at a site of DNA damage leading to uncoupling of the stalled replicative polymerase and the MCM helicase, resulting in the generation of long stretches of SSB-bound ssDNA (Byun et al., 2005). In addition to full-length PrimPol, we also analysed a truncation of the enzyme (PrimPol$^{24-354}$) that lacks both the ssDNA binding ZnF domain and the C-terminal region required for RPA interaction. This truncation allowed investigation into the effect of SSBs on the AEP domain of PrimPol alone, which has previously been shown to possess polymerase activity (Keen et al., 2014b).

In the absence of SSBs, the full length PrimPol and PrimPol$^{24-354}$ fully extended the majority of primers by the final time point (Figure 2.6.A. and B.). However, the presence of RPA, mtSSB, or T4 SSB, dramatically impeded primer extension by both enzymes (Figure 2.6.A. and B.). This inhibition caused both a reduction in the length of extended primers and an increase in the amount of unextended DNA substrate. The partial extension observed in the presence of SSBs suggests that PrimPol was unable to displace these proteins from the template DNA during primer extension. As a result of the dynamic nature of SSBs binding to DNA, any ssDNA un-bound by SSBs that is close to the primer-template junction would be available for extension by PrimPol until the enzyme was restricted by SSBs bound downstream or dissociated and was unable to bind again due to exclusion by SSB. The varying levels of inhibition observed in the presence of different SSBs may therefore be a result of the different binding footprints and affinities of the SSBs used, in conjunction with the length of the template. The inhibition of PrimPol$^{24-354}$, coupled with the inhibitory effect of T4 SSB on full-length PrimPol, indicates that this effect is not due to an interaction between PrimPol and the SSBs. Furthermore, inhibition of PrimPol$^{24-354}$ suggests that the inhibitory effect of SSBs is not only due to competition with PrimPol's ZnF domain for binding of ssDNA (Figure 2.6.B.).

In order to ensure that the inhibitory effect of RPA and mtSSB on PrimPol primer-extension is not simply a result of the amount of protein used, we repeated the assay in the presence of a large range of SSB concentrations. In each case, a similar level of inhibition was observed at protein concentrations sufficient to coat the single-stranded region of the DNA template (Figure 2.7. and 2.8.). A similar level of inhibition was also observed when PrimPol was pre-incubated with the template before adding dNTPs and SSBs (Figure 2.9.). In addition, Pol δ with PCNA, Pol γ^{exo-}, and T4 Pol were able to displace RPA, mtSSB, and T4 SSB respectively, and fully extend the primer in the same conditions in which PrimPol activity is limited (Figure 2.10.). This reveals a striking

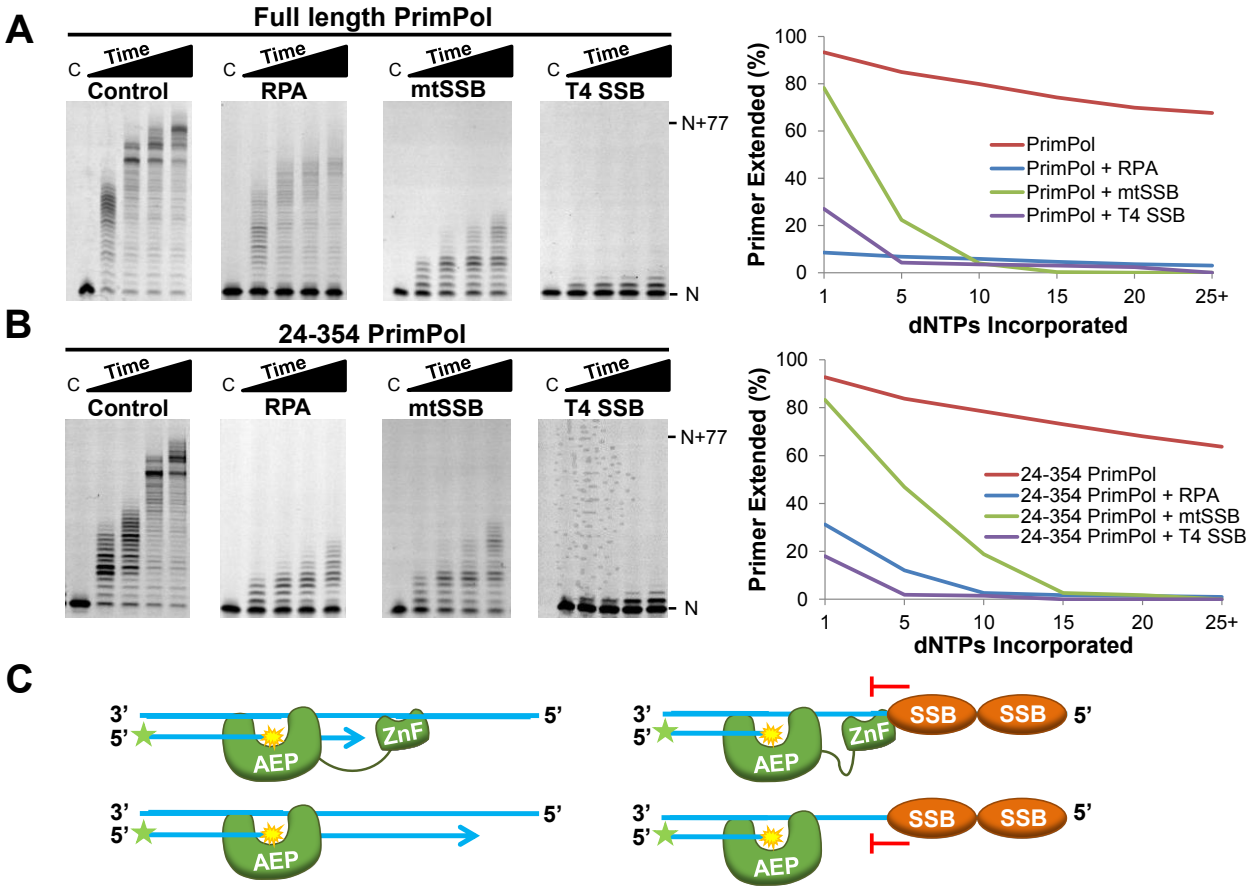


Figure 2.6. The effect of SSBs on the primer extension activity of PrimPol.
(A) Primer-template substrates (20 nM) and dNTPs (100 μM) were incubated with PrimPol (100 nM), either alone or with RPA (400 nM), mtSSB (200 nM) or T4 SSB (400 nM), for increasing times (1, 3, 5, 10 mins). In the presence of SSBs, full-length PrimPol's primer extension activity is severely impeded. The primer runs at the position indicated 'N', with full extension denoted by 'N+77'. 'C' indicates the no enzyme control. For each gel, the 10 minute time-point was quantified to identify the percentage of primers extended more than 1, 5, 10, 15, 20, or 25 bases. This quantification is shown on the right of the gels. **(B)** The primer extension activity of PrimPol$^{24\text{-}354}$ is also restricted in the presence of SSBs. Quantification of the 10 minute time-point for each gel is again shown to the right of the gels. **(C)** Schematic representation of the inhibitory effect of SSBs on the primer extension activity of PrimPol.

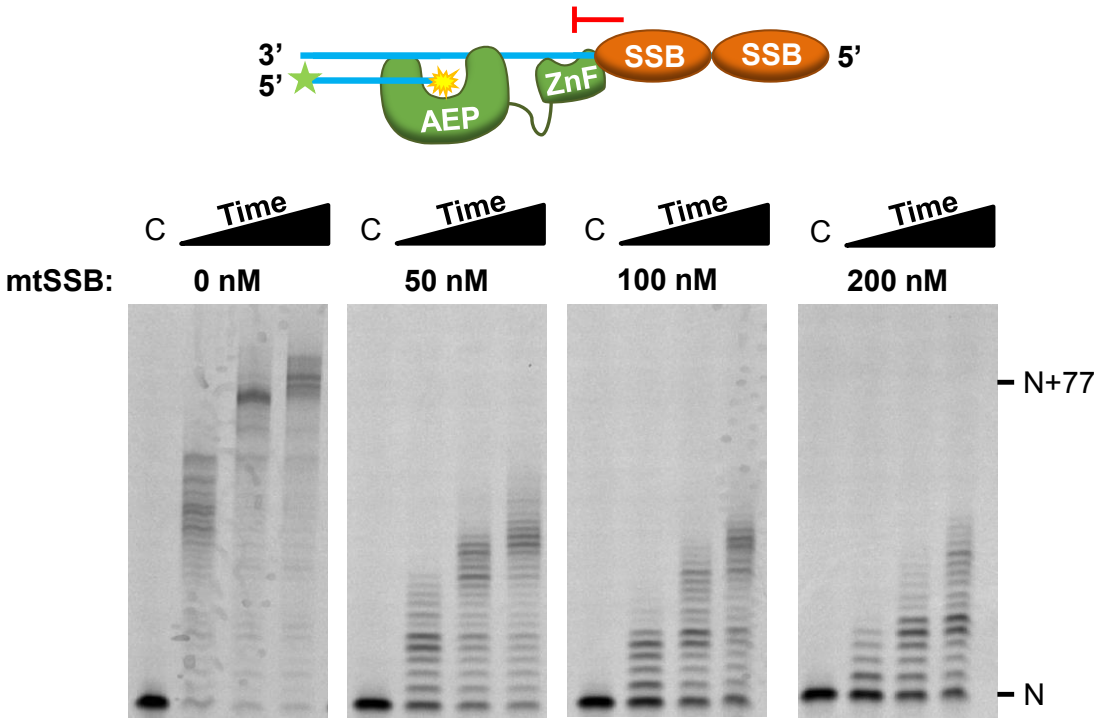


Figure 2.7. mtSSB inhibits PrimPol over a large range of protein concentrations.
The concentration of mtSSB used in each assay is indicated above each gel. The time-points used are 1, 5, and 10 mins 'C' indicates the no enzyme control.

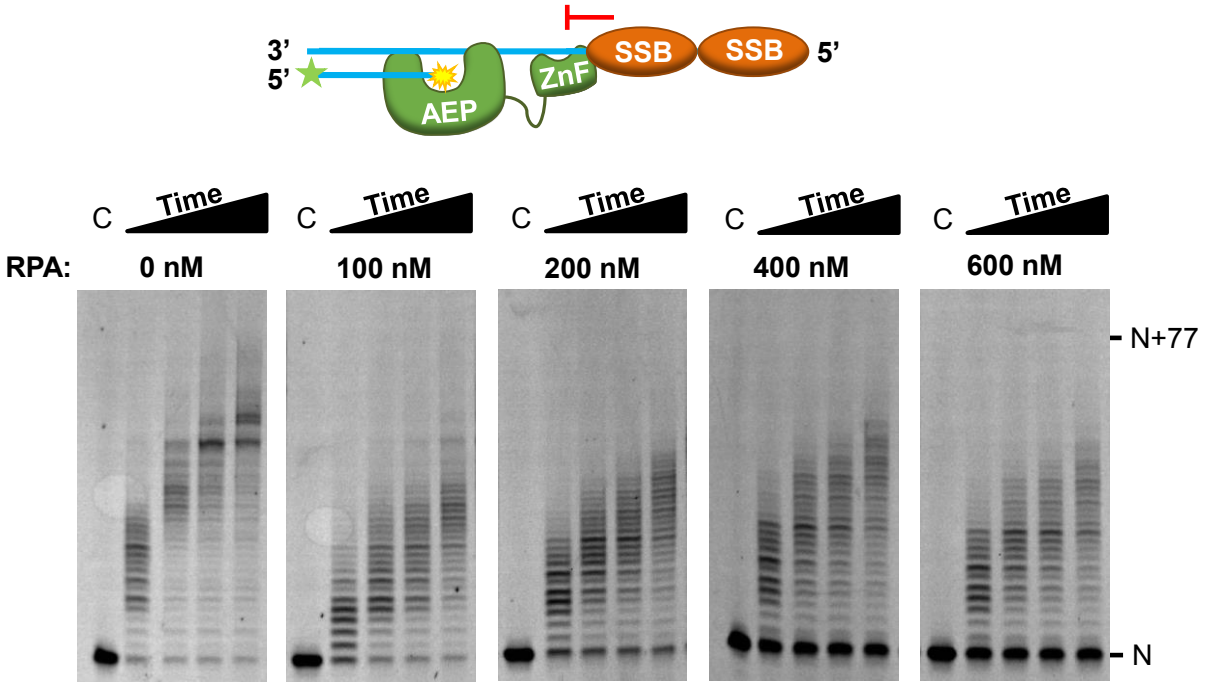


Figure 2.8. RPA inhibits PrimPol over a large range of protein concentrations.
The concentration of RPA used in each assay is indicated above each gel. 'C' indicates the no enzyme control.

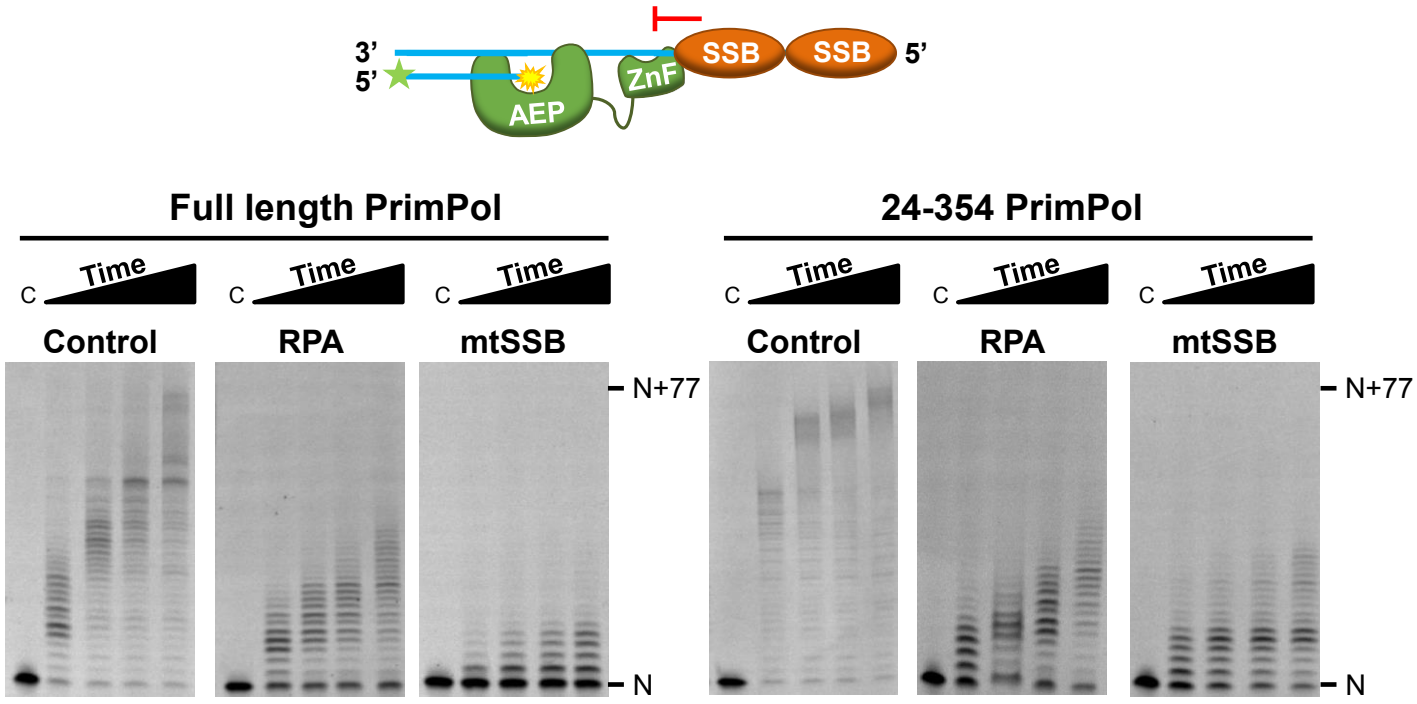


Figure 2.9. SSBs inhibit PrimPol when added second.
Full-length and 24-354 PrimPol were pre-incubated with the DNA template to allow binding before the addition of dNTPs and either RPA or mtSSB. In each case, inhibition of primer extension was observed. 'C' indicates the no enzyme control.

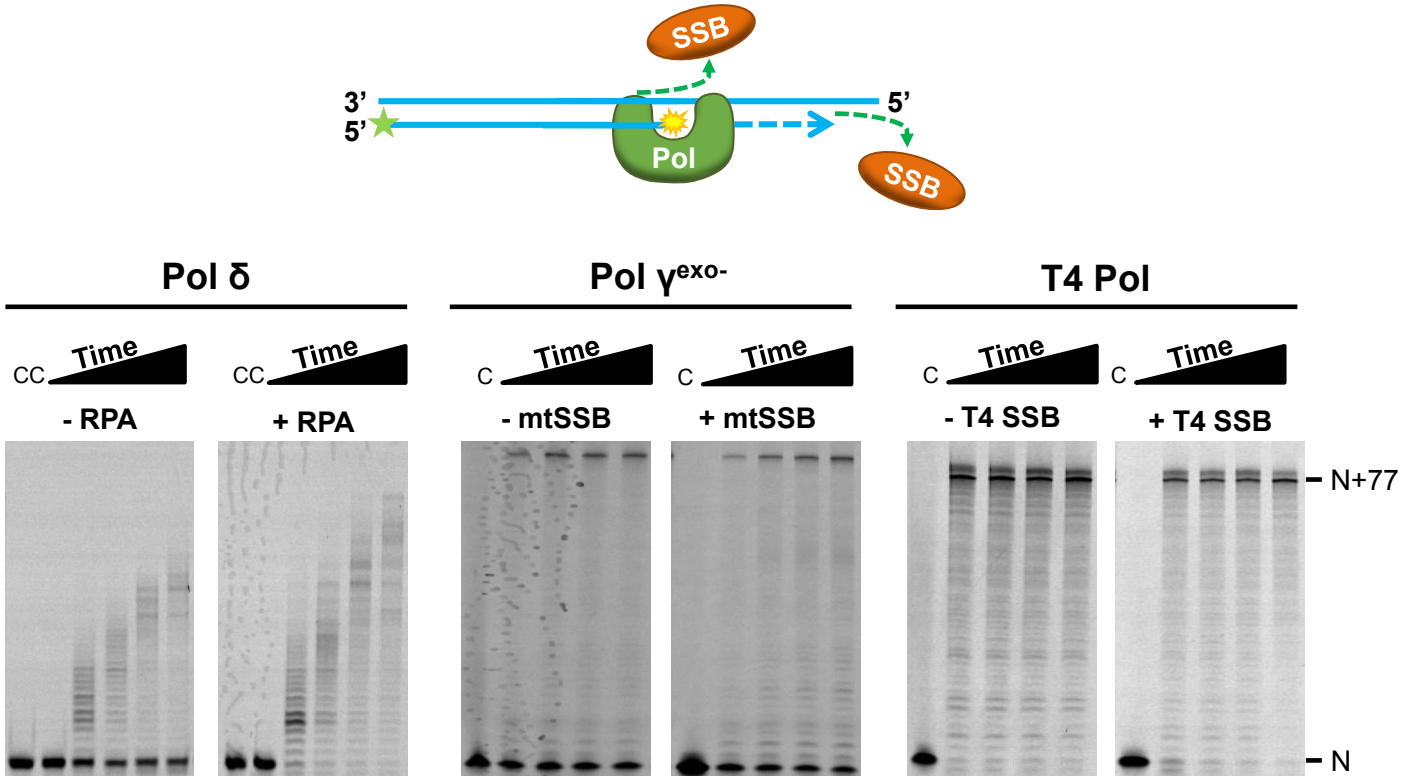


Figure 2.10. Replicative polymerases displace SSBs.
Pol δ with PCNA, Polγ^{exo-}, and T4 Pol, are able to displace RPA, mtSSB, and T4 SSB, from DNA, respectively. 'C' indicates the no enzyme control.

difference in the ability of PrimPol to displace SSBs in comparison to replicative polymerases.

Stimulation of Pol γ by mtSSB has previously been shown to be salt-dependent, maximally stimulated at 20 mM potassium chloride and inhibited at concentrations ~100 mM (Oliveira and Kaguni, 2010). We tested whether the inhibition of PrimPol by RPA and mtSSB was also salt-dependent by repeating the primer extension assays in a range of salt concentrations. The restraining impact of RPA and mtSSB on the polymerase activity of PrimPol was consistent across the range of salt concentrations tested, between 0 and 60 mM KCl (Figure 2.11. and 2.12.), ruling out the possibility that this effect is salt-dependent.

RPA has previously been implicated in modulating the fidelity of 8-oxo-dG bypass by Pol λ and Pol η, specifically acting to increase accurate dCTP incorporation opposite the lesion (Maga et al., 2007). We have recently reported that PrimPol is able to bypass 8-oxo-dG lesions by incorporation of either dCTP or dATP. In addition, PrimPol incorrectly incorporates dTTP opposite the first T of a 6-4PP and incorporates dATP opposite deoxyuracil, whilst the full-length enzyme is unable to bypass CPDs or Ap sites in magnesium (Keen et al., 2014b). We examined whether RPA or mtSSB affected the ability, or fidelity, of lesion bypass across a range of different templating lesions (sequences 11-15, Table 2.1) using single incorporation primer extension assays with the lesion immediately downstream of the primer-template junction. RPA and mtSSB did not appear to alter either the ability to bypass lesions or the fidelity of this bypass, except in the case of the 6-4PP where bypass was inhibited in both cases (Figure 2.13.). This may be due to the shorter length of the 6-4PP template compared to the length of the other lesion-containing templates. Alternatively, the SSBs might prevent the looping out mechanism which has been proposed to be employed by PrimPol for bypass of 6-4PP (Mourón et al., 2013).

These results demonstrate that PrimPol's polymerase activity is severely limited in the presence of SSBs. PrimPol has been confirmed as a competent TLS polymerase with roles in DDT *in vivo* (Rudd et al., 2014). As such, PrimPol is likely to be recruited to stalled replication forks, possibly by RPA (Wan et al., 2013), where it will encounter long stretches of RPA/mtSSB-bound ssDNA, a result of uncoupling of the replicative polymerase and helicase. Therefore, the inability of PrimPol to displace SSBs during primer elongation likely acts as a mechanism to limit PrimPol's contribution to DNA replication. In order to identify the necessity to restrict DNA synthesis by PrimPol during genome replication, we next investigated the fidelity and mutagenicity of this enzyme.

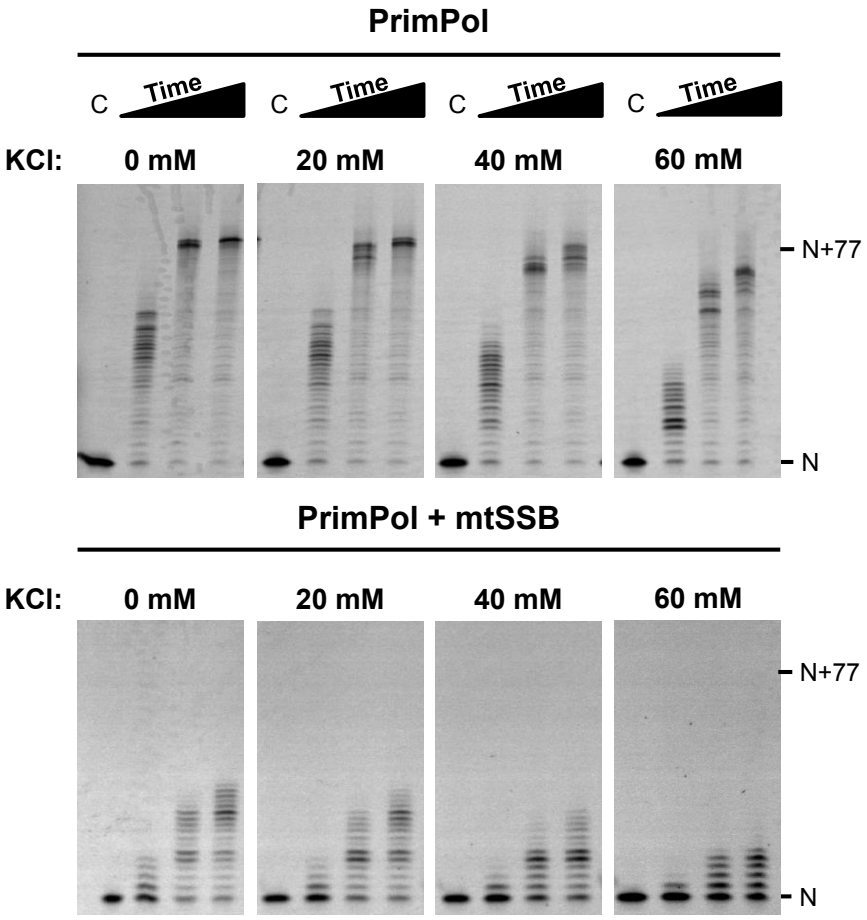


Figure 2.11. mtSSB inhibits PrimPol over a range of salt concentrations.
The concentration of KCl in the buffer for each assay is indicated above each gel. The time-points used are 1, 5, and 10mins. 'C' indicates the no enzyme control.

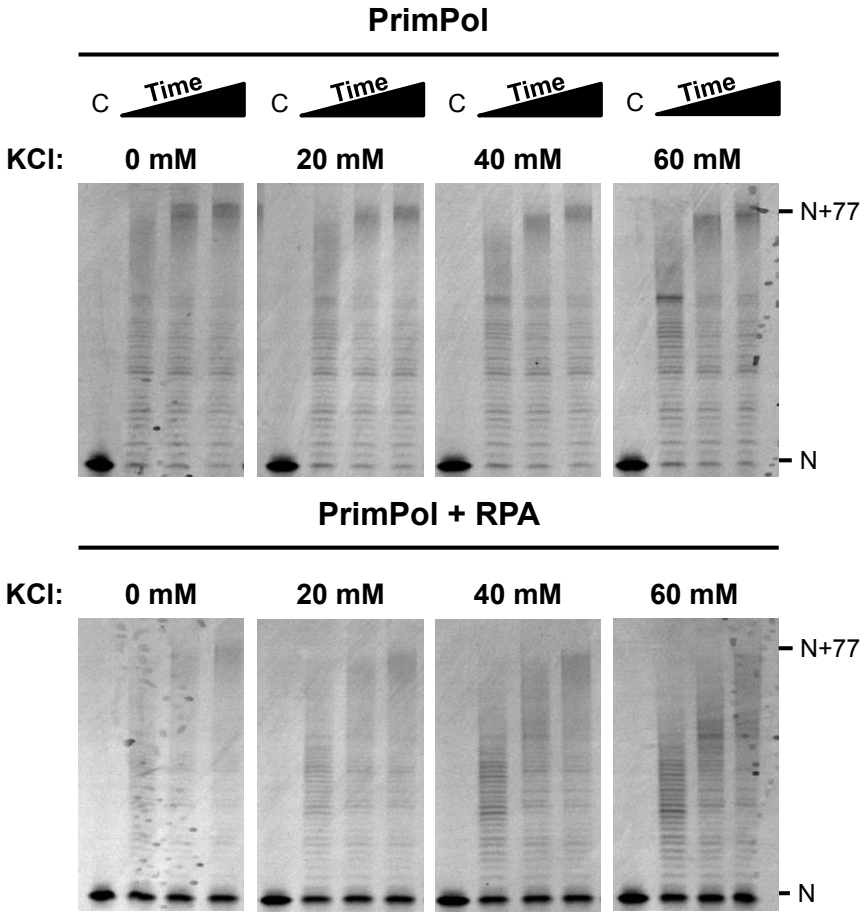


Figure 2.12. RPA inhibits PrimPol over a range of salt concentrations.
The concentration of KCl in the buffer for each assay is indicated above each gel. The time-points used are 1, 5, and 10 mins. 'C' indicates the no enzyme control.

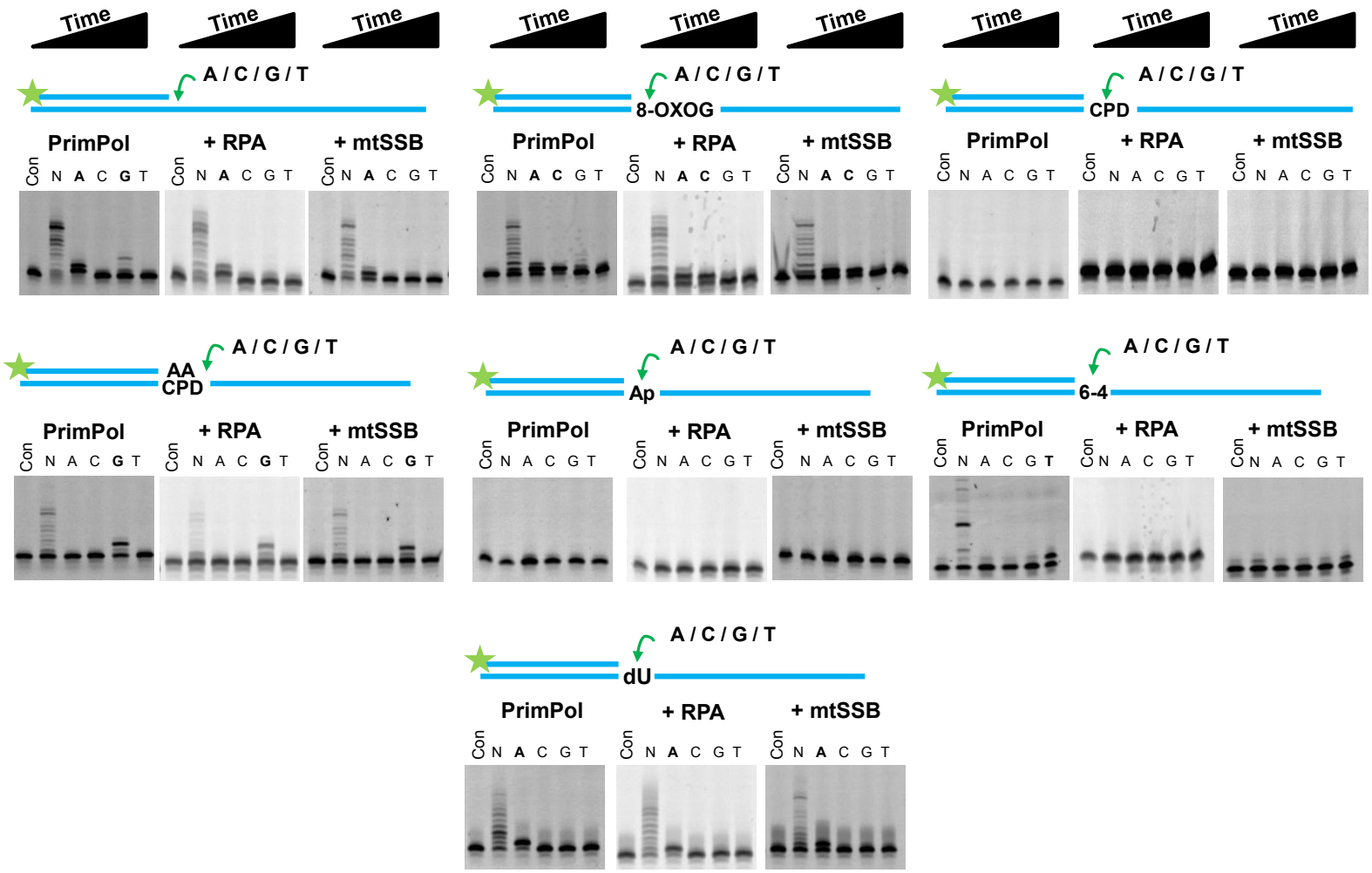


Figure 2.13. RPA and mtSSB do not affect the lesion bypass fidelity of PrimPol.
The lesion present in each assay is indicated above the respective gel. Each individual reaction was carried out for a single 10 minutes time-point. 'Con' indicates the no dTNP control.

2.4.6. PrimPol Shows a Propensity for Misincorporation and Mispair Extension

As an initial investigation into the base substitution fidelity of PrimPol, we used a primer extension assay based on single incorporation of either the correct or incorrect incoming base (Figure 2.14.). PrimPol was incubated with a short primer-template with either adenine (A), cytosine (C), guanine (G), or thymine (T), as the immediate templating base (N+1 position). This base was followed by two templating Cs (N+2 and N+3 positions), except where C was the base at N+1, in which case A and G were the upstream templating bases (N+2 and N+3 positions). For each primer-template substrate, the reaction was supplemented with only one of the four dNTPs (dATP, dCTP, dGTP, or dTTP). Quantification of these data and normalisation to correct incorporation suggest that PrimPol has a strong propensity to misincorporate dGTP, especially opposite a templating G (Figure 2.14.B.). However, when dGTP is the incoming base, significant product bands were visible at the N+2 and N+3 positions (Figure 2.14.A.). This increased N+2 and N+3 incorporation could result from PrimPol misaligning the N+1 templating base, rather than through misincorporation, which would in turn suggest a potential to generate base deletions.

PrimPol also showed a preference to misincorporate dCTP and dATP opposite a templating C (Figure 2.14.A. and B.). Consistently, a significant N+2 product was visible on the G template when dCTP was the incoming base (Figure 2.14.A.). Again, this corresponds to misincorporation of dCTP opposite a templating C at the N+2 position. A similar result was observed on the T template with the correct incoming base (dATP), indicating misincorporation of dATP opposite a templating dC at the N+2 position (Figure 2.14.A.). These results suggest that PrimPol has a propensity to misincorporate both dCTP and dATP opposite a templating C, which may be a potential error signature of PrimPol.

We also analysed the ability of PrimPol to extend different terminal mismatched base pairs (Figure 2.14.C. and D.). PrimPol was particularly proficient at extending C-C mismatches, with ~12% of the primers being extended relative to extension of a correctly matched C-G primer-template junction. The enzyme also showed a capacity to extend A-A, C-T, G-G, and A-C/C-A mismatches, whilst showing a much lower ability to extend from T-T and G-A/A-G mismatches (Figure 2.14.C. and D.). Together, these data reveal that PrimPol is able to facilitate both misincorporation of bases and perform extension of these base mispairs.

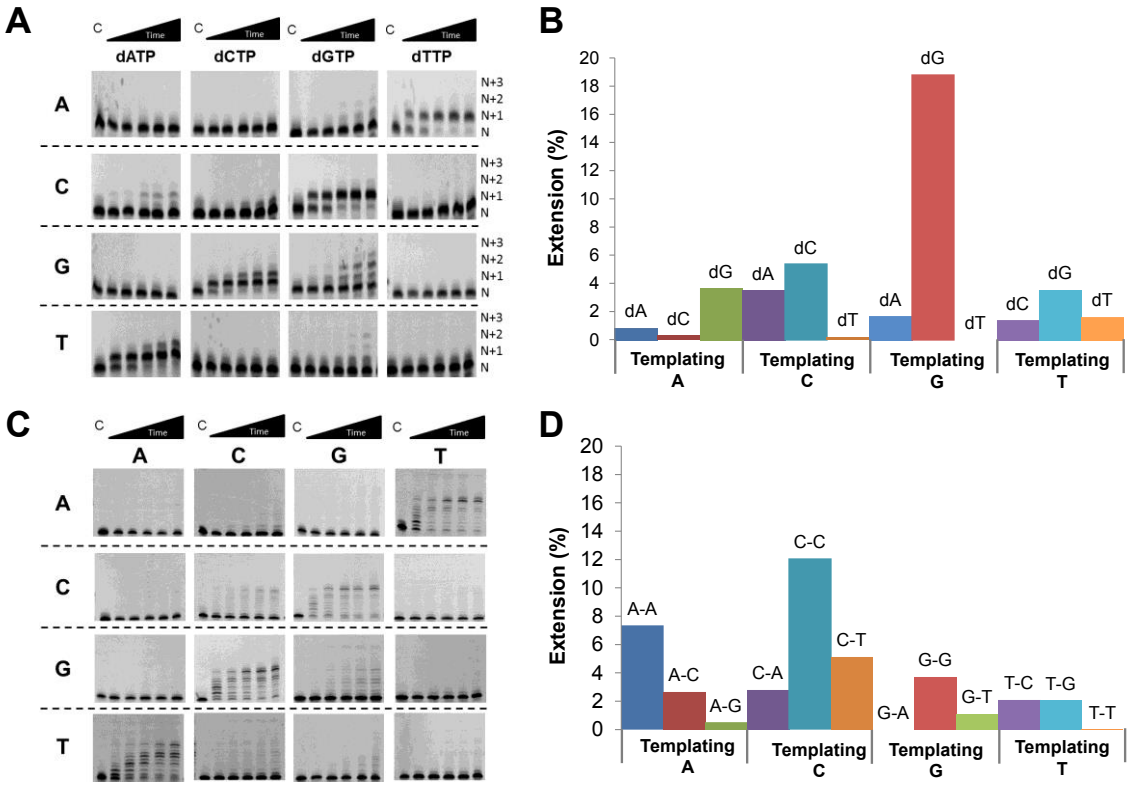


Figure 2.14. PrimPol can misincorporate bases and extend from mismatched bases.
(A) Analysis of misincorporation by PrimPol. PrimPol (100 nM) was incubated at 37ºC with primer-template substrates (20 nM) containing either A, C, G, or T, as the templating base immediately downstream of the primer for increasing times (1, 3, 5, 10 mins). Individual reactions contained either dATP, dCTP, dGTP, or dTTP (100 μM). The templating bases are indicated on the left, whilst the dNTP provided is shown above. Dotted lines separate reactions where the templating base is the same. 'N' denotes the position at which the primer runs whilst 'N+1', 'N+2', and 'N+3' indicate incorporation of 1, 2, or 3, bases respectively. **(B)** Quantification of the misincorporation assays at each 10 minute time-point. Data were normalised against the correct incoming base. **(C)** Analysis of mismatch extension by PrimPol. PrimPol was incubated at 37ºC for increasing times (1, 3, 5, 10 mins) in the presence of all four dTNPs and primer-template substrates containing a mismatched base at the 3' end of the primer The templating base for each gel is indicated on the left, whilst the corresponding mismatched primer base is shown above. **(D)** Quantification of the mismatch extension assays at each 10 minute time-point. Data were normalised against extension from correctly matched bases.

2.4.7. PrimPol is an Error-Prone Polymerase with a Preference to Generate Base Insertions and Deletions

Unlike the major human replicative polymerases (Pol δ and Pol ε), PrimPol lacks a 3' to 5' exonuclease proofreading domain and is therefore expected to be a potentially error-prone DNA polymerase. To characterise PrimPol's error frequency and spectrum, we employed a plasmid based *lacZα* reporter gap-filling assay (Keith et al., 2013). Due to the distributive nature of PrimPol we modified the recently developed pSJ3 plasmid to create a new plasmid (pSJ4) containing a shorter 64 nt long gapped region. Initially, the fidelity of two well characterised DNA polymerases (Klenow fragment of Taq-Pol A and exonuclease-deficient variant of Tgo-Pol B^{exo-}) was measured. From raw mutation frequencies an absolute error rate (number of mistakes made per base incorporated) was calculated as previously described (Keith et al., 2013). The Klenow fragment Taq-Pol A and Tgo-Pol B^{exo-} presented error rates of 3.6×10^{-5} and 1.6×10^{-5}, respectively (Table 2.2.; Figure 2.15.A.), agreeing with the previously reported fidelities of these enzymes (Jozwiakowski et al., 2014; Keith et al., 2013). Analysis of human PrimPol revealed an error-prone phenotype with a calculated error rate of 1×10^{-4}, essentially an order of magnitude lower than the exonuclease-deficient variants of *S. cerevesiae* replicative DNA polymerases δ, ε and the TLS specialised DNA polymerase ζ (Fortune et al., 2005; Shcherbakova et al., 2003; Zhong et al., 2006).

The mutations generated by PrimPol whilst copying the 64 nt long fragment of the *lacZα* reporter are visualised in Figure 2.2.C. PrimPol exhibited a tenfold preference for base substitution mutations when C or G were the templating bases, in comparison to errors introduced when copying A or T. Perhaps more intriguing however, was the finding that more than half of the mutations observed were base deletions/insertions, rather than the expected base substitutions (Figure 2.15.B.; Table 2.3). This apparent propensity of PrimPol to generate a very high proportion of insertion-deletion (indel) errors seems to support the previously proposed template scrunching mechanism, which PrimPol can employ to skip damaged nucleobases (Mourón et al., 2013).

To investigate this phenomenon further and cross-validate the findings we also measured the fidelity of PrimPol using the HSV-tk forward mutation assay (Eckert et al., 2002). We engineered the HSV-tk gene substrates to contain additional short tandem repeat (STR) sequences within the 5' region of the gene, in order to study polymerase fidelity within repetitive sequences. Using $[T]_8$ and $[A]_8$ STR-containing substrates, the observed HSV-tk frequency resulting from PrimPol DNA synthesis is $1400 \pm 360 \times 10^{-4}$ and $900 \pm 210 \times 10^{-4}$, respectively (Table 2.4.). In comparison to the replicative

Polymerase	Total Number of Colonies[a]	Number of White (Mutant) Colonies	Corrected Mutation Frequency[b]	Error Rate[c]
Klenow Taq-Pol A	58,555	96	1.6×10^{-3}	3.6×10^{-5}
Tgo-Pol B (exo-, D215A)	48,167	33	0.7×10^{-3}	1.6×10^{-5}
PrimPol	54,667	264	4.6×10^{-3}	1.0×10^{-4}

[a]The fidelity of each polymerase was determined in three separate experiments, each of which involved scoring *E. coli* colonies on 9 separate plates. The aggregated numbers are given.

[b]Corrected mutation frequency equals: ({number of white colonies/total number of colonies} – background mutation rate). A background mutation rate of 1.7×10^{-5} was used for gapped pSJ4.

[c]Error rate is the number of mistakes made per base incorporated. The corrected mutation frequency was converted to error rate as previously described (Keith et al., 2013). An expression frequency (P) of 0.3 was used. Due to the limited amount of sequencing data an Ni/N value of 1 was used and the number of detectable sites (D) was the sum of the number of determined base substitutions plus insertion/deletions, that is, 147 for pSJ4.

Table 2.2. Mutation frequencies observed for the Klenow fragment Taq-Pol A, Tgo-Pol B (exo-) and human PrimPol enzymes.
The fidelity measurements were determined using a plasmid-based gap filling assay (pSJ4-*lacZ*α). The pSJ4 plasmid assay was developed to study the fidelity of TLS DNA polymerases and it is modified a version of the previously described pSJ3 plasmid (Keith et al., 2013).

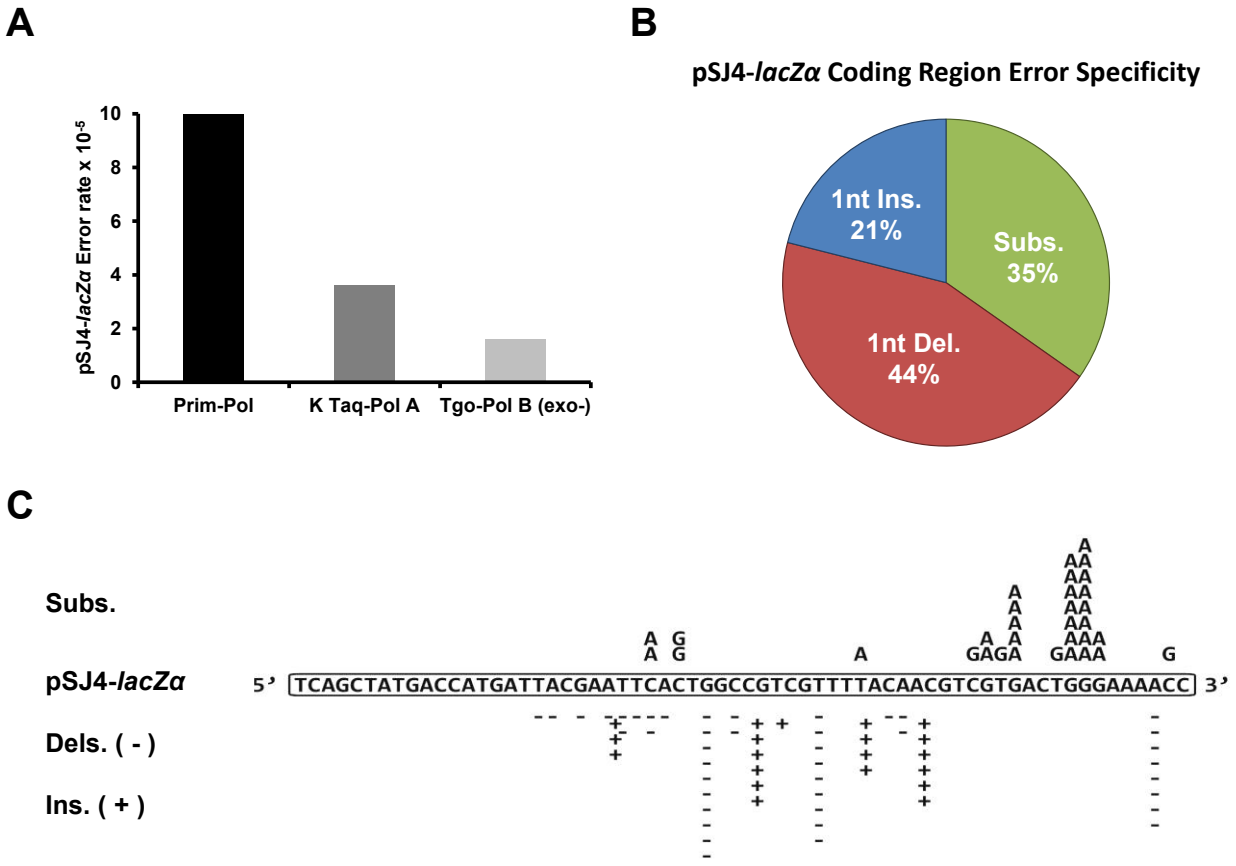


Figure 2.15. Mutational signature of human PrimPol at the lacZα sequence.
(A) The average pSJ4-*lacZα* coding region error rate of PrimPol, relative to other polymerases. The fidelity of each polymerase was determined in three separate experiments, each of which involved scoring *E. coli* colonies on 9 separate plates. The aggregated numbers are given. **(B)** Proportions of base substitutions, insertions, and deletions generated in the pSJ4-*lacZα* coding region. Pie chart depicts percentages from 95 total mutational events observed. **(C)** The diagram shows the 64 nucleotide long sequence of the *lacZα* reporter synthesised by human PrimPol *in vitro* in the pSJ4 gap-filling fidelity assay. A total of 95 white colonies were sequenced revealing a unique mutation signature of human PrimPol.

Mutation Type	Number	Frequency (%)
A>T / T>A	1	1.05
A>C / T>G	2	2.1
A>G / T>C	0	0
G>A / C>T	24	25.3
G>C / C>G	4	4.2
G>T / C>A	2	2.1
Insertions	20	21.05
Deletions	42	44.2
Total	95	100
A>N / T>N	3	3.16
G>N / C>N	30	31.6
Transitions	24	25.3
Transversions	9	9.5

Table 2.3. Mutations generated by PrimPol at the lacZα sequence.
Individual changes introduced into the *lacZ*α indicator gene of pSJ4 by human PrimPol are shown. A total of 95 white colonies were sequenced in order to reveal the types and frequency of mutations.

Mutational Target	Frequency x 10^{-4}	
	T_8	A_8
Observed HSV-tk Frequency ± SD[a]	1400 ± 360	900 ± 210
ssDNA Background	0.7	0.7
Pol EF_{est}[b]	1300 (59)[c]	770 (40)
STR Region	470 (21)	210 (11)
HSV-tk Coding Region	860 (38)	560 (29)
Frameshifts	590 (26)	380 (20)
Large Deletions	110 (5)	100 (5)
Complex	160 (7)	80 (4)
Base Substitutions	<23 (0)	<19 (0)

[a]Mutant frequencies are mean ± standard deviation of three independent reactions

[b]Pol EF_{est} is calculated as described in methods

[c]Number of independent error from two reactions

Table 2.4. PrimPol error rates within STR and HSV-tk coding sequences.

polymerases α and δ, PrimPol displays a 16 to 28-fold increase in HSV-tk frequency (Table 2.5.) (Ananda et al., 2014; Hile and Eckert, 2008). Furthermore, PrimPol displays an error frequency that is more than 2-fold higher than the repair and specialized Pols η, κ, and β.

PrimPol creates errors within the artificial STR sequence at a rate comparable to Pols η, κ, and β (Table 2.5.). Additionally, PrimPol's error specificity within the STR region is in-line with what we have previously observed for other polymerases at these repeats (Figure 2.16.), suggesting that STR errors are not what is driving PrimPol's low fidelity. In contrast to the STR region, PrimPol's error frequency within the coding-region of the HSV-tk gene is higher than any other polymerase analyzed to date (Figure 2.17.A.). In agreement with the *lacZ*α gap-filling assay, PrimPol's coding region error spectrum is almost entirely indel based, with a bias towards deletions (Figure 2.17.B.; Figure 2.18.). The very high proportion of insertion errors (36%) is a phenotype never observed in this assay for other DNA polymerases, and is tremendously different from error-prone Pol η which creates indel frameshift and base-substitution errors at similar rates (Ananda et al., 2014). Although we have observed Pol κ insertion errors, the vast majority of Pol κ indels were deletions (Baptiste and Eckert, 2012), and all of the Pol β indel errors we have observed in the HSV-tk gene were deletion events (Eckert et al., 2002). These findings show that PrimPol has an error specificity unique amongst DNA polymerases.

We observed a pronounced mutational hotspot that included both indels and complex errors within a sequence that can potentially from a hairpin structure (Figure 2.17.C.). Both the complex events and the insertion events result in changes to the template sequence that expand the $[T]_2$ template sequence to a $[T]_3$ or $[T]_4$ sequence. Intriguingly, while such an observation would suggest PrimPol is prone to expand repeated sequences, only a single expansion event was observed at both the $[T]_8$ and $[A]_8$ STR sequences (Figure 2.16.).

2.5. Discussion

PrimPol is a recently discovered primase-polymerase that is potentially important for TLS and repriming during DNA replication in eukaryotic cells. Unlike other TLS enzymes, PrimPol does not interact with, nor is it stimulated by, unmodified or mono-ubuitylated PCNA, indicating that other factors potentially regulate its activities *in vivo*. Nevertheless, our results do not rule out the possibility that PrimPol might interact with PCNA indirectly through an additional bridging partner. In this report we identify a potential regulatory mechanism employed to limit the contribution of PrimPol to DNA replication that is distinct from that used by other TLS polymerases. This mechanism involves SSBs that

Polymerase	Pol EF_{est} x 10^{-4}		
	Overall	HSV-tk Coding Region	STR Region
PrimPol			
T_8 Template	1300 (59)[a]	860 (38)	470 (21)
A_8 Template	770 (40)	560 (29)	210 (11)
Pol η[b,c]			
T_8 Template	440 (35)	210 (17)	230 (18)
A_8 Template	270 (67)	87 (22)	180 (45)
Pol κ[c]			
T_8 Template	460 (49)	140 (15)	320 (34)
Pol α[b]			
T_8 Template	47 (77)	1.8 (3)	45 (74)
A_8 Template	55 (77)	4.3 (6)	51 (71)
Pol δ[c]			
T_8 Template	33 (32)	2.1 (2)	31 (30)
Pol β[b]			
T_8 Template	380 (83)	32 (7)	350 (76)
A_8 Template	400 (82)	15 (3)	390 (79)

[a]Number of independent errors from two reactions

[b]Published data from Ananda et al., 2014.

[c]Published data from Hile et al., 2012

Table 2.5. Comparison of PrimPol error rates within STR and HSV-tk coding sequences with TLS and replicative polymerases.

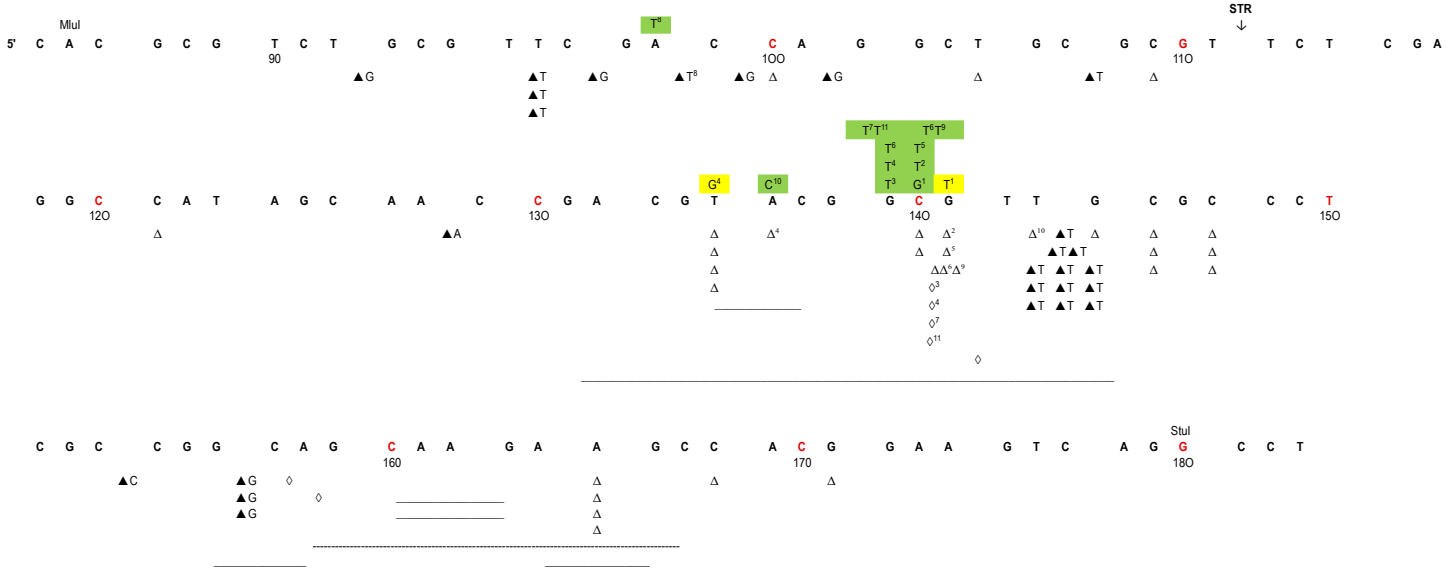


Figure 2.16. PrimPol HSV-tk coding region error spectrum.
Base substitutions are shown above the template sequence, highlighted green for detectable events or yellow for non-detectable. Single deletion and insertion events are shown below the template with open and closed triangles respectively, while diamonds indicate a tandem deletion. Superscripts mark the errors found within an individual complex event. Long lines under the template sequence note >2nt deletion events.

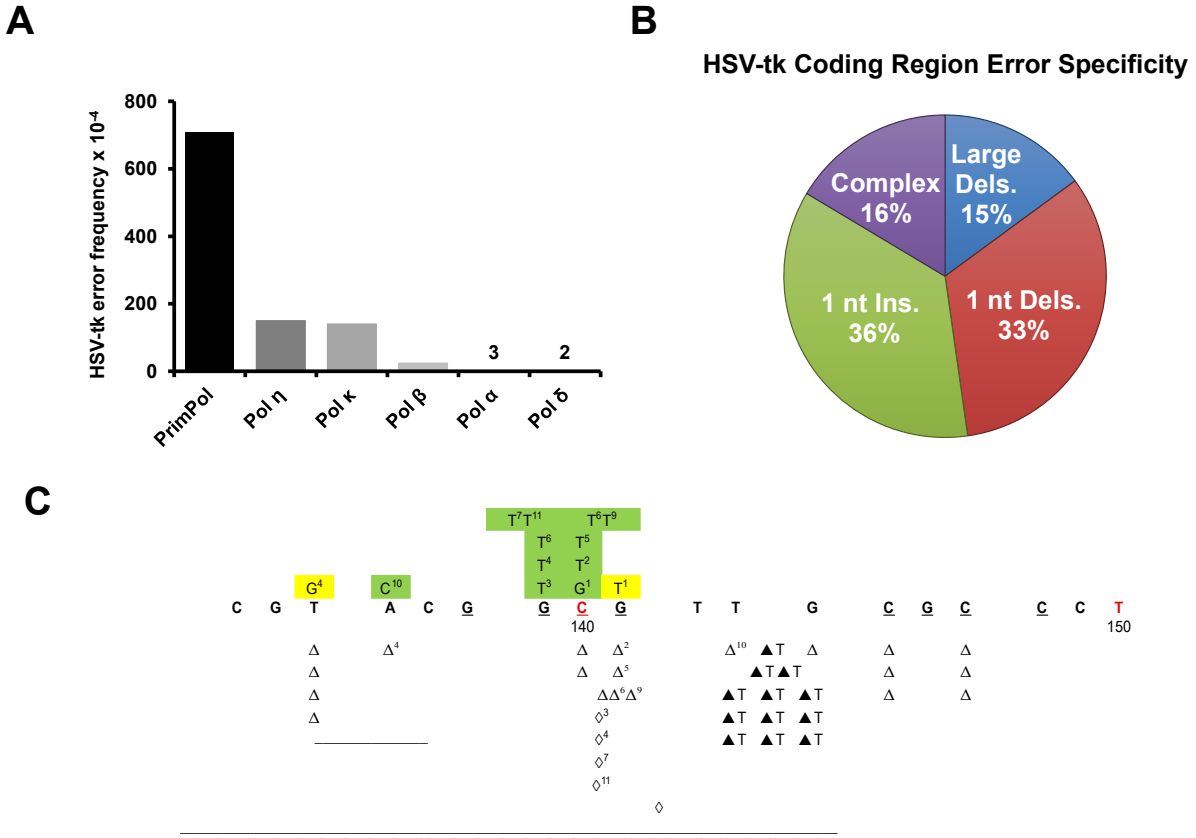


Figure 2.17. PrimPol's replication of the HSV-tk coding region is error-prone and unique.
(A) The average HSV-tk coding region error frequency of PrimPol, relative to other human polymerases. Data generated from Table 2.5. **(B)** Types of errors created by PrimPol in the HSV-tk coding region. Pie chart depicts percentages from 67 total mutational events observed. **(C)** A mutation hot-spot in the coding region is shown to highlight PrimPol's unique error specificity. Base substitutions are shown above the template sequence, and are highlighted green for phenotypically detectable events or yellow for non-detectable events. Single deletion and insertion events are shown below the template with open and closed triangles respectively, while diamonds indicate a tandem deletion. Superscripts mark the errors found within an individual complex event. The underlined bases within the template denote the sequence with hairpin-forming potential.

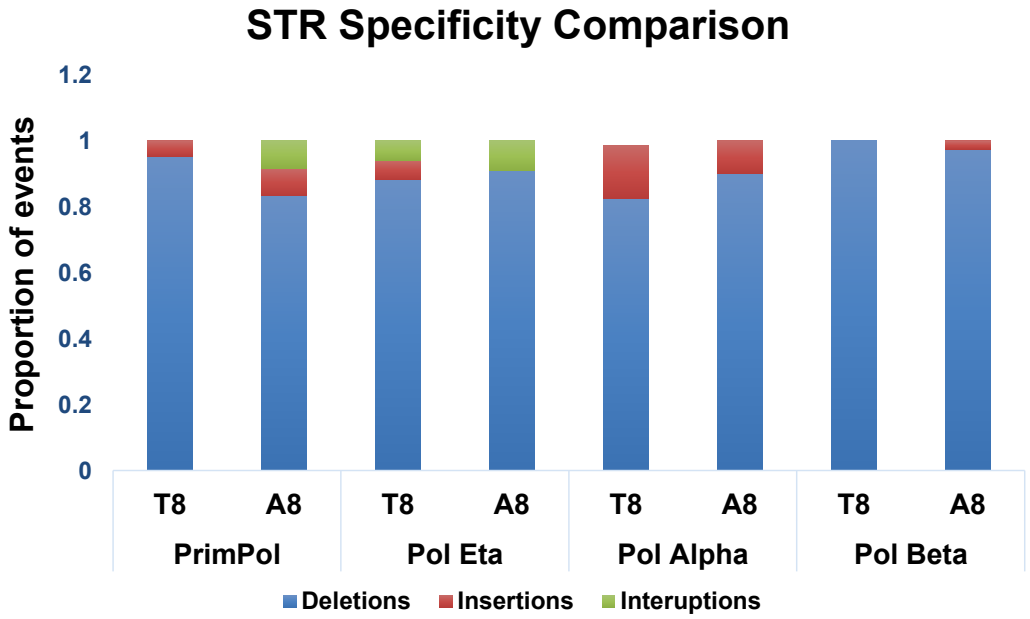


Figure 2.18. The STR is not driving PrimPol's elevated mutation frequency.
Error specificity at the STR is shown for PrimPol in comparison to the other polymerases examined on both templates.

directly restrict DNA synthesis by PrimPol by limiting the availability of ssDNA template, downstream of the stalled replisome, thus preventing re-binding of PrimPol.

It was recently reported that PrimPol interacts with RPA1 and that this interaction is required for foci formation *in vivo* (Wan et al., 2013). These initial studies suggested that RPA may be involved in the recruitment/regulation of PrimPol at sites of DNA damage. Here, the RPA binding domain of PrimPol, in conjunction with the ssDNA binding ZnF domain, may act as a docking mechanism for recruitment of the enzyme. This would allow tethering of PrimPol to stretches of ssDNA partially coated with RPA, following stalling of the replication fork. In this report, we further explored the interaction between PrimPol and SSBs, in addition to the impact that these proteins have on the activity of PrimPol *in vitro.* We identified that PrimPol interacts with mtSSB, as well as RPA (Figure 2.2.C), pertaining to the role of the enzyme in mitochondrial, as well as nuclear, DNA replication (Bianchi et al., 2013; García-Gómez et al., 2013). Furthermore, we establish that PrimPol interacts with the RPA70N protein recruitment domain (Figure 2.4.). This is in contrast to previous reports suggesting that PrimPol interacts with RPA70C (Wan et al., 2013). This revision is consistent with the absence of any other published data suggesting RPA70C is involved in interactions with other proteins. However, it remains possible that PrimPol may have two different sites of interaction on the RPA70 subunit.

It is surprising that both RPA and mtSSB, in addition to the non-interacting T4 SSB, act to significantly impede the primase and polymerase activities of PrimPol, (Figure 2.5. and 2.6.). Previously, RPA has been shown to suppress the ability of the Pol α complex to synthesise primers, identifying a role for RPA in preventing non-specific priming events (Collins and Kelly, 1991). Our results establish that this role also extends to regulating priming by PrimPol, with mtSSB fulfilling an analogous role in mitochondria. Interestingly, this suggests that PrimPol is only able to synthesise primers at regions of the genome not occupied by SSBs, either where SSBs have been displaced by other replication factors, or where the topology of the DNA template prevents SSB binding, for example where DNA secondary structures occur.

Previous studies have shown that RPA can stimulate the polymerase activity of Pol α and Pol δ (Kenny et al., 1989; Tsurimoto and Stillman, 1989), with mtSSB acting to stimulate Pol γ (Oliveira and Kaguni, 2010). In stark contrast, our results show that both RPA and mtSSB severely restrict the polymerase activity of PrimPol (Figure 2.6.). Interestingly, in *E. coli,* SSB inhibits the progression of the TLS polymerase, Pol II (Indiani et al., 2013), and additionally, Pol IV when the interaction between the two proteins is disrupted (Furukohri et al., 2012). We have previously shown that PrimPol displays very

low processivity as a polymerase, incorporating only ~4 bases per binding event, suggesting that the enzyme is only required for very short stretches of DNA synthesis (Keen et al., 2014b). Notably, the ZnF domain of PrimPol, which only binds ssDNA, is involved in modulation of the enzymes processivity (Keen et al., 2014b). This domain is believed to be spatially separated from the polymerase domain, fulfilling a role in regulating both the primase and polymerase activities of PrimPol. This distributive nature of PrimPol likely acts as the primary level of regulation to limit the involvement of PrimPol in DNA synthesis. Indeed, the limiting effect of SSBs on PrimPol's polymerase activity may be in part due to the prevention of rebinding of the enzyme to ssDNA, following its dissociation as a result of its low processivity. However, interestingly, we also find that the 24-354 truncation of PrimPol, lacking the ssDNA binding ZnF domain, is also inhibited by SSBs. These results suggest that, in addition to the low processivity of PrimPol, RPA and mtSSB contribute to restraining the enzyme to limit its potentially mutagenic DNA synthesis during replication restart. However, it is possible that *in vivo* additional remodeling factors may permit synthesis by PrimPol on SSB-coated DNA. The contrasting effects of SSBs on replicative polymerases and PrimPol do, however, suggest a potential mechanism of regulation represented by a model summarised in Figure 2.19.

The stimulatory effect of SSBs on the progression of replicative polymerases permits DNA synthesis on SSB-coated DNA until a lesion is encountered on the template strand (Figure 2.19.A.). The intolerant replicative polymerase stalls at the lesion and idles upstream, as a result of its 3'-5' exonuclease activity (Garg et al., 2004), consequently displacing any surrounding SSBs and generating a ssDNA interface for access of PrimPol (Figure 2.19.B.). PrimPol is then recruited to the SSB bound immediately downstream of the lesion via its RBD, additionally binding the exposed ssDNA interface through its ZnF domain. Subsequently, PrimPol utilises its AEP domain to perform a TLS reaction, extending the primer terminus over the DNA lesion, before further extension is prevented by the downstream SSB (Figure 2.19.C.). Bypass of the lesion allows replication to proceed, with the previously stalled replicative polymerase continuing extension (Figure 2.19.D.).

Importantly, DNA synthesis by PrimPol is limited by SSBs which likely act to ensure that the enzyme only participates in the synthesis of short stretches of DNA. This level of regulation may be necessary due to the mutagenic potential of PrimPol. We provide experimental evidence that PrimPol is a highly error-prone DNA polymerase. Strikingly, in two forward mutation assays, PrimPol shows a strong propensity to indel errors. Within the HSV-tk sequence, PrimPol created indel errors almost exclusively, with base

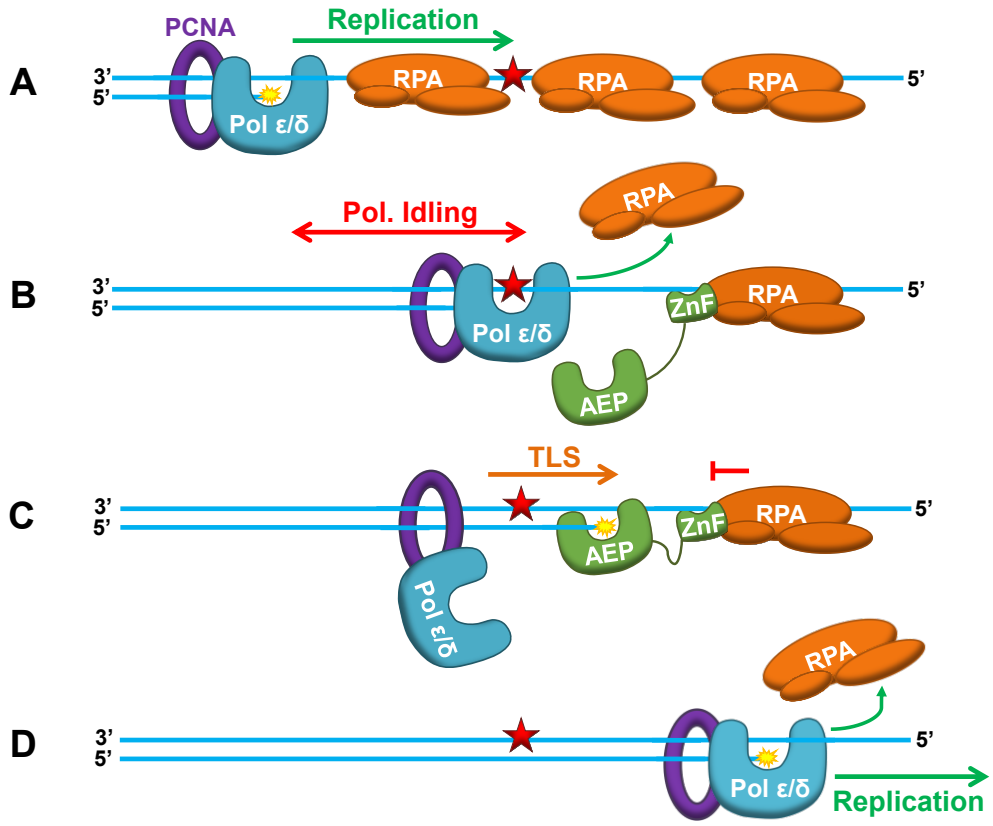


Figure 2.19. Model for regulation of PrimPol synthesis by SSBs during DNA replication.
(A) Unperturbed replication proceeds up to the lesion on the SSB-bound ssDNA template, with the replicative DNA polymerase displacing the bound SSB as it synthesises the daughter DNA strand. **(B)** Replication continues until a lesion is encountered, here the intolerant replicative polymerase stalls at the lesion and idles upstream, displacing any surrounding SSBs and generating a ssDNA interface. This allows recruitment of PrimPol to the downstream SSB, with additional binding to the exposed ssDNA interface through the ZnF domain. **(C**) PrimPol utilises its AEP domain to catalyse a TLS reaction, here the primer terminus is extended over the lesion before further synthesis is prevented by the downstream SSB. **(D)** Following bypass of the lesion, the replicative polymerase again proceeds with replication.

insertions accounting for more than one-third of PrimPol's error spectrum (Figure 2.17.B.). We observed a unique mutational hot-spot for PrimPol (Figure 2.17.C.). This region is rich with [CG] repeats, and contains several sequences with the potential to form hairpin structures. The single base insertion events can be explained by classic primer strand misalignment, while the complex events are more difficult to dissect. However, the common result of both insertion and complex errors within this hotspot is the expansion $[T]_2$ to $[T]_3$, suggesting that the PrimPol complex errors are generated primarily within the loop of the hairpin. PrimPol can displace the template strand when copying sequences with microhomology (Mourón et al., 2013). Possibly, at the HSV-tk hotspot sequence, PrimPol is forced to slip both the primer and template strand to get through the hairpin. While this mechanism is speculative, the data we present do confirm that PrimPol's error specificity is unique from other human polymerases. Together, our observations support the template scrunching mechanism, which PrimPol can employ to skip damaged nucleobases (Mourón et al., 2013). Initially, the scrunching mechanism was observed in the presence of manganese ions during translesion bypass of Ap sites and UV damage lesions (Mourón et al., 2013). Our data suggest that the same scrunching mechanism is utilised by this enzyme when copying non-damaged DNA, in the presence of magnesium ions.

Therefore, PrimPol's *modus operandi* during synthesis appears to be a double-edged sword, facilitating replication restart at bulky lesions (e.g. 6-4PPs) but potentially introducing base insertions and deletions into undamaged templates. This threat to genomic integrity requires tight regulation of the activity of PrimPol during DNA replication and we propose that PrimPol's inability to displace SSBs at the replication fork ensures that the mutagenic potential of this enzyme is greatly limited. In addition, nuclear PrimPol is active primarily in S-phase (Mourón et al., 2013). Intermolecular proofreading of PrimPol synthesis products by either Pols δ or ε could limit mutagenesis. Similarly, in mitochondria, pol γ, which has a very active exonuclease domain, could correct PrimPol's errors. Finally, the error-prone bypass of lesions by PrimPol produces a structure that is readily detected by the post-replicative MMR machinery. Clearly, PrimPol's important role in replication and the prevention of chromosomal instability must be balanced with its potential mutagenic activity.

Chapter 3

PolDIP2 Interacts with Human PrimPol and Enhances its DNA Polymerase Activities

3.1. Abstract

TLS employs specialised DNA polymerases to bypass replication fork stalling lesions. PrimPol was recently identified as a primase and TLS polymerase involved in DDT. Here, we identify a novel PrimPol binding partner, PolDIP2, and describe how it regulates PrimPol's enzymatic activities. PolDIP2 stimulates the polymerase activity of PrimPol, enhancing both its capacity to bind DNA and the processivity of the catalytic domain. In addition, PolDIP2 stimulates both the efficiency and error-free bypass of 8-oxo-dG lesions by PrimPol. We show that PolDIP2 binds to PrimPol's catalytic domain and identify potential binding sites. Finally, we demonstrate that depletion of PolDIP2 in human cells causes a decrease in replication fork rates, similar to that observed in PrimPol$^{-/-}$ cells. However, depletion of PolDIP2 in PrimPol$^{-/-}$ cells does not produce a further decrease in replication fork rates. Together, these findings establish that PolDIP2 can regulate the TLS polymerase and primer extension activities of PrimPol, further enhancing our understanding of the roles of PolDIP2 and PrimPol in eukaryotic DDT.

3.2. Introduction

In eukaryotes, the replicative polymerases α, δ, and ε are primarily responsible for bulk DNA replication. These enzymes, which duplicate DNA with extremely high efficiency and accuracy, are prone to stalling upon encountering helix-distorting DNA lesions generated by DNA damage (Aguilera and Gómez-González, 2008). The inability of the replicative polymerases to synthesise across damaged nucleobases in turn causes replication fork stalling and requires DDT mechanisms in order to proceed with replication and prevent fork collapse (Li and Heyer, 2008; Sale et al., 2012).

A number of distinct replication restart mechanisms exist in order to permit continued replication in the presence of damage. These include the firing of dormant origins downstream of the damage, the generation of new Okazaki fragments on the lagging strand or repriming on the leading strand, the use of an alternative sister template to bypass the damage via HR, and direct synthesis past the damage through TLS (Heller and Marians, 2006; Li and Heyer, 2008; Sale et al., 2012). Whilst it has long been appreciated that specialised DNA polymerases, particularly those of the Y-family, play a key role in eukaryotic DDT by TLS, the role of DNA primases in this process has until recently been mostly overlooked. However, novel roles for primases in DNA repair and damage tolerance are emerging from both archaea and eukarya (Guilliam et al., 2015b). Notably, archaeal replicative primases are now known to display TLS activity (Jozwiakowski et al., 2015), whilst most eukaryotes possess a specialised primase-polymerase (PrimPol) that appears to play roles in TLS and repriming (Rudd et al., 2014).

PrimPol is a member of the AEP superfamily (Guilliam et al., 2015b) and demonstrates primer synthesis capabilities with both rNTPs and dNTPs (Bianchi et al., 2013; García-Gómez et al., 2013; Rudd et al., 2013). In addition, the enzyme displays robust template-dependent TLS polymerase activity, which it utilises to bypass 6-4PPs and 8-oxo-dG lesions (Bianchi et al., 2013; García-Gómez et al., 2013). These activities have been shown to be relevant *in vivo* as cells lacking PrimPol show increased sensitivity to DNA damaging agents and decreased replication fork speeds (Bianchi et al., 2013; Keen et al., 2014b). *In vivo* PrimPol localises to both the nucleus and mitochondria, indeed PrimPol$^{-/-}$ cells also present mtDNA replication defects (Bianchi, 2013; García-Gómez et al., 2013). Unlike canonical Y-family polymerases, PrimPol does not seem to be regulated through interactions with PCNA (Guilliam et al., 2015a). Despite this, PrimPol is a low fidelity polymerase and alternative mechanisms must exist to regulate its activity *in vivo* (Guilliam et al., 2015a). One such regulator is the inherent distributive nature of the enzyme, which limits incorporation to ~4 nucleotides per binding event (Keen et al.,

2014b). In addition, PrimPol's activities are also regulated by its association with SSBs (Guilliam et al., 2015a). Interactions with these proteins may also be involved in the recruitment of PrimPol to the replisome (Wan et al., 2013). Nevertheless, it is likely that additional replication factors also regulate the activity of PrimPol during replication.

In addition to SSBs, PolDIP2 (also known as PDIP38) was also identified in a pull-down screen as a possible cellular binding partner of PrimPol (Guilliam et al., 2015a). Recently, it was reported that PolDIP2 may play a role in DDT, specifically through the regulation of TLS (Maga et al., 2013; Tissier et al., 2010). However, PolDIP2 is a relatively understudied protein, which has been ascribed multiple roles *in vivo* and its function in DNA replication is still unclear. This protein was first identified through yeast two-hybrid screening as a binding partner of the p50 subunit of Pol δ, as well as PCNA (Liu et al., 2003). Further characterisation suggested that PolDIP2 is a mitochondrial protein (Cheng et al., 2005), which inhibits Pol δ and might be involved in Pol δ-mediated viral DNA replication (Xie et al., 2005). However, in contrast to this initial characterisation, more recent studies have identified that PolDIP2 also localises to the nucleus (Wong et al., 2013) and actually stimulates the activity of Pol δ *in vitro* (Maga et al., 2013). Additionally, PolDIP2 has been shown to increase the processivity and fidelity of lesion bypass by Pols λ and η (Maga et al., 2013). The protein was previously found to interact with Pols η, ζ, and Rev1, with depletion causing persistent Pol η nuclear foci and decreased cell survival following UV (Tissier et al., 2010).

Aside from a potential role in DNA replication, PolDIP2 has also been reported to have roles in regulating the nuclear redox environment (Lyle et al., 2009), mitotic spindle formation (Klaile et al., 2008), and in pre-mRNA processing in the spliceosome (Wong et al., 2013). The multitude of roles assigned to PolDIP2 highlights the multi-functional nature of the protein but also obscures the interpretation of many results. This has brought into question the role of PolDIP2 in TLS and DNA replication (Wong et al., 2013), thus necessitating further study to properly characterise its function in these areas.

In this report, we aimed to further explore the regulation of PrimPol, and the role of PolDIP2 in TLS, by investigating the relationship between these two proteins. We observed that PolDIP2 stimulates the polymerase activity of PrimPol. This stimulation appears to be the result of an increase in DNA binding by PrimPol in the presence of PolDIP2, consequently producing an increase in the processivity of the enzyme to levels not previously observed. Furthermore, we found that PolDIP2 alone is sufficient to stimulate the efficiency and fidelity of 8-oxo-dG bypass by PrimPol, an effect similar to that seen with Pols λ and η in the presence of PCNA, RPA and PolDIP2 (Maga et al.,

2013). We used cross-linking and MS analysis to investigate the interaction between PrimPol and PolDIP2. We found that PolDIP2 binds to the catalytic domain of PrimPol and identify potential binding sites, including a region displaying homology to the previously identified PolDIP2-interacting region of Pol η (Tissier et al., 2010). Finally, we explored the role of PolDIP2 in replication *in vivo*. We observed that depletion of PolDIP2 decreased replication fork rates in human cells following UV irradiation. The level of decrease in replication fork rates was similar to that observed in the absence of PrimPol and, additionally, no further decrease in fork speeds was evident when PolDIP2 was depleted in PrimPol$^{-/-}$ cells. Together, these findings support a role for PolDIP2 in regulating TLS and enhancing the primer extension activities of PrimPol during DNA replication. We propose that PolDIP2 acts specifically to enhance PrimPol's primer extension and TLS activities, whilst having minimal effect on its primase activity. Overall, this study provides further evidence for the involvement of both PrimPol and PolDIP2 in DDT during replication in higher eukaryotes.

3.3. Materials and Methods

3.3.1. Expression and Purification of Recombinant Proteins

Full-length Human PrimPol and PrimPol$^{24-354}$ were cloned and purified as described previously (Guilliam et al., 2015a; Keen et al., 2014b). Recombinant GST-PolDIP2, Pol η, PCNA, RPA, and mtSSB, were expressed and purified as previously reported (Biertümpfel et al., 2011; Longley et al., 2009; Maga et al., 2013; Masuda et al., 2007). Untagged PolDIP2 was purified from GST-PolDIP2 through cleavage of the GST tag using PreScission protease before further purification using a GSTrap column and ion exchange chromatography to remove the cleaved GST tag and protease.

PolDIP2$^{51-368}$ was constructed by PCR using the following forward and reverse primers; FWD: 5'-GTTTCTTCATATGCTCTCGTCCCGAAACCGACCAGAGGGCAAA-3', REV: 5'-CAAAGAAGCGGCCGCCTACCAGTGAAGGCCTGAGGG-3', followed by cloning into pET28a via the introduced NdeI and NotI restriction sites. The resulting construct was expressed in *E. coli* at 20°C overnight. Cells were then pelleted before resuspension in buffer A (50 mM Tris–HCl (pH 7.5), 200 mM NaCl, 30 mM imidazole, 10% (v/v) glycerol, 17 μg/ml PMSF, 34 μg/ml benzamidine) supplemented with IGEPAL to a final concentration of 0.5%. Cells were lysed by sonication and the lysate clarified by centrifugation. The clarified lysate was applied to a Ni^{2+}-NTA agarose chromatography column (Qiagen) pre-equilibrated with buffer A. The protein was eluted with buffer A supplemented with 300 mM imidazole following sufficient washing. The resulting eluate was diluted into buffer B (50 mM Tris–HCl (pH 8.5), 50 mM NaCl, and 10% (v/v) glycerol)

and subject to ion exchange chromatography using a 5 ml MonoS column (GE Healthcare) prior to a gradient elution with buffer B containing 1 M NaCl. Fractions containing PolDIP2$^{51-368}$ were further purified by size exclusion chromatography on a Superdex S-75 analytical gel-filtration column (GE Healthcare) equilibrated with buffer C (50 mM Tris–HCl (pH 7.5), 300 mM NaCl, 10% (v/v) glycerol).

Following purification, all proteins were snap frozen in liquid nitrogen and stored at -80°C. Protein concentrations were calculated based on sample absorbance at 280nm and corrected to the protein specific extinction coefficient as determined using the ProtParam tool (ExPASy).

3.3.2. Electrophoretic Mobility Shift Assays

Assays were performed as previously described (Guilliam et al., 2015a) in 20 µl reactions containing 10 mM Bis-Tris-Propane-HCl (pH 7.0), 10 mM $MgCl_2$, 100 µM DTT, 20 nM primer/template substrate (sequences 2 and 9, Table 2.1.), and varying concentrations of PrimPol and PolDIP2 (as indicated on individual figures). Reactions were resolved on 5% (v/v) native polyacrylamide gels at 75 V for 60min in 0.5x TBE buffer. Fluorescently labelled DNA was detected using a FujiFilm FLA-5100 image reader.

3.3.3. DNA Primase Assays

Reactions were assembled in buffer containing 10 mM Bis-Tris-Propane-HCl (pH 7.0), 10 mM $MgCl_2$, and 100 µM DTT, supplemented with 2 µM PrimPol, 250 µM dNTPs, 2.5 µM FAM dNTPs (dATP, dCTP, dUTP) (Jena-Biosciences), 1 µM single-stranded template (sequence 17, Table 2.1), and varying concentrations of PolDIP2 or GST (as indicated on individual figures). Reactions were incubated at 37°C for 15 mins before quenching in binding-washing (B-W) buffer (10 mM Tris-HCl pH 7.5, 500 mM NaCl, 10 mM EDTA). Quenched reactions were incubated with ~20 µl streptavidin coated beads for 1 hour at 4°C to allow capture of the DNA templates and primase reaction products. Following capture, beads were washed 3 times with 1 mL volumes of B-W buffer before final suspension in 20 µL stop buffer (95% formamide solution with 0.25% bromophenol blue and xylene cyanol dyes). Samples were then boiled for 5 mins and products detected by resolution on a 15% (v/v) polyacrylamide gel containing 7 M urea and 1x TBE buffer run at 850 V for 2.5 hours in 1x TBE. Reaction products were visualised using a FujiFilm FLA-5100 image reader.

3.3.4. DNA Primer Extension Assays

Reactions (20 µl) were assembled containing; 20 nM 5’ Hexa-chlorofluorescein (HEX)-labelled DNA primers annealed to the appropriate DNA templates (Sequences 1-5, 7, 9,

11-15, and 17, Table 2.2.), 10 mM Bis-Tris-Propane-HCl (pH 7.0), 10 mM $MgCl_2$, 100 µM DTT, and 100 µM dNTPs. Reactions were supplemented with varying amounts of PrimPol or Pol η (as indicated in figure legends), and incubated at 37ºC (time points are shown in figure legends). Where present, accessory proteins were added prior to the addition of the enzyme at the concentrations indicated on figures. In the case of single nucleotide incorporation analysis, dNTPs were substituted for 100 µM of the single dNTP being analysed (dATP, dCTP, dGTP, or dCTP). Reactions were quenched with buffer containing 95% formamide, 0.05% bromophenol blue, 0.09% xylene cyanol, and 200 nM competitor oligonucleotide. Quenched reactions were heated to 95ºC for 5 mins before electrophoresis on a 15% (v/v) polyacrylamide/7 M urea gel. Extended fluorescent primers were imaged using a FujiFilm FLA-5100 image reader. All quantification was performed using ImageQuant TL software (GE Life Sciences). Data were plotted and analysed using GraphPad Prism 6.

3.3.5. Polymerase Processivity Assays

PrimPol's processivity in the presence of varying amounts of PolDIP2 was analysed using the method previously described (Keen et al., 2014b). Extension reactions were assembled containing 100 nM PrimPol, varying concentrations of PolDIP2, 40 nM primer/template substrate (sequences 2 and 9, Table 2.2.), 10 mM Bis-Tris-Propane-HCl (pH 7.0), 10 mM $MgCl_2$, and 100 µM DTT, and incubated at 37ºC. Reactions were initiated by supplementation with 100 µM dNTPs and 1 mg/ml sonicated herring sperm trap DNA. Reaction products were monitored over a time course and quenched at various time points (as indicated in figure legends) using buffer containing 95% formamide, 0.05% bromophenol blue, 0.09% xylene cyanol, and 200 nM competitor oligonucleotide. The efficiency of the trap DNA was analysed using control reactions containing the trap DNA in the initial reaction assembly to ensure no extension was observed. Extension products were resolved and imaged as described in 'DNA primer extension assays'. Reaction products were quantified using ImageQuant TL software (GE Life Sciences) and the previously described method (Keen et al., 2014b).

3.3.6. Crosslinking and Mass Spectrometry Analysis

Purified untagged-PolDIP2 and $PrimPol^{24-354}$ were mixed at equimolar concentrations in buffer containing 50 mM HEPES (pH 7.5), 150 mM NaCl, and 10% (v/v) glycerol, for 30 mins on ice to allow binding. Following this, protein samples were supplemented with bis(sulfosuccinimidyl)suberate (BS^3) crosslinker at increasing concentrations (from 1:1 to 20:1 crosslinker:protein molar ratios). Samples were incubated on ice for 45 mins to 1 h to allow crosslinking reactions to proceed, before quenching with 50 mM Tris and

further incubation for 15 mins. Crosslinked samples were supplemented with Laemmli buffer and resolved by electrophoresis on a polyacrylamide gel. Bands corresponding to a 1:1 PrimPol:PolDIP2 complex molecular weight were excised and processed for MS as described (Shevchenko et al., 2007).

MS samples were analysed using a nano-LC-MS (ThermoFisher U3000 nanoLC and Orbitrap XL mass spectrometer) as previously described (Hatimy et al., 2015). Raw MS and MS/MS spectra were converted to the .mgf format using Compass (Wenger et al., 2011) and searched against the SwissProt database with Mascot (Matrix Science). Search parameters employed a precursor tolerance of 5 ppm and a fragment ion tolerance of 0.6 Da. Crosslinked peptides were identified by analysing .mgf files using StavroX crosslinking analysis software as described previously (Götze et al., 2012).

3.3.7. Cell Culture and DNA Fibre Analysis

MRC5 cells were cultured in MEM supplemented with 15% FCS, 1% L-glutamine and 1% PenStrep (v/v), at 37°C in CO_2 controlled incubators. Cells were transfected with PolDIP2 siRNA (SMARTpool ON-TARGET plus siRNA Thermo Fisher Scientific), or mock siRNA treated, using Lipofectamine RNAiMAX (Invitrogen) according to the manufacturer's instructions. 72 h following siRNA transfection cells were subject to DNA fibre analysis as described previously (Bianchi et al., 2013). All DNA fibre analysis was performed in triplicate. MRC5 PrimPol$^{-/-}$ cells were generated using the zinc finger nuclease knockout method according to the manufacturer's recommendations (Sigma-Aldrich).

3.4. Results

3.4.1. PolDIP2 Stimulates the Polymerase Activity of PrimPol

PolDIP2 was originally identified in a large-scale pull-down/MS screen previously performed to identify cellular binding partners of PrimPol (Guilliam et al., 2015a; Rudd, 2013). This screen also identified the SSBs, RPA and mtSSB, as interacting partners of PrimPol, whilst an interaction with PCNA was not established. Further studies suggested that RPA and mtSSB regulate PrimPol's enzymatic activities and revealed that, unlike canonical TLS polymerases, PrimPol is not stimulated by the presence of PCNA *in vitro* (Guilliam et al., 2015a). In light of reports implicating PolDIP2 in the regulation of TLS (Maga et al., 2013; Tissier et al., 2010), we aimed to analyse whether PolDIP2 might also act as a PrimPol regulator. To do this, we employed primer extension assays on a 20/50-mer DNA primer/template substrate (sequences 2 and 5, Table 2.2.), in the presence of increasing concentrations of GST-PolDIP2. When titrated into primer

extension reactions containing PrimPol, we observed that PolDIP2 stimulated the activity of the enzyme, producing an increase in the amount of extended primers in a dose-dependent manner (Figure 3.1.A.). Notably, PolDIP2 generated a similar effect when titrated into reactions containing PrimPol$^{1-354}$, a truncation of the enzyme comprising the catalytic AEP domain only. Intriguingly, when plotted, these data produced a sigmoidal kinetic profile, suggesting that PrimPol may bind multiple PolDIP2 molecules (Figure 3.1.B.). Indeed, assays using both full-length PrimPol and PrimPol$^{1-354}$, generated Hill coefficients of 5.176 ± 1.481 and 5.258 ± 1.466 respectively, revealing positive cooperativity in PolDIP2 binding. The K_{half} values for both PrimPol and PrimPol$^{1-354}$ were 41 ± 1.817 nM and 33.93 ± 2.112 nM, respectively, with stimulation slightly more pronounced for the truncated enzyme.

Furthermore, at higher PrimPol and PolDIP2 concentrations an increase in the length of extended primers, with a significant increase in the amount of fully extended primers, was observed (Figure 3.1.C. and D. and Figure 3.2.). Notably, the GST-tag was not responsible for this stimulation as GST alone did not affect the polymerase activity of PrimPol (Figure 3.2.). In line with previous reports (Maga et al., 2013), GST-PolDIP2 was used for these assays due to the ease of purification and increased solubility over the untagged protein. These results establish that PolDIP2 is able to stimulate the polymerase activity of PrimPol, increasing both the amount and length of extended primers.

Given the stimulatory effects of PolDIP2 on the polymerase activity of PrimPol, we next sought to assess if it also modulated PrimPol's primase activity. To determine this, we analysed the primase activity of the enzyme on a 66-mer mixed sequence ssDNA template (sequence 17, Table 2.1.) in the presence of increasing concentrations of GST-PolDIP2 or GST alone. As observed previously (Schiavone et al., 2016), PrimPol was able to synthesise primers on this ssDNA template in the absence of PolDIP2 or GST. When present, PolDIP2 did not significantly increase the amount of primers synthesised (Figure 3.3.). However, in the presence of PolDIP2, PrimPol did appear to extend generated primers further. This supports a scenario where PolDIP2 is unable to increase the rate at which primers are synthesised but is able to increase the rate and length to which these primers are extended.

3.4.2. PolDIP2 Enhances PrimPol's DNA Binding

PrimPol has previously been shown to bind relatively poorly to DNA (Keen et al., 2014b), thus it seems likely that additional factors assist it in the coordination of DNA *in vivo.* Previously, it was reported that PolDIP2 increases the DNA binding affinity of Pol λ, whilst

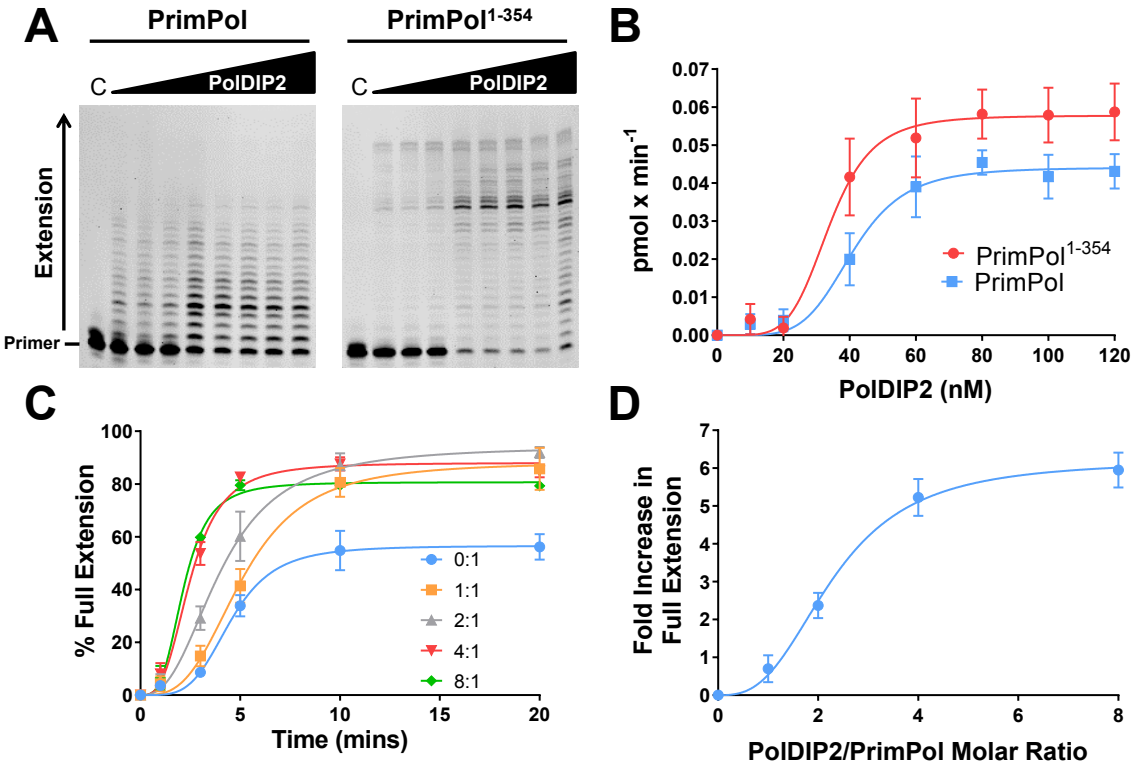


Figure 3.1. PolDIP2 stimulates the polymerase activity of PrimPol.
(A) PrimPol or PrimPol$^{1-354}$ (20 nM) were incubated with 5' labelled primer/template DNA substrates (20/50-mer; 20 nM) and dNTPs (100 μM) in the presence of increasing concentrations of GST-PolDIP2 (0, 10, 20, 40, 60, 80, 100, and 120 nM) for a single 5 minute timepoint. 'C' indicates the no enzyme control reaction. **(B)** Relative increase in the rate of primer extension by PrimPol and PrimPol$^{1-354}$ in the presence of increasing GST-PolDIP2 concentrations (0, 10, 20, 40, 60, 80, 100, and 120 nM). Data were fitted using GraphPad Prism 6 software. Values are the means of four independent experiments. Error bars are ± SD. **(C)** PrimPol generates a greater proportion of fully extended primers in the presence of PolDIP2. PrimPol (100 nM) was incubated with 5' labelled primer/template DNA substrates (20/50-mer; 20 nM) and dNTPs (100 μM) in the presence of increasing GST-PolDIP2 concentrations (0, 0.1, 0.2, 0.4, and 0.8 μM) for increasing timepoints (0, 1, 3, 5, 10, and 20 mins) Fully extended primers (as indicated on Figure 3.2.) were quantified for each timepoint as a percentage with respect to the total primers present. Data were fitted using GraphPad Prism 6 software. Values are the means of three independent experiments. Error bars are ± SD. Representative gels used for quantification are shown in Figure 3.2. **(D)** Fold increase in fully extended primers by PrimPol in the presence of increasing PolDIP2/PrimPol molar ratios at a single 3 min timepoint. Data were fitted using GraphPad Prism 6 software. Values are the means of three independent experiments. Error bars are ± SD. Representative gels used for quantification are shown in Figure 3.2.

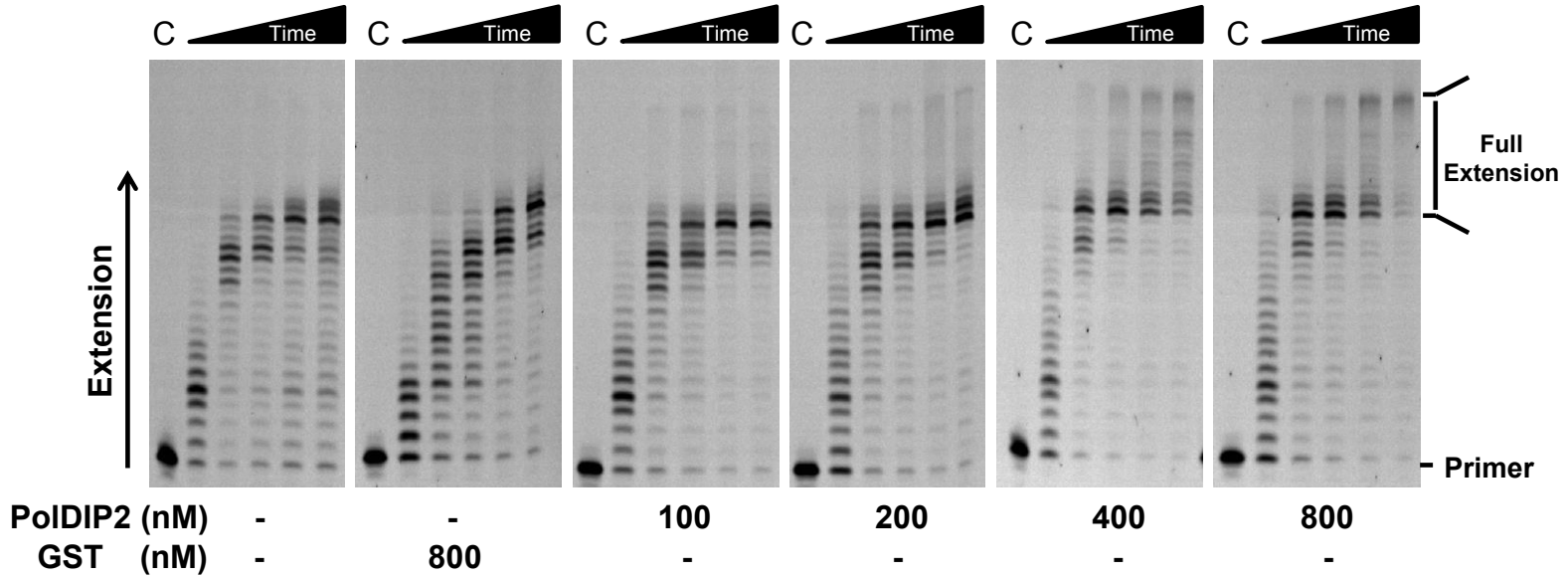


Figure 3.2. PolDIP2 stimulates full-length primer extension by PrimPol.
PrimPol (100 nM) was incubated with 5' labelled 20/50-mer primer/template substrates (20 nM) and dNTPs (100 μM) in the absence or presence of increasing concentrations of GST-PolDIP2 over a time course (1, 3, 5, 10, 20 mins). 'C' indicates the no enzyme control. Quantification of the data is shown in Figure 3.2.C. and D.

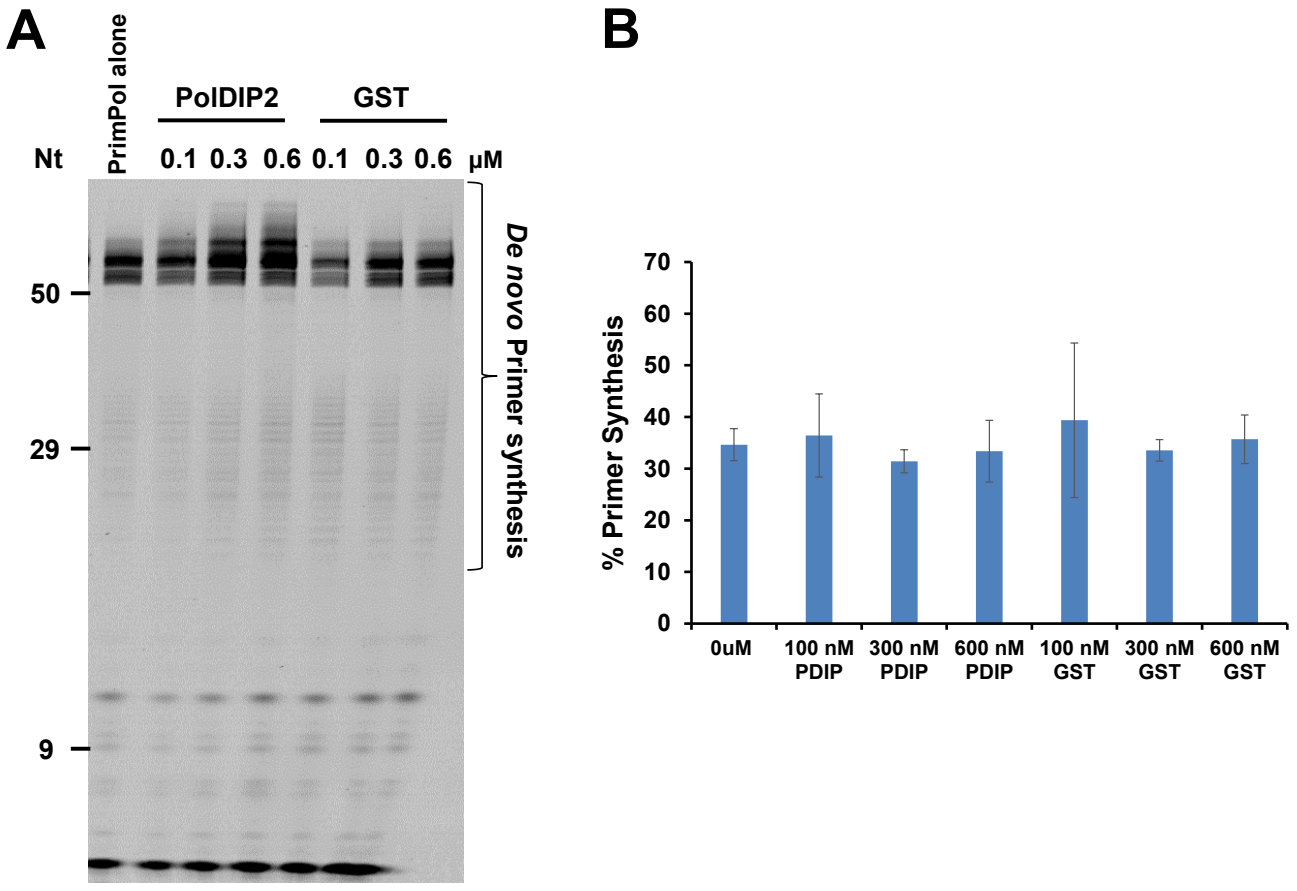


Figure 3.3. PolDIP2 does not increase primer synthesis by PrimPol.
(A) The primase activity of PrimPol was analysed in the absence and presence of GST-PolDIP2 or GST only over a range of concentrations (as indicated). **(B)** Quantification of the data shown in A, showing the percentage of primer synthesis reaction products generated in the presence of varying concentrations of GST-PolDIP2 or GST only. Nucleotide (Nt) size markers are shown on the left of the figure.

lacking the capacity to bind DNA itself (Maga et al., 2013). Consequently, the effect of PolDIP2 on the DNA binding of PrimPol was analysed. To this end, EMSAs were performed using a 20/97-mer DNA primer/template substrate (sequences 2 and 9, Table 2.1.) in the presence of varying amounts of untagged-PolDIP2 and PrimPol. In order to analyse the effect of PolDIP2 on the DNA binding affinity of the catalytic domain of PrimPol only, a truncation of the enzyme ($PrimPol^{24-354}$) was used that contained only the AEP domain. Importantly, this eliminates possible binding artefacts being introduced by the presence of the ssDNA-binding ZnF domain (Keen et al., 2014b).

When titrated into EMSAs supplemented with a fixed concentration of PolDIP2, PrimPol bound to a significantly increased amount of DNA compared to EMSAs with PrimPol alone (Figure 3.4.A. and B.). Consistent with previous reports, PolDIP2 alone showed no ability to bind the DNA substrate (Maga et al., 2013), suggesting that this increase in binding was due to PolDIP2's effect on PrimPol. A smaller increase in binding was also observed when PolDIP2 was titrated into EMSAs with a fixed concentration of PrimPol (Figure 3.4.C and D.). Together, these data show that PolDIP2 exerts a similar influence on the DNA binding avidity of PrimPol, as previously reported for Pol λ (Maga et al., 2013). Furthermore, this confirms that PrimPol forms a complex with PolDIP2 on the DNA. As these assays were conducted with the AEP domain alone, it suggests that PolDIP2 binds PrimPol via the catalytic domain. This is in agreement with the stimulation observed in primer extension assays, as is also the case with Pol λ (Maga et al., 2013).

3.4.3. PolDIP2 Increases the Processivity of PrimPol

PrimPol is a poorly processive polymerase, incorporating only ~4 nt per binding event (Keen et al., 2014b). Interestingly, it has previously been shown that this low processivity is partially due to the restraining effect of the ZnF domain, in combination with the enzyme's weak affinity for DNA binding (Keen et al., 2014b). The ssDNA binding activity of the ZnF domain is thought to produce inter-domain collisions with the catalytic AEP domain following synthesis of ~4 nt, thus limiting PrimPol's contribution to DNA synthesis to very short stretches. However, in light of the fact that PolDIP2 can increase PrimPol's DNA binding ability, we hypothesised that the protein may also increase PrimPol's processivity. To investigate this, we employed a standard primer extension assay on DNA primer/template substrates (20/97-mer) in the presence of excess unlabelled trap DNA. Pre-incubation of PrimPol and the DNA template before initiation with dNTPs and a DNA trap allowed incorporation during a single association/dissociation event to be analysed and thus enabled us to determine the processivity of the enzyme.

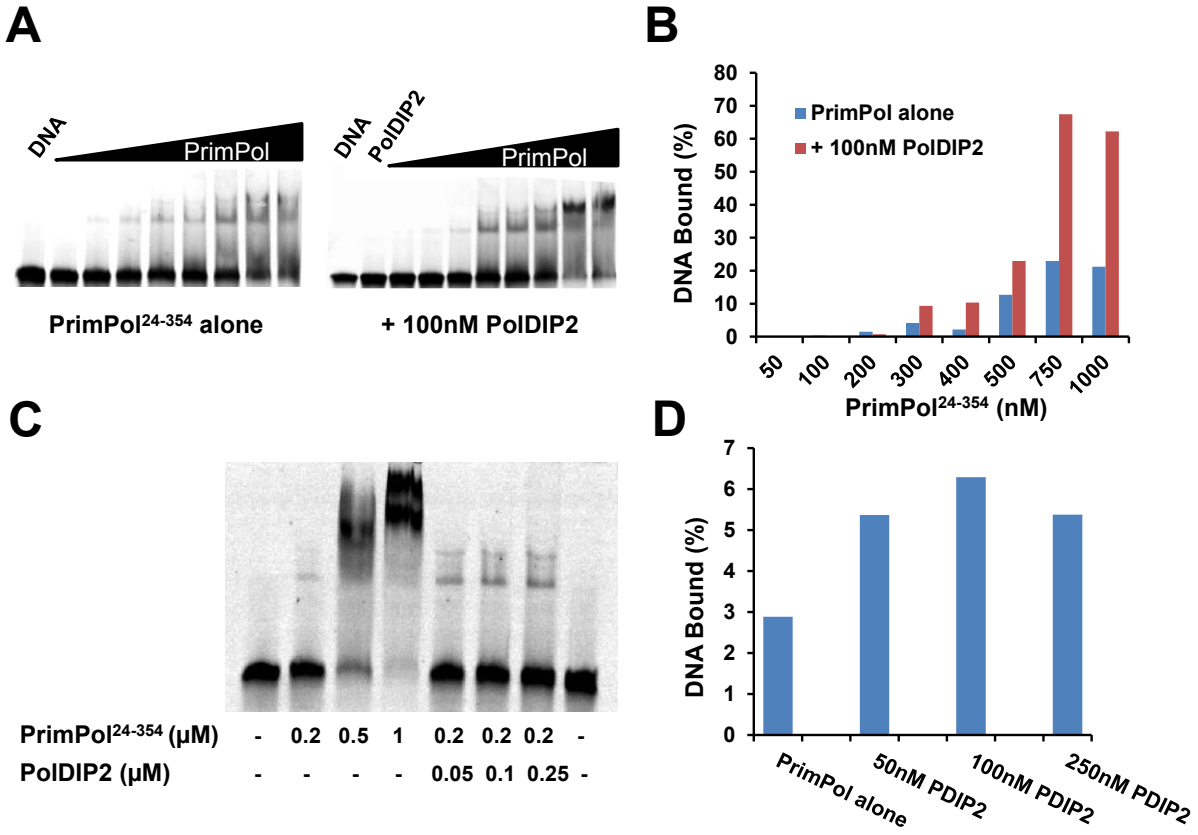


Figure 3.4. PolDIP2 enhances the DNA binding of the AEP domain of PrimPol.
(A) Increasing concentrations of PrimPol$^{24-354}$ (0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75, 1 μM) were incubated in EMSA reactions containing 5'-labelled 20/97-mer primer/template substrates in the absence or presence of untagged-PolDIP2 (100 nM). The 'PolDIP2' lane indicates the no PrimPol control, showing that PolDIP2 (100 nM) alone does not bind to the primer-template substrate. **(B)** Quantification of the data presented in A. For each PrimPol concentration the percentage of DNA bound (in relation to the total DNA) was calculated and compared for EMSAs containing PrimPol only, or PrimPol and PolDIP2. **(C)** PrimPol$^{24-354}$ alone (0, 0.2, 0.5 and 1 μM) and with increasing concentrations of PolDIP2 (200 nM PrimPol; 0, 50, 100 and 250 nM PolDIP2) was incubated in EMSA reactions containing 5'-labelled 20/97-mer primer-template substrates. **(D)** Quantification of the data shown in 'C'. For each PolDIP2 concentration the percentage of DNA bound (in relation to the total DNA) was calculated. Reactions containing PolDIP2 only again confirm that PolDIP2 alone does not bind to the primer-template substrate.

In the absence of PolDIP2, PrimPol's processivity was in line with the previously determined levels, confirming the efficiency of the DNA trap (Keen et al., 2014b). However, when titrated in identical reaction conditions, PolDIP2 produced a significant dosage-dependent increase in the processivity of PrimPol (Figure 3.5.A., B., and C.). At the highest concentration of PolDIP2 (1.6 μM) and longest time point (2 mins), products of more than 16 nt in length were visible, representing a > 4-fold increase in PrimPol's processivity. In addition to the increased length of the synthesised products, PrimPol also produced longer products more rapidly in the presence of PolDIP2. This is evidenced by analysing reaction products from the shortest time point where, in the absence of PolDIP2, PrimPol had still not synthesised products of 4 nt in length. In contrast, in the presence of PolDIP2 products > 8 nt in length were apparent. Again, no stimulation of processivity was observed in the presence of GST alone, confirming that the GST tag is not causing this effect. Furthermore, untagged PolDIP2 was able to produce similar increases in processivity when used at higher concentrations (Figure 3.6.). Higher concentrations were probably required due to the decreased solubility of the protein in the absence of the GST tag. Together, these data reveal that in the presence of PolDIP2, PrimPol produces longer products more efficiently in a single binding event compared to PrimPol alone. Importantly, these results suggest that PrimPol is potentially involved in synthesis of longer stretches of DNA than previously thought.

3.4.4. PolDIP2 Does Not Allow PrimPol to Displace SSBs

We previously observed that, unlike many replicative polymerases, PrimPol was unable to displace both RPA and mtSSB from DNA during synthesis and we proposed a model whereby these SSBs regulate PrimPol's activity to restrict the enzyme's potentially mutagenic contribution to DNA replication (Guilliam et al., 2015a). However, in light of the increased processivity and DNA binding potential of PrimPol when in complex with PolDIP2, we postulated that this complex might be able to overcome negative regulation by SSBs. To test this hypothesis, we employed standard primer extension assays with PrimPol in the presence of PolDIP2 and RPA or mtSSB. Here, PolDIP2 was unable to relieve the inhibitory effects of RPA and mtSSB on the primer extension activity of PrimPol (Figure 3.7.). In each case, primer extension was significantly inhibited when compared to reactions in the absence of accessory proteins. These results show that even in the presence of PolDIP2, PrimPol is unable to displace SSBs from DNA and is therefore unable to overcome their negative regulatory effects.

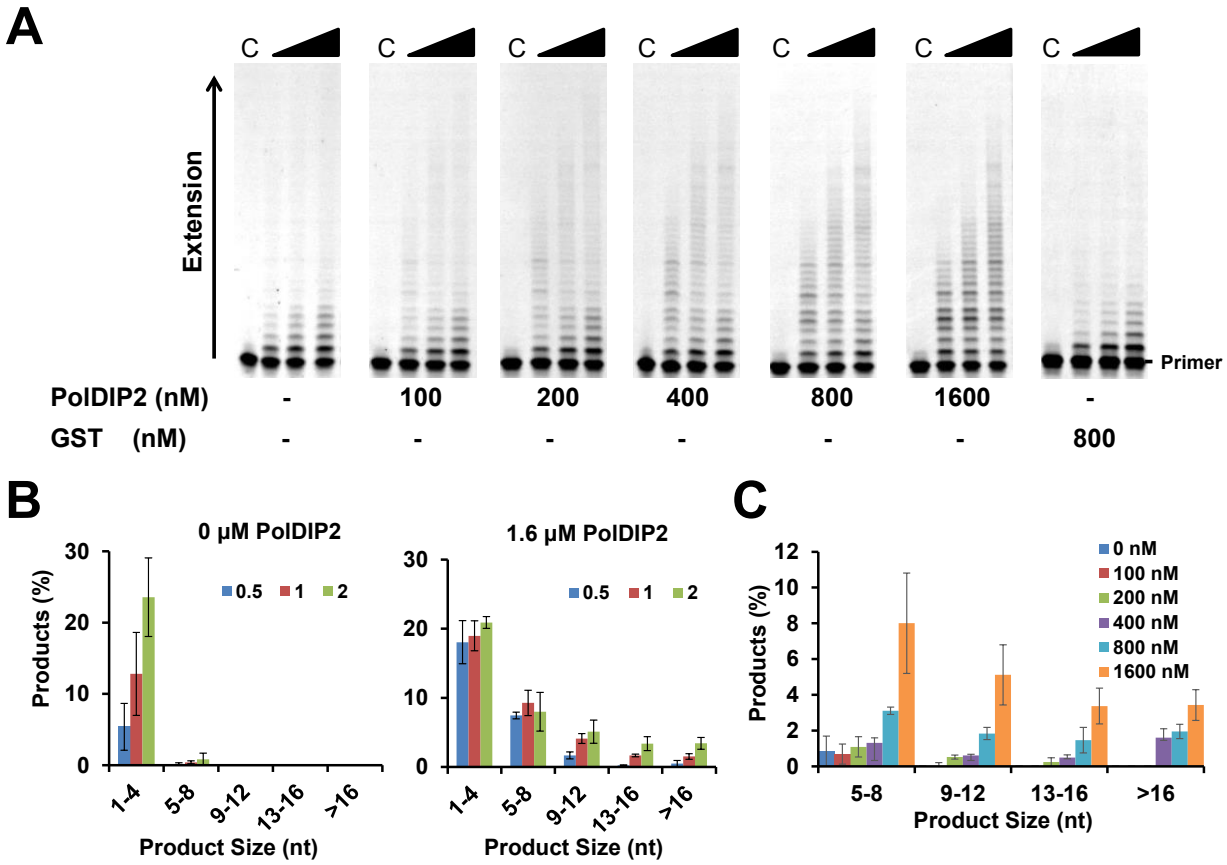


Figure 3.5. PolDIP2 enhances the processivity of PrimPol.
(A) PolDIP2 was titrated into reactions containing PrimPol (100 nM) and 5' labelled 20/97-mer primer/template substrates (20 nM). Reactions were initiated with dNTPs (100 μM) and excess trap DNA and quenched at 0.5, 1, and 2 min time-points. Reactions containing GST only show no increase in PrimPol's processivity. 'C' indicates the control reaction with trap DNA added before the enzyme. **(B)** Quantification of processivity reactions containing either PrimPol alone, or PrimPol and PolDIP2 (1.6 μM). Reaction products were quantified as a function of their size in relation to the total primers present for each time-point. Data represent the means of three independent experiments. Error bars are ± SD. **(C)** Quantification of PrimPol processivity in the presence of increasing GST-PolDIP2 concentrations (as shown) at the 2 min timepoint. Reaction products were quantified as a function of their size in relation to the total primers present for each time-point. Data represent the means of three independent experiments. Error bars are ± SD.

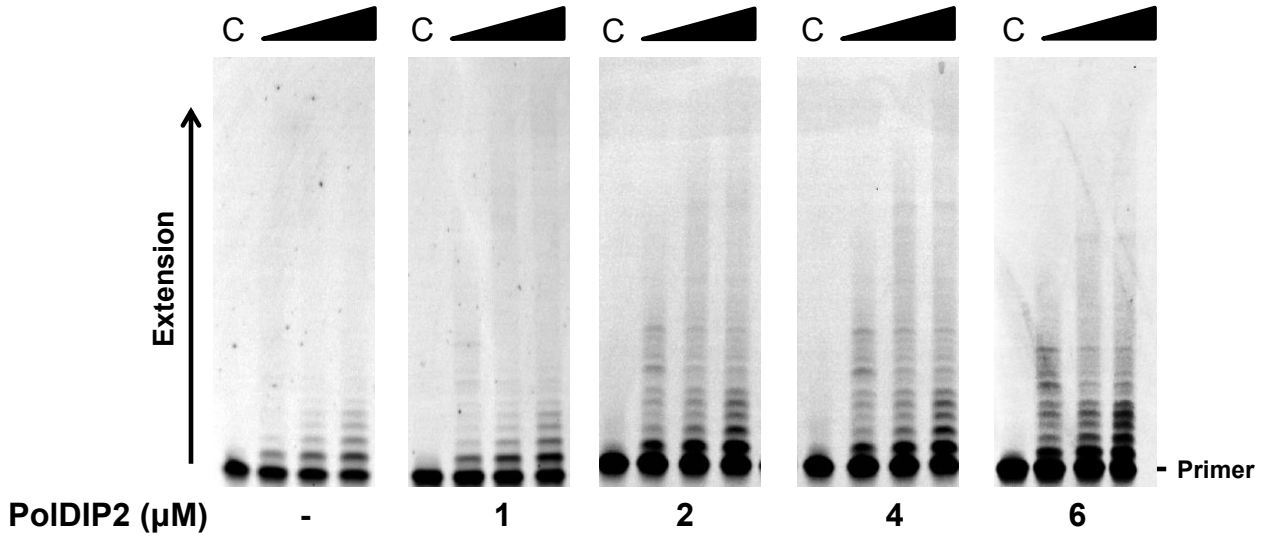


Figure 3.6. Untagged PolDIP2 increases PrimPol's processivity at higher concentrations.
Untagged-PolDIP2 was titrated into reactions containing PrimPol (100 nM) and 5' labelled 20/97-mer primer/template substrates (20 nM). Reactions were initiated with dNTPs (100 μM) and excess trap DNA and quenched at 0.5, 1, and 2 min time-points.

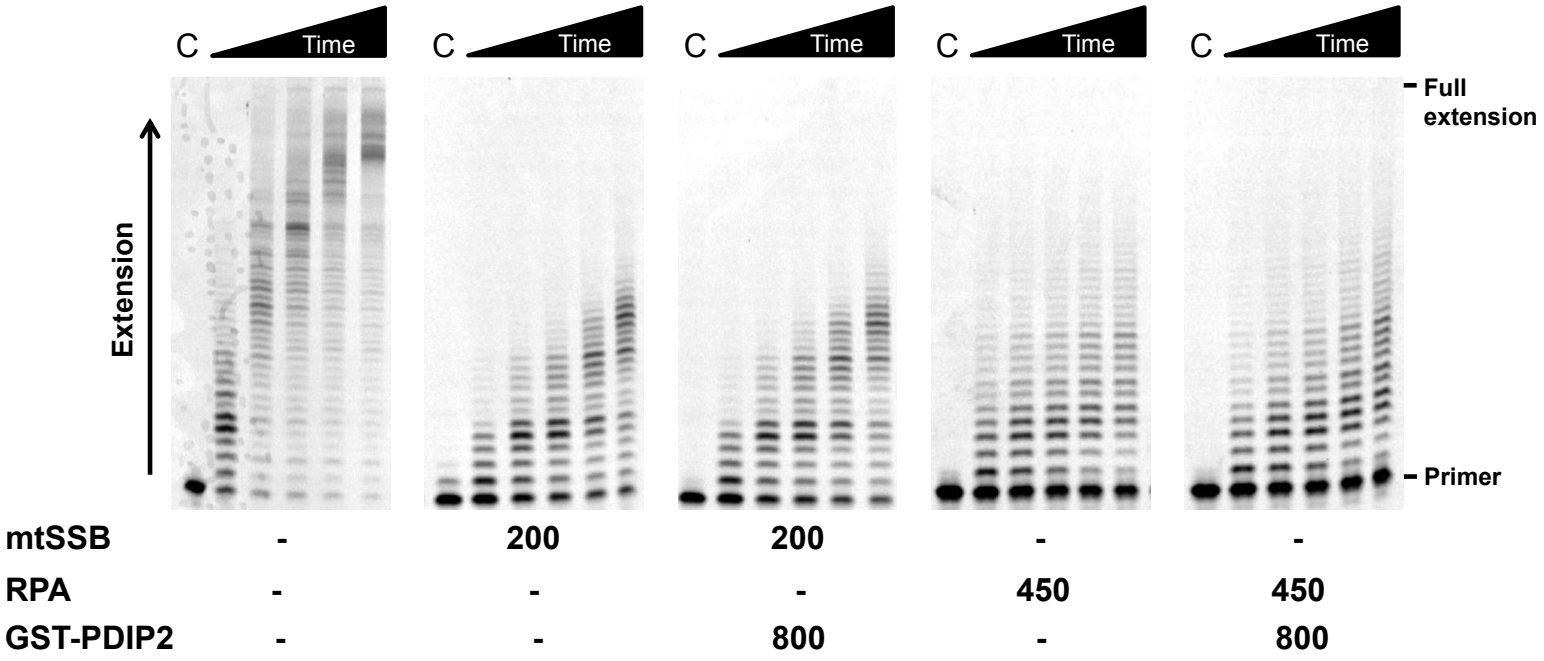


Figure 3.7. PolDIP2 does not allow PrimPol to displace mtSSB or RPA.
The primer extension activity of PrimPol was analysed in the absence and presence of PolDIP2 and mtSSB or RPA using a 5' labelled 20/97-mer primer/template substrate. 'C' indicates the no enzyme control.

3.4.5. PrimPol is Inhibited in the Presence of Both PolDIP2 and PCNA

PolDIP2 has previously been found to interact with PCNA (Liu et al., 2003). Coupled with this, it has been shown that the protein can increase the affinity of Pol δ for PCNA, resulting in increased stimulation of the enzyme's polymerase activity in presence of PolDIP2 and PCNA over either factor alone (Maga et al., 2013). These studies suggest that PolDIP2 is able to act as a bridging factor to help tether polymerases to PCNA, leading to further stimulation of their activity. It was previously found that PrimPol does not interact with, and is not stimulated in the presence of, PCNA (Guilliam et al., 2015a). This apparent lack of interaction and stimulation by PCNA is in contrast with canonical TLS polymerases, leading to speculation that PrimPol is not regulated by the PCNA-mediated polymerase switch mechanism, which regulates the activity of the Y-family TLS polymerases. However, given that PolDIP2 can interact with PCNA, it is possible that the protein might act as a bridging partner between PrimPol and PCNA, thus allowing PrimPol to be regulated by the classical PCNA-mediated polymerase switch model. To test this, we again used standard primer extension assays with PrimPol in the presence of PCNA alone or PCNA and PolDIP2. As shown previously, in the presence of PCNA PrimPol's activity is not affected, with no stimulation or inhibition observed compared to PrimPol alone (Figure 3.8.). However, somewhat unexpectedly, in the presence of both PCNA and PolDIP2, PrimPol's polymerase activity was actually inhibited when compared to reactions with the enzyme alone, or with PCNA only. This inhibitory effect may be due to PolDIP2 associating with, and stabilising, PCNA on the primer/template substrate and in turn blocking access by PrimPol. Nevertheless, this suggests that PolDIP2 does not act to tether PrimPol to PCNA, and in opposition to what has previously been reported with Pol δ, PCNA prevents stimulation of PrimPol by PolDIP2. These results further support the proposal that PrimPol is regulated by a mechanism distinct from that employed by canonical Y-family TLS polymerases.

3.4.6. PolDIP2 Increases the Efficiency and Fidelity of 8-oxo-dG Bypass by PrimPol

PrimPol has previously been shown to possess TLS polymerase activity, displaying an ability to tolerate templating 8-oxo-dG lesions and 6-4PPs (Bianchi et al., 2013; García-Gómez et al., 2013; Keen et al., 2014b). It has been postulated that the ability of PrimPol to bypass 8-oxo-dG lesions may be of particular importance in the mitochondria given the localisation of PrimPol there and the fact that Pol γ deals poorly with these lesions (Bianchi, 2013; García-Gómez et al., 2013). Intriguingly, PolDIP2 has also been shown to localise to the mitochondria (Cheng et al., 2005; Xie et al., 2005), in addition to

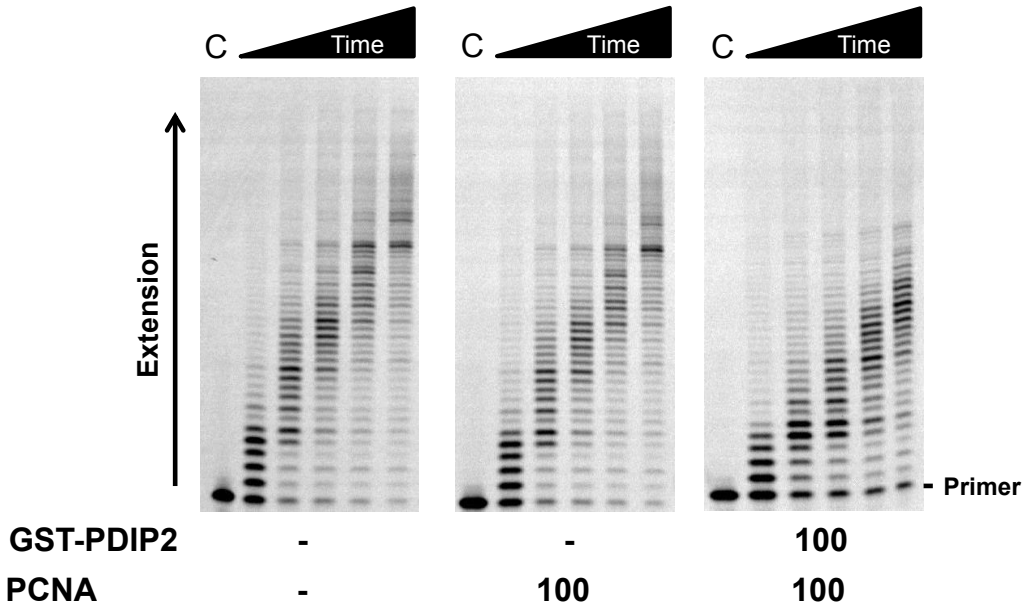


Figure 3.8. PrimPol is inhibited in the presence of PolDIP2 and PCNA in combination.
PrimPol's (100 nM) polymerase activity was examined using a 5' labelled 20/97-mer primer/template substrate in the presence of PCNA alone, or a combination of PCNA and PolDIP2 over a time course (1, 3, 5, 10, 20 mins). 'C' indicates the no enzyme control.

stimulating 8-oxo-dG bypass by Pols λ and η (Maga et al., 2013). Therefore, we set out to analyse the influence of PolDIP2 on the activity of PrimPol during 8-oxo-dG bypass.

PrimPol's 8-oxo-dG bypass efficiency, in the absence and presence of PolDIP2, was initially investigated using a primer/template (20/50-mer) containing a single 8-oxo-dG lesion located 8 nucleotides downstream from the primer/template junction (sequences 2 and 13, Table 2.2.). We observed that in the presence of increasing PolDIP2 concentrations, PrimPol and PrimPol$^{1-354}$ synthesised a greater number of post-8-oxo-dG extension products in a dose-dependent manner, compared to reactions containing PrimPol only (Figure 3.9.A., B., and C.). However, it is important to note that, unlike Pol λ (Maga et al., 2013), this stimulation was not significantly greater than that observed on non-damaged DNA templates (Figure 3.9.B., and C.). Nevertheless, the enhancement of PrimPol's polymerase activity by PolDIP2 does increase the efficiency of 8-oxo-dG bypass compared to the enzyme alone. Together, these results suggest that PolDIP2 stimulates PrimPol-mediated 8-oxo-dG bypass by enhancing the polymerase activity of the enzyme rather than the ability of PrimPol to traverse the lesion.

Despite possessing the ability to bypass 8-oxo-dG lesions, PrimPol's inherent bypass fidelity is relatively poor, displaying around 1:1 error-prone (dATP) to error-free (dCTP) incorporations opposite the 8-oxo-dG lesion (Keen et al., 2014b). In addition to stimulating 8-oxo-dG bypass, PolDIP2 has also been shown to increase bypass fidelity by Pols λ and η, but only in the added presence of PCNA and RPA, with the protein alone not affecting lesion bypass fidelity (Maga et al., 2013). Therefore, we next analysed the effect of PCNA, RPA, and PolDIP2 on PrimPol's fidelity when bypassing 8-oxo-dG lesions. To examine this, we employed primer extension assays with a primer/template (28/50-mer), where the immediate templating base was an 8-oxo-dG lesion (position 29 on the template) (sequences 4 and 13, Table 2.1.). Reactions were supplemented with either dATP or dCTP to allow analysis and quantification of error-prone and error-free bypass. As demonstrated previously (Keen et al., 2014b), in the absence of auxiliary proteins PrimPol incorporates both dATP and dCTP opposite the 8-oxoG lesion at ~ 1:1 ratio (Figure 3.9.D.). This ratio was largely unchanged in the presence of RPA, which has previously been shown to increase the fidelity of 8-oxo-dG bypass by Pol λ and Pol η (Maga et al., 2007). It should be noted that RPA was used at low concentrations to prevent inhibition of PrimPol (Guilliam et al., 2015a). PCNA also did not significantly affect the fidelity of lesion bypass by PrimPol. However, in the presence of PolDIP2 alone, PrimPol's fidelity opposite 8-oxo-dG was significantly improved, with the enzyme demonstrating an almost 2-fold increase in dCTP incorporation, whilst dATP incorporation remained largely the same. However, in the presence of both PolDIP2 and

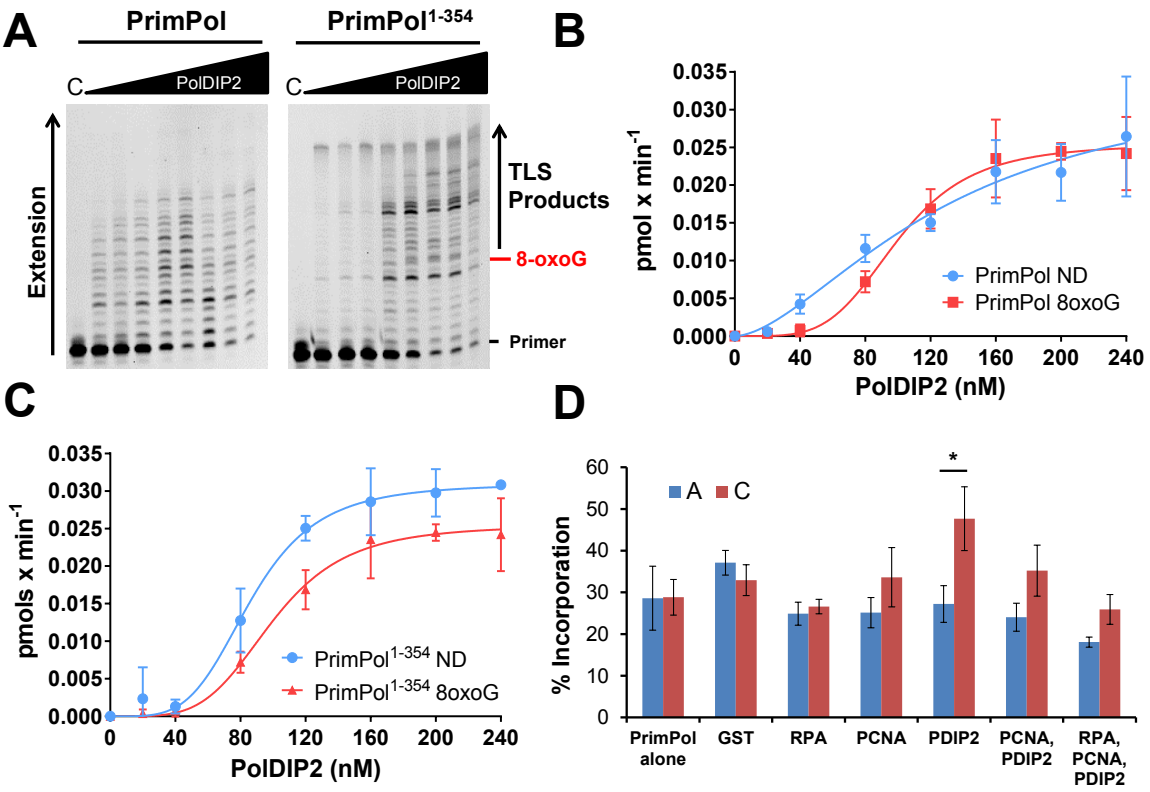


Figure 3.9. PolDIP2 enhance the efficiency and fidelity of 8-oxo-dG bypass by PrimPol. **(A)** PrimPol or PrimPol$^{1-354}$ (40 nM) were incubated with dNTPs (100 μM) and 5'-labelled 20/50-mer primer/template substrates containing a single 8-oxoG lesion 8 nt downstream of the primer-template junction in the presence of increasing GST-PolDIP2 concentrations (0, 20, 40, 80, 120, 160, 200, 240 nM) for a single 10 min timepoint. 'C' indicates the no enzyme control reaction. **(B)** Relative increase in the rate of TLS product synthesis (as indicated on figure) by PrimPol on non-damaged (ND) and 8-oxo-dG containing templates in the presence of increasing GST-PolDIP2 concentrations (0, 20, 40, 80, 120, 160, 200, 240 nM). Data were fitted using GraphPad Prism 6 software. Values are the means of four independent experiments. Error bars are ± SD. **(C)** Relative increase in the rate of TLS product synthesis (as indicated on figure) by PrimPol$^{1-354}$ on non-damaged (ND) and 8-oxo-dG containing templates in the presence of increasing GST-PolDIP2 concentrations (0, 20, 40, 80, 120, 160, 200, 240 nM). Data were fitted using GraphPad Prism 6 software. Values are the means of four independent experiments. Error bars are ± SD. **(D)** PrimPol (100 nM) was incubated with either dATP or dCTP (100 μM) and 5'labelled 28/50-mer primer/template substrates with a single 8-oxoG lesion as the immediate templating base (position 29 on the template) in the absence and presence of GST-PolDIP2 (300 nM), PCNA (100 nM), and RPA (25 nM), or a combination of each. Reaction products were quantified to give the relative amounts of correct (dCTP, red) and incorrect (dATP, blue) incorporation. Data represent the mean of 3 independent experiments. Error bars are ± SD. Data were subject to an unpaired t-test, PolDIP2 alone data $p<0.05$.

PCNA, or PolDIP2, PCNA, and RPA, in combination, this increase in fidelity was reduced and the overall amount of incorporation (both dATP and dCTP) decreased. These data therefore demonstrate that, unlike Pols λ and η (Maga et al., 2013), PrimPol's fidelity opposite 8-oxo-dG is increased in the presence of PolDIP2 alone. Furthermore, addition of RPA and PCNA actually act to lessen the effect of PolDIP2 on PrimPol's lesion bypass fidelity.

The catalytic domain of PrimPol alone has the ability to bypass CPDs, however the full-length enzyme is stalled by these lesions. In the presence of magnesium, PrimPol is also stalled by Ap sites (Bianchi et al., 2013; Keen et al., 2014b). We also tested whether PolDIP2 permits bypass of these lesions by PrimPol. In each case, the presence of PolDIP2 did not allow PrimPol to synthesise across the damaged nucleotide (Figure 3.10.). In addition, we analysed the lesion bypass fidelity of PrimPol in the presence of PolDIP2 when traversing a uracil and 6-4PP lesion. Again, PrimPol's fidelity was in line with the previously published results, incorporating dATP opposite uracil, and dTTP opposite the 6-4PP (Keen et al., 2014b).

PrimPol is an error-prone polymerase, which has previously been shown to misincorporate bases and extend base mismatches (Guilliam et al., 2015a). In particular, the enzyme shows a propensity to misincorporate and extend mismatched bases opposite a templating dC (Guilliam et al., 2015a). Since PolDIP2 increases PrimPol's fidelity when synthesising past an 8-oxo-dG lesion, we also tested whether the protein affects PrimPol's fidelity on non-damaged DNA. To measure this, we analysed PrimPol's level of misincorporation opposite a templating dC in the absence and presence of PolDIP2 (Figure 3.11.). However, we observed that PolDIP2 does not reduce PrimPol's level of misincorporation, suggesting that the protein does not improve the enzyme's fidelity on non-damaged DNA. Together, these results suggest that PolDIP2 acts to increase PrimPol's efficiency and fidelity when specifically bypassing an 8-oxo-dG lesion, rather than improving the enzymes overall fidelity rates.

3.4.7. Analysis of the Interaction of PolDIP2 with PrimPol

PolDIP2 was originally identified as a potential PrimPol interacting protein in a large-scale pull-down screen performed previously (Rudd, 2013). In order to analyse this interaction in more detail, we employed BS_3 cross-linking and MS analysis. This type of analysis allows non-covalent interactions between proteins to be converted into covalent bonds, specifically BS^3 is able to cross-link primary amines on the side chains of lysine residues, in addition to the N-terminus of proteins. Further protease digestion and MS analysis of cross-linked protein complexes allows the covalently attached regions of each

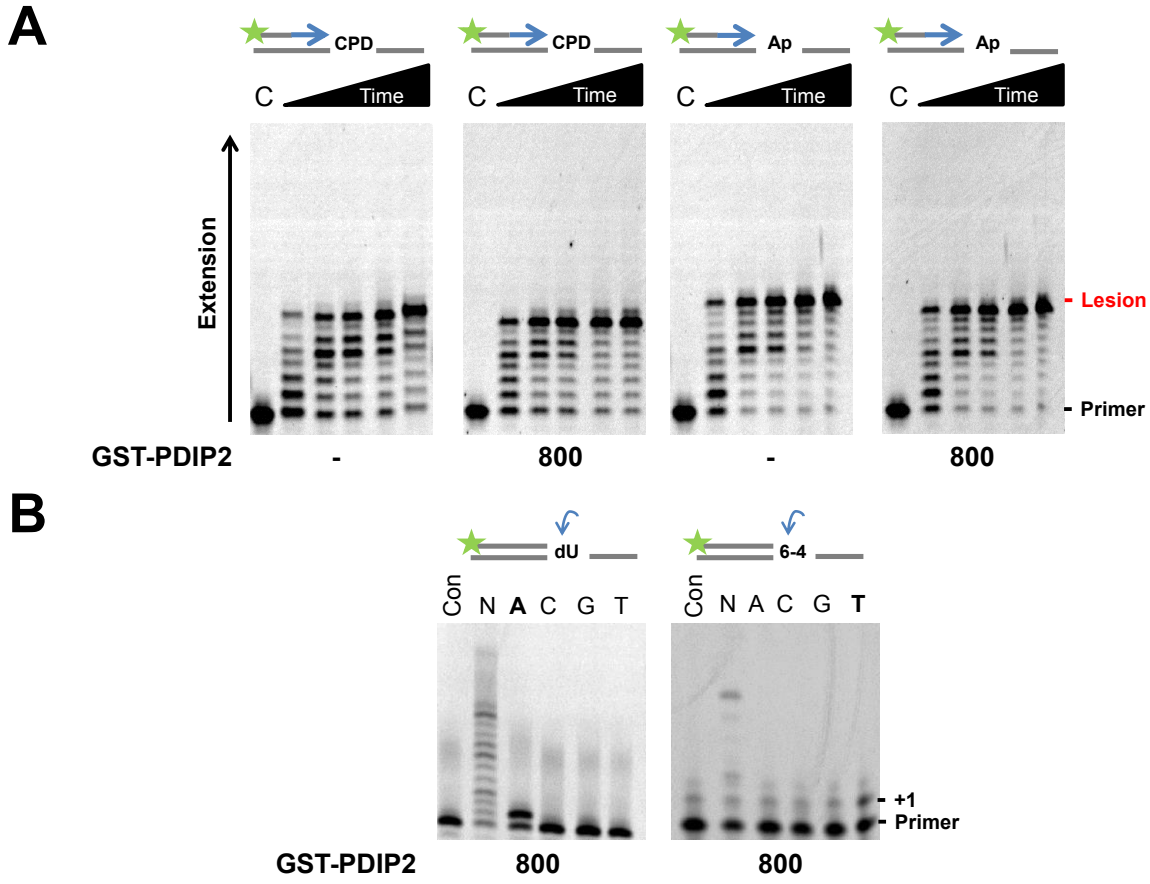


Figure 3.10. PolDIP2 does not enable PrimPol to bypass CPDs or Ap sites, or alter the fidelity of dU or 6-4PP bypass.
(A) PrimPol was incubated with dNTPs (100 μM) and 5'-labelled 20/50-mer primer/template substrates containing a single CPD or Ap site 8 nt downstream of the primer-template junction in the absence or presence of PolDIP2 (800 nM). Reactions were monitored over a time-course of 1, 3, 5, 10, and 20 mins. 'C' indicates the no enzyme control reaction. **(B)** PrimPol was incubated with single dNTPs (dATP, dCTP, dGTP, or dTTP) (100 μM) and 5'-labelled primer/template substrates with a single dU or 6-4PP as the immediate templating base in the presence of GST-PolDIP2. 'C' indicates the no dNTP control.

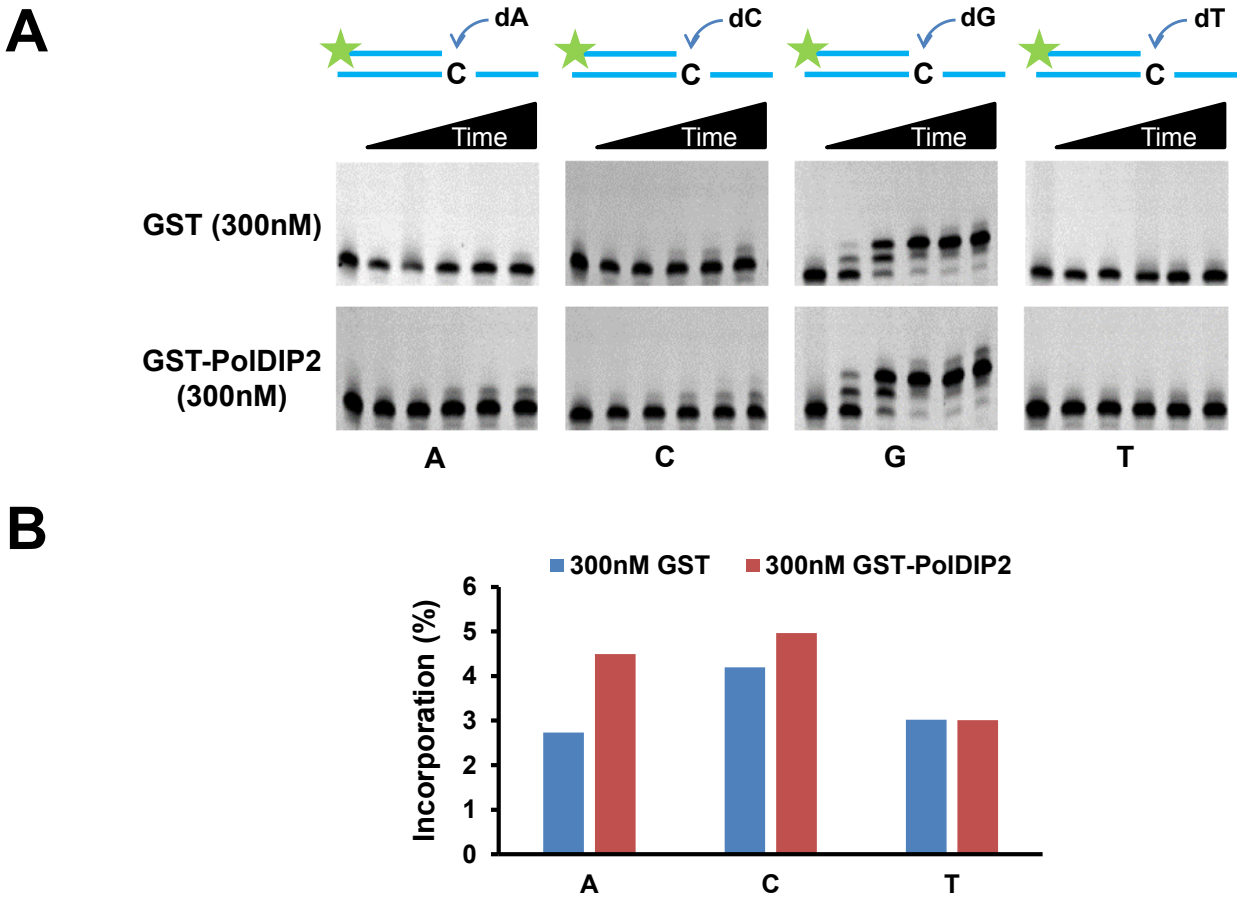


Figure 3.11. PolDIP2 does not affect PrimPol's fidelity on non-damaged DNA.
(A) PrimPol was incubated with single dNTPs (dATP, dCTP, dGTP, or dTTP) (100 μM) and 5'-labelled 27/50-mer primer/template substrates containing dC as the immediate templating base. Reactions were monitored over a time-course of 1, 3, 5, 10, and 20 mins. 'C' indicates the no dNTP control reaction. **(B)** The data shown in A were normalised against the correct incoming base (dGTP) and quantified to give the relative misincorporation of each base opposite the templating dC.

protein to be recognised and thereby interacting regions to be identified (Rappsilber, 2011). StavroX cross-linking analysis software was used to identify and score cross-linked peptides, as detailed previously (Götze et al., 2012). Since EMSA data suggested an interaction between PolDIP2 and the catalytic domain of PrimPol (Figure 3.4.), untagged PolDIP2 and PrimPol$^{24-354}$ were used for this analysis.

Intriguingly, the vast majority of PrimPol-PolDIP2 cross-links identified were mediated by the N-terminus of PolDIP2 (residues 1-8), with additional secondary crosslinks also identified (Table 3.1., Figure 3.12.A. and B.). The most abundant and highest scoring cross-linked peptide identified on PrimPol, which cross-linked to the N-terminus of PolDIP2, was located between amino acid positions 60 and 70 on the full-length protein (EDVHVFALECK), with the cross-linked residue identified as lysine 70 (Table 3.1). Notably, this peptide displays strong homology to a region of Pol η previously found to mediate the enzyme's interaction with PolDIP2 (Figure 3.12.C.) (Tissier et al., 2010), potentially suggesting that PrimPol and Pol η share a similar mode of binding to PolDIP2. A number of other cross-linked peptides were also identified on PrimPol. The majority of these were located towards the C-terminus of the truncated protein. However, analysis of intra-PrimPol cross-links, suggests that these regions are in close proximity to the EDVHVFALECK peptide in the folded protein (Table 3.2., Figure 3.12.A. and B.).

Given that the N-terminal 50 amino acids of PolDIP2 comprise a mitochondrial targeting sequence (MTS), which is likely cleaved off upon entry to the mitochondria (Xie et al., 2005), it was somewhat surprising to identify this region as the mediator of the PrimPol interaction. In order to validate the findings of the crosslinking and MS analysis, we generated a truncated form of PolDIP2 lacking the first 50 amino acids (PolDIP2$^{51-368}$) and assayed its ability to stimulate PrimPol's processivity in comparison to the full-length protein. In addition, we also analysed the effect of PolDIP2$^{51-368}$ on the processivity of Pol η. In line with previous results, we find that full-length PolDIP2 is able to stimulate the processivity of both PrimPol and Pol η (Figure 3.12.D.) (Maga et al., 2013). However, in contrast, PolDIP2$^{51-368}$ failed to stimulate the processivity of either enzyme. This assay was repeated across a range of PolDIP2$^{51-368}$ concentrations with no increase in PrimPol's processivity identified (Figure 3.13.). Furthermore, PolDIP2$^{51-368}$ failed to produce an increase in the DNA binding of PrimPol, which was previously observed with the full-length protein (Figure 3.13.). These results further support the findings of the MS analysis, suggesting that the interaction between PolDIP2 and PrimPol is mediated by the N-terminus of PolDIP2. Furthermore, these findings suggest that Pol η may also interact with the N-terminus of PolDIP2.

Score	m/z	z	Measured Mass	Calculated Mass	PolDIP2 Peptide	Peptide Location	PrimPol Peptide	Peptide Location
131	807	3	2420	2420	**M**AABTARR	0-8	EDVHVFALEB**K**	58-68
105	411	4	1643	1643	MAAB**T**AR	0-7	LYKSS**K**	290-295
101	755	3	2263	2263	**M**AABTARR	0-8	IL**T**BEPSQNK	341-350
90	567	4	2263	2263	MAAB**T**ARR	0-8	ILTBEPSQN**K**	341-350
87	768	3	2301	2301	AENPAGHGSK EVKG**K**	115-129	MF**T**EK	231-235
86	411	4	1643	1643	MAAB**T**AR	0-7	LY**K**SSK	290-295
84	532	4	2125	2125	ET**L**RAWQEK	215-223	QGFSFN**K**	224-230
81	709	3	2125	2125	ET**L**RAWQEK	215-223	QGFSFN**K**	223-230
63	1132	2	2263	2263	**M**AABTARR	0-8	IL**T**BEPSQNK	341-350
43	548	3	1643	1643	**M**AABTAR	0-7	L**Y**KSSK	290-295
40	554	3	1659	1659	**M**AABTAR	0-7	LYKSS**K**	290-295
32	567	4	2263	2263	MAAB**T**ARR	0-8	ILTBEPSQN**K**	341-350
29	411	4	1643	1643	MAAB**T**AR	0-7	LY**K**SSK	290-295
24	617	5	3083	3083	MAAB**T**AR	0-7	FSDTLRILTBEP SQNKQ**K**	335-353
23	771	4	3083	3083	**M**AABTAR	0-7	FSDTLRIL**T**BEP SQNKQK	335-353
22	411	4	1643	1643	**M**AABTAR	0-7	LYKSS**K**	290-295

Table 3.1. MS analysis of PolDIP2-PrimPol cross-linked peptides.
The amino acid sequence of cross-linked PolDIP2 (blue) and PrimPol (green) peptides identified in the MS analysis are shown. Crosslinked residues in each case are shown in bold. Peptide location indicates the location of peptide in the specific constructs used for the analysis, not the wild-type protein. Scores are based on the probability of the occurrence of an ion in addition to the intensity of the signal in the fragment spectrum, as given by the StavroX analysis software (Götze et al., 2012). 'm/z' indicates the mass to charge ratio, 'z' shows the charge of the precursor.

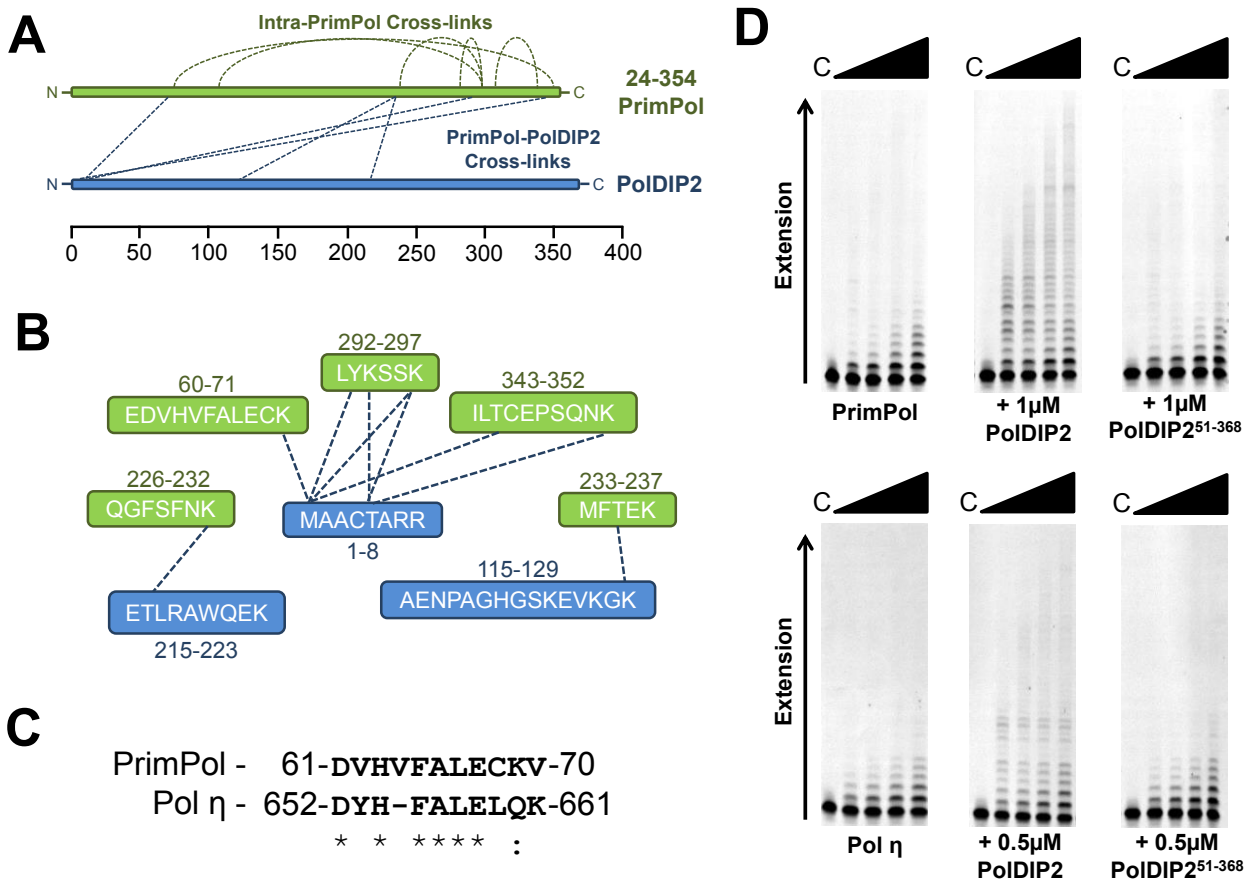


Figure 3.12. Analysis of the PrimPol-PolDIP2 interaction.
(A) PrimPol$^{24-354}$ (green) and untagged-PolDIP2 (blue) cross-links were analysed by digest and MS revealing potential interacting regions on each protein (Table 3.1.). The locations of intra-PrimPol and inter-PrimPol/PolDIP2 cross-links are indicated by dotted lines. The relative amino acid positions are shown below. **(B)** The amino acid sequences of PrimPol-PolDIP2 cross-linked peptides. Dotted lines between peptides indicate the specific residues cross-linked in each case. Cross-linked PrimPol peptides are shown in green and PolDIP2 peptides in blue. **(C)** Alignment of the PolDIP2-interacting regions of PrimPol and Pol η showing the high degree of homology between the two peptides. **(D)** PrimPol (100 nM) and Pol η were incubated with 5'-labelled 20/97-mer primer/template substrates and dNTPs (100 μM) in the absence and presence of GST-PolDIP2 or PolDIP2$^{51-368}$. GST-PolDIP2 stimulated the processivity of both PrimPol and Pol η, however PolDIP2$^{51-368}$ failed to stimulate the processivity of either enzyme.

Score	m/z	z	Measured Mass	Calculated Mass	Peptide 1	Peptide location	Peptide 2	Peptide location
172	665	3	1993	1993	F**S**DTLR	335-340	VALEVTEDN**K**	300-309
137	712	3	2134	2134	Q**K**	351-353	SB**K**EDVHVFAL EBK	55-68
30	496	5	2478	2478	**S**SK	293-295	MFTE**K**ATEESW TSNSKK	231-247
27	542	6	3245	3245	NFRLYK**S**SK	287-295	VALEVTEDNKF FPIQS**K**	300-316
25	1311	3	3932	3932	PANPGADGK**K**	118-127	DVSDEYQYFLS SLVSNVRFSD**T** LR	317-340
20	579	5	2890	2890	IG**K**R	296-299	NNMGE**K**HLFVD LGVYTRNR	268-286
15	1475	3	4423	4423	IYLVTTYAEFW FYY**K**	75-89	NNMGEKHLFVD LGVY**T**RNR	268-286
11	712	3	2134	2134	Q**K**	351-353	SB**K**EDVHVFAL EBK	55-68

Table 3.2. MS analysis of intra-PrimPol cross-links.
The amino acid sequence of cross-linked PrimPol (green) peptides identified in the MS analysis are shown. Cross-linked residues in each case are shown in bold. Peptide location indicates the location of peptide in the specific constructs used for the analysis, not the wild-type protein. Scores are based on the probability of the occurrence of an ion in addition to the intensity of the signal in the fragment spectrum, as given by the StavroX analysis software (Götze et al., 2012). 'm/z' indicates the mass to charge ratio, 'z' shows the charge of the precursor.

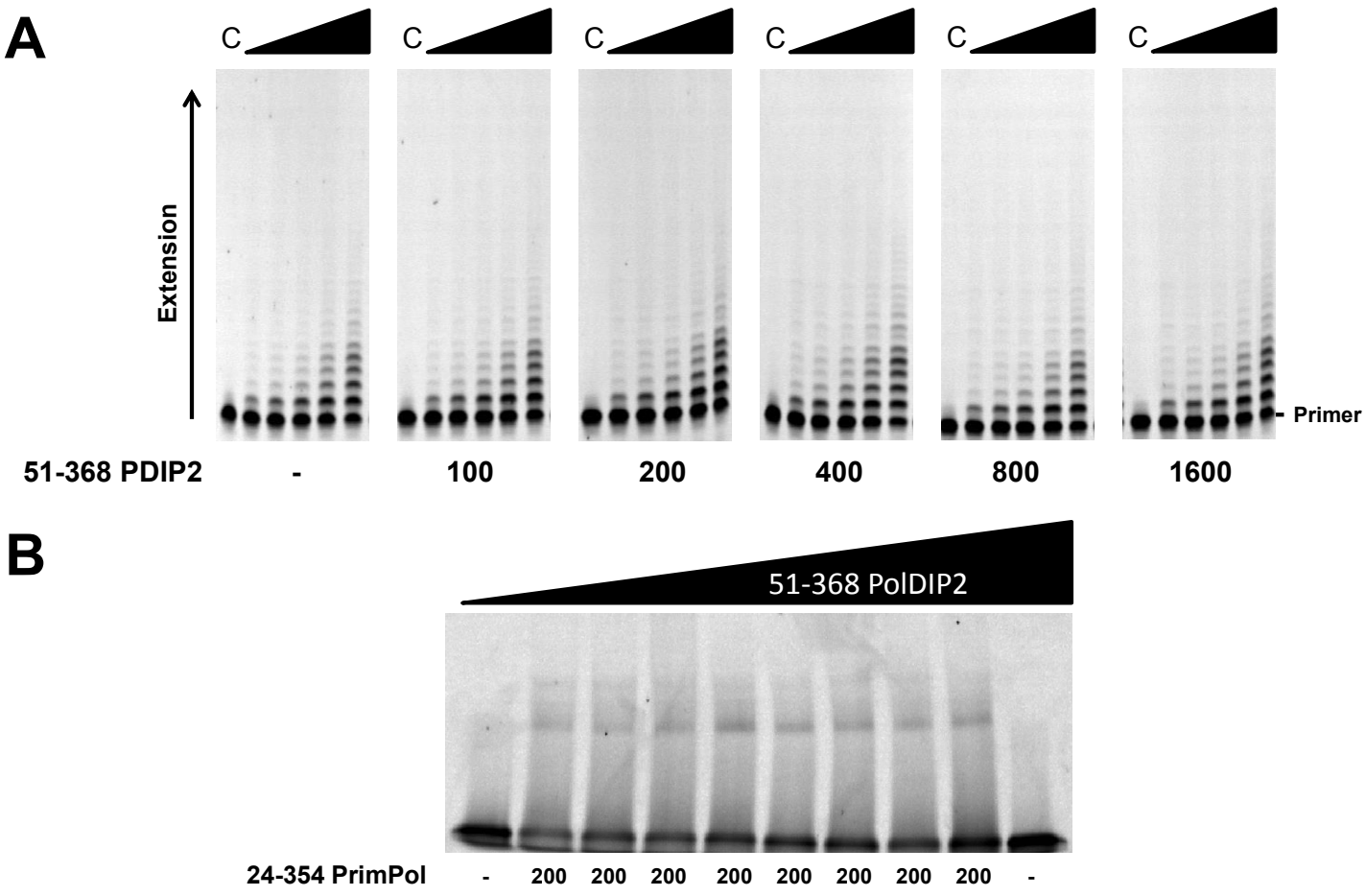


Figure 3.13. PolDIP2$^{51-368}$ does not stimulate PrimPol's processivity or DNA binding.
(A) PolDIP2$^{51-368}$ was titrated into reactions containing PrimPol (100 nM) and 5' labelled 20/97-mer primer/template substrates (20 nM). Reactions were initiated with dNTPs (100 μM) and excess trap DNA and quenched at 0.5, 1, 2, 5, and 10 min time-points. **(B)** PrimPol$^{24-354}$ was incubated with increasing concentrations of PolDIP2$^{51-368}$ (0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75, 1 μM) in EMSA reactions containing 5'-labelled 20/97-mer primer-template substrates.

3.4.8. Depletion of PolDIP2 Causes Slowed Replication Fork Rates after UV Damage

Despite the inability of PolDIP2$^{51-368}$ to stimulate the processivity of Pol η, existing data suggests that these proteins do share a functional interaction *in vivo*. Specifically, it has been shown that depletion of PolDIP2 causes persistent Pol η foci in the absence of damage. Furthermore, PolDIP2 depleted cells showed increased UV sensitivity to a similar level as cells lacking Pol η (XPV cells) however, no further increase in sensitivity was observed when PolDIP2 was depleted in XPV cells (Tissier et al., 2010). Additionally, cells depleted of PolDIP2 showed an increased sensitivity to oxidative damage (Maga et al., 2013). These studies implicate PolDIP2 in the regulation of TLS *in vivo*, although the direct impact of depletion of PolDIP2 on DNA replication has not previously been examined.

To analyse the impact of depletion of PolDIP2 on replication fork rates following DNA damage, both wild-type and PrimPol$^{-/-}$ MRC5 cells were either PolDIP2 siRNA or mock treated before DNA fibre analysis was conducted (Figure 3.14.A.). Cells were pulse labelled with chlorodeoxyuridine (CldU) for 20 mins before UV irradiation (20 J/m^2), following this, cells were pulse labelled again with iododeoxyuridine (IdU) for an additional 20 mins and the ratios of the two labels determined. Significantly, depletion of PolDIP2 in wild-type MRC5 cells causes a significant decrease in replication fork rates following UV-C irradiation (Figure 3.14.B., C., and D.). This suggests that PolDIP2 is involved in DNA replication and, more specifically, in DDT, supporting published studies implicating it in TLS processes. Although to a lesser extent than observed in the previously studied PrimPol$^{-/-}$ DT40 cells (Bianchi et al., 2013), MRC5 cells lacking PrimPol also display a decrease in replication fork rates following UV-C irradiation. However, intriguingly, depletion of PolDIP2 in PrimPol$^{-/-}$ cells did not produce a further decrease in replication fork rates, suggesting that PrimPol and PolDIP2 work epistatically in the same pathway to promote continued DNA replication in the presence of UV damage (Figure 3.14.B., C., and D.). This also suggests that PolDIP2 may also operate in a post-replicative manner during gap-filling by other TLS polymerases, potentially explaining why a further decrease in replication fork rates was not observed, despite PolDIP2 likely partnering other TLS enzymes.

3.5. Discussion

Previous studies have implicated PolDIP2 in TLS damage tolerance processes through the regulation of Pols λ and η (Maga et al., 2013; Tissier et al., 2010). In addition, this

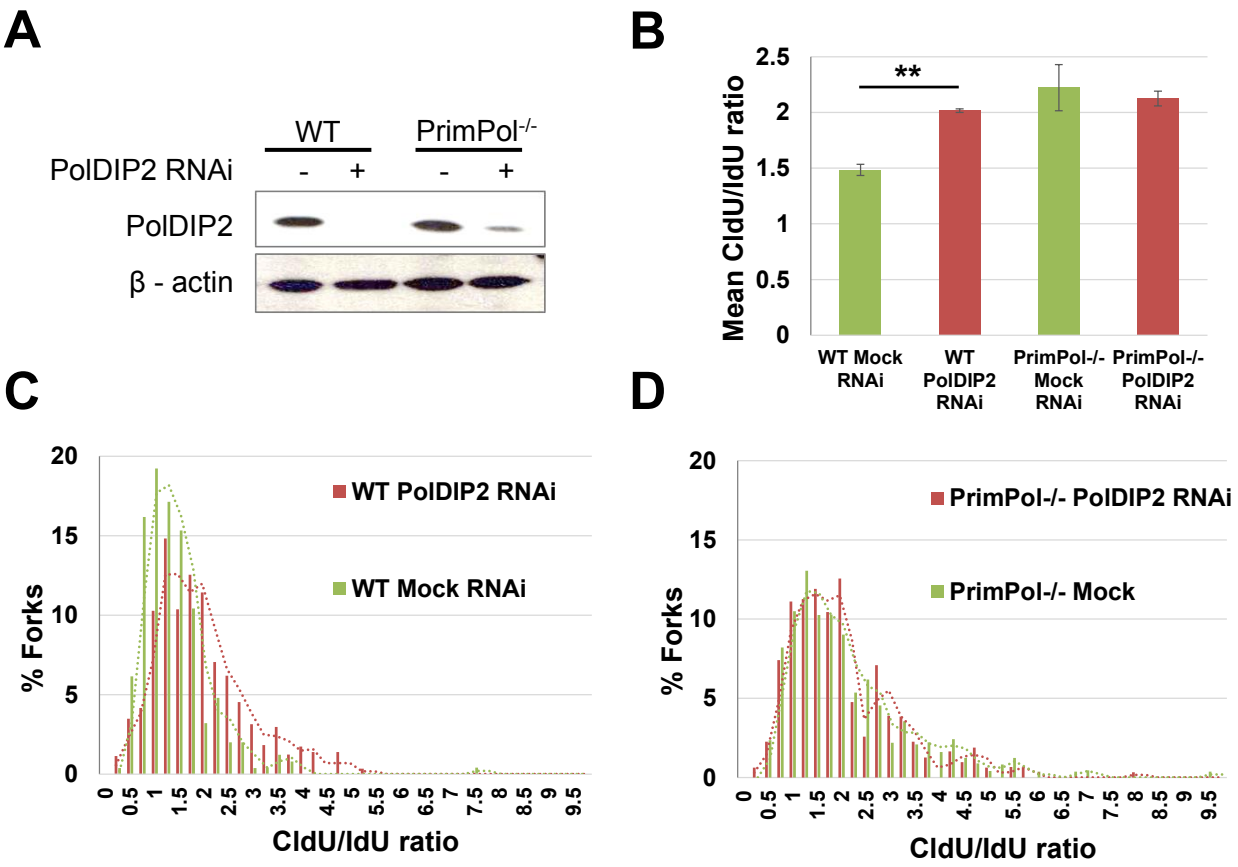


Figure 3.14. Depletion of PolDIP2 causes decreased replication fork speeds following UV damage.
(A) Western blot analysis of PolDIP2 silencing by siRNA in wild-type (WT) and PrimPol$^{-/-}$ MRC5 cells, compared to mock depleted cells. **(B)** DNA replication fork rates in PolDIP2 siRNA or mock treated wild-type and PrimPol$^{-/-}$ MRC5 cells were analysed by DNA fibre analysis. Cells were pulsed with CldU for 20 mins followed by UV irradiation (20 J/m^2) and pulse labelled again with IdU for a further 20 mins. The mean CldU/IdU ratio of wild-type and PrimPol$^{-/-}$ MRC5 cells either mock or PolDIP2 siRNA treated is shown. Error bars indicate ± SE (n = 3). **(C)** DNA replication fork rate data for wild-type MRC5 cells shown as the ratio of CldU to IdU (n = 3). Green lines indicate analysis of mock treated cells, whilst red lines indicated analysis of PolDIP2 siRNA treated cells. **(D)** DNA replication fork rate data for PrimPol$^{-/-}$ MRC5 cells shown as the ratio of CldU to IdU (n = 3). Green lines indicate analysis of mock treated cells, whilst red lines indicated analysis of PolDIP2 siRNA treated cells.

protein has also been shown to interact with PCNA, Pol δ, Pol ζ, and Rev1 (Tissier et al., 2010). In this current study, we show that PolDIP2 is also involved in the regulation of PrimPol's polymerase activity. Specifically, PolDIP2 is able to increase the polymerase activity of PrimPol by increasing the enzyme's DNA binding capacity and processivity. In addition, PolDIP2 also acts as a fidelity factor for PrimPol during the bypass of 8-oxo-dG, enhancing dCTP incorporation opposite this oxidative lesion. In contrast, PolDIP2 has a minimal effect on PrimPol's primase activity. This suggests that the protein acts specifically to promote PrimPol's polymerase activity. Previously, it has been shown that PrimPol's ZnF domain is required for the primase activity of the enzyme, and additionally, is involved in negatively regulating its processivity (Keen et al., 2014b). This raises the possibility that binding of PolDIP2 may alleviate this negative regulation by the ZnF domain, in turn promoting increased processivity of the enzyme. Together, the *in vitro* data presented here suggests that PolDIP2 increases the processivity and polymerisation rates of PrimPol by stabilising the enzyme on DNA and improving PrimPol's inherently poor DNA binding capacity (Keen et al., 2014b).

Previously published results, and data presented here, suggest that PrimPol is not regulated through a canonical PCNA-mediated polymerase switch mechanism (Guilliam et al., 2015a). Furthermore, the presence of PCNA inhibited the positive impact of PolDIP2 on PrimPol's primer extension activity. It has previously been suggested that PolDIP2 might act as a bridging factor to enhance polymerase-PCNA interactions and thereby further stimulate polymerase activity (Maga et al., 2013). However, our data imply that PolDIP2 acts alone, in the absence of PCNA, to enhance PrimPol's activity. In further support of this, it was previously found that PolDIP2 only enhances Pol η and λ bypass fidelity opposite 8-oxo-dG in the presence of RPA and PCNA (Maga et al., 2013). However, in the case of PrimPol, we observed that PolDIP2 alone is sufficient to increase its 8-oxo-dG bypass fidelity and the further presence of RPA and PCNA actually reduces this effect.

Additionally, we found that PrimPol possesses a potential PolDIP2 binding motif with homology to that of Pol η (Tissier et al., 2010). Interestingly, we identified that this motif appears to bind to the very N-terminus of PolDIP2. However, the first ~50 amino acids of PolDIP2 are thought to comprise a MTS and are likely cleaved off upon entry to the mitochondria (Xie et al., 2005). PolDIP2 lacking this MTS was not able to stimulate the processivity of either PrimPol or Pol η. Importantly, previous studies reporting stimulation of Pols η, λ, and δ by PolDIP2 only employed full-length PolDIP2 with the N-terminal 50 amino acids intact (Maga et al., 2013). Additionally, it was originally reported that PolDIP2 inhibits Pol δ activity at higher concentrations, however this study was

performed using truncated PolDIP2 without the MTS (Xie et al., 2005). Therefore, it seems that these contradictory reports can be explained by the different PolDIP2 constructs used in each case. Furthermore, these reports support data presented here that the first 50 amino acids of PolDIP2 are required for stimulation of, and likely the interaction with, polymerases including not only PrimPol but also Pol η and Pol δ.

Despite this, *in vivo* data does support a role for PolDIP2 in DNA replication, and more specifically DDT. Here we have shown that depletion of PolDIP2 from cells causes a decrease in replication fork rates following UV irradiation to a similar level as that seen with PrimPol$^{-/-}$ cells. Furthermore, depletion of PolDIP2 in PrimPol$^{-/-}$ cells does not produce a further decrease in fork speed. Therefore, it appears that PolDIP2 and PrimPol cooperate to promote continued replication in the presence of DNA damage. Importantly, this does not rule out the possibility that PolDIP2 also assists other TLS polymerases in a post-replicative gap filling manner. Indeed, this is supported by previous studies suggesting that PolDIP2 acts to promote interactions between canonical TLS polymerases and PCNA. In support of this, previous reports suggest that PCNA ubiquitylation is not required to maintain normal fork progression on damaged DNA but is instead essential for filling-in post-replicative gaps (Edmunds et al., 2008).

Importantly, the initial characterisation of PolDIP2 suggested that the protein was primarily mitochondrial (Xie et al., 2005). However, it was also acknowledged that PolDIP2 may also be present in the nucleus in small amounts and that interactions between PolDIP2 and Pol δ may only occur during specific cellular events, such as following DNA damage. This study also suggested that additional isoforms of PolDIP2 may exist, which localise to the nucleus and possibly serve different functions to those in the mitochondria (Xie et al., 2005). Since this initial characterisation, additional reports indicate that PolDIP2 does indeed localise to the nucleus, with an increase following UV damage (Wong et al., 2013). Therefore, it is possible that nuclear and mitochondrial PolDIP2 fulfil different functions *in vivo*. Whilst PolDIP2 localised to the mitochondria will likely have its MTS removed, it is possible that a small amount of PolDIP2, which localises to the nucleus does not. This suggests that the stimulatory effects of PolDIP2 on PrimPol may be primarily nuclear and in response to DNA damage, rather than mitochondrial. PrimPol is a highly error-prone enzyme and must be strictly regulated to restrict its contribution to DNA synthesis (Guilliam et al., 2015a). Thus, it seems likely that the PrimPol-PolDIP2 interaction may also be mediated by post-translational modifications in response to DNA damage, this would prevent PrimPol's DNA binding and processivity from being constantly enhanced and therefore limit its involvement in DNA synthesis. However, further studies are required to assess potential PrimPol and

PolDIP2 post-translational modifications and their influence on the interactions between these proteins.

Overall, the findings presented here establish that PolDIP2 is able to enhance the polymerase activities of PrimPol *in vitro*. In support of these findings, we also demonstrate that cells depleted of PolDIP2 show replication defects similar to PrimPol$^{-/-}$ cells after UV irradiation. These effects are not further increased when PolDIP2 is depleted in PrimPol$^{-/-}$ cells, clearly suggesting that PolDIP2 is a binding partner of PrimPol *in vivo* and likely mediates its primer extension activities in response to DNA damage. These findings further support the accumulating published evidence implicating PolDIP2 as a general DDT factor involved in promoting TLS by a number of different polymerases. Together, our work describes a new regulatory partner of PrimPol and enhances our understanding of the emerging role of PolDIP2 in DDT.

Chapter 4

The Development of a Fluorescent Gel-Based Primase Assay to Analyse PrimPol's Repriming Activity

4.1. Overview

Both Chapter 2 and Chapter 3 focussed on the characterisation of PrimPol's functional interaction with its binding partners, RPA, mtSSB, and PolDIP2. The articles presented in these chapters identified that PrimPol's activities are restricted by SSBs, potentially to reduce its contribution to DNA synthesis and thus the chance for mutagenesis, whilst PolDIP2 promotes PrimPol's primer extension capabilities, stabilising it on the DNA template and increasing its processivity. Chapter 4 switches focus to the primase activity of PrimPol. This chapter is a combination of work from three published articles. The first half of the chapter (section 4.2) describes the development of a fluorescence-based primase assay, following a detailed discussion of previous methods for characterising primases, covering their advantages and limitations. This part of the chapter is accepted for publication in *Methods in Enzymology*. The primary motivation for developing this assay was to allow a more accurate characterisation of PrimPol's repriming activities, without the use of radioactivity. Notably, the previous method used in the Doherty laboratory was time-consuming and relied on indirect observation of primase reaction products by labelling and extension using Klenow-Taq polymerase (Bianchi et al., 2013; Keen et al., 2014b). During the development of this assay it was observed that PrimPol can extend the 3' end of ssDNA templates. An investigation of the mechanism responsible for this extension is therefore included after the *Methods in Enzymology* article. Note that this data and discussion were not published in the article itself (section 4.3). The second half of the chapter (section 4.4) describes the use of this assay in analysing PrimPol's involvement in the bypass non-canonical replication impediments, including G4-quadruplexes and chain-terminating nucleotide analogues (CTNAs). The data produced from these studies are published in *Molecular Cell* (Schiavone et al., 2016) and *Cell Cycle* (Kobayashi et al., 2016), respectively.

4.2. Current and Emerging Assays for Studying the Primer Synthesis Activities of DNA Primases

4.2.1. Abstract

Primases play a crucial role in the initiation of DNA synthesis during replication by *de novo* synthesis of short RNA or DNA 'primers'. In recent years, evidence has accumulated which expands the essential roles of primases to include, not only the initiation of replication, but also other critical roles in DNA metabolism, including damage tolerance and repair. Despite the broadening roles for these enzymes, the methods used to identify and characterise primase activities are limited. Historically, biochemical analysis of primases has been based on the synthesis of radioactively-labelled primers and their detection on denaturing polyacrylamide gels. In the last two decades, a number of alternative primase assays have been developed in an effort to supersede radioactive methods. However, the radioactive gel-based assay, which has not significantly changed since its conception in the late 1970s, remains the most widely used and favoured method. In this chapter, we discuss the background to, and the advantages and disadvantages of, the current techniques used to characterise primase activity *in vitro.* Finally, we describe an alternative, gel-based, fluorescent primase assay, which we have successfully used in the characterisation of a recently identified primase-polymerase, PrimPol.

4.2.2. Introduction

Primases, unlike DNA polymerases, possess the unique ability to utilise ssDNA templates for the initiation of RNA or DNA synthesis *de novo*. The short RNA or DNA chains produced from this synthesis are termed primers and provide the 3' hydroxyl required for further extension by DNA polymerases during the initiation of DNA replication. Due to the semi-discontinuous nature of DNA replication, primase activity is essential not only during initiation, but also throughout replication to prime Okazaki fragment synthesis on the lagging strand. All domains of life employ primases in this crucial role, however two distinct primase superfamilies, DnaG primases and AEPs, facilitate bacterial and archaeal/eukaryotic DNA replication, respectively. Recently, evidence has accumulated suggesting that primase-polymerases of the AEP superfamily also play key roles in DNA damage tolerance and repair, here their primase activity is essential for replication restart mechanisms including repriming of replication downstream of lesions and secondary structures (Guilliam et al., 2015b; Guilliam and Doherty, 2017).

Historically, DNA polymerases were studied *in vitro* using radioactive primer extension assays. Here, a 5' radiolabelled primer annealed to a DNA template is extended by the DNA polymerase of interest. Subsequently, resolution of the reaction products on a denaturing polyacrylamide gel, followed by autoradiography or phosphorimaging allows visualisation and analysis of primer extension products and thus polymerase activity. These assays have been reliably used for many years to study DNA polymerase kinetics, fidelity, processivity, repair, and translesion synthesis. However, over recent years, fluorescent tags have begun to replace radioactive labels offering clear advantages in safety and speed, whilst maintaining sensitivity.

Despite the use of fluorescence in primer extension assays, gel-based primase assays still routinely make use of radiolabelled nucleotides. In these assays, primase activity is determined by the quantification of radiolabelled nucleotide containing primers, visualised on denaturing polyacrylamide gels. In this chapter, we discuss the advantages and limitations of existing methods used to study primases *in vitro*, including both traditional radioactive gel-based assays and more recently developed non-radioactive high-throughput screening (HTS) approaches. Finally, we describe a gel-based primase assay of particular use in the analysis of primase-polymerases. This assay, which utilises fluorescently labelled nucleotides, removes the need for potentially hazardous radioactivity, allowing the assay to be performed in any laboratory without requiring training in the handling of radioactivity. Furthermore, this assay can be used in the same

way as traditional radioactive primase assays to study inherent primase activity, as well as to determine the effect of binding partners and reaction conditions, in addition to sequence preference and the location of priming, on the assayed primase. To demonstrate the effectiveness of this assay, we make use of purified human PrimPol, a recently characterised primase-polymerase involved in DNA damage tolerance in higher eukaryotes.

4.2.3. Radioactive-Based Primase Assays

4.2.3.1. Traditional Radioactive Primase Assays

In the early 1970s, *in vivo* studies of T7 DNA replication implicated the phage gene 4 protein in priming DNA synthesis (Strätling and Knippers, 1973; Wolfson and Dressler, 1972). Subsequent *in vitro* studies developed a DNA primase complementation assay to analyse the T7 gene 4 product, this assay helped to confirm that the protein acts to synthesise primers required for the initiation of DNA synthesis by T7 DNA polymerase (Hinkle and Richardson, 1975; Romano and Richardson, 1979; Scherzinger et al., 1977; Scherzinger and Litfin, 1974). Consequently, the T7 gene 4 product became the first designated DNA primase (Scherzinger et al., 1977).

The assay used in these early studies measured the ability of purified T7 primase to stimulate DNA synthesis in extracts prepared from *Escherichia coli* infected with T7 lacking the gene 4 protein (Hinkle and Richardson, 1975). Here, reactions were assembled containing rNTPs and dNTPs, one of which was ^{3}H or ^{32}P labelled, extracts from T7 infected *E. coli*, purified T7 primase and T7 linear duplex DNA. Following incubation, the reaction products were precipitated and washed on filter paper and radioactivity was measured by a liquid scintillation counter (Hinkle and Richardson, 1975). Thus, the ability of T7 primase to stimulate DNA synthesis in the extracts could be determined by the increase in acid insoluble radioactivity produced following incubation. In addition to analysing the effect of T7 primase on DNA synthesis in extracts, these reports also examined the effect of the enzyme on T7 polymerase activity using a similar approach. In these experiments, reactions were assembled in the same manner, however extract was omitted and replaced with purified T7 DNA polymerase. Again, it was identified that T7 primase markedly stimulated DNA synthesis by the polymerase on duplex T7 DNA (Hinkle and Richardson, 1975).

Despite indicating that the stimulation observed in these early studies was due to *de novo* primer synthesis by T7 primase, interpretation of results was somewhat limited due to the duplex linear DNA template used. In order to gain further clarity, pycnographic analysis of reaction products was required. In this case, template T7 DNA was ^{3}H, ^{13}C,

and ^{15}N labelled and α-^{32}P dATP was provided for primer synthesis. Subsequent CsCl density gradient centrifugation confirmed that the ^{32}P labelled reaction products separated from the heavy template DNA, thereby indicating that the reaction products and template DNA were not covalently linked and allowing inference that *de novo* initiation of DNA synthesis had occurred (Hinkle and Richardson, 1975). Follow-up studies of T7 primase were able to avoid this issue by using circular ss phage ΦX174 DNA as a template, thereby removing the possibility of synthesis being initiated by a loop mechanism (Scherzinger et al., 1977). Interestingly, it was later found that the T7 gene 4 protein is also a DNA helicase, thus explaining the initial observation of stimulation of DNA synthesis on duplex linear DNA (Bernstein and Richardson, 1988).

A similar complementation based primase assay was used in the initial characterisation of the *E. coli* replicative primase DnaG (Bouché et al., 1975; Rowen and Kornberg, 1978). Here, it was determined that DnaG is required for stimulation of DNA synthesis in extract prepared from *E. coli* expressing a temperature-sensitive DnaG mutant (Rowen and Kornberg, 1978). Furthermore, in the absence of extract, DnaG was able to initiate DNA synthesis on phage G4 DNA in the presence of *E. coli* SSB and DNA polymerase III holoenzyme (Rowen and Kornberg, 1978). Additionally, priming by DnaG was also observed on G4 and M13 DNA, as well as poly(dT) templates, in the presence of the DnaB helicase but in the absence of SSB or DNA polymerase III (Arai and Kornberg, 1979).

Similar assays have also been used in the basic characterisation of other prokaryotic and phage primases (Krevolin and Calendar, 1985; Lanka et al., 1979; Morris et al., 1975). These assays were highly useful in the early analysis and identification of primases. Importantly, they allowed potential primases, initially identified through *in vivo* studies, to be confirmed, in addition to permitting investigation into the requirements of either rNTPs or dNTPs for synthesis, the importance of other replication machinery components on priming and the requirement of different DNA templates and initiation sites.

Nevertheless, the information gleaned from these assays was limited. Notably, primer synthesis could often only be detected in the presence of additional replisome components and results could be obscured by the DNA template used. Furthermore, analysis of reaction products by liquid scintillation counting was not able to provide qualitative information about the length and sequence of the synthesised primers. Consequently, follow-up studies of phage and prokaryotic primases, as well as early analyses of the replicative eukaryotic primase Prim1/2 from yeast and Drosophila, further

resolved primase reaction products by polyacrylamide gel electrophoresis (PAGE) (Biswas et al., 1987; Bouché et al., 1978; Conaway and Lehman, 1982a, 1982b; Romano and Richardson, 1979; Scherzinger et al., 1977; Wu et al., 1992). Here, reactions were assembled in essentially the same manner as described previously using ss phage or poly(dT) DNA templates, however other replisome components such as DNA polymerases and helicases were omitted, allowing direct synthesis by the primase to be examined (Biswas et al., 1987). Following incubation, reactions were typically stopped by supplementation with EDTA and ionic surfactants such as SDS or sarkosyl. Additionally, reactions were often incubated with proteinase K for 1-2 hrs before phenol chloroform extraction or ethanol precipitation. Subsequently, pellets were resuspended in buffer containing formamide, in addition to the dyes bromophenol blue and xylene cyanol, and denatured by heating before loading onto the gel. Electrophoresis was performed on urea-polyacrylamide gels which were autoradiographed using x-ray film (Biswas et al., 1987; Conaway and Lehman, 1982a).

By resolving reaction products on polyacrylamide gels, the synthesised primers could be analysed to single nucleotide resolution, thus providing more direct information about primase activity than liquid scintillation counting alone. Many studies have utilised this approach in the basic characterisation of viral, prokaryotic, and eukaryotic primases in order to determine the length of synthesised primers (Frick and Richardson, 2001). Further processing of RNA primase reaction products by limited alkali or ribonuclease digestion, prior to resolution on urea-polyacrylamide gels, has also been used to determine the sequence of these primers (Romano and Richardson, 1979).

Since the development of gel-based radioactive primase assays in the late 1970s and early 1980s, the basic outline of this technique has largely remained unchanged. However, clean-up steps including proteinase K digestion and phenol chloroform extraction are now generally omitted. This greatly increases the speed and ease of the technique without significantly affecting the quality of results. In addition, advances in oligonucleotide synthesis technology have allowed specific DNA templates of known sequence to be generated and analysed in the assay, thus increasing the applicability of the technique. One example of this is in the analysis of T7 primase recognition sites. In this report, numerous oligonucleotides containing different modified primase recognition sites were synthesised and tested for their ability to act as primer synthesis templates (Frick and Richardson, 1999). This allowed the authors to identify the requirements for sequence-specific DNA binding and primer synthesis by T7 primase. Synthetic oligonucleotides were also used in the early characterisation of human Prim1/2 (Kirk and Kuchta, 1999b). Here, oligonucleotides containing increasing amounts of deoxycytidine

were used to analyse the effect of manganese (Mn^{2+}) ions on primase activity. It was found that the level of stimulation of processivity by Mn^{2+} decreased on templates containing higher amounts of deoxycytidine. The authors suggested that this was due to the inherently higher processivity of Prim1/2 on deoxycytidine-rich templates (Kirk and Kuchta, 1999b). Thus, the use of synthetic oligonucleotides in primase assays has a number of advantages over traditional phage DNA templates. These include the ability to use shorter oligonucleotides of known sequence, allowing primase initiation sites and product sequences to be much more easily determined. Furthermore, synthetic templates can be easily modified in order to examine the effect of different sequences and DNA secondary structures on primase activity. Nevertheless, phage DNA, in particular ss M13, is often still used as a general template to identify primase activity, due to its affordability and circular nature, which prevents snap-back and extension of the 3' end of the template.

In addition to advances in oligonucleotide synthesis, developments in other accessory components of the radioactive primase assay have increased the speed and ease with which it can be performed. One particular advancement is in the methods used to detect radiolabelled primase products. Traditionally, autoradiography was used to detect radioactivity by exposing X-ray films to polyacrylamide gels following resolution. The resulting images could then be quantified by densiometry, permitting kinetic analysis of the reaction products. However, these techniques can be time-consuming and require long exposure times for high sensitivity to be achieved. In addition, X-ray film is easily over exposed, resulting in signal lying outside the linear dynamic range and thus preventing accurate quantification. Generally, radioactive gels are now analysed by phosporimaging. Here, storage-phosphor screens replace X-ray films, subsequent scanning of these with a helium-neon laser causes emission of luminescence proportional to the level of radiation which is quantified using a photomultiplier, consequently producing a digital image. The resulting image and reaction products can then be quantified using image analysis software. In comparison to X-ray film detection, phosphorimaging has greatly increased, sensitivity, linear dynamic range, and speed of image development, despite increased expense and lower resolution (Van Kirk et al., 2010).

In summary, the radioactive gel-based primase assay has remained largely unchanged since its initial development in the late 1970s, however refinements of the technique and advances in accessary components have greatly increased the speed and ease with which it can be performed. Typically reactions are now assembled in optimum buffer conditions containing, the primase of interest, activating metal ion cofactors, a synthetic

or phage ssDNA template of known length and sequence, unmodified dNTPs or rNTPs, and radiolabelled dNTPs or rNTPs. Following incubation, reactions are stopped by addition of formamide and EDTA, before heat denaturation and loading onto a urea-polyacylamide gel. After resolution, the gel is visualised by phosphorimaging, the resulting digital image can then be analysed and quantified using image analysis software (outlined in Figure 4.1.).

Despite relatively few changes to the original gel-based primase assay, the technique is still the "go-to" option when analysing primase activity due to its excellent sensitivity and ability to provide information on the size and yield of reaction products. This is highlighted by the use of the assay in the characterisation of recently discovered archaeal and viral primases (De Silva et al., 2009; Galal et al., 2012; Lipps et al., 2004; Silva et al., 2007; Zuo et al., 2010). However, the assay also has a number of limitations and disadvantages. Perhaps the most significant of these is the reliance on potentially hazardous radioactivity. This requires training in the handling and disposure of radioactivity, in addition to the implementation of rigorous safety measures, before the assay can be performed. Coupled with this, radiolabelled dNTPs have a short half-life, limiting the amount of time that they can be stored and increasing the cost of the assay if used over a long period of time. Radioactivity imaging methods can also suffer from poor linear dynamic range, indeed over exposure can result in a signal outside of the linear range, therefore preventing accurate quantification. Additionally, in spite of refinements of the original assay, the technique is still relatively time consuming thus making it unsuitable for certain applications, such as HTS of primase inhibitors.

4.2.3.2. A High-Throughput Radioactive-Based Primase Assay

Due to the limited applicability of the traditional gel-based radioactive primase assay in HTS approaches, a modified 96-well plate scintillation proximity assay (SPA) was developed to screen *E. coli* DnaG inhibitors (summarised in Figure 4.2.) (Zhang et al., 2002). In this assay, reactions are assembled in a 96-well plate in the typical manner, containing an appropriate buffer, metal ion cofactors, unlabelled dNTPs, [^{3}H]CTP, a ssDNA template, DnaG, DnaB, and the test compound or DMSO. Following incubation, a suspension of polyvinyl toluene-polyethyl-eneimine (PVT-PEI) coated SPA beads are added and after 1 hour the plates are read on a Topcount instrument (Packard) (Zhang et al., 2002). The assay is based upon the capture of primase reaction products on the PVT-PEI SPA beads. This capture brings the ^{3}H-labelled reaction products in close proximity to the SPA beads. Consequently, decay of the ^{3}H releases β-particles, which stimulate the scintillant in the beads to emit photons. The emitted photons can then be

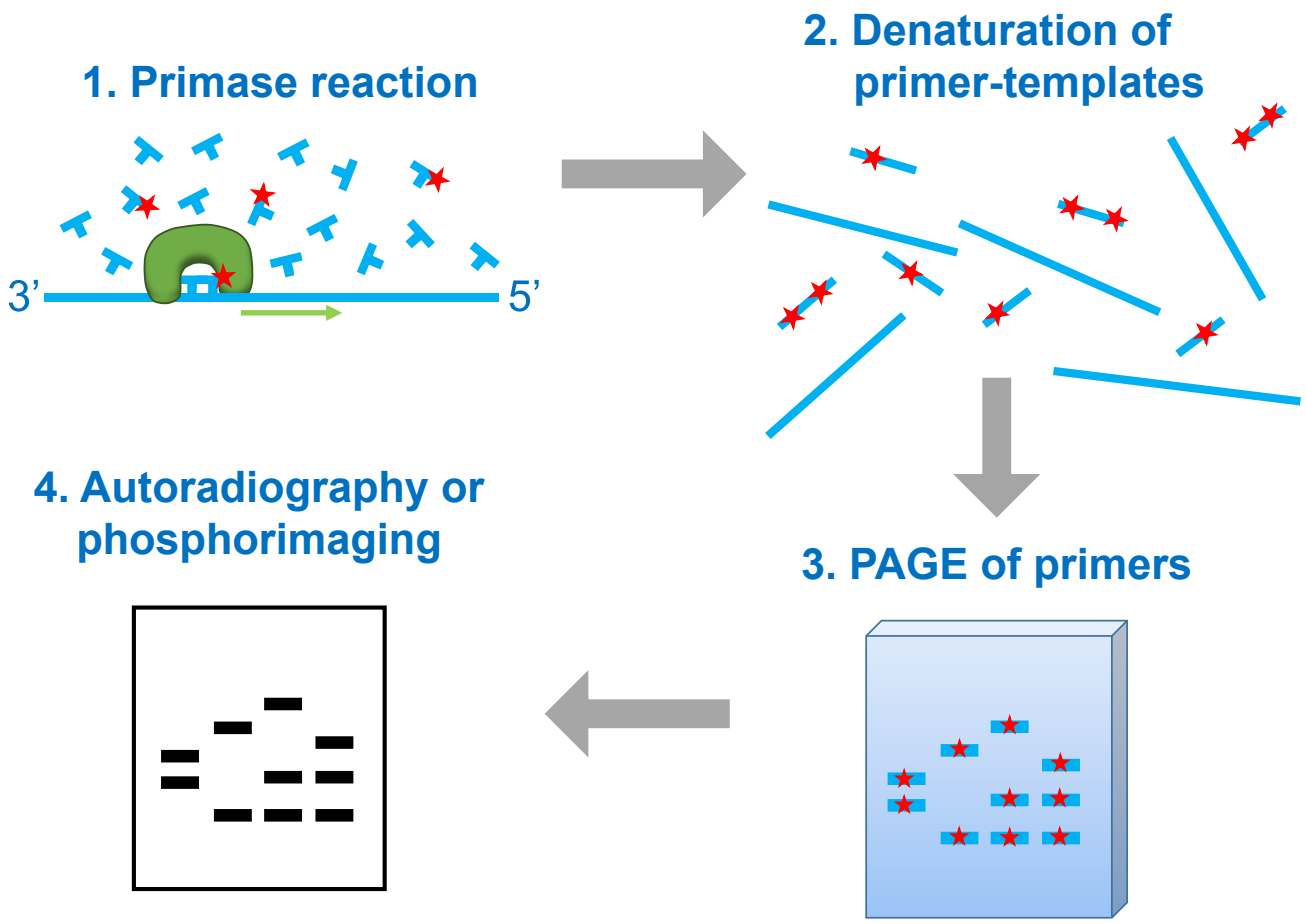


Figure 4.1. Outline of the radioactive gel-based primase assay.
1. Primase reactions are assembled containing the primase of interest (shown in green), a ssDNA template, native rNTPs or dNTPs, and radiolabelled rNTPs or dNTPs (indicated by red star) in an appropriate buffer. Incorporation of radiolabelled nucleotides during synthesis and extension generates radiolabelled primers. 2. Following primer synthesis, reactions are quenched through addition of EDTA and primer-template duplexes are denatured by heating in buffer containing formamide. 3. Reaction products are loaded onto a denaturing urea-polyacrylamide gel and electrophoresis is performed to resolve the radiolabelled primers. 4. Synthesised primers are visualised by autoradiography using X-ray film, or phosphorimaging.

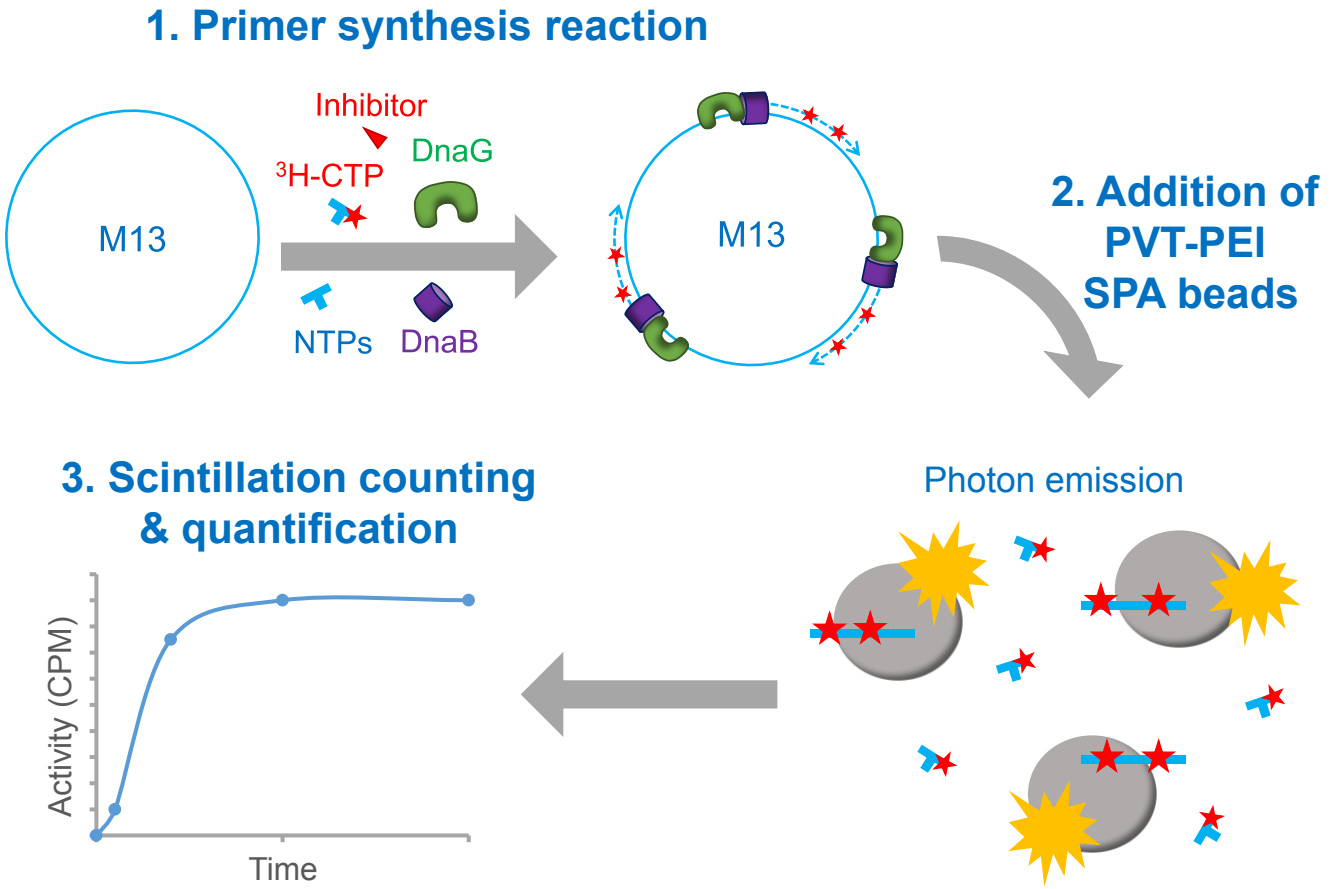


Figure 4.2. Overview of the high throughput radioactive-based primase assay.
1. Primer synthesis reactions are assembled in 96-well plates containing *E. coli* DnaG primase (shown in green), DnaB helicase (shown in purple), native rNTPs, ^{3}H-CTP (indicated by red star), ssM13 template DNA, and the test compound (shown as red triangle) or DMSO, in an appropriate buffer. Incubation of reactions permits synthesis of ^{3}H-CTP-labelled primers on the M13 template. 2. Polyvinyl toluene-polyethyl-eneimine (PVT-PEI) scintillation proximity assay (SPA) beads are added to the reaction. Capture of the radiolabelled primers on the PVT-PEI SPA beads stimulates photon emission from the scintillant in the beads. Free ^{3}H-CTP does not bind the beads, and therefore does not stimulate photon emission. 3. Photon emission is detected and quantified using a photomultiplier tube based scintillation counter.

detected and quantified using a photomultiplier tube based scintillation counter. Free ^{3}H[CTP], not bound to the SPA beads, does not stimulate photon emission due to the insufficient energy of the β-particles to reach the beads. Thus, SPA has the advantage that no separation or washing step is required to remove the free ^{3}H[CTP].

This assay overcomes the time-consuming nature of the traditional gel-based primase assay and extends the application of the technique to HTS. The assay provides a sensitive and efficient method to quantify primase activity and screen DnaG inhibitors that, when coupled with a DnaB helicase assay, can provide insights into the mechanism of action of those compounds (Zhang et al., 2002). However, the technique is still reliant on the use of radioactivity and, due the HTS approach, is costly and may generate large amounts of liquid waste, potentially making it unsuitable in an academic setting. Additionally, unlike gel-based approaches, the assay does not provide qualitative information about the size or sequence of the synthesised primers, which may have important implications for the interpretation of the mechanism of action of inhibitors being screened.

4.2.4. Non-Radioactive Primase Assays

Given the drawbacks of radioactive-based primase assays, a number of alternative non-radioactive assays have been developed, some of which are applicable to HTS. In this section, the background to these methods and their advantages and disadvantages will be discussed.

4.2.4.1. Thermally Denaturing High-Performance Liquid Chromatography Primase Assay

In order to avoid the cost and safety issues associated with radioactive based assays, an alternative primase assay was developed, based upon high-performance liquid chromatography (HPLC) analysis of reaction products (outlined in Figure 4.3.) (Koepsell et al., 2004). To develop the assay, the authors made use of *E. coli* DnaG and enzymatic reactions were assembled in appropriate buffer conditions containing the primase, ssDNA template (< 30 nt in length) blocked at the 3' end, and native rNTPs. Importantly, by blocking the 3' end of the template, the user can directly examine *de novo* synthesis by the primase, rather than elongation from a 3'-end hairpin on the synthetic ssDNA template, which can be produced by template snap-back. Notably, other assays, including the radioactive HTS method, did not control for this phenomenon, potentially generating misleading data and interpretation of results. Indeed, Koepsell *et al.* found that when the 3' end of the template was not blocked, 10-fold more primase was required for *de novo* primer synthesis and the rate constant of primer synthesis was three times

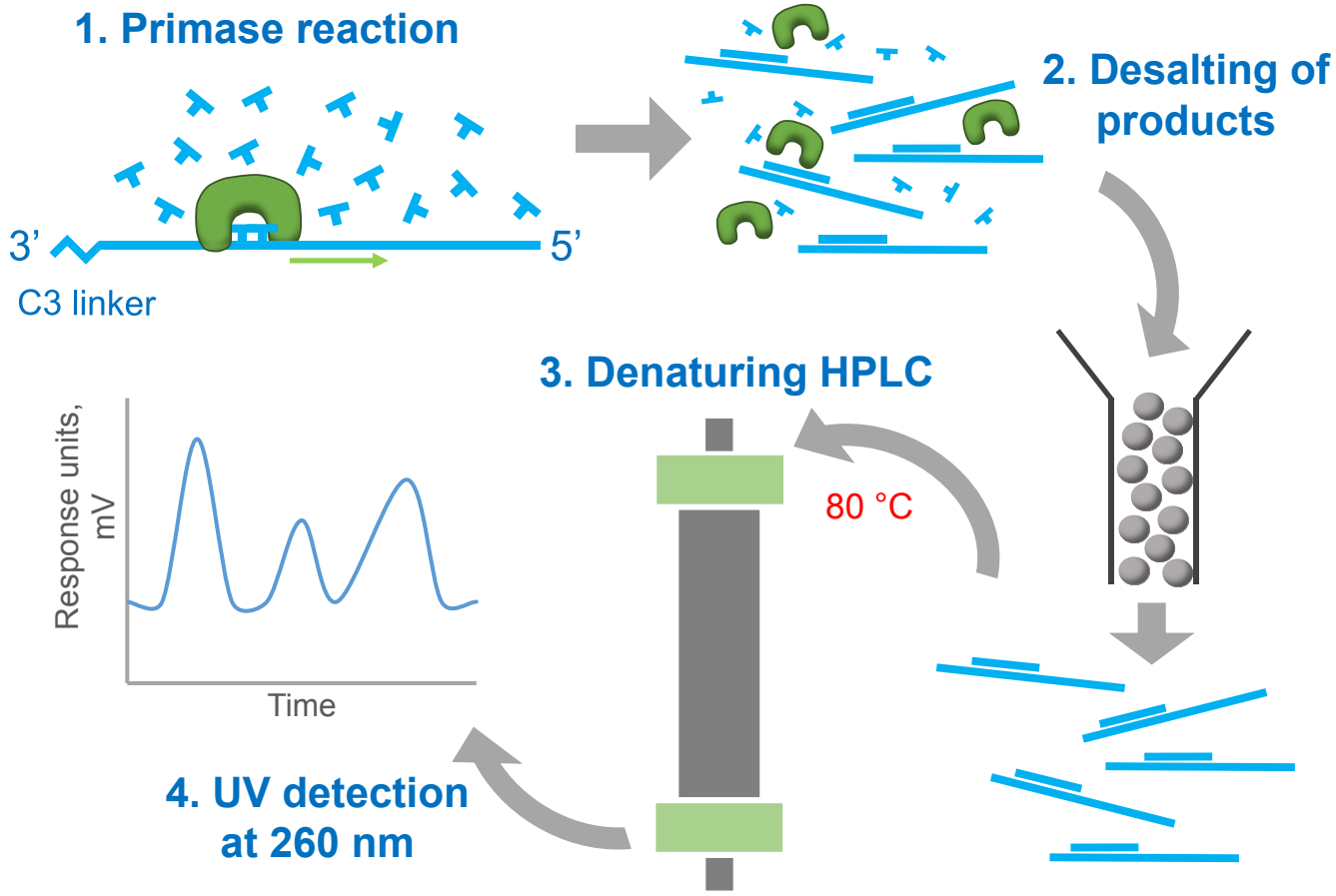


Figure 4.3. Summary of the thermally denaturing HPLC primase assay.
1. Primase reactions are assembled containing the primase of interest (shown in green), native dNTPs or rNTPs, and a ssDNA template blocked at the 3' end by a C3 linker, in an appropriate buffer. Assembled reactions are incubated allowing synthesis of unlabelled primers. 2. Reactions are stopped by heating, before being desalted, dried, and resuspended in water. 3. Samples are then loaded onto an alkylated non-porous polystyrene-divinylbenzene copolymer microsphere bead column and analysed by HPLC under thermally denaturing conditions (80 °C). 4. Primase reaction products and template DNA are detected upon elution by monitoring UV absorbance at 260 nm. The resulting chromatogram can be used for quantification of primer synthesis by comparison to a standard curve, taking into account the extinction coefficient of the oligonucleotide.

greater than that reported when using the radioactive HTS method (Koepsell et al., 2004; Zhang et al., 2002). Following incubation, reactions were stopped by heat denaturation, and desalted and dried, before resuspension in 1/10th the reaction volume of water. Subsequently, reaction products were analysed by HPLC on an alkylated non-porous polystyrene-divinylbenzene co-polymer microsphere bead column under thermally denaturing conditions. UV detection of the eluted oligonucleotides at 260 nm produced chromatograms with peaks corresponding to the template and various smaller products (Koepsell et al., 2004). Reaction products could then be quantified by analysing the area under each peak, taking into account variations in extinction coefficients between oligonucleotides, and compared to a standard curve. Importantly, analysis by denaturing HPLC allows reaction products to be separated by both size and hydrophobicity, thus producing qualitative, as well as quantitative, information about the synthesised primers. The assay was used to determine the kinetics of *de novo* primer synthesis by DnaG, as well as to identify the IC50 for dNTP inhibition of primase activity (Koepsell et al., 2004).

The denaturing HPLC primase assay has a number of advantages over traditional gel-based radioactive assays. Perhaps the greatest of these is that the assay is performed with native rNTPs/dNTPs, removing all hazards associated with radioactivity. Additionally, the removal of radioactivity also decreases the cost of the assay, discounting initial costs for equipment. Like the traditional gel-based assays, this method is able to provide sensitive qualitative and quantitative information about primer synthesis. Furthermore, the HPLC analysis is automated and scalable to a degree, with each run taking ~ 20 mins (Koepsell et al., 2004). This makes the assay much quicker for individual experiments requiring only a short number of runs, when compared to gel based assays. However, with larger experiments requiring multiple runs, the analysis time can be much greater and it is here that gel based assays have the advantage of resolving multiple samples at the same. Likewise, this method lacks sufficient throughput for HTS approaches, making it unsuitable for screening large chemical libraries to identify inhibitors. Quantification of reaction products using chromatogram peaks can also be more difficult than analysing gels, notably variations in extinction coefficients between products requires knowledge of the nucleotide content of the peak, in addition to the generation of a standard curve. Coupled with this, HPLC analysis requires optimisation to each specific ssDNA template being used, prior to performing experiments.

4.2.4.2. A Fluorometric High-Throughput Primase Assay

Given the limitations of the HPLC primase assay in HTS, Koepsell *et al.* then developed a high throughput microplate-based fluorescent primase assay, adaptable to robotic

screening methods (summarised in Figure 4.4.) (Koepsell et al., 2005). This assay is based on PicoGreen nucleic acid dye, a fluorochrome that binds specifically to dsDNA. When bound, PicoGreen fluoresces at an excitation maximum of 480 nm, with an emission peak at 520 nm. The dye was previously found to offer an effective and sensitive way to quantify dsDNA due to its high level of fluorescence enhancement upon DNA binding, thus making it suitable for the detection of primer-template duplexes (Ahn et al., 1996). In this high throughput fluorescent primase assay, reactions are assembled in a 96-well microplate and incubated for the desired time to allow primer synthesis. Following incubation, PicoGreen dye is added, which binds to the RNA-DNA duplexes generated from primer synthesis and fluoresces upon excitation, allowing detection and quantification of primase reaction products using a spectrofluorometer. Additionally, PicoGreen dye quenches the primase reaction, removing the need for quenching with EDTA (Koepsell et al., 2005).

This high-throughput fluorometric primase assay, therefore, offers a non-radioactive alternative for HTS studies of potential primase inhibitors. The assay is able to provide quantitative information on primer synthesis but, like the radioactive HTS method, does not generate qualitative information, such as primer length or sequence. Although the microplate format, fast analysis time, and ability to function in the presence of DMSO, makes the method an attractive option for HTS, there are a number of potential drawbacks to the technique. Firstly, the assay is effective in detecting primers longer than ~ 6 nt due to their stable association with the template DNA. However, shorter primers may not provide the stable duplex required for PicoGreen binding and fluorescence enhancement. Secondly, a potential issue in using fluorometric assays for HTS is the interaction of non-polar and aromatic compounds with the fluorescent label, which may interfere with the signal and obscure results (Biswas et al., 2012).

4.2.4.3. A High-Throughput Primase-Pyrophosphatase Activity Assay

A more recently developed alternative to the fluorometric primase assay, with similar HTS applications, is the primase-pyrophosphatase assay (outlined in Figure 4.5.) (Biswas et al., 2012). In this assay, primase activity is coupled to inorganic pyrophosphatase (PPase), which cleaves the pyrophosphate (PPi) released during the priming reaction into phosphate (Pi). The Pi concentration can then be measured using malachite green reagent (MGR), which displays increasing absorbance at 620 nm as the concentration of Pi increases, producing a colour change from yellow to green that is quantifiable using a plate reader. Importantly, PPase does not exhibit any cleavage activity on NTPs, consequently making the enzyme's activity dependent upon PPi

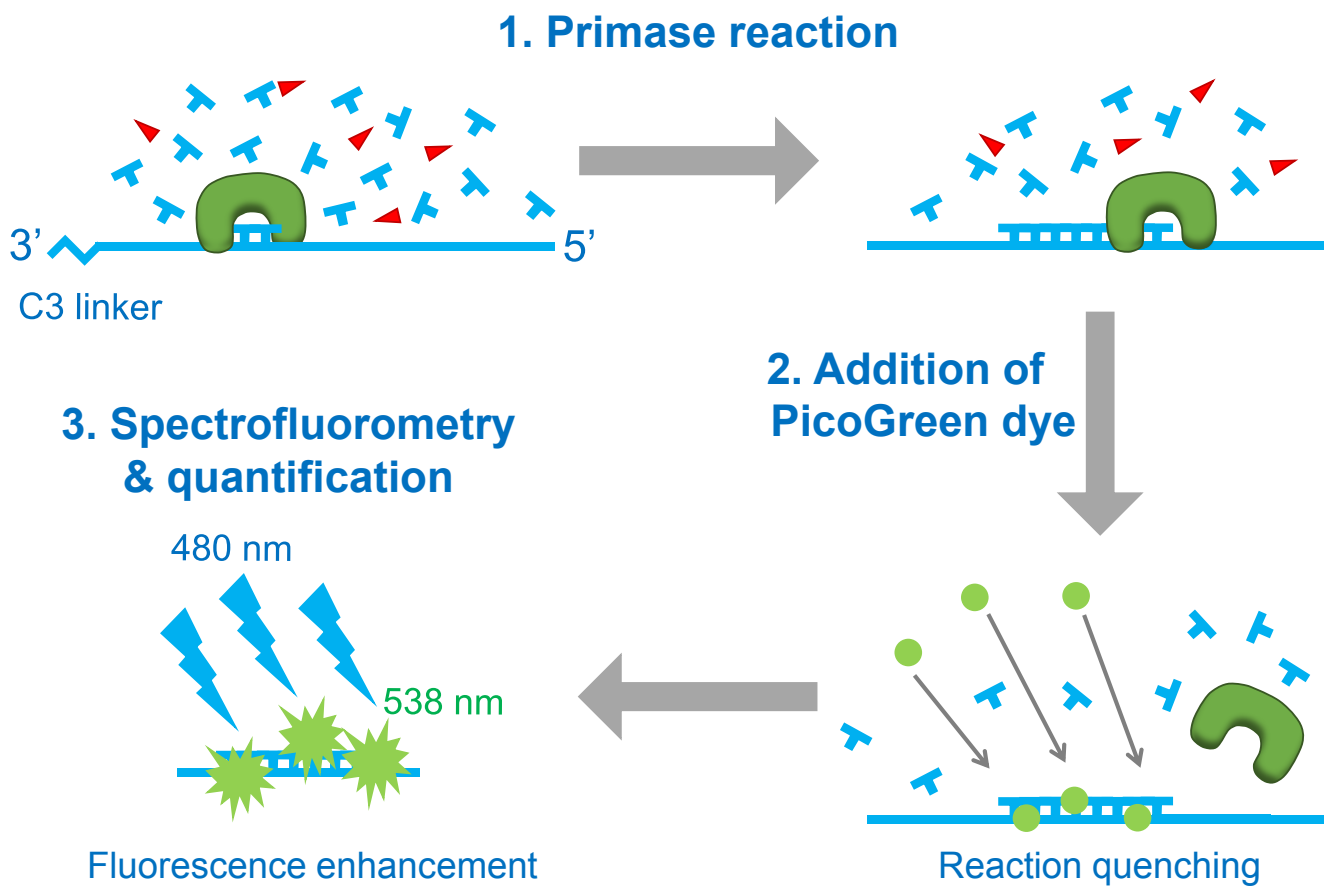


Figure 4.4. Overview of the fluorometric high-throughput primase assay.
1. Reactions are assembled in a 96-well plate containing the primase of interest (shown in green), a ssDNA template with a blocked 3’ end, unlabelled dNTPs, and the chemical compound to be screened (indicated by red triangle) or DMSO, in a suitable buffer. Incubation of reactions facilitates primer synthesis and extension, generating primer-template duplexes. 2. PicoGreen dye (shown as a green circle) is added to the reaction which binds to the primer-template duplexes, producing fluorescence enhancement. Addition of PicoGreen dye also quenches the reaction, removing the need to add EDTA. 3. The 96-well plate is scanned using a spectrofluorometer with an excitation at 485 nm and an emission at 538 nm. Thus, the level of fluorescence is determined by the amount of PicoGreen dye bound to dsDNA, which is dependent upon the amount of primer synthesis.

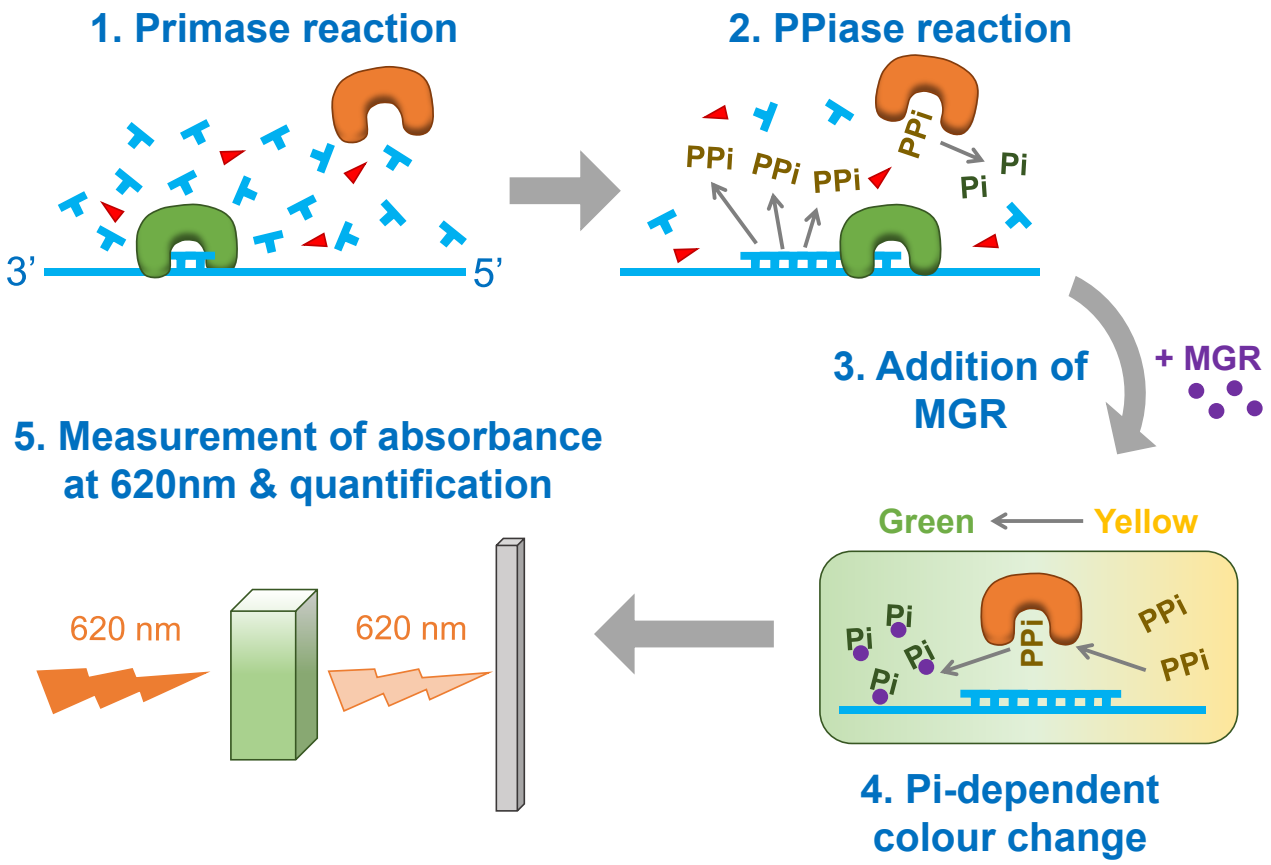


Figure 4.5. Outline of the high-throughput primase-pyrophosphatase activity assay.
1. For HTS, primase reactions are assembled in 384-well plates containing *Mtb* DnaG (shown in green), *Mtb* pyrophosphatase (PPase) (shown in orange), unlabelled rNTPs, a ssDNA template, and the inhibitors to be tested (shown as a red triangle) or DMSO, in a suitable buffer. Incubation of the reactions permits primer synthesis and extension. 2. Incorporation of nucleotides by the primase during synthesis releases pyrophosphate (PPi). The released pyrophosphate is then cleaved by PPase into phosphate (Pi). 3. Malachite green reagent (MGR) (indicated by purple circle) and sodium citrate are added to the reaction. MGR forms a complex with the phosphate which produces a colour change from yellow to green. Thus, the colour change is dependent on the level of free Pi, which is consequently dependent upon PPi release, and ultimately the level of primase activity. 4. The colour change is detected and quantified by measuring absorbance at 620 nm in a plate reader.

released during primer synthesis (Biswas et al., 2012). Additionally, ssDNA templates lacking thymidine nucleotides in their 5' half are used, allowing ATP to be omitted and thus preventing any background signal from being generated by ATPase activity present in the protein preparations.

Biswas *et al*. were able to successfully use this assay to screen 2560 small molecules for *Mycobacterium tuberculosis* DnaG (*Mtb* DnaG) inhibition (Biswas et al., 2012). For HTS, reactions were assembled in 384-well plates containing the inhibitors, rNTPs, ssDNA template, *Mtb* DnaG, and *Mtb* PPase, in the appropriate buffer. After incubation, MGR and sodium citrate were added and absorbance at 620 nm was measured in a plate reader (Biswas et al., 2012). Thus, PPi release can be used to provide a quantitative measure of primer synthesis, permitting kinetic analyses of *Mtb* DnaG. Using this method for HTS, the authors identified a number of hits for DnaG inhibitors, including suramin and doxorubicin. Further kinetic analysis of primase activity, in the presence of these inhibitors, measured the release of PPi as a function of inhibitor concentration under various DNA and NTP concentrations. This analysis provided insights into the mode of inhibition, which suggested that the inhibitors may act by blocking the binding of DnaG to DNA (Biswas et al., 2012).

The primase-pyrophosphate assay is thus a proven method for HTS of primase inhibitors. Like the previously discussed assays, the method benefits from being non-radioactive, significantly reducing potential hazards and making it more suitable for HTS in an academic setting. Additionally, the assay is quick to perform and directly measures NTP incorporation, allowing any primase activity to be detected regardless of primer length. Importantly, however, the use of PPase in the method requires further analysis to confirm that the screened inhibitors are acting upon the primase itself and not PPase. Indeed, the authors used the traditional gel-based radioactive primase assay to confirm that DnaG was inhibited by the identified compounds (Biswas et al., 2012).

4.2.5. A Fluorescent Gel-Based Primase Assay

Given the disadvantages of radioactive assays, and the limited ability of non-radioactive HTS methods to generate qualitative information (e.g. primer length or sequence), we aimed to develop a gel-based non-radioactive primase assay with the same benefits and basic set-up as the classic radioactive assay but without the hazards and time-consuming nature of radioactive work. To this end, we have developed a fluorescent gel-based primase assay, which utilises 6-carboxyfluoroscein (6-FAM) labelled nucleotides in place of radioactivity. In this section, the background and outline of the method will be described, before moving onto provide a detailed method for the assay using PrimPol as

an example primase and, finally, a discussion of the advantages and disadvantages of the technique.

4.2.5.1. Theory and Overview of the Fluorescence-Based Primase Assay

The fluorescence-based primase assay described here has the same basic layout as the traditional radioactive-based primase assay (Figure 4.6.). Reactions are assembled in appropriate buffer conditions containing the purified primase to be studied, a ssDNA template, native dNTPs or rNTPs, 6-FAM dNTPs or rNTPs (~ 100 fold lower concentration than unmodified nucleotides), and the required activating divalent metal ions. Incubation of the assembled reactions permits primer synthesis and extension on the ssDNA template. The majority of synthesis occurs using unlabelled nucleotides, however incorporation of 6-FAM dNTPs or rNTPs during extension of the primers allows the reaction products to be visualised. Following quenching of the reactions, background given by the fluorescent nucleotides can be reduced using a number of optional DNA precipitation and clean-up approaches. The resulting reaction products are subsequently resolved on a denaturing urea-polyacrylamide gel and visualised on a fluorescence image reader.

4.2.5.2. Preparation of Primase Assay Reagents

Prior to performing the assay, the primase of interest must be expressed and isolated to a high level of purity. Importantly, the purified primase must be free from contaminating primases, polymerases, and nucleases, which can interfere with the interpretation of results. In addition to preparing the primase of interest, a suitable purified ssDNA template must be obtained. Both synthetic linear and circular phage templates are suitable for this assay, although reaction products smaller than ~10 nucleotides can be difficult to distinguish from background without further extension, therefore templates >10 nt are recommended.

4.2.5.3. Primer Synthesis Reaction

4.2.5.3.1. Buffers and Reagents

- 10 x TBE: 1 M Tris (pH 7.6), 1 M boric acid, 20 mM EDTA
- 7 M urea 15 % polyacrylamide gel mix 60 ml: 28.8 g urea, 22.5 ml acrylamide : bisacrylamide (19:1), 199.2 µl APS, 24 µl TEMED, 6 ml 10 x TBE (add APS and TEMED immediately before pouring gel) (see tip 1 and 2)
- 10 x reaction buffer: 100 mM Bis-Tris-Propane-HCl (pH 7.0), 100 mM $MgCl_2$, 10 mM DTT (buffer available from NEB as NEBuffer 1)

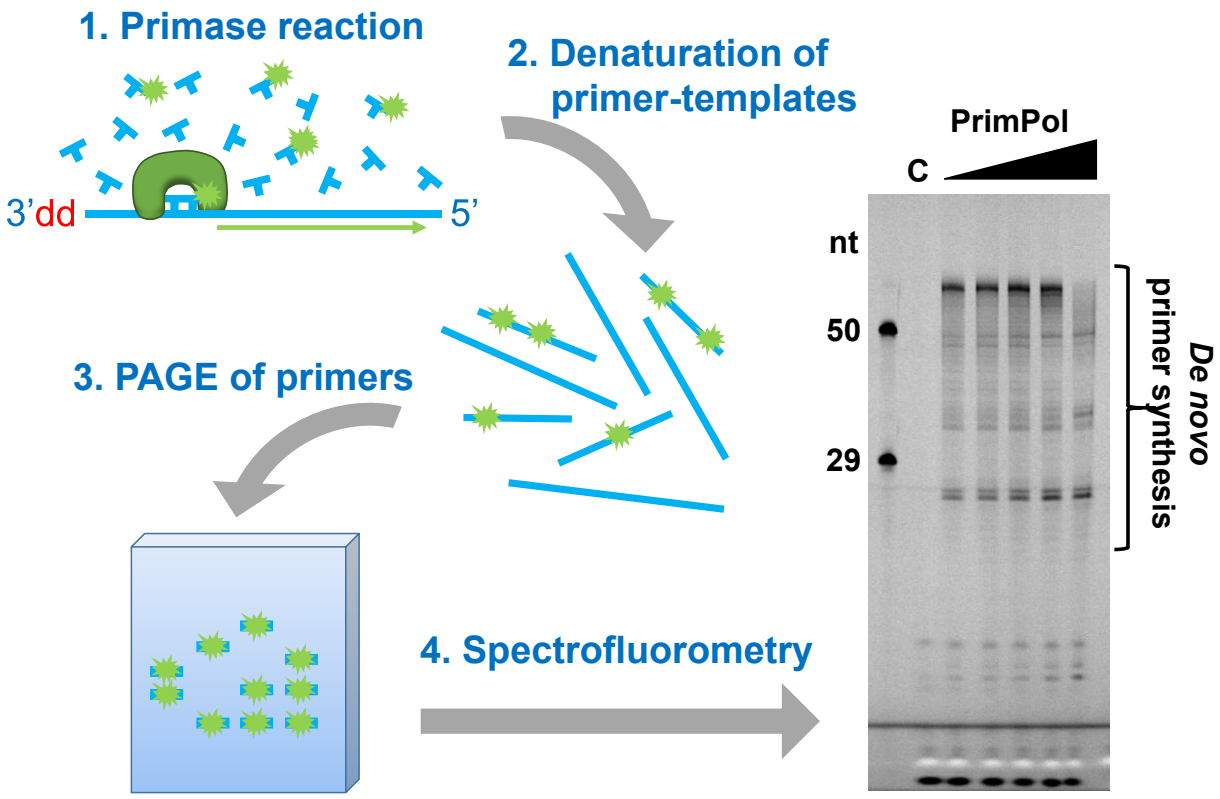


Figure 4.6. Summary of the fluorescence gel-based primase assay.
1. Primase reactions are assembled containing the primase of interest (shown in green), a ssDNA template blocked at the 3' end with a dideoxynucleotide, unlabelled dNTPs or rNTPs, and 6-FAM labelled dNTPs or rNTPs (indicated by green star), in an appropriate buffer. Incubation of the reaction facilitates primer synthesis and extension with incorporation of 6-FAM dNTPs/rNTPs during synthesis by the primase generating fluorescently labelled primers. 2. Reactions are quenched in buffer containing EDTA and formamide, and primer-template duplexes are denatured by heating. 3. Reaction products are resolved on a denaturing urea-polyacrylamide gel. 4. Following electrophoresis, gels are scanned with a fluorescent image reader. An example gel image is shown. Here, the fluorescence gel-based primase assay was performed as described in 'experimental procedure' using increasing concentrations of PrimPol (0.25, 0.5, 1, 2, and 4 μM) (indicated by black triangle) on a 66 nt template (sequence shown in 'buffers and reagents'), blocked at the 3' end with a dideoxynucleotide. Reactions were incubated for a single 15 min time point. C indicates the 'no enzyme control'. Nucleotide size markers of 50 and 29 nt are shown to the left of the image.

- 10 µM ssDNA template (5'-Biot-GTCTTCTATCTCGTCTATATTCTATTGTCTCT ATGAATACCTTCATCAGTCTCACATAGATGCAT-dideoxyC-3' or another suitable ssDNA template)
- 2.5 mM dNTP stock solution: 2.5 mM of each dNTP (NEB), diluted in ddH_2O (dNTPs can be replaced with rNTPs if required)
- 25 µM FAM dNTP stock solution: 25 µM N^6-(6-amino)hexyl-dATP-6-FAM, 25 µM 5-Propargylamino-dCTP-6-FAM, 25 µM Aminoallyl-dUTP-6-FAM (JenaBioscience), diluted in ddH_2O (can also be replaced with 6-FAM rNTPs)
- Purified Primase: PrimPol (or another primase of interest)
- 2 x stop buffer: 95 % formamide, 20 mM EDTA, 0.25 % bromophenol blue and xylene cyanol (see tip 3)
- FAM-labelled oligonucleotide size marker

4.2.5.3.2. Equipment

- Vertical nucleic acid PAGE set up (adjustable gel slab system, 165 mm x 280 mm glass plates, 0.75 mm spacers, 20 well combs) (C.B.S.Scientific) or suitable equivalent (see tip 4)
- Incubator, water bath or dry heating block
- Standard microcentrifuge
- FLA-5100 fluorescent image analyser (Fujifilm)
- ImageQuant TL for image analysis

4.2.5.3.3. Experimental Procedure

1. Before assembling the primase assay reactions, urea-polyacrylamide gels should be prepared and poured according to the manufacturer's instructions. This will allow sufficient time for setting. Note that the gel should be pre-run in 1 x TBE for 0.5 - 1 hr before loading (see tip 5).
2. On ice, assemble 10 µl reactions for each variable in the following order: 5 µl ddH_2O, 1 µl 10 x reaction buffer (1 x final), 1 µl ssDNA template (1 µM final), 1 µl dNTP stock solution (250 µM final), and 1 µl FAM dNTP stock solution (2.5 µM final) (see tip 6). If taking multiple time points, make one stock reaction with 10 µl per time point and an additional 10 µl to account for pipetting errors, e.g. if taking

5 time points make a 60 µl reaction. A “no enzyme” control should also be prepared. Exposure of the FAM dNTP stock, and reactions containing FAM dNTPs, to light should be kept to a bare minimum.

3. On ice, make a 10 x stock of PrimPol using 1 x reaction buffer for dilution. For each primase, the concentration required to give the desired level of activity should be determined by testing a range of concentrations. In this case, the amount of ddH_2O added to the reaction can be adjusted to account for the changing volume.

4. Pre-incubate the assembled reactions at 37 °C for 5 mins (see tip 7).

5. Initiate the reaction by adding 1 µl of 10 x PrimPol stock and mix by pipetting.

6. Incubate the reactions for the desired time point or time course at 37 °C (see tip 7).

7. Stop the reaction by adding 10 µl 2 x stop buffer. If using a stock reaction and taking multiple time points, 10 µl of the reaction should be removed and added to 10 µl of 2 x stop buffer.

8. Incubate the quenched reactions at 90 °C for 3 mins and spin briefly in a microcentrifuge.

9. Load each 20 µl sample onto the pre-run urea-polyacrylamide gel and resolve according to the manufacturers recommendations (see tip 5 and 8). Observe the migration of the bromophenol blue and xylene cyanol dyes to monitor progression of the samples.

10. A FAM-labelled oligonucleotide size marker should be run alongside the samples to allow determination of product sizes.

11. Before imaging, the gel system should be disassembled and the plates thoroughly cleaned with dH_2O and ethanol. Failure to do this will affect the image quality. Note that it is not necessary to remove the gel from the glass plates before scanning.

12. Visualise the gel using an FLA-5100 image reader, or an equivalent imager.

13. The resulting digital image can be analysed using image analysis software, such as ImageQuant TL, if quantification of primase reaction products is desired.

4.2.5.3.4. Tips

1. For optimal resolution the concentration of acrylamide should be adjusted according to the expected product size. A higher polyacrylamide concentration will resolve smaller oligonucleotide products.
2. Following the addition of APS and again after the addition of TEMED, the gel mix should be mixed gently by inverting to avoid aeration and the generation of bubbles.
3. The inclusion of these dyes in the stop buffer can interfere with the fluorescent signal if migrating at the same size as the reaction product. In this case, the dyes can be omitted from the stop buffer and run in an empty well to still allow monitoring of sample progression.
4. Using a large sequencing PAGE set up allows much great resolution of reaction products compared to smaller gels. Addition of a metal heat dispersion plate in the set-up can help prevent 'smiling' of the samples.
5. Prior to pre-running the gel and again before loading samples, the wells should be washed 2 x using a 1 x TBE filled syringe and needle to remove any gel pieces and urea.
6. If assembling a large number of reactions, it is beneficial to make a single reaction stock and aliquot this for each sample, taking into account the different amount of ddH_2O which might be required for each reaction depending on the variable being analysed. This also minimises variation between reactions caused by pipetting error.
7. Although a bench top dry heating block is sufficient for short time-courses, an incubator should be used for long time points to prevent evaporation and condensation on the inside of the Eppendorf tube lid.
8. If all of the wells of the gel are not being used, loading of 'blank samples', containing 1 x TBE and 1 x stop buffer, into the empty wells can help prevent gel 'smiling'.

4.2.5.3.5. Methods to Reduce Background

Although the primase assay can be used reliably without any clean-up steps to remove unincorporated 6-FAM dNTPs, background signal from the dNTPs is visible lower down the gel. To remove this background and improve the quality of results, a number of DNA

precipitation techniques can be implemented. These clean-up steps should be performed following incubation of the reactions, but before the addition of stop buffer (between Steps 6 and 7).

We have reliably precipitated primer-template duplexes and removed free 6-FAM dNTPs using magnetic streptavidin beads (Roche) (note that the ssDNA template must be biotin labelled to use this method), Oligo Clean & Concentrator columns (Zymo Research), and ethanol precipitation (Figure 4.7.). Of these techniques, ethanol precipitation, following a standard protocol (note that incubation on ice for 10-15 mins is sufficient for precipitation), is the most cost-effective and efficient method. If using this method, note that the pellet will likely not be visible. Use of the Oligo Clean & Concentrator columns (Zymo Research) or equivalent, following the manufacturer's protocol, may be desirable in the interest of time when analysing a large number of samples. Following precipitation, the DNA should be suspended in 1 x stop buffer and the protocol continued as described above.

4.2.5.4. Considerations when Performing the Fluorescence-Based Primase Assay

The fluorescent gel-based primase assay described above has been successfully used in both published and unpublished primase studies from our group (Kobayashi et al., 2016; Schiavone et al., 2016). However, there are a number of points which must be considered before performing the assay. Firstly, we have used this assay to study both human PrimPol and archaeal replicative primases with success. These primases are able to incorporate the 6-FAM labelled dNTPs during primer synthesis and extension (Figure 4.8.). However, it must be noted that 6-FAM labelled dNTPs are significantly modified in comparison to native or radiolabelled dNTPs and thus care must be taken to ensure that the primase of interest is able to efficiently incorporate these modified nucleotides. Note that PrimPol and replicative archaeal primases preferentially synthesise primers using dNTPs over rNTPs, hence the use of 6-FAM dNTPs here. In the case that the primase of interest synthesises RNA primers, 6-FAM dNTPs can be substituted for 6-FAM rNTPs, which are also commercially available.

Secondly, when using a linear ssDNA template with a free 3' end, products larger than the template may be observed (Figure 4.9.). We have determined that this is likely due to formation of a hairpin at the 3' end of the template, produced by snap-back (see section 4.3.). Extension of this 3' end hairpin by the primase consequently produces reaction products much larger than expected and may additionally sequester the enzyme away from performing *de novo* primer synthesis. This can be avoided by blocking the 3' end of the template with a 3' C3 spacer, dideoxynucleotide, or other suitable modification.

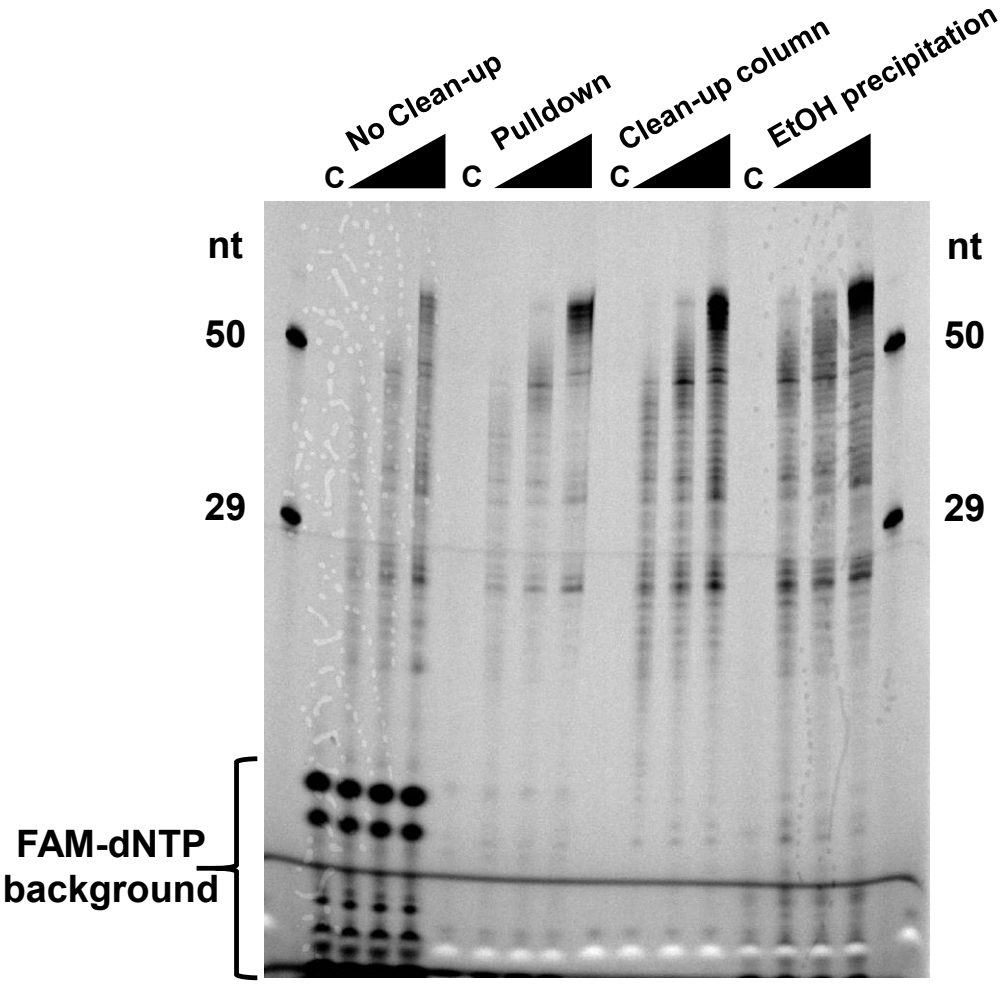


Figure 4.7. DNA precipitation methods used to reduce FAM dNTP background.
The assay was performed as described in 'experimental procedure' using 250 nM PrimPol. The black triangle indicates increasing time points of 1, 5, and 10 mins. Following incubation, reactions were either immediately quenched with stop buffer (as indicated in 'buffers and reagents') (shown on the left as 'no clean-up') or were subject to DNA precipitation clean-up techniques. 'Pull down' reaction samples were quenched with binding-washing buffer (10 mM Tris-HCl pH 7.5, 500 mM NaCl, 10mM EDTA) and supplemented with 20 µl streptavidin coated beads (Roche). Binding was performed for 1 hr at 4 °C and samples were subsequently washed 3 x 1 ml with binding-washing buffer, before resuspension in 1 x stop buffer. 'Clean-up column' reaction samples were bound, washed, and eluted from Oligo Clean & Concentrator columns (Zymo Research) according to the manufacturer's instructions, before resuspension in 1 x stop buffer. 'EtOH precipitation' samples were supplemented with 1/10 volume 3M NaOAc and 3 volumes of 100% EtOH, before incubation for 15 mins on ice. Following incubation, samples were spun in a microcentrifuge at top speed for 30 mins at 4 °C, washed with 70 % EtOH, centrifuged again for 15 mins, dried, and resuspended in 1 x stop buffer. FAM-dNTP background can be seen at bottom of the gel for the 'no clean-up' samples, but not in any of the DNA precipitation sample lanes. 'C' indicates the no enzyme control. Nucleotide (nt) size markers of 50 and 29 nt are shown on either side of the image.

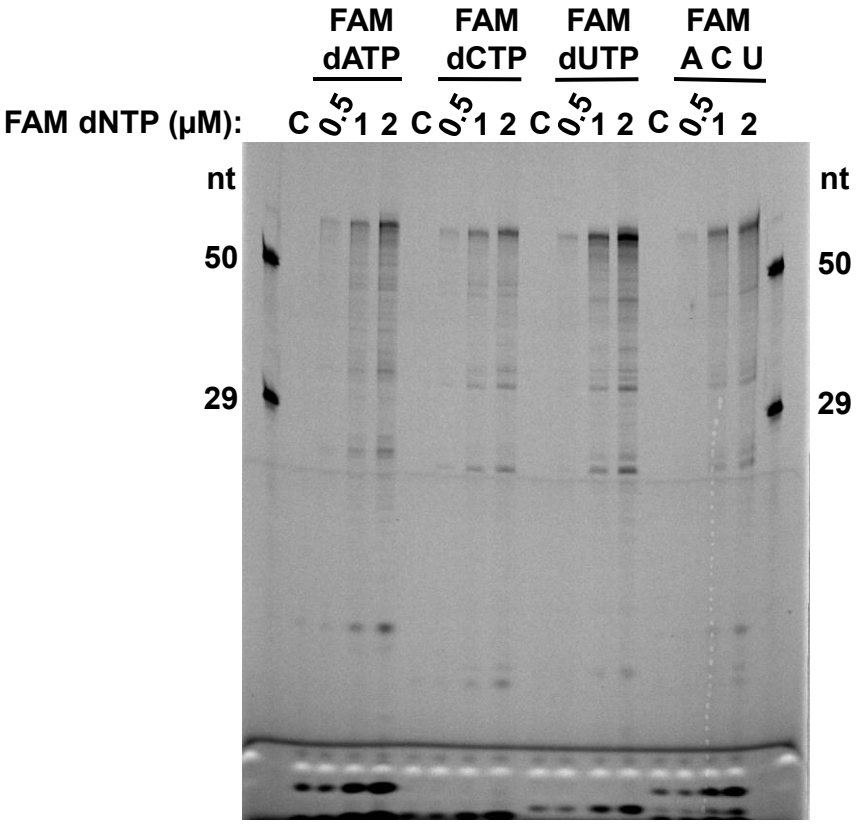


Figure 4.8. PrimPol can incorporate 6-FAM dATP, dCTP, and dUTP.
Reactions were assembled containing 250 nM PrimPol and either 6-FAM dATP, dCTP, dUTP, or all three, at increasing concentrations (0.5, 1, and 2 μM), and performed as outlined in 'experimental procedure' for a single 10 min time point. 'C' indicates the no enzyme control. Nucleotide (nt) size markers of 50 and 29 nt are shown on either side of the image.

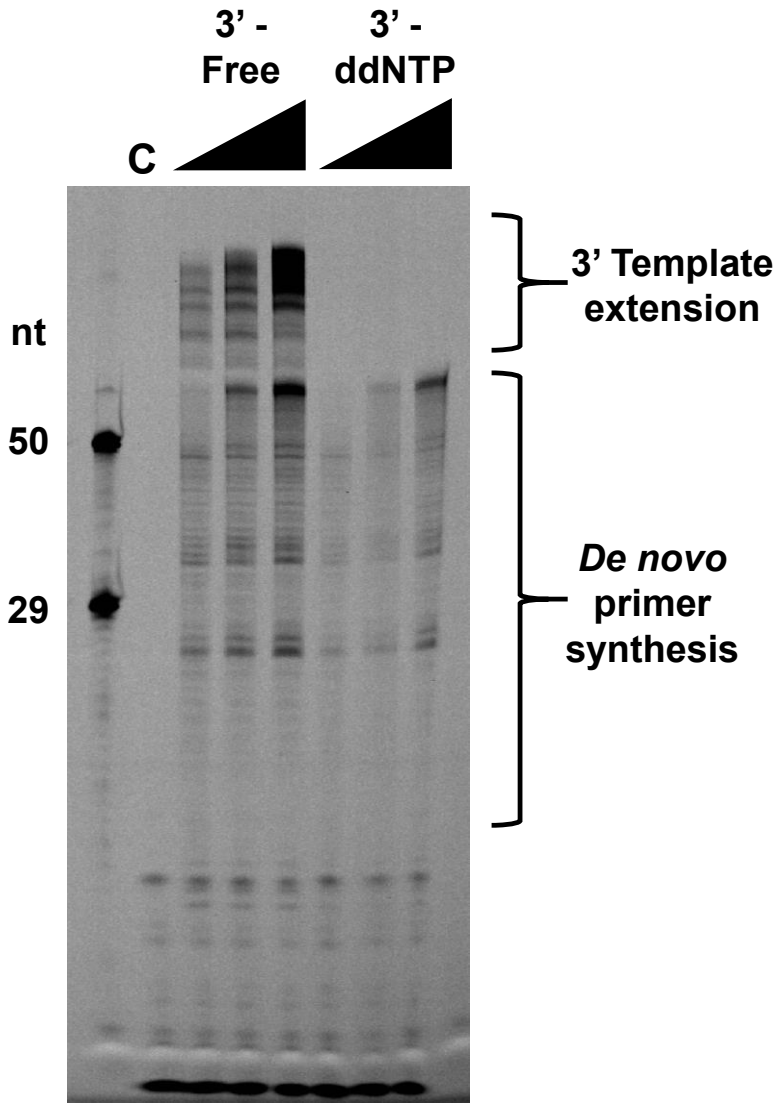


Figure 4.9. Comparison of the fluorescence gel-based primase assay on a linear ssDNA template with either a free 3' end, or a 3' end blocked with a dideoxynucleotide.
Assays were performed as detailed in 'experimental procedure' containing 250 nM PrimPol, and a 66 nt template with a free 3' end or containing a dideoxynucleotide at the 3' end (sequence in 'buffers and reagents'). When the 3' end is free, reaction products larger than the template are observed ('3'-free' sample lanes), however, addition of a 3' dideoxynucleotide removes these products ('3'-ddNTP' sample lanes), suggesting they are produced by extension of the 3' end of the template due to template snap-back. The black triangle indicates increasing time points of 5, 10, and 15 mins. 'C' indicates the no enzyme control. Nucleotide (nt) size markers of 50 and 29 nt are shown on the left side of the image.

Alternatively, a circular ssDNA template may be used. Although often overlooked, 3' end snap-back is a general problem when analysing primase activity using linear ssDNA templates and was briefly discussed by Koepsell *et al.* in their denaturing HPLC primase assay method (Koepsell et al., 2004).

Lastly, this assay is most applicable for the analysis of primase-polymerases, such as PrimPol, which are able to synthesise and extend their own primers. When using linear ssDNA templates < 100 nt, these enzymes can perform extension up to the end of the template. By analysing these reaction products, taking into account the length of the oligonucleotide product and template, the primase initiation site can be determined. This is particularly useful when assessing activities such as repriming (Kobayashi et al., 2016; Schiavone et al., 2016). Very short primase reaction products, which aren't extended after the initial synthesis, may be difficult to distinguish from the background given by 6-FAM dNTPs without any additional clean-up steps. However, the short nature of these primers may make them more liable to being washed away during clean-up due to their less stable duplex formation with the template DNA. Coupled with this, a higher concentration of 6-FAM dNTPs may need to be used in assays where the product size is very small, in order to increase the probability of the primase incorporating the 6-FAM dNTP into, and therefore labelling, its primer. Alternatively, a single native dNTP may be omitted and replaced with the equivalent 6-FAM dNTP. This problem can also be overcome by coupling the primase with a processive polymerase, as has been described previously in both fluorescent and radioactive assays (Bianchi et al., 2013; Galal et al., 2012; Keen et al., 2014a, 2014b). Here, primers synthesised by the primase of interest are extended by the polymerase, producing larger and more easily distinguishable reaction products.

4.2.5.5. Advantages and Limitations

This assay is intended to be used in place of the classic radioactive gel-based primase assay, which is still widely used in the identification and characterisation of primases. By utilising fluorescence, rather than radioactivity, the assay has a number of major advantages. Firstly, the potential hazards, rigorous safety measures, training, and cost of waste disposal, associated with handling radioactivity are avoided. Secondly, substitution of radioactivity with fluorescence permits more accurate quantification of reaction products due to the improved linear dynamic range. Radioactive dNTPs also have a short half-life, for example in the case of [α-^{32}P]dATP and [α-^{32}P]dCTP this is only 14 days. In contrast, 6-FAM dNTPs can be stored at -20 °C for up to 1 year before performance decreases. Consequently, despite their initial cost, fluorescent dNTPs can

be a more affordable option if assays are to be performed over a long period of time. Additionally, fluorescent gels can be immediately and rapidly scanned (~10 mins scanning time) following electrophoresis. This avoids the lengthy phosphor screen exposure times required for radioactivity detection, which can often take up to 12 hrs. Coupled with this, the gel-based nature of the assay allows a large number of samples (up to 20 per gel) to be resolved and imaged at the same time, taking in total 2-3 hrs. This potentially makes the technique faster than alternatives, such as the denaturing HPLC primase assay (20 min run time per sample), if a large number of samples are to be analysed.

Another major advantage of the assay is its similar set-up and readout to the traditional radioactive method. This allows the technique to be easily adopted by laboratories used to performing gel-based radioactive primase assays without extensive alterations to the method and equipment, or additional training. Indeed, despite alternatives, such as denaturing HPLC, gel-based primase assays are still most commonly used due to the ease of interpretation and lack of requirement for extensive optimisation when changing templates and enzymes.

Although possessing a number of advantages over similar qualitative primase assays, the fluorescence-based primase assay also shares some drawbacks with these techniques. Most notably, similar to the gel-based radioactive and HPLC primase assays, the fluorescence primase assay is not yet amenable to HTS. However, gel-based radioactive assays have previously been used to confirm inhibitors identified from large HTS methods, such as the primase-pyrophosphatase activity assay (Biswas et al., 2012). In these instances, fluorescence could be used to replace radioactivity, in order to confirm hits from HTS, and thus make the approach completely non-radioactive. Neverthless, it must be noted that the fluorescence does not provide the same level of sensitivity as radioactivity, requiring micromolar, in comparison to nanomolar, concentrations of labelled dTNPs. Lastly, it is possible that some primases may not tolerate the FAM-labelled nucleotides and we have yet to test the assay with FAM-rNTPs.

4.2.6. Summary and Conclusion

Since the identification of the first primases in the 1970s, characterisation of these enzymes has largely relied upon radioactive gel-based methods. Despite possessing excellent sensitivity and generating valuable qualitative primer synthesis information, these assays have major disadvantages, primarily due to their use of radioactivity. Consequently, in the last two decades a number of alternative non-radioactive primase assays have been developed. In most cases, these techniques have focused on enabling

HTS of potential primase inhibitor compounds. Although these assays suit this purpose well, they generally lack the capability to provide qualitative information about reaction products and often require large-scale optimisation prior to being performed. As a consequence, in spite of its time-consuming and hazardous nature, the gel-based radioactive primase assay remains the go-to option in the identification and basic characterisation of primases.

In this chapter, we have described how fluorescence provides a reliable alternative to radioactivity in the traditional gel-based primase assay, without requiring significant changes to the procedure or set-up. By replacing radioactive dNTPs with 6-FAM labelled dNTPs, all the disadvantages associated with radioactive work are eliminated. Furthermore, this substitution also offers clear advantages in speed over the traditional technique. We have used this fluorescent primase assay in published studies of the recently discovered eukaryotic primase-polymerase, PrimPol, thereby highlighting the general applicability of this technique in primase characterisation. This assay can be used in place of radioactive techniques to characterise basic primase activity, identify initiation sites, assess the impact of binding partners and accessory proteins, determine the effect of different reaction conditions, and to confirm primase inhibitor compounds identified through HTS. In summary, the fluorescent gel-based primase assay described here offers a safer and faster alternative to the classic, but still widely used, radioactive assay.

4.3. A Brief Investigation into the Mechanism Permitting Extension of Free 3’ Template Termini by PrimPol

4.3.1. Introduction

During the development of the fluorescent primase assay, it was observed that PrimPol produced reaction products longer than the ssDNA template (Figure 4.9.). These products could be eliminated by blocking the 3’ end of the template with a dideoxynucleotide. Thus, it was clear that these larger products were being generated by elongation of the template rather than through *de novo* primer synthesis. However, there exist a number of possible mechanisms by which PrimPol could facilitate this extension. Firstly, PrimPol could extend the 3’ termini in a template independent fashion. Secondly, PrimPol could promote synapsis and extension of the template by annealing it to another template molecule, which may be suggestive of an end-joining-like role. Lastly, the 3’ termini of the template may “snap-back” and anneal with itself, potentially mediated by PrimPol, consequently producing a pseudo primer-template for extension

by the enzyme. Intriguingly, it has previously been reported that PrimPol has a capacity to promote template synapsis and template-independent extension (Keen et al., 2014b; Martínez-Jiménez et al., 2015). However, these activities were only observed in the presence of manganese (Mn^{2+}) as a cofactor and the possibility that they were produced by template snap-back was not considered. Nevertheless, a number of reports have shown that bypass of some lesions, including 6-4PPs and Ap sites, by PrimPol, is achieved through a pseudo-TLS mechanism, although again these experiment were performed in the presence of Mn^{2+}. In this mechanism, PrimPol realigns the primer downstream of the lesion, mediated by microhomologies in the template strand. This effectively loops out the lesion to permit further extension (García-Gómez et al., 2013; Mourón et al., 2013). These observations are consistent with the preference of the enzyme to generate deletion mutations (Guilliam et al., 2015a). Therefore, it is not inconceivable that this ability to promote primer realignment might also allow PrimPol to perform template synapsis and extension. In this section, the mechanism by which PrimPol extends template 3'-ends is investigated and discussed.

4.3.2. Materials and Methods

4.3.2.1. Terminal Transferase Assay

PrimPol and Pol η were expressed and purified as previously described (Biertümpfel et al., 2011; Keen et al., 2014b). The terminal transferase assay was performed in the same manner as a standard primer extension assay, however a 5'-FAM labelled ssDNA template (sequence 1, Table 4.1.) or dsDNA template (sequences 1 and 2, Table 4.1.) was provided for extension. Reactions were assembled containing 100 nM PrimPol or Pol η, 20 nM template DNA, 100 μM dNTPs (Roche), 10 mM $MgCl_2$, 10 mM Bis-Tris-Propane-HCl (pH 7.0), and 1 mM DTT. Assembled reactions were incubated at 37ºC for a time course of 1, 3, 5, 10, and 20 minutes before quenching with 2x stop buffer (95% formamide, 20 mM EDTA, 200 nM competitor oligonucleotide) and heated to 95 ºC before resolution on a 15% (v/v) polyacrylamide gel containing 7 M urea in 1x TBE buffer. Reaction products were visualised using a Fujifilm FLA-5100 image reader.

4.3.2.2. MMEJ Assays

MMEJ assays were performed as described previously (Kent et al., 2015). Reactions were assembled containing 100 nM of 5'-FAM labelled ssDNA or partially ssDNA (pssDNA) templates with varying degrees of 3' homology (sequences 3-7, Table 4.1.), 100 nM polymerase, and 500 μM dNTPs, in primase assay buffer (10 mM Bis-Tris-Propane-HCl (pH 7.0), 10 mM $MgCl_2$, 1 mM DTT, with or without 1 mM $MnCl_2$) or the buffer used by Kent et al. (25 mM Tris-HCl pH 8.8, 10% glycerol, 1 mM DTT, 0.01% NP-

#	Oligonucleotide	Modification	Sequence (5'-3')
1	Terminal Transferase 1	5'-FAM	CATATCCGTGTCGCCCCTTATTCCGATAGTGACTACA
2	Terminal Transferase 2	N/A	GTATAGGCACAGCGGGGAATAAGGCTATCACTGATGT
3	MMEJ DNA-6	5'-FAM	CACTGTGAGCTTAGGGTTAGCCCGGG
4	MMEJ DNA-4	5'-FAM	CACTGTGAGCTTAGGGTTAGCCGG
5	MMEJ DNA-2	5'-FAM	CACTGTGAGCTTAGGGTTAGCG
6	MMEJ DNA-0	5'-FAM	CACTGTGAGCTTAGGGTTAGATAC
7	MMEJ Comp 14 mer	5'-Phosphate	CTAAGCTCACAGTG

Table 4.1. Terminal transferase and MMEJ oligonucleotides.
Sequences of oligonucleotides used for terminal transferase and micohomology-mediated end-joining assays. Underlined sequences indicate regions of 3' microhomology.

40, 10 mM $MgCl_2$, and 0.1 mg/ml BSA). Experiments were performed and samples processed as described by Kent et al. (Kent et al., 2015), before resolution on a 15% nondenaturing or denaturing (urea) polymacrylamide gel. Reaction products were visualised using a Fujifilm FLA-5100 image reader.

4.3.3. PrimPol can Extend Single-Stranded but not Double-Stranded DNA Templates

AEPs involved in NHEJ possess a limited ability to perform template-independent extension of both ssDNA and blunt dsDNA ends (Della et al., 2004; Pitcher et al., 2005). Previously, in the presence of high concentrations of Mn^{2+}, PrimPol was demonstrated to be capable of template-independent extension of ssDNA but not dsDNA ends (Keen et al., 2014b). As an initial investigation in to the mechanism of 3'-end template extension by PrimPol, this ability was analysed in the same buffer as that used in the fluorescent primase assay. To this end, a standard primer extension assay was performed using a 5'-FAM labelled 37-mer template, either with (dsDNA) or without (ssDNA) an unlabelled complementary 37-mer strand (sequences 1 and 2, table 4.1.). Pol η, which is not capable of template-independent extension and was previously reported to be unable to perform template synapsis, was also analysed for comparison (Kent et al., 2015).

Both PrimPol and Pol η were unable to extend the blunt dsDNA template, confirming the absence of any terminal transferase-like activity in these conditions (Figure 4.10.). However, both enzymes showed some capacity to extend the ssDNA template, with notable differences in the reaction products produced. In the case of Pol η, a single predominant product of ~8 nt in length was observed. However, with PrimPol, both an 8 nt product and a ~29 nt product were generated (Figure 4.10.). Analysis of the DNA template sequence reveals that the 8 nt product could be produced by snap-back of the dA and dC nucleotides, present at positions 37 and 36, to the dT and dG nucleotides at positions 9 and 10 on the ssDNA template strand, respectively. Thus, extension of this snap-backed 3' end would produce an 8 nt product. Given the distance between the 3' end of the template and the two complementary nucleotides at positions 9 and 10, it is unlikely that this product would be produced by polymerase-mediated end synapsis of two template strands. The ~29 nt product, however, could be produced by either snap-back or synapsis and extension. Here, the dA and dC at the very 3' end of the template could base-pair with the dT and dG nucleotides at positions 30 and 31 on the template strand, respectively (Figure 4.10.). Given the short distance between these sets of nucleotides, it is not inconceivable that this base-pairing was mediated in an end-synapsis fashion. Thus, the appearance of the additional 29 nt product in the PrimPol

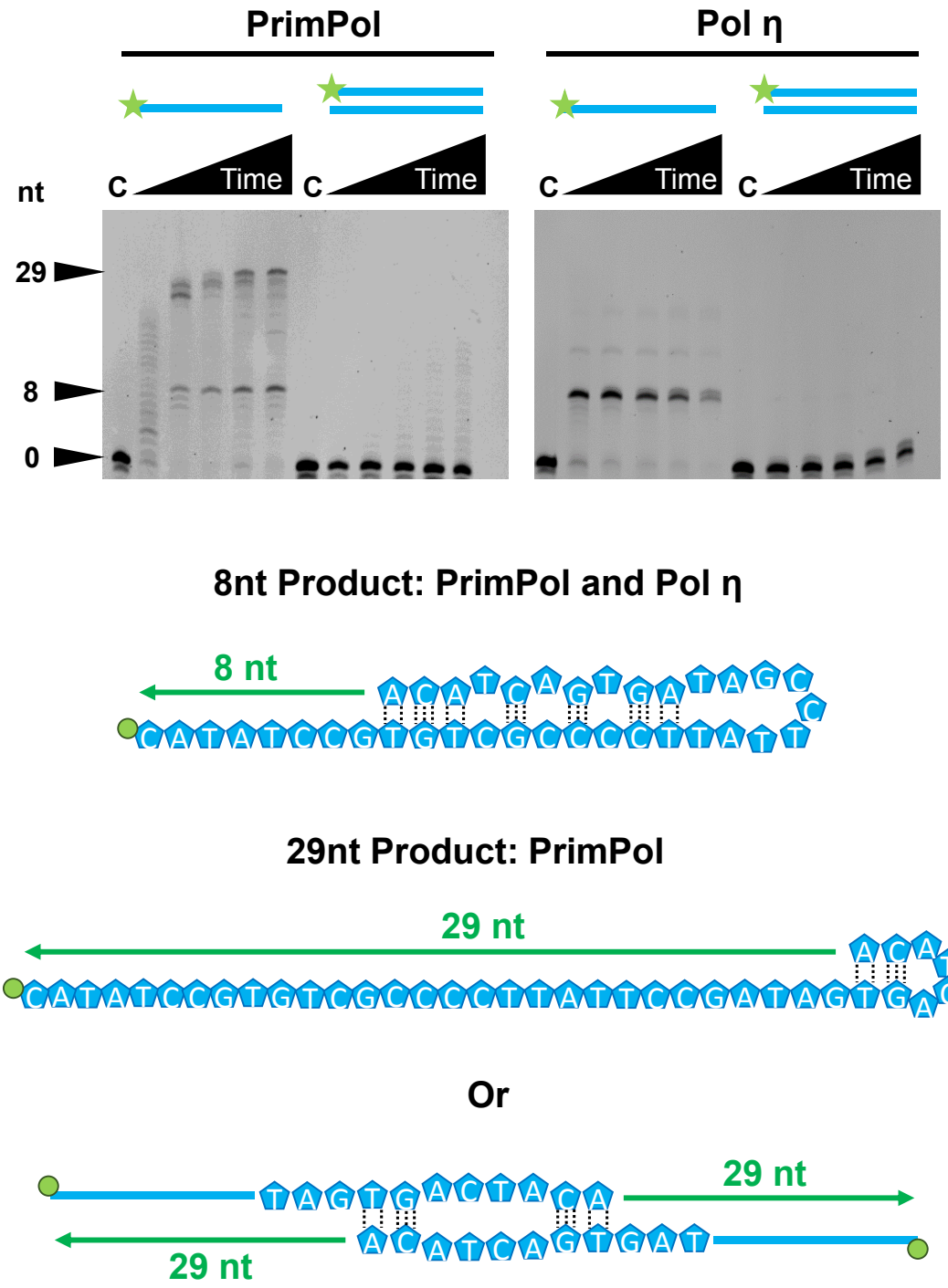


Figure 4.10. PrimPol extends ssDNA but not dsDNA templates.
PrimPol and Pol η were incubated in standard primer extension assay conditions containing either ssDNA templates or blunt-ended dsDNA templates for increasing time-points. Both PrimPol and Pol η were unable to extend blunt dsDNA templates. However, each enzyme generated reaction products by extending ssDNA templates. PrimPol and Pol η produced a predominant product of 8 nt, with PrimPol generating a second 29 nt product. The potential template configurations permitting generation of these products are shown in the schematic below. Here, the 8 nt product is likely generated by snap-back and extension of the 3'-end of the ssDNA template. The 29 nt product could be produced by either snap-back or MMEJ.

but not Pol η reactions, could be due to PrimPol-mediated synapsis of two ssDNA template molecules, followed by extension.

4.3.4. Can PrimPol Perform Microhomology-Mediated End-Joining?

The results obtained from the initial experiments described above were potentially consistent with PrimPol harbouring an ability to promote end-synapsis and extension of ssDNA templates based on microhomologies within the template sequence. Over a decade since the discovery of the NHEJ repair polymerase PolDom in prokaryotes (Della et al., 2004; Weller et al., 2002), evidence has emerged which is suggestive of a role for Pol θ in an analogous microhomology-mediated end joining (MMEJ) pathway in eukaryotes (Chan et al., 2010; Koole et al., 2014; Roerink et al., 2014; Yousefzadeh et al., 2014). MMEJ is an alternative pathway for the repair of DSBs, here limited resection of DNA ends generates a 3' ssDNA overhang and exposes microhomologies on each side of the break. PolDom or Pol θ dimers are then able to facilitate DNA synapse formation by promoting the annealing of these microhomologies, before extension and strand displacement (Della et al., 2004; Kent et al., 2015; Weller et al., 2002). The ability of Pol θ to perform this activity was recently characterised *in vitro* by analysing its capacity to promote MMEJ of DNA templates with either 6 (DNA-6), 4 (DNA-4), 2 (DNA-2), or 0 (DNA-0), base-pairs of microhomology at their 3' end (Kent et al., 2015). Experiments were performed using these templates alone (ssDNA) or with a 14 nt complementary strand, with a 5'-terminal phosphate, annealed to 3' end of the template, thus leaving a 3' ssDNA overhang (pssDNA). Analysis of reaction products on non-denaturing polyacrylamide gels revealed that Pol θ was able to efficiently promote synapsis of the pssDNA templates containing 6, 4, and 2, but not 0, base-pairs of microhomology. Synapsis of ssDNA templates without the 14 nt complementary strand was inefficient, with the major product being snap-back extension.

In order to examine the ability of PrimPol to perform MMEJ, we performed the same assay used in the studies of Pol θ, on identical DNA templates (sequences 3-7, Table 4.1.). Reactions were performed in both the buffer used for the fluorescent-primase assay (Figure 4.11.A.), and the buffer used in the Pol θ report (Figure 4.11.B.). The assay was also performed in reactions containing a mix of Mg^{2+} and Mn^{2+} (Figure 4.11.C.). Pol η was used for comparison (Figure 4.12.), in the same conditions as those in the Pol θ report, where it was unable to promote synapsis of templates containing 4 base-pairs of microhomology (Kent et al., 2015).

In each case, PrimPol displayed a limited ability to perform MMEJ of templates containing either 6 or 4 nt of homology at their 3' end (Figure 4.11.). A slight increase in

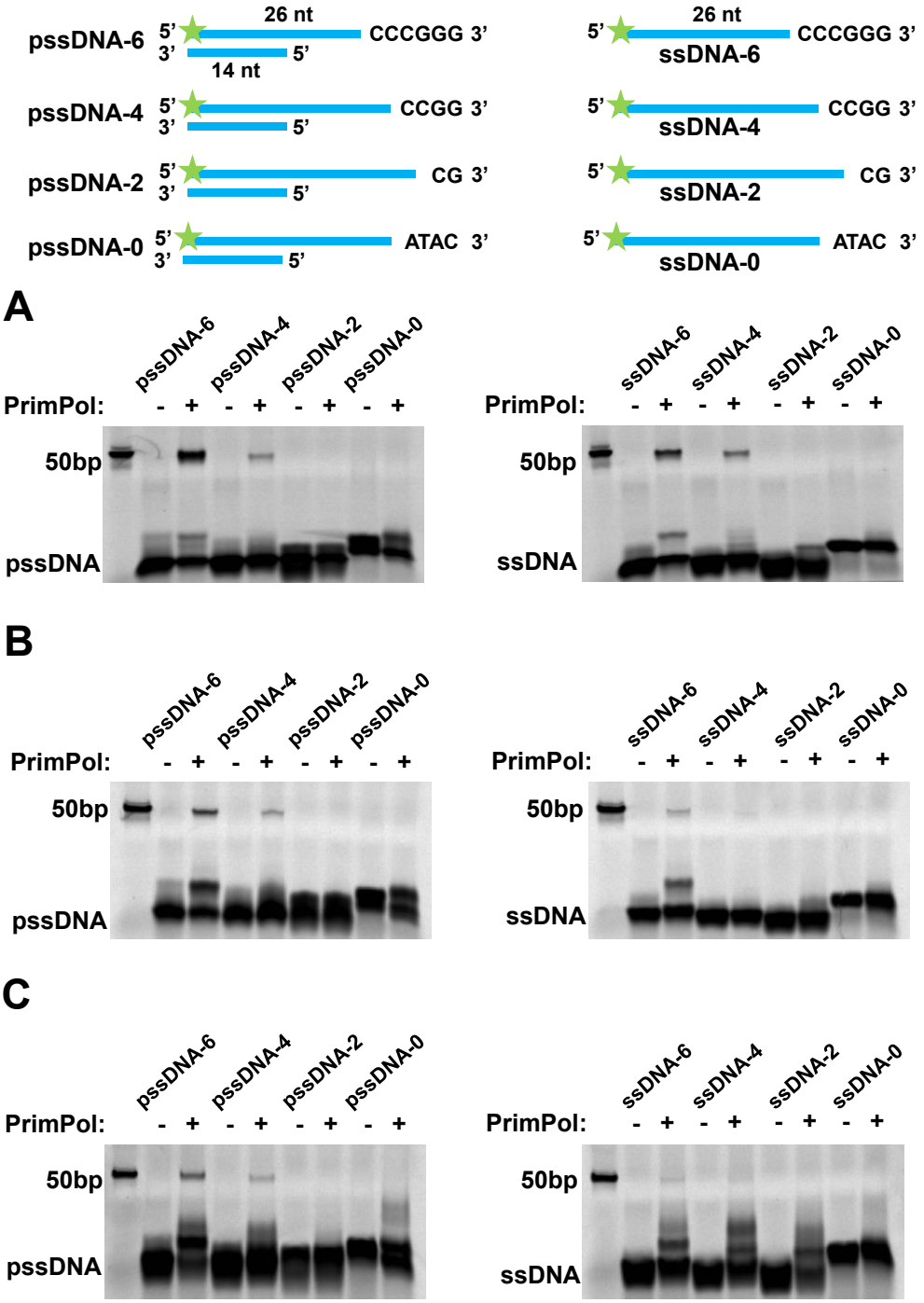


Figure 4.11. Analysis of PrimPol's ability to perform MMEJ.
Non denaturing gels analysing PrimPol's ability to promote MMEJ of templates with varying lengths of 3' microhomology either single-stranded (ssDNA) or annealed to a 14 nt complementary strand (pssDNA), as shown above. **(A)** PrimPol MMEJ assays performed in primase assay buffer (10 mM Bis-Tris-Propane-HCl (pH 7.0), 10 mM $MgCl_2$, and 1 mM DTT). **(B)** PrimPol MMEJ assays performed in the buffer used by Kent et al. (2015) (25 mM Tris-HCl pH 8.8, 10% glycerol, 1 mM DTT, 0.01% NP-40, 10 mM $MgCl_2$, and 0.1 mg/ml BSA). **(C)** PrimPol MMEJ assays performed in primase assay buffer containing 1 mM $MnCl_2$. '50bp' indicates the migration position of a ds50-mer. PrimPol displayed a limited ability to promote MMEJ of templates containing 6 and 4 nt of 3' microhomology.

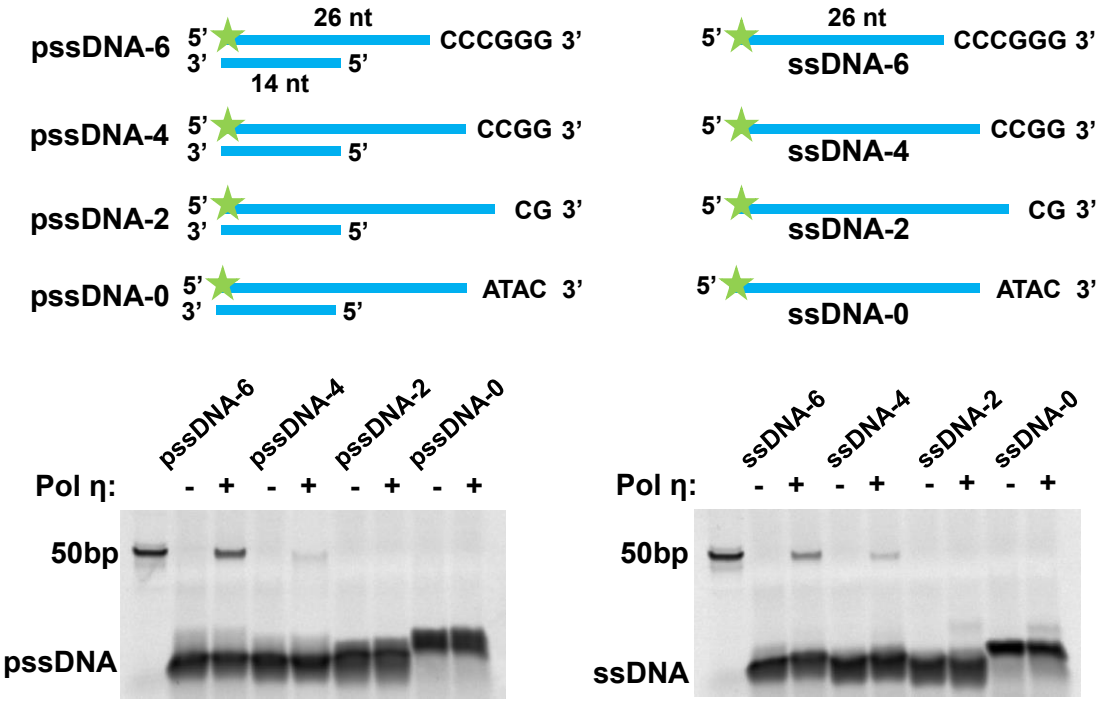


Figure 4.12. Analysis of Pol η's ability to perform MMEJ.
Non denaturing gels analysing Pol η's ability to promote MMEJ of templates with varying lengths of 3' microhomology either single-stranded (ssDNA) or annealed to a 14 nt complementary strand (pssDNA), as shown above. Reactions were performed in primase assay buffer (10 mM Bis-Tris-Propane-HCl (pH 7.0), 10 mM $MgCl_2$, and 1 mM DTT). Pol η showed an ability to promote MMEJ of of templates containing 6 and 4 nt of 3' microhomology.

this activity was apparent on pssDNA templates. No significant difference was observed when using primase assay buffer (Figure 4.11.A.) in comparison to the buffer used by Kent *et al*. (Figure 4.11.B.), additionally, no increase in MMEJ products was seen when using buffer containing Mn^{2+} (Figure 4.11.C.). Despite this limited MMEJ activity, Pol η produced near identical results to PrimPol, with reaction products generated when using templates with 6 or 4 nt of homology (Figure 4.12.). Therefore, PrimPol does not appear to be proficient at MMEJ at a level above that observed with other lesion bypass polymerases. Moreover, a significant amount of snap-back products were observed in these reactions, indicated by a slight shift in the template band but much less than that observed for MMEJ products, as was previously determined (Kent et al., 2015).

4.3.5. PrimPol is Proficient at Snap-Back Synthesis

To better analyse the overall activity of PrimPol in the reactions described above, products were resolved on denaturing urea-polyacrylamide gels. Here, PrimPol displayed a significant level of activity on all templates except pssDNA-2 (Figure 4.13.). The lack of activity on this particular template is likely a result of the complementary strand blocking snap-back of the 3' end. Notably, in the absence of the complementary strand, significant activity was observed (template ssDNA-2). Furthermore, a significant level of extension was apparent on templates ssDNA-0 and pssDNA-0, which display no microhomology and which didn't produce any MMEJ products, therefore confirming that this activity is a result of snap-back extension (Figure 4.13.). The predominant sizes of the reaction products in each case are also consistent with potential snap-back locations on each template. Interestingly, the size of the products produced on pssDNA templates is not significantly limited by the complementary strand, suggesting that PrimPol may possess some strand-displacement activity.

In summary, it is likely that the template extension products observed in the primase assay are a result of 3' end snap-back and extension by PrimPol. Although the enzyme appears to display some MMEJ activity, the level of this activity is not significantly higher than that observed with Pol η. Furthermore, there is currently a lack of *in vivo* evidence supporting a role for PrimPol in MMEJ or DSB repair. Wan *et al*., observed the formation of PrimPol foci in response to ionising radiation (IR), however another report failed to observe the same effect (Bianchi et al., 2013; Wan et al., 2013). Nevertheless, both reports agree that $PrimPol^{-/-}$ cells are not sensitive to IR (Bianchi et al., 2013; Wan et al., 2013). The experiments presented here do, however, reveal that PrimPol is proficient at snap-back extension and exhibits some strand-displacement activity. It is currently unclear what the *in vivo* relevance of these activities are and it remains possible that they

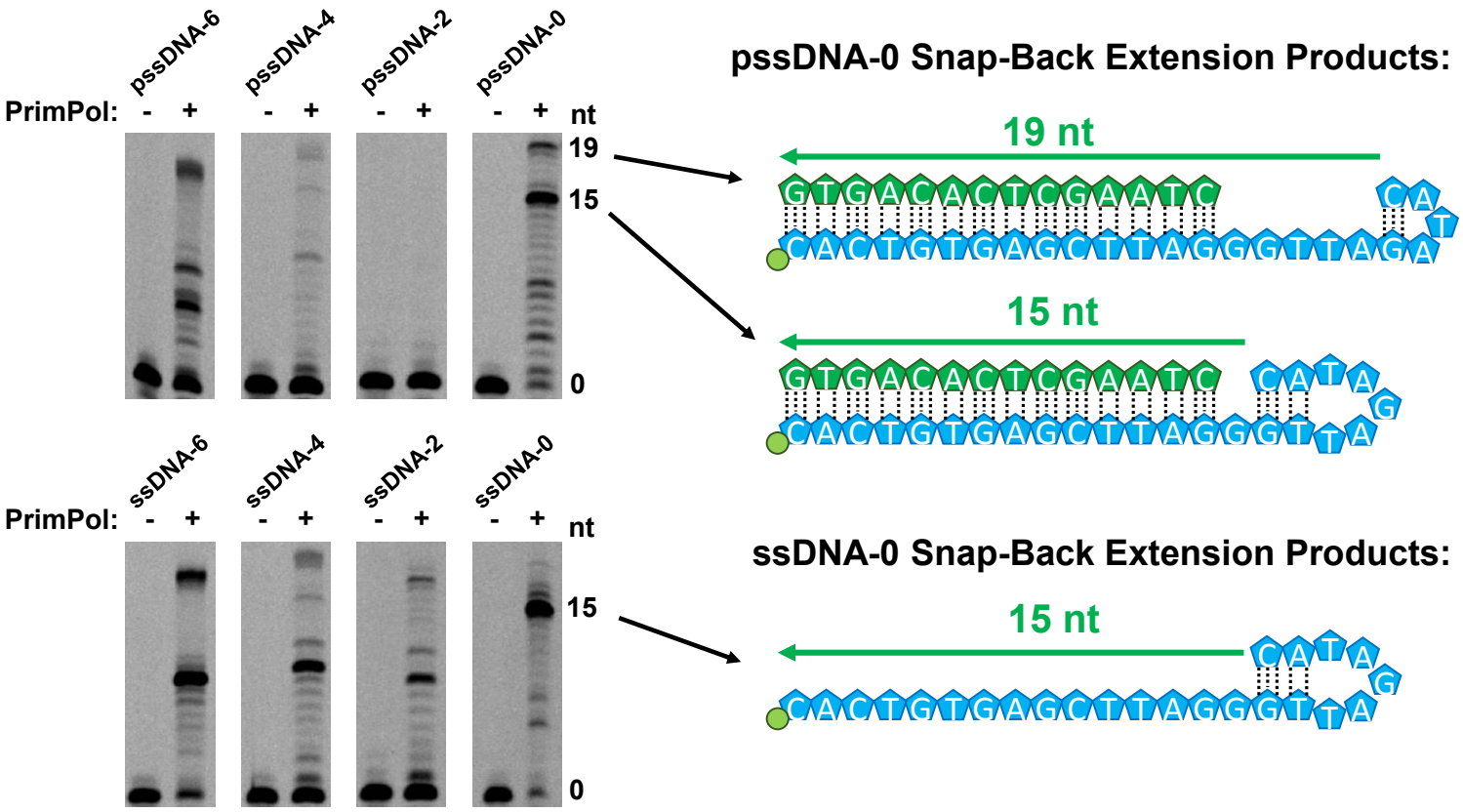


Figure 4.13. Denaturing gel analysis of PrimPol MMEJ reaction products.
Denaturing gel analysis of PrimPol MMEJ reaction products shown in Figure 4.11.A. Reaction products generated on templates lacking 3' micohomology (pssDNA-0 and ssDNA-0) due to template 3'-end snap-back are indicated. The intra-template homology permitting the synthesis of these products is displayed in the schematic to the right of the gels. The results reveal that PrimPol is proficient at template snap-back extension.

are simply the by-product of the accommodating catalytic requirements of priming. However, Pol θ has also been shown to efficiently perform snap-back synthesis, termed templated-extension *in cis* by the relevant reports, which was suggested to support a role in end-joining, although the usefulness of this activity *in vivo* is not immediately obvious (Black et al., 2016; Kent et al., 2016). At the least, PrimPol's propensity to perform snap-back synthesis and to some degree strand-displacement and MMEJ, highlights the versatility of the enzyme and opens up the possibility of additional, as yet unknown, roles.

4.4. PrimPol Bypasses Non-Canonical Replication Impediments by Repriming Downstream

4.4.1. Introduction

The identification of a role for PrimPol in the tolerance of DNA damage lesions generated interest in the potential of the enzyme to bypass other non-canonical replication impediments. In collaboration with the Sale and Hirota laboratories, the role of PrimPol in bypassing G-quadruplex structures and CTNAs, respectively, was investigated. The remainder of Chapter 4 outlines the *in vivo* findings of our collaborators and describe the complementary *in vitro* investigations performed. These *in vitro* studies, permitted by the development of the fluorescent primase assay, support the *in vivo* results suggesting that PrimPol reprimes, and consequently, restarts replication downstream of G-quadruplexes and CTNAs, in addition to more canonical DNA damage lesions.

4.4.2. PrimPol is Required for Replicative Tolerance of G-Quadruplexes in Vertebrate Cells

G-quadruplex structures, formed by the stacking of quartets of Hoogsteen-bonded sequences of guanines, present a potent obstacle to replisome progression. Indeed, *in vitro* replicative polymerases are stalled by these structures (Woodford et al., 1994). A number of specialised helicases including, Fanconi Anemia Group J Protein (FANCJ), ATP-dependent DNA helicase PIF1, Bloom syndrome helicase (BLM), and Werner syndrome helicase (WRN), in addition to Pol η, Pol κ, and Rev1, have been implicated in enabling G-quadruplex replication (León-Ortiz et al., 2014; Wickramasinghe et al., 2015). The Sale laboratory previously developed an assay to monitor G-quadruplex replication in chicken DT40 cells (Sarkies et al., 2011; Schiavone et al., 2014). This assay measures the expression of the Bu-1a cell surface glycoprotein, the locus of which contains a G-quadruplex structure downstream of the transcription start site. When cells

lack replication machinery components required for G-quadruplex replication, the replisome stalls at the site of the structure, consequently leading to stochastic loss of Bu-1a expression following cell division. This loss of expression arises due to the uncoupling of DNA synthesis and histone recycling which causes a loss of parental histone modifications. Thus, loss of Bu-1a expression is as a result of epigenetic, rather than genetic, instability.

Using this assay, the Sale laboratory observed that PrimPol$^{-/-}$ DT40 cells exhibited three levels of Bu-1a expression, termed Bu-1a^{high}, Bu-1a^{medium}, and Bu-1a^{low} (Schiavone et al., 2016). In comparion, wild-type DT40 cells uniformally expressed Bu-1a at a high level. It was determined that the Bu-1a^{medium} and Bu-1a^{low} populations of PrimPol$^{-/-}$ cells arose stochastically from the Bu-1a^{high} population. The decrease in Bu-1a expression in the absence of PrimPol was found to be due to epigenetic instability at the *BU-1A* locus, importantly deletion of the G-quadruplex forming sequence stabilised Bu-1a expression and reintroduction caused instability again. Additionally, when reintroduced in a manner which would cause lagging-strand stalling, no instability was observed. These observations were consistent with a role for PrimPol in facilitating replisome bypass of the G-quadruplex structure on the leading strand. In light of these results, the potential mechanism employed by PrimPol to bypass G-quadruplexes was investigated biochemically.

4.4.2.3. Materials and Methods

4.4.2.3.1. Primer Extension Assays

PrimPol, Pol η, and *A. fulgidus* Pol B were purified as previously described (Biertümpfel et al., 2011; Jozwiakowski et al., 2014; Keen et al., 2014b). Primer extension assays were performed using 5' Hexachlorofluorescein labelled DNA primers annealed to their complementary DNA templates (sequences 1-13, Table 4.2.). Extensions were carried out at 37°C in 20 µL volumes, typically containing 10 mM Bis-Tris-Propane-HCl pH 7.0, 50 mM KCl (except where otherwise stated), 10 mM $MgCl_2$, 1 mM DTT, 20 nM primer-template substrate, 100 µM dNTPs (NEB), 100 µg/mL BSA (NEB), and 100 nM of the assayed polymerase. Extension reactions were monitored over a time course and quenched with 20µL stop buffer (200 nM competitor oligonucleotide in a 95% formamide solution with 0.25% bromophenol blue and xylene cyanol). Products were boiled at 95°C for 5 min before resolution on a 15% (v/v) polyacrylamide/7M urea gel. Gels were scanned using an FLA-5100 image reader (Fujifilm).

#	Oligonucleotide	Modification	Sequence (5'-3')
1	HP-20 Primer	5'-HEX	TGTCGTCTGTTCGGTCGTTC
2	HIV Integrase Template	3'-Biotin	ACCGCGAACTTGAATTCTATGGGTGGGTGGGTGGGTCAATGCACAACATATGGC TTTCGAAGACCGAACGACCGAACAGACGACA
3	G4#1 Template	3'-Biotin	ACCGCGAACTTGAATTCTATTGGTTTTGGTTTTGGTTTTGGTCAATGCACAACA TATGGCTTTCGAAGACCGAACGACCGAACAGACGACA
4	G4#2 Template	3'-Biotin	ACCGCGAACTTGAATTCTATGGGTTTGGGTTTGGGTTTGGGTCAATGCACAACA TATGGCTTTCGAAGACCGAACGACCGAACAGACGACA
5	G4#3 Template	3'-Biotin	ACCGCGAACTTGAATTCTATGGGGTTGGGGTTGGGGTTGGGGCAATGCACAACA TATGGCTTTCGAAGACCGAACGACCGAACAGACGACA
6	G4#4 Template	3'-Biotin	ACCGCGAACTTGAATTCTATTTTGGGTGGGTGGGTGGGTTTTCAATGCACAACA TATGGCTTTCGAAGACCGAACGACCGAACAGACGACA
7	ND-97 Template	3'-Biotin	ACCGCGAACTTGAATTCTAGTTCAGTCTAAATGCTCTCAAGCACTGAGCAATTC ACAACATATGGCTTTCGATTACCGAACGACCGAACAGACGACA
8	P-Globin Template	3'-Biotin	ACCGCGAACTTGAATTCTAGGGGAGTAAAAGGGAGCGGGGTGCTGGGGCAATGC ACAACATATGGCTTTCGAAGACCGAACGACCGAACAGACGACA
9	Bu-1a Template	3'-Biotin	ACCGCGAACTTGAATTCTAGGGCTGGGTGGGTGCTGTCAAGGGCTGGGCAATGC ACAACATATGGCTTTCGAAGACCGAACGACCGAACAGACGACA
10	Mt-G4 Template	3'-Biotin	ACCGCGAACTTGAATTCTAGGGTGGGATGGGCGGGGCAATGCACAACATATGG CTTTCGAAGACCGAACGACCGAACAGACGACA
11	T4 Telomeric G4 Template	3'-Biotin	ACCGCGAACTTGAATTCTATTAGGGTTAGGGTTAGGGTTAGGGCAATGCACAAC ATATGGCTTTCGAAGACCGAACGACCGAACAGACGACA
12	MtOL Template	3'-Biotin	ACCGCGAACTTGAATTCTGCCGGGGCTTCTCCCGCCTTTTTTCCCGGCGGCGGG AGAAGTAGATTAACATATGGCTGAACGACCGAACAGACGACA
13	*T.Bu* Sb. Tel. Template	3'-Biotin	ACCGCGAACTTGAATTCAGCTTTCGGGTTAGGGTGTTTCGGGTTAGCGGGCAAA CAACATATGGCTTTCGAAGACCGAACGACCGAACAGACGACA

Table 4.2. G-quadruplex primer-template oligonucleotides.
Sequences of the primers and G-quadruplex containing templates used in primer-extension assays. Sequences predicted to form G-quadruplexes and DNA secondary structures are shown in red.

4.4.2.3.2. Primase Assay

For G-quadruplex repriming assays, 5' biotin-labelled DNA templates (sequences 1-4, Table 4.3.) was incubated at 95ºC in annealing buffer with 50 mM KCl and cooled to allow G4 structure formation and primer annealing. PrimPol (2 μM) was incubated with templates (1 μM) for 30 min at 37ºC in primase assay buffer (10 mM Bis-Tris-Propane-HCl [pH 7.0], 10 mM $MgCl_2$, 1 mM DTT, 250 μM dNTPs or rNTPs, and 2.5 μM FAM dNTPs). Reactions were quenched with Bindng-Washing (B-W) buffer (10 mM Tris-HCl [pH 7.5], 500 mM NaCl, 10 mM EDTA) and incubated with streptavidin beads (Roche) for 1 hr at 4ºC. Beads were washed with B-W buffer and suspended in stop buffer. Samples were heated to 95ºC and resolved on a 15% polyacrylamide / 7 M urea gel for 90 min. Products were visualized on an FLA-5100 imager. Lesion and CTNA repriming assays were performed in a similar manner using 1 μM PrimPol and a 15 min incubation time (sequences 4-9, Table 4.3.). Oligonuclotides containing Carbovir or Acyclovir were chemically synthesised as previously described (sequences 7 and 8, Table 4.3.) (Yamamoto et al., 2011).

4.4.2.3.3. Circular Dichroism Analysis of G-Quadruplexes

Circular dichroism (CD) spectra were measured using a Jasco spectropolarimeter. The HIV integrase quadruplex containing template (sequence 2 without biotin tag, Table 4.2.) was diluted to a final concentration of 1 μM in buffer containing 10mM Bis-Tris-Propane-HCl pH 7.0, 10 mM $MgCl_2$, and varying levels of KCl. CD spectra of oligomers were recorded in a 0.5 cm quartz cuvette from wavelengths of 320 nm to 220 nm with a 0.2 nm data pitch, a speed of 50 nm min^{-1}, and a band width of 1 nm. The average of 4 accumulations for each condition was recorded and the CD spectrum of the buffer alone was subtracted from samples containing DNA templates. In melting experiments the temperature was increased incrementally from 25°C to 90°C, the cuvette was incubated at each temperature for 10 mins before spectra were recorded.

4.4.2.4. Structure Stability and Loop Length Influence Polymerase Bypass of G-Quadruplexes

Given that PrimPol has both primase and TLS polymerase activities, it could potentially facilitate G-quadruplex bypass in two different ways, either directly through a TLS-like mechanism, or indirectly by repriming and restarting replication downstream of the structure. Initial investigations focussed on analysing PrimPol's ability to directly read through a G-quadruplex in a TLS-like manner.

#	Oligonucleotide	Modification	Sequence (5'-3')
1	G4 Repriming Primer	3'-ddG	GATGCATCTATGTddG
2	G4 Repriming Template 1	5'-Biotin	GTCTTCTATCTCGTCTATATTCTATTGTCTCTATGGGTGGGTGGGTGGGTCTCA CATAGATGCATC
3	G4 Repriming Template 2	5'-Biotin	GTCTTCTATCTCGTCTATATTCTATTGTCTCTATGGGTGGGTGGGTGGGTCTCA
4	Repriming Control Template	5'-Biotin	GTCTTCTATCTCGTCTATATTCTATTGTCTCTATGAATACCTTCATCAGTCTCA CATAGATGCATC
5	Ap Repriming Template	5'-Biotin	GTCTTCTATCTCGTCTATATTCTATTGTCTCTATGAATACCTTCATCAApTCTC ACATAGATGCATC
6	Tg Repriming Template	5'-Biotin	GTCTTCTATCTCGTCTATATTCTATTGTCTCTATGAATACCTTCATCATgTCTC ACATAGATGCATC
7	ACV primer	3'-ACV	TCCGTTGAAGCCTGCTTTACV
8	CBV primer	3'-CBV	TCCGTTGAAGCCTGCTTTCBV
9	ACV/CBV Repriming Template	5'-Biotin	GTCTTCTATCTCGTCTATATTCTATTGTCTCTATGAATACCTTCATCCAAAGCA GGCTTCAACGGA

Table 4.3. Template and primer sequences used for repriming assays.
Sequences of primers and templates used in G-quadruplex, CTNA, and lesion repriming assays. Modified nucleotides, lesions, and G-quadruplex forming sequences are shown in red.

Prior to analysing the ability of PrimPol to replicate through different G-quadruplex structures *in vitro*, it was first necessary to ensure the correct folding of these structures under the experimental conditions to be used. The HIV-integrase aptamer was chosen as a model for this purpose. This template has an 'ideal' quadruplex forming sequence of $d(GGGT)_4$ which previous studies have shown folds into a parallel quadruplex structure with a high level of stability (Kelley et al., 2011). To investigate the stability of this structure in different KCl concentrations and temperatures CD was employed. This analysis confirmed that the HIV-integrase template correctly folds into a parallel G-quadruplex structure *in vitro*. This is revealed by a maximum at ~260 nm and a minimum at ~240 nm, with the structure becoming less stable as the temperature is increased (Figure 4.14.A. and B.). In addition, the stability of the structure was dependent on the concentration of KCl present, with 50 mM KCl promoting a more stable structure than 5mM KCl (Figure 4.14.A. and B.). These results therefore confirmed that the G-quadruplex sequence correctly folds in to the expected structure in the context of the synthetic substrate and conditions used. This permitted further analysis of polymerase bypass of the structure.

PrimPol and Pol η were subject to primer extension assays using the HIV integrase template in the presence of a range of monovalent cations (50 mM K^+, Na^+, Li^+, or no salt). It has previously been reported that different monovalent cations affect quadruplex structure stability with K^+ being the most stabilising, whilst Li^+ has a destabilising effect (Hardin et al., 1992; Prakash et al., 2011). The ability of both PrimPol and Pol η to replicate through the HIV-integrase quadruplex was dependent upon on the monovalent cation present (Figure 4.14.C. and D.). In the presence of NaCl, LiCl, and when no salt was present, both Pol η and PrimPol were able to fully extend the primer with only slight pausing being observed at the G-quadruplex forming region. However, in the presence of KCl, full primer extension by both Pol η and PrimPol was inhibited, supporting previous reports of the ability of K^+ to stabilise the DNA secondary structure.

To further establish the ability of PrimPol to read through these structures, templates containing G-quadruplex-forming sequences with different loop lengths (Figure 4.15.A.) were used in primer extension assays. In addition to PrimPol (Figure 4.15.B.) and Pol η (Figure 4.15.C.), *A. fulgidus* Pol B (*A.fu* Pol B) (Figure 4.15.D.) was analysed for comparison of the ability of replicative polymerases to bypass these structures. The results of the assays revealed that all three polymerases exhibit a similar inability to read through the structures, indicating that PrimPol does not possess an increased ability to synthesise through these G-quadruplexes (Figure 4.15.). Here, loop length was correlated with bypass, with structures which possess longer loops being more easily

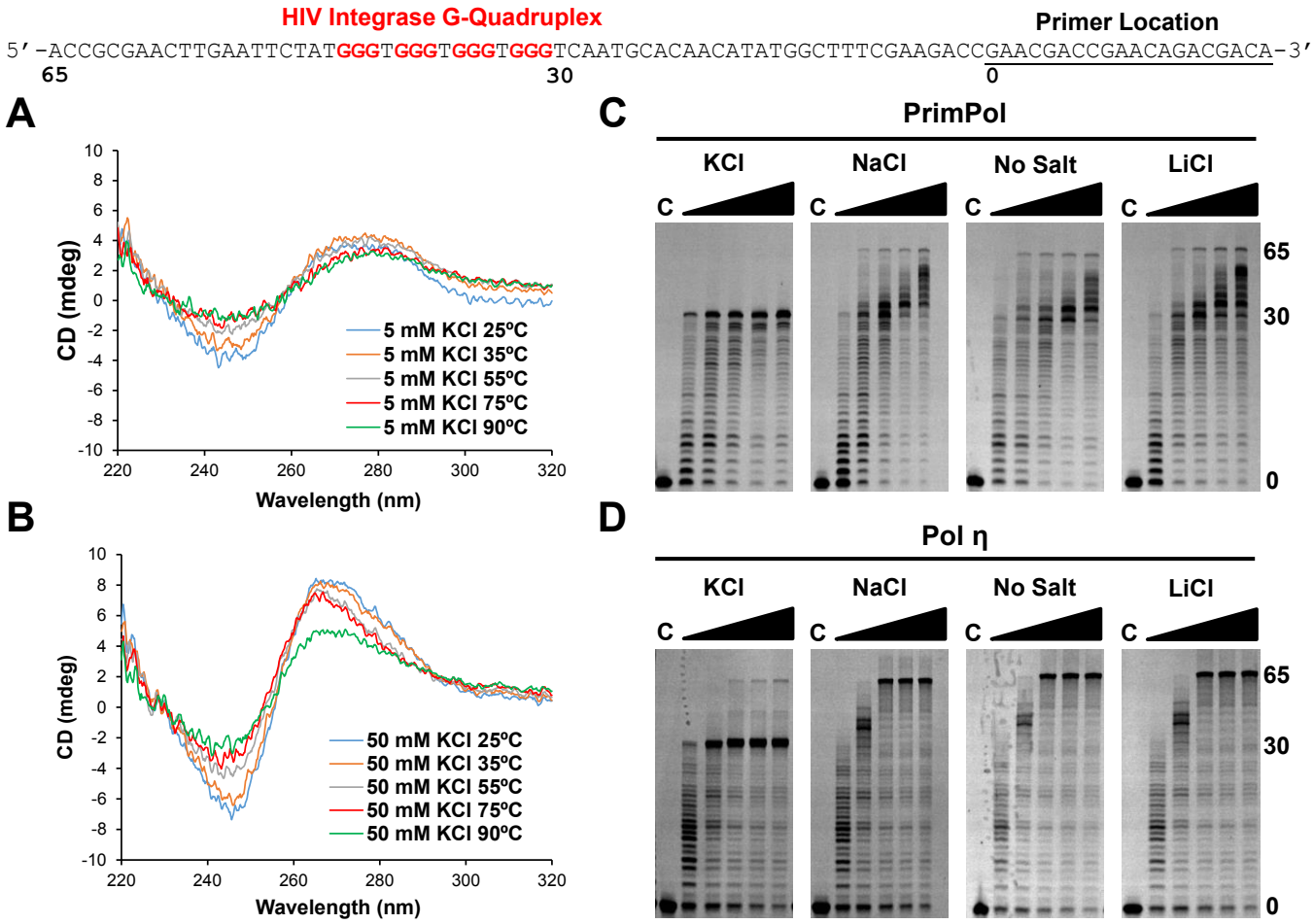


Figure 4.14. G-quadruplex structure stability governs DNA polymerase bypass ability.
(A/B) CD spectra of the HIV integrase G-quadruplex containing template in the presence of 5 mM **(A)** or 50 mM **(B)** KCl at increasing temperatures (25, 35, 55, 75, and 90°C). **(C/D)** Primer extension reactions performed using PrimPol **(C)** or Pol η **(D)** on the HIV-integrase G-quadruplex containing template in the presence of 50 mM KCl, NaCl, LiCl, or in the absence of salt. The sequence of the HIV integrase G-quadruplex containing template is shown above with the primer location and positions of stalling at the structure (30) and full-length extension (65) indicated.

A

	G-Quadruplex Sequence (5'-3')	Tm (°C)	Topology	Non-G Loop
G1	TTGGTTTTGGTTTTGG TTTTGGT	27.6	Antiparallel	4
G2	TGGGTTTGGGTTTGGG TTTGGGT	60.8	Hybrid	3
G3	TGGGGTTGGGGTTGGG GTTGGGG	86.2	Hybrid	2
G4	TTTTGGGTGGGTGGGT GGGTTTT	>95	Parallel	1

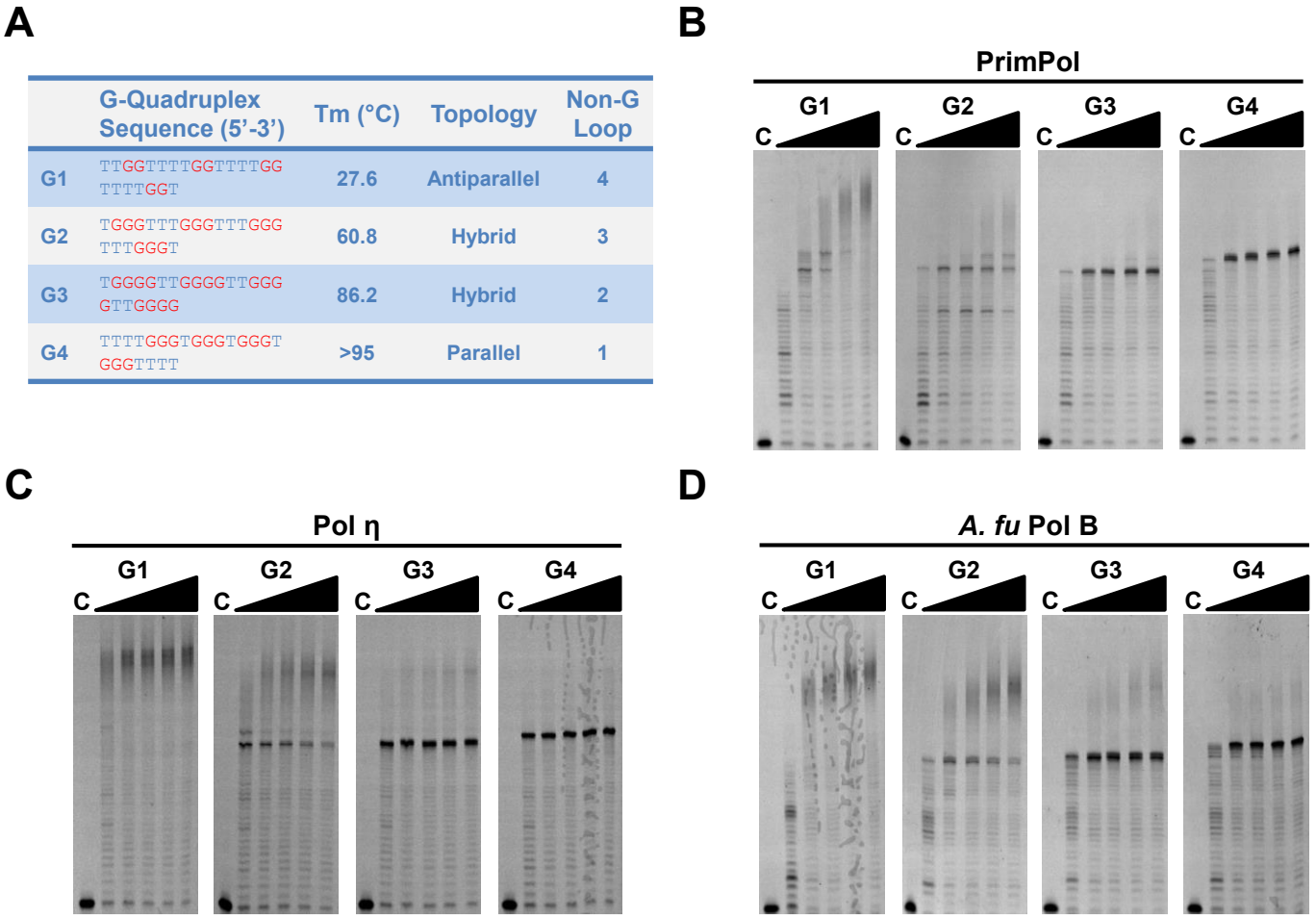


Figure 4.15. G-quadruplex loop length affects polymerase bypass ability.
(A) Sequences of the G-quadruplexes used in primer extension assays, showing the melting temperature (Tm), topology, and loop length. **(B/C/D)** Primer extension reactions with PrimPol **(B),** Pol η **(C),** and *A. fulgidus* Pol B **(D)** on templates containing G-quadruplexes with different non-G loop lengths. All polymerases show increased levels of inhibition on templates containing G-quadruplexes with shorter loop lengths and a higher Tm.

replicated. Loop length also negatively correlates with melting temperature (T_m) (i.e. those structures with longer loops display a lower *in vitro* T_m), suggesting that structures with shorter loop lengths are more stable. Interestingly, the Sale group previously found that in Rev1-deficient cells, Bu-1a instability correlated with G-quadruplex loop length but not *in vitro* T_m (Schiavone et al., 2014). However, when analysing PrimPol-deficient cells the opposite effect was observed, with the greatest instability being observed with G-quadruplexes possessing short loops and a high *in vitro* T_m (Schiavone et al., 2016). The inability of PrimPol to directly read through structures with these characteristics *in vitro* would suggest that it does not employ a TLS-like bypass mechanism *in vivo*. This observation may be indicative of two different G-quadruplex bypass mechanisms. Less stable structures with longer loop lengths are bypassed in a Rev1 dependent TLS-like mechanism, whereas more stable structures with shorter loop lengths potentially require PrimPol-mediated repriming.

4.4.2.5. PrimPol does not Bypass G-Quadruplexes Through a TLS-like Mechanism

The results presented above suggested that PrimPol does not possess an increased ability to perform TLS bypass of G-quadruplexes in comparison to Pol η or *A.fu* Pol B. However, the possibility remained that this was due to the limited range of G-quadruplex structures tested. To rule out this possibility, bypass by PrimPol, Pol η and *A.fu* Pol B was compared on templates containing various G-quadruplex motifs using primer extension assays (sequences 1-2 and 7-13, Table 4.2.). Again, PrimPol did not exhibit an increased ability to bypass any of the structures tested in comparison to Pol η and *A.fu* Pol B (Figure 4.16). Together, these results strongly suggest that PrimPol does not perform TLS bypass of G-quadruplex structures *in vivo*. Note that these experiments were repeated and followed-up by Dr Stanislaw Jozwiakowski using a range of primer lengths, templates, and accessory factors. The same results were observed and can be found in the final published manuscript (Schiavone et al., 2016). Additionally, these follow-up experiments identified that PrimPol shows a preference for binding to G-quadruplex structures and homopolymeric dG sequences in EMSAs.

4.4.2.6. PrimPol can Catalyse Close-Coupled Repriming Downstream of a G-Quadruplex Structure

Given the inability of PrimPol to extend through G-quadruplex structures *in vitro*, its capacity to reprime on the distal side of these structural barriers was next examined using the fluorescent primase assay. G4#4, which forms a highly stable G-quadruplex (Schiavone et al., 2014) and potent block to PrimPol (Figure 4.15.B.), was incorporated into a mixed sequence template strand. In order to analyse repriming downstream of the

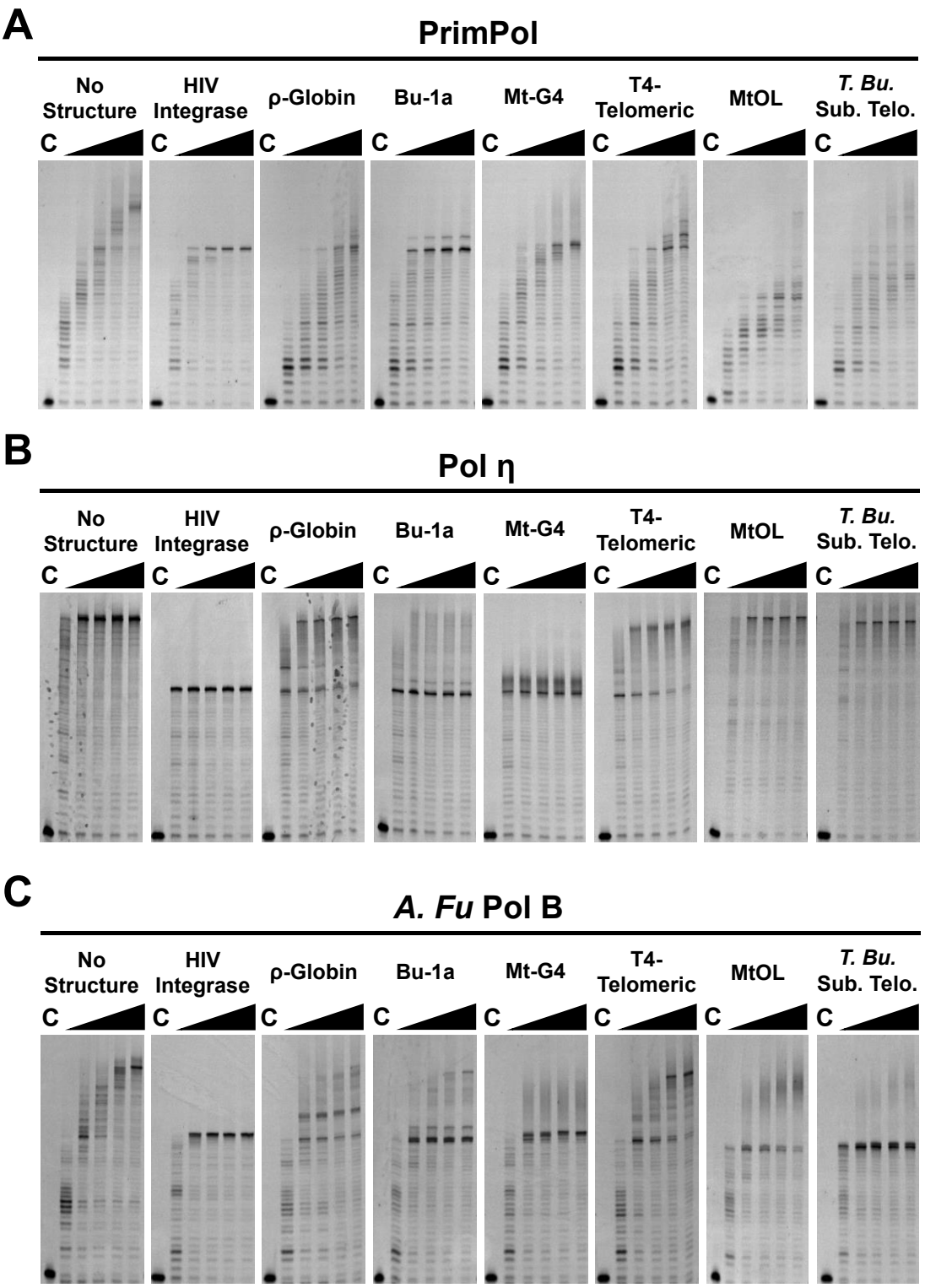


Figure 4.16. Comparison of bypass of various DNA secondary structures by PrimPol, Pol η, and A. fulgidus Pol B.
(A/B/C) Primer extension assays using PrimPol **(A)**, Pol η **(B)**, and *A. fulgidus* Pol B **(C)** on templates containing various G-quadruplexes and DNA secondary structures. The 'no structure' template was included for comparison of polymerase activity in the absence of DNA secondary structures.

G-quadruplex, and better represent a situation where replication has stalled at the structure, a primer containing a 3'-dideoxynucleotide (3'-dd) (sequence 1, Table 4.3.) was annealed upstream of the G-quadruplex motif (sequence 2, Table 4.3) (Figure 4.17). Additionally, templates containing only a short sequence (5 bases) upstream of the G-quadruplex, and no 3'-dd primer, were used to eliminate any artifactual results caused by the primer (sequence 3, Table 4.3.). Although PrimPol was unable to synthesise through this G-quadruplex, it did catalyse *de novo* synthesis of primer strands on the G-quadrupelx templates (Figure 4.17). The size of the extended products, both on templates with and without a 3'-dd primer, were consistent with repriming ~6 nucleotides downstream of the G-quadruplex structure. When tested on the equivalent templates containing no G-quadruplex structure (sequence 4, Table 4.3.), PrimPol synthesised longer and more variable products, suggesting priming in multiple locations further upstream. Based on template configurations and the lengths of fully extended primers, it is apparent that repriming on the G-quadruplex templates is occurring almost immediately after the structure, leaving only a minimal sized gap before the restart of replication is resumed. Although PrimPol has a preference for primer synthesis on pyrimidine tracts (Bianchi et al., 2013), repriming is initiated on these mixed sequence templates. This mechanism is consistent with a role in reinitiating DNA synthesis immediately after G-quadruplexes during replication.

As previously mentioned, PrimPol's primase, but not polymerase, activity is dependent upon an intact ZnF domain (Keen et al., 2014b). In order to further investigate the repriming mechism indicated by *in vitro* experiments, the Sale laboratory analysed the ability of a ZnF-knockout (ZfKO) mutant of PrimPol to permit G-quadruplex replication *in vivo*. Here, complementation of PrimPol$^{-/-}$ DT40 cells with the ZfKO PrimPol mutant, in contrast to the wild-type protein, was not able to restore Bu-1a expression (Schiavone et al., 2016). These findings, therefore, support the *in vitro* results and together strongly suggest that PrimPol reprimes downstream of G-quadruplex structures following stalling of the replisome on the leading strand.

4.4.3. Repriming by PrimPol is Critical for DNA Replication Restart Downstream of Lesions and Chain-Terminating Nucleosides

The identification of the importance of PrimPol in tolerating UV-induced lesions generated interest into what other types of damage the enzyme may be required for the tolerance of (Bianchi et al., 2013). In collaboration with the Hirota group, it was revealed that PrimPol$^{-/-}$ DT40 cells are sensitive to methymethane sulfonate (MMS), cisplatin, and hydroxyurea (HU), which cause Ap sites, crosslinks, and reduction of the dNTP pool,

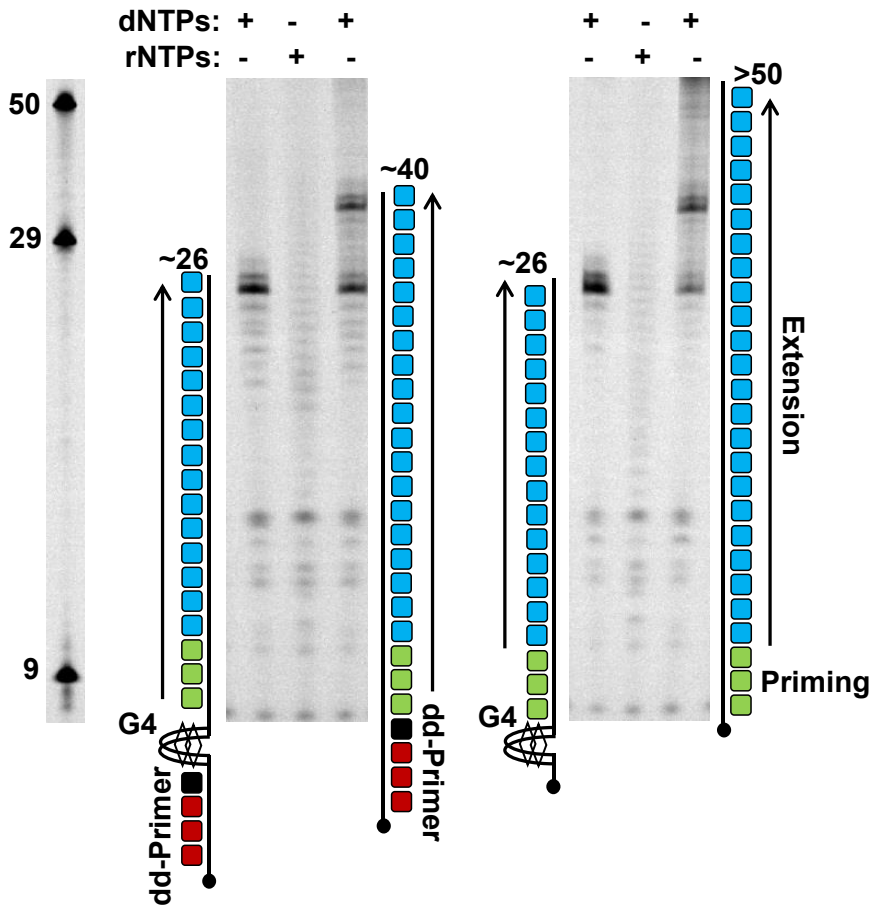


Figure 4.17. PrimPol can catalyse close-coupled repriming downstream of a G-quadruplex.
PrimPol (2 μM) was incubated for 30 min at 37°C with dNTPs or rNTPs (250 μM), FAM-dNTPs (dATP, dCTP, dUTP) (2.5 μM), and mixed sequence G-quadruplex-containing or control templates (1 μM) (as shown in the schematic). Identical reactions were also performed with rNTPs (250 μM) instead of dNTPs on the G4 containing templates only (middle lanes). Templates were either annealed to primers containing a 3' dideoxynucleotide (shown in red) upstream of the G-quadruplex structure or contained only a short sequence (5 nt) before the structure. Priming and extension are represented as green and blue, respectively. The length of products extended to the end of the template by PrimPol allows analysis of the priming location; identically sized extension products on both G-quadruplex-containing templates reveals close-coupled repriming downstream of the structure in each case. Nucleotide (Nt) length markers are shown in the left panel.

respectively. However, loss of PrimPol did not sensitise cells to camtothecin, ICRF193, or γ rays, all agents which cause strand breaks (Kobayashi et al., 2016). These results therefore support a role for PrimPol in DDT, but not repair. Moreover, as had previously been observed with UV sensitivity (Bianchi et al., 2013), further loss of Pol η and Pol ζ increased the sensitivity of PrimPol$^{-/-}$ cells to replisome stalling DNA damage, indicating that PrimPol is employed in a distinct pathway to these TLS polymerases.

In addition to the canonical DNA damaging agents examined above, the sensitivity of PrimPol$^{-/-}$ cells to CTNAs was also explored. CTNAs prevent extension of DNA polymers following incorporation into the 3'-termini, often because they lack a 3' hydroxyl group, consequently causing replicases to stall. Importantly, these obstacles cannot be overcome by conventional TLS or TS as they prevent polymerase-mediated elongation of the daughter strand. Intriguingly, PrimPol$^{-/-}$ cells displayed sensitivity to a number of CTNAs, including Abacavir, Zidovudine, and acyclovir, suggesting a role in restarting replication downstream by repriming (Kobayashi et al., 2016).

4.4.3.1. PrimPol Reprimes Replication Downstream of CTNA Incorporated Sites and DNA Damage Lesions *In vitro*

Given the critical role of PrimPol in cellular tolerance to CTNAs and the presumed requirement of the enzyme's primase activity for this tolerance, its capacity to reprime downstream of the 3' side of an incorporated CTNA was investigated. Since Abacavir is phosphorylated in a unique stepwise anabolism and converted to the triphosphated guanine analog, carbovir, in cells, repriming downstream of carbovir, in addition to acyclovir, was tested *in vitro* (Faletto et al., 1997). In order to analyse repriming downstream of a CTNA incorporation site, a primer containing a CTNA (carbovir or acyclovir) at its terminal 3' end was annealed to a biotinylated DNA template (sequences 7-9, Table 4.3.). In addition, we analysed the ability of PrimPol to reprime downstream of an Ap site and Tg lesion located in the template strand (sequences 5 and 6, Table 4.3.), both of which PrimPol is unable to bypass through TLS in the presence of magnesium (Keen et al., 2014b). In this case, a 3' dideoxynucleotide primer was annealed upstream of the templating lesion to represent a situation where replication has stalled at the damage site (sequence 1, Table 4.3.). The 3' dideoxy moiety also prevents template-independent primer extension that interferes with the evaluation of PrimPol's repriming activity. Although PrimPol was unable to synthesise through the lesions or extend from 3' terminal CTNAs, the enzyme displayed a capacity to perform close-coupled *de novo* synthesis of primer strands downstream in each case (Figure 4.18.). The size of the extended products, both with 3' carbovir and 3' acyclovir primers, in

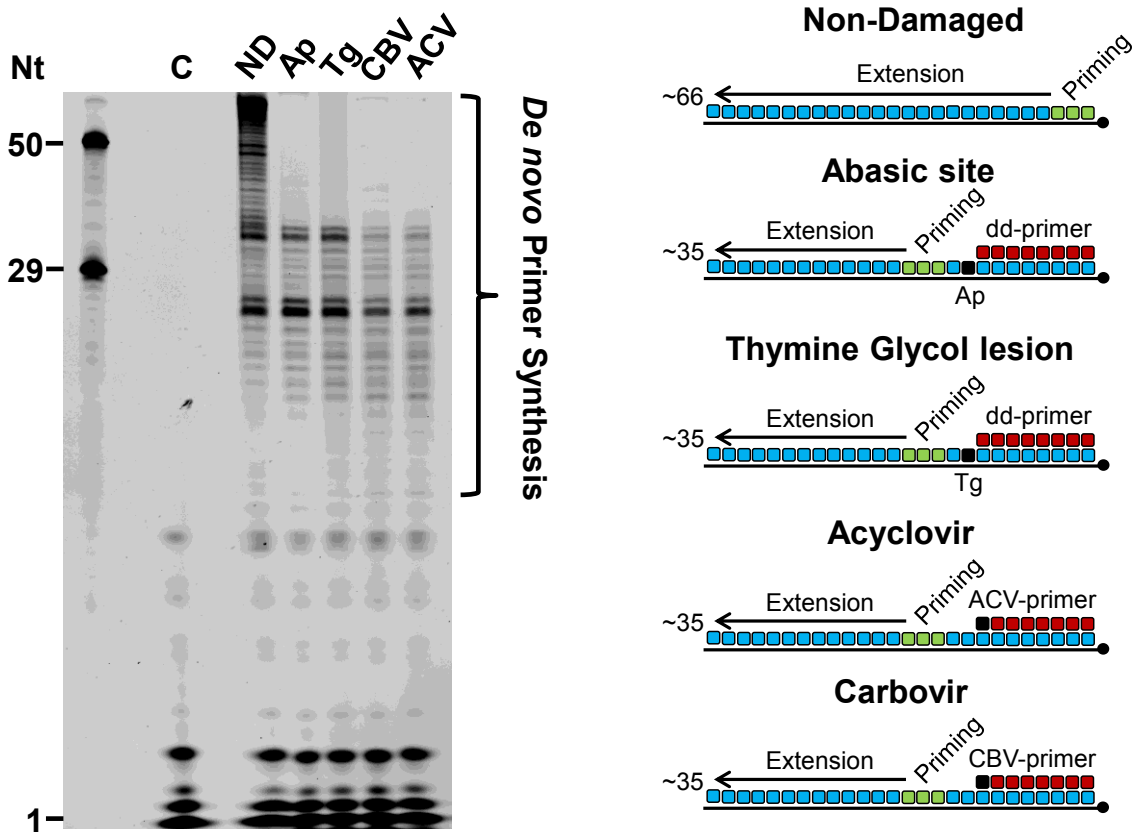


Figure 4.18. PrimPol catalyses repriming downstream of 3' incorporated CTNAs and templating abasic sites or thymine glycol lesions. PrimPol (1 μM) was incubated for 15 min at 37°C with dNTPs (250 μM), FAM-dNTPs (dATP, dCTP, dUTP) (2.5 μM), and mixed sequence primer-templates (1 μM) (as shown in the schematic). Primers containing a 3' dideoxynucleotide were annealed upstream of the lesion on templates containing a single Ap site (Ap) or thymine glycol lesion (Tg) to allow analysis of repriming, rather than TLS, activity. In the case of CTNAs, a single CTNA (acyclovir (ACV) or carbovir (CBV)) was located at the 3' end of the primer in place of the dideoxnucleotide. The length of primase reaction products extended to the end of the template allows analysis of the priming location by PrimPol; the near identical extension products produced in each case show close-coupled repriming by PrimPol downstream of the lesion or CTNA. Nucleotide (Nt) length markers are shown on the left. Priming and extension are represented in the schematic as green and blue, respectively. "C" indicates the no enzyme control. "ND" indicates the non-damaged template without an annealed primer.

addition to the templating Ap site and Tg lesion, were consistent with repriming ~14 nt downstream of the CTNAs or lesion site. Importantly, in the absence of the CTNA primer or lesion, PrimPol generated longer and more variable synthesis products, indicating that the enzyme is capable of performing close-coupled repriming downstream of a stalled replication fork.

These *in vitro* observations of repriming were supported by further *in vivo* work. PrimPol$^{-/-}$ cells were complemented with wild-type PrimPol, the ZfKO mutant, or PrimPol^{Y89D} and damage sensitivity was analysed. Previously, it was shown that mutation of tyrosine 89 to aspartic acid (PrimPol^{Y89D}) causes significantly reduced polymerase activity, but primase activity is retained (Keen et al., 2014a). As previously mentioned, mutation of the ZnF abolishes primase activity. It was found that wild-type PrimPol and PrimPol^{Y89D}, but not the ZfKO mutant, were able to suppress hypersensitivity to MMS, UV, cisplatin, and CTNAs (Kobayashi et al., 2016). This confirms that PrimPol's primase activity is critical for the tolerance of these DNA damaging agents *in vivo*, thus supporting the *in vitro* results.

4.5. Discussion

Taken together, the results generated in this chapter, enabled by the development of the fluorescence-based primase assay, support a model whereby PrimPol's primary role is to reprime leading strand replication downstream of a range of replisome-stalling obstacles. Coupled with the *in vivo* results of the Sale and Hirota groups, it is clear that PrimPol's primase activity is sufficient and critical for the tolerance of non-canonical replication impediments, including G-quadruplexes and CTNAs. Moreover, the enzyme's primase activity is required for bypass of canonical DNA damage, such as Ap sites and Tg lesions, which it cannot bypass by TLS. In this regard, the range of fork stalling damage types which PrimPol participates in the tolerance of, is at odds with a specialised role in TLS bypass of specific lesions. Indeed, the enzyme appears to be more important for overcoming obstacles which it cannot directly read through, including DNA secondary structures, lesions, and CTNAs. This is strongly indicative that PrimPol is employed as a general DDT enzyme, required to reprime downstream of a range of replisome impediments, rather than as a TLS polymerase. Importantly, to be able to reprime downstream of a stalled replicase, the mechanism controlling PrimPol's recruitment must be distinct from that of TLS and replicative polymerases. Here, PrimPol would need to contact ssDNA downstream of the impediment and not the primer template junction, potentially explaining the lack of an interaction with PCNA reported in Chapter 2. In

Chapter 5 the molecular basis and potential mechanism responsible for recuiting PrimPol to reprime replication is investigated.

Chapter 5

Molecular Basis for PrimPol Recruitment to Replication Forks by RPA

5.1. Abstract

DNA damage and secondary structures can act as potent obstacles to the replication machinery. Persistent stalling events lead to genomic instability and therefore numerous cellular tolerance mechanisms exist to complete genome duplication in the presence of such impediments. In addition to TLS polymerases, eukaryotic cells contain a multi-functional replicative enzyme called PrimPol that is capable of directly bypassing DNA damage by TLS, as well as repriming replication downstream of lesions and secondary structures. Here, we report that PrimPol is recruited to reprime replication through its interaction with RPA. Using crystallographic and biophysical approaches, we identify that PrimPol possesses two RPA-interacting motifs and identify the key residues required for these interactions. We demonstrate that one of these motifs is critical for PrimPol's recruitment to stalled replication forks *in vivo* thus facilitating its role in repriming DNA synthesis. In addition, biochemical analysis reveals that RPA serves to stimulate the primase activity of PrimPol. Together, these findings provide unprecedented molecular insights into PrimPol's mode of recruitment to stalled forks that enables it to efficiently reprime and restart DNA replication.

5.2. Introduction

An intricate complex of molecular machines, known collectively as the replisome, duplicate the genome during DNA replication. At the heart of the replisome are the replicative polymerases, which synthesise DNA with a high degree of accuracy and efficiency during this process. Nevertheless, these enzymes are vulnerable to aberrations in the DNA template, including DNA damage lesions and secondary structures, which lead to the stalling of replication at these sites. A number of mechanisms exist to permit the resumption of replication during these events (Blow et al., 2011; Li and Heyer, 2008; Sale et al., 2012). One such mechanism is the generation of a nascent primer downstream of the obstacle, termed repriming (Yeeles et al., 2013). This allows the replisome to effectively skip over the impediment and restart replication.

Until recently, Pol α-primase was thought to be the only eukaryotic primase, we now know eukaryotes possess a second primase, known as PrimPol, which has been discovered and characterised by a number of groups (Bianchi et al., 2013; García-Gómez et al., 2013; Wan et al., 2013). PrimPol is a member of the archaeo-eukaryotic primase (AEP) superfamily, whose members fulfil a range of roles in DNA replication, repair and damage tolerance (Guilliam et al., 2015b). In line with this, PrimPol displays both primase and TLS polymerase activity (Bianchi et al., 2013; García-Gómez et al., 2013). Evidence is accumulating that suggests the primary role of PrimPol *in vivo* is to reprime DNA replication downstream of DNA damage lesions and secondary structures (Keen et al., 2014b; Kobayashi et al., 2016; Mourón et al., 2013; Schiavone et al., 2016). Despite assisting the replisome through this role, PrimPol could be potentially deleterious to genomic integrity due to its low fidelity and penchant for generating frame-shift mutations (Guilliam et al., 2015a). As a result, the enzyme must be tightly regulated and only allowed to contribute to DNA synthesis when absolutely required.

We previously identified the nuclear and mitochondrial single-stranded DNA binding proteins, RPA and mtSSB, as PrimPol interacting partners *in vivo*. Using biochemical approaches, we demonstrated that both of these binding-partners could serve to restrict the contribution of PrimPol to DNA synthesis during replication and therefore limit the opportunity for mutagenesis (Guilliam et al., 2015a). PrimPol was also identified as an RPA-binding partner by another group (Wan et al., 2013). In this study, it was suggested that RPA may act to recruit PrimPol to stalled replication forks *in vivo*, although interpretation of these results is limited as large deletion-mutants (~80 amino acids removed) were used for analysis.

In this study, we present an in-depth interrogation of the interaction between PrimPol and RPA, identifying that PrimPol possesses two **R**PA-**b**inding **m**otifs (RBMs) in its C-terminal domain (RBM-A and RBM-B). Both of these are able to bind directly to RPA70N, a primary recruitment domain of RPA that mediates interactions with a number of DNA damage response proteins, including p53, ATRIP, RAD9 and MRE11 (Xu et al., 2008). Using crystallographic and biophysical approaches, we elucidated the molecular basis of each of the PrimPol RBM-RPA interactions and identified the critical residues involved in each case. We generated PrimPol RBM mutants *in vivo* and analysed the importance of each of these sites for PrimPol's role in DNA damage tolerance. We identify that RBM-A is the primary mediator of PrimPol's interaction with RPA *in vivo*, with RBM-B potentially playing a more secondary role. The interaction between RBM-A and RPA70N is critical for the recruitment of PrimPol to chromatin and for stimulating the enzyme's role in repriming DNA replication. Notably, mutations in both RBMs affecting key residues involved in binding (e.g. F522V and I554T) have been identified in cancer patient cell lines and these mutations are sufficient to abrogate binding of RPA70N to the affected RBM. Collectively, these results describe the molecular and cellular basis for PrimPol's recruitment by RPA to stalled replication forks and demonstrate the importance of these interactions for maintaining PrimPol's function in replication fork progression *in vivo*.

5.3. Materials and Methods

5.3.1. Construction and Expression of Human PrimPol and RPA Truncation Variants

Full-length PrimPol was expressed and purified as previously described (Bianchi et al., 2013). PrimPol amino acids 480-560 ($PrimPol_{480-560}$) was cloned into pET28a by polymerase chain reaction using wild-type PrimPol as a template via standard methods (primers 1 and 2, Table 5.1.). $PrimPol_{480-560}$ was expressed in BL21(pLysS) cells overnight at 25 °C and purified using Ni Sepharose (Qiagen), followed by Q Sepharose (GE Healthcare) and gel filtration using a Superdex 75 10/300 GL column (GE Healthcare) according to the manufacturer's instructions. PrimPol residues 480-546 ($PrimPol_{480-546}$), PrimPol D514R, D518R, D519R ($PrimPol_{RBM-A-KO}$), PrimPol D514R, D518R, D519R, D551R, I554A, I555A ($PrimPol_{RBM-A-KO/RBM-B-KO}$), PrimPol F522V on an RBM-B-KO background, and PrimPol I554T on an RBM-A-KO background were cloned by site-directed mutagenesis (primers 3-12, Table 5.1.). PrimPol 480-546 with the D514R, D518R, D519R mutations ($PrimPol_{480-546/RBM-A-KO}$) was also cloned, using the 480-546 construct DNA as a template. All these proteins were expressed and purified as described for $PrimPol_{480-560}$.

#	PCR Primer	Sequence (5'-3')
1	480 FWD	GTTTCTTCATATGACAACAGATGAAGCAGATGAAAC
2	560 REV	CAAAGAAGCGGCCGCTTACTCTTGTAATACTTCTATAATTAGTTC
3	1-546 FWD	GTGAAGTGTAGTAAATTCCTGATGAACTAATTATAG
4	1-546 REV	CAGGAATTTACTACACTTCACTGTTATAACTGAG
5	D514R/D518R/D519R FWD	GGATCCGCCGCGCTTATTTTTTAGAAGCTACTGAAGATGCTGAATTAG
6	D514R/D518R/D519R REV	AAGCGCGGCGGATCCCATTTCTCCAGACAGCATCAGCAGATG
7	D551R/I554A/I555A FWD	TCTCAGTTATAACAGTGAAGTGGATGAAATTCCTCGCGAACTAGCGGCGGAAGTACTGCAGGAG
8	D551R/I554A/I555A REV	GGTGGTGGTGCTCGAGTGCGGCCGCTTACTCCTGCAGTACTTCCGCCGCTAGTTCGCGAGGAATTTC
9	F522V FWD	GATGATGCTTATGTTTTAGAAGCTACTGAAGATGCTGAATTAGCTGAAGC
10	F522V REV	CTTCTAAAACATAAGCATCATCAATGCCATTATCCCAGACAGCATC
11	I554T FWD	CTGATGAACTAACTATAGAAGTATTACAAGAGTAAGATCCGAATTCGAGCTC
12	I554T REV	ATACTTCTATAGTTAGTTCATCAGGAATTTCATCCACTTCACTGTTATAACTGAGAAG
13	N-FLAG 1 FWD	GTTTCTTGGATCCATGGATTACAAGGATGACGACGATAAGGGAAGCCATGGAAGCCATATGAATAGAAAATGGGAAGCAAAACTG
14	N-FLAG 480 FWD	GTTTCTTGGATCCATGGATTACAAGGATGACGACGATAAGGGAAGCCATGGAAGCCATATGACAGATGAAGCAGATGAAAC
15	D519R/F522A FWD	GGCATTGATCGTGCTTATGCTTTAGAAGCTACTGAAGATGC
16	D519R/F522A REV	GCTTCTAAAGCATAAGCACGATCAATGCCATTATCCCAGAC
17	D551R/I554A FWD	GAAATTCCTCGTGAACTAGCTATAGAAGTATTACAAGAG
18	D551R/I554A REV	CTTCTATAGCTAGTTCACGAGGAATTTCATCCACTTCAC

Table 5.1. PCR primers used to construct PrimPol RBM mutants.
Sequences of the primers used in PCR for site-directed mutagenesis and cloning of various PrimPol constructs.

The full RPA timer was expressed and purified as previously described (Masuda et al., 2007). RPA70N ($RPA70_{1-120}$) was cloned as described previously (Souza-Fagundes et al., 2012; Xu et al., 2008). The $RPA70N^{E7R}$ variant that readily forms crystals with basic-site ligands was utilised in the experiments shown here (Feldkamp et al., 2013); the properties of this protein variant are not affected in any way apart from in its crystal lattice contacts. Protein concentrations were determined based on absorbance at 280 nm corrected with the protein-specific extinction coefficient. Extinction coefficient values for each of the recombinant proteins were calculated using ProtParam tool (ExPASy).

RBM-A (510-528) and RBM-B (546-560) peptides for NMR were synthesised (GenScript), purified with a Waters Delta 600 HPLC using a Proto 300 C4 column (Higgins Analytical, Inc.), and confirmed using mass spectrometry. The $PrimPol_{514-528}$ peptide used for co-crystallisation experiments was synthesised (Genscript) and used as supplied

5.3.2. Nuclear Magnetic Resonance Methods

^{15}N-^{1}H HSQC experiments were performed at 25 °C on a Bruker Avance III 800 or 900 MHz NMR spectrometer with a cryogenically cooled probe. Spectra were acquired for 100 µM samples of ^{15}N-enriched RPA70N or ^{15}N-enriched $PrimPol_{480-560}$ alone and in the presence of 200 µM unlabeled binding partner. All samples were equilibrated in a buffer containing 20 mM HEPES (pH 7.5), 80 mM NaCl, 2 mM DTT and 5% deuterium oxide.

5.3.3. Analytical Size-Exclusion Chromatography and SEC-MALS

Protein interactions were analysed by size-exclusion chromatography (SEC) on a Superdex S75 10/300 GL gel filtration column (GE Healthcare). The column was calibrated using albumin (67,000 Da), ovalbumin (43,000 Da), chymotrypsinogen A (25,000 Da), ribonuclease A (13,700 Da), and aprotinin (6,512 Da). The protein was loaded at 0.5 mL min^{-1}. Retention volume of the proteins were plotted against the molecular weight of each protein to reliably predict protein molecular weights. The column was pre-equilibrated in a buffer containing 50 mM Tris-HCl (pH 7.5), 100 mM NaCl and 2 mM TCEP that had been sterile-filtered using a 0.2 µm pore size vacuum filtration system (Nalgene). 0.5 mL of protein was loaded at a concentration of 35 µM. Protein interactions were determined by a shift in the chromatograph peaks relative to individual protein peaks.

SEC multi angle light scattering (MALS) was performed on an AKTA Purifier FPLC system (GE) connected to an Agilent Technologies 1200 Series Refractive Index (RI) Detector and a Wyatt Technologies Dawn Helios 8+ MALS unit. A Superdex 75 increase

10/300 GL (24 mL) column was equilibrated in running buffer consisting of 20 mM HEPES (pH 7.1), 80 mM NaCl, 0.5 mM TCEP. The flow was maintained at a consistent 0.5 $mLmin^{-1}$ and sample injections of 100 µL from a static loop were initiated at the 0 mL point of each run. UV, RI, Quasi-Elastic LS, and LS values were recorded using ASTRA 6.1 (Wyatt) software. Data were collected using samples of RBD at 185 µM with RPA70N E7R added at 0, 1, 2 and 4x molar ratios. Estimated molecular weights for RBD and its saturated complexes were calculated using the Zimm algorithm surrounding the peak maximum.

5.3.4. Crystallisation and X-ray Structure Solution

Crystals of the RPA70N-PrimPol complex were grown at 293 K by vapour diffusion as sitting drops. The protein complex was screened at a 2.5:1 ratio of 1.75 mM $PrimPol_{514-528}$ peptide: 0.70 mM of $RPA70N^{E7R}$ in drops containing 0.5 µL of protein complex mixed with 0.5 µL of crystallization buffer (0.2 M Ammonium acetate 0.1 M Sodium acetate 4.5 20 % w/v PEG 3350). Prior to data collection, crystals were cryoprotected by soaking in mother liquor containing 20% ethylene glycol. 1.542 Å X-ray diffraction data was collected in-house at 100 K using a Rigaku MicroMax 007-HF. The diffraction data were processed with SCALA (Evans, 2006) with additional processing by programs from the CCP4 suite (Collaborative Computational Project, Number 4, 1994).

Crystals of the RPA70N-PrimPol complex were grown at 293 K by vapour diffusion as sitting drops. The protein complex was screened at a 1:1 ratio with 700 µM of each of $RPA70N^{E7R}$ and $PrimPol_{480-560}$; 0.5 µL of protein complex was mixed with 0.5 µL of crystallisation buffer (0.2 M imidazole malate (pH 6.0), 30% (w/v) PEG 4000). Prior to data collection, crystals were soaked in mother liquor containing 20% ethylene glycol. 0.914 Å X-ray diffraction data was collected at 100 K using a synchrotron source at station I03 Diamond Light Source, Didcot, UK. The diffraction data were processed with xia2 (Winter et al., 2013) with additional processing by programs from the CCP4 suite (Collaborative Computational Project, Number 4, 1994). The statistics for data processing are summarized in Table 1. For both models, initial phases were obtained by molecular replacement with PHASER (McCoy et al., 2007) (using $RPA70N^{E7R}$ (4IPC) (Feldkamp et al., 2013) as a search model). Iterative cycles of model building and refinement were performed using Coot (Emsley et al., 2010) and Phenix (Adams et al., 2010). A final refined model at 2.0 Å resolution, with an R_{factor} of 18.73% and R_{free} of 22.86% was obtained for the RPA70N-$PrimPol_{514-528}$ peptide complex. The Ramachandran statistics for this complex place 97.9% of residues in the favoured region and 2.1% in the allowed region. For the RPA70N-PrimPol complex a refined model at

1.28 Å resolution, with an R_{factor} of 15.37% and R_{free} of 17.85 % was obtained with Ramachandran statistics of 98.6% of residues in the favoured region and 1.4% in the allowed region. Structural images were prepared with CCP4mg (McNicholas et al., 2011). Stereo images for portions of the electron density of RPA70N-PrimPol$_{514-528}$ and RPA70N-PrimPol$_{480-560}$ are shown in Figure 5.1. The structures of the RPA70N-PrimPol$_{514-528}$ peptide complex and the RPA70N-PrimPol$_{480-560}$ complex are deposited in the Protein Data Bank under accession codes 5N85 and 5N8A, respectively.

5.3.5. Circular Dichroism

PrimPol RBD samples for CD were equilibrated in 20 mM HEPES (pH 7.5), 80 mM NaCl, and 2 mM DTT, and then diluted 1:10 with ultrapure water to a final concentration of 20 µM. A JASCO J-810 spectrophotometer equilibrated at 25 °C was used to collect 5 scans over the spectral width 190-250 nm. Molar ellipticity was calculated based on the final protein concentration of 20 µM.

5.3.6. Dynamic Light Scattering

Light scattering experiments were performed using a Wyatt Technology DynaPro NanoStar instrument. PrimPol RBD was equilibrated in 20 mM HEPES (pH 7.5), 80 mM NaCl, and 2 mM DTT at a concentration of 200 µM. A 5 µL sample was then equilibrated in a COC cuvette at 25 °C for 5 minutes prior to acquisition. Ten data points were acquired and fitted using the coils protein shape model using Wyatt Dynamics software. The resulting regularization graph was plotted as a function of %Mass and Mw-R calculated based on observed sample radius.

5.3.7. Isothermal Titration Calorimetry

Isothermograms were recorded using a MicroCal VP-ITC instrument. 10 µL injections of 400 µM PrimPol RBD (either WT or variants) were added to a 1.4 mL cell with RPA70N at 20 µM. The system was equilibrated for 5 minutes between injections. Both proteins were dialysed in the same pool of 20 mM HEPES (pH 7.5), 80 mM NaCl, and 3 mM β-mercaptoethanol buffer. Dissociation constants were calculated with MicroCal Origin software using a single site binding model.

5.3.8. Fluorescent M13 Primase Assay

Full-length PrimPol (400 nM) was incubated in 20 µL reactions containing 10 mM Bis-Tris-Propane-HCl (pH 7.0), 10 mM $MgCl_2$, 1mM DTT, 250 µM dNTPs, 2.5 µM FAM dTNPs (dATP, dCTP, dUTP), and 20 ng/µL single-stranded M13 template, at 37 °C for 15 mins. Individual reactions were supplemented with increasing concentrations of RPA

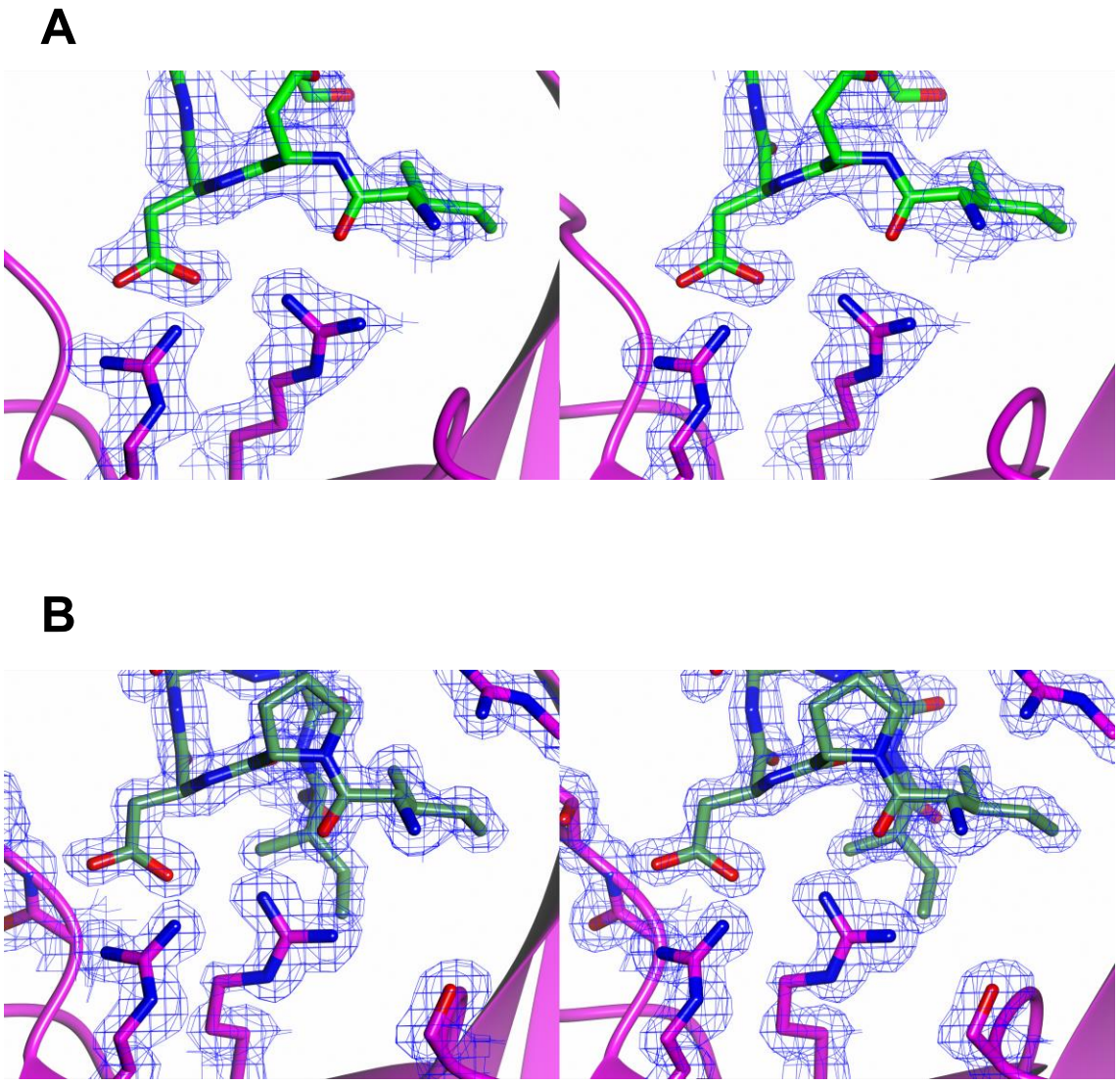


Figure 5.1. Stereo views of electron density for RBM-A and RBM-B.
(A) A stick representation of the residues I517-A520 of RBM-A (light green). Residues R43 and R31 from RPA70N that form the ionic interactions with RBM-A are also depicted (magenta). Density from a weighted 2Fo-Fc map scaled at 0.6σ is depicted in blue. **(B)** Residues I549-I554 of RBM-B are depicted (dark green). Residues R91, S54, R43, R31 and T34 from RPA70N involved in ionic interactions with RBM-B are also shown (purple). Density from a weighted 2Fo-Fc map scaled at 0.8σ is also depicted in blue.

(0, 0.0625, 0.125, 0.25, 0.5, 1, 2, 4, and 8 μM) prior to the addition of PrimPol. Following primer synthesis, remaining free FAM dNTPs were removed using an Oligo Clean and Concentrator kit (Zymo Research) according to the manufacturer's instructions. Eluted primers were supplemented with loading buffer (95% formamide with 0.25% bromophenol blue and xylene cyanol dyes) (total volume 20 μL). Samples were boiled and resolved on a 15% polyacrylamide/7M urea gel for 90 mins. Products were visualised on an FLA-5100 imager.

5.3.9. Maintenance and Generation of Stable HEK293 Flp-In T-Rex Cells

HEK293 Flp-In T-REx (Invitrogen) cells were cultured in DMEM containing 10% foetal calf serum (FCS), 1% L-glutamine, and 1% PenStrep. For the generation of stable inducible N-terminal FLAG-tagged PrimPol HEK293 Flp-In T-REx, cells were grown in medium containing 15 μg/mL Blasticidin (Invitrogen) and 100 μg/mL Zeocin prior to transfection. Cells were transfected with pOG44 plasmid and pcDNA5/FRT/TO plasmid (1:9 ratio) encoding FLAG-PrimPol (WT, D519R/F522A, D551R/I554A, D519R/F522A/D551R/I554A, and PrimPol$^{480-560}$) using Lipofectamine 2000 following the manufacturer's instructions. pcDNA5/FRT/TO constructs encoding N-terminal FLAG-PrimPol and FLAG-PrimPol$^{480-560}$ were generated by standard PCR and cloning procedures (primers 13, 14 and 2, Table 5.1.). RBM mutant N-FLAG PrimPol constructs were produced by site-directed mutagenesis (primers 15-18, Table 5.1.). 48 hours after transfection selective medium containing 15 μg/mL Blasticidin and 100 μg/mL Hygromycin (Invitrogen) was added. Selective medium was replaced every 2-3 days, until resistant clones appeared. Clones were then pooled, expanded, and stocks made.

5.3.10. Co-Immunoprecipitation in HEK-293 Cells Expressing FLAG-PrimPol

HEK293 Flp-In T-REx cells engineered for inducible expression of FLAG-PrimPol (WT, D519R/F522A, D551R/I554A, D519R/F522A/D551R/I554A, and PrimPol$^{480-560}$) were grown in the presence or absence of doxycycline (10 ng/mL) 24 hours before harvesting. Cell pellets were re-suspended in 1 mL lysis buffer (150 mM NaCl, 30 mM Tris-HCl (pH 7.5), 0.5% NP40, 2.5 mM $MgCl_2$, 100 μg/mL DNase I) and incubated at 4 °C for 30 mins. The resulting lysate was clarified by centrifugation at 10,000xg for 10 mins at 4 °C. The supernatant was retained (sample taken as 'input'), added to 100 μL pre-washed anti-FLAG magnetic beads (Sigma) and incubated at 4 °C overnight. Unbound material was removed and the beads were washed 3x5 mins with 1 mL wash buffer (Lysis buffer without DNase I and with 0.1% NP40). Three successive 5 min elutions were performed using 200 μL elution buffer (25 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM PMSF, and

200 µg/mL 3xFLAG peptide (Sigma)). Eluted samples were boiled in Laemmli buffer and analysed by western blot using the following antibodies; Anti-FLAG (Sigma F3165; 1:1000 dilution), Anti-RPA2 (Calbiochem NA18; 1:500 dilution), HRP (horseradish peroxidase) conjugated Anti-mouse IgG (Abcam ab6728; 1:5000 dilution).

5.3.11. Triton X-100 Fractionation of HEK-293 Cells

HEK-293 cellular fractionation was performed as previously described (Bianchi et al., 2013). Briefly, protein expression was induced (10 ng/mL doxycycline, 24 hours) in HEK293 Flp-In T-REx cells stably transfected with various FLAG-PrimPol constructs (WT, D519R/F522A, D551R/I554A, and D519R/F522A/D551R/I554A). The following day cells were either mock or UV-C (30 J/m^2) irradiated and allowed to recover for 3 hours. Cells were harvested and pellets resuspended in cytoskeletal buffer (100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 10 mM PIPES (pH 6.8), 1 mM EGTA, 0.2% Triton X-100, and protease and phosphatase inhibitors (Roche)), followed by incubation on ice for 5 mins. Samples were then centrifuged at 13,000 rpm for 10 mins. Supernatant was retained as the soluble fraction. The insoluble pellet was washed three times in PBS and boiled in Laemmli buffer. Whole cell extract, soluble, and insoluble, samples were analysed by western blot using Anti-FLAG, Anti-RPA2, and Anti-mouse IgG antibodies sourced and used as described above, in addition to Anti-Histone H3 (Abcam ab1791; 1:5000 dilution) and HRP conjugated Anti-Rabbit IgG (Abcam ab6721; 1:3000 dilution).

5.3.12. Chromatin Isolation from *Xenopus* egg Extract

Demembranated sperm nuclei were prepared by lysolecithin treatment as previously described (Murray, 1991). Preparation of Xenopus egg extracts and the isolation of chromatin from egg extract were carried out as previously described (Taylor et al., 2013). Western blot analysis was performed using the following antibodies; Anti-GST (Abcam ab92; 1:2000 dilution), Anti-Orc2 (gift from Julia Blow; 1:2000 dilution), HRP conjugated Anti-mouse IgG (DAKO P0260; 1:5000 dilution), and HRP conjugated Anti-rabbit IgG (DAKO P0448; 1:5000 dilution).

5.3.13. Complementation and DNA Fibre Assays in PrimPol$^{-/-}$ DT40 Cells

DT40 cells were cultured in RPMI 1640 medium containing 10% FCS, 1% chicken serum, 10 µM β-mercaptoethanol, 1% L-glutamine, and 1% PenStrep. Mutant DT40 cell lines were derived from DT40 Clone 653 from Prof. S. Takeda's group (Kyoto University). PrimPol$^{-/-}$ DT40 cells (previously generated (Bianchi et al., 2013)) were stably complemented with pCI-neo plasmid encoding WT PrimPol, PrimPol$^{D519R/F522A}$, and PrimPol$^{D551R/I554A}$ by electroporation as previously detailed (Bianchi et al., 2013). pCI-neo

contructs encoding RBM-mutant PrimPol were generated by site-directed mutagenesis (primers 15-18, Table 5.1.). Positive clones were selected using medium containing 2 mg/mL G418 (Sigma) and expression was confirmed by western blot using Anti-PrimPol (raised against recombinant purified PrimPol, 1:1000), Anti-α-Tubulin (Sigma T5168, 1:3000 dilution), and HRP conjugated Anti-Rabbit IgG and Anti-Mouse IgG (sourced and used as described above). All DNA fibre analysis was performed as described previously in triplicate (Bianchi et al., 2013).

5.3.14. Yeast Two-Hybrid Assay

Full-length PrimPol and its C-terminal RBM domain (PP-RBM – a.a. 480-560) were cloned into the NdeI site of the pGADT7 vector using polymerase chain reaction with wild-type PrimPol as a template, and T4 polymerase to process the DNA ends. PrimPol mutants were prepared by site-directed mutagenesis. RPA70N (a.a 1-120) was cloned into NdeI site of the pGBKT7 vector. Plasmids containing the GAL4 activation domain (pGADT7) fused to the PrimPol variants or the empty vector were transformed into the *S. cerevisiae* strain PJ69-4a. Plasmids containing the GAL4 DNA binding domain (pGBKT7) fused to RPA70N or the empty vector were transformed into PJ69-4α strain. The haploid strains were mated on a YPD plate and replica plated on selective medium lacking leucine and tryptophan. The resulting diploid strains were grown to A600 ~ 1 and spotted as 10-fold serial dilutions on media lacking leucine, tryptophan, histidine, or adenine. 1 mM 3-Amino-1,2,4-triazole (3AT) was added to decrease the background HIS3 expression. Plates were scanned after 3 days of incubation at 30 °C

5.4. Results

5.4.1. PrimPol's CTD Interacts with RPA70N

Previously, we identified that full-length PrimPol interacts directly with RPA70N (Guilliam et al., 2015a). In contrast, deletion of the C-terminal **R**PA **b**inding **d**omain (RBD) ablated this interaction. In order to determine if PrimPol's RBD (480-560) is sufficient to mediate binding, we performed analytical gel filtration chromatography (GFC) on $PrimPol_{RBD}$ titrated with RPA70N (Figure 5.2.A.). With one equivalent of RPA70N added, a bimodal peak appears with broadened densities between a position near free $PrimPol_{RBD}$ and a peak presumably of the complex (blue dot trace). With two equivalents of RPA70N added, the peak at the $PrimPol_{RBD}$ position is much weaker, while the complex elutes slightly earlier and increases in intensity (blue dash trace). With four equivalents of RPA70N added, the complex peak increases in intensity, the free $PrimPol_{RBD}$ peak disappears, and a peak at the free RPA70N position becomes visible (blue solid trace).

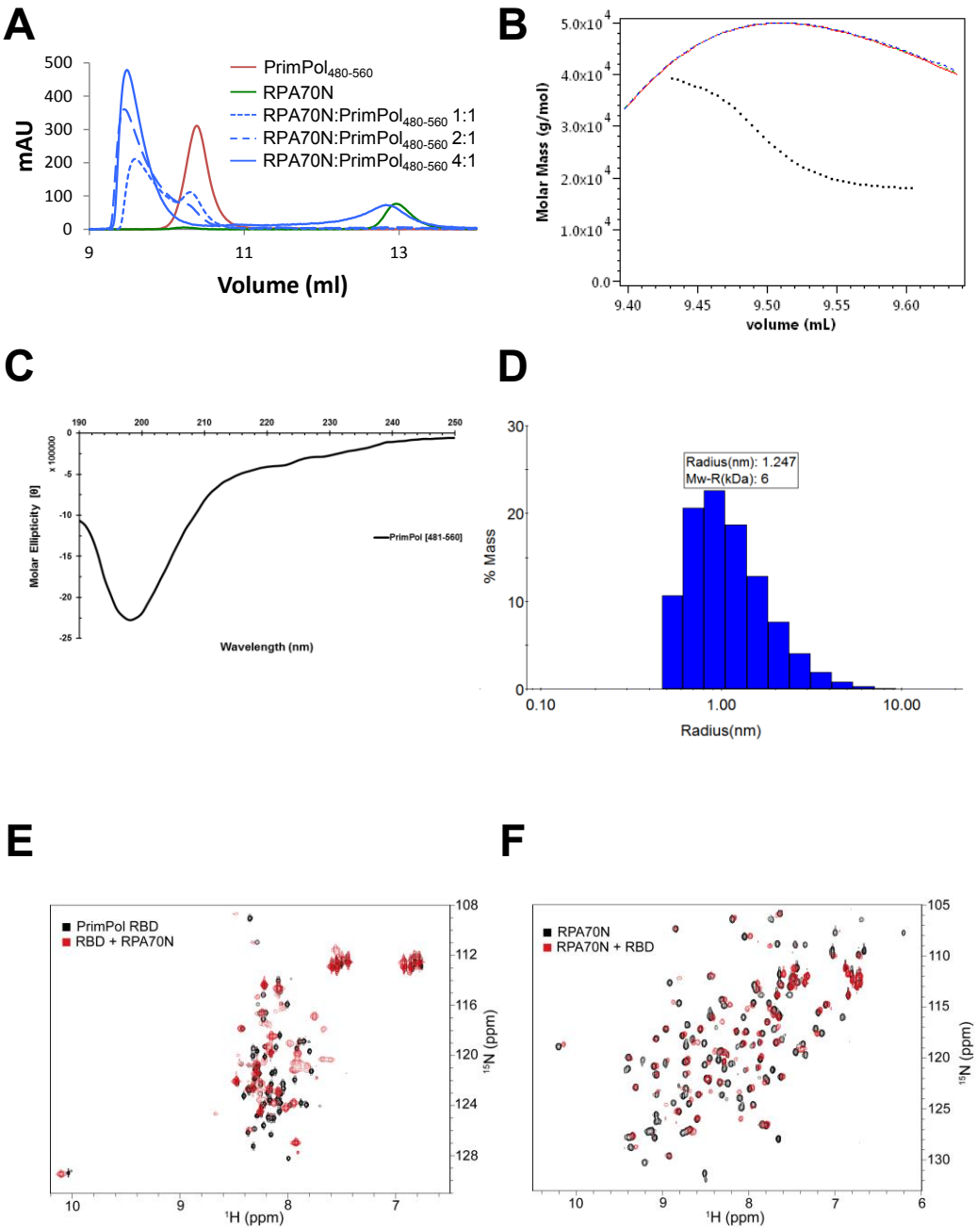


Figure 5.2. PrimPol's RBD interacts with RPA70N.
(A) Chromatograph showing the retention volumes of the PrimPol RBD (residues 480-560) (Red), RPA70N (Green) and RBD titrated with varying molar ratios of RPA70N (Blue). **(B)** MALS analysis of the major peak eluted from the saturated RBD-RPA70N E7R sample at a 1:4 molar ratio. The observed Refractive Index (blue), UV (green), and Light Scattering (red) readings were used to calculate the molecular weight of the complex over the course of elution (black). **(C)** Circular dichroism spectrum of 20 μM $PrimPol_{RBD}$ collected between 190-250 nm showing a spectral shape characteristic of an unstructured protein lacking any significant α-helical or β-strand propensity. **(D)** Regularisation graph of 10 accumulations of dynamic light scattering data collected on $PrimPol_{RBD}$ samples at 200 μM. A single peak was observed with a radius of 1.25 nm that corresponds to a predicted molecular weight of ~6 kDa when modelled with random coil protein shape algorithms. This is close to the expected 8.8 kDa of the monomeric protein. **(E)** ^{15}N-^{1}H HSQC spectra of ^{15}N-enriched PrimPol RBD in the absence (black) and presence (red) of unlabelled RPA70N (70N) titrated at a 2:1 ratio. **(F)** ^{15}N-^{1}H HSQC spectra of ^{15}N-enriched RPA70N in the absence (black) or presence (red) of 2-fold molar excess of unlabelled PrimPol RBD.

This data indicates a heterogeneous interaction, most likely from two binding sites of similar affininty (Figure 5.2.A.). The stoichiometry of the binding is most likely 2:1 RPA70N:PrimPol$_{RBD}$ due to the complete disappearance of the individual RPA70N peak at this ratio (Figure 5.2.A.). This stoichiometry was further confirmed by multiangle light scattering (MALS) analysis of the eluted peak fractions, identifying a heterogenous mix of both 1:1 and 2:1 RPA70N:PrimPol$_{RBD}$ complexes (Figure 5.2.B.). PrimPol$_{RBD}$ had a much lower retention volume (10.39 mL) than expected for an 8.8 kDa protein, corresponding to a predicted molecular weight of ~42 kDa if the protein was globular. Nevertheless, circular dichroism (CD) and Dynamic Light Scattering (DLS) revealed that PrimPol$_{RBD}$ is monomeric in solution with a largely non-globular structure (Figure 5.2.C. and D.).

NMR spectroscopy was next utilised to cross-validate this interaction. To this end, ^{15}N-enriched PrimPol$_{RBD}$ was produced and analysed by two-dimensional (2D) ^{15}N-^{1}H heteronuclear single quantum coherence (HSQC) NMR (Figure 5.2.E.). The low dispersion observed in the ^{1}H dimension of the spectrum is characteristic of a protein with non-globular structure. Upon addition of unlabelled RPA70N to a two-fold molar excess, there was a significant effect on the spectrum, with peaks attenuating, broadening or shifting. These observations confirm that there is an interaction between the two proteins. We also observed significant peak shifting and disappearance in the corresponding spectrum of ^{15}N-enriched RPA70N in the presence of a 2-fold excess of unlabeled PrimPol$_{RBD}$ (Figure 5.2.F.). The large number of peaks affected and the variety of effects on the signals suggest the interaction is not mediated by a single high-affinity site, but rather some form of heterogeneous binding.

5.4.2. PrimPol's RBD Contains a Conserved RPA Binding Motif

RPA70N contains a prominent surface cleft that binds many interacting partners, including RAD9, MRE11, ATRIP and p53 (Xu et al., 2008). These partners utilize a similar highly negatively charged motif, which interacts with the exposed basic residues in the RPA70N cleft (Xu et al., 2008). Examination of the sequence of human PrimPol revealed a divergent acidic motif within its RBD (residues 513-527; Figure 5.3.A.), we termed this motif RPA binding motif A (RBM-A).

To investigate the potential interaction between PrimPol's RBM-A and RPA70N, we employed NMR spectroscopy using an RBM-A peptide (RBM-A$_{510-528}$). An overlay of 2D ^{15}N-^{1}H HSQC spectra of ^{15}N-enriched RPA70N acquired in the absence (black) and presence (red) of a 2-fold excess of RBM-A$_{510-528}$ revealed significant chemical shift perturbations (CSPs) induced by binding of the peptide (Figure 5.3.B.). The CSPs above

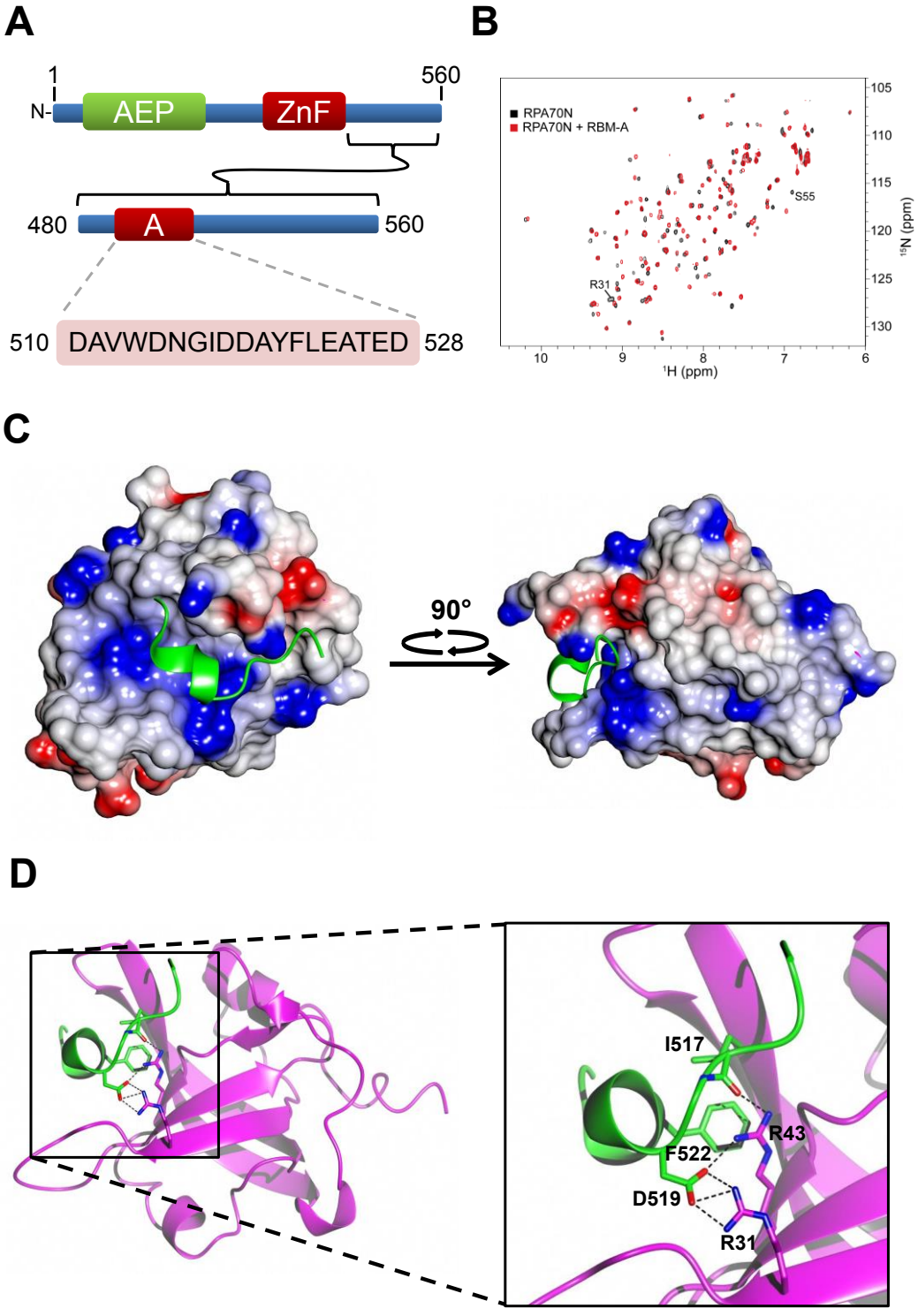


Figure 5.3. PrimPol possesses a conserved RPA binding motif that binds to the basic cleft of RPA70N.
(A) Schematic showing the sequence of PrimPol's RBM-A (residues 510-528), located in the C-terminal RBD (residues 480-560). **(B)** ^{15}N-^{1}H HSQC spectra showing RPA70N in the absence (black) or presence (red) of 2-fold molar excess of unlabelled RBM-A peptide. **(C)** Electrostatic surface model of RPA70N with RBM-A (green) bound in the basic cleft. Basic and acidic surfaces are coloured blue and red, respectively. **(D)** Key stabilising interactions of RBM-A (green) in the RPA70N basic cleft (purple). RBM-A binds between β sheets in the β barrel of RPA70N. Of particular importance for binding are the electrostatic interactions of D519 with the side chains of two arginines (R31 and R43) in the RPA70N basic cleft.

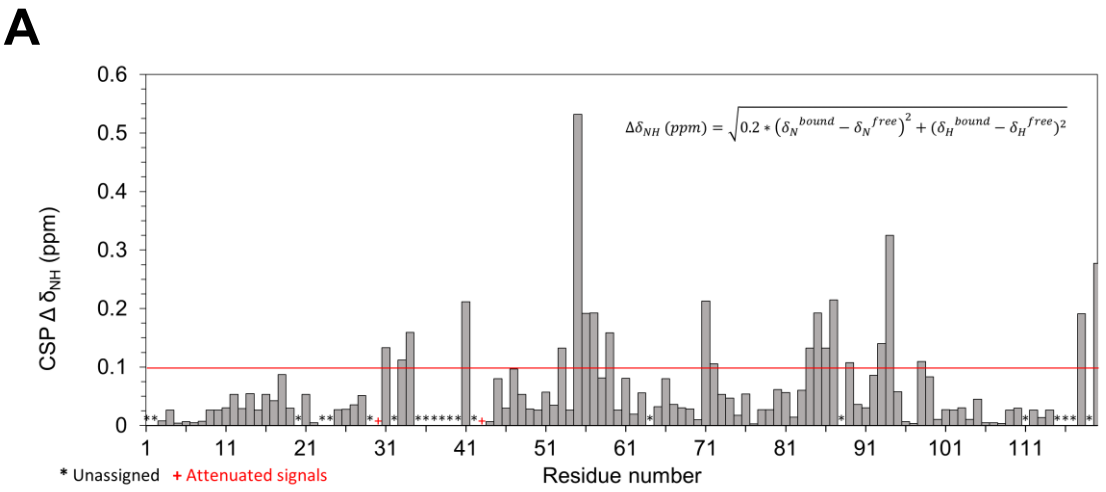
a defined threshold (Δδ > 0.1) were mapped onto the RPA70N structure and compared with the corresponding CSPs caused by the binding of other RPA-interacting proteins; ATRIP, Rad9 and MRE11 (Figure 5.4.A. and B.) (Xu et al., 2008). Similar to these binding partners, RBM-A bound within the basic cleft of RPA70N. Together, these studies establish that RBM-A interacts with RPA via the basic cleft of RPA70N.

5.4.3. Molecular Basis for RBM-A / RPA70N Interaction

To determine the molecular basis for RPA70N binding to the RBM-A site of PrimPol, RPA70N^{E7R} (an RPA70N mutant optimized for crystallization of complexes (Feldkamp et al., 2013) and the RBM-A peptide residues (PrimPol$_{514-528}$) were co-crystallised. Co-crystals contained a 1:1 molar ratio in a $P2_12_12_1$ crystal lattice (Figure 5.3.C. and D.). The statistics for data processing are summarised in Table 5.2. Continuous electron density covers the entirety of RPA70N^{E7R} and 12 residues (514-525) of the 15-mer PrimPol$_{514-528}$ peptide are visible in the electron density maps. Within this short peptide, residues aspartate 519 to leucine 523 are α-helical in content. Given that no α-helices were identified from CD of the free RBD, it is likely that the α-helical peptide identified here is induced upon binding. A striking feature of this α-helix is that the primary interactions with the basic cleft of RPA70N^{E7R} are via salt bridges between aspartate 519 of PrimPol and RPA70N^{E7R} arginines R31 and R43. Hydrogen bonds are also found between isoleucine 517 of PrimPol and RPA70N^{E7R} arginine 43. In addition to the ionic interactions, PrimPol phenyalalnine 522 sits in a hydrophobic pocket made up of RPA70N^{E7R} serine 55, methionine 57 and valine 93. Isoleucine 517 of PrimPol also has an aliphatic interaction with the side chain of RPA70N^{E7R} arginine 91 (Figure 5.3.C. and D.). The combination of these electrostatic and hydrophobic interactions drives the stabilisation of this complex.

5.4.4. PrimPol's RBD Contains a Second RPA Binding Motif

To determine the molecular basis for binding of PrimPol RBM-A to RPA70N, in the wider context of the whole RBD, we co-crystallised a complex of PrimPol$_{480-560}$ bound to RPA70N^{E7R}. Again, co-crystals contained a 1:1 molar ratio in an orthorhombic $P2_12_12_1$ crystal lattice (Figure 5.5.B., C. and E.). The statistics for data processing are summarised in Table 5.2. Similar to the RBM-A peptide, continuous electron density covers the entirety of RPA70N^{E7R} and nine amino acids of an α-helical peptide from PrimPol$_{480-560}$ are visible in the electron density maps. Surprisingly, model building and density refinement revealed that RPA70N bound to PrimPol residues 546-560 (Figure 5.5.A. and B.) rather than RMB-A. An excellent fit to the high-resolution (1.28 Å) electron density is evident for amino acids 548-556, despite residues 480-547 not being visible in



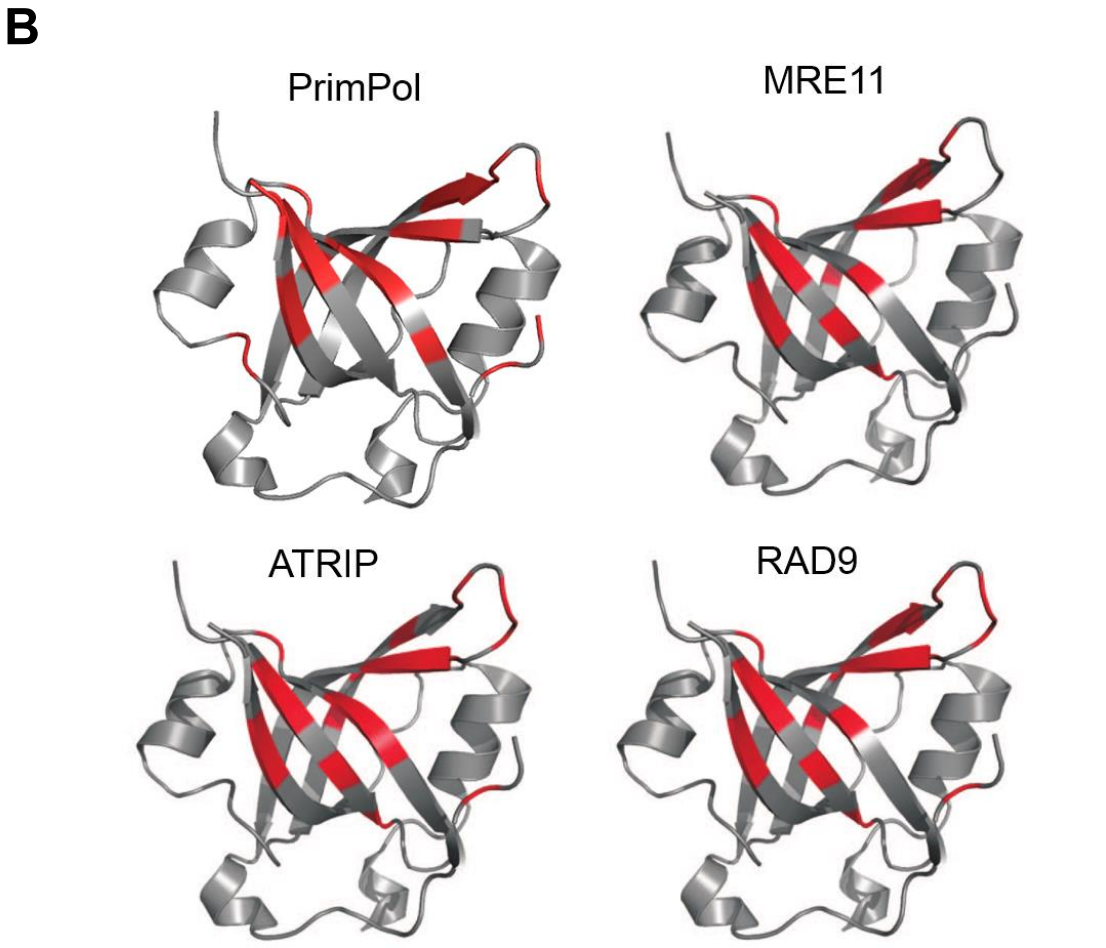


Figure 5.4. PrimPol interacts with RPA70N in the same region as other RPA70N binding partners.
(A) Chemical shift perturbations (CSPs) observed in the ^{15}N-^{1}H HSQC NMR spectra following titration of PrimPol RBM-A peptide quantified and plotted versus RPA70N residue number. **(B)** Structure of RPA70N mapped with CSPs above a Δδ threshold of 0.1 ppm shown in red. The interacting region of PrimPol is similar to that of other binding partners such as ATRIP, RAD9, and MRE11, which bind mostly on one side of the β-barrel.

Data Collection	RPA70N^{E7R}/A-pep	RPA70N^{E7R}/RBD
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	37.86, 53.09, 54.63	38.05, 53.49, 53.9
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Resolution (Å)	31.12 (2.00) *	16.25 (1.28)*
R_{sym} or R_{merge}	0.217 (0.751)	0.044 (0.655)
$I / \sigma I$	10.9 (3.0)	26.9 (2.7)
Completeness (%)	99.6 (98.1)	99.8 (98.2)
Redundancy	12.8 (10.4)	12.1 (7.4)
Refinement		
Resolution (Å)	31.12 (2.00)	16.25 (1.28)
No. reflections	7825	28966
R_{work} / R_{free}	0.1873/0.2286	0.1537/0.1785
No. atoms	1148	1210
Protein	1074	1070
Ligand/ion		
Water	74	140
B-factors		
Protein	25.25	19.73
Ligand/ion		
Water	31.23	32.95
R.m.s. deviations		
Bond lengths (Å)	0.004	0.007
Bond angles (°)	0.785	0.912

*Data from one crystal for each structure. *Values in parentheses are for highest-resolution shell.

Table 5.2. Data collection and refinement statistics (molecular replacement).

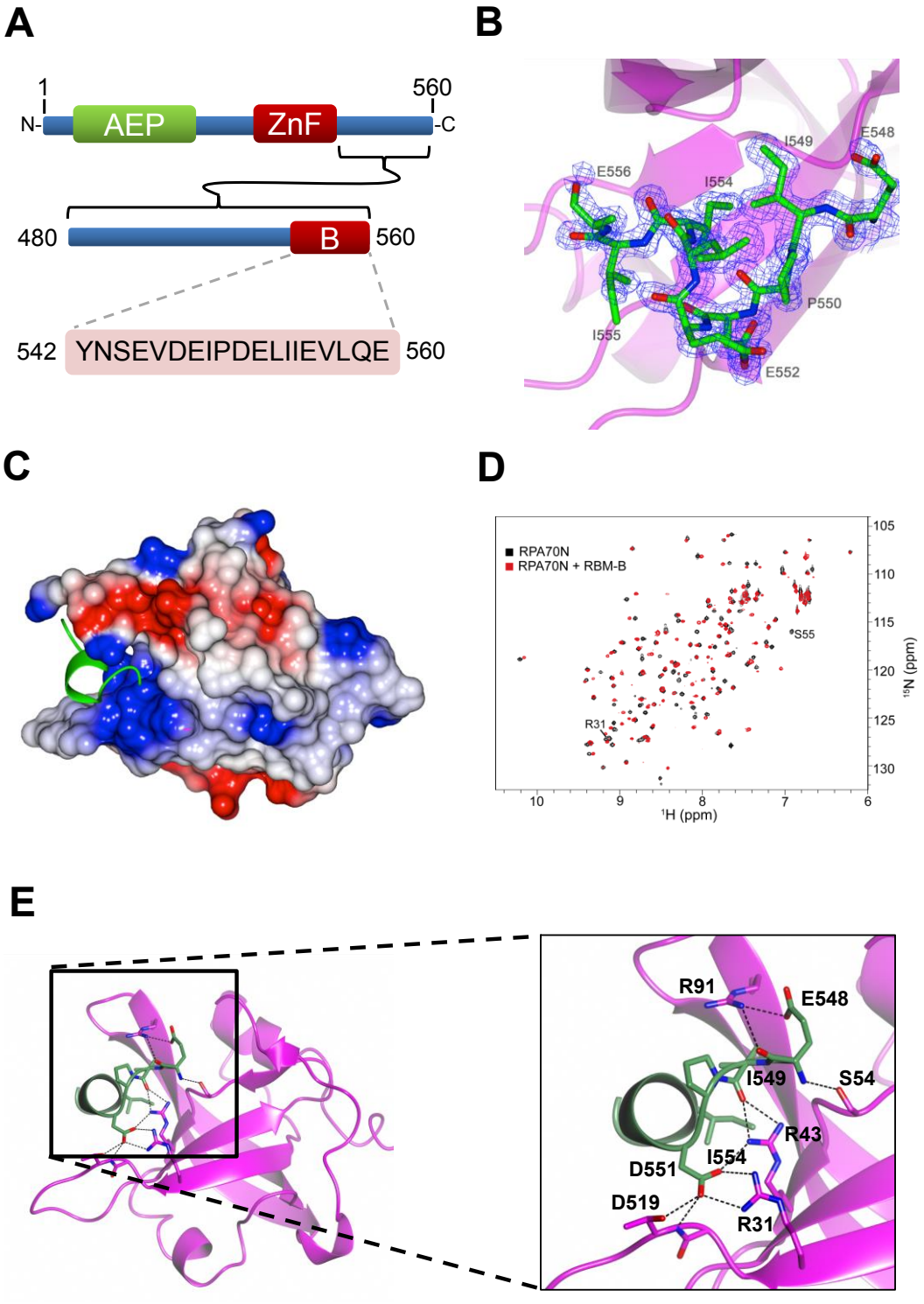


Figure 5.5. PrimPol possesses a second RPA binding motif that also binds to the basic cleft of RPA70N.
(A) The sequence of PrimPol's RBM-B (residues 542-560), located in the C-terminal RBD (residues 480-560). **(B)** The continuous electron density of RBM-B residues 548-556 in the complex with RPA70N. **(C)** Electrostatic surface model of RPA70N with RBM-B (green) bound in the basic cleft. Basic and acidic surfaces are coloured blue and red, respectively. **(D)** ^{15}N-^{1}H HSQC spectra showing RPA70N in the absence (black) or presence (red) of a 2-fold molar excess of unlabelled RBM-B peptide. **(E)** Key stabilising interactions of RBM-B (green) in the RPA70N basic cleft (purple). RBM-B binds between β sheets in the β barrel of RPA70N. D551 is of particular importance as it forms a number of electrostatic interactions with both the side chains and a backbone amide NH of the RPA70N peptide.

the area of contiguous electron density (Figure 5.5.B.). As PrimPol's RBD lacks significant secondary structure or globular fold, and residues 480-547 are not tethered to RPA in the lattice, we expect that these residues remain flexible in the crystal and this disorder inhibits their resolution.

The crystal structure revealed that the second RBM, termed RBM-B, also binds to the basic cleft of RPA70N (Figure 5.5.C.). Like RBM-A, RBM-B has a low pI (pI=3.25) but this motif contains two adjacent Asp-Glu motifs instead of the typical di-Asp motif (Figure 5.5.A.), not previously identified in the RPA70N binding motifs of other RPA partner proteins. To confirm that the interaction observed in the crystal is a *bona fide* RPA70N binding motif, we examined the binding to RPA70N of a $PrimPol_{542-560}$ peptide using ^{15}N-^{1}H HSQC NMR. The spectrum of ^{15}N-enriched RPA70N in the absence and presence of a 2-fold molar excess of the RBM-B peptide reveals significant CSPs induced by the binding of PrimPol RBM-B (Figure 5.5.D.). As observed for the RBM-A titration, the RBM-B peptide causes chemical shift perturbations of residues in RPA70N's basic cleft, including characteristic residues S55 and R31 (Figure 5.5.D.). Together, these data demonstrate that PrimPol's RBD contains a second independent RPA70N binding motif.

5.4.5. Molecular Basis for RBM-B / RPA70N Interaction

Notably, the RBM-A sequence and the structure of its complex with RPA70N is at odds with the well-defined "canonical" RBMs (e.g. p53, ATRIP) and likewise, the structure of RBM-B bound to RPA70N in the crystal of PrimPol RBD confirms these distinct features (Figure 5.3.D., 5.5.E. and Figure 5.6.A.). Notably, these differences arose despite the absence of any significant effects on the structure of RPA70N. The orientation of the RBM-B helix is stabilised by a number of electrostatic interactions (Figure 5.5.E.). The aspartate at position 551 of PrimPol is perhaps the most important point of contact as it interacts with the two arginines of RPA70N (R31 and R43) and a threonine (T34) side chain, as well as the backbone amide N-H of T34. The carbonyl group of PrimPol's isoleucine at position 549 likely acts as a hydrogen bond acceptor for the RPA's R43. The glutamate at position 548 forms an electrostatic interaction with an arginine (R91) on the other side of RPA70N's β-barrel, acting to secure the helix of PrimPol in this orientation. These electrostatic interactions are of paramount importance in the binding of PrimPol's RBM-B to RPA70N *in vitro* (Figure 5.5.B. and E.).

Comparison of the RBM-A and RBM-B structures reveals that the peptides adopt almost identical helical conformations that occupy the basic cleft in a similar fashion (Figure 5.6.A., B. and C.). Intruigingly, the interactions between PrimPol's RBM-A / B and RPA70N are significantly different from the interactions reported for either a modified

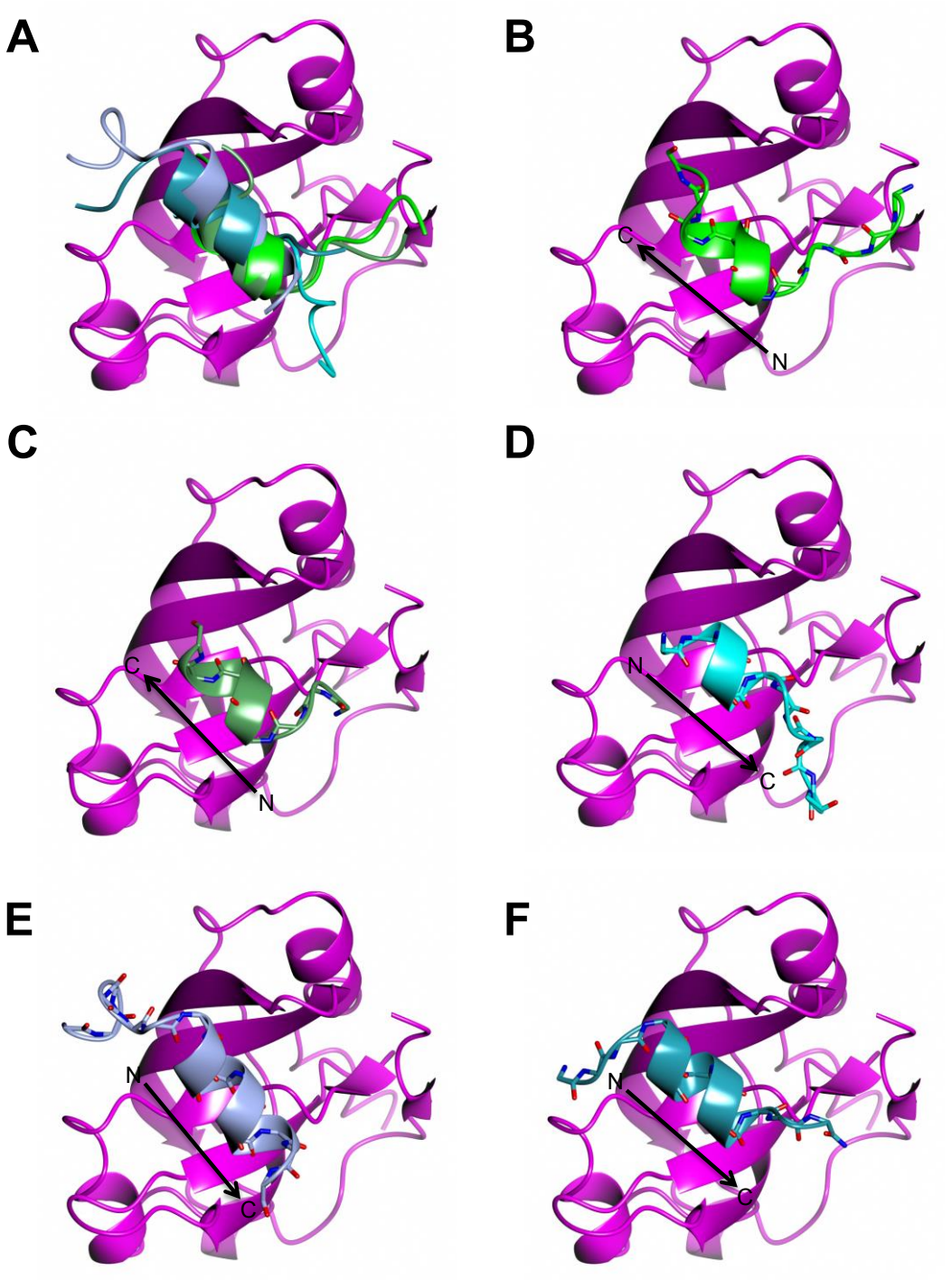


Figure 5.6. RBMs bind with a reverse polarity to the basic cleft of RPA70N.
Ribbon representations of RPA70N^{E7R} from the current PrimPol$_{514-528}$ peptide complex superposed with helical elements from previously established RPA70N protein complexes. The RBM helices are seen to bind in reverse polarity to these established complexes. The RPA70N is coloured magenta. **(A)** All the helices from the different RPA70N complexes superposed: PrimPol$_{514-528}$ peptide complex coloured light green, PrimPol$_{480-560}$ coloured dark green, peptide from Dna2 (PDBID: 5EAY) coloured cyan, p53N (fragment 33-60) (PDBID: 2B3G) coloured sky blue, 3,4 dichlorophenyalanine ATRIP derived peptide (PDBID: 4NB3) coloured turquoise. **(B-F)** Combined ribbon and main chain representations of: **(B)** The current PrimPol$_{514-528}$ peptide complex. **(C)** The PrimPol$_{480-560}$ complex. **(D)** RPA70N binding a peptide from Dna2 (PDBID: 5EAY). **(E)** RPA70N binding p53N (fragment 33-60) (PDBID: 2B3G). **(F)** RPA70N binding 3,4 dichlorophenyalanine ATRIP derived peptide (PDBID: 4NB3).

ATRIP stapled peptide or a p53 peptide bound to RPA70N (Bochkareva et al., 2005; Frank et al., 2014). A superposition of the modified ATRIP peptide with RBMs shows that the two helices bind in a similar region to RPA70N however, they are in opposite orientations (Figure 5.6.). In addition, the main interaction of the modified ATRIP peptide is of its modified 3,4-dichlorophenyl amino acid into a hydrophobic pocket on RPA70N, and in p53 there is a phenylalanine residue that extends into this pocket. This pocket is also the region where a RPA70N binding inhibitor (VUO79104) bound to a co-crystal structure (Feldkamp et al., 2013). PrimPol's RBM-A and RBM-B have hydrophobic residues phenylalanine (F522) and isoleucine (I554) that occupy the hydrophobic pocket on RPA70N (Figure 5.3.D. and 5.5.E.). F522 forms hydrophobic non-bonding contacts with a serine (S55) methionine (M57) and a valine (V93) of RPA70N in this pocket. Whereas, I554 forms hydrophobic non-bonding contacts with the methionine and valine only. We propose that the RPA70N binding modes observed for PrimPol may be more "physiological" as the bound motifs are not modified in any way, unlike p53 and ATRIP where co-crystals could only be obtained by altering the peptides (Frank et al., 2014).

5.4.6. Exchangeable Binding of PrimPol RBMs to RPA70N

As both RBM-A and RBM-B interact in the basic cleft, we next analysed whether these sites bind co-ordinately or competitively. To this end, we constructed RBM-A (D514R/D518R/D519R) and B (480-546 truncation) knock-out (K.O.) mutants in the $PrimPol_{RBD}$ construct. Both NMR and GFC were used to analyse the binding of these mutants to RPA70N. Similar to results observed with $PrimPol_{RBD,}$ $PrimPol_{A\text{-}K.O.}$ and RPA70N eluted together as a well-defined multimeric complex from GFC (Figure 5.7.A.). Additionally, HSQC titrations of 2-fold molar addition of $PrimPol_{A\text{-}K.O.}$ into ^{15}N-enriched RPA70N produced clear evidence of binding (Figure 5.7.B.). Likewise, $PrimPol_{B\text{-}K.O.}$ was found to bind RPA70N in both GFC and NMR analyses (Figure 5.7.C. and D.). Notably, in each GFC analysis a small fraction of unbound RPA70N was observed, unlike GFC using the wild-type RBD. Similarly, NMR analysis produced results mimicking those of the isolated motifs, suggesting that both mutants retain binding activity characteristic of the unaltered RBM-A and B. By overlaying the HSQC spectra of RPA70N in the presence of WT, or mutant RBM-A/B, RBD, we identified that, whilst most of the signals from the complex with mutant RBM-A/B RBD are identical to the complex with WT RBD, (Figure 5.8.A.), some peaks from the RBM-A/B-bound spectra do not overlap. These signals correspond to residues that attenuate or disappear in the complex with WT RBD. Analysis of this phenomenon suggests that RPA70N binds to both sites in solution and this process is exchangeable. This is consistent with ITC data showing that $PrimPol_{A\text{-}K.O}$

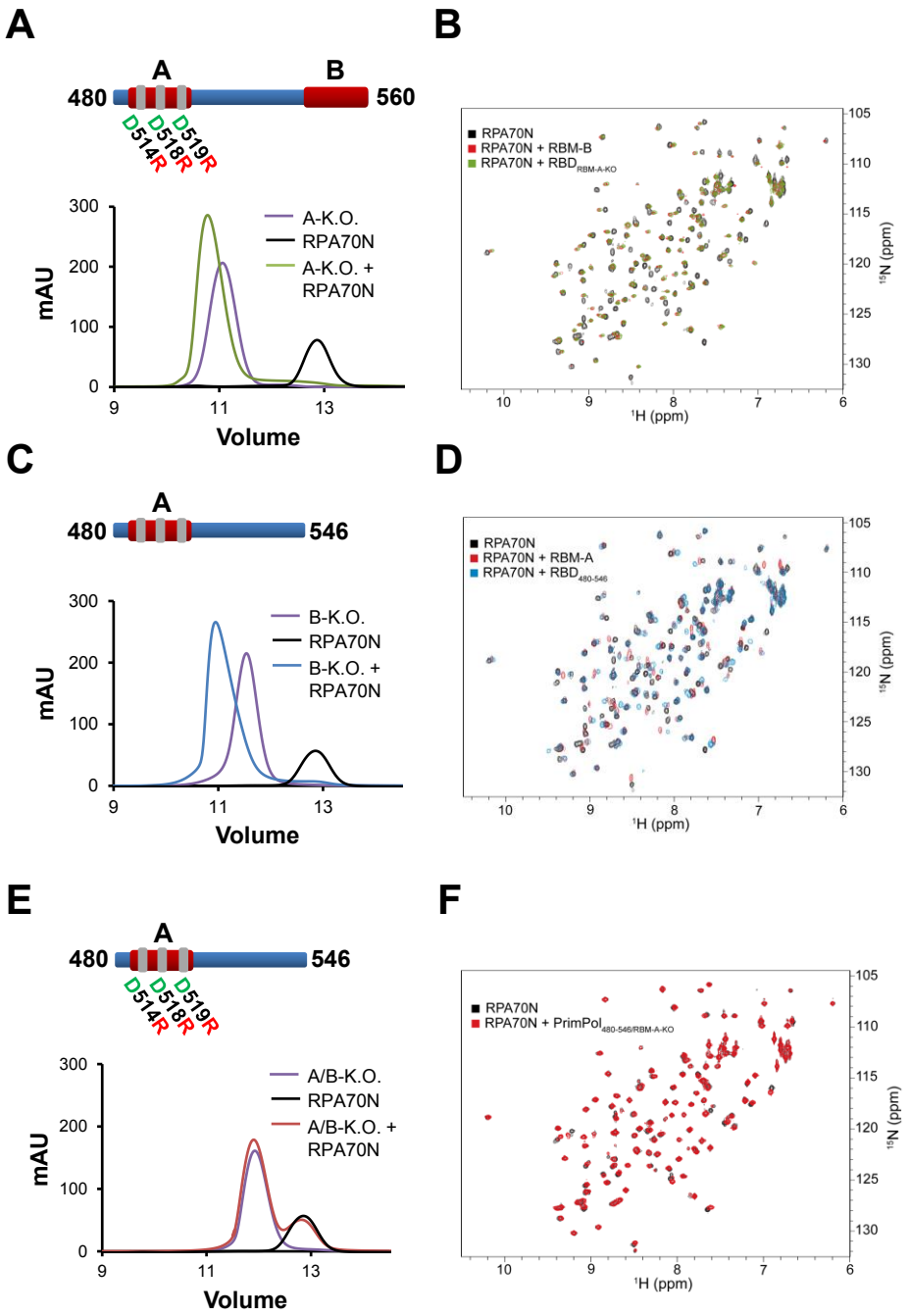


Figure 5.7. RPA70N dynamically interacts with both RBM-A and RBM-B.
(A) Mutation of RBM-A does not abolish binding of PrimPol's RBD to RPA70N. Chromatographs showing the retention volumes of $RBD_{A\text{-}K.O.}$ (purple), RPA70N (black), and $RBD_{A\text{-}K.O.}$ with RPA70N in a 1:1 ratio (green). **(B)** ^{15}N-^{1}H HSQC spectra showing RPA70N alone (black), in the presence of 2-fold molar excess of either $RBD_{A\text{-}K.O.}$ (green) or RBM-B peptide (residues 542-560) (red). The perturbations observed for $RBD_{A\text{-}K.O.}$ are similar to those induced by the RBM-B peptide. **(C)** Truncation of RBM-B does not prevent binding of PrimPol's RBD to RPA70N. Chromatographs showing the retention volumes of $RBD_{B\text{-}K.O.}$ (purple), RPA70N (black), and $RBD_{B\text{-}K.O.}$ with RPA70N in a 1:1 ratio (blue). **(D)** ^{15}N-^{1}H HSQC spectra showing RPA70N alone (black) or in the presence of 2-fold molar excess of $RBD_{B\text{-}K.O.}$ (blue) or RBM-A peptide (residues 510-528) (red). The perturbations observed for $RBD_{B\text{-}K.O.}$ are similar to those induced by the RBM-A peptide. **(E)** Mutation of both RBM-A and RBM-B abolishes the binding of PrimPol's RBD to RPA70N. Chromatographs showing the retention volumes of $RBD_{A/B\text{-}K.O.}$ (purple), RPA70N (black), and $RBD_{A/B\text{-}K.O.}$ with RPA70N in a 1:1 ratio (red). **(F)** ^{15}N-^{1}H HSQC spectra showing RPA70N alone (black) or in the presence of 2-fold molar excess of $RBD_{A/B\text{-}K.O.}$ (red). The near identity of the two spectra indicates there is no interaction.

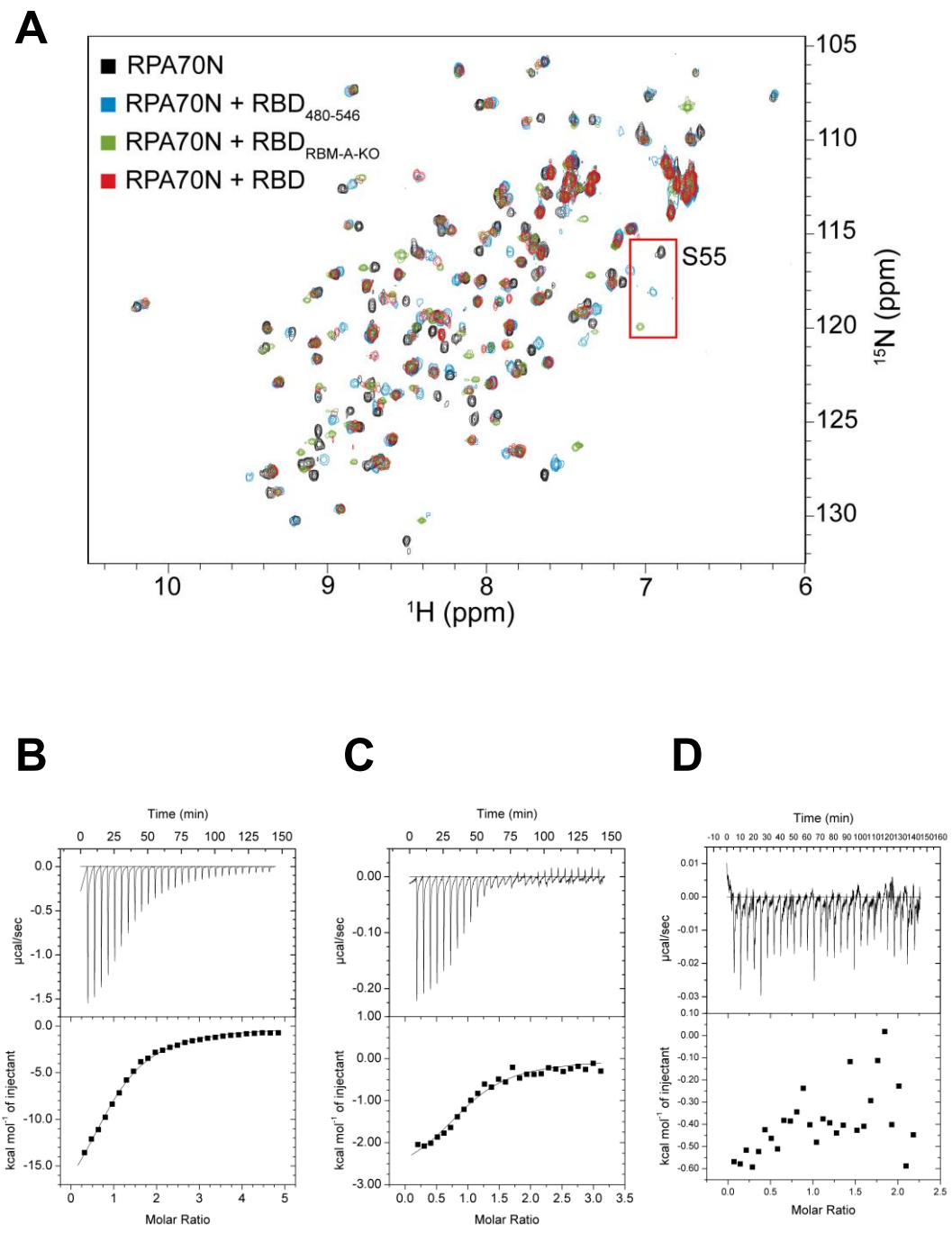


Figure 5.8. RPA70N dynamically interacts with both RBM-A and RBM-B.
(A) ^{15}N-^{1}H HSQC overlay comparing RPA70N bound to wild type RBD or constructs mutated to inhibit binding to RBM-A or B. The RPA70N spectrum (black) shows distinct chemical shift perturbations (CSPs) when titrated with 2-fold molar excess of RBD constructs that select for binding to only RBM-A (blue) or RBM-B (green). These are different from the combination of signal shifting and broadening induced by the wild type RBD (red), which is indicative of dynamic exchange between RBMs. **(B-D)** Isothermal titration calorimetry data showing the heat of binding evolved upon titrating a cell containing 1.4 mL of 20 µM RPA70N with iterative 10 µL injections of 400 µM PrimPolRBD mutants **(B)** PrimPol$_{480-546}$ (B-K.O.), **(C)** RBM-A-K.O. and **(D)** RBM-A/B-K.O. Dissociation constants were calculated with a single site binding model to give statistically equivalent values of 7.8 ± 0.6 µM and 6.7 ± 1.5 µM for the A-K.O. and B-K.O. mutant domains, respectively. The double knockout RBM-A/B-K.O. showed no significant heat evolved upon titration, indicating that no interaction was observed.

and PrimPol$_{B\text{-}K.O}$ bind to RPA70N with statistically identical affinities of 7.8 ± 0.6 μM and 6.7 ± 1.5 μM, respectively (Figure 5.8.A. and B.).

In contrast, there was no observed binding in the GFC or NMR when the “double” mutant (PrimPol$_{A/B\text{-}K.O}$) was incubated with RPA70N (Figure 5.7.E. and F.). Additionally, no heat of binding was observed by ITC (Figure 5.8.C.). Therefore, whilst retaining either one of these domains is sufficient to maintain RPA70N binding *in vitro*, knocking out both RBM-A and RBM-B completely abrogates binding. This indicates that there are no additional RPA70N binding sites beyond RBM-A and RBM-B.

To obtain less perturbing mutants for experiments *in vivo*, we analysed ‘finer’ point mutants of both RBM-A and B, based on the crystallographic data. We found that the PrimPol$_{A\text{-}RA}$ (D519R/F522A) and PrimPol$_{B\text{-}RA}$ (D551R/I554A) mutants retained the ability to bind RPA70N in GFC. However, binding was lost when all four residues were mutated (Figure 5.9.A.). We additionally analysed these mutations in the context of the full-length protein and RBD (480-560) using the yeast two-hybrid assay. Here, PrimPol$_{A\text{-}RA}$ and PrimPol$_{B\text{-}RA}$ exhibited decreased binding to RPA70N, with an additional decrease when both sites were mutated. Near identical results were observed when analysing both the full-length enzyme and RBD, confirming that both RBM-A and RBM-B are able to bind RPA70N when outside their innate vertebrate cell environment (Figure 5.9.B.). These results, therefore, confirmed that each RBM is accessible for RPA70N binding in the context of the full-length protein. Furthermore, they provided minimally perturbing PrimPol variants to probe the functional significance of the RPA interaction, and the contributions of the two RPA-binding motifs, *in vivo*.

5.4.7. RBM-A Mediates the PrimPol-RPA interaction *in vivo*

To ascertain the importance of each RBM in mediating PrimPol’s interaction with RPA *in vivo*, we introduced doxycycline-inducible N-terminal FLAG-tagged PrimPol variants lacking either RBM, or both, into HEK-293 derivative cells (Flp-In T-Rex-293) (Figure 5.10.A. and B.) and performed co-immunoprecipitation experiments. We find that RPA co-precipitates with FLAG-PrimPol *in vivo* when both RBMs are unmodified (Figure 5.10.C.), confirming that FLAG-PrimPol interacts with RPA in a damage-independent manner, as observed previously (Guilliam et al., 2015a; Wan et al., 2013). Additionally, FLAG-PrimPol$_{RBD}$ (the CTD only) also co-precipitates with RPA, supporting our *in vitro* data and previous reports that PrimPol interacts with RPA via its CTD (Figure 5.9.C.) (Guilliam et al., 2015a; Wan et al., 2013). Interestingly, we find that mutation of RBM-A (D519R/F522A) alone abolishes this interaction, despite the protein possessing an intact RBM-B (Figure 5.10.D.). Furthermore, when RBM-B is mutated (D551R/I554A), but

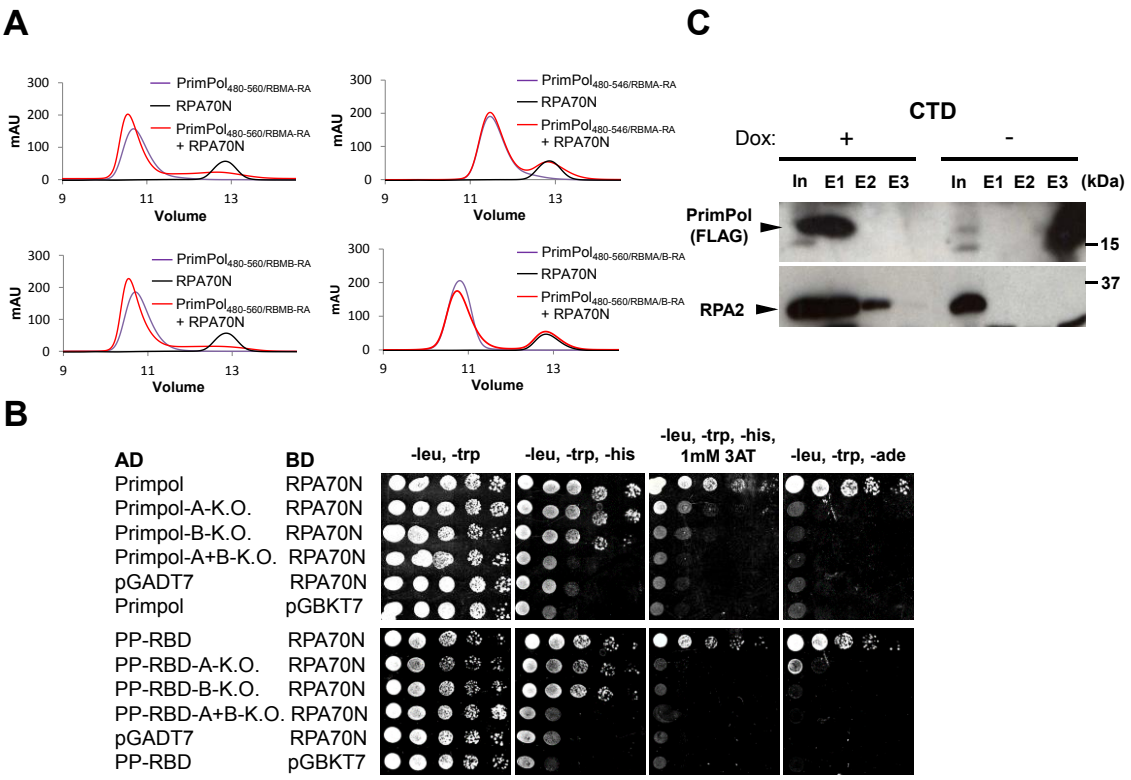


Figure 5.9. PrimPol's RBD interacts with RPA in vivo.
(A) Validation of RBMA/B-RA mutants used for *in vivo* analysis. Chromatographs showing the retention volumes of PrimPol$_{480-560}$ RBMA/B-RA mutants (D519R/F522A and D551R/I554A) in the presence (red) and absence (purple) of RPA70N (black) in a 1:1 ratio. **(B)** PrimPol interacts with RPA70N in the yeast two-hybrid assay. To study the interaction with RPA70N, either full-length PrimPol (upper panels) or its RBD (PP-RBD – a.a. 480-560; lower panels) were used. The following amino acids of PrimPol were mutated - D519R and F522A in A-K.O.; D551R and I554A in B-K.O.; D519R, F522A, D551R, I554A in A+B-K.O. Diploid strains containing plasmids with the indicated genes fused to the GAL4 activation domain (AD) and GAL4 DNA binding domain (BD), were spotted as 10-fold serial dilutions on media lacking leucine, tryptophan, histidine, or adenine. 1mM 3-Amino-1,2,4-triazole (3AT) was added to decrease the background HIS3 expression (panels in 3rd. row). **(C)** RPA co-precipitates with PrimPol's RBD. Flp-In T-Rex-293 cells transfected with FLAG-tagged PrimPol$^{480-560}$ were grown in the presence or absence of doxycycline (10 ng/mL, 24 hrs), FLAG-PrimPol$^{480-560}$ was immunoprecipitated from the soluble cell lysate using anti-FLAG antibody and western blotted for PrimPol (anti-FLAG) and RPA (anti-RPA2). The presence and absence of doxycycline is indicated by +/- Dox, 'In' indicates the input, 'E1', 'E2', and 'E3', indicate elutions 1, 2, and 3, respectively.

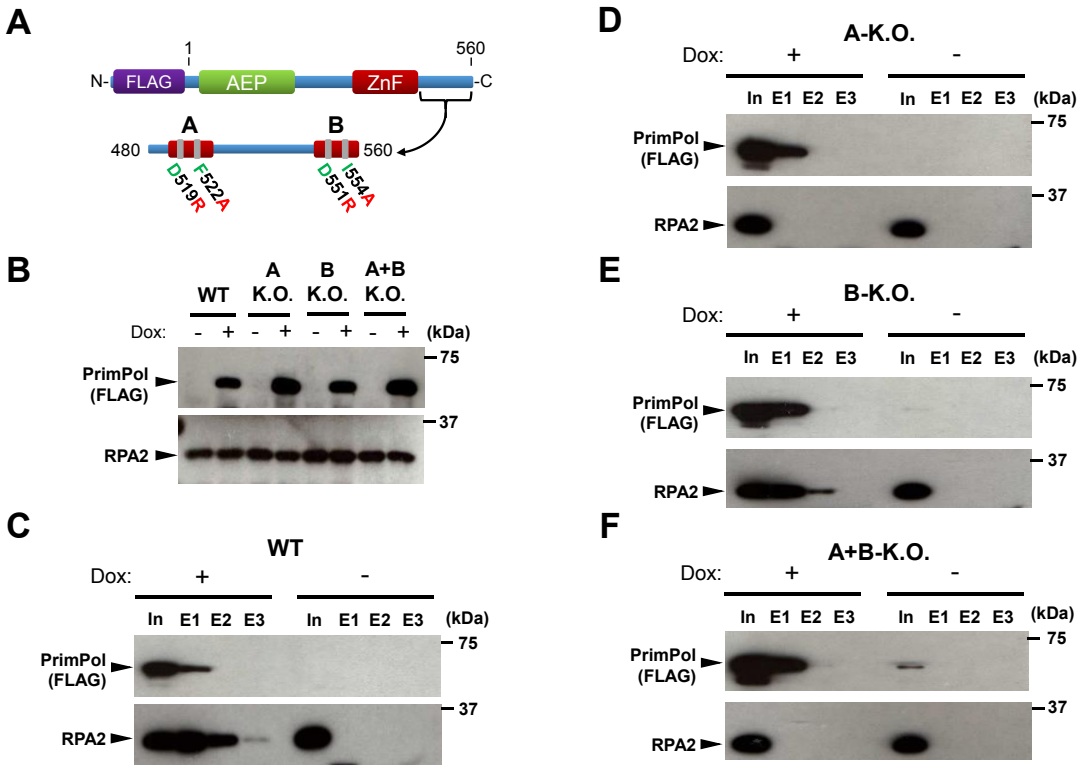


Figure 5.10. PrimPol's RBM-A is critical for RPA binding in vivo.
(A) Schematic detailing the domain architecture of N-terminal FLAG-tagged PrimPol transfected into HEK-293 derivative cells (Flp-In T-Rex-293). The RBD (480-560) containing the RBM-A and B sites is shown below with the mutations forming the A-K.O. (D519R and F522A) and B-K.O. (D551R and I554A) highlighted. **(B)** Flp-In T-Rex-293 cells were transfected with wild-type and RBM-A and B mutated PrimPol. Expression was confirmed by addition of 10 ngmL^{-1} doxycycline (indicated by Dox +/- on figure) for 24 hrs and subsequent western blotting. **(C)** Flp-In T-Rex-293 cells transfected with FLAG-tagged wild-type (WT) PrimPol were grown in the presence or absence of doxycycline (10 ngmL^{-1}, 24 hrs), FLAG-PrimPol was immunoprecipitated from the soluble cell lysate using anti-FLAG antibody and western blotted for PrimPol (anti-FLAG) and RPA (anti-RPA2). The presence and absence of doxycycline is indicated by +/- Dox, 'In' indicates the input, 'E1', 'E2', and 'E3', indicate elutions 1, 2, and 3, respectively. **(D)** Immunoprecipitation of FLAG-PrimPol$_{A\text{-}K.O.}$ (D519R/F522A) from Flp-In T-Rex-293 cells grown in the presence and absence of doxycycline. **(E)** Immunoprecipitation of FLAG-PrimPol$_{B\text{-}K.O.}$ (D551R/I554A) from Flp-In T-Rex-293 cells grown in the presence and absence of doxycycline. **(F)** Immunoprecipitation of FLAG-PrimPol$_{A+B\text{-}K.O.}$ (D519R/F522A and D551R/I554A) from Flp-In T-Rex-293 cells grown in the presence and absence of doxycycline.

RBM-A is intact, a reduced, but significant, amount of RPA still co-precipitates with FLAG-PrimPol (Figure 5.10.E.). Unsurprisingly, when both RBMs are mutated, the interaction with RPA is again lost (Figure 5.10.F.). Together, these findings identify that RBM-A is the primary mediator of PrimPol's interaction with RPA *in vivo* and residues D519 and F522 as essential for forming the complex. In contrast, RBM-B appears to play a more secondary role in RPA binding *in vivo*.

5.4.8. PrimPol's Interaction with RPA is Required for its Role in Replication Restart

PrimPol has previously been shown to promote DNA replication fork restart following UV damage by repriming (Bianchi et al., 2013; Keen et al., 2014b; Mourón et al., 2013). To define the importance of each RBM on PrimPol's role during this process, we complemented PrimPol$^{-/-}$ DT40 cells with RBM-A (D519R/F522A) and RBM-B (D551R/I554A) mutants (Figure 5.11.A.) and performed DNA fibre analysis on these cells in the presence of UV damage. We labelled replicating cells with the nucleotide analogue chlorodeoxyuridine (CldU) for 20 minutes, cells were then UV-C irradiated (20 J/m^2) and labelled with a second nucleotide analogue, iododeoxyuridine (IdU), for an additional 20 minutes (Figure 5.11.B.). Following detection by immunofluorescence, the degree of fork stalling after UV damage in the PrimPol RBM-mutant cells was determined by analysing the CldU:IdU tract length ratios. An increase in this ratio indicates a shorter IdU tract and therefore an increase in the amount of fork stalling or slowing following UV-C irradiation.

Cells expressing RBM-A-mutant PrimPol presented a significant increase in the mean CldU:IdU tract length ratio when compared to cells complemented with wild-type PrimPol (Figure 5.11.C. and D.). Additionally, these cells displayed more variation in CldU:IdU ratios with an increase in the percentage of forks with higher ratios (Figure 5.11.C. and D.). This indicates that there was an increase in fork stalling events, or a decreased ability to restart stalled forks, in these cells. The observed effect was not as severe as that seen in PrimPol$^{-/-}$ cells, however given that RBM-A-mutant PrimPol is catalytically identical to wild-type PrimPol, and over-expressed in these cells, this was not surprising. This result suggests that mutation of RBM-A affects PrimPol's recruitment to stalled replication forks, and therefore causes an impairment in the ability to restart these forks. Given the level of over-expression of RBM-A-mutant PrimPol in these cells, we expect some PrimPol would still localise to where it is required, resulting in a delay rather than total block to fork restart.

In contrast, RBM-B mutant complemented PrimPol$^{-/-}$ cells did not display a significant increase in the mean CldU:IdU ratio when compared to cells expressing wild-type

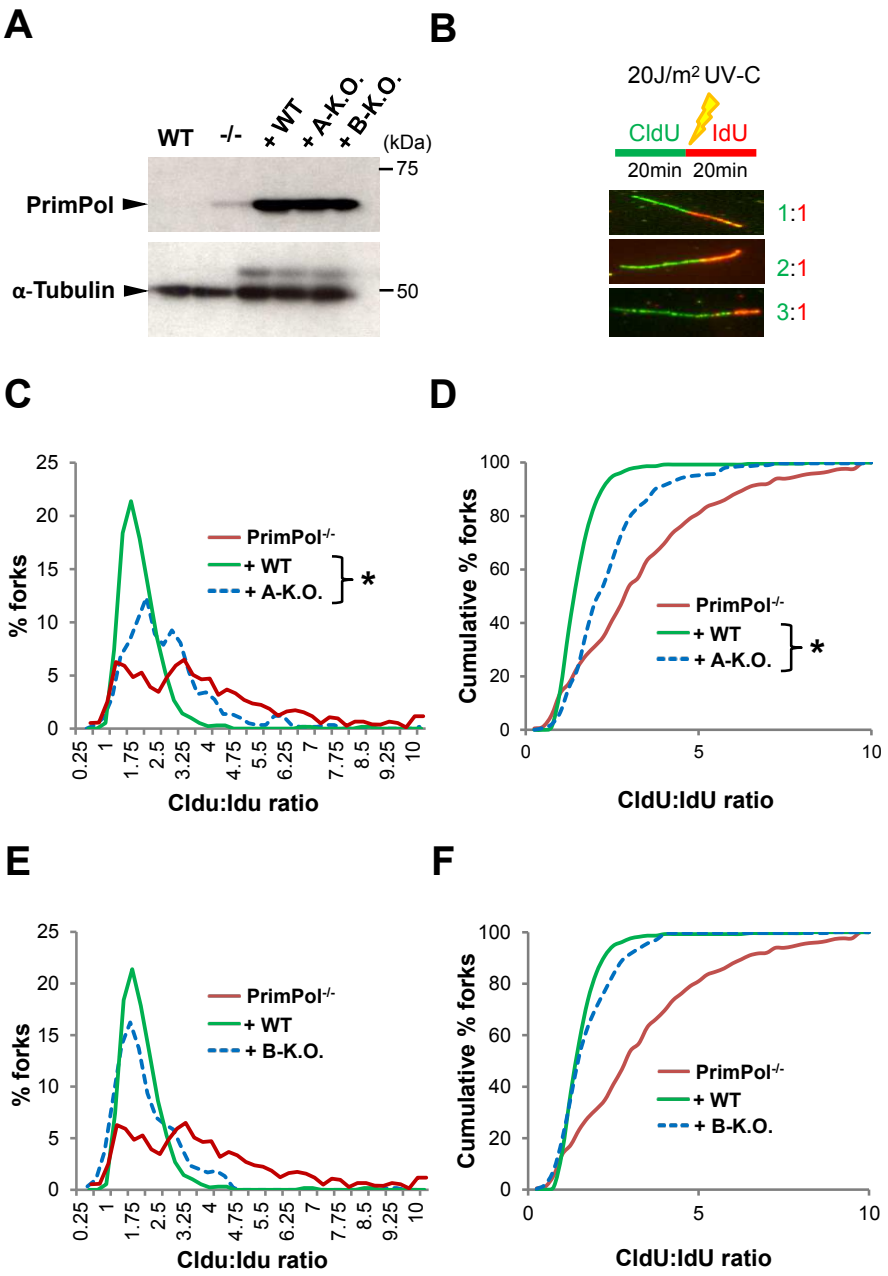


Figure 5.11. RBM-A is required for PrimPol function in DNA replication restart.
(A) PrimPol$^{-/-}$ DT40 cells were complemented with un-tagged human PrimPol constructs; wild-type hPrimPol (+ WT), hPrimPol$_{D519R/F522A}$ (+ A-K.O.), and hPrimPol$_{D551R/I554A}$ (+ B-K.O.). 'WT' indicates lysate from wild-type DT40 cells, '-/-' indicates lysate from PrimPol$^{-/-}$ DT40 cells. **(B)** DNA fibre analysis was performed on DT40 cells expressing each PrimPol construct. Cells were UV-C irradiated (20 Jm^{-2}) between the CldU and IdU labelling periods (each 20 mins). Representative DNA fibres showing 1:1, 2:1, and 3:1 CldU:IdU ratios are presented; >100 individual DNA fibres were scored for each experiment. **(C)** Mutation of RBM-A causes increased fork stalling following UV-C irradiation. Data are representative of the means of three individual experiments and were subject to an unpaired *t*-test showing a significant difference between the mean CldU/IdU ratio for the '+ WT hPrimPol' and '+ A-K.O. hPrimPol' data sets ($P < 0.05$). **(D)** DNA fibre analysis from the '+ A-K.O. hPrimPol' DT40 cells presented as a cumulative percentage of forks at each ratio. **(E)** Mutation of RBM-B does not significantly alter the level of fork stalling following UV-C irradiation. DNA fibre analysis of the '+ B-K.O. hPrimPol' DT40 cells, showing the percentage of forks at each CldU:IdU ratio. Data are representative of the means of three individual experiments. **(F)** DNA fibre analysis from the '+ B-K.O. hPrimPol' DT40 cells presented as a cumulative percentage of forks at each ratio.

PrimPol (Figure 5.11.E. and F.). There was a slight increase in the variation of CldU:IdU ratios, however the majority of forks conformed to wild-type ratios (Figure 5.11.E.). Again, given that PrimPol is over-expressed in these cells, a more significant effect may be observed upon mutation of the endogenous protein, with over-expression potentially masking subtle impacts on PrimPol recruitment. Nevertheless, this suggests that RBM-B is not essential for PrimPol's role in replication restart *in vivo*. Together, these results show that PrimPol's interaction with RPA, primarily mediated by RBM-A, is important for the enzyme's role in repriming and restarting stalled replication forks following DNA damage.

5.4.9. RBM-A is Essential for Recruitment of PrimPol to Chromatin

We previously reported that PrimPol is recruited to chromatin in response to UV damage (Bianchi et al., 2013). Given the effect of mutating PrimPol's RBM-A on the enzyme's role in replication restart, we aimed to confirm if this was due to a defect in recruitment. To this end, we prepared detergent-insoluble chromatin-rich fractions from HEK-293 cells, expressing RBM-mutant PrimPol constructs, 3 hours following mock or UV-C irradiation (30 J/m^2). As previously observed, wild-type PrimPol partitioned to the detergent-insoluble chromatin-enriched fraction following UV irradiation (Figure 5.12.A and 5.13.A. and B.). A similar increase in the level of RPA enrichment was observed in the insoluble fraction, confirming that replication forks were stalled by the damage and an increase in RPA binding had occurred (Figure 5.12.A.). In contrast, we find that mutation of RBM-A, either alone or in combination with RBM-B, abolishes the localisation of PrimPol to chromatin, both in the absence or presence of UV damage. Mutation of RBM-B, however, did not affect the level of enrichment of PrimPol following UV irradiation (Figure 5.12.A.). This suggests that PrimPol's recruitment to chromatin is dependent upon its RBM-A, which is the primary mediator of the interaction of the enzyme with RPA *in vivo.*

To confirm these findings and examine the role played by the RBD of PrimPol in the recruitment of the protein to replicating chromatin, we employed a Xenopus synchronous cell-free extract system. We previously showed that recombinant hPrimPol accumulates on chromatin when the elongation phase of DNA replication is inhibited with aphidicolin (Bianchi et al., 2013). Similarly the presence of PrimPol's RBD (480-560) is sufficient to allow recruitment of a GST fusion protein to chromatin in aphidicolin-treated extracts (Figure 5.12.B.). RBD recruitment is severely reduced by mutation of the D519 and F522 residues within RBM-A. Mutation of the corresponding residues in RBM-B (D551, I554) also results in a modest reduction in the level of protein recruited to the chromatin,

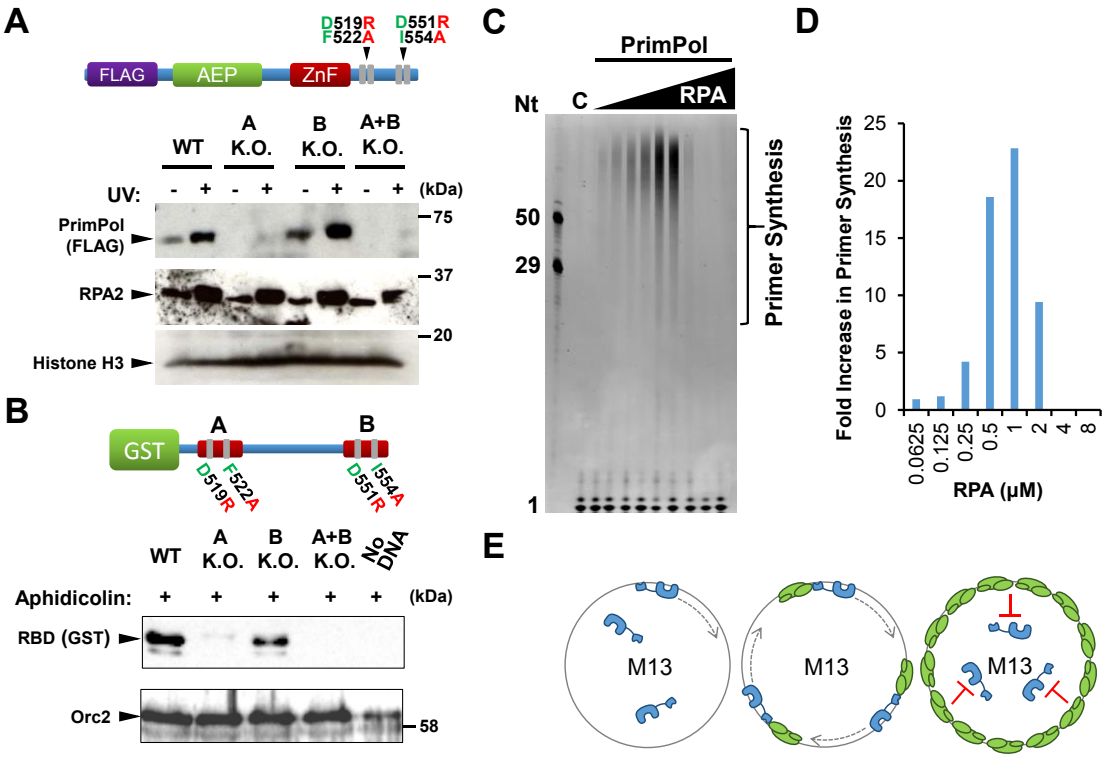


Figure 5.12. RPA recruits PrimPol to stalled replication forks in vivo.
(A) PrimPol's RBM-A, but not RBM-B, is critical for recruitment to chromatin. Flp-In T-Rex-293 cells transfected with WT and RBM-A and B mutant PrimPol constructs were either mock (-) or UV-C (30 Jm^{-2}) (+) irradiated before separation into Triton X-100 (0.5%) soluble and insoluble fractions. Samples were analysed by western blot alongside whole-cell extracts (WCE). Only insoluble samples are presented here, WCE and soluble blots can be found in Figure 5.13. **(B)** PrimPol's RBD is recruited to *Xenopus* egg extract chromatin in response to aphidicolin treatment, RBM-A is critical for this recruitment. Recombinant hPrimPol-GST constructs (4 $ngμl^{-1}$) were added to Xenopus egg extract supplemented with sperm nuclei (3 x $10^3 μl^{-1}$). Extract was treated with aphidicolin 100 $μgml^{-1}$ and incubated at 21 °C for 80 minutes. Chromatin was isolated and associated proteins analysed by SDS-PAGE and western blotting using the antibodies indicated. **(C)** Low concentrations of RPA stimulate PrimPol's primase activity, high concentrations inhibit. Primer synthesis by WT hPrimPol (400 nM) on M13 ssDNA templates (20 $ngμL^{-1}$) in the presence of increasing concentrations of RPA. 'C' indicates the no enzyme control, oligonucleotide (Nt) length markers are shown on the left of the gel. **(D)** Quantification of data shown in 'c'. For each RPA concentration the fold increase in primer synthesis relative to reactions containing no RPA was calculated. Data are representative of three repeat experiments. **(E)** Schematic showing the effect of increasing RPA concentrations on PrimPol's primase activity. When no RPA is present a proportion of PrimPol binds to the M13 template and facilitates primer synthesis (left). When low/moderate concentrations of RPA are present PrimPol is recruited to the RPA/ssDNA interface causing an increase in primer synthesis activity (middle). At high RPA concentrations the M13 DNA template is fully saturated, blocking access of PrimPol to the DNA and inhibiting primer synthesis (right).

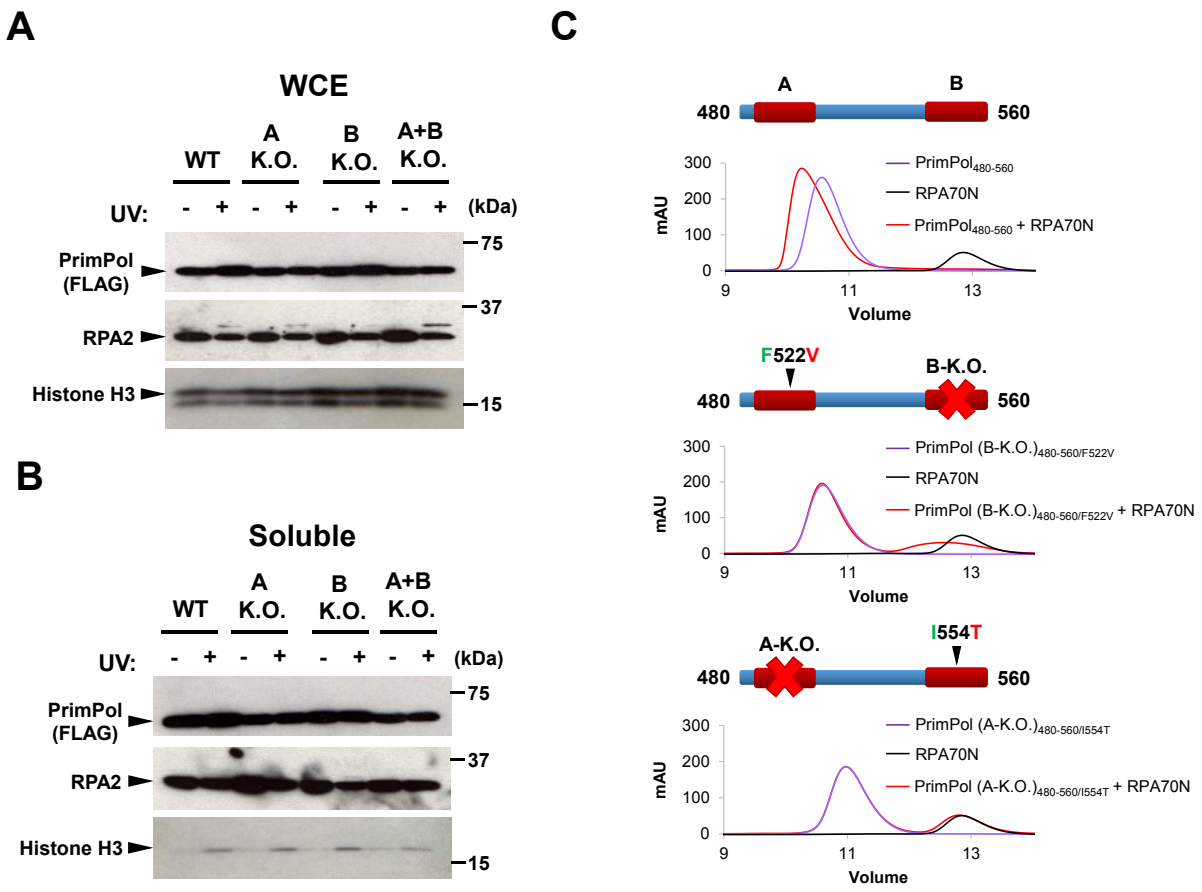


Figure 5.13. Whole-cell extracts and soluble samples from PrimPol chromatin recruitment experiments.
(A) Whole cell extract samples corresponding to the chromatin extraction experiment presented in Figure 5.12.A. **(B)** Soluble samples from the chromatin extraction experiment presented in Figure 5.12.A. **(C)** Analysis of the effect of PrimPol RBM mutations F522V and I554T, identified in cancer patient cell lines, on the RPA70N interaction. Chromatographs showing the retention volumes of unmodified PrimPol$_{480-560}$, PrimPol$_{480-560}$ F522V (on a B-K.O. construct), and PrimPol$_{480-560}$ I554T (on an A-K.O. construct), in the presence (red) and absence (purple) of RPA70N (black) in a 1:1 ratio. Unmodified PrimPol$_{480-560}$ is able to bind RPA70N (top panel). However, introduction of the mutations F522V (middle panel) or I554T (bottom) panel, in the absence of a second functional RBM, significantly abrogates binding to RPA70N.

although this reduction is much less severe than that observed with the RBM-A mutations. Consistent with these observations, a construct carrying mutations in both RBM-A and RBM-B is not detectable on the chromatin. These results demonstrate that RBM-A plays the major role in recruiting PrimPol to chromatin, with a relatively minor contribution from RBM-B.

Intriguingly, some of the key residues involved in binding of both RBM-A and RBM-B to RPA70N have been found to be mutated (F522V and I554T) in cancer patient cell lines (see COSMIC, CBioportal, CIGC repositories). We therefore generated these cancer-related PrimPol RBD mutants (F522V and I554T) in RBM-B K.O. and RBM-A K.O. backgrounds, respectively, and analysed their binding to RPA70N using GFC (Figure 5.13.C.). In each case, we identify that these mutations significantly abrogate binding of the affected RBM to RPA70N, potentially suggesting that both sites play an important role in appropriate PrimPol function *in vivo*.

5.4.10. RPA Stimulates the Primase Activity of PrimPol

In light of the role for RPA in recruiting PrimPol to stalled replication forks *in vivo*, we next assessed the impact of RPA on the primase activity of the enzyme *in vitro*. Using an indirect fluorescent primase assay, we previously identified that saturating concentrations of RPA are able to block primer synthesis by PrimPol on 60-mer poly-dT linear templates (Guilliam et al., 2015a). To better determine the effect of RPA on PrimPol's primase activity, we performed direct fluorescent primase assays using single-stranded M13 templates in the presence of increasing concentrations of RPA. Here, we observe that sub-saturating concentrations of RPA act to significantly increase the amount of primer-synthesis by PrimPol, when compared to reactions containing the enzyme only (Figure 5.12.C. and D.). Above concentrations calculated to fully coat the M13 template (~1.6 μM), the level of stimulation by RPA decreases and at higher concentrations severely inhibits primer synthesis (Figure 5.12.C. and D.). This demonstrates that lower concentrations of RPA significantly stimulate the primase activity of PrimPol, presumably by recruiting the enzyme and mediating binding to the DNA template. In contrast, high concentrations of RPA saturate the DNA template and block access of PrimPol, thus inhibiting primase activity (Figure 5.12.E.). These results suggest that PrimPol requires a ssDNA interface adjacent to the bound RPA in order to be recruited and facilitate primer synthesis.

5.5. Discussion

Despite possessing the ability to perform TLS, recent studies suggest that PrimPol's primary role in replication restart is to reprime downstream of DNA damage lesions and secondary structures (Keen et al., 2014b; Kobayashi et al., 2016; Mourón et al., 2013; Schiavone et al., 2016). The data presented here support a model whereby PrimPol is recruited to fulfil this repriming role through its interaction with RPA (Figure 5.14.). This interaction is primarily mediated by residues D519 and F522 of PrimPol's RBM-A, which bind to the basic cleft of RPA70N, with RBM-B potentially playing a secondary role in RPA binding *in vivo*. In this regard, an intriguing possibility, consistent with our findings, is that RBM-B binds a second RPA molecule following initial recruitment through RBM-A *in vivo*, potentially contributing to the stabilisation of PrimPol on the template DNA to further promote repriming. In addition to ATRIP, Mre11 and p53, we identified divergent RBM-like acidic motifs in a wide range of other DNA repair, replication and checkpoint proteins, many of which are known to interact with RPA e.g. Werner helicase (Figure 5.15.) (Oakley and Patrick, 2010).

Notably, it has been shown through crystallographic and biochemical analyses that RPA binds to ssDNA with a defined polarity (Fan and Pavletich, 2012; Iftode and Borowiec, 2000; Kolpashchikov et al., 2001; Laat et al., 1998). Initial binding is mediated by the tandem DNA-binding domain A (DBD-A) and DBD-B OB folds of RPA70, forming an 8-nt binding complex. A 20-30-nt binding mode is subsequently generated by the binding of RPA's DBD-C and DBD-D (Brosey et al., 2013). This occurs in a strict 5'-3' direction on the template strand, which likely positions the PrimPol-recruiting RPA70N domain 5' relative to the other OB-folds (Figure 5.14.A.). This polarity suggests that PrimPol may bind downstream of RPA following recruitment through RPA70N on the leading strand. In a previous scenario (Guilliam et al., 2015a), we speculated that PrimPol may bind upstream of RPA during TLS, due to the requirement of the ssDNA-binding ZnF domain to contact the template downstream. However, during primer-synthesis the ZnF domain can access ssDNA both upstream and downstream of the AEP domain. Recent studies highlighting the importance of PrimPol's primase activity *in vivo* (Keen et al., 2014b; Kobayashi et al., 2016; Mourón et al., 2013; Schiavone et al., 2016), coupled with the recruitment of the enzyme via RPA70N shown here, argue that PrimPol more likely binds downstream of RPA, with the ZnF bound to ssDNA upstream of the AEP domain, during primer synthesis (Figure 5.14.B.).

PrimPol displays low processivity, only extending primers 1-5 nt in a single binding event (Keen et al., 2014b). This processivity is in part regulated by the ZnF domain, which

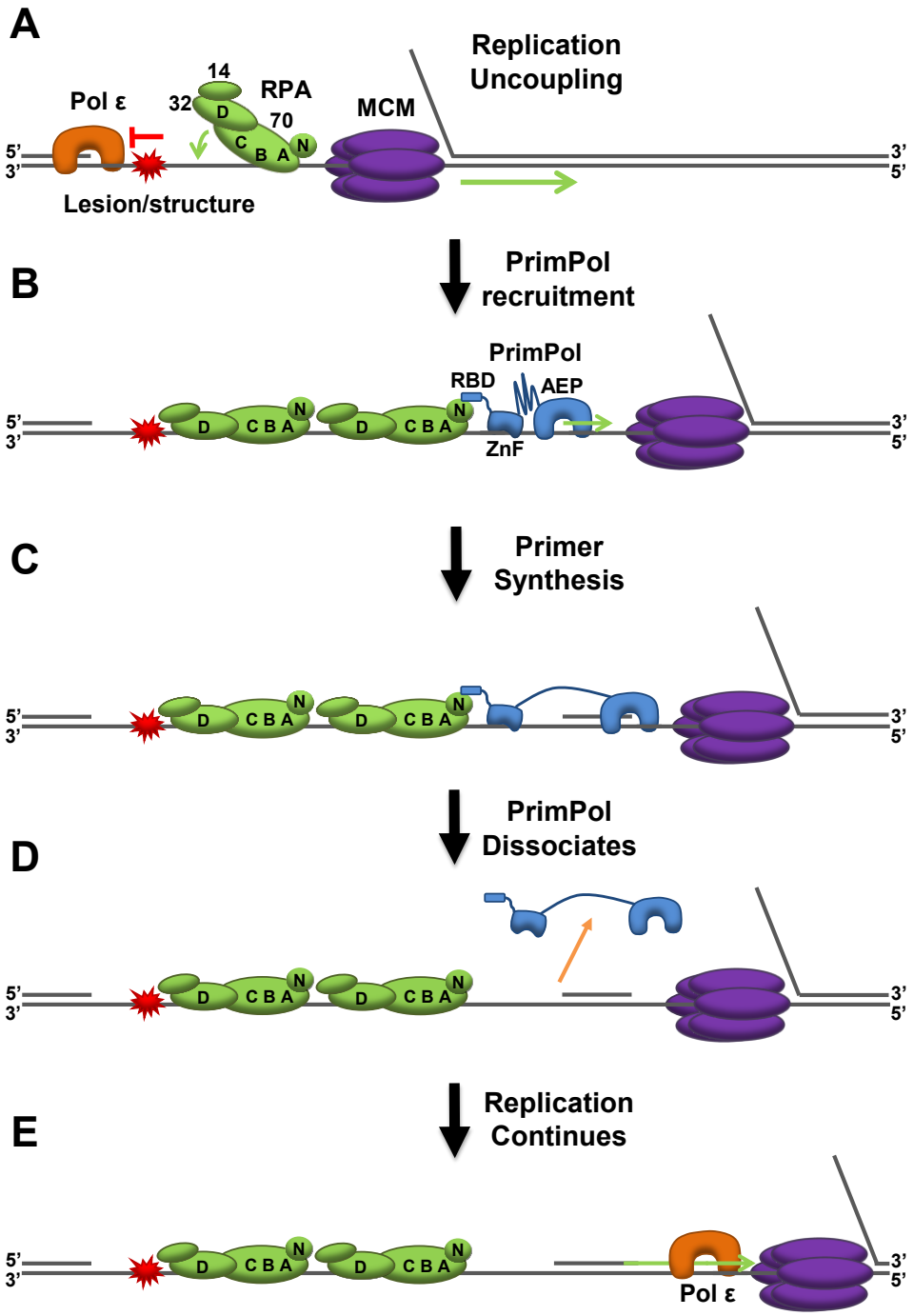


Figure 5.14. Model for PrimPol recruitment to stalled replication forks by RPA.
(A) Unrepaired DNA damage lesions, or DNA secondary structures, in the leading strand template lead to stalling of polymerase ε. This causes uncoupling of replication, generating ssDNA downstream of the DNA damage lesion/structure and facilitating binding of RPA. Note that for simplicity other replisome components and lagging strand synthesis machinery are not shown. **(B)** PrimPol is recruited to the ssDNA interface uncovered by the replicative helicase through the interaction of its RBM-A with RPA70N. This interaction is stabilised by the binding of the ZnF and AEP domains to ssDNA. **(C)** PrimPol catalyses the synthesis of a new DNA primer, before further extension is prevented by the restraining effect of the RPA-interaction and ZnF domain, coupled with the enzyme's low processivity. **(D)** Unable to continue with primer extension, PrimPol dissociates from the template strand. Re-binding upstream is prevented by RPA. **(E)** The nascent primer is utilised by the replicative polymerase for continued DNA replication. This leaves behind a short RPA-coated ssDNA region opposite the lesion to be filled in by template switching or TLS.

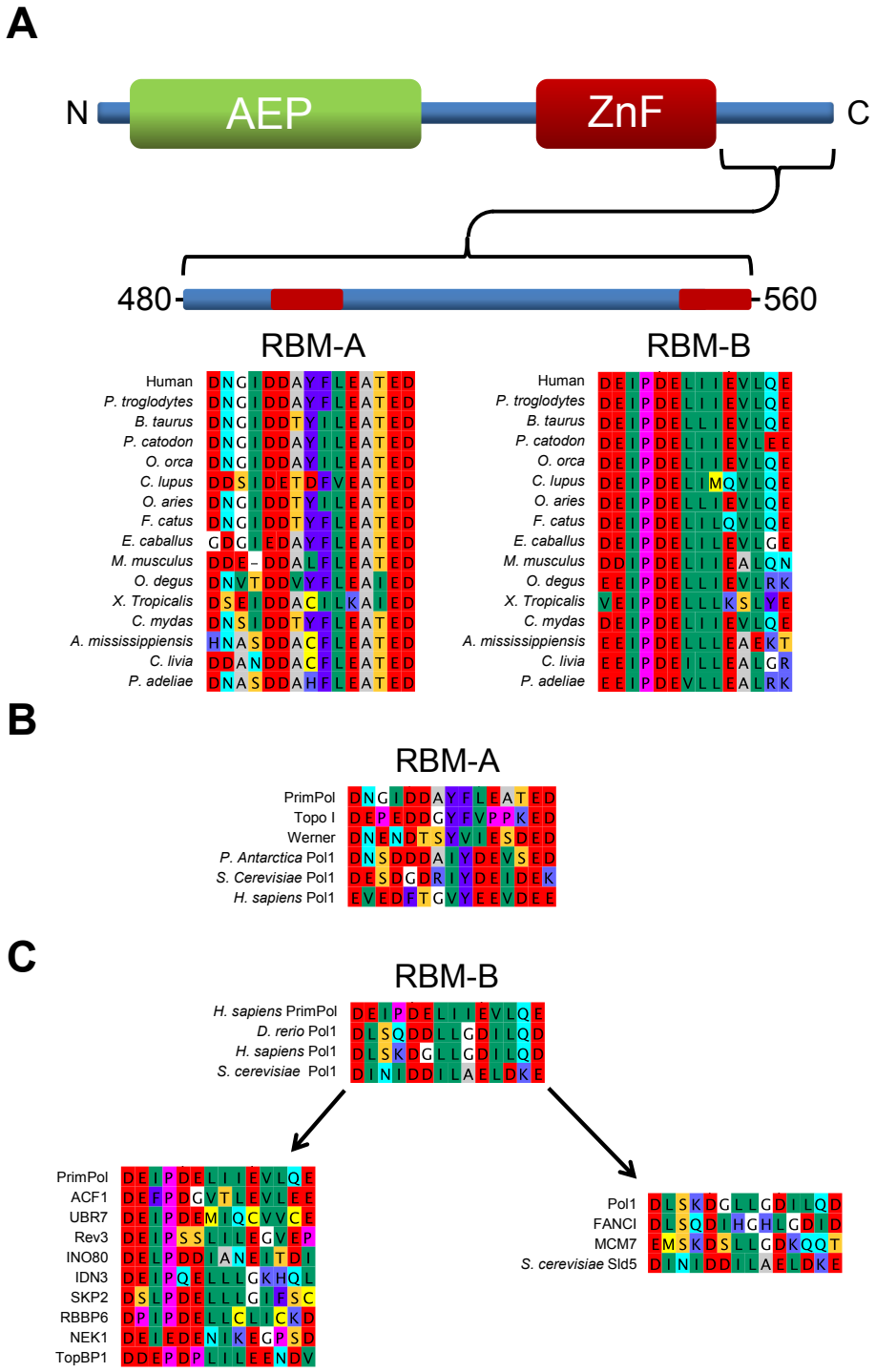


Figure 5.15. RBM-A and RBM-B represent common protein interaction motifs.
(A) The domain architecture of PrimPol showing the positions of RBM-A and RBM-B relative to the functional domains of the protein. RBM-A and RBM-B motifs are conserved across chordate species. **(B)** Sequences similar to that of RBM-A are identified in Topoisomerase I and the Werner helicase. RBM-A of PrimPol aligns to a short acidic sequence in the N-terminus catalytic subunit of Pol α of *Pseudozyma antarctica*. A comparable sequence is found in the human homologue suggesting that this may be an RPA-interacting region of Pol α (Pol 1). **(C)** PrimPol RBM-B resembles the Ctf4(AND1)-binding motif. The Ctf4-binding motif that has previously been identified in the catalytic polymerase subunit of Pol α and Sld5 of GINS is present in PrimPol at the C-terminus in RBM-B. This region of PrimPol interacts with RPA70N and represents a potential multiple protein binding motif. A PSI-BLAST of this motif from human PrimPol and Pol α identifies a number of other proteins involved in the metabolism of DNA.

serves as a ‘counting mechanism’ to limit primer extension by the AEP domain (Keen et al., 2014b), as has been observed with other primases (Kuchta and Stengel, 2010a). The ZnF and AEP domains therefore likely form a hinge-like structure with the ZnF domain limiting extension by PrimPol following initial primer synthesis (Keen et al., 2014b; Kuchta and Stengel, 2010a). Given that PrimPol is recruited by RPA *in vivo*, it is likely that the enzyme initially binds the ssDNA downstream of RPA in a ‘closed hinge’ mode (Figure 5.14.B.). Primer synthesis and polymerisation then proceed until PrimPol reaches its maximum open conformation, dictated by the ZnF domain and interaction with RPA, which thereby prohibit further extension (Figure 5.14.C. and D.). The newly synthesised primer can then be utilised by the replicative polymerase for continued extension (Figure 5.14.E.).

We previously reported that, in contrast to the effect on replicative polymerases, RPA acts to inhibit the polymerase activity of PrimPol (Guilliam et al., 2015a). This phenomenon may be explained by the polarity of RPA when bound to ssDNA, in addition to the protein’s interaction with PrimPol. It has been suggested that replicative polymerases are able to readily displace RPA from DNA because they encounter the protein from the 3’ side (Iftode et al., 1999). As the replicative polymerase synthesises DNA, moving 3’-5’ on the template strand, it would first encounter the relatively weakly bound RPA32 DBD-D and RPA70 DBD-C, before RPA70 DBD-B and A. By approaching RPA in this orientation and making specific protein-protein interactions, the replicative polymerase may shift the equilibrium from the stronger 20-30-nt RPA binding mode to the weaker 8-nt mode (Brosey et al., 2013), and in turn, the more weakly bound RPA can be displaced by further DNA synthesis. In contrast, recruitment of PrimPol to the 5’ side of RPA would result in the enzyme moving away from the protein, making it unable to displace RPA in the same manner as canonical polymerases.

In addition, we show that RPA stimulates the primase activity of PrimPol at sub-saturating concentrations. However, when the template is fully coated with RPA the primase activity of PrimPol is inhibited. This suggests that PrimPol requires a ssDNA interface adjacent to RPA to be efficiently recruited for priming. Given that PrimPol likely binds downstream of RPA on the leading strand during replication, this ssDNA interface could be formed following uncoupling of leading and lagging strand replication upon stalling at a DNA lesion or secondary structure (Lopes et al., 2006). Continued unwinding of duplex DNA by MCM may generate the leading strand ssDNA interface necessary for PrimPol to reprime, following recruitment by RPA. It was recently reported that the mitochondrial replicative helicase Twinkle is able to stimulate DNA synthesis by PrimPol (Stojkovič et al., 2016), potentially suggesting that replicative helicases can facilitate

synthesis by the enzyme *in vivo.* However, the exact interplay between RPA, PrimPol, and other PrimPol-interacting partners, such as PolDIP2 requires further examination (Guilliam et al., 2016). The necessity of a ssDNA interface for PrimPol activity, in conjunction with the enzyme's inability to displace RPA, may act as an important regulatory mechanism to prevent un-scheduled primer synthesis during replication. Intriguingly, it has been hypothesised that recruitment of DNA damage response proteins to RPA70N may be regulated by phosphorylation of RPA32C (Oakley and Patrick, 2010). In support of this, it has been shown that binding of Mre11 and Rad9 to RPA is increased upon RPA32C phosphorylation (Robison et al., 2004; Wu et al., 2005), it remains to be determined if this is the case for PrimPol.

Together, the findings presented here describe the molecular basis of PrimPol's interaction with RPA and provide insights into its biological roles. We found that the PrimPol-RPA interaction, mediated primarily by RBM-A, is essential for PrimPol recruitment and its function as a repriming enzyme during DNA replication. Notably, mutations of critical residues in both RBMs have been identified in the genomes of some cancer patient cell lines and we have shown that these mutations are sufficient to abrogate the functionality of their respective RBMs. Further studies are underway to more precisely define how PrimPol is recruited to stalled replication forks and regulated by interactions with other replisome components to better understand the critical roles of PrimPol in the restart of stalled replication forks.

5.6. Further Work

5.6.1. Initial Investigations into the Polarity of RPA-Mediated PrimPol Recruitment and Repriming

An important question to arise from the work described in the above article is, which side of RPA does PrimPol get recruited to and reprime on? This question has important mechanistic implications for repriming. If PrimPol is recruited to the 3' side of the bound RPA molecule on the ssDNA template, repriming would likely occur close to the replication impediment and stalled replicase. Whereas, repriming on the 5' side of RPA would permit replication restart close to the replicative helicase, preventing a situation where the replicase has to "catch-up" with the CMG complex following leading/lagging strand replication uncoupling and repriming. Moreover, each scenario has implications for the domain organisation of PrimPol during repriming, due to the requirement of the ZnF to contact the ssDNA template (Keen et al., 2014b). Recruitment to the 3' side of RPA would require the ZnF to bind downstream of the AEP domain, with initial binding

of PrimPol in an “open” conformation. In comparison, recruitment to the 5’ side of RPA would orientate the ZnF upstream of the AEP domain, with initial binding occurring in a “closed” conformation. In the article presented in Chapter 2 we proposed a model whereby PrimPol was recruited to the 3’ side of RPA, allowing the enzyme to contact the primer-template junction to facilitate TLS. This orientation was necessary to permit the ZnF to contact the ssDNA template and the AEP domain to access the primer-template junction. However, since this publication, increasing evidence suggests PrimPol primarily functions as a repriming enzyme, rather than a TLS polymerase (Kobayashi et al., 2016; Pilzecker et al., 2016; Rechkoblit et al., 2016; Schiavone et al., 2016; Vallerga et al., 2015). Importantly, during repriming the ZnF can contact ssDNA on either side of the AEP domain.

As previously discussed, RPA binds to DNA in a manner which positions the PrimPol-recruiting RPA70N domain 5’ relative to the rest of the protein. This would suggest that PrimPol is recruited to the 5’ side of RPA. However, RPA70N and PrimPol’s RBD are attached to the rest of their respective proteins by flexible linkers which may allow PrimPol recruitment to the 3’ side of RPA. RPA binds to polypyrimidine sequences with ~50-fold higher affinity that polypurine sequences, specifically, in order of decreasing affinity RPA binds to dC > dT > mixed ssDNA > dA/dG (Kim et al., 1992). As an initial investigation into the polarity of PrimPol recruitment to RPA during repriming, mixed sequence DNA templates (96 nt in length) containing a 30 nt polydC or polydT tract at their 5’ or 3’ end were synthesised (Figure 5.16.A.). These templates were used for standard fluorescence-based primase assays to analyse the primase activity of PrimPol in the presence of increasing concentrations of RPA.

Here, it would be expected that if PrimPol is recruited to the 5’ side of RPA, a polypyrimidine tract at the 3’ end of the template, and therefore RPA binding to the 3’ end of the template, should stimulate primase activity as the concentration of RPA is increased. Moreover, priming should be initiated immediately after the RPA molecule, producing reaction products of ~66 nt in its presence. With templates containing a 5’ polypyrimidine tract, RPA would not recruit PrimPol and longer products should be observed from primer synthesis and extension from the 3’ end of the template. Alternatively, if PrimPol is recruited to the 3’ side of RPA, templates with a 3’ polypyrimidine tract should not stimulate primase activity and may sequester PrimPol away from the ssDNA template, causing inhibition. In this situation, templates with a 5’ polypyriminde tract should produce short reaction products of around 30-40 nt due to recruitment of PrimPol to the ssDNA on the 3’ side of the RPA molecule bound at the 5’ end of the template.

The results of these assays show that when the polypyrimidine tract (both polydC and polydT) is located at the 3' end of the template, PrimPol's primase activity is stimulated as the concentration of RPA is increased. At the highest concentration, inhibition occurs due to saturation of the template, as previously determined (Figure 5.16.B.). In each case, products of ~60-70 nt are produced, indicative of priming immediately after the polypyrimdine tract. Notably, similar size products were produced in the absence of RPA, likely due to the inability of PrimPol to prime in the polypyrimidine tract. Nevertheless, in the presence of RPA a more consistent 60-70 nt product is observable, compared to reactions lacking RPA. Interestingly, on the 3' polydC template, a second predominant product of ~30-40 nt is produced at higher RPA concentrations. This is consistent with PrimPol binding downstream (on the 5' side) of a second RPA molecule (Figure 5.16.B. indicated by red asterisk). The stimulation of primase activity, coupled with the size of reaction products, observed with increasing RPA on these templates suggests that PrimPol can be efficiently recruited to the 5' side of RPA to reprime (Figure 5.16.C.).

Interestingly, when the polypyrimidine tract is located at the 5' end of the template, reaction products are much longer, equivalent to the size of the full template (Figure 5.16.B.). This is indicative of priming at the very 3' end of the template and not at the polypyrimdine tract where RPA is expected to be bound. As the concentration of RPA is increased, more full-length reaction products are produced and stalling of PrimPol at the start of the polypyrimidine tract is reduced. Intriguingly, this suggests that PrimPol is priming at the 3' end of the template and extending the primer to the polypyrimidine tract (Figure 5.16.C.). In the absence of RPA PrimPol stalls here, producing ~60-70 nt products. However, in the presence of RPA synthesis continues to the end of the template. This suggests that RPA is binding to the polypyrimdine tract and aiding primer extension through this region, potentially by relaxing any DNA secondary structures, which it has previously been reported to do (Chen et al., 2013; Safa et al., 2016). This allows inference that RPA is binding where expected and PrimPol is priming at the opposite end of the template, in contrast to what is seen on 3' polypyrimidine templates (Figure 5.16.C.). This further supports recruitment of PrimPol to the 5', rather than 3', side of RPA during primer synthesis.

These preliminary results support the model suggested in the article presented above, whereby PrimPol binds and primes downstream of RPA on 5' side (Figure 5.16.C.). Indeed, this is consistent with the polarity of RPA when bound to ssDNA. In light of this, the inhibitory effect of RPA on PrimPol during primer-extension assays, as shown in Chapter 2, may be partially explained by recruitment of the enzyme to the 5' side of the protein, consequently preventing access to the primer. It is not inconceivable that this

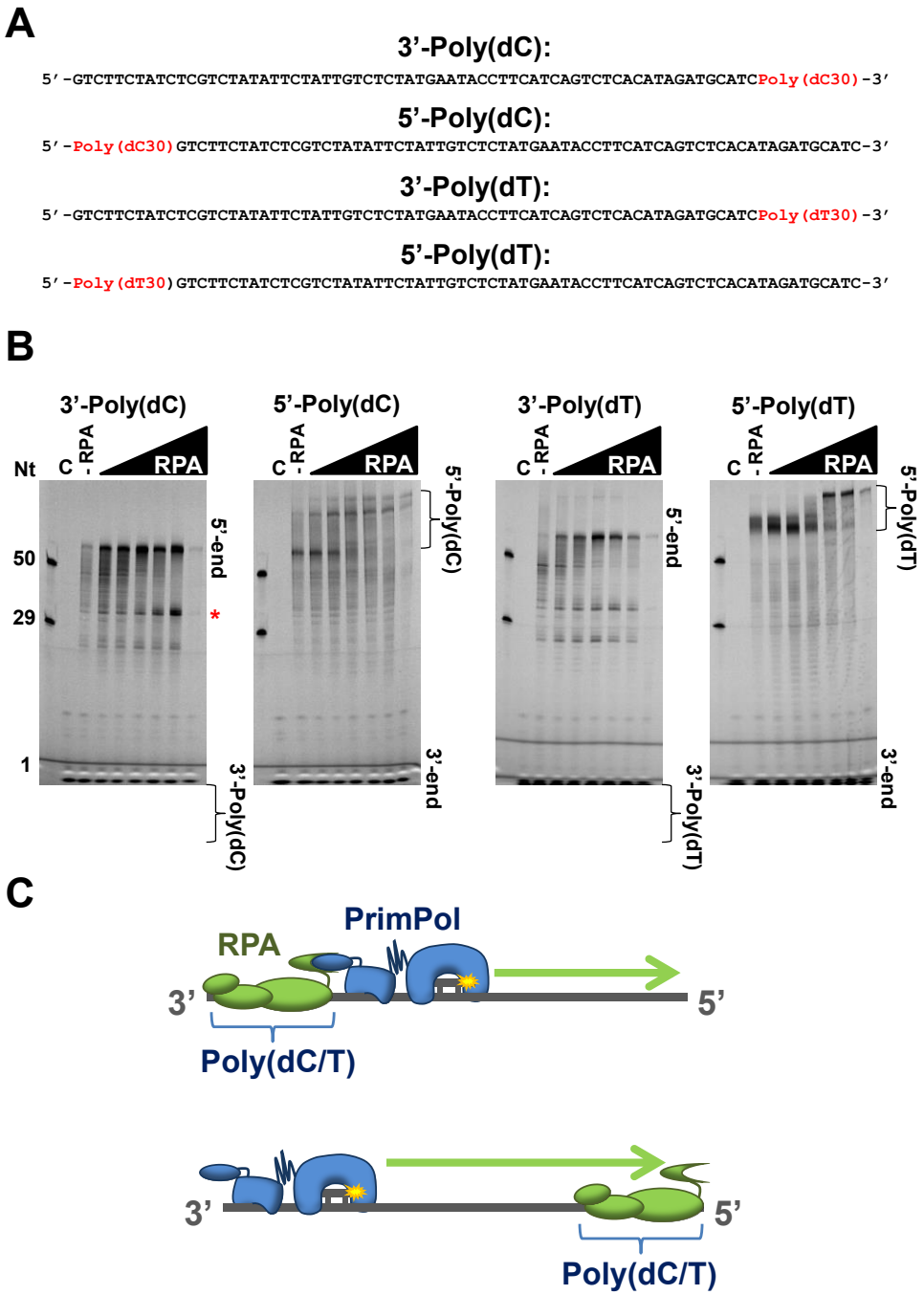


Figure 5.16. Preliminary investigations into the polarity of RPA-mediated PrimPol recruitment.
(A) Sequences of the ssDNA templates used for the primase assays presented in B. Templates contained a 30 nt Poly(dC) or Poly(dT) region at either the 3' or 5' end (indicated in red) and a 3' biotin tag. **(B)** Analysis of PrimPol's primase activity on the templates listed in A in the presence of increasing concentrations of RPA (0, 0.5, 1, 2, 3, 4, 8 μM). PrimPol (1 μM) was incubated with the respective template (1 μM) and RPA for a single 10 min time-point. 'C' indicates the no PrimPol control, '-RPA' indicates reactions containing PrimPol only, the red asterisk indicates the location of a second predominant reaction product at higher RPA concentrations on the 3'-Poly(dC) template, presumably a result of priming on the 5' side of a second RPA molecule. Nucleotide size markers (Nt) are shown on the left of the gels. **(C)** Schematic showing the location of PrimPol (blue) priming on both 3' and 5' polypyrimidine templates based upon the size of the reaction products produced in each case. The reactions products produced in B allow inference that RPA (green) is binding to the polypyrimidine tract of each template and suggests PrimPol is recruited to the 5' side of RPA during primer synthesis.

could prevent PrimPol from performing TLS *in vivo* and instead promote a repriming role. Nevertheless, these studies are in the early stages and further experiments and controls are required to confirm or negate this model.

5.6.2. Studies to Identify Additonal PrimPol Interacting Partners

As shown in Figure 5.15.C. PrimPol's RBM-B displays a high level of homology with the previously identified And-1/Ctf4-interacting motifs of Pol α and the Sld5 subunit of GINS (Simon et al., 2014). Ctf4 forms a homotrimer which has been suggested to tether Pol α to the CMG complex, however in reconstitution experiments Pol α functioned distributively in the presence of Ctf4 (Yeeles et al., 2017). Alternatively, it has also been suggested that Pol α is tethered to the replication fork by shared interactions with FACT and nucleosomes, rather than by Ctf4 (Kurat et al., 2017). Nevertheless, the similarity between RBM-B and the Ctf4 interacting motif of Pol α, coupled with the dispensability of RBM-B in binding RPA *in vivo*, raise the possibility that RBM-B instead interacts with Ctf4 during replication. In this section, efforts to identify additional PrimPol interacting partners, including Ctf4, will be described. These investigations have focused on identifying binding partners through co-immunoprecipitation and MS analysis with western blot validation, using a number of FLAG-tagged PrimPol constructs.

5.6.2.1. Materials and Methods

5.6.2.1.1. Co-Immunoprecipitation and MS Analysis

Large-scale immunoprecipitation experiments for MS analysis were performed using HEK293 Flp-In T-REx cells engineered for inducible expression of FLAG-PrimPol (WT, 480-560, and full-length (FL) RBM-B-K.O.). The methodology used was identical to that described in section 5.3.10. except five 175 cm^2 flasks of confluent HEK293 cells, grown in the presence or absence of 10 ng/mL of doxycycline for 24 hrs, were used. Following immunoprecipitation, elutions were pooled and processed in-solution for MS. Samples were concentrated to 40 µL and supplemented with 160 µL of 8 M urea, 100 mM Tris, 10 mM $CaCl_2$, and 5mM DTT, before incubation at 37°C for 1 hr. Following incubation, 16 µL of 200 mM iodoactemide (final concentration 15 mM) was added and samples incubated for a further 20 mins in the dark at room temperature. Samples were then diluted 7x using 100 mM Tris and 10 mM $CaCl_2$. Digestion was performed using 3 µg of sequencing grade modified trypsin (Promega) at 37°C overnight. Subsequently, samples were concentrated to 100 µL using a SpeedVac concentrator and pH adjusted using trifluoroacetic acid to a final concentration of 1%. Samples were then desalted using Pierce C18 tips (ThermoFisher) according to the manufacturer's protocol before concentrating to 10 µL using a SpeedVac.

Prepared samples were analysed by MS using a nano-LC-MS (ThermoFisher U3000 nanoLC and Orbitrap XL mass spectrometer) (performed by Dr. Peter Kolesar) as previously described (Hatimy et al., 2015). Data were analysed and quantified by MaxQuant using standard settings to determine label-free quantification (LFQ) intensities. Four independent FLAG-PrimPol$_{480-560}$ experiments were performed, whilst the results presented for WT and FL B-RA are each from a single analysis. For experiments using HU, cells were treated for 3 hrs with 1 mM HU prior to harvesting. Small-scale immunoprecipitation and western blot experiments were performed as described in section 5.3.10.

5.6.2.1.2. Isolation of Mitochondria from HEK293 Cells

HEK293 Flp-In T-REx cells engineered for inducible expression of FL WT and RBM-A-K.O. FLAG-PrimPol were induced with 10 ng/mL doxycycline 24 hrs prior to harvesting. Mitochondria were isolated using a mitochondria isolation kit (ThermoFisher) following the manufacturers instructions. Nuclear and cytoplasmic samples were kept and analysed alongside isolated mitochondria by western blot using anti-FLAG, anti-RPA2, and anti-mtSSB antibodies.

5.6.2.2. MS Analysis of PrimPol CTD Interacting Partners

In order to identify additional PrimPol CTD interacting partners, potentially including Ctf4/And1, large-scale co-immunoprecipitation and MS experiments were performed using HEK293 cells engineered for inducible expression of N-terminal FLAG-tagged PrimPol$_{480-560}$. Previous pull-down and MS screens of PrimPol interacting partners used C-terminal *Strep*-tagged full-length PrimPol (Rudd, 2013). Importantly, as RBM-B is located at the C-terminus of PrimPol, tagging of the enzyme at this region may prevent binding of RBM-B interacting partners. Use of FLAG-PrimPol$_{480-560}$ has the additional benefit of removing the DNA-binding AEP and ZnF domains, potentially helping to reduce background from DNA bridging between proteins. Despite lacking the functional domains required for activity, Figure 5.12.B. demonstrates that PrimPol$_{480-560}$ is still recruited to chromatin, validating its use in these studies.

A summary of the results of four independent FLAG-PrimPol$_{480-560}$ co-immunoprecipitations are highlighted in Table 5.3. In addition to all three RPA subunits, mtSSB was enriched 2.2-fold in induced cells, over those grown without doxycycline. This suggests that mtSSB, like RPA, interacts with the CTD of PrimPol. Aside from these previously identified binding partners, a number of potential novel PrimPol-interacting partners, involved in DNA replication and repair, were identified. Perhaps most significant was the 104-fold enrichment of DNA ligase 3 (LIG3) and 13-fold enrichment of PARP1

Protein Name	LFQ +Dox	LFQ -Dox	Fold Enrichment	Only in +Dox	In all expts.	Functional Group
LIG3	2.38×10^9	2.28×10^7	104.4	N	N	DNA Repair
BANF1	1.06×10^9	6.81×10^7	15.6	N	Y	DDR
PARP1	1.36×10^{10}	1.03×10^9	13.2	N	Y	DNA Repair
Histone H2A	4.37×10^8	6.03×10^7	7.2	N	Y	Chromatin Component
RPA14	1.70×10^9	3.20×10^8	5	N	Y	DNA Replication
RPA70	3.37×10^{10}	7.6×10^9	4.4	N	Y	DNA Replication
FACT subunit SSRP1	2.42×10^8	5.55×10^7	4.4	N	N	Chromatin Processing
RPA32	1.00×10^{10}	2.82×10^9	3.6	N	Y	DNA Replication
mtSSB	2.67×10^9	1.20×10^9	2.2	N	Y	mtDNA Replication
DNA-PKcs	8.62×10^8	0	N/A	Y	N	DNA Repair
XRCC1	6.50×10^8	0	N/A	Y	N	DNA Repair
Histone H2B	1.93×10^8	0	N/A	Y	N	Chromatin Component
XRCC5	1.08×10^8	0	N/A	Y	Y	DNA Repair
Histone H4	6.95×10^7	0	N/A	Y	N	Chromatin Component
PCNA	4.20×10^7	0	N/A	Y	N	DNA Replication
RuvB-like 2	3.20×10^7	0	N/A	Y	N	Chromatin Remodelling
XRCC6	3.07×10^7	0	N/A	Y	Y	DNA Repair
APTX	2.72×10^7	0	N/A	Y	N	DNA Repair
PrimPol	2.66×10^7	0	N/A	Y	Y	DNA Replication / DDT
PNKP	2.06×10^7	0	N/A	Y	N	DNA Repair
hNRP	1.73×10^7	0	N/A	Y	N	Chromatin Formation
DDB1	1.47×10^7	0	N/A	Y	N	DNA Repair
PARP2	3.30×10^6	0	N/A	Y	N	DNA Repair
RBBP4	2.58×10^6	0	N/A	Y	N	Chromatin Remodelling
TFAM	2.51×10^6	0	N/A	Y	N	mtDNA Replication

Table 5.3. Mass spectrometry analysis of PrimPol480-560 co-immunoprecipitation experiment elutions.
The results presented are a summary of four independent co-immunoprecipitation analyses. Label-free quantification (LFQ) intensities were determined by MaxQuant and represent the means of the four analyses. For proteins identified in both + Dox and – Dox samples, the fold enrichment was calculated by dividing LFQ + Dox intensities by LFQ – Dox intensities.

in samples from induced cells. Likewise, DNA-dependent protein kinase catalytic subunit (DNA-PKcs), Bifunctional polynucleotide/kinase (PNKP), Aprataxin (APTX), DNA damage-binding protein 1 (DDB1), PARP2, and X-ray repair cross-complementing protein (XRCC) 1, 5, and 6, were all only identified from cells expressing PrimPol$_{480\text{-}560}$. All of these proteins play a role in DNA repair, specifically XRCC5/6, DNA-PKcs, and APTX are involved in NHEJ, whilst PARP1, XRCC1, and LIG3, are thought to function in MMEJ (Ceccaldi et al., 2016). PNKP assists in each pathway by processing DNA termini for extension or ligation, while DDB1 functions in the recognition of DNA damage and PARP2 is implicated in a number of DNA repair processes including BER/SSBR and NHEJ. Given the role of AEPs in prokaryotic DSBR, it is an alluring possibility that PrimPol plays an analogous role in eukaryotes. Indeed, in section 4.3. it was observed that PrimPol displays a degree of catalytic versatility which might lend itself to such a process. However, many of the DNA repair proteins identified here have previously been found in pull-down and MS studies of RPA, these include PARP1, LIG3, DNA-PKcs, XRCC1, XRCC5/6, PNKP, and DDB1 (Maréchal et al., 2014). In light of this, it seems likely that these interactions are simply mediated by RPA, rather than a direct interaction with PrimPol. Likewise, to date, no *in vivo* studies support a role for PrimPol in either NHEJ or MMEJ. Nevertheless, PARP1, DNA-PKcs, and XRCC1, have recently been implicated in the restart of stalled replication forks, suggesting that PrimPol could coordinate with these proteins in a novel restart pathway, rather than a DSBR process (Ying et al., 2016).

In addition to the DNA repair proteins described above, the co-immunoprecipitation studies of FLAG-PrimPol$_{480\text{-}560}$ also identified a number of chromatin constituents and remodeling factors including, FACT subunit SSRP1, Histone H2A, H2B, and H4, RuvB-like 2, Nucleosome assembly protein 1-like 1 (hNRP), and Histone-binding protein RBBP4 (Table 5.3.). This observation is intriguing given that functional interactions between FACT, Pol α, and nucleosomes, have recently been suggested to regulate priming of Okazaki fragments during lagging strand synthesis (Kurat et al., 2017). Similarly, TFAM, which has roles in both DNA packaging and replication in the mitochondria was also identified. However, further work is required to assess the affect of chromatin and DNA packaging factors on priming by PrimPol.

Notably, although known interacting partners including RPA and mtSSB were identified in these studies, Ctf4/AND1 was not detected in any of the analyses. This suggests that RBM-B is not involved in binding to Ctf4 *in vivo*, further supporting the previous suggestion that it instead functions to bind a second RPA molecule following recruitment by RBM-A.

5.6.2.3. mtSSB, but not AND-1, Co-Purifies with PrimPol's CTD

A rather surprising observation from the MS analyses presented in Table 5.3. was the detection of PCNA only in samples from induced cells. This is at odds with the previously described *Strep*-PrimPol pull-down MS analysis which failed to identify PCNA as an interacting partner (section 2.4.2.). In order to further investigate this and confirm the observed interaction with mtSSB, western blot analysis of elutions from immunoprecipitation of FL WT, RBM-A-K.O., RBM-B-K.O. and RBM-A+B-K.O. FLAG-PrimPol, in addition to $PrimPol_{480-560}$ (CTD), was performed (Figure 5.17.).

Here, analysis of WT FLAG-PrimPol elutions with anti-PCNA, identified the protein in both + and – Dox samples at similar levels, despite the absence of any detectable PrimPol in – Dox elutions (Figure 5.17.A.). This suggests that PCNA binds to the anti-FLAG antibodies used in the immunoprecipitation studies, supporting previous observations that the protein does not interact with PrimPol. Blots were also analysed with anti-AND1. Consistent, with the MS results shown in Table 5.3., AND-1 was not detectable in any elutions from the PrimPol variants used, further suggesting that the proteins do not share an interaction *in vivo* (Figure 5.17.A-E.).

In contrast, mtSSB was indentified in elutions from both WT and $PrimPol_{480-560}$-expressing cells and did not appear in elutions from uninduced cells (Figure 5.17.A. and B.). Moreover, mtSSB did not co-elute with PrimPol when RBM-A alone, or in combination with RBM-B, was mutated (Figure 5.17.C. and E.). When RBM-B was mutated alone, mtSSB binding was significantly reduced, however the protein was detectable when using longer exposure times (Figure 5.17.D.). This is very similar to the results observed with RPA in Figure 5.10., potentially revealing that the mtSSB interaction, like the RPA interaction, is mediated by PrimPol's RBM-A, with a secondary contribution from RBM-B. This would not be surprising given the analogous function of mtSSB in the mitochondria. However, unlike RPA, we have failed to identify an interaction between PrimPol and mtSSB in SEC analysis (data not shown). Similarly, mtSSB does not stimulate the primase activity of PrimPol *in vitro* (Figure 5.18.A.). These results are also consistent with the mtSSB-interaction potentially being mediated by RPA and/or ssDNA upon cell lysis during pull-down studies. Nevertheless, it is likely that mitochondrially-localised PrimPol is modified in comparison to nuclear PrimPol and these modifications may be required to observe an interaction *in vitro*. One possibility is that an intact RBM-A is required for localisation of PrimPol to the mitochondria. To test this, mitochondria were isolated from WT FLAG-PrimPol and A-RA FLAG-PrimPol expressing HEK293 cells and analysed by western blot (Figure 5.18.B.). Here, there was

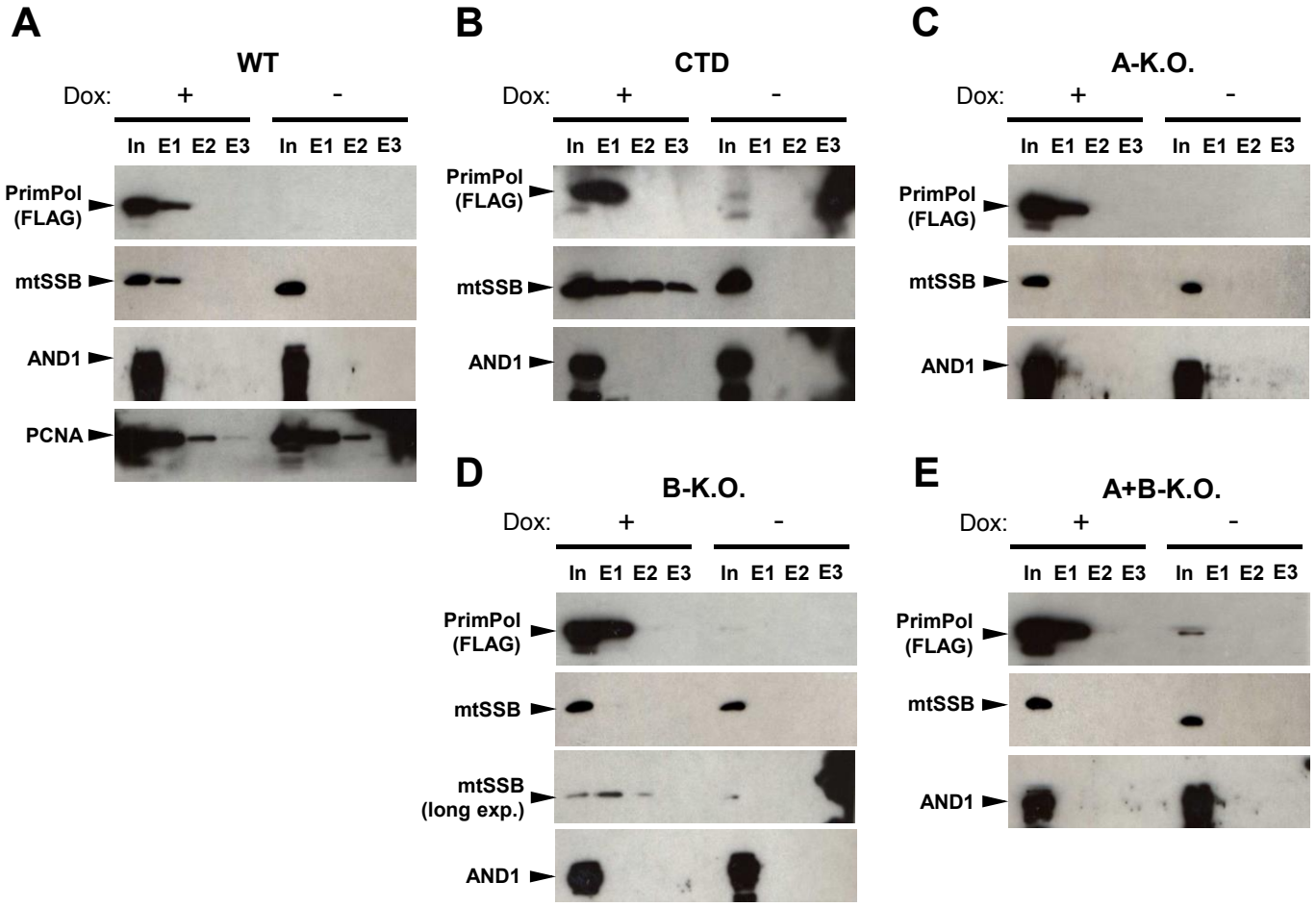


Figure 5.17. mtSSB co-purifies with PrimPol$_{480\text{-}560}$.
(A-E) Flp-In T-Rex-293 cells transfected with FLAG-tagged wild-type (WT) **(A)**, 480-560 (CTD) **(B)**, RBM-A-K.O. **(C)**, RBM-B-K.O. **(D)**, and RBM-A+B-K.O. **(E)** PrimPol were grown in the presence or absence of doxycycline (10 ngmL^{-1}, 24 hrs), FLAG-PrimPol was immunoprecipitated from the soluble cell lysate using anti-FLAG antibodies and western blotted for PrimPol (anti-FLAG), mtSSB (anti-mtSSB), AND1 (anti-AND1) and PCNA (anti-PCNA, A only). The presence and absence of doxycycline is indicated by +/- Dox, 'In' indicates the input, 'E1', 'E2', and 'E3', indicate elutions 1, 2, and 3, respectively.

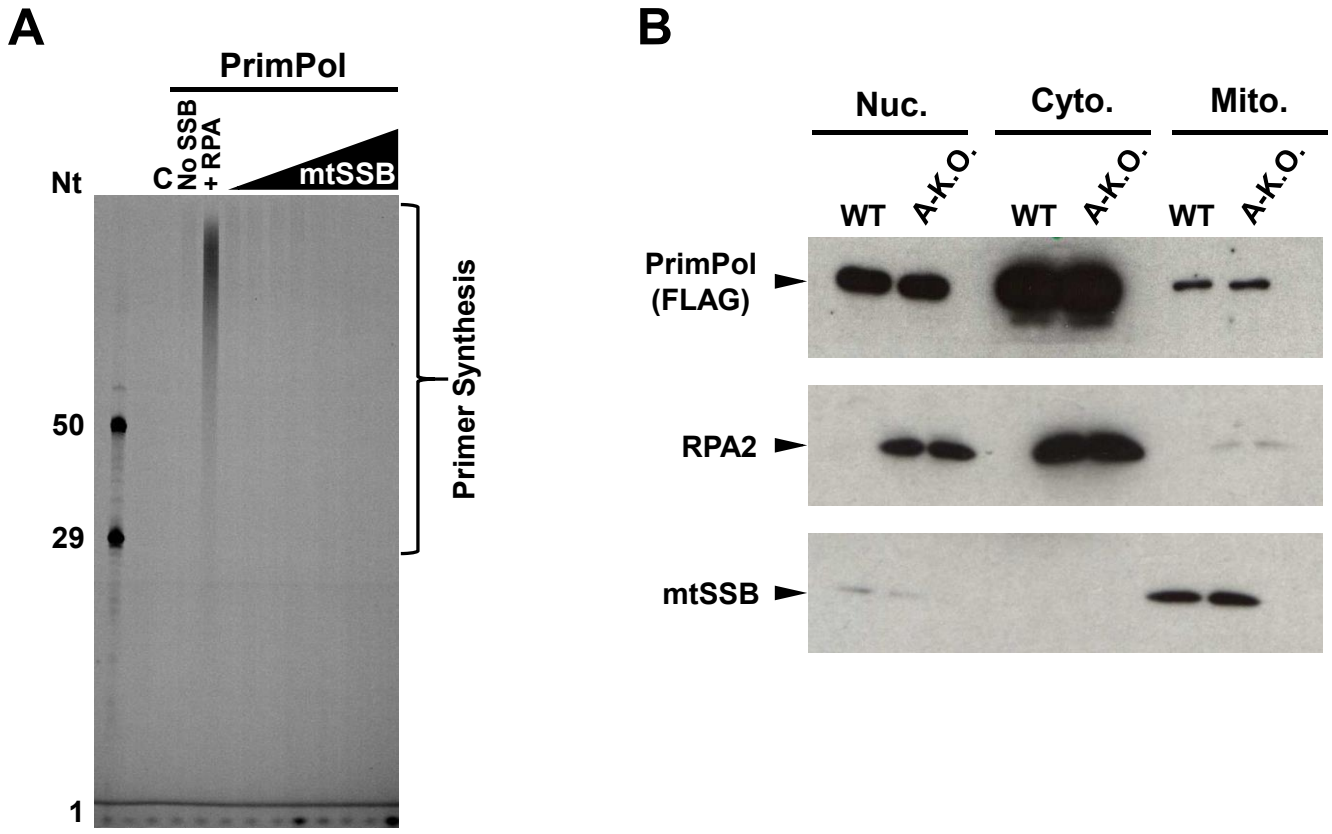


Figure 5.18. mtSSB does not stimulate PrimPol's primase activity in vitro and is not required for localisation of PrimPol to mitochondria in vivo. **(A)** Primer synthesis by WT hPrimPol (400 nM) on M13 ssDNA templates (20 ngμL^{-1}) in the presence of increasing concentrations of mtSSB (0.0625, 0.125, 0.25, 0.5, 1, 2, 4, and 8 μM). Reactions containing PrimPol only are indicated by 'no SSB', a positive control containing 1 μM RPA in place of mtSSB is indicated by '+ RPA. 'C' indicates the no enzyme control, oligonucleotide (Nt) length markers are shown on the left of the gel. **(B)** Flp-In T-Rex-293 cells transfected with FLAG-tagged wild-type (WT) and RBM-A-K.O. PrimPol were grown in the presence of doxycycline (10 ngmL^{-1}, 24 hrs) prior to isolation of mitochondria. Nuclear (Nuc.), cytoplasmic (Cyto.) and mitochondrial (Mito.) samples were retained and analysed by western blot using anti-FLAG, anti-RPA2, and anti-mtSSB antibodies. Enrichment of mtSSB but not RPA in Mito. samples confirms the appropriate isolation of mitochondria.

no observable difference in the level of PrimPol in the mitochondria following mutation of RBM-A. This demonstrates that RBM-A is not required for localisation of PrimPol to the mitochondria.

5.6.2.4. MS Analysis of Full-length PrimPol Interacting Partners

In order to identify if the potential CTD-interacting partners observed in the MS analysis are relevant in the context of the full-length protein, co-immunoprecipitation and MS studies were repeated using WT FLAG-PrimPol-expressing HEK293 cells (Table 5.4.). PrimPol has previously been shown to form foci in HEK293 cells in response to HU treatment (Bianchi et al., 2013). To identify potential new interactions induced upon fork stalling, analysis was also performed on cells treated with HU for 3 hrs (Table 5.5.). Additionally, analysis using cells expressing full-length RBM-B-K.O. PrimPol was also performed to identify any differences in comparison to the WT (Table 5.6.).

In each case RPA and mtSSB were only identified in elutions from induced cells (Table 5.4., 5.5. and 5.6.). This is in agreement with the results observed in Figure 5.10. and 5.17. showing that RBM-B, unlike RBM-A, is not critical for PrimPol's interaction with RPA or mtSSB *in vivo*. Analysis of elutions from WT FLAG-PrimPol cells treated with HU, and untreated RBM-B-K.O. FLAG-PrimPol cells, also identified a number of repair factors found in the CTD analyses, these include PARP1, XRCC6, DNA-PKcs, and LIG3 (Table 5.5. and 5.6.). Likewise a number of chromatin constituents and remodeling factors were identified in all three full-length PrimPol analyses, including hNRP, Histone H2A and H2B, and RuvB-like 1 and 2 (Table 5.4., 5.5. and 5.6.). Although FACT was not detected in any of the elutions.

Aside from factors already identified in the CTD analyses, a number of extra proteins were identified when using full-length PrimPol, including MCM 4, 6, and 7, and CDK1, suggesting that any potential interaction with these proteins is not mediated by the CTD (Table 5.4., 5.5. and 5.6.). Following treatment of WT FLAG-PrimPol cells with HU, additional proteins were detected, such as High mobility group protein (HMG) B1 and B2, and MMR protein MSH2 (Table 5.5.). Further work is required to confirm or negate these potential interacting partners.

5.6.2.4. Western Blot Analysis of Full-length PrimPol Interacting Partners

In order to validate some of the interactions observed in the MS analysis, small scale immunoprecipitations of full-length WT, RBM-A-K.O., RBM-B-K.O. and RBM-A+B-K.O. PrimPol, in addition to $PrimPol_{480\text{-}560}$ (CTD), were performed and analysed by western blot using anti-PARP1, anti-XRCC1, and anti-MCM4 antibodies (Figure 5.19.). PARP1

Protein Name	LFQ +Dox	LFQ -Dox	Fold Enrichment	Only in +Dox	Functional Group
PrimPol	1.05×10^9	1.19×10^7	88	N	DNA Replication / DDR
XRN1	2.16×10^7	0	N/A	Y	RNA/DNA Processing
RPA70	1.38×10^7	0	N/A	Y	DNA Replication
mtSSB	1.18×10^7	0	N/A	Y	mtDNA Replication
MCM4	1.01×10^7	0	N/A	Y	DNA Replication
RuvB-like 1	7.37×10^6	0	N/A	Y	Chromatin Remodelling
MCM7	7.31×10^6	0	N/A	Y	Replication
RuvB-like 2	6.54×10^6	0	N/A	Y	Chromatin Remodelling
hNRP	5.93×10^6	0	N/A	Y	Chromatin Formation
RPA14	4.95×10^6	0	N/A	Y	DNA Replication
RPA32	4.90×10^6	0	N/A	Y	DNA Replication
MCM6	2.56×10^6	0	N/A	Y	DNA Replication

Table 5.4. Mass spectrometry analysis of WT PrimPol co-immunoprecipitation experiment elutions.
Label-free quantification (LFQ) intensities were determined by MaxQuant. For proteins identified in both + Dox and – Dox samples, the fold enrichment was calculated by dividing LFQ + Dox intensities by LFQ – Dox intensities.

Protein Name	LFQ +Dox	LFQ -Dox	Fold Enrichment	Only in +Dox	Functional Group
PrimPol	5.46×10^8	1.15×10^7	48	N	DNA Replication / DDT
PARP1	1.68×10^7	5.44×10^6	3	N	DNA Repair
MCM7	1.07×10^7	0	N/A	Y	DNA Replication
mtSSB	7.17×10^6	0	N/A	Y	mtDNA Replication
RPA14	6.14×10^6	0	N/A	Y	DNA Replication
XRCC6	6.03×10^6	0	N/A	Y	DNA Repair
HMGB2	5.30×10^6	0	N/A	Y	DNA Replication / Repair
RuvB-like 1	5.12×10^6	0	N/A	Y	Chromatin Remodelling
Histone H2A	5.05×10^6	0	N/A	Y	Chromatin Constituent
HMGB1	5.04×10^6	0	N/A	Y	DNA Replication / Repair
Histone H2B	4.76×10^6	0	N/A	Y	Chromatin Constituent
RPA70	4.39×10^6	0	N/A	Y	DNA Replication
PCNA	2.47×10^6	0	N/A	Y	DNA Replication
CDK1	8.25×10^5	0	N/A	Y	Cell-Cycle Regulation
DNA-PKcs	4.93×10^5	0	N/A	Y	DNA Repair
MSH2	2.94×10^5	0	N/A	Y	DNA Repair

Table 5.5. Mass spectrometry analysis of WT PrimPol co-immunoprecipitation experiment elutions following hydroxyurea treatment.
Cells were treated with 1 mM HU for 3 hrs prior to harvesting and immunoprecipitation. Label-free quantification (LFQ) intensities were determined by MaxQuant. For proteins identified in both + Dox and – Dox samples, the fold enrichment was calculated by dividing LFQ + Dox intensities by LFQ – Dox intensities.

Protein Name	LFQ +Dox	LFQ -Dox	Fold Enrichment	Only in +Dox	Functional Group
PrimPol	5.90×10^8	4.08×10^6	145	N	DNA Replication / DDT
PARP1	3.00×10^7	1.49×10^7	2	N	DNA Repair
XRCC6	7.04×10^6	5.38×10^6	1.3	N	DNA Repair
RPA70	3.00×10^7	0	N/A	Y	DNA Replication
RPA32	1.32×10^7	0	N/A	Y	DNA Replication
RPA14	9.47×10^6	0	N/A	Y	DNA Replication
MCM7	8.33×10^6	0	N/A	Y	DNA Replication
BANF1	7.11×10^6	0	N/A	Y	DDR
mtSSB	7.05×10^6	0	N/A	Y	mtDNA Replication
Histone H2A	4.20×10^6	0	N/A	Y	Chromatin Constituent
LIG3	1.56×10^6	0	N/A	Y	DNA Repair
CDK1	7.36×10^5	0	N/A	Y	Cell-Cycle Regulation

Table 5.6. Mass spectrometry analysis of full-length RBM-B-K.O. PrimPol co-immunoprecipitation experiment elutions.
Label-free quantification (LFQ) intensities were determined by MaxQuant. For proteins identified in both + Dox and – Dox samples, the fold enrichment was calculated by dividing LFQ + Dox intensities by LFQ – Dox intensities.

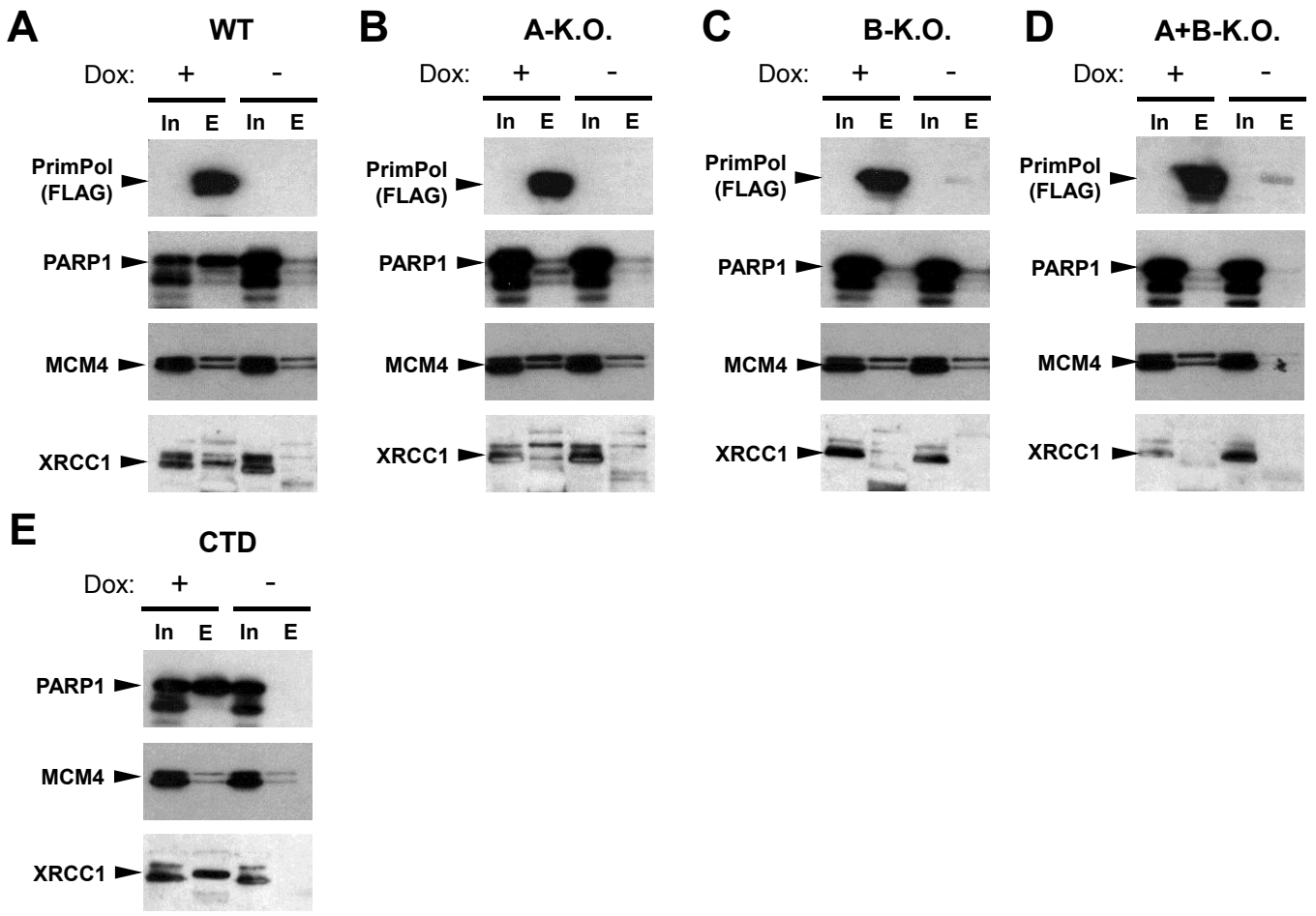


Figure 5.19. Western blot analysis of PARP1, MCM4, and XRCC1 interactions.
(A-E) Flp-In T-Rex-293 cells transfected with FLAG-tagged wild-type (WT) **(A)**, RBM-A-K.O. **(B)**, RBM-B-K.O. **(C)**, RBM-A+B-K.O. **(D)**, and 480-560 (CTD) **(E)** PrimPol, were grown in the presence or absence of doxycycline (10 $ngmL^{-1}$, 24 hrs), FLAG-PrimPol was immunoprecipitated from the soluble cell lysate using anti-FLAG antibodies and western blotted for PrimPol (anti-FLAG), PARP1 (anti-PARP1), MCM4 (anti-MCM4) and XRCC1 (anti-XRCC1). The presence and absence of doxycycline is indicated by +/- Dox, 'In' indicates the input, 'E' indicates elution.

was enriched in elutions from WT and CTD expressing cells (Figure 5.19.A. and E). However, this interaction was significantly reduced upon mutation of RBM-A or RBM-B in isolation and together (Figure 5.19.B-D.). Although notably, PARP1 was enriched 2-fold in the MS analysis of RBM-K.O. elutions (Table 5.6.). Likewise, XRCC1 was detected in CTD and WT elutions, but was lost upon mutation of either RBM. Analysis using anti-MCM4 identified the protein in both + and – Dox elutions, suggesting that MCM like PCNA, may bind the anti-FLAG beads used for the experiment. However, there was a slight enrichment of the protein in + Dox elutions of the FL PrimPol constructs (Figure 5.19.A-D.), which was not observed in the CTD samples (Figure 5.19.E.). This is potentially consistent with the MS results which identified MCM subunits in the FL PrimPol, but not CTD, elutions. Further work is required to examine if this is a genuine interaction.

These results give further credence to the possibility that many of the interactions identified in these analyses are mediated by RPA. Given that mutation of RBM-A prevents recruitment of PrimPol to chromatin and a loss of the RPA-interaction, it is difficult to distinguish if potential interactions with other partners are lost due to PrimPol not being recruited to chromatin, or because the interactions are mediated by RPA. The observation that many of the factors identified in the MS analyses have also been identified in RPA pull-down studies, suggests that the latter is more likely (Maréchal et al., 2014). Nevertheless, interactions mediated by RPA might still be functionally relevant *in vivo* by localising PrimPol to other DNA replication and repair factors. Importantly, the interaction studies presented here are in the preliminary stages and further work is required to rule out or confirm a direct interaction between PrimPol and the identified proteins. However, the repeated observation that many of these interactions are lost upon mutation of PrimPol's RBMs further highlights the importance of the RPA-interaction for the enzyme's role *in vivo*. The identification of RPA in all of the MS analyses strengthens the work presented earlier in this Chapter and supports RPA as the primary PrimPol-interacting partner.

Chapter 6

The Role, Recruitment, and Regulation of PrimPol During DNA Replication

6.1. Abstract

The complex molecular machines responsible for genome replication encounter many obstacles during their progression along DNA. Tolerance of these obstructions is critical for efficient and timely genome duplication. In recent years, PrimPol has emerged as a new player involved in maintaining eukaryotic replication fork progression. This versatile replicative enzyme, a member of the AEP superfamily, has the capacity to perform a range of template-dependent and independent synthesis activities. However, accumulating evidence suggests that PrimPol's primary role is to reprime and restart DNA replication downstream of a range of replicase-stalling impediments. The following Chapter is a modified version of a review of the current PrimPol literature which was recently published in *Genes* (Guilliam and Doherty, 2017). Here, the insights into the role, recruitment, and regulation of PrimPol, garnered in part from the work presented in this thesis, are discussed in the context of the wider literature. This review provides an overview of the complete body of PrimPol studies performed to date, allowing new interpretations of some of the data presented in the preceding Chapters. In this respect, there is some repetition in the themes and topics presented, however this is necessary to consider the initial PrimPol reports in light of more recent studies. The review begins with an overview of the evolutionary history of PrimPol, before proceeding to discuss recent advances in our understanding of the enzyme's role, recruitment, and regulation, as well as highlighting unanswered questions and potential future avenues of investigation.

6.2. The Discovery and Evolutionary History of PrimPol

As discussed in Chapter 1, the AEP superfamily is evolutionarily and structurally distinct from the bacterial DnaG-type primases which, like AEPs in archaea and eukarya, are absolutely required for DNA replication initiation in bacteria (Guilliam et al., 2015b). Nevertheless, DnaG-like primases are also present in archaea, and likewise, AEPs have been identified in bacteria (Aravind et al., 1998b). In each case, these enzymes have diverged to fulfil alternative roles, for example in bacteria a member of the AEP superfamily is employed, together with Ku and DNA ligase homologues, in an NHEJ DNA break repair pathway (Della et al., 2004; Koonin et al., 2000). It is likely that the presence of AEPs in bacteria is a result of HGT, with the enzymes originally being recruited for replication initiation by the archaeo-eukaryotic lineage following their divergence from bacteria (Guilliam et al., 2015b; Iyer et al., 2005). The catalytic core of AEPs is defined by two structural modules; an N-terminal module with an $(\alpha\beta)_2$ unit, and a C-terminal RRM-like fold. These two modules pack together, with the active site residues located in between them (Iyer et al., 2005).

In 2005, detailed *in silico* analyses divided the AEP superfamily into 13 major families which were further organised into three higher order clades; the AEP proper clade, the NCLDV-herpesvirus primase clade, and the primpol clade (Iyer et al., 2005). These analyses also identified PrimPol and assigned it to the NCLDV-herpesvirus clade, whose members are only present in eukaryotes and their viruses. This clade encompasses the iridovirus primase and herpes-pox primase families, PrimPol belonging to the latter. Members of the herpes-pox primase family possess a conserved C-terminal β-strand-rich region, which replaces the PriCT domains of the iridovirus primase family (Iyer et al., 2005). The NCLDV-herpesvirus primase clade is suggested to have originated from bacteriophage or bacterial proteins possessing a fused AEP and PriCT-2 domain. Herpes viruses likely acquired their primase from the NCLDV class, before replacing the C-terminal PriCT domain with the characteristic β-strand-rich region (Iyer et al., 2005).

PrimPol orthologues are conserved across vertebrates, plants, and primitive eukaryotes including species of fungi, algae, and protists, such as apicomplexans and the slime mold *Dictyostelium*. PrimPol is, however, notably absent from prokaryotes and a number of fungi and animal species, including *Caenorhabditis elegans* and *Drosophila* (Bianchi et al., 2013; García-Gómez et al., 2013; Iyer et al., 2005). This interrupted distribution of PrimPol, coupled with the diversity of AEPs observed in mobile elements, such as viruses and plasmids (Guilliam et al., 2015b), suggests that PrimPol was originally obtained through HGT by an early eukaryote and then lost on multiple separate

occasions. Importantly, PrimPol is not closely related to the eukaryotic replicative DNA primase small subunit, Prim1, a member of the AEP-proper clade, and is dispensable for DNA replication in higher eukaryotes (Bianchi et al., 2013; Iyer et al., 2005). It has been speculated that PrimPol may have originated as a DNA repair enzyme in NCLDVs, potentially required due to their large genome size and lack of access to cellular DNA repair enzymes during replication (Iyer et al., 2005). Likewise, PrimPol may have played a role in DNA replication initiation in these viruses.

6.3. What can PrimPol do? The Domain Architecture and Catalytic Activities of PrimPol

Since the initial identification of PrimPol in 2005 (Iyer et al., 2005), a number of groups have purified and characterised the recombinant protein, permitting insight into the architectural and biochemical properties of the enzyme (Bianchi et al., 2013; García-Gómez et al., 2013; Wan et al., 2013). These studies revealed PrimPol's impressive range of nucleotidyl transferase activities, suggesting a number of potential roles *in vivo*. In this section, we will describe these activities and the domain architecture of the protein, which underpins its catalytic flexibility (summarised in Figure 6.1.).

6.3.1. Domain Architecture and Structure

Previously, an alignment of PrimPol homologues identified 14 conserved regions within the protein, including three characteristic AEP catalytic motifs (motifs I, II, and III) towards the N-terminus, forming the AEP domain (Iyer et al., 2005). Interestingly, motif I displays the variant DxE, rather than the typical DxD motif possessed by most AEPs. Motif I and motif III (xD) together form the divalent metal ion binding site and are essential for the catalytic activity of the enzyme. Motif II (SxH) was predicted to form part of the nucleotide binding site, and is again required for all catalytic activity (Bianchi et al., 2013; García-Gómez et al., 2013; Keen et al., 2014b; Rudd et al., 2013). Recently, the crystal structure of the AEP domain of PrimPol (residues 1-354) in ternary complex with a DNA template-primer and incoming nucleotide was solved, confirming the existence and role of these motifs and two additional motifs, Ia (RQ) and Ib (QRhY/F), which interact with the template DNA strand (Rechkoblit et al., 2016). The structure reveals that PrimPol's catalytic core encloses the 3'-end of the primer with two α/β modules, ModN and ModC, lining the cavity. ModN primarily interacts with the template strand, whilst ModC contains the catalytic residues and interacts with the incoming nucleotide, as well as the template strand. Intriguingly, the structure of PrimPol's AEP domain does not resemble a typical polymerase fold in any way. There is no thumb domain to hold the primer-template, in

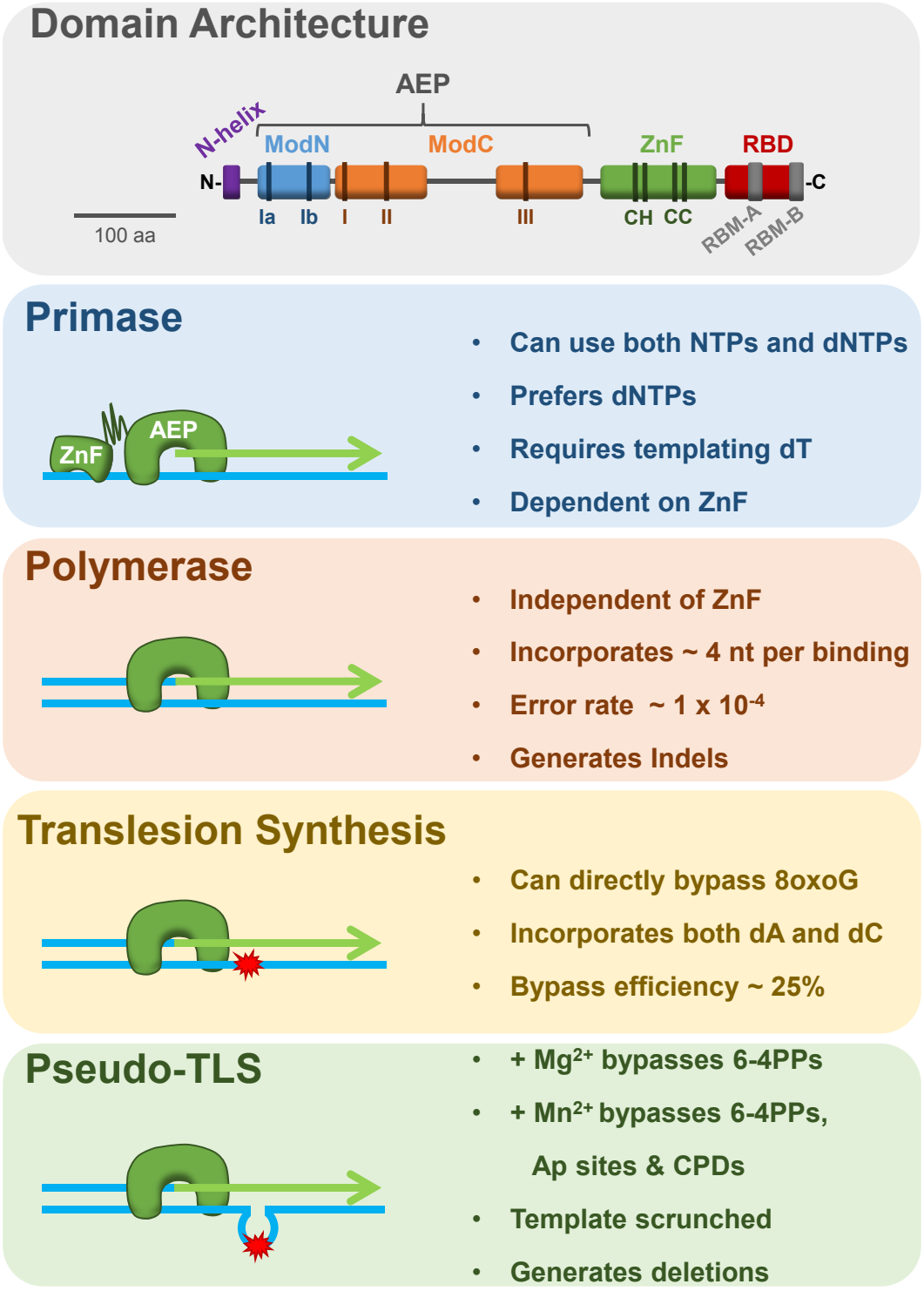


Figure 6.1. Domain architecture and catalytic activities of PrimPol.
The domain architecture of PrimPol is depicted in the top panel. A helix (purple) located at the N-terminus is connected to ModN by a flexible linker and contacts the DNA major groove. ModN (blue) and ModC (orange) comprise the archaeo-eukaryotic primase (AEP) domain and contain motifs Ia, Ib, I, II, and III, required for template binding and catalytic activity. The zinc finger (ZnF) (green) contains three conserved cysteines and a histidine which coordinate a zinc ion and are required for primase, but not polymerase, activity. The RPA binding domain (RBD) (red) containing RPA binding motif-A (RBM-A) and RBM-B (grey) is located at the C-terminus. A 100 amino acid (aa) scale bar is shown to the right. The catalytic activities of PrimPol are displayed below.

fact the primer DNA strand almost completely lacks protein contacts, and ModC was shown to function as both the finger and palm domains (Rechkoblit et al., 2016).

PrimPol's second conserved domain, the C-terminal UL52-like ZnF contains three conserved cysteines and a histidine, as is typical for herpes-pox primase family members (Iyer et al., 2005). The first conserved cysteine and histidine residues of this domain coordinate a zinc ion and are critical for the primase, but not polymerase, activity of the enzyme (Keen et al., 2014b; Mourón et al., 2013).

6.3.2. Primase Activity

As predicted by the initial *in silico* identification, PrimPol is an active primase, able to utilise both NTPs and dNTPs for primer synthesis, a unique ability amongst eukaryotic enzymes (Bianchi et al., 2013; García-Gómez et al., 2013; Wan et al., 2013). Surprisingly, PrimPol actually displays a preference for dNTPs over NTPs during primer synthesis, a feature more typically associated with archaeal primases (Bianchi et al., 2013). This primase activity is dependent upon an intact ZnF domain, which is consistent with previous studies on the Herpes Simplex Virus Type 1 (HSV1) helicase/primase complex. Here, primase activity was lost when key residues in the UL52 zinc-binding domain were mutated (Biswas and Weller, 1999; Chen et al., 2005). Interestingly, the ZnF domain of PrimPol has been shown to bind ssDNA but not dsDNA, suggesting that this module may be important for stabilising PrimPol on ssDNA templates to allow synthesis of the initial dinucleotide (Keen et al., 2014b). The ability of PrimPol to synthesise DNA primers *de novo* gives it the potential to reprime and restart replication downstream of DNA damage lesions and fork-stalling obstacles *in vivo*.

6.3.3. Polymerase and Lesion Bypass Activities

In addition to its DNA and RNA primase activity, PrimPol is also a template-dependent DNA polymerase, with an ability to bypass a number of DNA damage lesions. Notably, PrimPol can bypass both oxidative and UV-induced lesions, including 8-oxo-dG, and 6-4PPs (Bianchi et al., 2013; García-Gómez et al., 2013). A recent study analysing the kinetics of 8-oxo-dG bypass by PrimPol found that the enzyme incorporates dC (error free) opposite the lesion with 6-fold higher efficiency than dA (error prone). Incorporation of dC opposite 8-oxo-dG occurred at ~25% efficiency compared to an unmodified templating dG, suggesting that PrimPol has the potential to function as an efficient TLS polymerase *in vivo* (Zafar et al., 2014). However, the accuracy of bypass differs in other reports, in some instances being only 50% error-free (Bianchi et al., 2013; Guilliam et al., 2016; Keen et al., 2014b; Stojkovič et al., 2016). In the case of 6-4PPs, PrimPol

bypasses the lesion in an error-prone manner (Bianchi et al., 2013). Although unable to directly traverse a CPD, PrimPol can extend from mis-matched bases opposite a CPD (Bianchi et al., 2013). Additionally, a truncated form of PrimPol, lacking the ZnF domain, can facilitate TLS past a CPD (Keen et al., 2014b). In contrast, in the presence of manganese, PrimPol's TLS activity is altered allowing the full-length enzyme to extend past CPDs and Ap sites, in addition to 6-4PPs and 8-oxo-dG lesions (Mourón et al., 2013). However, the usage of either magnesium or manganese as the primary cofactor for PrimPol *in vivo* remains unclear.

6.3.4. Lesion Skipping and Template Independent Extension

Despite displaying the ability to directly read through some damaged nucleobases, such as 8-oxo-dG, it appears that PrimPol's bypass of more bulky or distorting lesions is facilitated through a pseudo-TLS mechanism. Here, PrimPol is able to re-anneal the primer to a new position downstream of the lesion prior to extension, thus looping out the templating lesion and generating a shorter extension product than would be produced from strict template-dependent extension (García-Gómez et al., 2013; Martínez-Jiménez et al., 2015; Mourón et al., 2013). This activity is enhanced in the presence of manganese, permitting bypass of 6-4PPs, CPDs, and Ap sites by pseudo-TLS (García-Gómez et al., 2013; Mourón et al., 2013). Intriguingly, this characteristic is reminiscent of the Ap site bypass strategy employed by PolDom (Pitcher et al., 2007; Yakovleva and Shuman, 2006). The ability of manganese to stimulate primer-realignment and template scrunching by PrimPol offers a clear explanation for the altered TLS ability of the enzyme in the presence of this metal ion. It has also been reported that manganese increases both PrimPol's polymerase activity and affinity for DNA, when compared to magnesium (Zafar et al., 2014).

Notably however, manganese also promotes promiscuous template-independent extension by PrimPol, resulting in the generation of non-complementary homopolymeric strands (Keen et al., 2014b). The mutagenic effect of manganese on polymerase activity, through increased reactivity and promotion of non-template-directed nucleotidyl transfer, has been clear for several decades (El-Deiry et al., 1984; Goodman et al., 1983; Pelletier et al., 1996; Vaisman et al., 2005; Wang et al., 1977). Moreover, the bypass of lesions via template scrunching is potentially more detrimental than beneficial to genomic integrity, due to the high risk of generating frame-shift mutations. Therefore, it seems likely that more low-risk mechanisms would be employed *in vivo* where available.

The lower affinity of PrimPol for DNA and incoming nucleotides in the presence of magnesium is often taken as support for manganese as the enzyme's primary metal ion

cofactor *in vivo* (Zafar et al., 2014). However, PrimPol's inherent low affinity for DNA and dNTPs, when using magnesium as a cofactor, may actually act as an important mechanism to regulate its activity. In support of this, it has previously been shown that dNTP levels in yeast are increased 6-8 fold in the presence of DNA damage (Chabes et al., 2003). Importantly, TLS polymerases often require ~10 times greater dNTP concentrations for nucleotide binding opposite a lesion, compared to a replicative polymerase at an undamaged site (Minko et al., 2003; Shimizu et al., 2002). Increased intracellular dNTP concentrations have been found to correlate with an increase in damage tolerance, but also increased mutation rates, potentially due to the unregulated participation of TLS polymerases in "normal" replication (Chabes et al., 2003). Thus, in yeast it appears that the *in vivo* activity of TLS polymerases is partly regulated by dNTP levels, which increase after DNA damage, consequently restricting the contribution of these polymerases to "normal" DNA replication. Intriguingly, ribonucleotide reductase has been found to be up-regulated in response to DNA damage in all studied organisms, suggesting that increased dNTP synthesis in response to damage may be a conserved mechanism across all domains of life (Sabouri et al., 2008). Similarly, PrimPol's relatively poor affinity for DNA may be overcome *in vivo* by association with other factors, such as RPA and PolDIP2, again acting to regulate the enzyme by only recruiting it to loci where it is actually required (Guilliam et al., 2015a, 2016, 2017; Wan et al., 2013).

Additionally, it is not clear whether the relatively low intracellular concentrations of manganese (0.1 to 40 μM) (Ash and Schramm, 1982; Markesbery et al., 1984; Versieck and McCall, 1985), compared to magnesium (0.21 to 0.24 mM) (Gee II et al., 2001; Goldschmidt et al., 2006), are sufficient to support the manganese-dependent TLS activities of PrimPol *in vivo.* Indeed, PrimPol required manganese concentrations of 200-1000 μM to facilitate pseudo-TLS bypass of an abasic site *in vitro*, whilst 100 μM did not permit any observable bypass (Martínez-Jiménez et al., 2015). Thus, the cellular relevance of these activities is not immediately clear. One intriguing possibility is that PrimPol utilises manganese in the mitochondria only (Zafar et al., 2014). Here, dNTP concentrations are lower than those in the cytosol, there is a dearth of TLS polymerases, manganese uptake is increased in response to oxidative stress, and the high copy number nature of mtDNA may allow more promiscuous lesion bypass mechanisms to be employed (Rampazzo et al., 2004; Zafar et al., 2014). Although, more recent *in vitro* reconstitution experiments argue against a TLS-like role for PrimPol in oxidative damage bypass during mitochondrial DNA replication (Stojkovič et al., 2016).

6.3.5. Fidelity, Mutagenic Signature, and Processivity

Typically, the price paid by polymerases for DNA damage tolerance is a significant decrease in both fidelity and processivity. Whilst the structural features of replicative polymerases confer extremely efficient and high fidelity DNA synthesis, TLS polymerases possess more spacious active sites, altered finger and thumb domains, and lack proofreading exonuclease capabilities. These characteristics permit bypass of bulky lesions, but result in greatly decreased fidelity and processivity on undamaged DNA templates (Sale et al., 2012). Likewise, the eukaryotic replicative primase exhibits poor fidelity compared to replicative polymerases (Cotterill et al., 1987; Sheaff and Kuchta, 1994; Zhang and Grosse, 1990). Rather unsurprisingly, PrimPol, which combines both TLS and primase capabilities, exhibits high error rates of ~ 1 x 10^{-4}, comparable with Y and X-family polymerases, in assays described in Chapter 2 (Guilliam et al., 2015a). Unlike these polymerases however, PrimPol generates indel errors at a much higher frequency than substitution mutations, which may be a result of its template scrunching ability (Guilliam et al., 2015a). Manganese acts to further decrease PrimPol's fidelity on undamaged DNA and even more so on 8-oxo-dG containing templates (Zafar et al., 2014). In addition to poor fidelity, PrimPol shares the characteristic of low processivity with canonical TLS polymerases, incorporating only 1-4 nucleotides per binding event (Keen et al., 2014b). Intriguingly, the enzyme's processivity was found to be negatively regulated by its ZnF domain, which may act to stabilise DNA binding and allow primer synthesis, whilst additionally limiting primer extension. Removal of the ZnF domain has also been found to lower PrimPol's fidelity, suggesting the domain acts to regulate processivity and fidelity, as well as enabling primase activity (Keen et al., 2014b).

6.4. What does PrimPol do? The Role of PrimPol in DNA Replication

The biochemical classification of PrimPol as both an RNA/DNA primase and a TLS polymerase clearly suggests a role in DNA replication and damage tolerance. Moreover, these two characteristics give PrimPol the potential to assist the replisome in two different ways; through TLS or repriming. In this section, the *in vivo* characterisation of the enzyme, as well as the consequences of its deletion on the cell, are described. Using this information, recent advances in our understanding of the cellular roles of PrimPol, including work presented in the preceding chapters of this thesis, will be discussed. These emerging roles for the enzyme are outlined in Figure 6.2.).

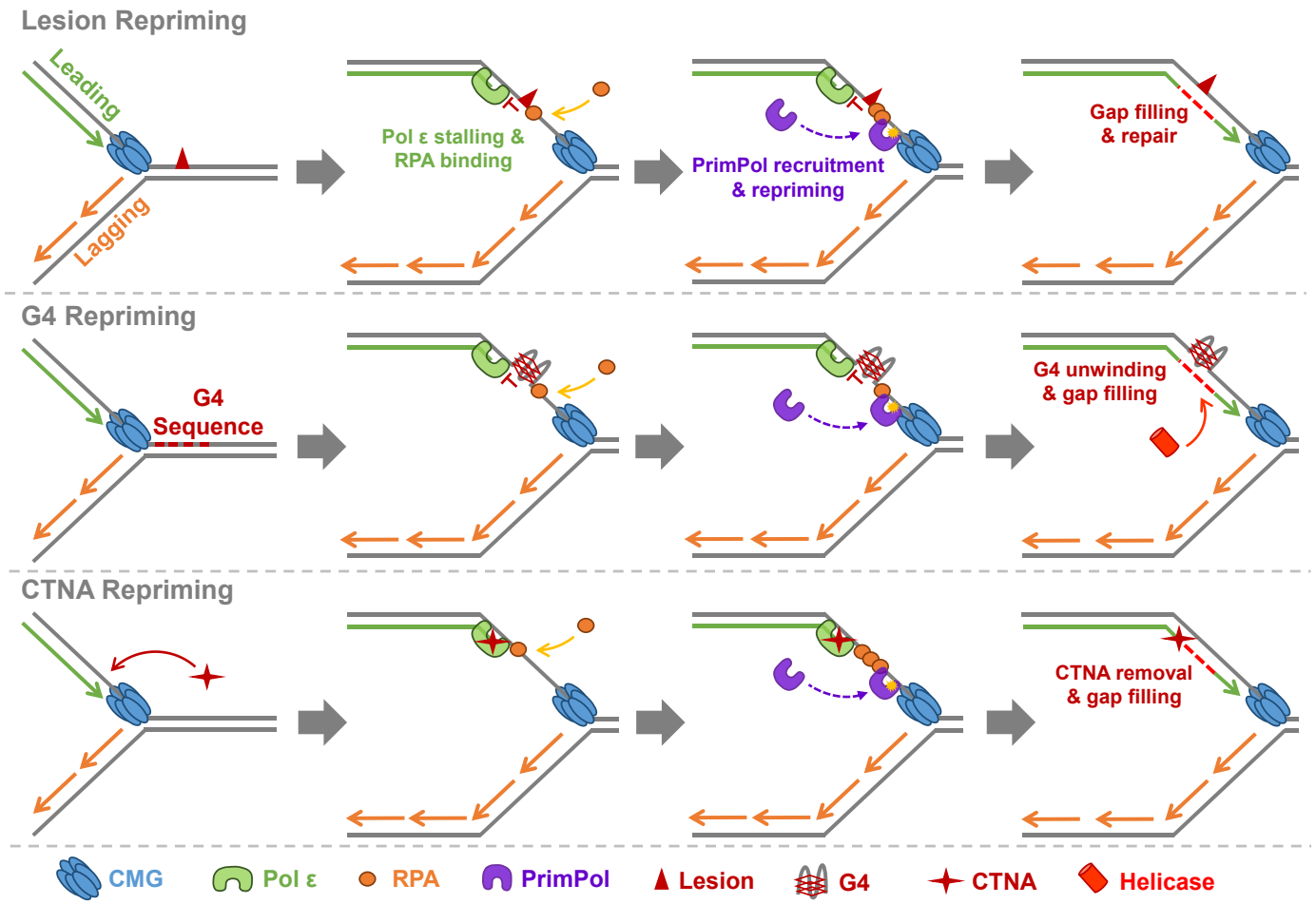


Figure 6.2. Repriming roles of PrimPol in nuclear DNA replication.
PrimPol is able to reprime and reinitiate leading strand replication downstream of a range of replicase stalling obstacles. Here, the ability of PrimPol to reprime downstream of DNA lesions, G4 secondary structures, and chain-terminating nucleotide analogues, is highlighted. Following repriming, replication can proceed and the resulting ssDNA gap is filled through TLS or template switching mechanisms, permitting subsequent repair or removal of the obstacle. Only the CMG complex, Pol ε, PrimPol, and RPA, are shown for simplicity. A key for identifying each factor is shown below.

6.4.1. PrimPol – a DNA Damage Tolerance Enzyme

DDT is critical to support continued replisome progression in the presence of unrepaired DNA damage. An inability to tolerate this damage can lead to prolonged fork stalling, collapse and, ultimately, genome instability and/or cell death. The importance of DDT in preserving genomic integrity is highlighted by the consequences on human health of dysfunction in these mechanisms. An obvious example is the variant form of xeroderma pigmentosum (XPV). Here, mutation of Pol η causes increased sensitivity to sunlight and a predisposition to skin cancer (Sale et al., 2012). This is thought to occur due to mutagenic bypass of UV-induced CPDs by alternative TLS polymerases.

Interestingly, loss of PrimPol in human XPV cells leads to a synergistic increase in UV sensitivity, with the enzyme performing a distinct role from Pol η during this process (Bianchi et al., 2013). In line with this, PrimPol forms sub-nuclear foci, and is recruited to chromatin, in response to UV irradiation (Bianchi et al., 2013; Mourón et al., 2013). Both human MRC5 and avian DT40 cells lacking PrimPol (PrimPol$^{-/-}$) also accumulate an increased number of stalled forks, or a reduced ability to restart stalled forks, following UV damage (Bianchi et al., 2013; Guilliam et al., 2016; Mourón et al., 2013; Wan et al., 2013). Unlike human cells, DT40 cells are hypersensitive to UV irradiation in the absence of PrimPol only, potentially due to the faster doubling times and increased S-phase population of these cells (Bianchi et al., 2013; Kobayashi et al., 2016). Interestingly, it has recently been shown that PrimPol$^{-/-}$ DT40 cells are even more sensitive to UV damage than previously appreciated (Bailey et al., 2016). In fact, these cells were found to be more sensitive than those lacking Pol η when analysed by colony formation assays. This effect was determined to be due to an extended G2 arrest, which prevented cell cycle progression, rather than an increase in apoptosis (Bailey et al., 2016). These reports clearly implicate PrimPol in the maintenance of replisome progression, or restart of stalled replication forks, in the presence of UV damage lesions.

However, PrimPol is also involved in the tolerance of other types of DNA damage. As discussed in Chapter 4, PrimPol$^{-/-}$ DT40 cells are hypersensitivity to MMS, cisplatin, and HU (Kobayashi et al., 2016). Further deletion of Pols ζ and η in these cells leads to an additional increase in damage sensitivity to a similar extent as in wild-type cells, again indicating an independent role for PrimPol in DNA damage tolerance (Kobayashi et al., 2016). PrimPol is also required for recovery of stalled replication forks following HU treatment in HeLa cells (Mourón et al., 2013; Wan et al., 2013). Notably, each of these DNA damaging agents acts to stall the progression of replication fork. In contrast, loss of PrimPol does not sensitise cells to ICRF193, camptothecin, or γ-rays, agents which

lead to DNA strand breaks. This is suggestive of a broad role for the enzyme in damage tolerance, but not in break repair (Kobayashi et al., 2016). PrimPol associates with chromatin during G1 and S-phase and PrimPol$^{-/-}$ mouse embryonic fibroblasts (MEFs) present chromosome aberrations indicative of S-phase defects, which are enhanced after aphidicolin treatment (Bianchi et al., 2013; Mourón et al., 2013). Collectively, these findings place PrimPol at the replication fork during S-phase and indicate a role in the tolerance of replicase-stalling DNA damage.

6.4.2. PrimPol Reprimes and Restarts Stalled Replication Forks

The DDT defects observed in the absence of PrimPol are potentially indicative of both a role as a TLS polymerase and a repriming enzyme. However, more recent reports clearly support the latter (Keen et al., 2014b; Kobayashi et al., 2016; Mourón et al., 2013; Pilzecker et al., 2016; Schiavone et al., 2016). Although PrimPol is described as a TLS polymerase, the spectrum of DNA damage types it can traverse by "true" TLS is actually rather limited. Discounting pseudo-TLS bypass, which may or may not be relevant *in vivo*, PrimPol is essentially only able to directly bypass 8-oxo-dG lesions (Bianchi et al., 2013; García-Gómez et al., 2013; Keen et al., 2014b; Mourón et al., 2013). Moreover, a number of other polymerases are also able to efficiently and accurately bypass these lesions (Haracska et al., 2000; Maga et al., 2013; Zahn et al., 2011). If PrimPol's primary role were as a TLS polymerase, this observation would be at odds with the range of replicase-stalling DNA damaging agents it is involved in tolerating (Kobayashi et al., 2016). This implies that PrimPol most likely acts as a repriming enzyme for the tolerance of DNA damage and this is supported by the study of separation of function mutants (Keen et al., 2014b; Kobayashi et al., 2016; Mourón et al., 2013; Schiavone et al., 2016).

Mutation of PrimPol's ZnF domain abolishes primase, but actually enhances polymerase activity (Keen et al., 2014b; Mourón et al., 2013). This important observation has permitted investigation into the requirement of primase activity for the enzyme's role in DNA replication. In each case, when PrimPol$^{-/-}$ cells are complemented with the primase-deficient / polymerase-proficient ZnF mutant, it is unable to rescue any of the observed damage tolerance defects (Keen et al., 2014b; Kobayashi et al., 2016; Mourón et al., 2013). In contrast, complementation with a primase-proficient / reduced-polymerase mutant of PrimPol restored DNA damage tolerance to wild-type levels (Keen et al., 2014a; Kobayashi et al., 2016). In agreement with this, in Chapter 4 it was shown that PrimPol is able to facilitate close-coupled repriming downstream of lesions which it cannot bypass by TLS *in* vitro (Kobayashi et al., 2016). Aside from increased sensitivity to DNA damaging agents and decreased replication fork rates, PrimPol$^{-/-}$ and knockdown

cells exhibit persistent RPA foci and increased phosphorylation of Chk1 (Mourón et al., 2013; Wan et al., 2013). Both of these stress response markers are indicative of the generation of ssDNA stretches (Choi et al., 2010; Lopes et al., 2006). This would be an expected consequence of a lack of repriming by PrimPol, resulting in the uncoupling of leading / lagging strand replication and excessive strand-specific unwinding by MCM (Lopes et al., 2006). In agreement, cells compensate for the loss of PrimPol by increasing both HR mediated fork rescue and dormant origin firing (Bianchi et al., 2013; Mourón et al., 2013). These compensatory back-up mechanisms, in addition to redundancy between PrimPol and TLS polymerases, may explain why PrimPol is dispensable for viability in human cells and mouse models (García-Gómez et al., 2013). These observations give further credibility to a place for PrimPol at the progressing replication fork during S-phase, which might not necessarily be the case if a TLS-like role was being performed.

As previously mentioned, TLS can potentially occur both at the replication fork, as well as post-replicatively, to fill in gaps left opposite lesions following repriming or dormant origin firing (Sale et al., 2012). Each of these possibilities are not mutually exclusive, but a number of studies point to post-replicative gap-filling as the predominant role for TLS. In yeast, DDT mechanisms, including TLS, have been found to operate effectively in a post-replicative manner, and ssDNA gaps, indicative of repriming, accumulate following UV-damage (Daigaku et al., 2010; Karras and Jentsch, 2010; Lopes et al., 2006). Likewise, in human cells DNA replication fork progression in the presence of UV damage was found to be independent of TLS and ssDNA gaps opposite UV lesions were identified. It was concluded that these gaps were likely a result of repriming downstream of lesions rather than dormant origin firing (Elvers et al., 2011). Importantly, mutation of Pol η or other TLS factors does not appear to significantly alter replication fork rates in the presence of damage (Edmunds et al., 2008; Elvers et al., 2011). This is in stark contrast to the effect of loss of PrimPol on replication fork progression following damage, further supporting a repriming, rather than TLS, role for this enzyme *in vivo*.

6.4.3. PrimPol Bypasses Non-Canonical Replication Impediments

Whilst DNA damage lesions are some of the best characterised replication impediments, they are not the only obstacles replication forks must overcome during their progression. In addition to the right-handed B-form of dsDNA, we have become familiar with since Watson and Crick's famous model (Watson and Crick, 1953), genomic DNA can also adopt a number of other secondary structures as a result of specific sequence motifs and protein interactions (Bochman et al., 2012). One alternative DNA secondary

structure, which has received increasing attention as evidence for its formation *in vivo* grows, is the G-quadruplex (Murat and Balasubramanian, 2014). As previously mentioned, G-quadruplexes are produced by the stacking of G-quartets, which form through alternative Hoogsten base-pairing between guanine bases. These structures may potentially play an important role in transcription and DNA replication in the cell, but they can also pose as major impediments to replisome progression (Cheung et al., 2002; Maizels and Gray, 2013; Ribeyre et al., 2009; Sarkies et al., 2010). Consequently, cells possess a number of specialised helicases and polymerases to replicate past G-quadruplexes (León-Ortiz et al., 2014).

Previously, cells lacking FANCJ or Rev1 were found to stochastically lose Bu-1a protein expression (Sarkies et al., 2011; Schiavone et al., 2014). Importantly, the *BU-1A* locus contains a G-quadruplex, which was determined to stall replication in these cells. This stalling causes uncoupling of replication from histone recycling at the *BU-1A* locus and consequently leads to the deletion of epigenetic marks, manifesting in loss of Bu-1a expression. As described in Chapter 4, it was recently identified using Bu-1a read-out assays that PrimPol also plays a critical role in the bypass of these structures during DNA replication (Schiavone et al., 2016). Consistent with PrimPol's behaviour at most DNA damage lesions, *in vitro* analysis shown in Chapter 4 revealed that the enzyme is unable to directly read through G-quadruplexes, but can bind to and facilitate close-coupled repriming downstream of these structures. Vitally, close-coupled repriming ahead of the G-quadruplex would permit the appropriate recycling of histones, and thus maintain epigenetic marks and Bu-1a expression. Bypass of G-quadruplex structures through repriming by PrimPol was confirmed *in vivo* using the ZnF primase-deficient mutant discussed previously. Here, complementation of PrimPol$^{-/-}$ cells with the ZnF mutant failed to prevent instability of Bu-1a expression, in contrast to the wild-type protein, confirming that PrimPol's primase activity is critical for G-quadruplex bypass (Schiavone et al., 2016). Intriguingly, PrimPol was found to only be required for G-quadruplex bypass during leading strand replication. Presumably, this is because primers are constantly generated on the lagging strand anyway, due to the discontinuous nature of DNA synthesis there.

Further evidence of a general role for PrimPol in repriming replication downstream of fork-stalling obstacles is provided by studies of CTNAs which are detailed in Chapter 4 (Kobayashi et al., 2016). CTNAs cause replication to stall by being incorporated into the 3'-termini of growing DNA polymers and preventing further extension, as they lack the 3' hydroxyl required for phosphodiester bond formation (Berdis, 2008; De Clercq and Field, 2006). Loss of PrimPol has been shown to cause hypersensitivity to a wide range of

CTNAs (Kobayashi et al., 2016). Critically, the inability of polymerases to extend from CTNAs rules out bypass by direct extension. PrimPol was found to be important for the tolerance of CTNAs by repriming downstream. This role was confirmed by both *in vivo* characterisation of the ZnF mutant and *in vitro* analysis of repriming after CTNAs using the primase assay described in Chapter 4 (Kobayashi et al., 2016).

These critical findings not only identify that PrimPol bypasses G-quadruplexes and CTNAs in a similar manner to DNA damage lesions, through repriming. They point to the possibility that PrimPol is able to bypass a wide range of leading-strand obstacles during normal and perturbed replication. This is in contrast to canonical TLS polymerases which are typically highly specialised in the lesions they can bypass. Consequently, it is likely that PrimPol is broadly employed as a general mechanism to reprime and restart replication ahead of many different leading strand replication impediments.

6.4.4. A Role for PrimPol in Mitochondrial DNA Replication?

The majority of genetic information in mammalian cells is stored in the nucleus. However, a small proportion of DNA is also located in the mitochondria. Despite being only ~16.6 kb long and encoding just 13 polypeptides, mutation of the mitochondrial genome is responsible for a number of mitochondriopathies and is implicated in various other pathologies including cancer, cardiovascular diseases, and neurodegenerative disorders (Alexeyev et al., 2013). Unlike nuclear DNA, cells possess many copies of mtDNA making it highly redundant. In line with this, the rate of mutagenesis is ~10-fold greater in the mitochondria than the nucleus (Alexeyev et al., 2013). A major function of mitochondria is the generation of ATP through OXPHOS. This process produces reactive ROS which can induce damage lesions, including 8-oxo-dG and Ap sites, in mtDNA (Alexeyev et al., 2013).

A significant proportion of PrimPol has been found to localise to the mitochondria where it interacts with mtSSB, suggesting a potential role in the tolerance of mtDNA damage (Bianchi, 2013; Bianchi et al., 2013; García-Gómez et al., 2013; Guilliam et al., 2015a). This is supported by defects in mtDNA replication and copy number observed in cells lacking PrimPol (Bianchi, 2013; García-Gómez et al., 2013). However, the ability to generate viable PrimPol$^{-/-}$ mice demonstrates that this role is redundant. Indeed, PolRMT is likely responsible for generating the initial primers essential for mtDNA replication (Falkenberg et al., 2007). These primers are then extended by Pol γ, which until recently was thought to be the only mitochondrial DNA polymerase (Loeb and Monnat, 2008). In addition to PrimPol, more recent reports indicate that Pol θ and Pol ζ also function in human mtDNA replication (Singh et al., 2015; Wisnovsky et al., 2016).

Given that few TLS polymerases appear to localise to the mitochondria, in addition to the high levels of ROS there, it was speculated that PrimPol may be involved in TLS bypass of mitochondrial 8-oxo-dG lesions and Ap sites (García-Gómez et al., 2013). In order to investigate this, a recent study analysed the ability of PrimPol to assist the mitochondrial replisome in oxidative damage bypass by TLS (Stojkovič et al., 2016). Here, it was found that the mitochondrial replisome is completely stalled by Ap sites and pauses significantly at 8-oxo-dG lesions. PrimPol did not enhance the bypass of either of these lesions, disagreeing with a TLS role in oxidative damage bypass in the mitochondria (Stojkovič et al., 2016). Thus, it seems more likely that PrimPol functions to reprime mtDNA replication downstream of blocking lesions, similar to its role in the nucleus. In addition to oxidative damage, mtDNA is also subject to deletions. Intriguingly, these deletions map in close proximity to G4-forming sequences (Bharti et al., 2014). In light of the role of PrimPol in repriming after G-quadruplexes in nuclear DNA replication, it would not be surprising if the enzyme fulfilled the same role in the mitochondria. However, further work is required to confirm a repriming role for PrimPol here.

6.4.5. Is PrimPol Involved in Somatic Hypermutation?

Generally, mutagenesis during DNA replication is avoided at all costs in order to preserve genomic stability. However, an exception to this is during the development of the immune system. Here, mutagenesis occurs in immunoglobulin (Ig) genes to enable variation in the generated antibodies. This programmed mutagenesis is driven by activation-induced deaminase (AID), which deaminates dC to dU (Noia and Neuberger, 2007). Replication of dU facilitates C>T transitions. Additionally, dU may be further processed by uracil DNA glycosylase (UNG) to generate Ap sites. TLS bypass of these Ap sites can alternatively create C>A/G/T mutations due to the non-instructive nature of the lesion (Sale et al., 2009).

The involvement of TLS polymerases in somatic hypermutation (SHM) at Ap sites led to speculation that, if PrimPol functions as a TLS Pol *in vivo,* it might also modulate this mutagenesis. Analysis of DT40 cells found that hypermutation and gene-conversion events are similar in wild-type and PrimPol$^{-/-}$ cells. Moreover, loss of PrimPol in wild-type and Pol η$^{-/-}$/Pol ζ$^{-/-}$ cells did not significantly alter the mutation spectrum of the studied Ig gene (Kobayashi et al., 2016). Intriguingly, another report, which analysed large mutational data sets in mice, identified that PrimPol does have a subtle effect on SHM outcome (Pilzecker et al., 2016). In this analysis, loss of PrimPol was found to selectively increase C>G transversions, but did not affect other G/C or A/T mutations. Interestingly, PrimPol was found to specifically prevent the generation of C>G transversions in the

leading strand, potentially explaining the G>C over C>G strand bias of somatically mutated Igh loci (Pilzecker et al., 2016). However, this anti-mutagenic activity of PrimPol was attributed to the enzyme's primase, rather than TLS polymerase, activity. It was concluded that PrimPol preferentially reprimes downstream of Ap sites on the leading strand, therefore maintaining fork progression and preventing error-prone TLS. The resulting ssDNA gap opposite the Ap site could then be filled in by error-free homology directed repair. Fascinatingly, in the same report, studies of invasive breast cancers suggested that this leading strand anti-mutagenic activity of PrimPol may be genome wide.

Together, these reports establish that PrimPol does not act as a canonical TLS polymerase during SHM. Rather, PrimPol affects the mutational outcome of SHM by repriming downstream of Ap sites on the leading strand and thus preventing C>G transversions. These findings, therefore, further support mounting evidence that PrimPol's primary role in DNA damage tolerance is to reprime leading strand replication and not to perform TLS.

6.4.6. Why Doesn't Pol α-Primase Reprime Leading Strand Replication?

The emerging role for PrimPol in repriming leading strand replication begs the question; why doesn't the replicative Pol α-primase complex fulfil this role? In *E. coli,* DnaG, the replicative primase, efficiently reprimes replication ahead of replicase stalling DNA damage lesions, permitting bypass of the damage without dissociation of the replisome (Yeeles and Marians, 2011, 2013). Likewise in yeast, which lack PrimPol, leading strand repriming is presumably facilitated by Pol α-primase, suggesting that, at least in these organisms, the replicative primase has the capacity to fulfil this role.

Whilst the answer to this question is not completely clear, PrimPol does have one advantage over Pol α-primase; it preferentially primes using dNTPs. This minimises the amount of RNA processing required on the leading strand. Although ribonucleotides are routinely incorporated during the initiation of each Okazaki fragment on the lagging strand and at replication origins on the leading strand, their persistent presence in DNA can lead to genomic instability (Williams et al., 2016). Ribonucleotides incorporated during primer synthesis are routinely removed through Okazaki fragment maturation (Williams and Kunkel, 2014). However, it is not clear how a DNA secondary structure or lesion requiring bypass upstream of the primer would affect this process.

Ribonucleotides incorporated by the replicative polymerases are removed by ribonucleotide excision repair (RER). Intriguingly, in RER deficient yeast, leading strand ribonucleotides are removed through a Top1 mediated mechanism, which likely also

removes a subset of ribonucleotides in RER proficient cells (Williams et al., 2015; Williams and Kunkel, 2014). This mechanism of ribonucleotide removal, which doesn't appear to occur on the lagging strand, is susceptible to causing genome instability. This makes ribonucleotides present in the leading strand potentially more detrimental than those in the lagging strand. This is supported by observations that loss of RER and increased ribonucleotide incorporation by Pol ε, but not Pol α or Pol δ, is lethal (Williams et al., 2015).

Although RER deficient yeast are viable, loss of this pathway in mice results in embryonic lethality (Reijns et al., 2012). Thus, the greater pressure on higher eukaryotes to minimise ribonucleotide presence in the genome may explain why PrimPol is utilised for leading strand repriming with dNTPs in these organisms, but has been lost in some lower eukaryotes such as yeast.

6.4.7. Why is PrimPol Damage Tolerant *In Vitro*?

If PrimPol's primary role *in vivo* is to reprime DNA replication, why does the enzyme display TLS-like activity *in vitro*? Although it is possible that PrimPol's TLS-like activity is important in the cell, recent studies suggest that the enzyme's primase activity is more relevant for its *in vivo* role, as discussed previously. This opens up the possibility that this TLS activity is a "side effect" of being a primase and this is supported by a number of observations.

Recent studies of the RNA primase domains of human Pol α-primase provide insight into the unique way primases interact with their DNA template and primer (Baranovskiy et al., 2016a, 2016b). The RNA primase associated with Pol α is a heterodimer composed of a small catalytic subunit, Prim1, and a large regulatory subunit, Prim2. These reports identify that the CTD of Prim2 binds to the DNA/RNA junction at the 5'-end of the RNA primer, whereas Prim1 binds and extends the 3' end of the primer moving away from Prim2. Prim1 makes few contacts with the DNA/RNA, resulting in distributive activity. By only contacting the primer at the 5' and 3' ends, the primase is unable to sense modified nucleotides in the RNA strand, potentially explaining the propensity of primases to perform TLS-like extension (Baranovskiy et al., 2016a, 2016b). The authors suggest that this binding mechanism is broadly applicable to most primases. In the context of PrimPol, the ZnF is likely functionally equivalent of Prim2. Indeed, both are flexibly tethered to their respective catalytic domains and required for template recognition during priming, although PrimPol's ZnF has only been shown to bind ssDNA (Keen et al., 2014b; Liu and Huang, 2015). Nevertheless, the ZnF domain may bind the ssDNA immediately upstream of the 5' end of the primer.

The crystal structure of PrimPol's AEP domain potentially supports this model (Rechkoblit et al., 2016). Here, only the templating base is held in the active-site cleft, with the rest of the 5' template strand directed out of the active site. Additionally, PrimPol lacks a thumb domain and makes few contacts with the primer strand. This potentially prevents the enzyme from sensing damaged bases in the template and allows them to be looped out. Furthermore, unlike TLS polymerases, PrimPol does not possess an "open" active-site cleft and is unable to accommodate bulky lesions such as CPDs and 6-4PPs (Rechkoblit et al., 2016). This adds further evidence that PrimPol is not a "true" TLS polymerase, rather it loops out bulky lesions during bypass, resulting in a propensity to generate indels as identified in Chapter 2.

The ability of primase-polymerases to perform TLS-like extension is well documented (Guilliam et al., 2015b). Some AEPs have co-opted this inherent catalytic versatility for use in other processes such as NHEJ, becoming specialised and in some instances, losing their ability to prime (Guilliam et al., 2015b). However, PrimPol's primase activity is critical for its role *in vivo* and thus it is possible that the TLS-like activities observed *in vitro* simply arise as a by-product of the structural features necessary for priming.

6.5. How does PrimPol get to Where it's Needed? The Recruitment of PrimPol to Stalled Replication Forks

The studies outlined above and presented in Chapter 4 strongly indicate that PrimPol's main role in DNA replication is to reprime ahead of impediments on the leading strand. In order to fulfil this role, PrimPol must be efficiently recruited to ssDNA downstream of stalled replication forks. In this section, we will describe recent advances in our understanding of the interactions and mechanisms governing the recruitment of PrimPol.

6.5.1. PrimPol Interacts with Single-Strand Binding Proteins

Replication fork stalling can cause uncoupling of leading and lagging strand synthesis, consequently generating ssDNA stretches on either strand due to continued unwinding by the replicative helicase (Lopes et al., 2006). The impact of this on the lagging stand is likely limited by the generation of new Okazaki fragments. However, in the absence of leading strand fork restart, extended uncoupling can produce stretches of ssDNA. In nuclear DNA replication, the resulting ssDNA is bound by RPA, which in turn can trigger the S phase checkpoint response (Zou and Elledge, 2003).

Work described in Chapter 2 reveals that unlike TLS polymerases, PrimPol does not interact with PCNA (Guilliam et al., 2015a). However, it does interact with both RPA and

mtSSB (Guilliam et al., 2015a; Wan et al., 2013). PrimPol's interaction with RPA is mediated by its CTD, which binds to the N-terminus of RPA70, the largest subunit of the RPA heterotrimer (Guilliam et al., 2015a). In Chapter 5 the structural basis for PrimPol's interaction with RPA was elucidated (Guilliam et al., 2017), identifying that PrimPol possesses two RBMs in its CTD (RBM-A and RBM-B), both of which bind to the basic cleft of RPA70N independently of each other. Interestingly, this cleft has previously been shown to interact with, and recruit, a number of different DNA damage response proteins, including RAD9, MRE11, ATRIP, and p53 (Xu et al., 2008).

Together, these studies suggest that PrimPol may also be recruited to stalled replication forks through its interaction with RPA; with mtSSB potentially playing an analogous role in mitochondria.

6.5.2. RPA Recruits PrimPol to Stalled Replication Forks

Previously, it was identified that PrimPol's CTD is required for its function and co-localisation with RPA *in vivo* (Wan et al., 2013). However, interpretation of these results is limited as removal of the whole CTD has been shown to reduce primase activity *in vitro*, and may abrogate interactions with other binding partners (Keen et al., 2014b). Structural studies of the PrimPol-RPA complex, presented in Chapter 5, enabled the *in vivo* analysis of point mutants which disrupt this interaction.

This work identified that PrimPol's RBM-A is the primary mediator of the RPA interaction *in vivo*, whilst RBM-B appears to play a secondary role. RBM-A mutants were unable to restore replication fork rates following UV-damage, in comparison to the wild-type or RBM-B mutant protein (Guilliam et al., 2017). These findings revealed that PrimPol's interaction with RPA is required for its cellular role. Moreover, this study also showed that this interaction is responsible for the recruitment of PrimPol to chromatin, demonstrating that the enzyme is indeed recruited to stalled replication forks by RPA (Guilliam et al., 2017). Intriguingly, mutations of key residues at each RBM have been identified in cancer patient cell lines, these mutations are sufficient to abrogate binding of the affected RBM, adding further support that these motifs are important for PrimPol's function *in vivo* (Guilliam et al., 2017). The observation that RBM-A is required for binding of PrimPol to chromatin and its interaction with RPA *in vivo*, suggests that this motif mediates the initial recruitment of PrimPol. An intriguing possibility, consistent with the observations in Chapter 5, is that RBM-B then binds a second RPA molecule upstream of the first to further stabilise the enzyme and promote repriming.

Aside, from identifying the mechanism by which PrimPol is recruited to stalled replication forks; these studies also add to the growing evidence supporting a role for PrimPol as a

repriming enzyme. PrimPol's recruitment to RPA, and lack of interaction with PCNA, suggests it binds to ssDNA downstream of a stalled replicase on the leading strand, the ideal place to facilitate repriming following initial leading/lagging strand uncoupling to prevent excessive ssDNA generation. A recent report investigating the role of Rad51 recombinase in aiding replication across UV lesions supports this (Vallerga et al., 2015, p. 51). Here, Rad51 and Mre11 depletion was found to favour ssDNA accumulation at replication obstacles and subsequent PrimPol-dependent repriming. This also supports previous suggestions that excessive unwinding of DNA following stalling of the replicase is sufficient to promote ssDNA generation and repriming at replication impediments (Elvers et al., 2011).

Further work is required to elucidate the exact mechanisms controlling PrimPol's recruitment by RPA to ssDNA. Interestingly, binding of MRE11 and RAD9 to RPA is enhanced upon RPA32C phosphorylation (Robison et al., 2004; Wu et al., 2005). Thus, phosphorylation of RPA may act as a way to signal recruitment of DNA damage response proteins, potentially including PrimPol (Oakley and Patrick, 2010).

6.6. The Regulation of PrimPol During DNA Replication

Work presented in this thesis and elsewhere, strongly indicates that PrimPol is recruited by RPA to the leading strand, following replicase stalling, in order to reprime replication and prevent genome instability (Guilliam et al., 2015a, 2017; Wan et al., 2013). However, as shown in Chapter 2, PrimPol is an error-prone enzyme and unscheduled or dysregulated activity could lead to mutagenesis (Guilliam et al., 2015a). In this section, the proposed mechanisms which act to limit PrimPol's contribution to DNA synthesis during replication will be discussed (outlined in Figure 6.3.).

6.6.1. Regulation of the Cellular Concentration of PrimPol

The simplest way to regulate the activity of a protein is by controlling its intracellular concentration. This is especially true for proteins which are only required in response to specific stressors, for example DNA damage response proteins. This approach is utilised during the SOS response in *E. coli*. Here, ~40 DNA damage response genes are upregulated in response to DNA damage (Michel, 2005).

In comparison to Prim1, PrimPol is expressed at very low levels in human U2OS cells (< 500 protein copies per cell compared to ~ 13,300) (Beck et al., 2011). This is, however, similar to the expression level of TLS Polymerases, including Pols η and κ. PrimPol mRNA expression peaks in G1-S phase, although the total protein levels remain roughly constant throughout the cell cycle (Mourón et al., 2013). Thus, the increased association

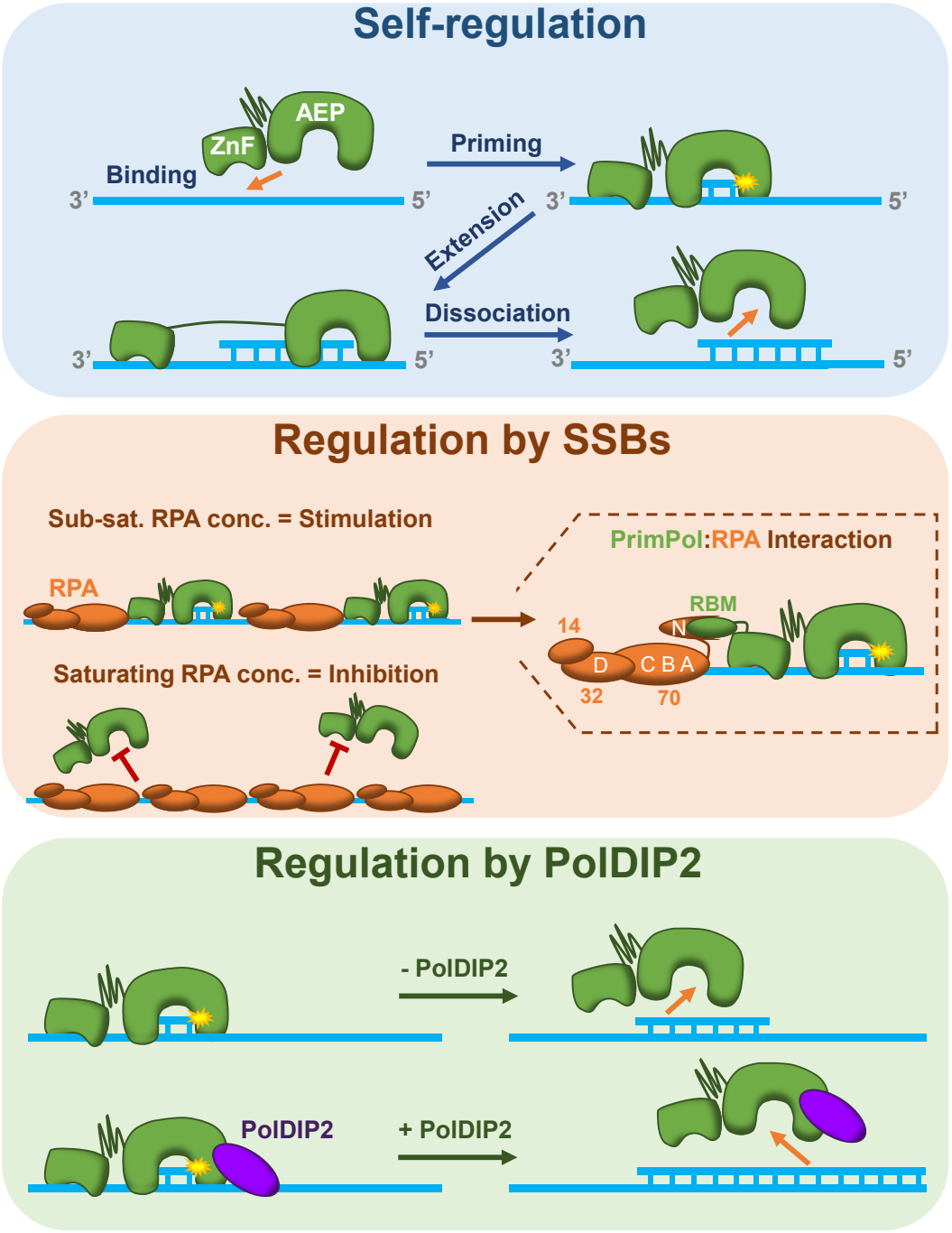


Figure 6.3. Regulation of PrimPol by its ZnF domain and interacting partners.
Top panel: PrimPol is inherently self-regulatory due to the restraining effect of its ZnF domain. The AEP and ZnF domains of PrimPol form a hinge-like structure, connected by a flexible linker. Binding of PrimPol to ssDNA is mediated by the ZnF domain which binds 3' relative to the AEP domain on the template strand. Binding of the ZnF stabilises the AEP domain, permitting primer synthesis. The AEP then extends the primer, but is restricted by the maximum distance it can move away from the ZnF. The enzyme subsequently dissociates leaving behind a short primer. This mechanism limits the processivity of the PrimPol. **Middle panel:** PrimPol is regulated by single-strand binding proteins (SSBs). At sub-saturating concentrations of RPA, the protein acts to recruit PrimPol to the ssDNA template, consequently stimulating primer synthesis. *In vivo* this interaction is primarily mediated by PrimPol's RBM-A which binds to the basic cleft of RPA70N. At saturating RPA concentrations, when the ssDNA template is fully coated, PrimPol cannot gain access and primer synthesis is inhibited. This serves to limit where PrimPol can prime. **Bottom panel:** PolDIP2 enhances PrimPol's primer extension activity by binding the AEP domain and stabilising it on the DNA.

of PrimPol with chromatin during the G1 and S phases of the cell cycle in unperturbed cells is a result of finer mechanisms controlling recruitment to DNA, rather than increased expression. This may also be the case with the increased recruitment of the enzyme to chromatin in response to DNA damage. Nevertheless, the low level of PrimPol expression, in comparison to the replicative primase, acts as the primary mechanism to restrict its contribution to “normal” replication.

6.6.2. PrimPol is Self-Regulating

The structural features afforded to PrimPol by virtue of being a primase also act as inherent regulatory mechanisms. As mentioned previously, PrimPol displays very low processivity. This distributive nature appears to be due to two key features. Firstly, the AEP catalytic domain has a much smaller ‘footprint’ than most polymerases, potentially explaining why the enzyme binds DNA so poorly (Rechkoblit et al., 2016). Secondly, the ZnF domain acts to negatively regulate PrimPol’s processivity (Keen et al., 2014b).

It has been suggested that the Prim2 subunit of the replicative eukaryotic primase enforces a strict counting mechanism on the enzyme (Kuchta and Stengel, 2010b). Here, the Prim2 and Prim1 subunits form a hinge-like structure. The enzyme binds to ssDNA in a “closed” conformation, with Prim2 facilitating template recognition. The Prim1 subunit then initiates primer synthesis, moving away from Prim2 which binds the 5’ end of the primer (Baranovskiy et al., 2016b; Kuchta and Stengel, 2010b). Thus, an inherent counting mechanism is conferred by the maximum distance Prim1 can elongate the primer away from Prim2. The ZnF domain of PrimPol likely act in a similar manner (Guilliam et al., 2017). In this scenario, the AEP domain and ZnF form a hinge structure, connected by a flexible linker. The enzyme binds to DNA in a closed conformation assisted by the ZnF domain, which binds on the 3’ side relative to the AEP domain on the template strand. The AEP domain can then synthesise and elongate the primer until further extension is restricted by the ZnF domain. PrimPol is, therefore, self-regulating. The supervisory effect of the ZnF domain, which permits priming but limits elongation, coupled with the AEP’s poor affinity for DNA, restricts the ability of PrimPol to partake in significant unregulated DNA synthesis during DNA replication.

6.6.3. Regulation by Single-Strand Binding Proteins

The ability of primases to bind and prime on ssDNA gives them the potential to facilitate un-scheduled priming *in vivo,* wherever ssDNA is available. Despite limiting the synthesis of long DNA tracts, PrimPol’s self-regulatory mechanisms do not restrict where it can prime. Dysregulated priming is potentially highly detrimental to the cell, as these primers

could be extended by other polymerases. To prevent this, PrimPol is also regulated by RPA and mtSSB. Both of these SSBs stimulate the activity of their respective replicative polymerases, δ and γ (Oliveira and Kaguni, 2010; Tsurimoto and Stillman, 1989). In contrast, as shown in Chapter 2, both RPA and mtSSB severely restrict the polymerase activity of PrimPol (Guilliam et al., 2015a). Additionally, these SSBs can also inhibit primase activity, as is the case with Pol α-primase (Collins and Kelly, 1991; Guilliam et al., 2015a). However, as demonstrated in Chapter 5, RPA's effect on PrimPol's primase activity is highly concentration-dependent. In fact, sub-saturating concentrations of RPA dramatically stimulate primer synthesis, but as the concentration increases inhibition occurs (Guilliam et al., 2017). It is likely that both RPA and mtSSB act to prevent un-scheduled priming events by blocking access to the ssDNA template. Thus, PrimPol requires a free ssDNA interface adjacent to the SSB in order to be recruited. This recruitment likely acts to enhance PrimPol's poor affinity for DNA, providing a platform for primer synthesis.

As previously mentioned, RPA binds ssDNA with a defined polarity (Fan and Pavletich, 2012; Iftode and Borowiec, 2000; Kolpashchikov et al., 2001; Laat et al., 1998). Initially, the DBD-A and DBD-B OB folds of RPA70 bind ssDNA in a tandem manner, forming an 8-nt binding complex. The interface in contact with DNA is then extended to 20-30 nts by the binding of DBD-C and DBD-D, which occurs in a defined 5'-3' direction on the template strand (Brosey et al., 2013). This would position the RPA70N domain, which recruits PrimPol, 5' relative to rest of the RPA molecule on the template strand. This suggests that PrimPol binds ahead of RPA *in vivo*, with the ZnF contacting ssDNA adjacent to RPA and the AEP bound downstream on the 5' side. Indeed, the binding of PrimPol in this manner is supported by preliminary investigations described in Chapter 5.

The orientation of PrimPol's interaction with RPA may explain the inhibition observed in primer extension assays in Chapter 2. By preferentially binding on the 5' side of RPA, PrimPol would not be able to access the primer strand at the 3' end of the template. This may also limit PrimPol's access of free 3' termini *in* vivo, therefore preventing pseudo-TLS bypass and mutagenesis. Additionally, replicative polymerases are thought to be able to easily displace RPA as they approach the protein from the 3' side, encountering the weakly bound DBD-D and DBD-C domains, before DBD-B and DBD-A (Iftode et al., 1999). This in turn shifts the equilibrium from the 20-30-nt RPA complex, to the more weakly bound 8-nt mode, thus permitting displacement. In contrast, if PrimPol binds to the 5' side of RPA, it would move away from the protein, preventing displacement in the same way. In support of this, primase assays presented in Chapter 5 reveal that PrimPol

can displace a single RPA molecule bound at the 5' end of a ssDNA template, when approaching from the 3' side (see section 5.4.10.). Presumably, in this scenario, PrimPol cannot be recruited to 5' side of RPA due to the lack of a ssDNA interface there. It is possible that the interaction between PrimPol and RPA also further enhances the regulation of the enzyme's processivity by "holding" the ZnF domain and preventing continued extension by the AEP domain.

6.6.4. What Generates the ssDNA Interface Required for PrimPol Recruitment?

The requirement of a ssDNA interface downstream of RPA for efficient PrimPol recruitment begs the question: how is this free ssDNA interface generated *in vivo*? Although the answer to this question is currently unknown, one obvious solution would be through the action of the replicative helicase. Following stalling of the leading strand replicase, leading and lagging strand replication can become uncoupled. Here, the replisome progresses in the absence of DNA synthesis on the leading strand. Continued unwinding of duplex parental DNA by MCM generates ssDNA on the leading strand, which is bound by RPA. Consequently, an RPA/ssDNA interface for PrimPol binding could be generated directly behind the progressing MCM. Subsequent repriming by PrimPol would prevent extended leading/lagging strand uncoupling, allowing leading strand replication to resume at the progressing replisome. The short RPA-bound ssDNA gap left behind could then be filled by TLS or template switching mechanisms.

In support of this, it has recently been shown that the mitochondrial replicative helicase, Twinkle, can stimulate DNA synthesis by PrimPol, indicating that replicative helicases can potentially facilitate PrimPol activity *in vivo* (Stojkovič et al., 2016). It is interesting to note that many DNA primases interact with replicative helicases, with some even possessing their own helicase domains (Kuchta and Stengel, 2010b).

6.6.5. Regulation by PolDIP2

PolDIP2 was originally identified as a binding partner of the p50 subunit of Pol δ, in addition to PCNA (Liu et al., 2003). More recently, PolDIP2 was shown to interact with Pols η, ζ, λ, and Rev1 (Maga et al., 2013; Tissier et al., 2010). *In vitro*, the protein stimulates the polymerase activity of Pol δ by increasing its affinity for PCNA, as well as enhancing TLS by Pols η and λ (Maga et al., 2013). These observations have led to suggestions that PolDIP2 may play an important role in the switch between Pol δ and TLS polymerases during DNA replication (Maga et al., 2013; Tissier et al., 2010).

As reported in Chapter 3, PolDIP2 also significantly enhances the DNA binding and processivity of PrimPol's AEP domain (Guilliam et al., 2016). Additionally, PolDIP2 appears to be important for PrimPol's function *in vivo*, suggesting it may act as a way to positively regulate the enzyme's activity. Notably, however, this was not sufficient to relieve the negative effect of RPA or mtSSB on PrimPol's polymerase activity. PolDIP2 was found to stimulate PrimPol's bypass of 8-oxo-dG lesions, however unlike previous studies of Pols λ and η, this stimulation was not greater than that seen on undamaged templates (Guilliam et al., 2016; Maga et al., 2013). Here, the increased bypass was a result of an overall increase in the efficiency of the enzyme, rather than an 8-oxo-dG specific effect, although fidelity opposite the lesion was improved. Given that PrimPol appears to primarily function in repriming, PolDIP2 is likely required for the enhancement of primer extension by PrimPol's AEP domain, following synthesis of the initial di-nucleotide. Interestingly, PolDIP2 binds to PrimPol at a region in close proximity to motifs Ia and Ib, identified in the recent crystal structure (Guilliam et al., 2016; Rechkoblit et al., 2016). These motifs harbour the majority of the residues responsible for mediating binding of the AEP domain to the DNA template. PolDIP2, therefore, potentially changes the conformation of this region to enhance PrimPol's affinity for the DNA template, resulting in the increased DNA binding and processivity described in Chapter 3. PrimPol's inherent low processivity only permits incorporation of 1-4 nt before dissociation, however it is unlikely that a primer this short would be stable or sufficient to facilitate restart by the replicative polymerases. As PrimPol lacks an equivalent of Pol α in the Pol α-primase complex, it must elongate its own primers to a viable length. This is likely the reason PrimPol displays both primase and polymerase activities, and suggests it requires a further processivity factor, such as PolDIP2, to function efficiently *in vivo.*

Aside from enhancing PrimPol's primer extension activity, another intriguing role for PolDIP2 is in the hand-off of the nascent primer strand from PrimPol to Pol δ. As mentioned in Chapter 1, recent biochemical and cellular studies have revealed that Pol δ initially extends primers synthesised by Pol α-primase on the leading strand during DNA replication in yeast (Daigaku et al., 2015; Yeeles et al., 2017). Assuming that this mechanism is conserved in higher eukaryotes, it is likely that primers generated by PrimPol during leading-strand repriming would be initially extended by Pol δ, before continued extension by Pol ε. Previous reports that PolDIP2 functions in the polymerase switch during TLS, coupled with the shared interaction of PrimPol, Pol δ, and PCNA, with the protein, raise the possibility that PolDIP2 may promote a switch from PrimPol to Pol δ following primer synthesis and extension (Maga et al., 2013) (Figure 6.4.). Nevertheless, more work is required to explore this potential mechanism.

6.7. Conclusions and Perspectives

Nearly half a century ago, Rupp and Howard-Flanders identified the presence of ssDNA gaps left opposite UV photoproducts following DNA replication in NER deficient *E. coli* (Rupp and Howard-Flanders, 1968). A model was proposed which envisaged reinitiation of replication downstream of the damage on both leading and lagging-strand templates; the first suggestion of repriming. The idea of leading strand reinitiation remained controversial until almost four decades later when origin-independent leading strand reinitiation was observed (Heller and Marians, 2006). Follow-up studies confirmed that the replicative primase, DnaG, could reprime leading strand replication downstream of a lesion, whilst the replisome remained associated with the template (Yeeles and Marians, 2011, 2013). Over recent years, evidence has accumulated to support leading strand repriming as a conserved mechanism for dealing with replisome-stalling impediments in eukaryotes (Elvers et al., 2011; Karras and Jentsch, 2010; Lopes et al., 2006). It now seems clear that PrimPol is employed in this mechanism in eukaryotic organisms for the bypass of a wide range of leading strand obstacles (Figure 6.4.) (Guilliam et al., 2017; Keen et al., 2014b; Kobayashi et al., 2016; Mourón et al., 2013; Pilzecker et al., 2016; Schiavone et al., 2016; Vallerga et al., 2015).

Since the initial reports describing PrimPol only four years ago, studies from a number of laboratories have greatly increased our understanding of the role, recruitment, and regulation of the enzyme during DNA replication (Bianchi et al., 2013; García-Gómez et al., 2013; Wan et al., 2013). However, we are only just beginning to appreciate the novel roles that PrimPol plays in DNA replication and damage tolerance. The exact interplay between leading strand repriming by PrimPol and other DDT mechanisms, such as TLS, is still not yet clear. It is possible that DDT mechanisms work to complement repriming by filling in the resulting ssDNA gaps. Alternatively, repriming could occur when TLS at the replication fork fails, in order to prevent extended leading/lagging strand replication uncoupling. The redundancy between Pol α-primase and PrimPol *in vivo* is also an interesting avenue of future study. The reason for the apparent requirement of PrimPol for leading strand repriming in higher eukaryotes, but not other organisms, is not yet completely clear. Although leading strand repriming is emerging as the primary role for PrimPol during DNA replication, the catalytic versatility of the enzyme may lend itself to disparate roles in other processes, such as transcription (Martínez-Jiménez et al., 2015).

We now know that RPA serves to recruit PrimPol to stalled replication forks in the nucleus (Guilliam et al., 2017). However, mtSSB has not yet been shown to play an analogous role in the mitochondria, although an interaction between the proteins *in vivo* is

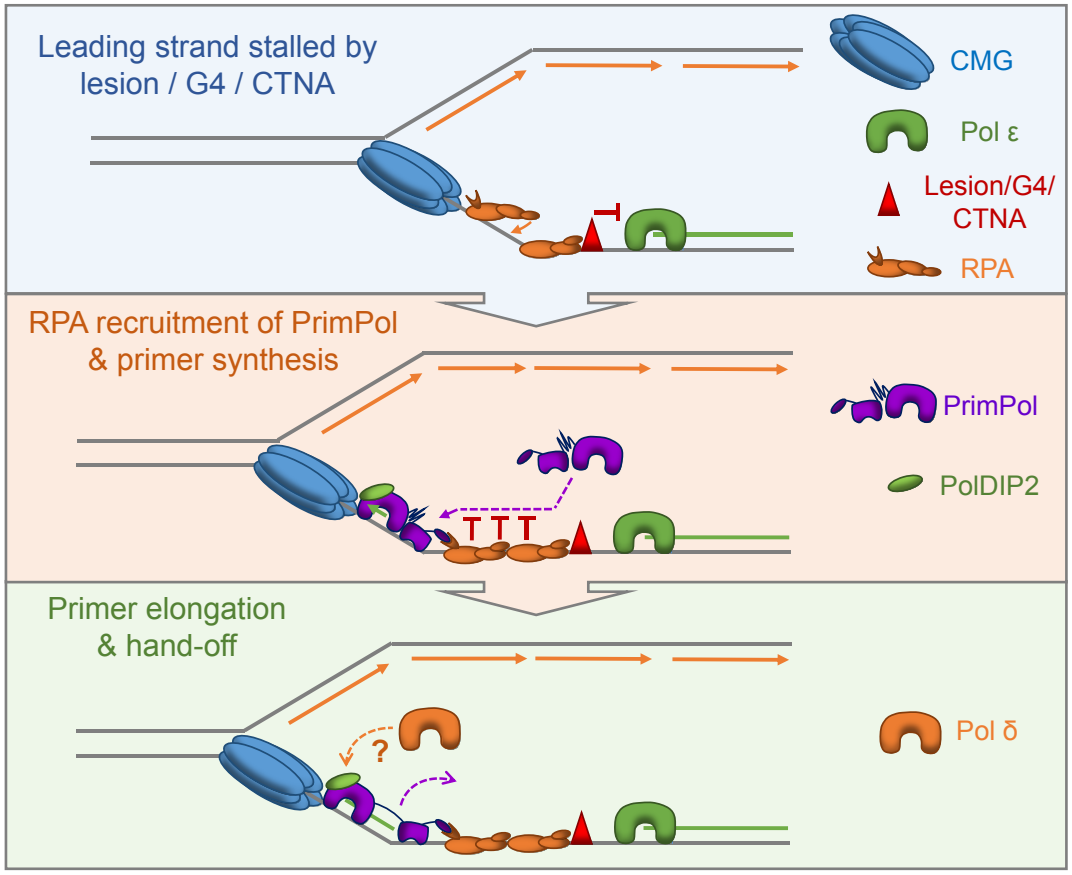


Figure 6.4. The role, recruitment, and regulation of PrimPol in DNA replication.
Top panel: Pol ε is stalled on the leading strand by a lesion, secondary structure, or CTNA. Lagging strand replication continues, subsequently generating ssDNA on the leading strand. This ssDNA is bound by RPA as the CMG complex progresses. **Middle panel:** The generation of an RPA / ssDNA interface provides a platform for PrimPol recruitment. PrimPol requires a free ssDNA region adjacent to RPA and thus is recruited to the exposed ssDNA behind the CMG complex. This recruitment is facilitated by the interaction between PrimPol's RBMs and RPA70N. Following recruitment, PrimPol reprimes the leading strand. **Bottom panel:** PrimPol elongates its primer, assisted by PolDIP2, before further extension is restricted by its ZnF and RPA interaction. The primer is then handed-off to the replicative polymerase, possibly Pol δ, mediated by each protein's interaction with PolDIP2.

documented (Guilliam et al., 2015a). Additionally, it is possible that post-translational modifications, as well as interactions with the replicative helicases, play a role in this process (Stojkovič et al., 2016). The necessity of appropriate recruitment and regulation of PrimPol in the cell is highlighted by the mutations of PrimPol's RBMs identified in cancer patient cell lines, which likely adversely affect recruitment of the enzyme (Guilliam et al., 2017). The regulation of PrimPol appears to walk a fine line between preventing and causing genetic instability, as PrimPol is inherently error-prone and also been found to be over-expressed in some cancers, such as glioma (Guilliam et al., 2015a; Yan et al., 2011). Although we have highlighted some of the known mechanisms regulating PrimPol's activity here, it is likely that additional layers of regulation remain to be discovered.

The hypersensitivity to DNA damaging agents observed in the absence of PrimPol legitimises the enzyme as a potential target for inhibition in combination with other DDT factors and DNA damaging chemotherapeutics (Bianchi et al., 2013; Kobayashi et al., 2016; Mourón et al., 2013). Similarly, PrimPol homologues in trypanosomes have been identified as essential for survival and thus PrimPol-like proteins in other species may also be potential targets for anti-parasitic drugs (Rudd et al., 2013). Further studies will be important in determining the viability and usefulness of manipulating PrimPol in treating cancer and other diseases.

References

Achar, Y.J., Balogh, D., Haracska, L., 2011. Coordinated protein and DNA remodeling by human HLTF on stalled replication fork. Proc. Natl. Acad. Sci. U. S. A. 108, 14073–14078. doi:10.1073/pnas.1101951108

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221. doi:10.1107/S0907444909052925

Aguilera, A., García-Muse, T., 2012. R loops: from transcription byproducts to threats to genome stability. Mol. Cell 46, 115–124. doi:10.1016/j.molcel.2012.04.009

Aguilera, A., Gómez-González, B., 2008. Genome instability: a mechanistic view of its causes and consequences. Nat. Rev. Genet. 9, 204–217. doi:10.1038/nrg2268

Ahn, S.J., Costa, J., Emanuel, J.R., 1996. PicoGreen Quantitation of DNA: Effective Evaluation of Samples Pre-or Psost-PCR. Nucleic Acids Res. 24, 2623–2625. doi:10.1093/nar/24.13.2623

Aksenova, A., Volkov, K., Maceluch, J., Pursell, Z.F., Rogozin, I.B., Kunkel, T.A., Pavlov, Y.I., Johansson, E., 2010. Mismatch repair-independent increase in spontaneous mutagenesis in yeast lacking non-essential subunits of DNA polymerase ε. PLoS Genet. 6, e1001209. doi:10.1371/journal.pgen.1001209

Alexeyev, M., Shokolenko, I., Wilson, G., LeDoux, S., 2013. The Maintenance of Mitochondrial DNA Integrity—Critical Analysis and Update. Cold Spring Harb. Perspect. Biol. 5, a012641. doi:10.1101/cshperspect.a012641

Almeida, K.H., Sobol, R.W., 2007. A unified view of base excision repair: lesion-dependent protein complexes regulated by post-translational modification. DNA Repair 6, 695–711. doi:10.1016/j.dnarep.2007.01.009

Alzu, A., Bermejo, R., Begnis, M., Lucca, C., Piccini, D., Carotenuto, W., Saponaro, M., Brambati, A., Cocito, A., Foiani, M., Liberi, G., 2012. Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. Cell 151, 835–846. doi:10.1016/j.cell.2012.09.041

Ananda, G., Hile, S.E., Breski, A., Wang, Y., Kelkar, Y., Makova, K.D., Eckert, K.A., 2014. Microsatellite Interruptions Stabilize Primate Genomes and Exist as Population-Specific Single Nucleotide Polymorphisms within Individual Human Genomes. PLOS Genet. 10, e1004498. doi:10.1371/journal.pgen.1004498

Arai, K., Kornberg, A., 1979. A general priming system employing only dnaB protein and primase for DNA replication. Proc. Natl. Acad. Sci. 76, 4308–4312.

Aravind, L., Koonin, E.V., 2001. Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. Genome Res. 11, 1365–1374. doi:10.1101/gr.181001

Aravind, L., Leipe, D.D., Koonin, E.V., 1998a. Toprim—a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. Nucleic Acids Res. 26, 4205–4213. doi:10.1093/nar/26.18.4205

Aravind, L., Leipe, D.D., Koonin, E.V., 1998b. Toprim--a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. Nucleic Acids Res. 26, 4205–4213.

Arunkumar, A.I., Klimovich, V., Jiang, X., Ott, R.D., Mizoue, L., Fanning, E., Chazin, W.J., 2005. Insights into hRPA32 C-terminal domain–mediated assembly of the simian virus 40 replisome. Nat. Struct. Mol. Biol. 12, 332–339. doi:10.1038/nsmb916

Ash, D.E., Schramm, V.L., 1982. Determination of free and bound manganese(II) in hepatocytes from fed and fasted rats. J. Biol. Chem. 257, 9261–9264.

Augustin, M.A., Huber, R., Kaiser, J.T., 2001. Crystal structure of a DNA-dependent RNA polymerase (DNA primase). Nat. Struct. Mol. Biol. 8, 57–61. doi:10.1038/83060

Avery, O.T., Macleod, C.M., McCarty, M., 1944. STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES : INDUCTION OF TRANSFORMATION BY A DESOXYRIBONUCLEIC ACID FRACTION ISOLATED FROM PNEUMOCOCCUS TYPE III. J. Exp. Med. 79, 137–158.

Bailey, L.J., Bianchi, J., Hégarat, N., Hochegger, H., Doherty, A.J., 2016. PrimPol-deficient cells exhibit a pronounced G2 checkpoint response following UV damage. Cell Cycle Georget. Tex 15, 908–918. doi:10.1080/15384101.2015.1128597

Bansbach, C.E., Bétous, R., Lovejoy, C.A., Glick, G.G., Cortez, D., 2009. The annealing helicase SMARCAL1 maintains genome integrity at stalled replication forks. Genes Dev. 23, 2405–2414. doi:10.1101/gad.1839909

Baptiste, B.A., Eckert, K.A., 2012. DNA polymerase kappa microsatellite synthesis: Two distinct mechanisms of slippage-mediated errors. Environ. Mol. Mutagen. 53, 787–796. doi:10.1002/em.21721

Baranovskiy, A.G., Babayeva, N.D., Zhang, Y., Gu, J., Suwa, Y., Pavlov, Y.I., Tahirov, T.H., 2016a. Mechanism of Concerted RNA-DNA Primer Synthesis by the Human Primosome. J. Biol. Chem. 291, 10006–10020. doi:10.1074/jbc.M116.717405

Baranovskiy, A.G., Zhang, Y., Suwa, Y., Gu, J., Babayeva, N.D., Pavlov, Y.I., Tahirov, T.H., 2016b. Insight into the Human DNA Primase Interaction with Template-Primer. J. Biol. Chem. 291, 4793–4802. doi:10.1074/jbc.M115.704064

Barlow, J.H., Faryabi, R.B., Callén, E., Wong, N., Malhowski, A., Chen, H.T., Gutierrez-Cruz, G., Sun, H.-W., McKinnon, P., Wright, G., Casellas, R., Robbiani, D.F., Staudt, L., Fernandez-Capetillo, O., Nussenzweig, A., 2013. Identification of Early Replicating Fragile Sites that Contribute to Genome Instability. Cell 152, 620–632. doi:10.1016/j.cell.2013.01.006

Barnes, D.E., Johnston, L.H., Kodama, K., Tomkinson, A.E., Lasko, D.D., Lindahl, T., 1990. Human DNA ligase I cDNA: cloning and functional expression in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. 87, 6679–6683.

Bartlett, E.J., Brissett, N.C., Doherty, A.J., 2013. Ribonucleolytic resection is required for repair of strand displaced nonhomologous end-joining intermediates. Proc. Natl. Acad. Sci. 110, E1984–E1991. doi:10.1073/pnas.1302616110

Baxter, J., 2015. “Breaking up is hard to do”: the formation and resolution of sister chromatid intertwines. J. Mol. Biol. 427, 590–607. doi:10.1016/j.jmb.2014.08.022

Beard, W.A., Wilson, S.H., 2001. DNA Lesion Bypass Polymerases Open Up. Structure 9, 759–764. doi:10.1016/S0969-2126(01)00646-3

Beck, H., Nähse-Kumpf, V., Larsen, M.S.Y., O’Hanlon, K.A., Patzke, S., Holmberg, C., Mejlvang, J., Groth, A., Nielsen, O., Syljuåsen, R.G., Sørensen, C.S., 2012. Cyclin-dependent kinase suppression by WEE1 kinase protects the genome through control of replication initiation and nucleotide consumption. Mol. Cell. Biol. 32, 4226–4236. doi:10.1128/MCB.00412-12

Beck, K., Lipps, G., 2007. Properties of an unusual DNA primase from an archaeal plasmid. Nucleic Acids Res. 35, 5635–5645. doi:10.1093/nar/gkm625

Beck, M., Schmidt, A., Malmstroem, J., Claassen, M., Ori, A., Szymborska, A., Herzog, F., Rinner, O., Ellenberg, J., Aebersold, R., 2011. The quantitative proteome of a human cell line. Mol. Syst. Biol. 7, 549. doi:10.1038/msb.2011.82

Bell, S.P., Kaguni, J.M., 2013. Helicase Loading at Chromosomal Origins of Replication. Cold Spring Harb. Perspect. Biol. 5. doi:10.1101/cshperspect.a010124

Berdis, A.J., 2008. DNA Polymerases as Therapeutic Targets. Biochemistry (Mosc.) 47, 8253–8260. doi:10.1021/bi801179f

Bermejo, R., Capra, T., Jossen, R., Colosio, A., Frattini, C., Carotenuto, W., Cocito, A., Doksani, Y., Klein, H., Gómez-González, B., Aguilera, A., Katou, Y., Shirahige, K., Foiani, M., 2011. The replication checkpoint protects fork stability by releasing

transcribed genes from nuclear pores. Cell 146, 233–246. doi:10.1016/j.cell.2011.06.033

Bermejo, R., Lai, M.S., Foiani, M., 2012. Preventing replication stress to maintain genome stability: resolving conflicts between replication and transcription. Mol. Cell 45, 710–718. doi:10.1016/j.molcel.2012.03.001

Bernstein, J.A., Richardson, C.C., 1988. Purification of the 56-kDa component of the bacteriophage T7 primase/helicase and characterization of its nucleoside 5'-triphosphatase activity. J. Biol. Chem. 263, 14891–14899.

Bharti, S.K., Sommers, J.A., Zhou, J., Kaplan, D.L., Spelbrink, J.N., Mergny, J.-L., Brosh, R.M., 2014. DNA Sequences Proximal to Human Mitochondrial DNA Deletion Breakpoints Prevalent in Human Disease Form G-quadruplexes, a Class of DNA Structures Inefficiently Unwound by the Mitochondrial Replicative Twinkle Helicase. J. Biol. Chem. 289, 29975–29993. doi:10.1074/jbc.M114.567073

Bianchi, J., 2013. Investigating the role of a novel primase-polymerase, PrimPol, in DNA damage tolerance in vertebrate cells (PhD Thesis). University of Sussex, Brighton.

Bianchi, J., Rudd, S.G., Jozwiakowski, S.K., Bailey, L.J., Soura, V., Taylor, E., Stevanovic, I., Green, A.J., Stracker, T.H., Lindsay, H.D., Doherty, A.J., 2013. PrimPol bypasses UV photoproducts during eukaryotic chromosomal DNA replication. Mol. Cell 52, 566–573. doi:10.1016/j.molcel.2013.10.035

Bienko, M., Green, C.M., Crosetto, N., Rudolf, F., Zapart, G., Coull, B., Kannouche, P., Wider, G., Peter, M., Lehmann, A.R., Hofmann, K., Dikic, I., 2005. Ubiquitin-binding domains in Y-family polymerases regulate translesion synthesis. Science 310, 1821–1824. doi:10.1126/science.1120615

Biertümpfel, C., Zhao, Y., Kondo, Y., Ramón-maiques, S., Gregory, M., Lee, J.Y., Masutani, C., Lehmann, A.R., Hanaoka, F., Yang, W., 2011. Structure and mechanism of human Dna polymerase η. Nature 476. doi:10.1038/nature10338

Biertümpfel, C., Zhao, Y., Kondo, Y., Ramón-Maiques, S., Gregory, M., Lee, J.Y., Masutani, C., Lehmann, A.R., Hanaoka, F., Yang, W., 2010. Structure and Mechanism of Human DNA Polymerase η. Nature 465, 1044–1048. doi:10.1038/nature09196

Biswas, E.E., Joseph, P.E., Biswas, S.B., 1987. Yeast DNA primase is encoded by a 59-kilodalton polypeptide: purification and immunochemical characterization. Biochemistry (Mosc.) 26, 5377–5382. doi:10.1021/bi00391a024

Biswas, N., Weller, S.K., 1999. A Mutation in the C-terminal Putative Zn2+ Finger Motif of UL52 Severely Affects the Biochemical Activities of the HSV-1 Helicase-Primase Subcomplex. J. Biol. Chem. 274, 8068–8076. doi:10.1074/jbc.274.12.8068

Biswas, T., Resto-Roldán, E., Sawyer, S.K., Artsimovitch, I., Tsodikov, O.V., 2012. A novel non-radioactive primase–pyrophosphatase activity assay and its application to the discovery of inhibitors of Mycobacterium tuberculosis primase DnaG. Nucleic Acids Res. gks1292. doi:10.1093/nar/gks1292

Black, S.J., Kashkina, E., Kent, T., Pomerantz, R.T., 2016. DNA Polymerase θ: A Unique Multifunctional End-Joining Machine. Genes 7, 67. doi:10.3390/genes7090067

Blastyák, A., Pintér, L., Unk, I., Prakash, L., Prakash, S., Haracska, L., 2007. Yeast Rad5 protein required for postreplication repair has a DNA helicase activity specific for replication fork regression. Mol. Cell 28, 167–175. doi:10.1016/j.molcel.2007.07.030

Blow, J.J., Ge, X.Q., Jackson, D.A., 2011. How dormant origins promote complete genome replication. Trends Biochem. Sci. 36, 405–414. doi:10.1016/j.tibs.2011.05.002

Bochkareva, E., Kaustov, L., Ayed, A., Yi, G.-S., Lu, Y., Pineda-Lucena, A., Liao, J.C.C., Okorokov, A.L., Milner, J., Arrowsmith, C.H., Bochkarev, A., 2005. Single-stranded DNA mimicry in the p53 transactivation domain interaction with

replication protein A. Proc. Natl. Acad. Sci. U. S. A. 102, 15412–15417. doi:10.1073/pnas.0504614102

Bochman, M.L., Paeschke, K., Zakian, V.A., 2012. DNA secondary structures: stability and function of G-quadruplex structures. Nat. Rev. Genet. 13, 770–780. doi:10.1038/nrg3296

Bocquier, A.A., Liu, L., Cann, I.K.O., Komori, K., Kohda, D., Ishino, Y., 2001. Archaeal primase: bridging the gap between RNA and DNA polymerases. Curr. Biol. 11, 452–456. doi:10.1016/S0960-9822(01)00119-1

Boehm, E.M., Spies, M., Washington, M.T., 2016. PCNA tool belts and polymerase bridges form during translesion synthesis. Nucleic Acids Res. 44, 8250–8260. doi:10.1093/nar/gkw563

Bohgaki, T., Bohgaki, M., Hakem, R., 2010. DNA double-strand break signaling and human disorders. Genome Integr. 1, 15. doi:10.1186/2041-9414-1-15

Bollum, F.J., Potter, V.R., 1957. Thymidine incorporation into deoxyribonucleic acid of rat liver homogenates. J. Am. Chem. Soc. 79, 3603–3604.

Bouché, J.P., Rowen, L., Kornberg, A., 1978. The RNA primer synthesized by primase to initiate phage G4 DNA replication. J. Biol. Chem. 253, 765–769.

Bouché, J.P., Zechel, K., Kornberg, A., 1975. dnaG gene product, a rifampicin-resistant RNA polymerase, initiates the conversion of a single-stranded coliphage DNA to its duplex replicative form. J. Biol. Chem. 250, 5995–6001.

Boudsocq, F., Kokoska, R.J., Plosky, B.S., Vaisman, A., Ling, H., Kunkel, T.A., Yang, W., Woodgate, R., 2004. Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication. J. Biol. Chem. 279, 32932–32940. doi:10.1074/jbc.M405249200

Braithwaite, D.K., Ito, J., 1993. Compilation, alignment, and phylogenetic relationships of DNA polymerases. Nucleic Acids Res. 21, 787–802.

Branzei, D., Seki, M., Enomoto, T., 2004. Rad18/Rad5/Mms2-mediated polyubiquitination of PCNA is implicated in replication completion during replication stress. Genes Cells Devoted Mol. Cell. Mech. 9, 1031–1042. doi:10.1111/j.1365-2443.2004.00787.x

Branzei, D., Szakal, B., 2016. DNA damage tolerance by recombination: Molecular pathways and DNA structures. DNA Repair 44, 68–75. doi:10.1016/j.dnarep.2016.05.008

Braun, K.A., Lao, Y., He, Z., Ingles, C.J., Wold, M.S., 1997. Role of Protein−Protein Interactions in the Function of Replication Protein A (RPA): RPA Modulates the Activity of DNA Polymerase α by Multiple Mechanisms. Biochemistry (Mosc.) 36, 8443–8454. doi:10.1021/bi970473r

Brautigam, C.A., Steitz, T.A., 1998. Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. Curr. Opin. Struct. Biol. 8, 54–63.

Brill, S.J., DiNardo, S., Voelkel-Meiman, K., Sternglanz, R., 1987. Need for DNA topoisomerase activity as a swivel for DNA replication for transcription of ribosomal RNA. Nature 326, 414–416. doi:10.1038/326414a0

Brissett, N.C., Martin, M.J., Bartlett, E.J., Bianchi, J., Blanco, L., Doherty, A.J., 2013. Molecular Basis for DNA Double-Strand Break Annealing and Primer Extension by an NHEJ DNA Polymerase. Cell Rep. 5, 1108–1120. doi:10.1016/j.celrep.2013.10.016

Brissett, N.C., Martin, M.J., Pitcher, R.S., Bianchi, J., Juarez, R., Green, A.J., Fox, G.C., Blanco, L., Doherty, A.J., 2011. Structure of a Preternary Complex Involving a Prokaryotic NHEJ DNA Polymerase. Mol. Cell 41, 221–231. doi:10.1016/j.molcel.2010.12.026

Brissett, N.C., Pitcher, R.S., Juarez, R., Picher, A.J., Green, A.J., Dafforn, T.R., Fox, G.C., Blanco, L., Doherty, A.J., 2007. Structure of a NHEJ Polymerase-Mediated DNA Synaptic Complex. Science 318, 456–459. doi:10.1126/science.1145112

Brosey, C.A., Yan, C., Tsutakawa, S.E., Heller, W.T., Rambo, R.P., Tainer, J.A., Ivanov, I., Chazin, W.J., 2013. A new structural framework for integrating replication protein A into DNA processing machinery. Nucleic Acids Res. 41, 2313–2327. doi:10.1093/nar/gks1332

Brun, R., Blum, J., Chappuis, F., Burri, C., 2010. Human African trypanosomiasis. The Lancet 375, 148–159. doi:10.1016/S0140-6736(09)60829-1

Burgers, P.M., 2009. Polymerase dynamics at the eukaryotic DNA replication fork. J. Biol. Chem. 284, 4041–4045.

Burgers, P.M., 1991. Saccharomyces cerevisiae replication factor C. II. Formation and activity of complexes with the proliferating cell nuclear antigen and with DNA polymerases delta and epsilon. J. Biol. Chem. 266, 22698–22706.

Burgers, P.M.J., Gordenin, D., Kunkel, T.A., 2016. Who Is Leading the Replication Fork, Pol ε or Pol δ? Mol. Cell 61, 492–493. doi:10.1016/j.molcel.2016.01.017

Burgers, P.M.J., Koonin, E.V., Bruford, E., Blanco, L., Burtis, K.C., Christman, M.F., Copeland, W.C., Friedberg, E.C., Hanaoka, F., Hinkle, D.C., Lawrence, C.W., Nakanishi, M., Ohmori, H., Prakash, L., Prakash, S., Reynaud, C.-A., Sugino, A., Todo, T., Wang, Z., Weill, J.-C., Woodgate, R., 2001. Eukaryotic DNA Polymerases: Proposal for a Revised Nomenclature. J. Biol. Chem. 276, 43487–43490. doi:10.1074/jbc.R100056200

Byrnes, J.J., Downey, K.M., Black, V.L., So, A.G., 1976. A new mammalian DNA polymerase with 3'to 5'exonuclease activity: DNA polymerase delta. Biochemistry (Mosc.) 15, 2817–2823.

Byun, T.S., Pacek, M., Yee, M., Walter, J.C., Cimprich, K.A., 2005. Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. Genes Dev. 19, 1040–1052. doi:10.1101/gad.1301205

Cadet, J., Douki, T., Gasparutto, D., Ravanat, J.-L., 2003. Oxidative damage to DNA: formation, measurement and biochemical features. Mutat. Res. Mol. Mech. Mutagen., Oxidative DNA Damage and its Repair Base Excision Repair 531, 5–23. doi:10.1016/j.mrfmmm.2003.09.001

Ceccaldi, R., Rondinelli, B., D'Andrea, A.D., 2016. Repair Pathway Choices and Consequences at the Double-Strand Break. Trends Cell Biol., Special Issue: Quality Control 26, 52–64. doi:10.1016/j.tcb.2015.07.009

Cerritelli, S.M., Frolova, E.G., Feng, C., Grinberg, A., Love, P.E., Crouch, R.J., 2003. Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. Mol. Cell 11, 807–815.

Chabes, A., Georgieva, B., Domkin, V., Zhao, X., Rothstein, R., Thelander, L., 2003. Survival of DNA damage in yeast directly depends on increased dNTP levels allowed by relaxed feedback inhibition of ribonucleotide reductase. Cell 112, 391–401.

Chan, S.H., Yu, A.M., McVey, M., 2010. Dual Roles for DNA Polymerase Theta in Alternative End-Joining Repair of Double-Strand Breaks in Drosophila. PLoS Genet. 6. doi:10.1371/journal.pgen.1001005

Chang, D.D., Clayton, D.A., 1985. Priming of human mitochondrial DNA replication occurs at the light-strand promoter. Proc. Natl. Acad. Sci. U. S. A. 82, 351–355.

Chang, D.D., Hauswirth, W.W., Clayton, D.A., 1985. Replication priming and transcription initiate from precisely the same site in mouse mitochondrial DNA. EMBO J. 4, 1559–1567.

Chargaff, E., Lipshitz, R., Green, C., 1952. Composition of the desoxypentose nucleic acids of four genera of sea-urchin. J. Biol. Chem. 195, 155–160.

Chen, H., Lisby, M., Symington, L.S., 2013. RPA Coordinates DNA End Resection and Prevents Formation of DNA Hairpins. Mol. Cell 50, 589–600. doi:10.1016/j.molcel.2013.04.032

Chen, Y., Carrington-Lawrence, S.D., Bai, P., Weller, S.K., 2005. Mutations in the Putative Zinc-Binding Motif of UL52 Demonstrate a Complex Interdependence between the UL5 and UL52 Subunits of the Human Herpes Simplex Virus Type

1 Helicase/Primase Complex. J. Virol. 79, 9088–9096. doi:10.1128/JVI.79.14.9088-9096.2005

Cheng, X., Kanki, T., Fukuoh, A., Ohgaki, K., Takeya, R., Aoki, Y., Hamasaki, N., Kang, D., 2005. PDIP38 Associates with Proteins Constituting the Mitochondrial DNA Nucleoid. J. Biochem. (Tokyo) 138, 673–678. doi:10.1093/jb/mvi169

Cheung, I., Schertzer, M., Rose, A., Lansdorp, P.M., 2002. Disruption of dog-1 in Caenorhabditis elegans triggers deletions upstream of guanine-rich DNA. Nat. Genet. 31, 405–409. doi:10.1038/ng928

Choi, J.-H., Lindsey-Boltz, L.A., Kemp, M., Mason, A.C., Wold, M.S., Sancar, A., 2010. Reconstitution of RPA-covered single-stranded DNA-activated ATR-Chk1 signaling. Proc. Natl. Acad. Sci. U. S. A. 107, 13660–13665. doi:10.1073/pnas.1007856107

Church, D.N., Briggs, S.E.W., Palles, C., Domingo, E., Kearsey, S.J., Grimes, J.M., Gorman, M., Martin, L., Howarth, K.M., Hodgson, S.V., NSECG Collaborators, Kaur, K., Taylor, J., Tomlinson, I.P.M., 2013. DNA polymerase ε and δ exonuclease domain mutations in endometrial cancer. Hum. Mol. Genet. 22, 2820–2828. doi:10.1093/hmg/ddt131

Ciccia, A., Elledge, S.J., 2010. The DNA Damage Response: Making It Safe to Play with Knives. Mol. Cell 40, 179–204. doi:10.1016/j.molcel.2010.09.019

Clausen, A.R., Lujan, S.A., Burkholder, A.B., Orebaugh, C.D., Williams, J.S., Clausen, M.F., Malc, E.P., Mieczkowski, P.A., Fargo, D.C., Smith, D.J., Kunkel, T.A., 2015. Tracking replication enzymology in vivo by genome-wide mapping of ribonucleotide incorporation. Nat. Struct. Mol. Biol. 22, 185–191. doi:10.1038/nsmb.2957

Collaborative Computational Project, Number 4, 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. 50, 760–763. doi:10.1107/S0907444994003112

Collins, K.L., Kelly, T.J., 1991. Effects of T antigen and replication protein A on the initiation of DNA synthesis by DNA polymerase alpha-primase. Mol. Cell. Biol. 11, 2108–2115. doi:10.1128/MCB.11.4.2108

Coloma, J., Johnson, R.E., Prakash, L., Prakash, S., Aggarwal, A.K., 2016. Human DNA polymerase α in binary complex with a DNA:DNA template-primer. Sci. Rep. 6, 23784. doi:10.1038/srep23784

Conaway, R.C., Lehman, I.R., 1982a. A DNA primase activity associated with DNA polymerase alpha from Drosophila melanogaster embryos. Proc. Natl. Acad. Sci. 79, 2523–2527.

Conaway, R.C., Lehman, I.R., 1982b. Synthesis by the DNA primase of Drosophila melanogaster of a primer with a unique chain length. Proc. Natl. Acad. Sci. 79, 4585–4588.

Copeland, W.C., Tan, X., 1995. Active Site Mapping of the Catalytic Mouse Primase Subunit by Alanine Scanning Mutagenesis. J. Biol. Chem. 270, 3905–3913.

Cotterill, S., Chui, G., Lehman, I.R., 1987. DNA polymerase-primase from embryos of Drosophila melanogaster. DNA primase subunits. J. Biol. Chem. 262, 16105–16108.

Courcelle, J., Donaldson, J.R., Chow, K.-H., Courcelle, C.T., 2003. DNA damage-induced replication fork regression and processing in Escherichia coli. Science 299, 1064–1067. doi:10.1126/science.1081328

Crick, F., 1970. Central dogma of molecular biology. Nature 227, 561–563.

Crick, F.H., Barnett, L., Brenner, S., Watts-Tobin, R.J., 1961. General nature of the genetic code for proteins. Nature 192, 1227–1232.

Crute, J.J., Lehman, I.R., 1991. Herpes simplex virus-1 helicase-primase. Physical and catalytic properties. J. Biol. Chem. 266, 4484–4488.

Crute, J.J., Tsurumi, T., Zhu, L.A., Weller, S.K., Olivo, P.D., Challberg, M.D., Mocarski, E.S., Lehman, I.R., 1989. Herpes simplex virus 1 helicase-primase: a complex of three herpes-encoded gene products. Proc. Natl. Acad. Sci. 86, 2186–2189.

Daigaku, Y., Davies, A.A., Ulrich, H.D., 2010. Ubiquitin-dependent DNA damage bypass is separable from genome replication. Nature 465, 951–955. doi:10.1038/nature09097

Daigaku, Y., Keszthelyi, A., Müller, C.A., Miyabe, I., Brooks, T., Retkute, R., Hubank, M., Nieduszynski, C.A., Carr, A.M., 2015. A global profile of replicative polymerase usage. Nat. Struct. Mol. Biol. 22, 192–198. doi:10.1038/nsmb.2962

Dairaghi, D.J., Shadel, G.S., Clayton, D.A., 1995. Addition of a 29 residue carboxyl-terminal tail converts a simple HMG box-containing protein into a transcriptional activator. J. Mol. Biol. 249, 11–28. doi:10.1006/jmbi.1995.9889

Dalgaard, J.Z., 2012. Causes and consequences of ribonucleotide incorporation into nuclear DNA. Trends Genet. TIG 28, 592–597. doi:10.1016/j.tig.2012.07.008

Davies, A.A., Huttner, D., Daigaku, Y., Chen, S., Ulrich, H.D., 2008. Activation of ubiquitin-dependent DNA damage bypass is mediated by replication protein a. Mol. Cell 29, 625–636. doi:10.1016/j.molcel.2007.12.016

De Clercq, E., Field, H.J., 2006. Antiviral prodrugs – the development of successful prodrug strategies for antiviral chemotherapy. Br. J. Pharmacol. 147, 1–11. doi:10.1038/sj.bjp.0706446

De Lucia, P., Cairns, J., 1969. Isolation of an E. coli strain with a mutation affecting DNA polymerase. Nature 224, 1164–1166.

De Silva, F.S., Paran, N., Moss, B., 2009. Products and substrate/template usage of vaccinia virus DNA primase. Virology 383, 136–141. doi:10.1016/j.virol.2008.10.008

Debatisse, M., Le Tallec, B., Letessier, A., Dutrillaux, B., Brison, O., 2012. Common fragile sites: mechanisms of instability revisited. Trends Genet. TIG 28, 22–32. doi:10.1016/j.tig.2011.10.003

Deegan, T.D., Yeeles, J.T., Diffley, J.F., 2016. Phosphopeptide binding by Sld3 links Dbf4-dependent kinase to MCM replicative helicase activation. EMBO J. 35, 961–973.

Della, M., Palmbos, P.L., Tseng, H.-M., Tonkin, L.M., Daley, J.M., Topper, L.M., Pitcher, R.S., Tomkinson, A.E., Wilson, T.E., Doherty, A.J., 2004. Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. Science 306, 683–685. doi:10.1126/science.1099824

Dennis, G., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., Lempicki, R.A., 2003. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 4, R60. doi:10.1186/gb-2003-4-9-r60

Despras, E., Delrieu, N., Garandeau, C., Ahmed-Seghir, S., Kannouche, P.L., 2012. Regulation of the specialized DNA polymerase eta: revisiting the biological relevance of its PCNA- and ubiquitin-binding motifs. Environ. Mol. Mutagen. 53, 752–765. doi:10.1002/em.21741

Devbhandari, S., Jiang, J., Kumar, C., Whitehouse, I., Remus, D., 2017. Chromatin Constrains the Initiation and Elongation of DNA Replication. Mol. Cell 65, 131–141. doi:10.1016/j.molcel.2016.10.035

Dewar, J.M., Budzowska, M., Walter, J.C., 2015. The mechanism of DNA replication termination in vertebrates. Nature 525, 345–350. doi:10.1038/nature14887

Dianova, I.I., Sleeth, K.M., Allinson, S.L., Parsons, J.L., Breslin, C., Caldecott, K.W., Dianov, G.L., 2004. XRCC1-DNA polymerase beta interaction is required for efficient base excision repair. Nucleic Acids Res. 32, 2550–2555. doi:10.1093/nar/gkh567

Doherty, A.J., Jackson, S.P., Weller, G.R., 2001. Identification of bacterial homologues of the Ku DNA repair proteins. FEBS Lett. 500, 186–188. doi:10.1016/S0014-5793(01)02589-3

Dornreiter, I., Erdile, L.F., Gilbert, I.U., von Winkler, D., Kelly, T.J., Fanning, E., 1992. Interaction of DNA polymerase alpha-primase with cellular replication protein A and SV40 T antigen. EMBO J. 11, 769–776.

Doublié, S., Zahn, K.E., 2014. Structural insights into eukaryotic DNA replication. Front. Microbiol. 5. doi:10.3389/fmicb.2014.00444

Duncan, B.K., Miller, J.H., 1980. Mutagenic deamination of cytosine residues in DNA. Nature 287, 560–561.

Durando, M., Tateishi, S., Vaziri, C., 2013. A non-catalytic role of DNA polymerase η in recruiting Rad18 and promoting PCNA monoubiquitination at stalled replication forks. Nucleic Acids Res. 41, 3079–3093. doi:10.1093/nar/gkt016

Eckert, K.A., Mowery, A., Hile, S.E., 2002. Misalignment-Mediated DNA Polymerase β Mutations: Comparison of Microsatellite and Frame-Shift Error Rates Using a Forward Mutation Assay. Biochemistry (Mosc.) 41, 10490–10498. doi:10.1021/bi025918c

Edmunds, C.E., Simpson, L.J., Sale, J.E., 2008. PCNA Ubiquitination and REV1 Define Temporally Distinct Mechanisms for Controlling Translesion Synthesis in the Avian Cell Line DT40. Mol. Cell 30, 519–529. doi:10.1016/j.molcel.2008.03.024

Edwards, M.C., Tutter, A.V., Cvetic, C., Gilbert, C.H., Prokhorova, T.A., Walter, J.C., 2002. MCM2-7 complexes bind chromatin in a distributed pattern surrounding the origin recognition complex in Xenopus egg extracts. J. Biol. Chem. 277, 33049–33057. doi:10.1074/jbc.M204438200

El-Deiry, W.S., Downey, K.M., So, A.G., 1984. Molecular mechanisms of manganese mutagenesis. Proc. Natl. Acad. Sci. 81, 7378–7382.

Elvers, I., Johansson, F., Groth, P., Erixon, K., Helleday, T., 2011. UV stalled replication forks restart by re-priming in human fibroblasts. Nucleic Acids Res. 39, 7049–7057. doi:10.1093/nar/gkr420

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501. doi:10.1107/S0907444910007493

Engelke, D.R., Krikos, A., Bruck, M.E., Ginsburg, D., 1990. Purification of Thermus aquaticus DNA polymerase expressed in Escherichia coli. Anal. Biochem. 191, 396–400.

Evans, P., 2006. Scaling and assessment of data quality. Acta Crystallogr. D Biol. Crystallogr. 62, 72–82. doi:10.1107/S0907444905036693

Evans, S.J., Fogg, M.J., Mamone, A., Davis, M., Pearl, L.H., Connolly, B.A., 2000. Improving dideoxynucleotide-triphosphate utilisation by the hyper-thermophilic DNA polymerase from the archaeon Pyrococcus furiosus. Nucleic Acids Res. 28, 1059–1066. doi:10.1093/nar/28.5.1059

Faletto, M.B., Miller, W.H., Garvey, E.P., Clair, M.H.S., Daluge, S.M., Good, S.S., 1997. Unique intracellular activation of the potent anti-human immunodeficiency virus agent 1592U89. Antimicrob. Agents Chemother. 41, 1099–1107.

Falkenberg, M., Larsson, N.-G., Gustafsson, C.M., 2007. DNA Replication and Transcription in Mammalian Mitochondria. Annu. Rev. Biochem. 76, 679–699. doi:10.1146/annurev.biochem.76.060305.152028

Fan, J., Pavletich, N.P., 2012. Structure and conformational change of a replication protein A heterotrimer bound to ssDNA. Genes Dev. 26, 2337–2347. doi:10.1101/gad.194787.112

Feldkamp, M.D., Frank, A.O., Kennedy, J.P., Patrone, J.D., Vangamudi, B., Waterson, A.G., Fesik, S.W., Chazin, W.J., 2013. Surface reengineering of RPA70N enables cocrystallization with an inhibitor of the replication protein A interaction motif of ATR interacting protein. Biochemistry (Mosc.) 52, 6515–6524. doi:10.1021/bi400542z

Fisher, R.P., Topper, J.N., Clayton, D.A., 1987. Promoter selection in human mitochondria involves binding of a transcription factor to orientation-independent upstream regulatory elements. Cell 50, 247–258.

Flood, C.L., Rodriguez, G.P., Bao, G., Shockley, A.H., Kow, Y.W., Crouse, G.F., 2015. Replicative DNA Polymerase δ but Not ε Proofreads Errors in Cis and in Trans. PLOS Genet. 11, e1005049. doi:10.1371/journal.pgen.1005049

Fortune, J.M., Pavlov, Y.I., Welch, C.M., Johansson, E., Burgers, P.M.J., Kunkel, T.A., 2005. Saccharomyces cerevisiae DNA Polymerase δ HIGH FIDELITY FOR BASE SUBSTITUTIONS BUT LOWER FIDELITY FOR SINGLE- AND MULTI-BASE DELETIONS. J. Biol. Chem. 280, 29980–29987. doi:10.1074/jbc.M505236200

Frank, A.O., Vangamudi, B., Feldkamp, M.D., Souza-Fagundes, E.M., Luzwick, J.W., Cortez, D., Olejniczak, E.T., Waterson, A.G., Rossanese, O.W., Chazin, W.J., Fesik, S.W., 2014. Discovery of a Potent Stapled Helix Peptide That Binds to the 70N Domain of Replication Protein A. J. Med. Chem. 57, 2455–2461. doi:10.1021/jm401730y

Franklin, M.C., Wang, J., Steitz, T.A., 2001. Structure of the replicating complex of a pol alpha family DNA polymerase. Cell 105, 657–667.

Franklin, R.E., Gosling, R.G., IUCr, 1953. The structure of sodium thymonucleate fibres. I. The influence of water content [WWW Document]. Acta Crystallogr. URL http://scripts.iucr.org/cgi-bin/paper?a00979 (accessed 1.25.17).

Frick, D.N., Richardson, C.C., 2001. DNA Primases. Annu. Rev. Biochem. 70, 39–80. doi:10.1146/annurev.biochem.70.1.39

Frick, D.N., Richardson, C.C., 1999. Interaction of Bacteriophage T7 Gene 4 Primase with Its Template Recognition Site. J. Biol. Chem. 274, 35889–35898. doi:10.1074/jbc.274.50.35889

Friedberg, E.C., 2006. The eureka enzyme: the discovery of DNA polymerase. Nat. Rev. Mol. Cell Biol. 7, 143–147. doi:10.1038/nrm1787

Friedberg, E.C., Lehmann, A.R., Fuchs, R.P.P., 2005. Trading Places: How Do DNA Polymerases Switch during Translesion DNA Synthesis? Mol. Cell 18, 499–505. doi:10.1016/j.molcel.2005.03.032

Fu, Y.V., Yardimci, H., Long, D.T., Ho, T.V., Guainazzi, A., Bermudez, V.P., Hurwitz, J., van Oijen, A., Schärer, O.D., Walter, J.C., 2011. Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. Cell 146, 931–941. doi:10.1016/j.cell.2011.07.045

Fumasoni, M., Zwicky, K., Vanoli, F., Lopes, M., Branzei, D., 2015. Error-Free DNA Damage Tolerance and Sister Chromatid Proximity during DNA Replication Rely on the Polα/Primase/Ctf4 Complex. Mol. Cell 57, 812–823. doi:10.1016/j.molcel.2014.12.038

Furukohri, A., Nishikawa, Y., Akiyama, M.T., Maki, H., 2012. Interaction between Escherichia coli DNA polymerase IV and single-stranded DNA-binding protein is required for DNA synthesis on SSB-coated DNA. Nucleic Acids Res. 40, 6039–6048. doi:10.1093/nar/gks264

Fusté, J.M., Wanrooij, S., Jemt, E., Granycome, C.E., Cluett, T.J., Shi, Y., Atanassova, N., Holt, I.J., Gustafsson, C.M., Falkenberg, M., 2010. Mitochondrial RNA polymerase is needed for activation of the origin of light-strand DNA replication. Mol. Cell 37, 67–78. doi:10.1016/j.molcel.2009.12.021

Galal, W.C., Pan, M., Kelman, Z., Hurwitz, J., 2012. Characterization of DNA Primase Complex Isolated from the Archaeon, Thermococcus kodakaraensis. J. Biol. Chem. 287, 16209–16219. doi:10.1074/jbc.M111.338145

Gambus, A., Jones, R.C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R.D., Labib, K., 2006. GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. Nat. Cell Biol. 8, 358–366. doi:10.1038/ncb1382

Gambus, A., Van Deursen, F., Polychronopoulos, D., Foltman, M., Jones, R.C., Edmondson, R.D., Calzada, A., Labib, K., 2009. A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase α within the eukaryotic replisome. EMBO J. 28, 2992–3004.

García-Gómez, S., Reyes, A., Martínez-Jiménez, M.I., Chocrón, E.S., Mourón, S., Terrados, G., Powell, C., Salido, E., Méndez, J., Holt, I.J., Blanco, L., 2013.

PrimPol, an archaic primase/polymerase operating in human cells. Mol. Cell 52, 541–553. doi:10.1016/j.molcel.2013.09.025

Garg, P., Burgers, P.M., 2005a. Ubiquitinated proliferating cell nuclear antigen activates translesion DNA polymerases η and REV1. Proc. Natl. Acad. Sci. U. S. A. 102, 18361–18366. doi:10.1073/pnas.0505949102

Garg, P., Burgers, P.M.J., 2005b. DNA Polymerases that Propagate the Eukaryotic DNA Replication Fork. Crit. Rev. Biochem. Mol. Biol. 40, 115–128. doi:10.1080/10409230590935433

Garg, P., Stith, C.M., Sabouri, N., Johansson, E., Burgers, P.M., 2004. Idling by DNA polymerase δ maintains a ligatable nick during lagging-strand DNA replication. Genes Dev. 18, 2764–2773. doi:10.1101/gad.1252304

Ge, X.Q., Jackson, D.A., Blow, J.J., 2007. Dormant origins licensed by excess Mcm2–7 are required for human cells to survive replicative stress. Genes Dev. 21, 3331–3341. doi:10.1101/gad.457807

Gee II, J.B., Corbett, R.J.T., Perlman, J.M., Laptook, A.R., 2001. Hypermagnesemia does not increase brain intracellular magnesium in newborn swine. Pediatr. Neurol. 25, 304–308. doi:10.1016/S0887-8994(01)00317-4

Geibel, S., Banchenko, S., Engel, M., Lanka, E., Saenger, W., 2009. Structure and function of primase RepB′ encoded by broad-host-range plasmid RSF1010 that replicates exclusively in leading-strand mode. Proc. Natl. Acad. Sci. 106, 7810–7815. doi:10.1073/pnas.0902910106

Georgescu, R.E., Langston, L., Yao, N.Y., Yurieva, O., Zhang, D., Finkelstein, J., Agarwal, T., O'Donnell, M.E., 2014. Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork. Nat. Struct. Mol. Biol. 21, 664–670. doi:10.1038/nsmb.2851

Ghosh, A., Joy, A., Schuster, G.B., Douki, T., Cadet, J., 2008. Selective one-electron oxidation of duplex DNA oligomers: reaction at thymines. Org. Biomol. Chem. 6, 916–928. doi:10.1039/b717437c

Giannattasio, M., Zwicky, K., Follonier, C., Foiani, M., Lopes, M., Branzei, D., 2014. Visualization of recombination-mediated damage bypass by template switching. Nat. Struct. Mol. Biol. 21, 884–892. doi:10.1038/nsmb.2888

Gilbert, D.M., 2007. Replication origin plasticity, Taylor-made: inhibition vs recruitment of origins under conditions of replication stress. Chromosoma 116, 341–347. doi:10.1007/s00412-007-0105-9

Gill, S., Krupovic, M., Desnoues, N., Béguin, P., Sezonov, G., Forterre, P., 2014. A highly divergent archaeo-eukaryotic primase from the Thermococcus nautilus plasmid, pTN2. Nucleic Acids Res. 42, 3707–3719. doi:10.1093/nar/gkt1385

Göhler, T., Sabbioneda, S., Green, C.M., Lehmann, A.R., 2011. ATR-mediated phosphorylation of DNA polymerase η is needed for efficient recovery from UV damage. J. Cell Biol. 192, 219–227. doi:10.1083/jcb.201008076

Goldschmidt, V., Didierjean, J., Ehresmann, B., Ehresmann, C., Isel, C., Marquet, R., 2006. Mg2+ dependency of HIV-1 reverse transcription, inhibition by nucleoside analogues and resistance. Nucleic Acids Res. 34, 42–52. doi:10.1093/nar/gkj411

Gong, C., Bongiorno, P., Martins, A., Stephanou, N.C., Zhu, H., Shuman, S., Glickman, M.S., 2005. Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C. Nat. Struct. Mol. Biol. 12, 304–312. doi:10.1038/nsmb915

Goodman, M.F., Keener, S., Guidotti, S., Branscomb, E.W., 1983. On the enzymatic basis for mutagenesis by manganese. J. Biol. Chem. 258, 3469–3475.

Gorbalenya, A.E., Koonin, E.V., Donchenko, A.P., Blinov, V.M., 1989. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. Nucleic Acids Res. 17, 4713–4730. doi:10.1093/nar/17.12.4713

Götze, M., Pettelkau, J., Schaks, S., Bosse, K., Ihling, C.H., Krauth, F., Fritzsche, R., Kühn, U., Sinz, A., 2012. StavroX—A Software for Analyzing Crosslinked

Products in Protein Interaction Studies. J. Am. Soc. Mass Spectrom. 23, 76–87. doi:10.1007/s13361-011-0261-2

Gouge, J., Rosario, S., Romain, F., Poitevin, F., Béguin, P., Delarue, M., 2015. Structural basis for a novel mechanism of DNA bridging and alignment in eukaryotic DSB DNA repair. EMBO J. 34, 1126–1142. doi:10.15252/embj.201489643

Greenfeder, S.A., Newlon, C.S., 1992. A replication map of a 61-kb circular derivative of Saccharomyces cerevisiae chromosome III. Mol. Biol. Cell 3, 999–1013.

Griffith, F., 1928. The Significance of Pneumococcal Types. J. Hyg. (Lond.) 27, 113–159.

Gueven, N., Becherel, O.J., Kijas, A.W., Chen, P., Howe, O., Rudolph, J.H., Gatti, R., Date, H., Onodera, O., Taucher-Scholz, G., Lavin, M.F., 2004. Aprataxin, a novel protein that protects against genotoxic stress. Hum. Mol. Genet. 13, 1081–1093. doi:10.1093/hmg/ddh122

Guilliam, T.A., Bailey, L.J., Brissett, N.C., Doherty, A.J., 2016. PolDIP2 interacts with human PrimPol and enhances its DNA polymerase activities. Nucleic Acids Res. 44, 3317–3329. doi:10.1093/nar/gkw175

Guilliam, T.A., Brissett, N.C., Ehlinger, A., Keen, B.A., Kolesar, P., Taylor, E., Bailey, L.J., Lindsay, H.D., Chazin, W.J., Doherty, A.J., 2017. Molecular basis for PrimPol recruitment to replication forks by RPA. Nat. Commun. (accepted).

Guilliam, T.A., Doherty, A.J., 2017. PrimPol—Prime Time to Reprime. Genes 8, 20. doi:10.3390/genes8010020

Guilliam, T.A., Jozwiakowski, S.K., Ehlinger, A., Barnes, R.P., Rudd, S.G., Bailey, L.J., Skehel, J.M., Eckert, K.A., Chazin, W.J., Doherty, A.J., 2015a. Human PrimPol is a highly error-prone polymerase regulated by single-stranded DNA binding proteins. Nucleic Acids Res. 43, 1056–1068. doi:10.1093/nar/gku1321

Guilliam, T.A., Keen, B.A., Brissett, N.C., Doherty, A.J., 2015b. Primase-polymerases are a functionally diverse superfamily of replication and repair enzymes. Nucleic Acids Res. 43, 6651–6664. doi:10.1093/nar/gkv625

Guo, C., Fischhaber, P.L., Luk-Paszyc, M.J., Masuda, Y., Zhou, J., Kamiya, K., Kisker, C., Friedberg, E.C., 2003. Mouse Rev1 protein interacts with multiple DNA polymerases involved in translesion DNA synthesis. EMBO J. 22, 6621–6630. doi:10.1093/emboj/cdg626

Halgasova, N., Mesarosova, I., Bukovska, G., 2012. Identification of a bifunctional primase–polymerase domain of corynephage BFK20 replication protein gp43. Virus Res. 163, 454–460. doi:10.1016/j.virusres.2011.11.005

Halliwell, B., Aruoma, O.I., 1991. DNA damage by oxygen-derived species. Its mechanism and measurement in mammalian systems. FEBS Lett. 281, 9–19.

Hance, N., Ekstrand, M.I., Trifunovic, A., 2005. Mitochondrial DNA polymerase gamma is essential for mammalian embryogenesis. Hum. Mol. Genet. 14, 1775–1783. doi:10.1093/hmg/ddi184

Haracska, L., Torres-Ramos, C.A., Johnson, R.E., Prakash, S., Prakash, L., 2004. Opposing effects of ubiquitin conjugation and SUMO modification of PCNA on replicational bypass of DNA lesions in Saccharomyces cerevisiae. Mol. Cell. Biol. 24, 4267–4274.

Haracska, L., Yu, S.-L., Johnson, R.E., Prakash, L., Prakash, S., 2000. Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase η. Nat. Genet. 25, 458–461. doi:10.1038/78169

Hardin, C.C., Watson, T., Corregan, M., Bailey, C., 1992. Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d(CGCG3GCG). Biochemistry (Mosc.) 31, 833–841. doi:10.1021/bi00118a028

Hashimoto, K., Shimizu, K., Nakashima, N., Sugino, A., 2003. Fidelity of DNA Polymerase δ Holoenzyme from Saccharomyces cerevisiae: The Sliding Clamp Proliferating Cell Nuclear Antigen Decreases Its Fidelity. Biochemistry (Mosc.) 42, 14207–14213. doi:10.1021/bi0348359

Hatimy, A.A., Browne, M.J.G., Flaus, A., Sweet, S.M.M., 2015. Histone H2AX Y142 phosphorylation is a low abundance modification. Int. J. Mass Spectrom., SI:Simon Gaskell honor issue 391, 139–145. doi:10.1016/j.ijms.2015.07.028

Hedglin, M., Pandey, B., Benkovic, S.J., 2016. Characterization of human translesion DNA synthesis across a UV-induced DNA lesion. eLife 5, e19788. doi:10.7554/eLife.19788

Heller, R.C., Marians, K.J., 2006. Replication fork reactivation downstream of a blocked nascent leading strand. Nature 439, 557–562. doi:10.1038/nature04329

Helmrich, A., Ballarino, M., Nudler, E., Tora, L., 2013. Transcription-replication encounters, consequences and genomic instability. Nat. Struct. Mol. Biol. 20, 412–418. doi:10.1038/nsmb.2543

Hendel, A., Krijger, P.H.L., Diamant, N., Goren, Z., Langerak, P., Kim, J., Reissner, T., Lee, K., Geacintov, N.E., Carell, T., Myung, K., Tateishi, S., D'Andrea, A., Jacobs, H., Livneh, Z., 2011. PCNA ubiquitination is important, but not essential for translesion DNA synthesis in mammalian cells. PLoS Genet. 7, e1002262. doi:10.1371/journal.pgen.1002262

Henninger, E.E., Pursell, Z.F., 2014. DNA polymerase ε and its roles in genome stability. IUBMB Life 66, 339–351. doi:10.1002/iub.1276

Hershey, A.D., Chase, M., 1952. Independent functions of viral protein and nucleic acid in growth of bacteriophage. J. Gen. Physiol. 36, 39–56.

Hiasa, H., Marians, K.J., 1994a. Topoisomerase IV can support oriC DNA replication in vitro. J. Biol. Chem. 269, 16371–16375.

Hiasa, H., Marians, K.J., 1994b. Topoisomerase III, but not topoisomerase I, can support nascent chain elongation during theta-type DNA replication. J. Biol. Chem. 269, 32655–32659.

Higgins, N.P., Kato, K., Strauss, B., 1976. A model for replication repair in mammalian cells. J. Mol. Biol. 101, 417–425.

Hile, S.E., Eckert, K.A., 2008. DNA polymerase kappa produces interrupted mutations and displays polar pausing within mononucleotide microsatellite sequences. Nucleic Acids Res. 36, 688–696. doi:10.1093/nar/gkm1089

Hines, J.C., Ray, D.S., 2011. A Second Mitochondrial DNA Primase is Essential for Cell Growth and Kinetoplast Minicircle DNA Replication in Trypanosoma brucei. Eukaryot. Cell. doi:10.1128/EC.00308-10

Hines, J.C., Ray, D.S., 2010. A Mitochondrial DNA Primase Is Essential for Cell Growth and Kinetoplast DNA Replication in Trypanosoma brucei. Mol. Cell. Biol. 30, 1319–1328. doi:10.1128/MCB.01231-09

Hinkle, D.C., Richardson, C.C., 1975. Bacteriophage T7 deoxyribonucleic acid replication in vitro. Purification and properties of the gene 4 protein of bacteriophage T7. J. Biol. Chem. 250, 5523–5529.

Hitomi, K., Iwai, S., Tainer, J.A., 2007. The intricate structural chemistry of base excision repair machinery: implications for DNA damage recognition, removal, and repair. DNA Repair 6, 410–428. doi:10.1016/j.dnarep.2006.10.004

Hoege, C., Pfander, B., Moldovan, G.-L., Pyrowolakis, G., Jentsch, S., 2002. RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature 419, 135–141. doi:10.1038/nature00991

Hoeijmakers, J.H.J., 2009. DNA damage, aging, and cancer. N. Engl. J. Med. 361, 1475–1485. doi:10.1056/NEJMra0804615

Hogg, M., Aller, P., Konigsberg, W., Wallace, S.S., Doublié, S., 2007. Structural and biochemical investigation of the role in proofreading of a beta hairpin loop found in the exonuclease domain of a replicative DNA polymerase of the B family. J. Biol. Chem. 282, 1432–1444. doi:10.1074/jbc.M605675200

Hogg, M., Osterman, P., Bylund, G.O., Ganai, R.A., Lundström, E.-B., Sauer-Eriksson, A.E., Johansson, E., 2014. Structural basis for processive DNA synthesis by yeast DNA polymerase ε. Nat. Struct. Mol. Biol. 21, 49–55. doi:10.1038/nsmb.2712

Hsieh, P., Yamane, K., 2008. DNA mismatch repair: molecular mechanism, cancer, and ageing. Mech. Ageing Dev. 129, 391–407. doi:10.1016/j.mad.2008.02.012

Hu, J., Guo, L., Wu, K., Liu, B., Lang, S., Huang, L., 2012. Template-dependent polymerization across discontinuous templates by the heterodimeric primase from the hyperthermophilic archaeon Sulfolobus solfataricus. Nucleic Acids Res. 40, 3470–3483. doi:10.1093/nar/gkr1256

Huang, D.W., Sherman, B.T., Lempicki, R.A., 2008. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57. doi:10.1038/nprot.2008.211

Huberman, J.A., Riggs, A.D., 1966. Autoradiography of chromosomal DNA fibers from Chinese hamster cells. Proc. Natl. Acad. Sci. 55, 599–606.

Hübscher, U., Maga, G., Spadari, S., 2002. Eukaryotic DNA polymerases. Annu. Rev. Biochem. 71, 133–163.

Huertas, P., Aguilera, A., 2003. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. Mol. Cell 12, 711–721.

Iftode, C., Borowiec, J.A., 2000. 5‘ → 3‘ Molecular Polarity of Human Replication Protein A (hRPA) Binding to Pseudo-Origin DNA Substrates. Biochemistry (Mosc.) 39, 11970–11981. doi:10.1021/bi0005761

Iftode, C., Daniely, Y., Borowiec, J.A., 1999. Replication Protein A (RPA): The Eukaryotic SSB. Crit. Rev. Biochem. Mol. Biol. 34, 141–180. doi:10.1080/10409239991209255

Indiani, C., Patel, M., Goodman, M.F., O’Donnell, M.E., 2013. RecA acts as a switch to regulate polymerase occupancy in a moving replication fork. Proc. Natl. Acad. Sci. 110, 5410–5415. doi:10.1073/pnas.1303301110

Ito, J., Braithwaite, D.K., 1991. Compilation and alignment of DNA polymerase sequences. Nucleic Acids Res. 19, 4045–4057.

Iyama, T., Wilson III, D.M., 2013. DNA repair mechanisms in dividing and non-dividing cells. DNA Repair, Genome maintenance in the Nervous system 12, 620–636. doi:10.1016/j.dnarep.2013.04.015

Iyer, L.M., Koonin, E.V., Leipe, D.D., Aravind, L., 2005. Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: structural insights and new members. Nucleic Acids Res. 33, 3875–3896. doi:10.1093/nar/gki702

Jackson, S.P., Bartek, J., 2009. The DNA-damage response in human biology and disease. Nature 461, 1071–1078. doi:10.1038/nature08467

Jemt, E., Farge, G., Bäckström, S., Holmlund, T., Gustafsson, C.M., Falkenberg, M., 2011. The mitochondrial DNA helicase TWINKLE can assemble on a closed circular template and support initiation of DNA synthesis. Nucleic Acids Res. 39, 9238–9249. doi:10.1093/nar/gkr653

Jiang, Y., Lucas, I., Young, D.J., Davis, E.M., Karrison, T., Rest, J.S., Le Beau, M.M., 2009. Common fragile sites are characterized by histone hypoacetylation. Hum. Mol. Genet. 18, 4501–4512. doi:10.1093/hmg/ddp410

Johansson, E., Dixon, N., 2013. Replicative DNA Polymerases. Cold Spring Harb. Perspect. Biol. 5, a012799. doi:10.1101/cshperspect.a012799

Johnson, R.E., Klassen, R., Prakash, L., Prakash, S., 2015. A Major Role of DNA Polymerase δ in Replication of Both the Leading and Lagging DNA Strands. Mol. Cell 59, 163–175. doi:10.1016/j.molcel.2015.05.038

Johnson, R.E., Prakash, S., Prakash, L., 1999. Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Poleta. Science 283, 1001–1004.

Jozwiakowski, S.K., Borazjani Gholami, F., Doherty, A.J., 2015. Archaeal replicative primases can perform translesion DNA synthesis. Proc. Natl. Acad. Sci. U. S. A. 112, E633-638. doi:10.1073/pnas.1412982112

Jozwiakowski, S.K., Connolly, B.A., 2011. A modified family-B archaeal DNA polymerase with reverse transcriptase activity. Chembiochem Eur. J. Chem. Biol. 12, 35–37. doi:10.1002/cbic.201000640

Jozwiakowski, S.K., Keith, B.J., Gilroy, L., Doherty, A.J., Connolly, B.A., 2014. An archaeal family-B DNA polymerase variant able to replicate past DNA damage: occurrence of replicative and translesion synthesis polymerases within the B family. Nucleic Acids Res. gku683. doi:10.1093/nar/gku683

Kainuma-Kuroda, R., Okazaki, R., 1975. Mechanism of DNA chain growth. J. Mol. Biol. 94, 213–228. doi:10.1016/0022-2836(75)90079-0

Karras, G.I., Fumasoni, M., Sienski, G., Vanoli, F., Branzei, D., Jentsch, S., 2013. Noncanonical role of the 9-1-1 clamp in the error-free DNA damage tolerance pathway. Mol. Cell 49, 536–546. doi:10.1016/j.molcel.2012.11.016

Karras, G.I., Jentsch, S., 2010. The RAD6 DNA Damage Tolerance Pathway Operates Uncoupled from the Replication Fork and Is Functional Beyond S Phase. Cell 141, 255–267. doi:10.1016/j.cell.2010.02.028

Keck, J.L., Roche, D.D., Lynch, A.S., Berger, J.M., 2000. Structure of the RNA Polymerase Domain of E. coli Primase. Science 287, 2482–2486. doi:10.1126/science.287.5462.2482

Keen, B.A., Bailey, L.J., Jozwiakowski, S.K., Doherty, A.J., 2014a. Human PrimPol mutation associated with high myopia has a DNA replication defect. Nucleic Acids Res. 42, 12102–12111. doi:10.1093/nar/gku879

Keen, B.A., Jozwiakowski, S.K., Bailey, L.J., Bianchi, J., Doherty, A.J., 2014b. Molecular dissection of the domain architecture and catalytic activities of human PrimPol. Nucleic Acids Res. 42, 5830–5845. doi:10.1093/nar/gku214

Keith, B.J., Jozwiakowski, S.K., Connolly, B.A., 2013. A plasmid-based lacZα gene assay for DNA polymerase fidelity measurement. Anal. Biochem. 433, 153–161. doi:10.1016/j.ab.2012.10.019

Keller, A., Nesvizhskii, A.I., Kolker, E., Aebersold, R., 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal. Chem. 74, 5383–5392.

Kelley, S., Boroda, S., Musier-Forsyth, K., Kankia, B.I., 2011. HIV-integrase aptamer folds into a parallel quadruplex: a thermodynamic study. Biophys. Chem. 155, 82–88. doi:10.1016/j.bpc.2011.03.004

Kelman, Z., White, M.F., 2005. Archaeal DNA replication and repair. Curr. Opin. Microbiol. 8, 669–676. doi:10.1016/j.mib.2005.10.001

Kenny, M.K., Lee, S.H., Hurwitz, J., 1989. Multiple functions of human single-stranded-DNA binding protein in simian virus 40 DNA replication: single-strand stabilization and stimulation of DNA polymerases alpha and delta. Proc. Natl. Acad. Sci. 86, 9757–9761.

Kent, T., Chandramouly, G., McDevitt, S.M., Ozdemir, A.Y., Pomerantz, R.T., 2015. Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase θ. Nat. Struct. Mol. Biol. 22, 230–237. doi:10.1038/nsmb.2961

Kent, T., Mateos-Gomez, P.A., Sfeir, A., Pomerantz, R.T., 2016. Polymerase θ is a robust terminal transferase that oscillates between three different mechanisms during end-joining. eLife 5, e13740. doi:10.7554/eLife.13740

Keszthelyi, A., Minchell, N.E., Baxter, J., 2016. The Causes and Consequences of Topological Stress during DNA Replication. Genes 7, 134. doi:10.3390/genes7120134

Kilkenny, M.L., Longo, M.A., Perera, R.L., Pellegrini, L., 2013. Structures of human primase reveal design of nucleotide elongation site and mode of Pol α tethering. Proc. Natl. Acad. Sci. 110, 15961–15966. doi:10.1073/pnas.1311185110

Kim, C., Snyder, R.O., Wold, M.S., 1992. Binding properties of replication protein A from human and yeast cells. Mol. Cell. Biol. 12, 3050–3059. doi:10.1128/MCB.12.7.3050

Kim, J.C., Mirkin, S.M., 2013. The balancing act of DNA repeat expansions. Curr. Opin. Genet. Dev. 23, 280–288. doi:10.1016/j.gde.2013.04.009

Kim, N., Huang, S.N., Williams, J.S., Li, Y.C., Clark, A.B., Cho, J.-E., Kunkel, T.A., Pommier, Y., Jinks-Robertson, S., 2011. Mutagenic processing of ribonucleotides in DNA by yeast topoisomerase I. Science 332, 1561–1564. doi:10.1126/science.1205016

Kim, R.A., Wang, J.C., 1989. Function of DNA topoisomerases as replication swivels in Saccharomyces cerevisiae. J. Mol. Biol. 208, 257–267.

Kirk, B.W., Kuchta, R.D., 1999a. Arg304 of Human DNA Primase Is a Key Contributor to Catalysis and NTP Binding: Primase and the Family X Polymerases Share Significant Sequence Homology. Biochemistry (Mosc.) 38, 7727–7736. doi:10.1021/bi990247c

Kirk, B.W., Kuchta, R.D., 1999b. Human DNA Primase: Anion Inhibition, Manganese Stimulation, and Their Effects on In Vitro Start-Site Selection. Biochemistry (Mosc.) 38, 10126–10134. doi:10.1021/bi990351u

Klaile, E., Kukalev, A., Öbrink, B., Müller, M.M., 2008. PDIP38 is a novel mitotic spindle-associated protein that affects spindle organization and chromosome segregation. Cell Cycle 7, 3180–3186. doi:10.4161/cc.7.20.6813

Klenow, H., Overgaard-Hansen, K., 1970. Proteolytic cleavage of DNA polymerase from Escherichia coli B into an exonuclease unit and a polymerase unit. FEBS Lett. 6, 25–27.

Klinedinst, D.K., Challberg, M.D., 1994. Helicase-primase complex of herpes simplex virus type 1: a mutation in the UL52 subunit abolishes primase activity. J. Virol. 68, 3693–3701.

Kobayashi, K., Guilliam, T.A., Tsuda, M., Yamamoto, J., Bailey, L.J., Iwai, S., Takeda, S., Doherty, A.J., Hirota, K., 2016. Repriming by PrimPol is critical for DNA replication restart downstream of lesions and chain-terminating nucleosides. Cell Cycle Georget. Tex 15, 1997–2008. doi:10.1080/15384101.2016.1191711

Koepsell, S., Bastola, D., Hinrichs, S.H., Griep, M.A., 2004. Thermally denaturing high-performance liquid chromatography analysis of primase activity. Anal. Biochem. 332, 330–336. doi:10.1016/j.ab.2004.06.019

Koepsell, S.A., Hanson, S., Hinrichs, S.H., Griep, M.A., 2005. Fluorometric assay for bacterial primases. Anal. Biochem. 339, 353–355. doi:10.1016/j.ab.2004.12.004

Koh, K.D., Balachander, S., Hesselberth, J.R., Storici, F., 2015. Ribose-seq: global mapping of ribonucleotides embedded in genomic DNA. Nat. Methods 12, 251–257. doi:10.1038/nmeth.3259

Kolpashchikov, D.M., Khodyreva, S.N., Khlimankov, D.Y., Wold, M.S., Favre, A., Lavrik, O.I., 2001. Polarity of human replication protein A binding to DNA. Nucleic Acids Res. 29, 373–379. doi:10.1093/nar/29.2.373

Koole, W., van Schendel, R., Karambelas, A.E., van Heteren, J.T., Okihara, K.L., Tijsterman, M., 2014. A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. Nat. Commun. 5, 3216. doi:10.1038/ncomms4216

Koonin, E.V., Wolf, Y.I., Kondrashov, A.S., Aravind, L., 2000. Bacterial homologs of the small subunit of eukaryotic DNA primase. J. Mol. Microbiol. Biotechnol. 2, 509–512.

Korhonen, J.A., Gaspari, M., Falkenberg, M., 2003. TWINKLE Has 5′ → 3′ DNA Helicase Activity and Is Specifically Stimulated by Mitochondrial Single-stranded DNA-binding Protein. J. Biol. Chem. 278, 48627–48632. doi:10.1074/jbc.M306981200

Kornberg, A., Lehman, I.R., Bessman, M.J., Simms, E.S., 1956. Enzymic synthesis of deoxyribonucleic acid. Biochim. Biophys. Acta 21, 197–198.

Kornberg, T., Gefter, M.L., 1971. Purification and DNA synthesis in cell-free extracts: properties of DNA polymerase II. Proc. Natl. Acad. Sci. U. S. A. 68, 761–764.

Kornberg, T., Gefter, M.L., 1970. DNA synthesis in cell-free extracts of a DNA polymerase-defective mutant. Biochem. Biophys. Res. Commun. 40, 1348–1355.

Krevolin, M.D., Calendar, R., 1985. The replication of bacteriophage P4 DNA in vitro. J. Mol. Biol. 182, 509–517. doi:10.1016/0022-2836(85)90237-2

Kuchta, R.D., Stengel, G., 2010a. Mechanism and Evolution of DNA Primases. Biochim. Biophys. Acta 1804, 1180–1189. doi:10.1016/j.bbapap.2009.06.011

Kuchta, R.D., Stengel, G., 2010b. Mechanism and evolution of DNA primases. Biochim. Biophys. Acta BBA - Proteins Proteomics, DNA Polymerase: Structure and Function 1804, 1180–1189. doi:10.1016/j.bbapap.2009.06.011

Kunkel, T.A., Burgers, P.M., 2008. Dividing the workload at a eukaryotic replication fork. Trends Cell Biol. 18, 521–527. doi:10.1016/j.tcb.2008.08.005

Kurat, C.F., Yeeles, J.T.P., Patel, H., Early, A., Diffley, J.F.X., 2017. Chromatin Controls DNA Replication Origin Selection, Lagging-Strand Synthesis, and Replication Fork Rates. Mol. Cell 65, 117–130. doi:10.1016/j.molcel.2016.11.016

Laat, W.L. de, Appeldoorn, E., Sugasawa, K., Weterings, E., Jaspers, N.G.J., Hoeijmakers, J.H.J., 1998. DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. Genes Dev. 12, 2598–2609. doi:10.1101/gad.12.16.2598

Lambert, S., Carr, A.M., 2013a. Impediments to replication fork movement: stabilisation, reactivation and genome instability. Chromosoma 122, 33–45. doi:10.1007/s00412-013-0398-9

Lambert, S., Carr, A.M., 2013b. Replication stress and genome rearrangements: lessons from yeast models. Curr. Opin. Genet. Dev. 23, 132–139. doi:10.1016/j.gde.2012.11.009

Lambert, S., Mizuno, K. 'ichi, Blaisonneau, J., Martineau, S., Chanet, R., Fréon, K., Murray, J.M., Carr, A.M., Baldacci, G., 2010. Homologous Recombination Restarts Blocked Replication Forks at the Expense of Genome Rearrangements by Template Exchange. Mol. Cell 39, 346–359. doi:10.1016/j.molcel.2010.07.015

Lange, S.S., Takata, K., Wood, R.D., 2011. DNA polymerases and cancer. Nat. Rev. Cancer 11, 96–110. doi:10.1038/nrc2998

Lanka, E., Scherzinger, E., Günther, E., Schuster, H., 1979. A DNA primase specified by I-like plasmids. Proc. Natl. Acad. Sci. 76, 3632–3636.

Lao-Sirieix, S., Bell, S.D., 2004. The Heterodimeric Primase of the Hyperthermophilic Archaeon Sulfolobus solfataricus Possesses DNA and RNA Primase, Polymerase and 3′-terminal Nucleotidyl Transferase Activities. J. Mol. Biol. 344, 1251–1263. doi:10.1016/j.jmb.2004.10.018

Lao-Sirieix, S., Pellegrini, L., Bell, S.D., 2005. The promiscuous primase. Trends Genet. 21, 568–572. doi:10.1016/j.tig.2005.07.010

Lao-Sirieix, S.-H., Nookala, R.K., Roversi, P., Bell, S.D., Pellegrini, L., 2005. Structure of the heterodimeric core primase. Nat. Struct. Mol. Biol. 12, 1137–1144. doi:10.1038/nsmb1013

Lazzaro, F., Novarina, D., Amara, F., Watt, D.L., Stone, J.E., Costanzo, V., Burgers, P.M., Kunkel, T.A., Plevani, P., Muzi-Falconi, M., 2012. RNase H and postreplication repair protect cells from ribonucleotides incorporated in DNA. Mol. Cell 45, 99–110. doi:10.1016/j.molcel.2011.12.019

Le Breton, M., Henneke, G., Norais, C., Flament, D., Myllykallio, H., Querellou, J., Raffin, J.-P., 2007. The Heterodimeric Primase from the Euryarchaeon Pyrococcus abyssi: A Multifunctional Enzyme for Initiation and Repair? J. Mol. Biol. 374, 1172–1185. doi:10.1016/j.jmb.2007.10.015

Lehman, I.R., 2003. Discovery of DNA Polymerase. J. Biol. Chem. 278, 34733–34738. doi:10.1074/jbc.X300002200

Lehman, I.R., 1974. DNA ligase: structure, mechanism, and function. Science 186, 790–797.

Lehman, I.R., Bessman, M.J., Simms, E.S., Kornberg, A., 1958. Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from Escherichia coli. J. Biol. Chem. 233, 163–170.

Lehmann, A.R., 2011. DNA polymerases and repair synthesis in NER in human cells. DNA Repair 10, 730–733. doi:10.1016/j.dnarep.2011.04.023

Lehmann, A.R., 1972. Postreplication repair of DNA in ultraviolet-irradiated mammalian cells. J. Mol. Biol. 66, 319–337. doi:10.1016/0022-2836(72)90418-4

Lehmann, A.R., Niimi, A., Ogi, T., Brown, S., Sabbioneda, S., Wing, J.F., Kannouche, P.L., Green, C.M., 2007. Translesion synthesis: Y-family polymerases and the polymerase switch. DNA Repair, Replication Fork Repair Processes 6, 891–899. doi:10.1016/j.dnarep.2007.02.003

Leipe, D.D., Aravind, L., Koonin, E.V., 1999. Did DNA replication evolve twice independently? Nucleic Acids Res. 27, 3389–3401. doi:10.1093/nar/27.17.3389

Leonard, A.C., Méchali, M., 2013. DNA Replication Origins. Cold Spring Harb. Perspect. Biol. 5, a010116. doi:10.1101/cshperspect.a010116

León-Ortiz, A.M., Svendsen, J., Boulton, S.J., 2014. Metabolism of DNA secondary structures at the eukaryotic replication fork. DNA Repair, Cutting-edge Perspectives in Genomic Maintenance 19, 152–162. doi:10.1016/j.dnarep.2014.03.016

Li, V., Hogg, M., Reha-Krantz, L.J., 2010. Identification of a new motif in family B DNA polymerases by mutational analyses of the bacteriophage t4 DNA polymerase. J. Mol. Biol. 400, 295–308. doi:10.1016/j.jmb.2010.05.030

Li, X., Heyer, W.-D., 2008. Homologous recombination in DNA repair and DNA damage tolerance. Cell Res. 18, 99–113. doi:10.1038/cr.2008.1

Li, X., Li, J., Harrington, J., Lieber, M.R., Burgers, P.M., 1995. Lagging strand DNA synthesis at the eukaryotic replication fork involves binding and stimulation of FEN-1 by proliferating cell nuclear antigen. J. Biol. Chem. 270, 22109–22112.

Li, X., Manley, J.L., 2005. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell 122, 365–378. doi:10.1016/j.cell.2005.06.008

Li, Y., Kong, Y., Korolev, S., Waksman, G., 1998a. Crystal structures of the Klenow fragment of Thermus aquaticus DNA polymerase I complexed with deoxyribonucleoside triphosphates. Protein Sci. 7, 1116–1123.

Li, Y., Korolev, S., Waksman, G., 1998b. Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of Thermus aquaticus DNA polymerase I: structural basis for nucleotide incorporation. EMBO J. 17, 7514–7525.

Li, Y., Waksman, G., 2001. Crystal structures of a ddATP-, ddTTP-, ddCTP, and ddGTP-trapped ternary complex of Klentaq1: Insights into nucleotide incorporation and selectivity. Protein Sci. 10, 1225–1233. doi:10.1110/ps.250101

Lieber, M.R., 2010. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu. Rev. Biochem. 79, 181–211. doi:10.1146/annurev.biochem.052308.093131

Lindahl, T., 1993. Instability and decay of the primary structure of DNA. Nature 362, 709–715. doi:10.1038/362709a0

Lipps, G., Röther, S., Hart, C., Krauss, G., 2003. A novel type of replicative enzyme harbouring ATPase, primase and DNA polymerase activity. EMBO J. 22, 2516–2525. doi:10.1093/emboj/cdg246

Lipps, G., Weinzierl, A.O., von Scheven, G., Buchen, C., Cramer, P., 2004. Structure of a bifunctional DNA primase-polymerase. Nat. Struct. Mol. Biol. 11, 157–162. doi:10.1038/nsmb723

Liu, L., Huang, M., 2015. Essential role of the iron-sulfur cluster binding domain of the primase regulatory subunit Pri2 in DNA replication initiation. Protein Cell 6, 194–210. doi:10.1007/s13238-015-0134-8

Liu, L., Komori, K., Ishino, S., Bocquier, A.A., Cann, I.K.O., Kohda, D., Ishino, Y., 2001. The Archaeal DNA Primase BIOCHEMICAL CHARACTERIZATION OF THE p41-p46 COMPLEX FROMPYROCOCCUS FURIOSUS. J. Biol. Chem. 276, 45484–45490. doi:10.1074/jbc.M106391200

Liu, L., Rodriguez-Belmonte, E.M., Mazloum, N., Xie, B., Lee, M.Y.W.T., 2003. Identification of a Novel Protein, PDIP38, That Interacts with the p50 Subunit of DNA Polymerase δ and Proliferating Cell Nuclear Antigen. J. Biol. Chem. 278, 10041–10047. doi:10.1074/jbc.M208694200

Loeb, L.A., Monnat, R.J., 2008. DNA polymerases and human disease. Nat. Rev. Genet. 9, 594–604. doi:10.1038/nrg2345

Loeb, L.A., Preston, B.D., 1986. Mutagenesis by apurinic/apyrimidinic sites. Annu. Rev. Genet. 20, 201–230. doi:10.1146/annurev.ge.20.120186.001221

Longley, M., Smith, L., Copeland, W., 2009. Preparation of Human Mitochondrial Single-Stranded DNA-Binding Protein, in: Stuart, J. (Ed.), Mitochondrial DNA, Methods in Molecular Biology™. Humana Press, pp. 73–85. doi:10.1007/978-1-59745-521-3_5

Lopes, M., Foiani, M., Sogo, J.M., 2006. Multiple Mechanisms Control Chromosome Integrity after Replication Fork Uncoupling and Restart at Irreparable UV Lesions. Mol. Cell 21, 15–27. doi:10.1016/j.molcel.2005.11.015

Lujan, S.A., Williams, J.S., Kunkel, T.A., 2016. DNA Polymerases Divide the Labor of Genome Replication. Trends Cell Biol. 26, 640–654. doi:10.1016/j.tcb.2016.04.012

Lutzmann, M., Maiorano, D., Méchali, M., 2006. A Cdt1-geminin complex licenses chromatin for DNA replication and prevents rereplication during S phase in Xenopus. EMBO J. 25, 5764–5774. doi:10.1038/sj.emboj.7601436

Lyle, A.N., Deshpande, N.N., Taniyama, Y., Seidel-Rogol, B., Pounkova, L., Du, P., Papaharalambus, C., Lassègue, B., Griendling, K.K., 2009. Poldip2, a Novel Regulator of Nox4 and Cytoskeletal Integrity in Vascular Smooth Muscle Cells. Circ. Res. 105, 249–259. doi:10.1161/CIRCRESAHA.109.193722

MacAlpine, D.M., 2016. ORChestrating the human DNA replication program. Proc. Natl. Acad. Sci. 113, 9136–9138. doi:10.1073/pnas.1610336113

MacAlpine, H.K., Gordân, R., Powell, S.K., Hartemink, A.J., MacAlpine, D.M., 2010. Drosophila ORC localizes to open chromatin and marks sites of cohesin complex loading. Genome Res. 20, 201–211. doi:10.1101/gr.097873.109

Maga, G., Crespan, E., Markkanen, E., Imhof, R., Furrer, A., Villani, G., Hübscher, U., Loon, B. van, 2013. DNA polymerase δ-interacting protein 2 is a processivity factor for DNA polymerase λ during 8-oxo-7,8-dihydroguanine bypass. Proc. Natl. Acad. Sci. 110, 18850–18855. doi:10.1073/pnas.1308760110

Maga, G., Villani, G., Crespan, E., Wimmer, U., Ferrari, E., Bertocci, B., Hübscher, U., 2007. 8-oxo-guanine bypass by human DNA polymerases in the presence of auxiliary proteins. Nature 447, 606–608. doi:10.1038/nature05843

Maga, G., Villani, G., Tillement, V., Stucki, M., Locatelli, G.A., Frouin, I., Spadari, S., Hübscher, U., 2001. Okazaki fragment processing: modulation of the strand displacement activity of DNA polymerase delta by the concerted action of replication protein A, proliferating cell nuclear antigen, and flap endonuclease-1. Proc. Natl. Acad. Sci. U. S. A. 98, 14298–14303. doi:10.1073/pnas.251193198

Mahbubani, H.M., Chong, J.P.J., Chevalier, S., Thömmes, P., Blow, J.J., 1997. Cell Cycle Regulation of the Replication Licensing System: Involvement of a Cdk-dependent Inhibitor. J. Cell Biol. 136, 125–135.

Maizels, N., Gray, L.T., 2013. The G4 Genome. PLOS Genet 9, e1003468. doi:10.1371/journal.pgen.1003468

Maréchal, A., Li, J.-M., Ji, X.Y., Wu, C.-S., Yazinski, S.A., Nguyen, H.D., Liu, S., Jiménez, A.E., Jin, J., Zou, L., 2014. PRP19 Transforms into a Sensor of RPA-ssDNA after DNA Damage and Drives ATR Activation via a Ubiquitin-Mediated Circuitry. Mol. Cell 53, 235–246. doi:10.1016/j.molcel.2013.11.002

Maric, M., Maculins, T., Piccoli, G.D., Labib, K., 2014. Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication. Science 346, 1253596. doi:10.1126/science.1253596

Markesbery, W.R., Ehmann, W.D., Alauddin, M., Hossain, T.I.M., 1984. Brain trace element concentrations in aging. Neurobiol. Aging 5, 19–28. doi:10.1016/0197-4580(84)90081-2

Marsden, H.S., McLean, G.W., Barnard, E.C., Francis, G.J., MacEachran, K., Murphy, M., McVey, G., Cross, A., Abbotts, A.P., Stow, N.D., 1997. The catalytic subunit of the DNA polymerase of herpes simplex virus type 1 interacts specifically with the C terminus of the UL8 component of the viral helicase-primase complex. J. Virol. 71, 6390–6397.

Martínez-Jiménez, M.I., García-Gómez, S., Bebenek, K., Sastre-Moreno, G., Calvo, P.A., Díaz-Talavera, A., Kunkel, T.A., Blanco, L., 2015. Alternative solutions and new scenarios for translesion DNA synthesis by human PrimPol. DNA Repair 29, 127–138. doi:10.1016/j.dnarep.2015.02.013

Masuda, Y., Suzuki, M., Piao, J., Gu, Y., Tsurimoto, T., Kamiya, K., 2007. Dynamics of human replication factors in the elongation phase of DNA replication. Nucleic Acids Res. 35, 6904–6916. doi:10.1093/nar/gkm822

Matsui, E., Nishio, M., Yokoyama, H., Harata, K., Darnis, S., Matsui, I., 2003. Distinct Domain Functions Regulating de Novo DNA Synthesis of Thermostable DNA Primase from Hyperthermophile Pyrococcus horikoshii. Biochemistry (Mosc.) 42, 14968–14976. doi:10.1021/bi035556o

Maya-Mendoza, A., Petermann, E., Gillespie, D.A.F., Caldecott, K.W., Jackson, D.A., 2007. Chk1 regulates the density of active replication origins during the vertebrate S phase. EMBO J. 26, 2719–2731. doi:10.1038/sj.emboj.7601714

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674. doi:10.1107/S0021889807021206

McCulloch, S.D., Kunkel, T.A., 2008. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. Cell Res. 18, 148–161. doi:10.1038/cr.2008.4

McElhinny, S.A.N., Kissling, G.E., Kunkel, T.A., 2010. Differential correction of lagging-strand replication errors made by DNA polymerases α and δ. Proc. Natl. Acad. Sci. 107, 21070–21075. doi:10.1073/pnas.1013048107

McGeoch, A.T., Bell, S.D., 2005. Eukaryotic/archaeal primase and MCM proteins encoded in a bacteriophage genome. Cell 120, 167–168. doi:10.1016/j.cell.2004.11.031

McMurray, C.T., 2010. Mechanisms of trinucleotide repeat instability during human development. Nat. Rev. Genet. 11, 786–799. doi:10.1038/nrg2828

McNicholas, S., Potterton, E., Wilson, K.S., Noble, M.E.M., 2011. Presenting your structures: the CCP4mg molecular-graphics software. Acta Crystallogr. D Biol. Crystallogr. 67, 386–394. doi:10.1107/S0907444911007281

Meselson, M., Stahl, F.W., 1958. The replication of DNA in Escherichia coli. Proc. Natl. Acad. Sci. 44, 671–682. doi:10.1073/pnas.44.7.671

Michel, B., 2005. After 30 Years of Study, the Bacterial SOS Response Still Surprises Us. PLOS Biol 3, e255. doi:10.1371/journal.pbio.0030255

Mikhailov, V.S., Rohrmann, G.F., 2002. Baculovirus Replication Factor LEF-1 Is a DNA Primase. J. Virol. 76, 2287–2297. doi:10.1128/jvi.76.5.2287-2297.2002

Minca, E.C., Kowalski, D., 2011. Replication fork stalling by bulky DNA damage: localization at active origins and checkpoint modulation. Nucleic Acids Res. 39, 2610–2623. doi:10.1093/nar/gkq1215

Minko, I.G., Washington, M.T., Kanuri, M., Prakash, L., Prakash, S., Lloyd, R.S., 2003. Translesion Synthesis past Acrolein-derived DNA Adduct, γ-Hydroxypropanodeoxyguanosine, by Yeast and Human DNA Polymerase η. J. Biol. Chem. 278, 784–790. doi:10.1074/jbc.M207774200

Minotti, G., Menna, P., Salvatorelli, E., Cairo, G., Gianni, L., 2004. Anthracyclines: molecular advances and pharmacologic developments in antitumor activity and cardiotoxicity. Pharmacol. Rev. 56, 185–229. doi:10.1124/pr.56.2.6

Miotto, B., Ji, Z., Struhl, K., 2016. Selectivity of ORC binding sites and the relation to replication timing, fragile sites, and deletions in cancers. Proc. Natl. Acad. Sci. 113, E4810–E4819. doi:10.1073/pnas.1609060113

Moreno, S.P., Bailey, R., Campion, N., Herron, S., Gambus, A., 2014. Polyubiquitylation drives replisome disassembly at the termination of DNA replication. Science 346, 477–481. doi:10.1126/science.1253585

Morris, C.F., Sinha, N.K., Alberts, B.M., 1975. Reconstruction of bacteriophage T4 DNA replication apparatus from purified components: rolling circle replication following de novo chain initiation on a single-stranded circular DNA template. Proc. Natl. Acad. Sci. 72, 4800–4804.

Moss, B., 2007. Fields virology, in: Fields Virology. Lippincott Williams & Wilkins, Philadelphia, pp. 2905–2946.

Mourón, S., Rodriguez-Acebes, S., Martínez-Jiménez, M.I., García-Gómez, S., Chocrón, S., Blanco, L., Méndez, J., 2013. Repriming of DNA synthesis at stalled replication forks by human PrimPol. Nat. Struct. Mol. Biol. 20, 1383–1389. doi:10.1038/nsmb.2719

Murakumo, Y., Ogura, Y., Ishii, H., Numata, S., Ichihara, M., Croce, C.M., Fishel, R., Takahashi, M., 2001. Interactions in the error-prone postreplication repair proteins hREV1, hREV3, and hREV7. J. Biol. Chem. 276, 35644–35651. doi:10.1074/jbc.M102051200

Murat, P., Balasubramanian, S., 2014. Existence and consequences of G-quadruplex structures in DNA. Curr. Opin. Genet. Dev., Genome architecture and expression 25, 22–29. doi:10.1016/j.gde.2013.10.012

Murray, A.W., 1991. Cell cycle extracts. Methods Cell Biol. 36, 581–605.

Muzi-Falconi, M., Giannattasio, M., Foiani, M., Plevani, P., 2003. The DNA Polymerase _-Primase Complex: Multiple Functions and Interactions. Sci. World J. 3, 21–33. doi:10.1100/tsw.2003.05

Myers, T.W., Gelfand, D.H., 1991. Reverse transcription and DNA amplification by a Thermus thermophilus DNA polymerase. Biochemistry (Mosc.) 30, 7661–7666. doi:10.1021/bi00245a001

Nakamura, J., Walker, V.E., Upton, P.B., Chiang, S.Y., Kow, Y.W., Swenberg, J.A., 1998. Highly sensitive apurinic/apyrimidinic site assay can detect spontaneous and chemically induced depurination under physiological conditions. Cancer Res. 58, 222–225.

Nelson, J.R., Gibbs, P.E., Nowicka, A.M., Hinkle, D.C., Lawrence, C.W., 2000. Evidence for a second function for Saccharomyces cerevisiae Rev1p. Mol. Microbiol. 37, 549–554.

Nick McElhinny, S.A., Kumar, D., Clark, A.B., Watt, D.L., Watts, B.E., Lundström, E.-B., Johansson, E., Chabes, A., Kunkel, T.A., 2010. Genome instability due to ribonucleotide incorporation into DNA. Nat. Chem. Biol. 6, 774–781. doi:10.1038/nchembio.424

Nirenberg, M., Leder, P., Bernfield, M., Brimacombe, R., Trupin, J., Rottman, F., O'Neal, C., 1965. RNA codewords and protein synthesis, VII. On the general nature of the RNA code. Proc. Natl. Acad. Sci. U. S. A. 53, 1161–1168.

Noia, J.M.D., Neuberger, M.S., 2007. Molecular Mechanisms of Antibody Somatic Hypermutation. Annu. Rev. Biochem. 76, 1–22. doi:10.1146/annurev.biochem.76.061705.090740

Oakley, G.G., Patrick, S.M., 2010. Replication protein A: directing traffic at the intersection of replication and repair. Front. Biosci. J. Virtual Libr. 15, 883–900.

Oehlmann, M., Score, A.J., Blow, J.J., 2004. The role of Cdc6 in ensuring complete genome licensing and S phase checkpoint activation. J. Cell Biol. 165, 181–190. doi:10.1083/jcb.200311044

Ohashi, E., Hanafusa, T., Kamei, K., Song, I., Tomida, J., Hashimoto, H., Vaziri, C., Ohmori, H., 2009. Identification of a novel REV1-interacting motif necessary for

DNA polymerase κ function. Genes Cells Devoted Mol. Cell. Mech. 14, 101–111. doi:10.1111/j.1365-2443.2008.01255.x

Ohashi, E., Murakumo, Y., Kanjo, N., Akagi, J.-I., Masutani, C., Hanaoka, F., Ohmori, H., 2004. Interaction of hREV1 with three human Y-family DNA polymerases. Genes Cells Devoted Mol. Cell. Mech. 9, 523–531. doi:10.1111/j.1356-9597.2004.00747.x

Oliveira, M.T., Kaguni, L.S., 2010. Functional Roles of the N- and C-Terminal Regions of the Human Mitochondrial Single-Stranded DNA-Binding Protein. PLOS ONE 5, e15379. doi:10.1371/journal.pone.0015379

Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G., Steitz, T.A., 1985. Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP. nature 313, 762–766.

Ong, J.L., Loakes, D., Jaroslawski, S., Too, K., Holliger, P., 2006. Directed Evolution of DNA Polymerase, RNA Polymerase and Reverse Transcriptase Activity in a Single Polypeptide. J. Mol. Biol. 361, 537–550. doi:10.1016/j.jmb.2006.06.050

Opresko, P.L., Shiman, R., Eckert, K.A., 2000. Hydrophobic Interactions in the Hinge Domain of DNA Polymerase β Are Important but Not Sufficient for Maintaining Fidelity of DNA Synthesis. Biochemistry (Mosc.) 39, 11399–11407. doi:10.1021/bi000698t

Paeschke, K., Bochman, M.L., Garcia, P.D., Cejka, P., Friedman, K.L., Kowalczykowski, S.C., Zakian, V.A., 2013. Pif1 family helicases suppress genome instability at G-quadruplex motifs. Nature 497, 458–462. doi:10.1038/nature12149

Pal, T., Permuth-Wey, J., Sellers, T.A., 2008. A review of the clinical relevance of mismatch-repair deficiency in ovarian cancer. Cancer 113, 733–742. doi:10.1002/cncr.23601

Parker, M.W., Botchan, M.R., Berger, J.M., 2017. Mechanisms and regulation of DNA replication initiation in eukaryotes. Crit. Rev. Biochem. Mol. Biol. 0, 1–41. doi:10.1080/10409238.2016.1274717

Pascal, J.M., O'Brien, P.J., Tomkinson, A.E., Ellenberger, T., 2004. Human DNA ligase I completely encircles and partially unwinds nicked DNA. Nature 432, 473–478. doi:10.1038/nature03082

Paulsen, R.D., Soni, D.V., Wollman, R., Hahn, A.T., Yee, M.-C., Guan, A., Hesley, J.A., Miller, S.C., Cromwell, E.F., Solow-Cordero, D.E., Meyer, T., Cimprich, K.A., 2009. A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. Mol. Cell 35, 228–239. doi:10.1016/j.molcel.2009.06.021

Pavlov, Y.I., Frahm, C., McElhinny, S.A.N., Niimi, A., Suzuki, M., Kunkel, T.A., 2006. Evidence that Errors Made by DNA Polymerase α are Corrected by DNA Polymerase δ. Curr. Biol. 16, 202–207. doi:10.1016/j.cub.2005.12.002

Pelletier, H., Sawaya, M.R., Wolfle, W., Wilson, S.H., Kraut, J., 1996. A Structural Basis for Metal Ion Mutagenicity and Nucleotide Selectivity in Human DNA Polymerase β,. Biochemistry (Mosc.) 35, 12762–12777. doi:10.1021/bi9529566

Perera, R.L., Torella, R., Klinge, S., Kilkenny, M.L., Maman, J.D., Pellegrini, L., 2013. Mechanism for priming DNA synthesis by yeast DNA Polymerase α. eLife 2, e00482. doi:10.7554/eLife.00482

Perkins, D.N., Pappin, D.J., Creasy, D.M., Cottrell, J.S., 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20, 3551–3567. doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2

Pfander, B., Moldovan, G.-L., Sacher, M., Hoege, C., Jentsch, S., 2005. SUMO-modified PCNA recruits Srs2 to prevent recombination during S phase. Nature 436, 428–433. doi:10.1038/nature03665

Pilzecker, B., Buoninfante, O.A., Pritchard, C., Blomberg, O.S., Huijbers, I.J., van den Berk, P.C.M., Jacobs, H., 2016. PrimPol prevents APOBEC/AID family mediated DNA mutagenesis. Nucleic Acids Res. 44, 4734–4744. doi:10.1093/nar/gkw123

Pitcher, R.S., Brissett, N.C., Picher, A.J., Andrade, P., Juarez, R., Thompson, D., Fox, G.C., Blanco, L., Doherty, A.J., 2007. Structure and Function of a Mycobacterial NHEJ DNA Repair Polymerase. J. Mol. Biol. 366, 391–405. doi:10.1016/j.jmb.2006.10.046

Pitcher, R.S., Tonkin, L.M., Green, A.J., Doherty, A.J., 2005. Domain Structure of a NHEJ DNA Repair Ligase from Mycobacterium tuberculosis. J. Mol. Biol. 351, 531–544. doi:10.1016/j.jmb.2005.06.038

Prakash, A., Kieken, F., Marky, L.A., Borgstahl, G.E.O., 2011. Stabilization of a G-Quadruplex from Unfolding by Replication Protein A Using Potassium and the Porphyrin TMPyP4. J. Nucleic Acids 2011, 529828. doi:10.4061/2011/529828

Prato, S., Vitale, R.M., Contursi, P., Lipps, G., Saviano, M., Rossi, M., Bartolucci, S., 2008. Molecular modeling and functional characterization of the monomeric primase–polymerase domain from the Sulfolobus solfataricus plasmid pIT3. FEBS J. 275, 4389–4402. doi:10.1111/j.1742-4658.2008.06585.x

Prindle, M.J., Schmitt, M.W., Parmeggiani, F., Loeb, L.A., 2013. A substitution in the fingers domain of DNA polymerase δ reduces fidelity by altering nucleotide discrimination in the catalytic site. J. Biol. Chem. 288, 5572–5580. doi:10.1074/jbc.M112.436410

Rampazzo, C., Ferraro, P., Pontarin, G., Fabris, S., Reichard, P., Bianchi, V., 2004. Mitochondrial Deoxyribonucleotides, Pool Sizes, Synthesis, and Regulation. J. Biol. Chem. 279, 17019–17026. doi:10.1074/jbc.M313957200

Rappsilber, J., 2011. The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. J. Struct. Biol., Combining computational modeling with sparse and low-resolution data 173, 530–540. doi:10.1016/j.jsb.2010.10.014

Rastogi, R.P., Richa, Kumar, A., Tyagi, M.B., Sinha, R.P., 2010. Molecular Mechanisms of Ultraviolet Radiation-Induced DNA Damage and Repair. J. Nucleic Acids 2010, e592980. doi:10.4061/2010/592980

Rechkoblit, O., Gupta, Y.K., Malik, R., Rajashankar, K.R., Johnson, R.E., Prakash, L., Prakash, S., Aggarwal, A.K., 2016. Structure and mechanism of human PrimPol, a DNA polymerase with primase activity. Sci. Adv. 2, e1601317. doi:10.1126/sciadv.1601317

Reijns, M.A.M., Kemp, H., Ding, J., de Procé, S.M., Jackson, A.P., Taylor, M.S., 2015. Lagging-strand replication shapes the mutational landscape of the genome. Nature 518, 502–506. doi:10.1038/nature14183

Reijns, M.A.M., Rabe, B., Rigby, R.E., Mill, P., Astell, K.R., Lettice, L.A., Boyle, S., Leitch, A., Keighren, M., Kilanowski, F., Devenney, P.S., Sexton, D., Grimes, G., Holt, I.J., Hill, R.E., Taylor, M.S., Lawson, K.A., Dorin, J.R., Jackson, A.P., 2012. Enzymatic Removal of Ribonucleotides from DNA Is Essential for Mammalian Genome Integrity and Development. Cell 149, 1008–1022. doi:10.1016/j.cell.2012.04.011

Remus, D., Beall, E.L., Botchan, M.R., 2004. DNA topology, not DNA sequence, is a critical determinant for Drosophila ORC–DNA binding. EMBO J. 23, 897–907. doi:10.1038/sj.emboj.7600077

Remus, D., Beuron, F., Tolun, G., Griffith, J.D., Morris, E.P., Diffley, J.F.X., 2009. Concerted Loading of Mcm2-7 Double Hexamers Around DNA during DNA Replication Origin Licensing. Cell 139, 719–730. doi:10.1016/j.cell.2009.10.015

Rhind, N., Gilbert, D.M., 2013. DNA Replication Timing. Cold Spring Harb. Perspect. Biol. 5. doi:10.1101/cshperspect.a010132

Ribeyre, C., Lopes, J., Boulé, J.-B., Piazza, A., Guédin, A., Zakian, V.A., Mergny, J.-L., Nicolas, A., 2009. The Yeast Pif1 Helicase Prevents Genomic Instability Caused by G-Quadruplex-Forming CEB1 Sequences In Vivo. PLOS Genet 5, e1000475. doi:10.1371/journal.pgen.1000475

Ricke, R.M., Bielinsky, A.-K., 2004. Mcm10 regulates the stability and chromatin association of DNA polymerase-α. Mol. Cell 16, 173–185.

Ritzi, M., Baack, M., Musahl, C., Romanowski, P., Laskey, R.A., Knippers, R., 1998. Human minichromosome maintenance proteins and human origin recognition complex 2 protein on chromatin. J. Biol. Chem. 273, 24543–24549.

Robison, J.G., Elliott, J., Dixon, K., Oakley, G.G., 2004. Replication Protein A and the Mre11·Rad50·Nbs1 Complex Co-localize and Interact at Sites of Stalled Replication Forks. J. Biol. Chem. 279, 34802–34810. doi:10.1074/jbc.M404750200

Roerink, S.F., van Schendel, R., Tijsterman, M., 2014. Polymerase theta-mediated end joining of replication-associated DNA breaks in C. elegans. Genome Res. 24, 954–962. doi:10.1101/gr.170431.113

Romano, L.J., Richardson, C.C., 1979. Characterization of the ribonucleic acid primers and the deoxyribonucleic acid product synthesized by the DNA polymerase and gene 4 protein of bacteriophage T7. J. Biol. Chem. 254, 10483–10489.

Ropp, P.A., Copeland, W.C., 1996. Cloning and characterization of the human mitochondrial DNA polymerase, DNA polymerase gamma. Genomics 36, 449–458. doi:10.1006/geno.1996.0490

Rothwell, P.J., Waksman, G., 2005. Structure and mechanism of DNA polymerases, in: Chemistry, B.-A. in P. (Ed.), Fibrous Proteins: Muscle and Molecular Motors. Academic Press, pp. 401–440.

Rowen, L., Kornberg, A., 1978. Primase, the dnaG protein of Escherichia coli. An enzyme which starts DNA chains. J. Biol. Chem. 253, 758–764.

Rudd, S.G., 2013. Cellular and biochemical characterisation of PrimPol, a novel eukaryotic primase-polymerase involved in DNA damage tolerance (Ph.D.). University of Sussex.

Rudd, S.G., Bianchi, J., Doherty, A.J., 2014. PrimPol-A new polymerase on the block. Mol. Cell. Oncol. 1, e960754. doi:10.4161/23723548.2014.960754

Rudd, S.G., Glover, L., Jozwiakowski, S.K., Horn, D., Doherty, A.J., 2013. PPL2 Translesion Polymerase Is Essential for the Completion of Chromosomal DNA Replication in the African Trypanosome. Mol. Cell 52, 554–565. doi:10.1016/j.molcel.2013.10.034

Rupp, W.D., Howard-Flanders, P., 1968. Discontinuities in the DNA synthesized in an excision-defective strain of Escherichia coli following ultraviolet irradiation. J. Mol. Biol. 31, 291–304.

Sabbioneda, S., Gourdin, A.M., Green, C.M., Zotter, A., Giglia-Mari, G., Houtsmuller, A., Vermeulen, W., Lehmann, A.R., 2008. Effect of proliferating cell nuclear antigen ubiquitination and chromatin structure on the dynamic properties of the Y-family DNA polymerases. Mol. Biol. Cell 19, 5193–5202. doi:10.1091/mbc.E08-07-0724

Sabbioneda, S., Green, C.M., Bienko, M., Kannouche, P., Dikic, I., Lehmann, A.R., 2009. Ubiquitin-binding motif of human DNA polymerase η is required for correct localization. Proc. Natl. Acad. Sci. U. S. A. 106, E20. doi:10.1073/pnas.0812744106

Sabouri, N., Viberg, J., Goyal, D.K., Johansson, E., Chabes, A., 2008. Evidence for lesion bypass by yeast replicative DNA polymerases during DNA damage. Nucleic Acids Res. 36, 5660–5667. doi:10.1093/nar/gkn555

Safa, L., Gueddouda, N.M., Thiébaut, F., Delagoutte, E., Petruseva, I., Lavrik, O., Mendoza, O., Bourdoncle, A., Alberti, P., Riou, J.-F., Saintomé, C., 2016. 5′ to 3′ Unfolding Directionality of DNA Secondary Structures by Replication Protein A G-QUADRUPLEXES AND DUPLEXES. J. Biol. Chem. 291, 21246–21256. doi:10.1074/jbc.M115.709667

Sale, J.E., Batters, C., Edmunds, C.E., Phillips, L.G., Simpson, L.J., Szüts, D., 2009. Timing matters: error-prone gap filling and translesion synthesis in immunoglobulin gene hypermutation. Philos. Trans. R. Soc. Lond. B Biol. Sci. 364, 595–603. doi:10.1098/rstb.2008.0197

Sale, J.E., Lehmann, A.R., Woodgate, R., 2012. Y-family DNA polymerases and their role in tolerance of cellular DNA damage. Nat. Rev. Mol. Cell Biol. 13, 141–152. doi:10.1038/nrm3289

Samuels, M., Gulati, G., Shin, J.-H., Opara, R., McSweeney, E., Sekedat, M., Long, S., Kelman, Z., Jeruzalmi, D., 2009. A biochemically active MCM-like helicase in Bacillus cereus. Nucleic Acids Res. gkp376. doi:10.1093/nar/gkp376

Sanchez-Berrondo, J., Mesa, P., Ibarra, A., Martínez-Jiménez, M.I., Blanco, L., Méndez, J., Boskovic, J., Montoya, G., 2012. Molecular architecture of a multifunctional MCM complex. Nucleic Acids Res. 40, 1366–1380. doi:10.1093/nar/gkr831

Sarkies, P., Murat, P., Phillips, L.G., Patel, K.J., Balasubramanian, S., Sale, J.E., 2011. FANCJ coordinates two pathways that maintain epigenetic stability at G-quadruplex DNA. Nucleic Acids Res. gkr868. doi:10.1093/nar/gkr868

Sarkies, P., Reams, C., Simpson, L.J., Sale, J.E., 2010. Epigenetic Instability due to Defective Replication of Structured DNA. Mol. Cell 40, 703–713. doi:10.1016/j.molcel.2010.11.009

Scherzinger, E., Bagdasarian, M.M., Scholz, P., Lurz, R., Rückert, B., Bagdasarian, M., 1984. Replication of the broad host range plasmid RSF1010: requirement for three plasmid-encoded proteins. Proc. Natl. Acad. Sci. 81, 654–658.

Scherzinger, E., Haring, V., Lurz, R., Otto, S., 1991. Plasmid RSF1010 DNA replication in vitro promoted by purified RSF1010 RepA, RepB and RepC proteins. Nucleic Acids Res. 19, 1203–1211. doi:10.1093/nar/19.6.1203

Scherzinger, E., Lanka, E., Morelli, G., Seiffert, D., Yuki, A., 1977. Bacteriophage-T7-Induced DNA-Priming Protein. Eur. J. Biochem. 72, 543–558. doi:10.1111/j.1432-1033.1977.tb11278.x

Scherzinger, E., Litfin, F., 1974. Initiation of the Replication of Single-Stranded DNA by Concerted Action of Phage T7 RNA and DNA Polymerases. Eur. J. Biochem. 46, 179–180. doi:10.1111/j.1432-1033.1974.tb03610.x

Schiavone, D., Guilbaud, G., Murat, P., Papadopoulou, C., Sarkies, P., Prioleau, M.-N., Balasubramanian, S., Sale, J.E., 2014. Determinants of G quadruplex-induced epigenetic instability in REV1-deficient cells. EMBO J. 33, 2507–2520. doi:10.15252/embj.201488398

Schiavone, D., Jozwiakowski, S.K., Romanello, M., Guilbaud, G., Guilliam, T.A., Bailey, L.J., Sale, J.E., Doherty, A.J., 2016. PrimPol Is Required for Replicative Tolerance of G Quadruplexes in Vertebrate Cells. Mol. Cell 61, 161–169. doi:10.1016/j.molcel.2015.10.038

Schmidt, T.G.M., Koepke, J., Frank, R., Skerra, A., 1996. Molecular Interaction Between the Strep-tag Affinity Peptide and its Cognate Target, Streptavidin. J. Mol. Biol. 255, 753–766. doi:10.1006/jmbi.1996.0061

Schmitt, M.W., Matsumoto, Y., Loeb, L.A., 2009. High fidelity and lesion bypass capability of human DNA polymerase delta. Biochimie 91, 1163–1172. doi:10.1016/j.biochi.2009.06.007

Shapiro, T.A., Englund, P.T., 1995. The structure and replication of kinetoplast DNA. Annu. Rev. Microbiol. 49, 117–143. doi:10.1146/annurev.mi.49.100195.001001

Shcherbakova, P.V., Pavlov, Y.I., Chilkova, O., Rogozin, I.B., Johansson, E., Kunkel, T.A., 2003. Unique Error Signature of the Four-subunit Yeast DNA Polymerase ϵ. J. Biol. Chem. 278, 43770–43780. doi:10.1074/jbc.M306893200

Sheaff, R.J., Kuchta, R.D., 1994. Misincorporation of nucleotides by calf thymus DNA primase and elongation of primers containing multiple noncognate nucleotides by DNA polymerase alpha. J. Biol. Chem. 269, 19225–19231.

Shen, J., Gilmore, E.C., Marshall, C.A., Haddadin, M., Reynolds, J.J., Eyaid, W., Bodell, A., Barry, B., Gleason, D., Allen, K., Ganesh, V.S., Chang, B.S., Grix, A., Hill, R.S., Topcu, M., Caldecott, K.W., Barkovich, A.J., Walsh, C.A., 2010. Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. Nat. Genet. 42, 245–249. doi:10.1038/ng.526

Sherman, G., Gottlieb, J., Challberg, M.D., 1992. The UL8 subunit of the herpes simplex virus helicase-primase complex is required for efficient primer utilization. J. Virol. 66, 4884–4892.

Shevchenko, A., Tomas, H., Havli, J., Olsen, J.V., Mann, M., 2007. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protoc. 1, 2856–2860. doi:10.1038/nprot.2006.468

Shima, N., Alcaraz, A., Liachko, I., Buske, T.R., Andrews, C.A., Munroe, R.J., Hartford, S.A., Tye, B.K., Schimenti, J.C., 2007. A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. Nat. Genet. 39, 93–98. doi:10.1038/ng1936

Shimizu, K., Hashimoto, K., Kirchner, J.M., Nakai, W., Nishikawa, H., Resnick, M.A., Sugino, A., 2002. Fidelity of DNA Polymerase ε Holoenzyme from Budding YeastSaccharomyces cerevisiae. J. Biol. Chem. 277, 37422–37429.

Siddiqui, K., On, K.F., Diffley, J.F.X., 2013. Regulating DNA Replication in Eukarya. Cold Spring Harb. Perspect. Biol. 5, a012930. doi:10.1101/cshperspect.a012930

Silva, F.S.D., Lewis, W., Berglund, P., Koonin, E.V., Moss, B., 2007. Poxvirus DNA primase. Proc. Natl. Acad. Sci. 104, 18724–18729. doi:10.1073/pnas.0709276104

Silverstein, T.D., Johnson, R.E., Jain, R., Prakash, L., Prakash, S., Aggarwal, A.K., 2010. Structural basis for the suppression of skin cancers by DNA polymerase η. Nature 465, 1039–1043. doi:10.1038/nature09104

Simon, A.C., Zhou, J.C., Perera, R.L., van Deursen, F., Evrin, C., Ivanova, M.E., Kilkenny, M.L., Renault, L., Kjaer, S., Matak-Vinković, D., Labib, K., Costa, A., Pellegrini, L., 2014. A Ctf4 trimer couples the CMG helicase to DNA polymerase α in the eukaryotic replisome. Nature 510, 293–297. doi:10.1038/nature13234

Simsek, D., Furda, A., Gao, Y., Artus, J., Brunet, E., Hadjantonakis, A.-K., Van Houten, B., Shuman, S., McKinnon, P.J., Jasin, M., 2011. Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. Nature 471, 245–248. doi:10.1038/nature09794

Singh, B., Li, X., Owens, K.M., Vanniarajan, A., Liang, P., Singh, K.K., 2015. Human REV3 DNA Polymerase Zeta Localizes to Mitochondria and Protects the Mitochondrial Genome. PLOS ONE 10, e0140409. doi:10.1371/journal.pone.0140409

Sørensen, C.S., Syljuåsen, R.G., 2012. Safeguarding genome integrity: the checkpoint kinases ATR, CHK1 and WEE1 restrain CDK activity during normal DNA replication. Nucleic Acids Res. 40, 477–486. doi:10.1093/nar/gkr697

Souza-Fagundes, E.M., Frank, A.O., Feldkamp, M.D., Dorset, D.C., Chazin, W.J., Rossanese, O.W., Olejniczak, E.T., Fesik, S.W., 2012. A high-throughput fluorescence polarization anisotropy assay for the 70N domain of replication protein A. Anal. Biochem. 421, 742–749. doi:10.1016/j.ab.2011.11.025

Spelbrink, J.N., Li, F.Y., Tiranti, V., Nikali, K., Yuan, Q.P., Tariq, M., Wanrooij, S., Garrido, N., Comi, G., Morandi, L., Santoro, L., Toscano, A., Fabrizi, G.M., Somer, H., Croxen, R., Beeson, D., Poulton, J., Suomalainen, A., Jacobs, H.T., Zeviani, M., Larsson, C., 2001. Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria. Nat. Genet. 28, 223–231. doi:10.1038/90058

Steitz, T.A., 1999. DNA Polymerases: Structural Diversity and Common Mechanisms. J. Biol. Chem. 274, 17395–17398. doi:10.1074/jbc.274.25.17395

Steitz, T.A., Smerdon, S.J., Jager, J., Joyce, C.M., 1994. A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. Science 266, 2022–2025. doi:10.1126/science.7528445

Stirling, P.C., Chan, Y.A., Minaker, S.W., Aristizabal, M.J., Barrett, I., Sipahimalani, P., Kobor, M.S., Hieter, P., 2012. R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. Genes Dev. 26, 163–175. doi:10.1101/gad.179721.111

Stojkovič, G., Makarova, A.V., Wanrooij, P.H., Forslund, J., Burgers, P.M., Wanrooij, S., 2016. Oxidative DNA damage stalls the human mitochondrial replisome. Sci. Rep. 6, 28942. doi:10.1038/srep28942

Strätling, W., Knippers, R., 1973. Function and Purification of Gene 4 Protein of Phage T7. Nature 245, 195–197. doi:10.1038/245195a0

Sundin, O., Varshavsky, A., 1981. Arrest of segregation leads to accumulation of highly intertwined catenated dimers: dissection of the final stages of SV40 DNA replication. Cell 25, 659–669.

Swan, M.K., Johnson, R.E., Prakash, L., Prakash, S., Aggarwal, A.K., 2009. Structural basis of high-fidelity DNA synthesis by yeast DNA polymerase delta. Nat. Struct. Mol. Biol. 16, 979–986. doi:10.1038/nsmb.1663

Sweetser, D., Nonet, M., Young, R.A., 1987. Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. Proc. Natl. Acad. Sci. 84, 1192–1196.

Takashima, H., Boerkoel, C.F., John, J., Saifi, G.M., Salih, M.A.M., Armstrong, D., Mao, Y., Quiocho, F.A., Roa, B.B., Nakagawa, M., Stockton, D.W., Lupski, J.R., 2002. Mutation of TDP1, encoding a topoisomerase I-dependent DNA damage repair enzyme, in spinocerebellar ataxia with axonal neuropathy. Nat. Genet. 32, 267–272. doi:10.1038/ng987

Takechi, S., Itoh, T., 1995. Initiation of unidirectional ColE2 DNA replication by a unique priming mechanism. Nucleic Acids Res. 23, 4196–4201. doi:10.1093/nar/23.20.4196

Takechi, S., Matsui, H., Itoh, T., 1995. Primer RNA synthesis by plasmid-specified Rep protein for initiation of ColE2 DNA replication. EMBO J. 14, 5141–5147.

Tanaka, S., Araki, H., 2013. Helicase Activation and Establishment of Replication Forks at Chromosomal Origins of Replication. Cold Spring Harb. Perspect. Biol. 5, a010371. doi:10.1101/cshperspect.a010371

Taylor, E.M., Bonsu, N.M., Price, R.J., Lindsay, H.D., 2013. Depletion of Uhrf1 inhibits chromosomal DNA replication in Xenopus egg extracts. Nucleic Acids Res. 41, 7725–7737. doi:10.1093/nar/gkt549

Tissier, A., Janel-Bintz, R., Coulon, S., Klaile, E., Kannouche, P., Fuchs, R.P., Cordonnier, A.M., 2010. Crosstalk between replicative and translesional DNA polymerases: PDIP38 interacts directly with Polη. DNA Repair 9, 922–928. doi:10.1016/j.dnarep.2010.04.010

Tissier, A., Kannouche, P., Reck, M.-P., Lehmann, A.R., Fuchs, R.P.P., Cordonnier, A., 2004. Co-localization in replication foci and interaction of human Y-family members, DNA polymerase pol eta and REVl protein. DNA Repair 3, 1503–1514. doi:10.1016/j.dnarep.2004.06.015

Topper, J.N., Clayton, D.A., 1990. Characterization of human MRP/Th RNA and its nuclear gene: full length MRP/Th RNA is an active endoribonuclease when assembled as an RNP. Nucleic Acids Res. 18, 793–799.

Tsao, C.-C., Geisen, C., Abraham, R.T., 2004. Interaction between human MCM7 and Rad17 proteins is required for replication checkpoint signaling. EMBO J. 23, 4660–4669. doi:10.1038/sj.emboj.7600463

Tsubota, T., Maki, S., Kubota, H., Sugino, A., Maki, H., 2003. Double-stranded DNA binding properties of Saccharomyces cerevisiae DNA polymerase epsilon and of the Dpb3p-Dpb4p subassembly. Genes Cells Devoted Mol. Cell. Mech. 8, 873–888.

Tsugita, A., Fraenkel-Conrat, H., 1960. THE AMINO ACID COMPOSITION AND C-TERMINAL SEQUENCE OF A CHEMICALLY EVOKED MUTANT OF TMV. Proc. Natl. Acad. Sci. U. S. A. 46, 636–642.

Tsurimoto, T., Stillman, B., 1989. Multiple replication factors augment DNA synthesis by the two eukaryotic DNA polymerases, alpha and delta. EMBO J. 8, 3883–3889.

Tyynismaa, H., Sembongi, H., Bokori-Brown, M., Granycome, C., Ashley, N., Poulton, J., Jalanko, A., Spelbrink, J.N., Holt, I.J., Suomalainen, A., 2004. Twinkle helicase

is essential for mtDNA maintenance and regulates mtDNA copy number. Hum. Mol. Genet. 13, 3219–3227. doi:10.1093/hmg/ddh342

Ulrich, H.D., Jentsch, S., 2000. Two RING finger proteins mediate cooperation between ubiquitin-conjugating enzymes in DNA repair. EMBO J. 19, 3388–3397. doi:10.1093/emboj/19.13.3388

Vaisman, A., Ling, H., Woodgate, R., Yang, W., 2005. Fidelity of Dpo4: effect of metal ions, nucleotide selection and pyrophosphorolysis. EMBO J. 24, 2957–2967. doi:10.1038/sj.emboj.7600786

Vaithiyalingam, S., Arnett, D.R., Aggarwal, A., Eichman, B.F., Fanning, E., Chazin, W.J., 2014. Insights into Eukaryotic Primer Synthesis from Structures of the p48 Subunit of Human DNA Primase. J. Mol. Biol. 426, 558–569. doi:10.1016/j.jmb.2013.11.007

Vallerga, M.B., Mansilla, S.F., Federico, M.B., Bertolin, A.P., Gottifredi, V., 2015. Rad51 recombinase prevents Mre11 nuclease-dependent degradation and excessive PrimPol-mediated elongation of nascent DNA after UV irradiation. Proc. Natl. Acad. Sci. U. S. A. 112, E6624-6633. doi:10.1073/pnas.1508543112

Van, C., Yan, S., Michael, W.M., Waga, S., Cimprich, K.A., 2010. Continued primer synthesis at stalled replication forks contributes to checkpoint activation. J. Cell Biol. 189, 233–246. doi:10.1083/jcb.200909105

Van Goethem, G., Dermaut, B., Löfgren, A., Martin, J.J., Van Broeckhoven, C., 2001. Mutation of POLG is associated with progressive external ophthalmoplegia characterized by mtDNA deletions. Nat. Genet. 28, 211–212. doi:10.1038/90034

Van Kirk, C., Feinberg, L., Robertson, D., Freeman, W., Vrana, K., 2010. Phosphorimager, in: John Wiley & Sons, Ltd (Ed.), Encyclopedia of Life Sciences. John Wiley & Sons, Ltd, Chichester, UK.

vanAnkeren, S.C., Murray, D., Meyn, R.E., 1988. Induction and rejoining of gamma-ray-induced DNA single- and double-strand breaks in Chinese hamster AA8 cells and in two radiosensitive clones. Radiat. Res. 116, 511–525.

Vanoli, F., Fumasoni, M., Szakal, B., Maloisel, L., Branzei, D., 2010. Replication and recombination factors contributing to recombination-dependent bypass of DNA lesions by template switch. PLoS Genet. 6, e1001205. doi:10.1371/journal.pgen.1001205

Vashee, S., Cvetic, C., Lu, W., Simancek, P., Kelly, T.J., Walter, J.C., 2003. Sequence-independent DNA binding and replication initiation by the human origin recognition complex. Genes Dev. 17, 1894–1908. doi:10.1101/gad.1084203

Versieck, J., McCall, J.T., 1985. Trace Elements in Human Body Fluids and Tissues. CRC Crit. Rev. Clin. Lab. Sci. 22, 97–184. doi:10.3109/10408368509165788

Voss, S., Skerra, A., 1997. Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the Strep-tag II peptide and improved performance in recombinant protein purification. Protein Eng. 10, 975–982. doi:10.1093/protein/10.8.975

Wahba, L., Amon, J.D., Koshland, D., Vuica-Ross, M., 2011. RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. Mol. Cell 44, 978–988. doi:10.1016/j.molcel.2011.10.017

Walter, P., Klein, F., Lorentzen, E., Ilchmann, A., Klug, G., Evguenieva-Hackenberg, E., 2006. Characterization of native and reconstituted exosome complexes from the hyperthermophilic archaeon Sulfolobus solfataricus. Mol. Microbiol. 62, 1076–1089. doi:10.1111/j.1365-2958.2006.05393.x

Wan, L., Lou, J., Xia, Y., Su, B., Liu, T., Cui, J., Sun, Y., Lou, H., Huang, J., 2013. hPrimpol1/CCDC111 is a human DNA primase-polymerase required for the maintenance of genome integrity. EMBO Rep. 14, 1104–1112. doi:10.1038/embor.2013.159

Wang, J.C., 1998. Moving one DNA double helix through another by a type II DNA topoisomerase: the story of a simple molecular machine. Q. Rev. Biophys. 31, 107–144.

Wang, J.C., 1996. DNA topoisomerases. Annu. Rev. Biochem. 65, 635–692. doi:10.1146/annurev.bi.65.070196.003223

Wang, T.S.-F., Eichler, D.C., Korn, D., 1977. Effect of manganese(2+) ions on the in vitro activity of human deoxyribonucleic acid polymerase .beta. Biochemistry (Mosc.) 16, 4927–4934. doi:10.1021/bi00641a029

Wanrooij, S., Miralles Fusté, J., Stewart, J.B., Wanrooij, P.H., Samuelsson, T., Larsson, N.-G., Gustafsson, C.M., Falkenberg, M., 2012. In vivo mutagenesis reveals that OriL is essential for mitochondrial DNA replication. EMBO Rep. 13, 1130–1137. doi:10.1038/embor.2012.161

Waters, L.S., Walker, G.C., 2006. The critical mutagenic translesion DNA polymerase Rev1 is highly expressed during G(2)/M phase rather than S phase. Proc. Natl. Acad. Sci. U. S. A. 103, 8971–8976. doi:10.1073/pnas.0510167103

Watson, J.D., Crick, F.H., 1953. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 171, 737–738.

Weller, G.R., Doherty, A.J., 2001. A family of DNA repair ligases in bacteria? FEBS Lett. 505, 340–342. doi:10.1016/S0014-5793(01)02831-9

Weller, G.R., Kysela, B., Roy, R., Tonkin, L.M., Scanlan, E., Della, M., Devine, S.K., Day, J.P., Wilkinson, A., Fagagna, F. d'Adda di, Devine, K.M., Bowater, R.P., Jeggo, P.A., Jackson, S.P., Doherty, A.J., 2002. Identification of a DNA Nonhomologous End-Joining Complex in Bacteria. Science 297, 1686–1689. doi:10.1126/science.1074584

Wenger, C.D., Phanstiel, D.H., Lee, M.V., Bailey, D.J., Coon, J.J., 2011. COMPASS: A suite of pre- and post-search proteomics software tools for OMSSA. PROTEOMICS 11, 1064–1074. doi:10.1002/pmic.201000616

Wickramasinghe, C.M., Arzouk, H., Frey, A., Maiter, A., Sale, J.E., 2015. Contributions of the specialised DNA polymerases to replication of structured DNA. DNA Repair, DNA polymerases 29, 83–90. doi:10.1016/j.dnarep.2015.01.004

Wilkins, M., Stokes, A., Wilson, H., 1953. Molecular Structure of Nucleic Acids: Molecular Structure of Deoxypentose Nucleic Acids. Nature 171, 738–740. doi:10.1038/171738a0

Williams, J.S., Clausen, A.R., Lujan, S.A., Marjavaara, L., Clark, A.B., Burgers, P.M., Chabes, A., Kunkel, T.A., 2015. Evidence that processing of ribonucleotides in DNA by topoisomerase 1 is leading-strand specific. Nat. Struct. Mol. Biol. 22, 291–297. doi:10.1038/nsmb.2989

Williams, J.S., Kunkel, T.A., 2014. Ribonucleotides in DNA: Origins, repair and consequences. DNA Repair, Cutting-edge Perspectives in Genomic Maintenance 19, 27–37. doi:10.1016/j.dnarep.2014.03.029

Williams, J.S., Lujan, S.A., Kunkel, T.A., 2016. Processing ribonucleotides incorporated during eukaryotic DNA replication. Nat. Rev. Mol. Cell Biol. 17, 350–363. doi:10.1038/nrm.2016.37

Williams, J.S., Smith, D.J., Marjavaara, L., Lujan, S.A., Chabes, A., Kunkel, T.A., 2013. Topoisomerase 1-mediated removal of ribonucleotides from nascent leading-strand DNA. Mol. Cell 49, 1010–1015. doi:10.1016/j.molcel.2012.12.021

Wilson, A.C., Cann, R.L., Carr, S.M., George, M., Gyllensten, U.B., Helm-Bychowski, K.M., Higuchi, R.G., Palumbi, S.R., Prager, E.M., Sage, R.D., Stoneking, M., 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linn. Soc. 26, 375–400. doi:10.1111/j.1095-8312.1985.tb02048.x

Winter, G., Lobley, C.M.C., Prince, S.M., 2013. Decision making in xia2. Acta Crystallogr. D Biol. Crystallogr. 69, 1260–1273.

Wisnovsky, S., Jean, S.R., Kelley, S.O., 2016. Mitochondrial DNA repair and replication proteins revealed by targeted chemical probes. Nat. Chem. Biol. 12, 567–573. doi:10.1038/nchembio.2102

Wolfson, J., Dressler, D., 1972. Regions of Single-Stranded DNA in the Growing Points of Replicating Bacteriophage T7 Chromosomes. Proc. Natl. Acad. Sci. 69, 2682–2686.

Wong, A., Zhang, S., Mordue, D., Wu, J.M., Zhang, Z., Darzynkiewicz, Z., Lee, E.Y., Lee, M.Y., 2013. PDIP38 is translocated to the spliceosomes/nuclear speckles in response to UV-induced DNA damage and is required for UV-induced alternative splicing of MDM2. Cell Cycle 12, 3373–3382. doi:10.4161/cc.26221

Woodford, K.J., Howell, R.M., Usdin, K., 1994. A novel K(+)-dependent DNA synthesis arrest site in a commonly occurring sequence motif in eukaryotes. J. Biol. Chem. 269, 27029–27035.

Woodward, A.M., Göhler, T., Luciani, M.G., Oehlmann, M., Ge, X., Gartner, A., Jackson, D.A., Blow, J.J., 2006. Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress. J. Cell Biol. 173, 673–683. doi:10.1083/jcb.200602108

Wu, C.A., Zechner, E.L., Marians, K.J., 1992. Coordinated leading- and lagging-strand synthesis at the Escherichia coli DNA replication fork. I. Multiple effectors act to modulate Okazaki fragment size. J. Biol. Chem. 267, 4030–4044.

Wu, X., Shell, S.M., Zou, Y., 2005. Interaction and colocalization of Rad9/Rad1/Hus1 checkpoint complex with replication protein A in human cells. Oncogene 24, 4728–4735. doi:10.1038/sj.onc.1208674

Xia, S., Konigsberg, W.H., 2014. RB69 DNA polymerase structure, kinetics, and fidelity. Biochemistry (Mosc.) 53, 2752–2767. doi:10.1021/bi4014215

Xie, B., Li, H., Wang, Q., Xie, S., Rahmeh, A., Dai, W., Lee, M.Y.W.T., 2005. Further Characterization of Human DNA Polymerase δ Interacting Protein 38. J. Biol. Chem. 280, 22375–22384. doi:10.1074/jbc.M414597200

Xu, X., Vaithiyalingam, S., Glick, G.G., Mordes, D.A., Chazin, W.J., Cortez, D., 2008. The Basic Cleft of RPA70N Binds Multiple Checkpoint Proteins, Including RAD9, To Regulate ATR Signaling. Mol. Cell. Biol. 28, 7345–7353. doi:10.1128/MCB.01079-08

Yakovleva, L., Shuman, S., 2006. Nucleotide Misincorporation, 3′-Mismatch Extension, and Responses to Abasic Sites and DNA Adducts by the Polymerase Component of Bacterial DNA Ligase D. J. Biol. Chem. 281, 25026–25040. doi:10.1074/jbc.M603302200

Yamamoto, K.N., Hirota, K., Kono, K., Takeda, S., Sakamuru, S., Xia, M., Huang, R., Austin, C.P., Witt, K.L., Tice, R.R., 2011. Characterization of environmental chemicals with potential for DNA damage using isogenic DNA repair-deficient chicken DT40 cell lines. Environ. Mol. Mutagen. 52, 547–561. doi:10.1002/em.20656

Yan, X., Ma, L., Yi, D., Yoon, J., Diercks, A., Foltz, G., Price, N.D., Hood, L.E., Tian, Q., 2011. A CD133-related gene expression signature identifies an aggressive glioblastoma subtype with excessive mutations. Proc. Natl. Acad. Sci. 108, 1591–1596. doi:10.1073/pnas.1018696108

Yeeles, J.T.P., Deegan, T.D., Janska, A., Early, A., Diffley, J.F.X., 2015. Regulated Eukaryotic DNA Replication Origin Firing with Purified Proteins. Nature 519, 431–435. doi:10.1038/nature14285

Yeeles, J.T.P., Janska, A., Early, A., Diffley, J.F.X., 2017. How the Eukaryotic Replisome Achieves Rapid and Efficient DNA Replication. Mol. Cell 65, 105–116. doi:10.1016/j.molcel.2016.11.017

Yeeles, J.T.P., Marians, K.J., 2013. Dynamics of Leading-Strand Lesion Skipping by the Replisome. Mol. Cell 52, 855–865. doi:10.1016/j.molcel.2013.10.020

Yeeles, J.T.P., Marians, K.J., 2011. The Escherichia coli Replisome Is Inherently DNA Damage Tolerant. Science 334, 235–238. doi:10.1126/science.1209111

Yeeles, J.T.P., Poli, J., Marians, K.J., Pasero, P., 2013. Rescuing Stalled or Damaged Replication Forks. Cold Spring Harb. Perspect. Biol. 5, a012815. doi:10.1101/cshperspect.a012815

Ying, S., Chen, Z., Medhurst, A.L., Neal, J.A., Bao, Z., Mortusewicz, O., McGouran, J., Song, X., Shen, H., Hamdy, F.C., Kessler, B.M., Meek, K., Helleday, T., 2016. DNA-PKcs and PARP1 Bind to Unresected Stalled DNA Replication Forks Where

They Recruit XRCC1 to Mediate Repair. Cancer Res. 76, 1078–1088. doi:10.1158/0008-5472.CAN-15-0608

Yousefzadeh, M.J., Wyatt, D.W., Takata, K., Mu, Y., Hensley, S.C., Tomida, J., Bylund, G.O., Doublié, S., Johansson, E., Ramsden, D.A., McBride, K.M., Wood, R.D., 2014. Mechanism of Suppression of Chromosomal Instability by DNA Polymerase POLQ. PLoS Genet. 10. doi:10.1371/journal.pgen.1004654

Yüce, Ö., West, S.C., 2013. Senataxin, defective in the neurodegenerative disorder ataxia with oculomotor apraxia 2, lies at the interface of transcription and the DNA damage response. Mol. Cell. Biol. 33, 406–417. doi:10.1128/MCB.01195-12

Yuzhakov, A., Kelman, Z., Hurwitz, J., O'Donnell, M., 1999. Multiple competition reactions for RPA order the assembly of the DNA polymerase delta holoenzyme. EMBO J. 18, 6189–6199. doi:10.1093/emboj/18.21.6189

Zafar, M.K., Ketkar, A., Lodeiro, M.F., Cameron, C.E., Eoff, R.L., 2014. Kinetic analysis of human PrimPol DNA polymerase activity reveals a generally error-prone enzyme capable of accurately bypassing 7,8-dihydro-8-oxo-2'-deoxyguanosine. Biochemistry (Mosc.) 53, 6584–6594. doi:10.1021/bi501024u

Zahn, K.E., Wallace, S.S., Doublié, S., 2011. DNA polymerases provide a canon of strategies for translesion synthesis past oxidatively generated lesions. Curr. Opin. Struct. Biol. 21, 358–369. doi:10.1016/j.sbi.2011.03.008

Zamenhof, S., Brawerman, G., Chargaff, E., 1952. On the desoxypentose nucleic acids from several microorganisms. Biochim. Biophys. Acta 9, 402–405.

Zeman, M.K., Cimprich, K.A., 2014. Causes and consequences of replication stress. Nat. Cell Biol. 16, 2–9. doi:10.1038/ncb2897

Zhang, H., Lawrence, C.W., 2005. The error-free component of the RAD6/RAD18 DNA damage tolerance pathway of budding yeast employs sister-strand recombination. Proc. Natl. Acad. Sci. U. S. A. 102, 15954–15959. doi:10.1073/pnas.0504586102

Zhang, S., Grosse, F., 1990. Accuracy of DNA primase. J. Mol. Biol. 216, 475–479. doi:10.1016/0022-2836(90)90370-2

Zhang, Y., Yang, F., Kao, Y.-C., Kurilla, M.G., Pompliano, D.L., Dicker, I.B., 2002. Homogenous Assays for Escherichia coli DnaB-Stimulated DnaG Primase and DnaB Helicase and Their Use in Screening for Chemical Inhibitors. Anal. Biochem. 304, 174–179. doi:10.1006/abio.2002.5627

Zhao, L., Washington, M.T., 2017. Translesion Synthesis: Insights into the Selection and Switching of DNA Polymerases. Genes 8, 24. doi:10.3390/genes8010024

Zhong, X., Garg, P., Stith, C.M., McElhinny, S.A.N., Kissling, G.E., Burgers, P.M.J., Kunkel, T.A., 2006. The fidelity of DNA synthesis by yeast DNA polymerase zeta alone and with accessory proteins. Nucleic Acids Res. 34, 4731–4742. doi:10.1093/nar/gkl465

Zou, L., Elledge, S.J., 2003. Sensing DNA Damage Through ATRIP Recognition of RPA-ssDNA Complexes. Science 300, 1542–1548. doi:10.1126/science.1083430

Zuo, Z., Rodgers, C.J., Mikheikin, A.L., Trakselis, M.A., 2010. Characterization of a Functional DnaG-Type Primase in Archaea: Implications for a Dual-Primase System. J. Mol. Biol. 397, 664–676. doi:10.1016/j.jmb.2010.01.057