PhD degree in Systems Medicine (curriculum in Molecular Oncology)

European School of Molecular Medicine (SEMM),

University of Milan and University of Naples "Federico II"

Settore disciplinare: XXXIII

Understanding the role of Polθ in chromosomal DNA replication under stressful conditions

Anjali Mann

Istituto FIRC di oncologia molecolare (IFOM), Milan

University of Milan

Matricola n. R12078

Tutor: Prof. Vincenzo Costanzo

IFOM, Milan

University of Milan

PhD Coordinator: Prof. Saverio Minucci

Anno accademico 2020-2021

Acknowledgements

I take this opportunity to express my gratitude to everyone whose contribution has enabled me to present this thesis.

I am grateful to my PhD supervisor **Prof. Vincenzo Costanzo** for believing in my abilities and for providing me the opportunities to grow personally and professionally. I would like to him for his patience and constant support.

I would like to thank my internal and external advisors, **Dr. Ylli Doksani** and **Dr. Raphaël Ceccaldi** for their valuable feedback on my project during the work presentations.

I am especially thankful to **Anna De Antoni**, **Vincenzo Sannino** and **Lucia Falbo** for teaching me how to design and perform experiments, analyse the data and most importantly troubleshoot when they fail. Thank you for the support in both academic and non-academic matters! I want to thank all my lab mates, **Andrea, Miguel, Yodharaudshani, Lucia S., Christelle, Cristina, Catiana,** and **Federica** for the fun and laughter in the lab. I would also like to thank the past members of VC group, **Giorgio, Maria, Negar, Hervé, Francesco** and **Erica** for passing on their experience and knowledge to me. I thank the master's students **Dario** and **Jessica** for the enjoyable conversations in the lab. I wish to acknowledge my friends in Milan - **Sara, Iman, Amir, Dipanjan, Ishani, Giuseppe, Chaithra, Priscilla, Camilla** – for making these four years memorable for me.

I sincerely thank the in-house facilities at IFOM – the kitchen team, EM facility, Cogentec, Warehouse, Canteen for their continued efforts in enabling us to work more efficiently. I sincerely acknowledge the funding agency – AIRC for the financial support.

I am grateful to my best friend, **Krishnendu** for his constant support and encouragement. Thank you for the cheers on my wins and having my back in the lows.

I express my respect and sincere tribute to my teachers, particularly **Prof. Sharat Chandra** and **Prof. Vidyanand Nanjundiah** for inspiring me and mentoring me to do research. Thank you for nourishing my curiosity and guiding me in this path!

Last but not the least, I am certainly grateful to my parents, **Suman Devi** and **Dalbir Singh Mann** for being my role models in life. Thank you for being my constant in this forever changing world! I am indebted to you for your unconditional love and support, patience, kindness and selflessness. I am thankful to my little brother, **Deepanshu Mann**, who brings me joy and laughter and inspires me to become a better human being. I would also acknowledge my dearest grandparents, **Dhanno Devi** and **Sultan Singh Mann**, for their loving gentleness and for their hard work over the years which enabled me to reach this stage in my life. Thank you!

Contents

List of Abbreviations

APC: Anaphase promoting complex

APH: Aphidicolin

bp: base pair

BER: Base Excision Repair

CMG: Cdc45, Mcm2–7, GINS

DDR: DNA Damage Response

dNTP: deoxyribonucleotide triphosphate

DSB: Double Strand Break

dsDNA: Double stranded DNA

EM: Electron Microscopy

FL: Full length

HCG: Human Chorionic Gonadotropin

HR: Homologous Recombination

MMEJ: Microhomology mediated end-joining

ORC: Origin Recombination Complex

Ori: Origin of replication

PARP1: Poly-ADP-ribose polymerase 1

PCNA: Proliferating cell nuclear antigen

PMSG: Pregnant Mare Serum Gonadotropin

Pre-IC: Pre-initiation complex

Pre-RC: Pre-replication complex

RI: Replication intermediate

RF: Reverse fork

RPA: replication factor A

RPC: replication factor C

SDS: Sodium Dodecyl Sulfate

ssDNA: single stranded DNA

TCA: Trichloroacetic acid

TMP: 4,5',8-TriMethylPsoralen

v/v: Volume/volume

w/v: weight/volume

Figure Index

Abstract

DNA polymerase-theta (Polθ) is a class A family DNA polymerase comprising of a helicase-like domain on the N-terminus and a DNA polymerase domain on the C-terminus. Polθ overexpression in breast and ovarian cancer patients correlate with high tumor grade and poor response to chemotherapeutic drugs. Consistently, Polθ inhibition is synthetically lethal with inactivation of BRCA1/2 genes, which are often found mutated in these tumors. The ability of Polθ to sustain viability of BRCA1/2 defective cancer cells has been attributed to its role in alternative end-joining repair of double strand breaks (DSBs) resulting from collapsed replication forks. In addition to DSBs a major role in BRCA1/2 activity is the suppression of defects associated with faulty DNA replication. Indeed, the occurrence of extensive DNA replication defects ranging from single stranded DNA gap accumulation to nascent DNA degradation in the absence of BRCA1/2 and RAD51 has been previously demonstrated. However, the role of Polθ in counteracting defective DNA replication in the absence of functional BRCA1/2 and RAD51 is poorly understood. To address this question, we cloned and purified the full length and different domains of *Xenopus laevis* Polθ and generated antibody to study Polθ function in replicating *Xenopus* egg extracts. Our preliminary findings indicate that Polθ has replication dependent and independent functions. Significantly, although dispensable for normal chromosomal DNA replication, Polθ is strongly enriched at stalled forks upon replication stress conditions induced by aphidicolin. Using DNA electron microscopy, we discovered that Polθ overexpression suppresses ssDNA gaps at the replication fork and replication fork reversal triggered by aphidicolin-induced fork stalling. Therefore, our results suggest that Polθ not only repairs DSBs but also prevents the

occurrence of potentially harmful DNA replication intermediates. We are currently investigating Polθ function in relation to replicative defects arising in the absence of BRCA1/2 and RAD51. Better understanding of Polθ function at stalled forks will help to target breast and ovarian cancers more effectively.

Chapter 1. Introduction

1.1 DNA replication in eukaryotes

DNA replication is a fundamental process and is highly conserved across the species. All organisms in the three kingdoms of life classified by Ernst Haeckel undergo semiconservative mode of DNA replication[1] as originally hypothesized by Watson and Crick after the discovery of structure of DNA[2]. DNA replication is a process by which the cell duplicates its genome in S phase prior to cell division. To ensure stable transmission of genetic information from one generation to the next one, DNA replication is tightly regulated to allow only one duplication event per cell cycle to give rise to two daughter cells with equal number of chromosomes. Therefore, replication of the eukaryotic genome is spatially and temporally regulated by a serious of steps[3].

Genomic DNA replication comprises of three steps – DNA replication initiation, elongation and termination (Fig.1.1). During replication initiation, DNA double helix unwinds at specific sites called origin of replication (ori) which act as replication start sites where specialized multi-protein complex called replisome is assembled. In the elongation step bulk DNA is replicated by the specialized multi-protein complex called replisome. During the final step of DNA replication termination, after complete replication of the whole genome, replisomes are disassembled and daughter molecules are resolved[4].

While most DNA replication happens with high fidelity, errors do happen which lead to mutations. The obstacles presented to the replication machinery are both intrinsic and extrinsic in nature and challenge replication accuracy[5, 6], that often leads to genome stability which is one of the hallmarks of cancer[7].

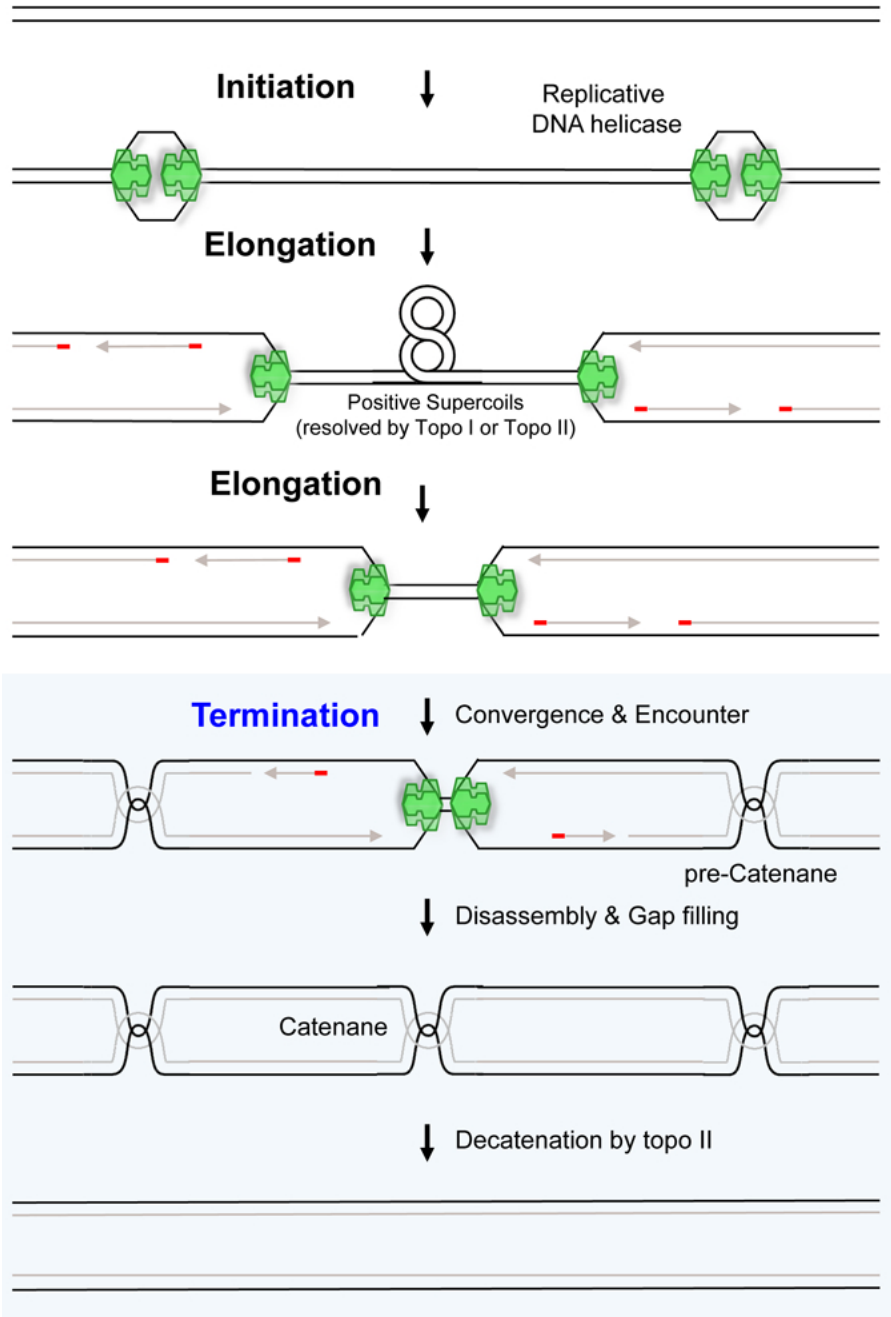


Figure1.1 Three key steps in genomic DNA replication. Graphical representation showing DNA replication initiation, elongation and termination. The termination step consists of fork convergence, replisome disassembly, gap filling, ligation and decatenation. Daughter strands are shown in grey lines and RNA primers are indicated in red. “Adapted from Dewar, J. M. et al., Nat Rev Mol Cell Biol, 2017”[4].

1.1.1 Replication origin licensing

Since the size of the eukaryotic genome is large, multiple replication origins (typically ~ 10^3 to 10^5) are utilized to ensure fast and complete duplication of the genome[8]. To ensure that the DNA is replicated only once, the cell has come up with means to differentiate between replicated and unreplicated DNA by controlling origin licensing and origin firing. The fact that replicated DNA differs from unreplicated DNA was experimentally first demonstrated by the classic cell fusion experiments by Rao, P.N. et. al., *Nature* 1970[9]. Replication origins establishment in eukaryotes is divided in two steps: first, origin licensing - the recognition of the pre replication assembly site, and second, origin firing - the activation of DNA synthesis. This dual check mechanism is essential to precisely prevent re-replication of DNA in one cell cycle[8, 10]. Replication origin licensing happens in G1 phase of the cell cycle and results from the sequential loading of the pre-replication complex (pre-RC) protein assembly on all the origins of the genome (Fig.1.2). The first step of origin licensing is the loading of hetero-hexameric Origin Recognition Complex (ORC) to potential replication origins. ORC, which has ATPase activity, is made up of six subunits, ORC1-6. ORC binding is followed by the binding of Cdc6 and Cdc10-dependent transcript 1 (Cdt1), which in turn triggers the loading of DNA helicases mini-chromosome maintenance (MCM), which is composed of six subunits MCM2-7. This is the last step of origin licensing. All these proteins are conserved from yeast to humans, suggesting that the basic features of origin licensing had been evolutionary conserved in all eukaryotes[11]. Although ORC is highly conserved from yeast to mammals, no specific motif for ORC binding has been identified yet except for some lower organisms such as *Saccharomyces cerevisiae*[12-15].

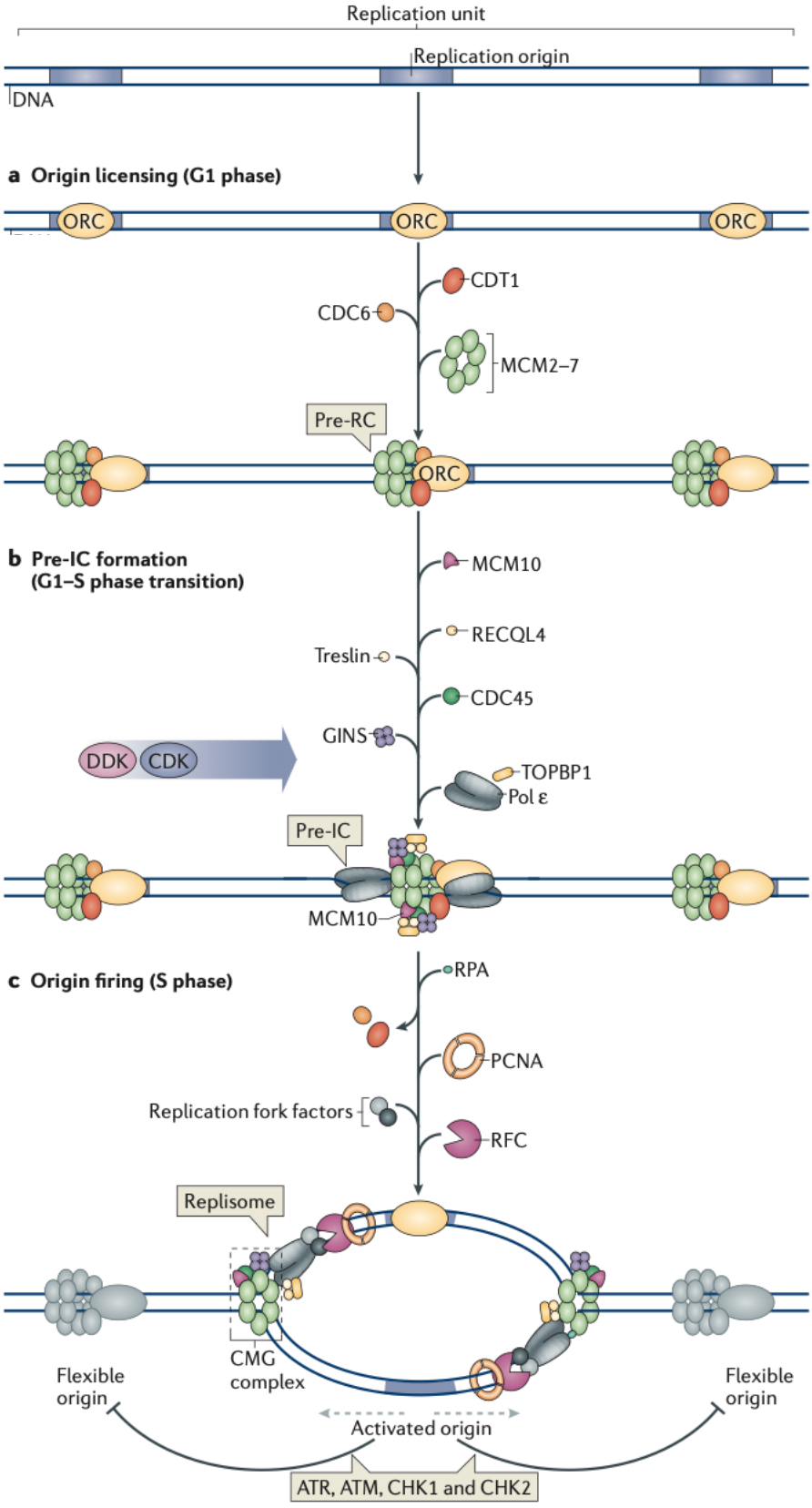


Figure1.2 DNA replication origin licensing and firing. Origin licensing occurs in G1 phase followed by pre-replication complex assembly in G1-S phase and origin firing to establish DNA synthesis in S phase. “Adapted from Fragkos, M. et al., Nat Rev Mol Cell Biol, 2015”[10].

1.1.2 Replication origin activation and firing

Origin activation happens at the G1-S phase transition and it involves the formation of pre-initiation complex (pre-IC) and activation of the MCM2-7 complex. G1-S phase transition is triggered by cyclin-dependent kinases (CDKs) and DBF4-dependent kinase (DDK) that causes a change in the structure of pre-RC and leads to the recruitment of additional replication factors (MCM10, RECQL4, Cdc45, Treslin, GINS, TOPBP1, Polε)[10, 11].

Cdc45 is an important factor in the transition from pre-RC to pre-IC because its association with MCM2-7 is a key step in DNA unwinding. Biochemical studies using *Xenopus* egg extracts have shown purified MCM2-7 free of Cdc45 does not have helicase activity whereas purified MCM2-7 from S-phase tightly bound to Cdc45 shows very high helicase activity[16, 17]. Helicase activation stimulates the recruitment of proliferating cell nuclear antigen (PCNA), replication factor A (RPA), replication factor C (RFC), DNA polymerase such as DNA Polymerase α and DNA Polymerase δ (Fig.1.2), this whole protein assembly on the chromatin is referred to as a replisome[18-20]. With this, each pre-IC is converted into two replication forks (with the replisome at each fork) that move in the opposite directions from the activated origin of replication (Fig.1.2) [21, 22].

However, to prevent re-replication of DNA in one cell cycle, further licensing must be blocked before cell enter S-phase. In metazoans, in S and G2 phase of cell cycle, re-replication is prevented by the interaction of geminin with Cdt1 which blocks the reloading of MCM2-7 complex on replication origins[23, 24]. Geminin is then degraded in the next G1 phase by Anaphase promoting complex (APC), to ensure new origin licensing[25, 26].

1.1.3 Replication fork progression

Upon replisome assembly, each pre-IC is converted into two replication forks (with the replisome at each fork) that move in the opposite directions from the activated origin of replication. Chromosomal replication takes place by the synthesis of nascent daughter DNA on both parental strand templates. Daughter strand synthesis in eukaryotes is performed by three DNA polymerases which are – DNA polymerase α, DNA polymerase δ and DNA polymerase ε[27, 28]. Upon unwinding of the parental strands by CMG helicase, DNA polymerase α-primase complex synthesizes a small RNA primer of 7-14 oxyribonucleotides and elongates it upto 20 deoxyribonucleotides (dNTPs) [29-32]. Subsequent elongation of these primers is carried out by DNA polymerase δ on the lagging strand and by DNA polymerase ε on the leading strand, in coordination with the replisome machinery[33-36]. Since the directionality of polymerization of DNA polymerase ε and δ is 5' to 3', 'leading strand' synthesis is carried out continuously and codirectional to the fork progression, while 'lagging strand' synthesis is carried out discontinuously and in the opposite direction of the fork progression (Fig.1.3). Lagging strand synthesis is divided into okazaki fragments of 100-300 nucleotides[37, 38]. Besides the 5' to 3' polymerization activity, polymerase δ also possesses 3' to 5' exonuclease activity to minimize any errors in DNA replication. Furthermore, Polymerase δ displaces okazaki fragments 5' RNA primer, which generates RNA-DNA single strand flaps which are removed by RNase H until the last few ribonucleotides and the remaining ribonucleotide flap is removed by 5' exonuclease activity of Fen1[39]. Subsequently Ligase I joins two okazaki fragment leading to an intact double strand [39, 40].

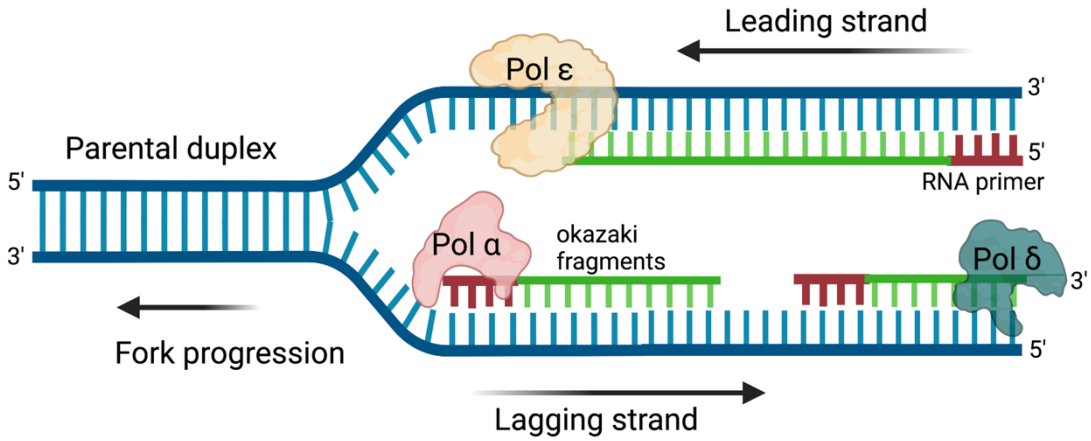


Figure1.3 DNA replication fork progression. A replication fork showing the leading and lagging strand synthesis in 5' to 3' direction. DNA synthesis on leading strand is continuous while it is discontinuous on the lagging strand. Polα-primase complex synthesizes a short RNA primer and extends ~20 deoxyribonucleotides by the polymerase activity of Polα. The primers are further extended by Polε on the leading strand and Polδ on the lagging strand[41]. Created with BioRender.com

1.1.4 Eukaryotic DNA polymerases

DNA polymerases are grouped into seven families (A, B, C, D, X, Y and RT) based on their sequence homology and structure analysis[42]. Replication in the nucleus of eukaryotic cells employs three DNA polymerases of the Family B – polymerase α, ε and δ[43–45]. DNA synthesis happens in the direction of 5' to 3' and all DNA polymerases require a free 3'-OH group to add new nucleotide[46]. All the family B polymerases are high fidelity polymerases, except Polα, and perform proofreading activity in 3' to 5' direction to correct any errors during DNA replication[47, 48]. All three eukaryotic DNA polymerases are multi-subunit enzymes as shown in Table1.1.

Polymerase	Species		Function
Polymerase α	***H. sapiens***	***S. cerevisiae***	
Catalytic or A-subunit	POLA1 (p180)	POL1	Catalytic subunit; polymerase activity; inactivated exonuclease
B-subunit	POLA2 (p70)	POL12	Regulatory subunit
Primase small subunit	PRIM1 (p49)	PRI1	Primase
Primase large subunit	PRIM2 (p58)	PRI2	Primase
Polymerase δ	***H. sapiens***	***S. cerevisiae***	
Catalytic or A-subunit	POLD1 (p125)	POL3	Catalytic subunit; has both polymerase and exonuclease activity
B-subunit	POLD2 (p50)	POL31	Accessory subunit
C-subunit	POLD3 (p66 or p68)	POL32	Accessory subunit
D-subunit	POLD4 (p12)	–	Accessory subunit
Polymerase ε	***H. sapiens***	***S. cerevisiae***	
Catalytic or A-subunit	POLE or POLE1	POL2	Catalytic subunit; has both polymerase and exonuclease activity
B-subunit	POLE2	DPB2	Accessory subunit
C-subunit	POLE3 (p17; CHRAC17)	DPB3	Accessory subunit
D-subunit	POLE4 (p12)	DPB4	Accessory subunit

Table1.1 Family B DNA polymerases and their subunits. All three eukaryotic DNA polymerases are multi-subunit enzymes. The table shows co-comparison of the subunits between *H. sapiens* and *S. cerevisiae* and the function of the individual subunits. "Adapted from Doublié, S. et al., *Front Microbiol*, 2014".

1.2 Family A polymerase – DNA Polymerase-theta

DNA polymerase-theta (Polθ) is a low fidelity class A family DNA polymerase encoded by *POLQ* gene [49, 50]. The N-terminal third of the protein is a helicase-like domain followed by three putative Rad51 binding motifs [51]. The C-terminal third of the protein is a DNA polymerase domain, which includes an exonuclease domain, though the functionality of the exonuclease domain is debatable [52, 53]. Adjoining the helicase-like and polymerase domain is a central region with no predicted function (Fig.1.4). Polθ has been implicated to play a role in the base excision repair pathway. It is a crucial player in MMEJ repair pathway, an alternate route that homologous-recombination (HR) deficient cells appear to use to defend against DNA damaging agents [51, 54]. Biochemical studies imply that HR and MMEJ pathways share the same substrate: resected DSBs with short single-stranded DNA (ssDNA) overhangs bound by RPA (Fig.1.6) [55, 56] . The helicase-like domain of Polθ removes RPA to promote annealing of ssDNA with microhomology [57]. The polymerase domain then likely fills ups the gaps left behind after ssDNA annealing to complete the MMEJ pathway (Fig.1.5). The role of Polθ has been well established in MMEJ, however it's function in DNA replication stress conditions has been poorly understood. DNA replication is a tightly controlled mechanism to ensure faithful duplication of the genome before cell division, and any aberration in it leads to replication stress[5, 58]. It has been proposed that HR deficient tumours are dependent on Polθ, but evidence has been pilling in the literature about the role of HR proteins in DNA replication beside DNA damage repair. It is therefore plausible to investigate the hypothesis that BRCA1/2 deficient tumours might be dependent on Polθ to deal with DNA replication stress.

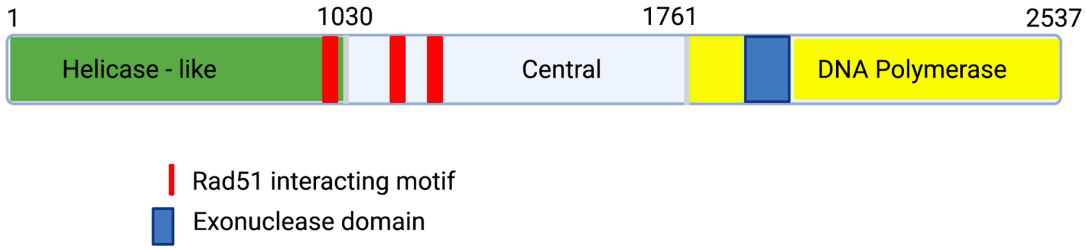


Figure1.4 Schematic representation of *Xenopus laevis* Polθ domain structure. The N-terminal region containing the helicase-like domain is followed by a central spacer region which harbors three putative Rad51 binding motifs, here shown in red. In the C-terminal region, there is a polymerase domain[51]. The DNA polymerase domain contains an inactive 3' to 5' exonuclease domain. Numbers above the protein structure indicate amino acid position. Created with BioRender.com

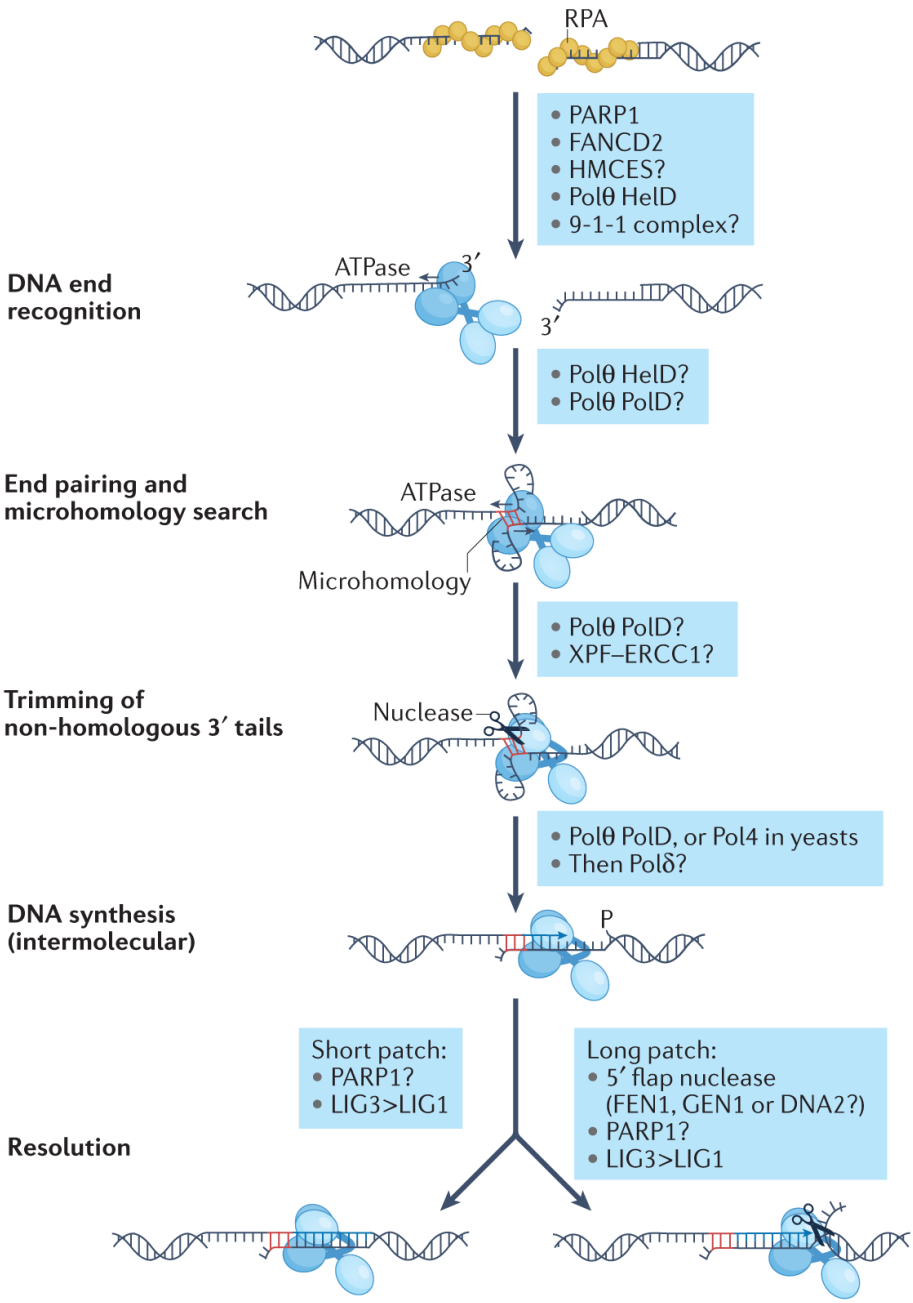


Figure1.5 Molecular mechanism of Polθ based microhomology mediated end-joining. Polθ is shown as a dimer, it utilizes resected dsDNA with 3' ssDNA overhang as a substrate bound by RPA which is removed by Polθ-helicase domain. Other factors such as HMCES, FANCD2, PARP1. May play a role in the recruitment of Polθ at DSBs. In the subsequent steps, microhomology based annealing takes place by Polθ, which typically leaves a flap which must be removed by exonucleases such as FEN1 or DNA2 followed by ligation by LIG1/3. "Adapted from Ramsden, D.A. et al., *Nat Rev Mol Cell Biol* 2021"[59].

1.3 Synthetic lethality between *POLQ* and *BRCA*

Genome instability is one of the hallmarks of cancer [7]. Mostly, the sources of genome instability arise from errors in DNA replication and repair. Loss of DNA repair is one of the earliest steps in tumorigenesis and is found in 40 to 50% of the cancers today [60]. Therefore, targeting DNA repair deficiencies has become an effective strategy to treat cancer. However, DNA repair deficient tumours often become highly dependent on alternative back up repair pathways like microhomology mediated end joining (MMEJ) (Fig.1.6) [51]. MMEJ is an error-prone double stand break repair (DSB) pathway and DNA polymerase-theta (Polθ) is the main protein involved in it. Studies from Alan D. D'Andrea and Agnel Sfeir demonstrated synthetic lethal relationship between Polθ-mediated repair and homologous recombination (HR) pathway, in particular BRCA genes [51, 61]. *POLQ* has also been shown to be synthetic lethal with BRCA1/2 loss of function, using BRCA2 synthetic lethal screening-based experiments[62]. Hence, identifying Polθ as a new druggable target for cancer therapy in tumours carrying mutations in HR pathway genes.

PARP inhibitors were the first clinically approved drugs exploiting the concept of synthetic lethality[63]. Poly-ADP-ribose polymerase 1 (PARP1) is an enzyme involved in DNA damage response[64]. Chemical inhibition of PARP1 causes unresolved damage which ends up in DSB formation[65]. This results in synthetic lethality for BRCA1/2 mutated cells that cannot efficiently repair DSB[66]. This approach was promising, but acquired resistance mechanisms have been observed lately. Among these are secondary BRCA mutations, PARP1 overexpression and elevated expression of P-glycoprotein efflux pumps, which enhance the intra-to-extracellular translocation of small molecules[67]. Optimism is growing that targeting Polθ will not only synergize

for PARP inhibitors but will have greater utility in cancer treatment[60]. One of the reasons why Polθ has become a major focus for cancer therapy is because its expression is largely absent in normal cells whereas it is highly increased in many cancers, both HR proficient and deficient tumors[68-73]. The reasons why Polθ overexpression is correlated with poor outcomes of the tumor are not very well understood but one of the possibilities could be that Polθ repairs spontaneous DNA damage.

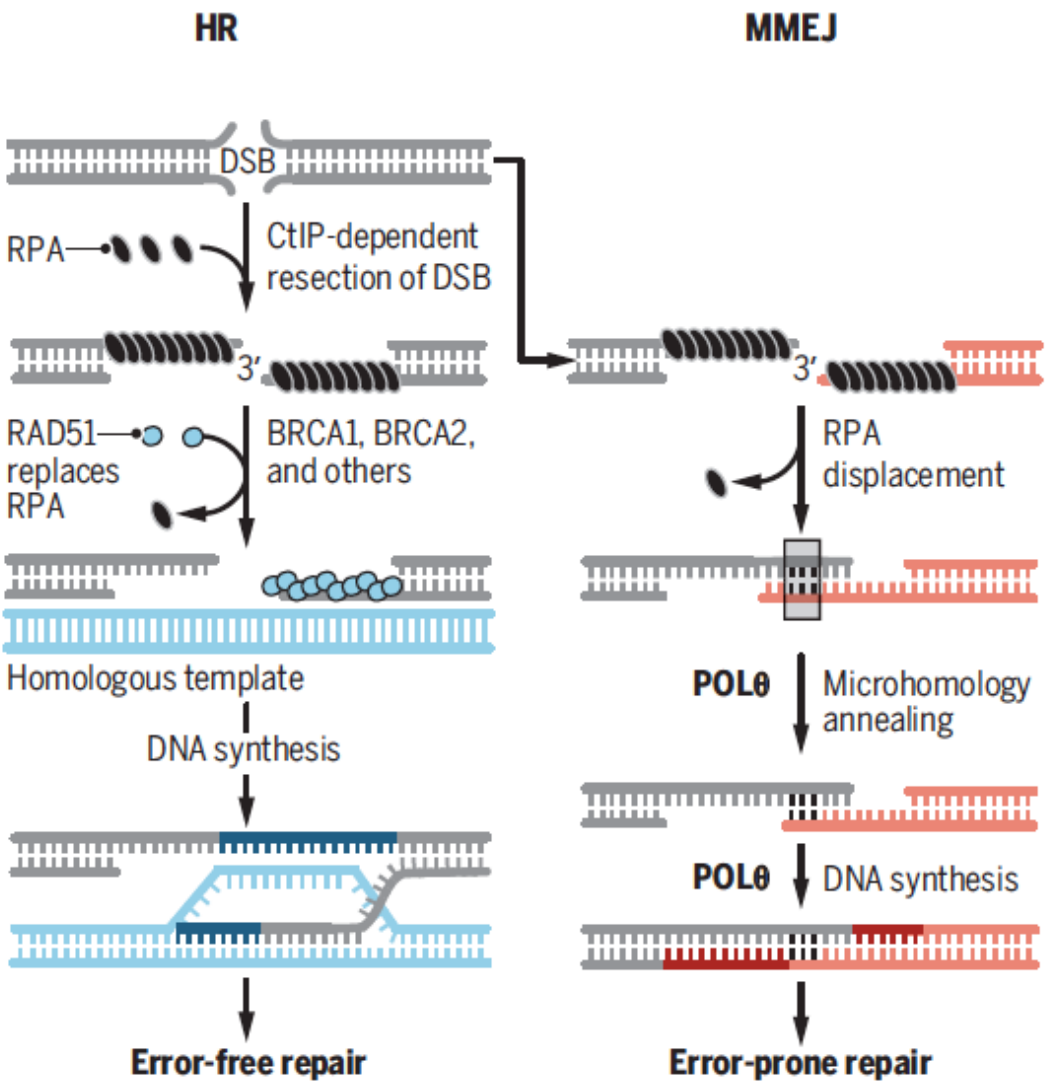


Figure1.6 Double strand break repair by HR and MMEJ. DSBs are predominantly repaired by HR in the S and G2 phase of the cell cycle because the homologous template is present. MMEJ is a "backup" repair pathway that HR deficient cells employ for DSB repair. Both HR and MMEJ share the same substrate, i.e., a resected double strand break with a 3' overhang bound by ssDNA binding protein RPA. HR is an error-free repair pathway whereas MMEJ is an error-prone repair pathway. "Adapted from Higgins, G.S. et. al., *Science* 2018[60]".

1.4 DNA replication stress and genome stability

In eukaryotes, DNA replication originates at multiple origin of replication and at each origin two replication forks are formed which run in two opposite directions. Replication origins are licensed prior to S-phase (for more detail see section 1.1.1). However not all replication origins which are licensed fire in an unperturbed replication S-phase. These 'dormant origins' can be activated to fire upon replication stress to ensure the completion of a faithful round of DNA replication at stalled replication forks[74-76].

Replication stress is defined as the slowing down or stalling of replication fork progression and DNA synthesis and this leads to an accumulation of persistent ssDNA at the fork[5]. These ssDNA gaps are formed mainly by the uncoupling of CMG helicase and DNA polymerase. Replication stress can be generated by a variety of sources such as: limiting nucleotides, DNA lesions, ribonucleotide incorporation, repetitive DNA, fragile sites, oncogene-induced stress, DNA secondary structures, ongoing transcription and RNA: DNA hybrids. Persistent ssDNA gaps bound by RPA adjacent to a stalled newly synthesized double stranded DNA, activates replication stress response (Fig.1.7) [77]. This leads to the recruitment of replication stress response proteins such as Ataxia-telangiectasia-mutated (ATM) and ataxia telangiectasia and Rad3-related (ATR), which in turn phosphorylates RPA (Ser33) or Chk1 (Ser345) to inhibit cell cycle progression and suppress late origin firing to maintain genome stability (Fig.1.8). In addition, ATR also helps to stabilize and restart the stalled forks. Fork collapse can also lead to formation of DNA double strand breaks at the stalled fork or it can be remodeled into a reversed fork by chromatin remodelers. Though it is not yet clear whether reverse fork formation is protective or pathological for the cell.

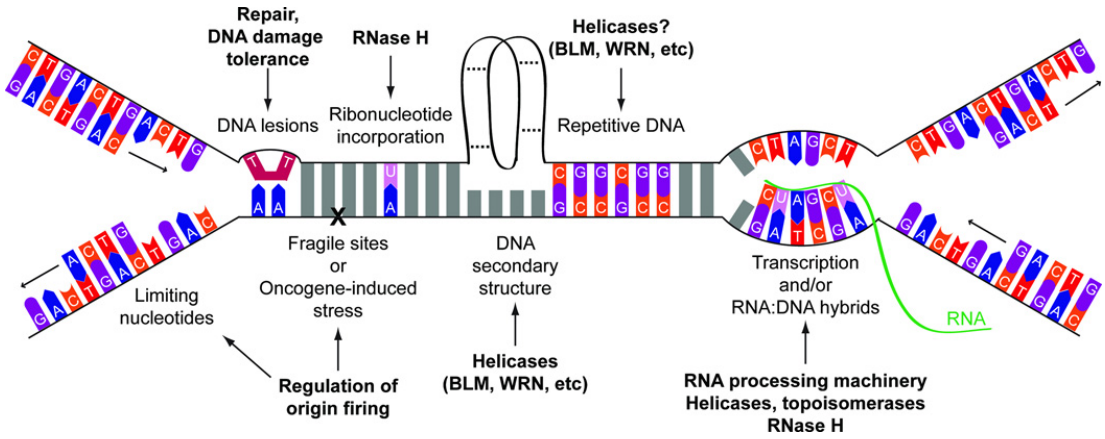


Figure1.7 Causes of replication stress and genomic instability. There are numerous intrinsic factors which pose a threat to the ongoing replication machinery and overall affecting genome stability. Some of the causes of DNA replication are limiting nucleotides, DNA lesions, ribonucleotide incorporation, repetitive DNA, fragile sites, oncogene-induced stress, DNA secondary structures, ongoing transcription and RNA:DNA hybrid. Some of the known pathways to resolve replication stress for each source of stress are indicated above in bold, marked with an arrow[5]. “Adapted from Zeman, M.K. et al., *Nat Cell Biol* 2014”.

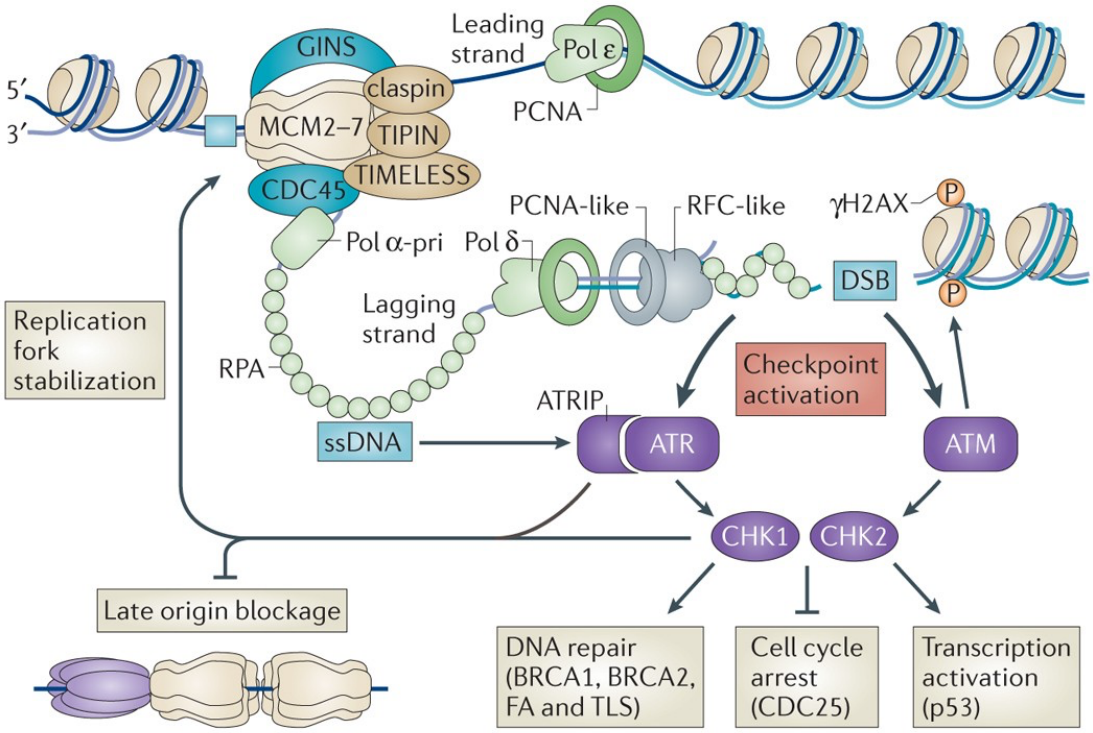


Figure1.8 Replication impairment and S-phase checkpoints. Upon unwinding of the parental strands by CMG helicase, DNA polymerase α-primase complex synthesizes a small RNA primer. Subsequent elongation of these primers is carried out by DNA polymerase δ on the lagging strand and by DNA polymerase ε on the leading strand, in coordination with the replisome machinery. Encountering an obstacle may lead to fork stalling or DSBs which triggers phosphorylation of CHK1 and CHK2 by ATM and ATR[58]. "Adapted from Gaillard, H. et al., *Nat Rev Cancer* 2015".

1.5 Replication fork reversal as an intermediate for DNA damage response

Ongoing DNA replication forks encounter several intrinsic and extrinsic DNA lesions which pose as obstacles for the replication machinery. Reverse fork formation is a key protective mechanism which allows the forks to reverse their course to promote DNA damage tolerance and prevent chromosomal breakage[78]. Reverse forks are formed by the coordinated annealing of the newly synthesized daughter strands into a four-way junction structure like a Holliday junction which can be restarted later (Fig.1.9)[79].

The first experimental evidence for reverse fork was reported in *E.coli* hyperrecombination regions where replication forks terminate at tus-ter protein-DNA complex[80]. This observation led to the hypothesis that reverse forks formation occur at termination sites due to their involvement in recombination events. For a long time, reverse forks were observed only in prokaryotes and certain yeast mutants, however recent visualization of replication intermediates in metazoans using electron microscopy shed light on the global presence of reverse forks[81].

When replication fork encounters a lesion and stalls, uncoupling of CMG helicase and DNA polymerase leads to an accumulation of ssDNA gaps behind the fork. ssDNA at the fork quickly gets coated by ssDNA binding protein RPA which in turn promotes the recruitment of E2-E3 ubiquitin conjugating enzymes which ubiquitinates PCNA. Mono or polyubiquitinated PCNA regulates the pathway choice between error-prone translesion DNA synthesis and error-free template switching mechanisms respectively[82, 83]. E2-E3 ubiquitin conjugating enzymes monoubiquitinates PCNA at lysine 164

(K164), as well as UBC13-dependednt polyubiquitination at lysine 164. Polyubiquitinated PCNA then interacts with chromatin remodeler ZRANB3 which promotes replication fork reversal (Fig.1.9)[79]. Persistent ssDNA gaps can also be remodeled into a reverse fork by Smarcal1[84, 85]. Different chromatin remodelers such Smarcal1, ZRANB3, HLTF recognize different fork structure to remodel into a reverse fork[86-88]. However, depletion of Smarcal1 or ZRANB3 does not fully abrogate fork reversal suggesting that fork reversal is not mediated by one fork remodeler and that different fork structure might require different chromatin remodelers or more than one fork remodeler[84, 89]. The central recombinase factor Rad51 which was previously thought to be limited to its function in homologous recombination, has been shown to be involved in replication fork reversal[90]. HR factors, in particular Rad51, BRCA1/2, are key regulators in maintaining the tight link between replication fork remodeling and degradation[91, 92].

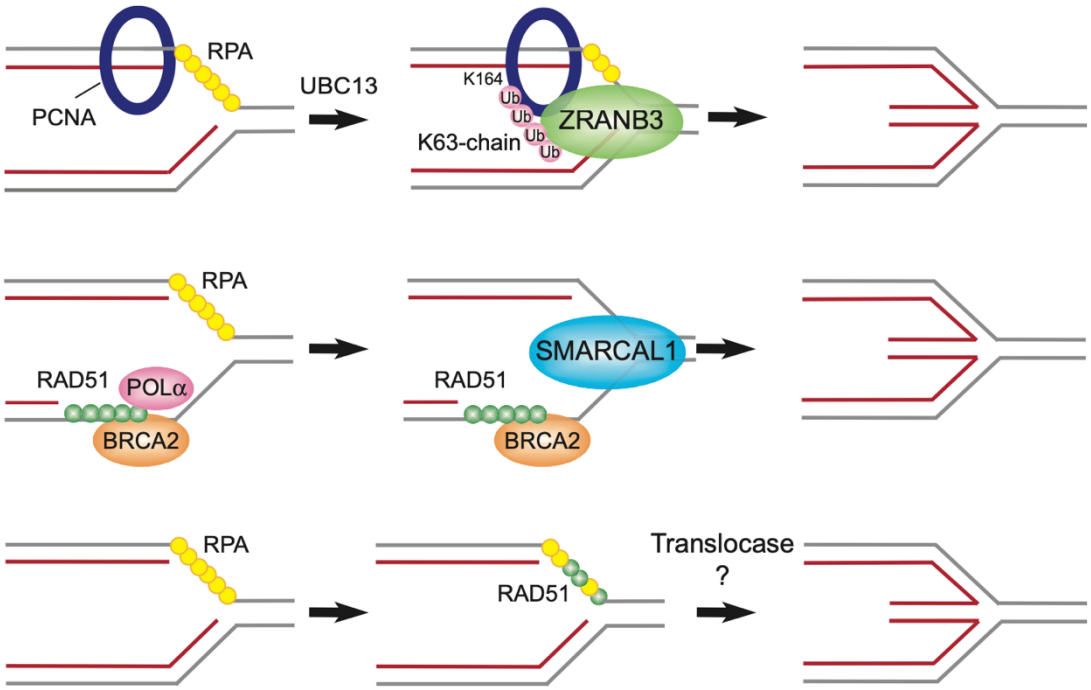


Figure1.9 Model for mechanisms of reverse fork formation. Uncoupling of CMG helicase and polymerase leads to ssDNA gaps behind the fork. Accumulation of ssDNA at the fork coated by RPA promotes the recruitment of E2-E3 ubiquitin conjugating enzymes which ubiquitinates PCNA, that interacts with ZRANB3 promoting fork reversal. Persistent ssDNA gaps can also be remodeled into a. reverse fork by Smarcal1[79]. "Adapted from Quinet, A. et al., *Mol Cell* 2017".

1.6 Role of homologous recombination factors in reverse fork protection

Homologous Recombination factors, in particular Rad51, BRCA1/2, are key regulators in maintaining the tight link between replication fork remodeling and degradation[91, 92]. The regressed arm of the reverse fork look like a one-ended double strand break which must be protected from nuclease cleavage. In normal cells, BRCA2 promotes Rad51 binding to replicating DNA and stabilizes Rad51 nucleofilament on the regressed arm of the reverse fork and thereby protecting it from DNA exonucleases CtIP, MRE11, EXO1 based resection. The RECQ1, DNA2 and WRN stimulate reverse fork restart.

Rad51 has two distinctive functions during DNA replication stress – a BRCA-independent function in promoting the first step of reverse fork formation, and a BRCA-dependent function in protecting the reverse fork degradation by forming stable Rad51 nucleofilament on the regressed arm to prevent nucleolytic degradation[84, 93, 94].

In BRCA2 deficient cells, regressed arm of the reverse arm is extensively degraded by CtIP, EXO1, MRE11[93]. MRE11 is recruited at the reverse fork by PTIP, MLL3/4 and RAD52[95]. Initial degradation of the regressed arm by MRE11 generates a reverse fork with 3' ssDNA flap which is a substrate of the endonuclease MUS81[79]. MUS81 cleaved product produces a migrating bubble which promotes DNA synthesis by POLD3[96] (Fig.1.10). Therefore, HR factors play a crucial role in protecting the reversed forks from nucleolytic cleavage thereby maintaining genome stability.

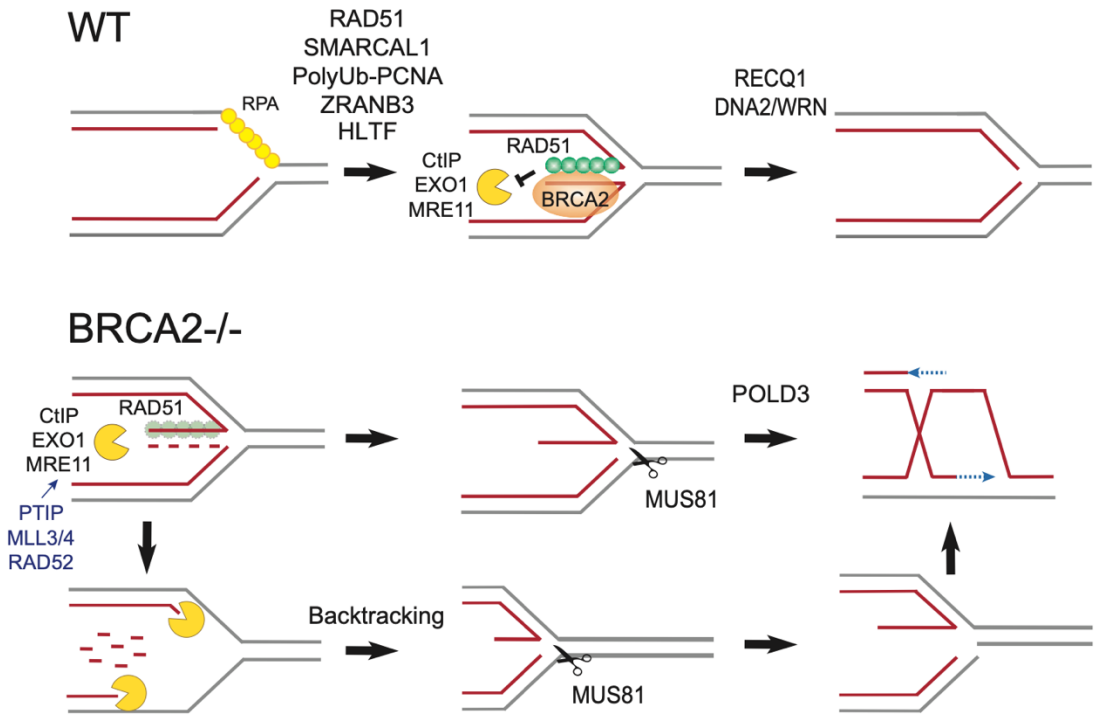


Figure1.10 Model for mechanisms of reverse fork protection and restart. In WT cells, BRCA2 stabilizes Rad51 nucleofilament on the regressed arm of the reverse fork thereby protecting it from CtIP, MRE11, EXO1 based resection. Then RECQ1, DNA2 and WRN stimulate reverse fork restart. In BRCA2 deficient cells, regressed arm of the reverse arm is extensively degraded by CtIP, EXO1, MRE11 which generates a reverse fork with 3' ssDNA arm which is a substrate of MUS81. MUS81 cleaved product produces a bubble which can be a template for DNA synthesis by POLD3 [79]"Adapted from Quinet, A. et al., *Mol Cell* 2017".

1.7 *Xenopus laevis* egg extract system as a model system to study DNA replication and DNA damage response

Cell free system based on vertebrate model system *Xenopus laevis* has been instrumental in elucidating biochemical basis of cell cycle check-point, DNA replication and repair [97]. Due to high levels of maternal proteins and mRNA molecules, *X. laevis* egg extract is able to recapitulate all cell cycle events such as nuclear assembly, semi-conservative DNA replication, chromosome condensation, spindle assembly and mitosis[98, 99]. Like other vertebrates, *Xenopus laevis* eggs are also arrested in Meiosis metaphase II, and the fertilized egg as well as egg extracts supplemented with sperm nuclei are able to initiate DNA replication and undergo 12 rounds of cell division until mid-blastula transition (MBT) without transcription[100]. Therefore, this system becomes a valuable tool to study protein complex and it allows to study DNA replication intermediates spatially and temporally independent of replication-transcription collision. Moreover, it offers an advantage to study essential proteins by depleting them from the protein-rich egg extract by immunodepletion, which in any other system would render them inviable. Owing to the ability of this embryonic system to replicate quickly, DNA replication dynamics can be monitored by chromatin binding assay of the different replication factors or by replication assays which is based on evaluating the incorporation of radiolabelled nucleotides. The use of cell free extracts can be combined with electron microscopy to visualize replication intermediates obtained from genomic DNA replicated in *Xenopus* extract[81, 92, 101]. In this study, we mainly used these approaches to study the synthetic lethality between *POLQ* and *BRCA* by investigating the role of Polθ in chromosomal DNA replication under stressful conditions. In particular, we

investigated the involvement of Polθ at the ssDNA gaps and reverse forks by biochemical analyses of replicated chromatin in *Xenopus* extract and visualization of replication intermediates by electron microscopy.

Chapter 2. Materials and Methods

2.1 Preparation of interphase *Xenopus laevis* egg extract

Xenopus laevis interphase egg extract were prepared as previously described[102, 103]. Adult *Xenopus laevis* females were injected with 250U human chorionic gonadotropin (HCG) in the morning and a second injection of 600U HCG six hours after the first injection. Each frog is placed in a single tank containing 100 mM NaCl for 12-16 hrs in an incubator at 20 °C. Next day morning, eggs are collected and freshly used for egg extract preparation. *Xenopus laevis* eggs are laid naturally arrested in metaphase II. The eggs were de-jellied in dejellying buffer (10 mM Tris pH8.0, 110 mM NaCl and 5 mM DTT) and rinsed three times in 50ml MMR (5 mM K-HEPES pH7.5, 100 mM NaCl, 0.5 mM KCl, 0.25 mM $MgSO_4$, 0.5 mM $CaCl_2$, 25 µM EDTA). De-jellied eggs were released in interphase in presence of 5 µM Calcium Ionophore (Sigma A23187) for 5-10 mins, when animal pole becomes smaller indicating egg activation, they were washed three times with MMR and rinsed twice with 20 ml ice cold S-buffer (50 mM K-HEPES pH 7.5, 50 mM KCl, 2.5 mM $MgCl_2$, 250 mM sucrose, 2 mM β-mercaptoethanol). Activated eggs were then compacked by centrifugation at 6000 x rpm for 8 secs and the excess of buffer was discarded. Eggs were then crushed at 13200 x rpm for 10 mins at 4 °C. The cytoplasmic fraction was collected and supplemented with 40 µg/mL cytochalasin B and gently mixed by inverting the tubes and centrifuged at 70.000 x rpm for 18 mins at 4 °C in a TLA100.3 rotor (Beckman 349622). The interphase extract was obtained by collecting and gently mixing the cleared cytoplasmic fraction plus the nuclear membranes. Aliquots of *Xenopus interphase* egg extracts were snap frozen with 3% glycerol and stored liquid nitrogen for later use.

2.2 Preparation of demembranated sperm nuclei

Demembranated sperm chromatin was prepared from testes of male frogs primed with 50 U of Pregnant Mare Serum Gonadotropin (PMSG) 7 days before and with 300 U of Human Chorionic Gonadotropin (HCG) the day before the sperm preparation. Testes were rinsed three times in EB buffer (50 mM Hepes-KOH pH 7.5, 50 mM KCl, 5 mM $MgCl_2$, 5mM EGTA, 2 mM ß-mercaptoethanol) and finely chopped with a razor. The obtained material was homogenized in a homogenizer, filtered through 25 µM Nylon membrane and centrifuged for 5min in a swinging-bucket rotor (JS 13.1, Beckman) at 4250 x g at 4 °C. The pellet was resuspended in 1.5 ml of SuNaSp buffer (15 mM Hepes-KOH, pH 7.5, 250 mM Sucrose, 75mM NaCl, 0.5 mM Spermidine, 0.15 mM Spermine) at room temperature and 50 µl of 10 mg/ml lysolecithin was added. The samples were incubated at room temperature for 5 mins. After the incubation sperm demembranation was tested by mixing 1.5 µl of sample with 1.5 µl of Hoechst stain 33258 (1 µl/ml). Following demembranation of more than 95% sperm population, 10 ml of cold SuNaSp buffer supplemented with 3% BSA was added to 1.5 ml sample and centrifuged for 5 mins in a swinging-bucket rotor (JS 13.1, Beckman) at 4250 x g at 4 °C. Obtained pellet was resuspended in 1 ml of cold SuNaSp buffer supplemented with 0.3% BSA and centrifuged again for 5 mins at 4250 x g. The sperm pellet was resuspended in EB buffer supplemented with 30% glycerol. The sperm density was then counted to reach a final concentration of 200,000 nuclei/µl and aliquots were quickly frozen in liquid nitrogen and stored at -80 °C.

2.3 Chromatin binding assay

For chromatin binding assays 30 µl *Xenopus laevis* egg extract was incubated with 4000 sperm nuclei/ µl for required time points. At each defined time point

chromatin was isolated at the indicated time points. To do so, samples were diluted with 10 volumes of EB (100 mM KCl, 2.5 mM $MgCl_2$, and 50 mM HEPES–KOH pH 7.5) containing 0.25% NP-40 (PanReac AppliChem) and centrifuged through a 0.5 M sucrose cushion at 10000 x g at 4 °C for 5 mins in a swinging backet rotor (TLA 100.3, Beckman). The supernatant and the dense sucrose layer were carefully removed without disturbing the pellet. Pellets were washed once with 300 µl EB and centrifuged in a benchtop refrigerated microcentrifuge at maximum speed for 5 mins. Resulting pellets were resuspended in Laemmli loading buffer. The samples were then denatured and resolved on a gradient SDS-PAGE and checked by WB.

2.4 iPOND (isolation of proteins on nascent DNA)

iPOND was performed as adapted from Sirbu, B. M. et. al., *Nat Protoc*, 2012[104]. 100 µl extract was used for each sample and supplemented with CP, CPK, ChX. Sperm nuclei were then added to reach a final concentration of 4000 nuclei/µl. 30 mins after sperm nuclei addition 10 mins DNA labelling pulses were carried out supplementing the extracts with 40 µM Biotin-16-dUTP (Roche). After 10 mins labelling, samples were supplemented with 1.5 mM APH or DMSO as control. DNA replication reactions were stopped by diluting 100 µl reactions with 200 µl cold EB-EDTA buffer (50 mM HEPES-KOH pH 7.5, 100 mM KCl, 2.5 mM $MgCl_2$, 1 mM EDTA). Samples were homogenized by using a cut p1000 tip and overlaid on 600 µl EB-EDTA-Sucrose buffer (EB-EDTA buffer + 30% w/v sucrose). Nuclei were collected by centrifugation at 10000rpm at 4 °C for 10 mins in a swinging-bucket rotor (TLA 100.3, Beckman). The supernatant was carefully removed and the nuclear pellet resuspended in 400 µl EB-NP40 buffer (50 mM HEPES-KOH pH 7.5, 100 mM KCl, 2.5 mM $MgCl_2$, 0.25% NP40) to lysate nuclei. Samples were then

subjected twice to a 10 mins sonication step (30 s ON / 40 s OFF cycle and Max Power with a Bioruptor device, Diagenode). After sonication, 20 µl from each sample were kept apart (5% Input to be loaded as control for SDS-PAGE). Biotinylated DNA fragments were then pulled-down by incubation with 40 µl Dynabeads M-280 Streptavidin (Thermo Fisher, 11205D) for 30 min at 4 °C. Dynabeads M-280 Streptavidin + the pull-down fractions were then washed three times with 200 µl EB-EDTA buffer and eventually resuspended with 30 µl of 1X denaturing loading buffer. The entire volume was eventually loaded on gel for SDS-PAGE and WB analysis.

2.5 Replication assay

DNA replication in *Xenopus* egg extract was performed as previously described[105]. Briefly, sperm nuclei (4000 nuclei/µl) were added to 10 µl of S-phase egg extract. Samples were supplemented with 0.1µl of α-^{32}P-dATP (3000Ci/mmol) and incubated at 23 °C for different time points. Replication reaction was stopped with stop buffer (8 mM EDTA, 80 mM Tris pH 8.0, 1% w/v SDS), supplemented with 1 mg/ml Proteinase K and incubated at 37 °C for 2 hrs. Samples were frozen on dry ice and again thawed at 37 °C for 5 mins for a freeze and thaw cycle and then separated by electrophoresis through a 0.8% agarose gel ran at 90V for 90 minutes. The agarose gel was then fixed in 30% TCA for 20 min, dried and exposed for autoradiography. For quantification of DNA replication efficiency, the gel was exposed to a phosphoscreen (GE Healthcare) for 12 hrs or longer if needed. The radioactive signal was monitored by phosphoImager (Typhoon) and quantified by Fiji software.

2.6 Cloning

The cDNA sequence encoding *Xenopus laevis* Polθ was obtained by retrotranscribing the RNA derived from *Xenopus laevis* eggs. The RNA was extracted from de-jellied eggs using the Trizol reagent (Thermo Fisher). cDNA first strand synthesis was done using oligo(dT)20 and superscript II reverse transcriptase (Thermo Fisher). The full-length sequence of Polθ was amplified by PCR using Phusion High-Fidelity DNA polymerase (New England Biolabs) and Polθ-N1 forward (CTGTGGATTATTATTGAGCCCCCG) and Polθ-C2 reverse (AAACCCTCTGGCCTCCTACAAGTC) primers. The PCR product was cloned into pCR-BLUNT II – TOPO (Invitrogen) obtaining the ADA438 plasmid. The sequence encoding the Polθ full length, N-terminal Helicase domain and C-terminal polymerase domain were amplified by PCR using Polθ-for-aa1-XhoI and Polθ-rev-aa2590-stop-NheI, Polθ-for-aa1-XhoI and Polθ-rev-aa1030-stop-NheI, Polθ-for-aa1761-XhoI and Polθ-rev-aa2590-stop-NheI respectively and cloned in pBAC-6H-MBP-TEV and pFH1 vectors obtaining the plasmids ADA444 (pBAC-6H-MBP-Polθ-FL$^{1-2590}$ for the expression of 6H-MBP-Polθ-FL), ADA445 (pFHis- Polθ-Helicase$^{1-1030}$ for the expression of 6H-Polθ-Helicase), ADA446 (pFHis- Polθ-Polymerase$^{1761-2590}$ 6H-Polθ-Polymerase), ADA447 (pBAC-6H-MBP-Polθ-Helicase$^{1-1030}$ for the expression of 6H-MBP-Polθ-Helicase) and ADA448 (pBAC-6H-MBP-Polθ-Polymerase$^{1761-2590}$ for the expression of 6H-MBP-Polθ-FL-Polymerase). All the sequences were checked by DNA sequencing. Cloning experiments were done in collaboration with Dr. Anna De Antoni.

2.7 Oligo extension assay

DNA polymerase oligo extension assay was performed as described previously[85, 106] . Reaction mixture for DNA polymerase assay was prepared by mixing pre annealed TET labelled forward (5' GCGGCTGTCATAAG 3') and reverse (3' GCGCCGACAGTATTCCGCCAG 5') template (final concentration 100 nM), 200 µM dNTPs, 2x reaction buffer (40 mM Tris-HCl pH 8, 50 mM KCl, 20 mM $MgCl_2$, 2 mM DTT), H_2O in a reaction volume of 5µl. The reactions were then supplemented with 5µl of 100 nM recombinant 6H-MBP-Polθ (FL) or 100 nM 6H-Polθ polymerase domain. The samples were incubated at 37 °C for defined time points. The reaction was stopped by adding 10 µl of 2x Gel Loading Buffer II (Thermo Fisher AM8546G) containing 95% formamide and denatured by boiling at 95 °C for 5 mins. The products were electrophoresed on a denaturing 15% TBE-Urea Polyacrylamide Gel (Bio-Rad 3450086) in 1x TBE.

2.8 Protein overexpression and purification

6H-MBP-Polθ (FL), 6H-MBP-Polθ-Helicase and 6H-Polθ-Polymerase from *Xenopus* were expressed in High Five insect cells (Invitrogen, B85502) infected with the respective recombinant baculoviruses by Silvia Monzani of the Crystallography Unit of the European Institute of Oncology. Briefly, for the MBP-fusion proteins, the cell pellet from 1L cell culture was resuspended in lysis buffer (50 mM HEPES, 200 mM NaCl, 1 mM EDTA, 10% glycerol, 2 mM 2-merceptethanol) supplemented with protease inhibitor (Calbiochem), lysed by sonication and cleared by centrifugation. The supernatant was added to 0.75 ml prewashed Amylose resin (New England Biolabs) for 3 hrs at 4 °C. The resin was then washed twice with lysis buffer-1 and the bound proteins were eluted with 10 mM maltose (Sigma, M5885). Relevant fractions were

concentrated in 50 kDa molecular mass cut-off Amicon ultra centrifugal filters (Millipore). The protein was further purified by SEC on Superdex-200 column (GE Healthcare) pre-equilibrated in SEC buffer (50 mM HEPES, 200 mM NaCl, 10% glycerol). The eluted peak fractions were pooled, concentrated and stored at −80 °C in small aliquots. For the 6H-Polθ-Polymerase proteins, the cell pellet from 1L cell culture was resuspended in lysis buffer (50mM HEPES, 300 mM NaCl, 10% glycerol, 2 mM 2-merceptethanol, 5 mM Imidazole) supplemented with protease inhibitor (Calbiochem), lysed by sonication and cleared by centrifugation at 14000 x g for 1hr at 4 °C. The supernatant was added to 1.5 ml prewashed Talon beads (New England Biolabs) for 2 hrs at 4 °C. The beads were then washed twice with lysis buffer and the bound proteins were eluted with 200-300 mM imidazole in lysis buffer. 5 mM EDTA was added to the relevant fractions and were concentrated in 50 kDa molecular mass cut-off Amicon ultra centrifugal filters (Millipore). The protein was further purified by SEC on Superdex-200 column (GE Healthcare) pre-equilibrated in SEC buffer (50 mM HEPES, 150 mM NaCl, 10% glycerol - filtered with 0.2 µm filter). The eluted peak fractions were pooled, concentrated and stored at −80 °C in small aliquots. All the protein purification experiments were performed in collaboration with Dr. Anna De Antoni.

2.9 Antibodies

Rabbit polyclonal antibodies against *Xenopus laevis* Polθ were raised by BioGenes GmBH using the recombinant protein 6H-TEV-Polθ$^{1761\text{-}2590}$ containing the C-terminal region of the protein. Antibodies against Xenopus ATM, Polα Polδ, Polε, ORC1, MCM7, Cdc45, RPA70, Smarcal1, Rad51, PCNA, H2B, γH2AX have been previously described [84, 92, 107].

2.10 Western Blotting

Unless otherwise specified throughout this study 4 - 20% Bis – Tris Poly Acrylamide gels (Biorad) were used for SDS-PAG Electrophoresis. All gels were run at 120V to allow proper migration of different molecular weight proteins until the dye front reached the bottom of the gel. Proteins from gel were transferred on Nitrocellulose membrane 0.2 µM (Bio-Rad) by semi-dry method using the TransBlot Turbo apparatus (Bio-Rad). Transferred membrane was washed twice with TBST (1x Tris Buffered Saline supplemented with 0.1% Tween-20). The membranes were incubated for 1 hr in 25 ml with 5% w/v nonfat dry milk blocking solution or 5% BSA in TBST (for detection of phosphorylated proteins). After blocking, the membranes were incubated with specific antibodies at 4 °C overnight. Again, the membranes were rinsed thrice for 15 mins with TBST and incubated with an HRP conjugated secondary antibody (1:10,000) in blocking solution for 1 hr at room temperature. Afterwards signals were detected using ECL-Western blot detection reagents according to the manufacture's guidelines (GE Healthcare).

2.11 Immunoprecipitation of nuclear proteins

100 µl extract was used for each sample and supplemented with energy mix. Sperm nuclei were then added to reach a final concentration of 4000 nuclei/µl. 45 mins after sperm addition samples were supplemented with 1.5 mM APH or DMSO. After 45 mins, the reaction was stopped by adding 1:10 cold EB buffer supplemented with protease inhibitors. Samples were then centrifuged through a 0.5 M sucrose cushion at 10000 x g at 4 °C for 5 mins in a swinging bucket rotor. Pellets were washed once with EB and resuspended in Resuspension buffer and sonicated using Bioruptor with cooling system (Diagenode) (15sec on/30 off), highest intensity, 3 cycles. Samples were

treated with benzonase (25U/ µl) and incubated for 1hr at 4 °C. Samples were then mixed with Dynabeads-ProteinA (Thermo Fisher, 10002D) conjugated with affinity purified rabbit Anti-Polθ antibody/Anti-Smarcal1 antibody and incubated at 4 °C for 2-3 hrs in a wheel at slow motion. Beads used for immunoprecipitation were extensively washed and suspended in Laemmli loading buffer and resolved through SDS-PAGE and probed for antibodies of interest.

2.12 Immunodepletion

To immunodeplete Polθ from the *Xenopus laevis* egg extract, 40 µg of affinity purified Polθ antibody (Rabbit 29047) was incubated with 150 µl Protein A Dynabeads slurry (Thermo Fisher 10002D) overnight at 4 °C**.** 150 µl of Protein A Dynabeads conjugated to the antibody were divided into two Eppendorf tubes and unbound IgGs were removed by placing the dynabeads on a magnetic rack and washing with PBS buffer. Then 400 µl of extract was added to the antibody bound dynabeads and incubated on a rotating wheel at 4 ºC. Two rounds of depletion were performed for 60 mins and 45 mins. Mock depletion was performed in parallel using the same protocol with Dynabeads-Protein A conjugated with affinity purified random rabbit IgG. The resulting supernatant was used as immunodepleted extract for downstream experiments. A chromatin binding reaction was performed simultaneously to follow the efficiency of Polθ depletion by western blotting.

2.13 hTRIM assay

hTRIM21 is an E3 ubiquitin ligase which binds strongly to the constant region of antibodies and recruit's ubiquitin-proteosome system to antibody bound antigens leading to their degradation[108]. It's a technique for rapidly depleting

the endogenous proteins as opposed to canonical dynabeads immunodepletion which takes longer and mechanically stressful for the egg extract, thereby reducing its replication efficiency. To deplete Polθ from the *Xenopus laevis* egg extract, 20 ng of affinity purified Polθ antibody/ µl extract (Rabbit 29047) was incubated with 40 µg 6H-hTRIM/ µl extract at 23 °C for 30 minutes. The immunodepleted extract was used to perform downstream experiments. A chromatin binding reaction was performed simultaneously to follow the efficiency of Polθ depletion by western blotting.

2.14 Electron microscopy sample preparation and analysis of replication intermediates

The protocol for DNA electron microscopy analysis was adapted from Hashimoto et al., 2010[92]. Briefly, to visualize replication intermediates, sperm nuclei (4000 nuclei/µl) were incubated at 23 °C in 200 µl egg extract for 60 mins. After 60 mins, the samples were diluted with 400 µl of EB buffer, layered onto 600 µl EB-EDTA (EB buffer + 1 mM EDTA) + 30% (w/v) sucrose and centrifuged at 10000rpm for 10 mins at 4 °C in a swinging bucket rotor. Supernatant was removed and nuclear pellets were resuspended in 100µl (final volume) cold EB-EDTA and then transferred to a pre-chilled 96-well microplate and subjected to four 4,5′,8-Trimethylpsoralen (TMP) crosslinking cycles. For each cycle, 5 µl of TMP stock (200 µg/ml) was added to each 100 µl nuclei suspension at 10 µg/ml TMP final concentration and mixed with a cut tip. Samples were incubated for 5 mins in the dark on ice and irradiated with UV-A (366 nm) for 7 mins (Max power with the Strata linker UV2400). Nuclei suspensions were then recovered and the wells washed with 300 µl of EB-EDTA buffer. Samples were then supplemented with 0.1% (w/v) SDS to lyse nuclei and 250 µg/ml RNase A and incubated for 1 hr at 37 °C, then further

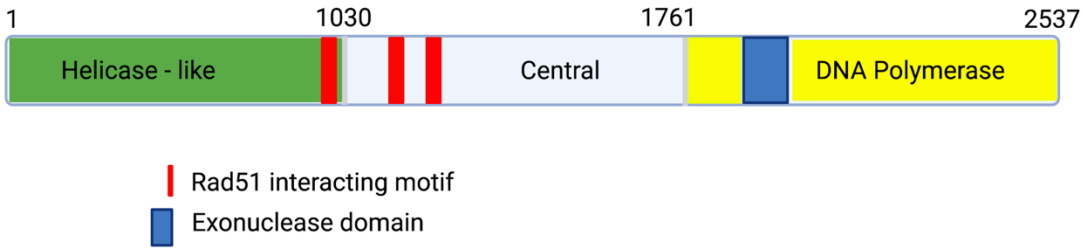
supplemented with proteinase K (1 mg/ml) and incubated for 2 hrs at 50 °C for complete protein digestion. Genomic DNA was extracted by adding one volume of 1:1 (v/v) chloroform-isoamylalcohol mixture, precipitated with 1 volume of isopropanol and 1:10 volume of 3M NaAc pH5.2, washed with prechilled 70% ethanol and resuspended overnight with 100 µl TE buffer. Samples were digested with the NdeI enzyme for 3 hrs at 37 °C in a final volume of 250 µl. Digested genomic preparations were then purified by means of Qiagen 20/G columns. Qiagen 20/G columns were equilibrated with 1ml of QBT buffer and then washed three times with 1 ml of 10 mM Tris pH 8.0, 1M NaCl and eventually equilibrated by washing three times with 1 ml of 10mM Tris pH 8.0, 300 mM NaCl. Each digested DNA mix was supplemented with 15 µl of 5M NaCl to bring the salt up to 300 mM NaCl and the final volume was adjusted to 1 ml by adding 10 mM Tris pH 8.0, 300 mM NaCl. DNA mix was loaded on the 20/G tip columns and let to flow by gravity. Columns were then washed 2 times with 1 ml of 10 mm Tris pH 8.0, 500 mM NaCl. DNA samples were eventually eluted with 2 x 600 µl of 10 mM Tris pH 8.0, 1M NaCl, 1.8% caffeine. DNA obtained from Qiagen 20/G column purification was then further cleaned and concentrated using Amicon 100k size-exclusion devices. Samples were washed twice with 200 µl TE buffer and then centrifuged for 10 mins at 10000 x g to minimally reduce sample volume. Concentrated DNA mix were recovered by spinning the Amicon devices upside down and 5 µl of each DNA sample were loaded on a 0.8% agarose gel to check for DNA concentration and quality. Purified DNA was later processed for electron microscopy. Electron microscopy grid shadowing was done with Leica MED20 and images were acquired using FEI Tecnai 20 EM microscope equipped with a GATAN high-resolution camera. All the electron microscopy experiments were performed in collaboration with Dr. Vincenzo Sannino.

Chapter3. Results

3.1 Protein sequence comparison between Human and *Xenopus laevis* Polθ

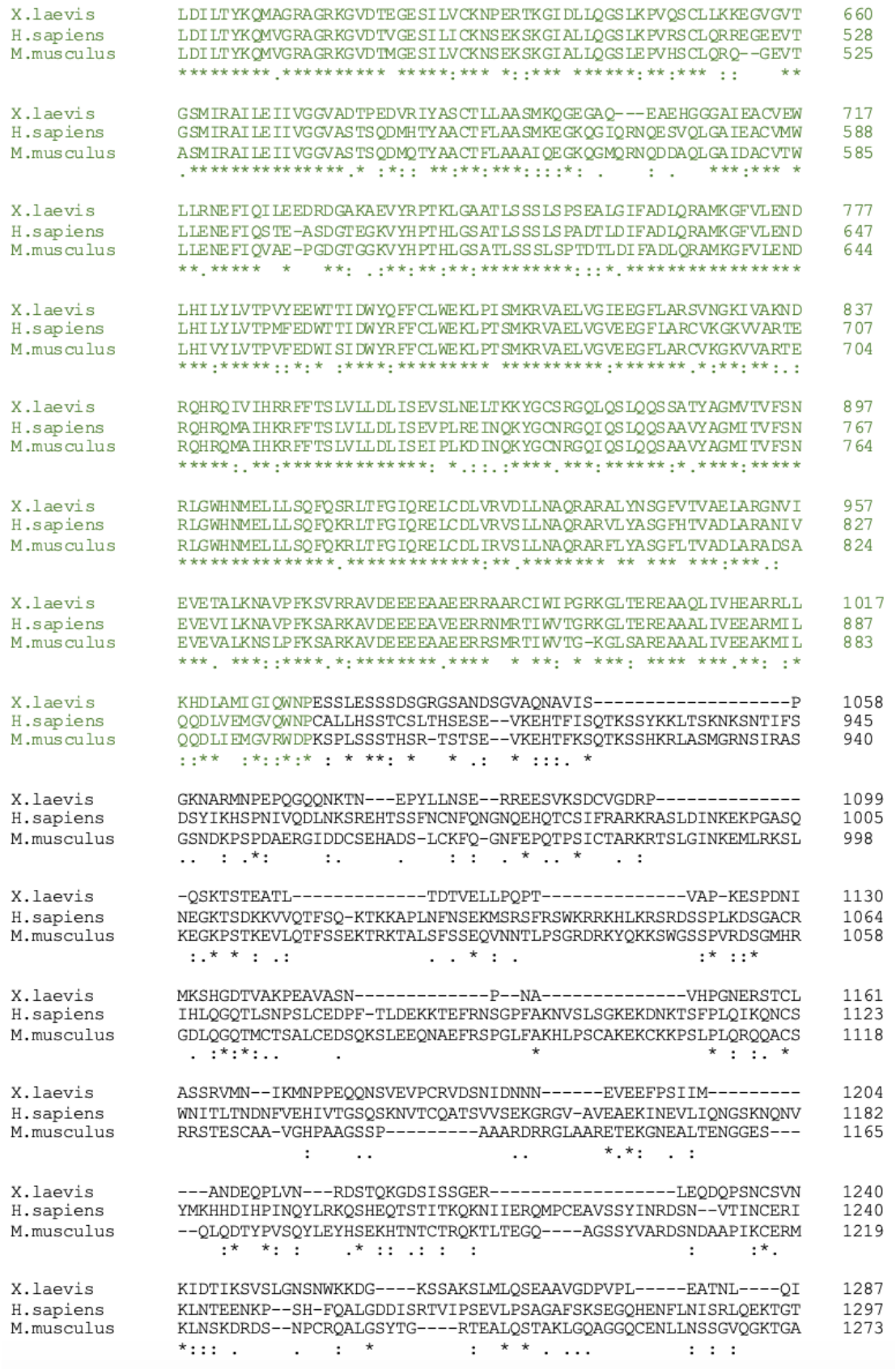
Proteins with high sequence identity posses functional similarity and are often evolutionary conserved[109, 110]. To understand the sequence conservation between *Xenopus laevis* Polθ and it's mammalian counterparts *Mus Musculus* and *Homo sapiens*, we retrived Polθ protein sequences from NCBI and subjected it to multiple sequence alignment tool on Clustal Omega program. *Xenopus laevis* Polθ shows 72% sequence identity with human Polθ, this percentage was calculated using the Basic Local Alignment Seach Tool (BLAST). In particluar, the Helicase-like and the polymerse domain of *Xenopus laevis* and *Homo sapiens* Polθ are higly identical (upto 85% sequence identity) and most of the sequence diversion is seen in the central spacer region whose functions are not yet well characterized. The human sequnece codes for a protein of 2590 amino acids, while that of Xenopus for a protein of 2541 amino acids. Fig.3.1A shows the tripartite domain structure of Xenopus laevis Polθ and Fig.3.1B shows the multiple sequence alignment between *Xenopus laevis*, *Homo sapiens* and *Mus Musculus* Polθ. The highlighted portion in green on the N-terminus is the Helicase-like domain, while the highlighted portion in yellow on the C-terminus is the polymerase domain.

A



B

```
CLUSTAL O(1.2.4) multiple sequence alignment

X.laevis        MQQIKKQPPLGGPSGSGASLSNILFGGLTNKENGRSSKGAPTVRGVLGPHNSRKRNRPTS    60
H.sapiens       ------------------------------------------------------------    0
M.musculus      ------------------------------------------------------------    0

X.laevis        SSSPGSEGSTSPSSPKRGPGGRQIRNRPRPQTSRTSQKRQRSESGHSTEAIHNASGHRDF    120
H.sapiens       ----------------------------MNLLRRSGKRRRSESGSDSFS----------    21
M.musculus      ----------------------------MSLPRRSRKRRRSSSGSDTFS----------    21
                                             .  * * **:**.** .: :

X.laevis        ILFSPAHQASILGKQKGQEPQVSRANLSVSVLTPPTGLERSLLDNSRFNDSVCASGLGPH    180
H.sapiens       -------------GSGGDSSASPQFLSGSVLSPPPGLGRCLKAAAA---------GECK    58
M.musculus      -------------GDG-DSFVSPQLRCGPVLSPPPGLGRGRRLTGT---------GTNK    57
                             .* :  .*    .  **:** ** *     .             :

X.laevis        MALPDALADKLMLASWGLPKTILDKYGSLGVTQMFEWQAECLMLGQVLEGRNLVYSAPTS    240
H.sapiens       PTVPDYERDKLLLANWGLPKAVLEKYHSFGVKKMFEWQAECLLLGQVLEGKNLVYSAPTS    118
M.musculus      RRVSDDQIDQLLLANWGLPKAVLEKYHSFGVRKMFEWQAECLLLGHVLEGKNLVYSAPTS    117
                  : *   *:*:**.*****::*:** *:** :*********:**:****:*********

X.laevis        AGKTLVAELLILKRVLETRRKALFILPFVSVAKEKTFYLQNLFQEVGVKVDGYMGSSSLA    300
H.sapiens       AGKTLVAELLILKRVLEMRKKALFILPFVSVAKEKKYYLQSLFQEVGIKVDGYMGSTSPS    178
M.musculus      AGKTLVAELLILKRVLETRKKALFILPFVSVAKEKKCYLQSLFQEVGLKVDGYMGSTSPT    177
                **************** *:***************. ***.******:********:* :

X.laevis        GGFSSLDVAVCTIEKANGLVNRLIEENKIELLGMMVVDELHMLGDSHRGYLLELLLTKVR    360
H.sapiens       RHFSSLDIAVCTIERANGLINRLIEENKMDLLGMVVVDELHMLGDSHRGYLLELLLTKIC    238
M.musculus      GQFSSLDIAVCTIERANGLVNRLIEENKMDLLGMVVVDELHMLGDSHRGYLLELLLTKIC    237
                  *****:******:****:*******::****:**********************:

X.laevis        YVTQKRAGGKPGAEVGGLGNDVQIVGMSATLPNLSLLATWLDAELYHTDFRPVPLLERVK    420
H.sapiens       YITRKSASCQA-DLASSLSNAVQIVGMSATLPNLELVASWLNAELYHTDFRPVPLLESVK    297
M.musculus      YVTRKSASHQA-ESASTLSNAVQIVGMSATLPNLQLVASWLNAELYHTDFRPVPLLESIK    296
                *:*:* *. :    .. *.* ************.*:*:**:*************** :*

X.laevis        IGKTVYDCNMVAVREFEPLLHVKGDDDHIVSLCYETVHGGHSILIFCPSKNWCEKLADTI    480
H.sapiens       VGNSIYDSSMKLVREFEPMLQVKGDEDHVVSLCYETICDNHSVLLFCPSKKWCEKLADII    357
M.musculus      IGNSIYDSSMKLVREFQPLLQVKGDEDHIVSLCYETIQDNHSVLIFCPSKKWCEKVADII    356
                :*:::**..*  ****:*:*:****:**:*******: ..**:*:*****:****:** *

X.laevis        AREFYNLFRRAEQQQQQQQGAAGGSDSSVSPVVLDRDGIRDVMDQLKRSPAGLDTVLGRT    540
H.sapiens       AREFYNLHHQAEGLV---------KPSECPPVILEQKELLEVMDQLRRLPSGLDSVLQKT    408
M.musculus      AREFYNLHHQPEGLV---------KSSEFPPVILDQKSLLEVMDQLKRSPSGLDSVLKNT    407
                *******.:: *            . *.  **:*::. : :*****:* *:***:** .*

X.laevis        VPWGVAFHHAGLTFDERDINEGAFRQGFVRVLAATSTLSSGVNLPARRVIIRSPLFNGRM    600
H.sapiens       VPWGVAFHHAGLTFEERDIIEGAFRQGLIRVLAATSTLSSGVNLPARRVIIRTPIFGGRP    468
M.musculus      VPWGVAFHHAGLTFEERDIIEGAFRQGFIRVLAATSTLSSGVNLPARRVIIRTPIFSGQP    467
                **************:**** *******::**********************:*:*.*:
```


X.laevis LDILTYKQMAGRAGRKGVDTEGESILVCKNPERTKGIDLLQGSLKPVQSCLLKKEGVGVT 660
H.sapiens LDILTYKQMVGRAGRKGVDTVGESILICKNSEKSKGIALLQGSLKPVRSCLQRREGEEVT 528
M.musculus LDILTYKQMVGRAGRKGVDTMGESILVCKNSEKSKGIALLQGSLEPVHSCLQRQ--GEVT 525
X.laevis GSMIRAILEIIVGGVADTPEDVRIYASCTLLAASMKQGEGAQ---EAEHGGGAIEACVEW 717
H.sapiens GSMIRAILEIIVGGVASTSQDMHTYAACTFLAASMKEGKQGIQRNQESVQLGAIEACVMW 588
M.musculus ASMIRAILEIIVGGVASTSQDMQTYAACTFLAAAIQEGKQGMQRNQDDAQLGAIDACVTW 585
X.laevis LLRNEFIQILEEDRDGAKAEVYRPTKLGAATLSSSLSPSEALGIFADLQRAMKGFVLEND 777
H.sapiens LLENEFIQSTE-ASDGTEGKVYHPTHLGSATLSSSLSPADTLDIFADLQRAMKGFVLEND 647
M.musculus LLENEFIQVAE-PGDGTGGKVYHPTHLGSATLSSSLSPTDTLDIFADLQRAMKGFVLEND 644
X.laevis LHILYLVTPVYEEWTTIDWYQFFCLWEKLPISMKRVAELVGIEEGFLARSVNGKIVAKND 837
H.sapiens LHILYLVTPMFEDWTTIDWYRFFCLWEKLPTSMKRVAELVGVEEGFLARCVKGKVVARTE 707
M.musculus LHIVYLVTPVFEDWISIDWYRFFCLWEKLPTSMKRVAELVGVEEGFLARCVKGKVVARTE 704
X.laevis RQHRQIVIHRRFFTSLVLLDLISEVSLNELTKKYGCSRGQLQSLQQSSATYAGMVTVFSN 897
H.sapiens RQHRQMAIHKRFFTSLVLLDLISEVPLREINQKYGCNRGQIQSLQQSAAVYAGMITVFSN 767
M.musculus RQHRQMAIHKRFFTSLVLLDLISEIPLKDINQKYGCNRGQIQSLQQSAAVYAGMITVFSN 764
X.laevis RLGWHNMELLLSQFQSRLTFGIQRELCDLVRVDLLNAQRARALYNSGFVTVAELARGNVI 957
H.sapiens RLGWHNMELLLSQFQKRLTFGIQRELCDLVRVSLLNAQRARVLYASGFHTVADLARANIV 827
M.musculus RLGWHNMELLLSQFQKRLTFGIQRELCDLIRVSLLNAQRARFLYASGFLTVADLARADSA 824
X.laevis EVETALKNAVPFKSVRRAVDEEEEAAEERRAARCIWIPGRKGLTEREAAQLIVHEARRLL 1017
H.sapiens EVEVILKNAVPFKSARKAVDEEEEAVEERRNMRTIWVTGRKGLTEREAAALIVEEARMIL 887
M.musculus EVEVALKNSLPFKSARKAVDEEEEAAEERRSMRTIWVTG-KGLSAREAAALIVEEAKMIL 883
X.laevis KHDLAMIGIQWNPESSLESSSDSGRGSANDSGVAQNAVIS----------P 1058
H.sapiens QQDLVEMGVQWNPCALLHSSTCSLTHSESE--VKEHTFISQTKSSYKKLTSKNKSNTIFS 945
M.musculus QQDLIEMGVRWDPKSPLSSSTHSR-TSTSE--VKEHTFKSQTKSSHKRLASMGRNSIRAS 940
X.laevis GKNARMNPEPQGQQNKTN---EPYLLNSE--RREESVKSDCVGDRP---------- 1099
H.sapiens DSYIKHSPNIVQDLNKSREHTSSFNCNFQNGNQEHQTCSIFRARKRASLDINKEKPGASQ 1005
M.musculus GSNDKPSPDAERGIDDCSEHADS-LCKFQ-GNFEPQTPSICTARKRTSLGINKEMLRKSL 998
X.laevis -QSKTSTEATL----------TDTVELLPQPT----------VAP-KESPDNI 1130
H.sapiens NEGKTSDKKVVQTFSQ-KTKKAPLNFNSEKMSRSFRSWKRRKHLKRSRDSSPLKDSGACR 1064
M.musculus KEGKPSTKEVLQTFSSEKTRKTALSFSSEQVNNTLPSGRDRKYQKKSWGSSPVRDSGMHR 1058
X.laevis MKSHGDTVAKPEAVASN----------P--NA----------VHPGNERSTCL 1161
H.sapiens IHLQGQTLSNPSLCEDPF-TLDEKKTEFRNSGPFAKNVSLSGKEKDNKTSFPLQIKQNCS 1123
M.musculus GDLQGQTMCTSALCEDSQKSLEEQNAEFRSPGLFAKHLPSCAKEKCKKPSLPLQRQQACS 1118
X.laevis ASSRVMN--IKMNPPEQQNSVEVPCRVDSNIDNNN------EVEEFPSIIM--------- 1204
H.sapiens WNITLTNDNFVEHIVTGSQSKNVTCQATSVVSEKGRGV-AVEAEKINEVLIQNGSKNQNV 1182
M.musculus RRSTESCAA-VGHPAAGSSP---------AAARDRRGLAARETEKGNEALTENGGES--- 1165
X.laevis ---ANDEQPLVN---RDSTQKGDSISSGER----------LEQDQPSNCSVN 1240
H.sapiens YMKHHDIHPINQYLRKQSHEQTSTITKQKNIIERQMPCEAVSSYINRDSN--VTINCERI 1240
M.musculus --QLQDTYPVSQYLEYHSEKHTNTCTRQKTLTEGQ----AGSSYVARDSNDAAPIKCERM 1219
X.laevis KIDTIKSVSLGNSNWKKDG----KSSAKSLMLQSEAAVGDPVPL-----EATNL----QI 1287
H.sapiens KLNTEENKP--SH-FQALGDDISRTVIPSEVLPSAGAFSKSEGQHENFLNISRLQEKTGT 1297
M.musculus KLNSKDRDS--NPCRQALGSYTG----RTEALQSTAKLGQAGGQCENLLNSSGVQGKTGA 1273

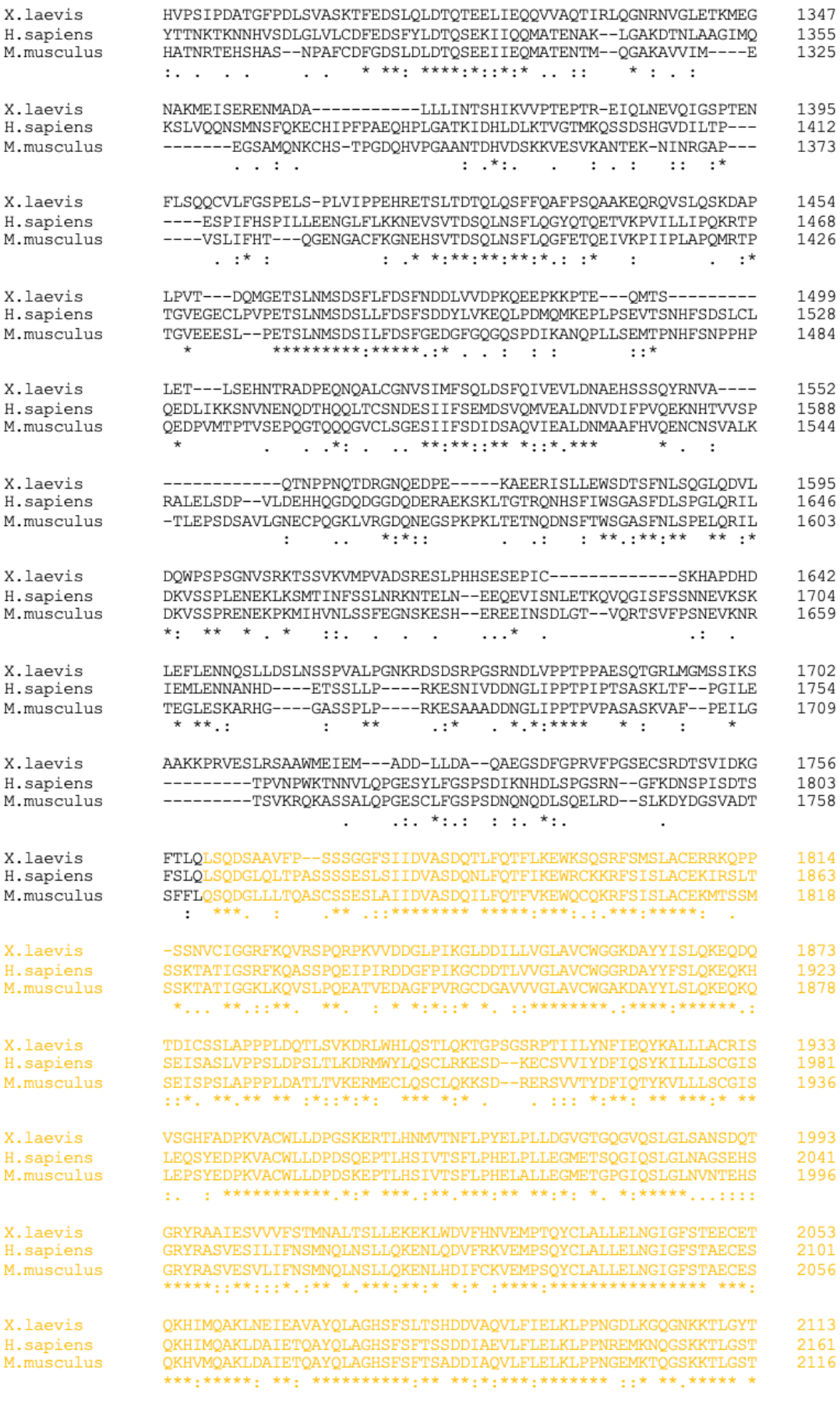

X.laevis HVPSIPDATGFPDLSVASKTFEDSLQLDTQTEELIEQQVVAQTIRLQGNRNVGLETKMEG 1347
H.sapiens YTTNKTKNNHVSDLGLVLCDFEDSFYLDTQSEKIIQQMATENAK--LGAKDTNLAAGIMQ 1355
M.musculus HATNRTEHSHAS--NPAFCDFGDSLDLDTQSEEIIEQMATENTM--QGAKAVVIM----E 1325
X.laevis NAKMEISERENMADA-----------LLLINTSHIKVVPTEPTR-EIQLNEVQIGSPTEN 1395
H.sapiens KSLVQQNSMNSFQKECHIPFPAEQHPLGATKIDHLDLKTVGTMKQSSDSHGVDILTP--- 1412
M.musculus -------EGSAMQNKCHS-TPGDQHVPGAANTDHVDSKKVESVKANTEK-NINRGAP--- 1373
X.laevis FLSQQCVLFGSPELS-PLVIPPEHRETSLTDTQLQSFFQAFPSQAAKEQRQVSLQSKDAP 1454
H.sapiens ----ESPIFHSPILLEENGLFLKKNEVSVTDSQLNSFLQGYQTQETVKPVILLIPQKRTP 1468
M.musculus ----VSLIFHT---QGENGACFKGNEHSVTDSQLNSFLQGFETQEIVKPIIPLAPQMRTP 1426
X.laevis LPVT---DQMGETSLNMSDSFLFDSFNDDLVVDPKQEEPKKPTE---QMTS--------- 1499
H.sapiens TGVEGECLPVPETSLNMSDSLLFDSFSDDYLVKEQLPDMQMKEPLPSEVTSNHFSDSLCL 1528
M.musculus TGVEEESL--PETSLNMSDSILFDSFGEDGFGQGQSPDIKANQPLLSEMTPNHFSNPPHP 1484
X.laevis LET---LSEHNTRADPEQNQALCGNVSIMFSQLDSFQIVEVLDNAEHSSSQYRNVA---- 1552
H.sapiens QEDLIKKSNVNENQDTHQQLTCSNDESIIFSEMDSVQMVEALDNVDIFPVQEKNHTVVSP 1588
M.musculus QEDPVMTPTVSEPQGTQQQGVCLSGESIIFSDIDSAQVIEALDNMAAFHVQENCNSVALK 1544
X.laevis ------------QTNPPNQTDRGNQEDPE-----KAEERISLLEWSDTSFNLSQGLQDVL 1595
H.sapiens RALELSDP--VLDEHHQGDQDGGDQDERAEKSKLTGTRQNHSFIWSGASFDLSPGLQRIL 1646
M.musculus -TLEPSDSAVLGNECPQGKLVRGDQNEGSPKPKLTETNQDNSFTWSGASFNLSPELQRIL 1603
X.laevis DQWPSPSGNVSRKTSSVKVMPVADSRESLPHHSESEPIC------------SKHAPDHD 1642
H.sapiens DKVSSPLENEKLKSMTINFSSLNRKNTELN--EEQEVISNLETKQVQGISFSSNNEVKSK 1704
M.musculus DKVSSPRENEKPKMIHVNLSSFEGNSKESH--EREEINSDLGT--VQRTSVFPSNEVKNR 1659
X.laevis LEFLENNQSLLDSLNSSPVALPGNKRDSDSRPGSRNDLVPPTPPAESQTGRLMGMSSIKS 1702
H.sapiens IEMLENNANHD----ETSSLLP----RKESNIVDDNGLIPPTPIPTSASKLTF--PGILE 1754
M.musculus TEGLESKARHG----GASSPLP----RKESAAADDNGLIPPTPVPASASKVAF--PEILG 1709
X.laevis AAKKPRVESLRSAAWMEIEM---ADD-LLDA--QAEGSDFGPRVFPGSECSRDTSVIDKG 1756
H.sapiens ---------TPVNPWKTNNVLQPGESYLFGSPSDIKNHDLSPGSRN--GFKDNSPISDTS 1803
M.musculus ---------TSVKRQKASSALQPGESCLFGSPSDNQNQDLSQELRD--SLKDYDGSVADT 1758
X.laevis FTLQLSQDSAAVFP--SSSGGFSIIDVASDQTLFQTFLKEWKSQSRFSMSLACERRKQPP 1814
H.sapiens FSLQLSQDGLQLTPASSSSESLSIIDVASDQNLFQTFIKEWRCKKRFSISLACEKIRSLT 1863
M.musculus SFFLQSQDGLLLTQASCSSESLAIIDVASDQILFQTFVKEWQCQKRFSISLACEKMTSSM 1818
X.laevis -SSNVCIGGRFKQVRSPQRPKVVDDGLPIKGLDDILLVGLAVCWGGKDAYYISLQKEQDQ 1873
H.sapiens SSKTATIGSRFKQASSPQEIPIRDDGFPIKGCDDTLVVGLAVCWGGRDAYYFSLQKEQKH 1923
M.musculus SSKTATIGGKLKQVSLPQEATVEDAGFPVRGCDGAVVVGLAVCWGAKDAYYLSLQKEQKQ 1878
X.laevis TDICSSLAPPPLDQTLSVKDRLWHLQSTLQKTGPSGSRPTIILYNFIEQYKALLLACRIS 1933
H.sapiens SEISASLVPPSLDPSLTLKDRMWYLQSCLRKESD--KECSVVIYDFIQSYKILLLSCGIS 1981
M.musculus SEISPSLAPPPLDATLTVKERMECLQSCLQKKSD--RERSVVTYDFIQTYKVLLLSCGIS 1936
X.laevis VSGHFADPKVACWLLDPGSKERTLHNMVTNFLPYELPLLDGVGTGQGVQSLGLSANSDQT 1993
H.sapiens LEQSYEDPKVACWLLDPDSQEPTLHSIVTSFLPHELPLLEGMETSQGIQSLGLNAGSEHS 2041
M.musculus LEPSYEDPKVACWLLDPDSKEPTLHSIVTSFLPHELALLEGMETGPGIQSLGLNVNTEHS 1996
X.laevis GRYRAAIESVVVFSTMNALTSLLEKEKLWDVFHNVEMPTQYCLALLELNGIGFSTEECET 2053
H.sapiens GRYRASVESILIFNSMNQLNSLLQKENLQDVFRKVEMPSQYCLALLELNGIGFSTAECES 2101
M.musculus GRYRASVESVLIFNSMNQLNSLLQKENLHDIFCKVEMPSQYCLALLELNGIGFSTAECES 2056
X.laevis QKHIMQAKLNEIEAVAYQLAGHSFSLTSHDDVAQVLFIELKLPPNGDLKGQGNKKTLGYT 2113
H.sapiens QKHIMQAKLDAIETQAYQLAGHSFSFTSSDDIAEVLFLELKLPPNREMKNQGSKKTLGST 2161
M.musculus QKHVMQAKLDAIETQAYQLAGHSFSFTSADDIAQVLFLELKLPPNGEMKTQGSKKTLGST 2116

Figure 3.1 Tripartite structure and sequence conservation in vertebrate Polθ. (**A**) Domain structure of *Xenopus laevis* Polθ, with a Helicase-like domain on the N-terminus harboring a Rad51intercating domain and, Polymerase domain on the C-terminus, conjoined by the central spacer region. (**B**) Multiple sequence alignment between *Xenopus laevis*, *Homo sapiens* and *Mus Musculus* Polθ showing indiviual domains. The highlighted portion in green on the N-terminus is the Helicase-like domain, while the highlighted portion in yellow on the C-terminus is the polymerase domain.

3.2 Cloning and protein purification of *Xenopus laevis* Polθ

Specific antibodies are an invaluable tool in a biochemistry lab to be used in multiple antibody-based techniques such as western blotting, immunodepletion, chromatin immunoprecipitation etc. We use *Xenopus laevis* as a model system, and since there were no commercial antibodies available against *Xenopus* Polθ, we decided to generate Polθ antibody to establish a toolkit for the upcoming experiments. We synthesized total cDNA from the RNA extracted from *Xenopus* eggs and cloned Polθ in a pCR™Blunt II-TOPO® vector. We then cloned Polθ full length (FL), helicase-like, and polymerase domain in pFH1 and pBAC-6H-MBP-TEV baculovirus vectors and verified it by DNA sequencing Fig.3.2. 6H-MBP-Polθ (FL), 6H-MBP-Polθ-Helicase and 6H-Polθ-Polymerase plasmids were used to transfect Sf9 insect cells to overexpress the proteins with the Baculovirus expression vector system. 6H-MBP-Polθ-FL, 6H-MBP-Polθ-Helicase were purified using amylose-affinity chromatography followed by size-exclusion chromatography, whereas 6H-Polθ-Polymerase was purified using talon-affinity chromatography followed by size-exclusion chromatography on Superdex-200 column. The detailed experimental methodology is summarized in Chapter2.7. The result of the purification of the recombinant protein 6H-MBP-Polθ (FL), 6H-MBP-Polθ-Helicase and 6H-Polθ-Polymerase are reported in Fig.3.3. the obtained recombinant proteins showed a molecular weight corresponding to their expected size of 323.8 kDa, 157.6 kDa and 91.3 kDa respectively. From one liter of insect cells 1 mg of 95% pure recombinant protein was obtained.

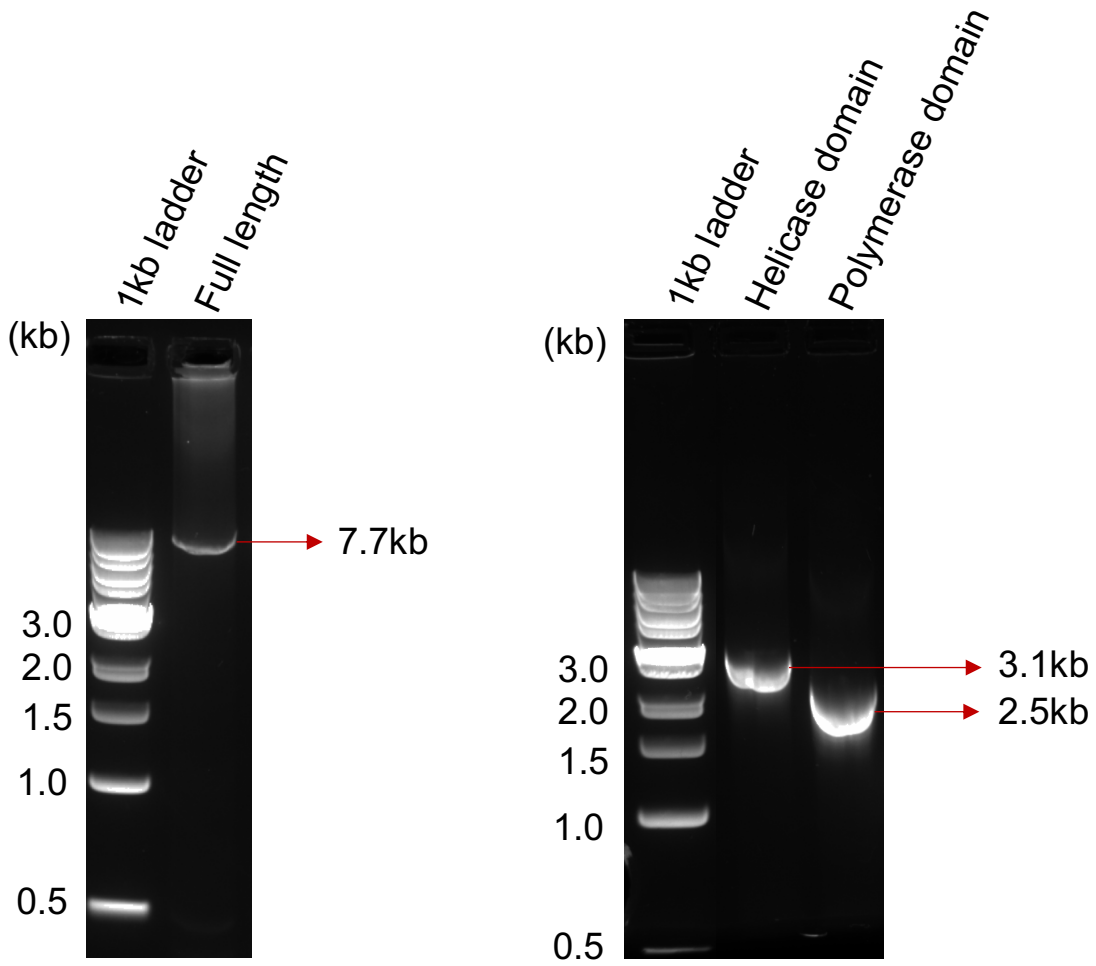


Figure 3.2 Cloning *Xenopus laevis* Polθ. Agarose gel electrophoresis of PCR products encompassing Polθ full length (FL), Polθ helicase-like domain and Polθ polymerase domain sequences, used for cloning in pFH1 and pBAC-6H-MBP-TEV vectors.

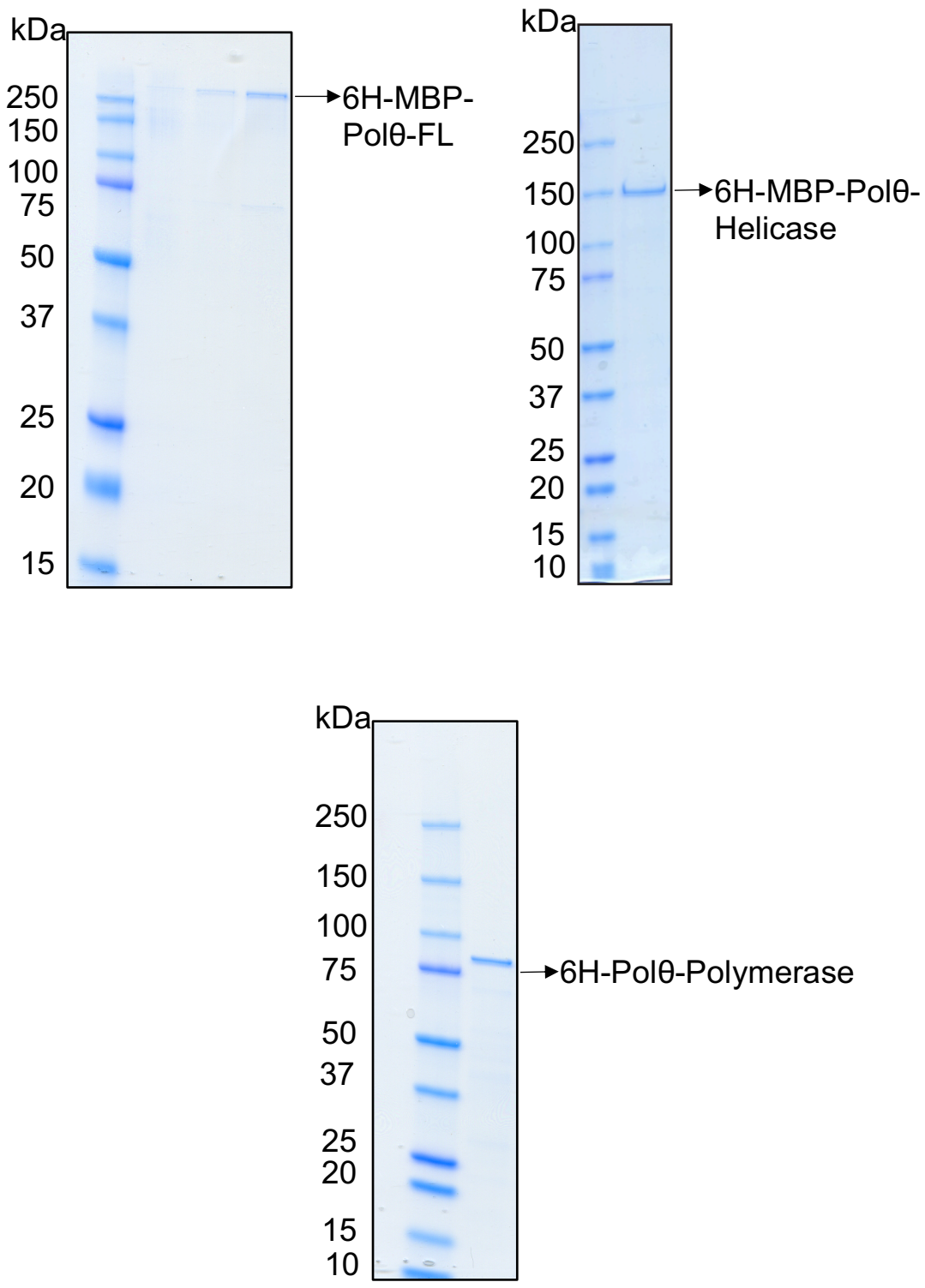


Figure 3.3 Purified recombinant *Xenopus laevis* Polθ. Coomassie Brillant blue stained SDS-PAGE gel showing pre-stained marker and purified recombinant protein 6H-MBP-Polθ-FL, 6H-MBP-Polθ-Helicase and 6H-Polθ-Polymerase of molecular weight 323.8 kDa, 157.6 kDa and 91.3 kDa respectively.

3.3 Antibody production and characterization for *Xenopus laevis* Polθ

The purified recombinant protein 6H-Polθ-Polymerase was used to immunize two rabbits (individuals 29046 and 29047) by BioGenes GmBH to raise polyclonal Anti-Polθ antibody (see section 2.8). All antibodies were purified by affinity chromatography (by Giuseppe Ossolengo, IFOM proteomic facility) against 6H-MBP-Polθ-Polymerse from final bleeds and tested for western blotting. The recombinant protein (20 ng) and the interphase extract from *Xenopus laevis* eggs (1µl) were loaded onto a 4-15% gradient polyacrylamide gel for protein electrophoresis and subsequently characterized by western blotting. Both anti-Polθ rabbit 29046 and 29047 could recognize the recombinant protein (6H-MBP-Polθ-FL, 323.8 kDa) but not the endogenous Polθ in the total extract (data not shown). Then we hypothesized it could be because the endogenous Polθ is less than 2nM in the Xenopus egg extract [111] and hence could possibly be below the detection limit in the total extract. So, we performed a chromatin binding experiment to enrich DNA-bound proteins and upon immunoblotting a band was observed at the predicted Polθ molecular weight >250 kDa in reference to the loading marker. To test the specificity of the antibody, we performed an immunodepletion of Polθ by both the canonical Protein A Dynabeads method (Fig.3.5), and by hTRIM21 based rapid degradation of endogenous proteins (Fig.3.4). hTRIM21 is an E3 ubiquitin ligase which binds strongly to the constant region of antibodies and recruits ubiquitin-proteosome system to antibody bound antigens leading to their degradation[108]. As expected, the antibody (Rabbit 29046) recognized a clear band both in the mock-depleted extracts and also the recombinant 6H-MBP-Polθ-FL while no band was detected in the Polθ depleted lanes. The

purified rabbit polyclonal anti-Polθ antibody exhibited high specificity and sensitivity against *Xenopus* Polθ. For all the experiments in this study, affinity purified anti-Polθ Rabbit 29046 was used for western blotting and affinity purified anti-Polθ Rabbit 29047 was used for immunodepletion.

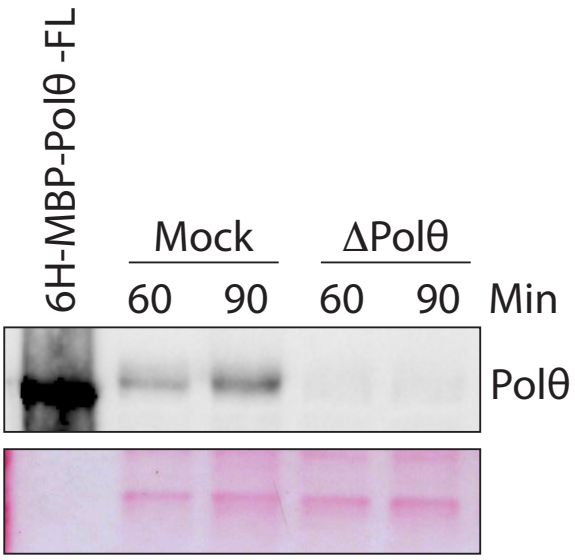


Figure 3.4 *Xenopus laevis* Polθ antibody characterization. Western blot analysis of hTRIM21 based depletion of Polθ using the affinity purified polyclonal anti-Polθ antibody (r29047). First lane shows 20ng of recombinant protein while the other lanes display proteins from the chromatin bound fractions. For the immunodetection the anti-Polθ antibody (r29046) was used. On the bottom part ponceau staining of the filter showing equal amount loading.

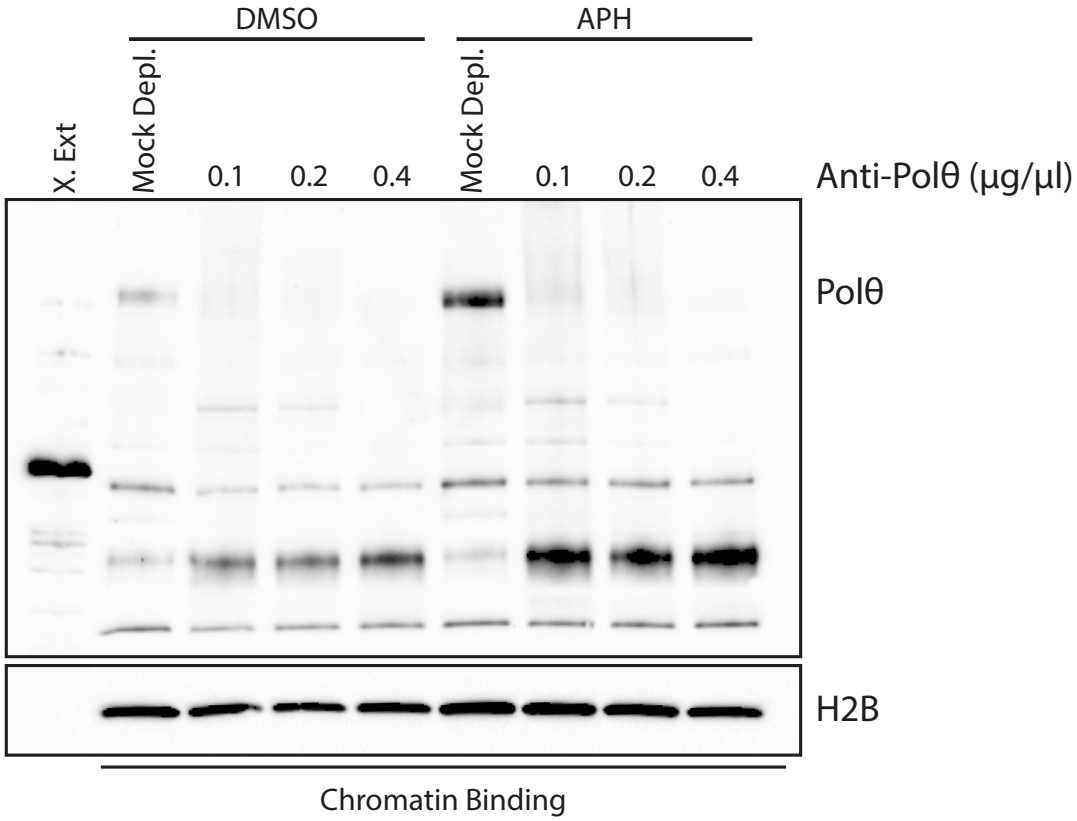


Figure 3.5 *Xenopus laevis* Polθ antibody characterization by Protein A Dynabeads based depletion. 1 μl of egg extract, chromatin bound fractions from mock depleted extract and Polθ depleted extract are loaded in 4-15% SDS-PAGE gel and immunoblotted to test the level of depletion.

3.4 DNA polymerase activity of Polθ

Polθ is a Family A polymerase which is classified as family of replication and repair polymerases[112]. To test the polymerase activity of the purified recombinant proteins 6H-MBP-Polθ-FL and 6H-Polθ-Polymerase, we performed fluorescent primer extension assay. The detailed protocol is summarized in section2.6. 6H-MBP-Polθ-FL extended the GC rich primer up to 6 nucleotide bases, though the extension length was limited by the length of the complementary strand. Then we examined if the Polθ polymerase domain alone can catalyse DNA synthesis on these primers. Indeed, 6H-Polθ-Polymerase was independently sufficient to elongate ssDNA. In fact, recombinant polymerase domain seems to catalyse the reaction faster when compared to the full length (Fig.3.6). We also observed degradation of this newly synthesized 20mer oligonucleotide at longer time points (data not shown), suggesting an inbuilt exonuclease function. Although, Maga, *et al*. 2002[53] has also suggested an intrinsic 3' to 5' exonuclease activity in hPolθ, in contrast to Seki, *et al*. 2003[52] where the authors ruled out any possible exonuclease activity in Polθ. Polθ has an exonuclease domain in the C-terminal third of the protein, the exonuclease activity we observed could be either because of the functionality of this domain or it could be due to co-purification of an associated exonuclease. Further studies are required to distinguish between the two cases. These experiments indicated that both 6H-MBP-Polθ-FL and 6H-Polθ-Polymerase could catalyse oligo extension. Polθ has an ability to add nucleotides on the 3' end of ssDNA, primed by a short stretch of dsDNA (Fig.3.6).

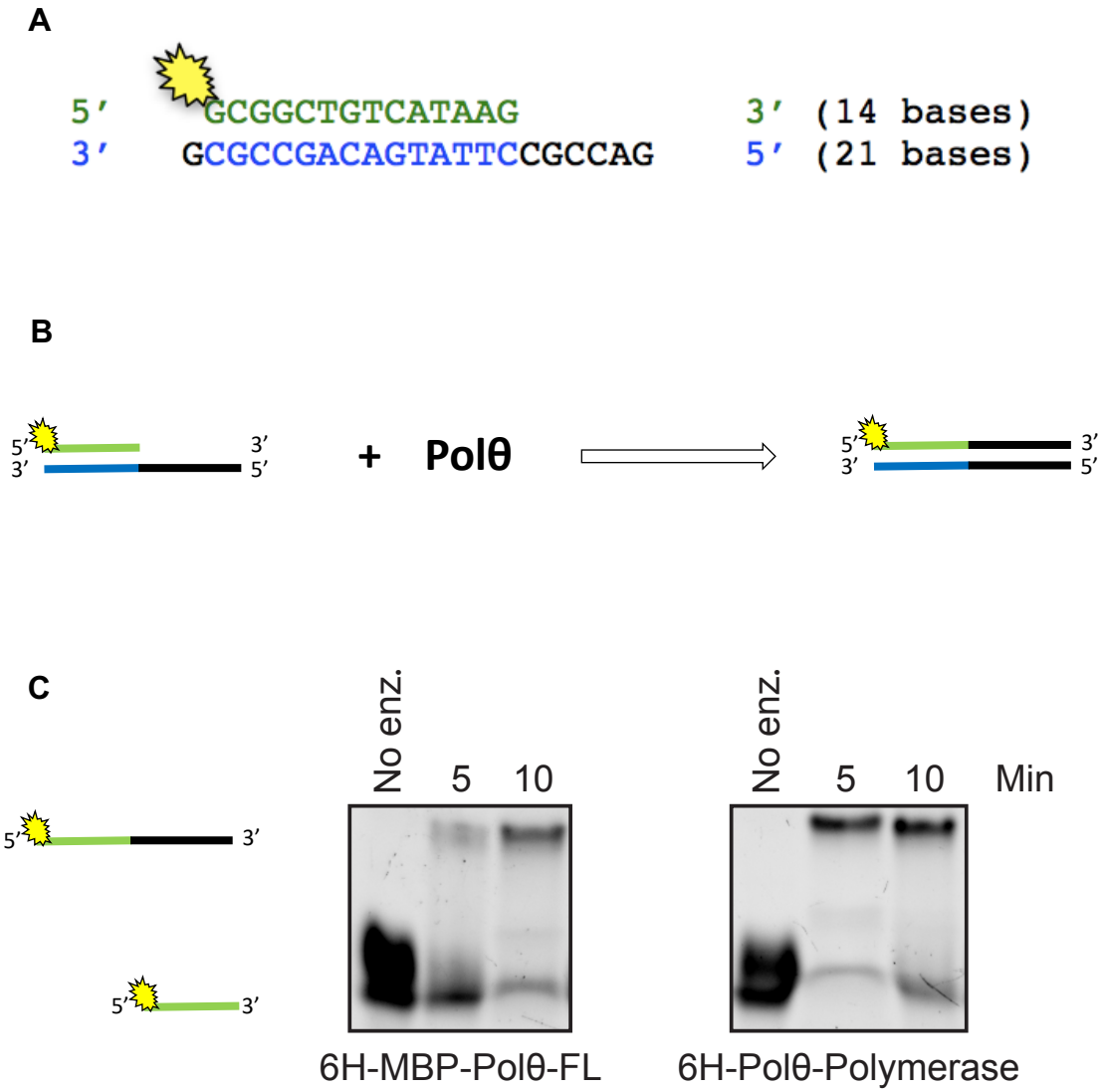


Figure 3.6 DNA polymerase activity of Polθ. (**A**) Template used for the primer extension assay. 5' of the forward primer was modified with a TET (tetrachlorofluorescein) fluorophore. (**B**) Schematic design of the the *in vitro* fluorescent-oligo extension assay. (**C**) TBE-urea-polyacrylamide gel electrophoresis showing elongated primer products. The gel was imaged using 520/30 filter on ChemidocXRS+ imager after UV excitation.

3.5 Recombinant 6H-MBP-Polθ-FL, 6H-MBP-Polθ-Helicase and 6H-Polθ-Polymerase bind to the replicating chromatin

Xenopus laevis egg extracts are a very powerful tool to investigate dynamics of chromatin bound proteins in various cellular processes such as DNA replication, chromatin remodelling and DNA damage response [98, 113]. In order to characterize the biochemical dynamics of loading of the recombinant Polθ proteins on chromatin, during DNA replication, chromatin binding experiments were performed (see section 2.3). As shown in Fig.3.7, Fig.3.8, Fig.3.9, 6H-MBP-Polθ-FL, 6H-MBP-Polθ-Helicase and 6H-Polθ-Polymerase associate with replicating chromatin and their binding is similar to Rad51, RPA70 and Smarcal1. These experiments showed that Polθ was already loaded on the chromatin at 30 mins after the start of the replication reaction and remains there for up to 120 minutes. We also observed a delay in replication timing upon 6H-MBP-Polθ-FL overexpression as previously shown by Lemeé, et al., 2010[71]. We also monitored the chromatin association of Polθ in the presence of replication fork stalling agents namely Aphidicolin (APH). Aphidicolin treatment led to an increase in Polθ loading onto the chromatin. APH is an inhibitor of replicative polymerases that stalls fork progression[114]. Hence the data confirmed that the recombinant 6H-MBP-Polθ-FL, 6H-MBP-Polθ-Helicase and 6H-Polθ-Polymerase bind to the replicating chromatin.

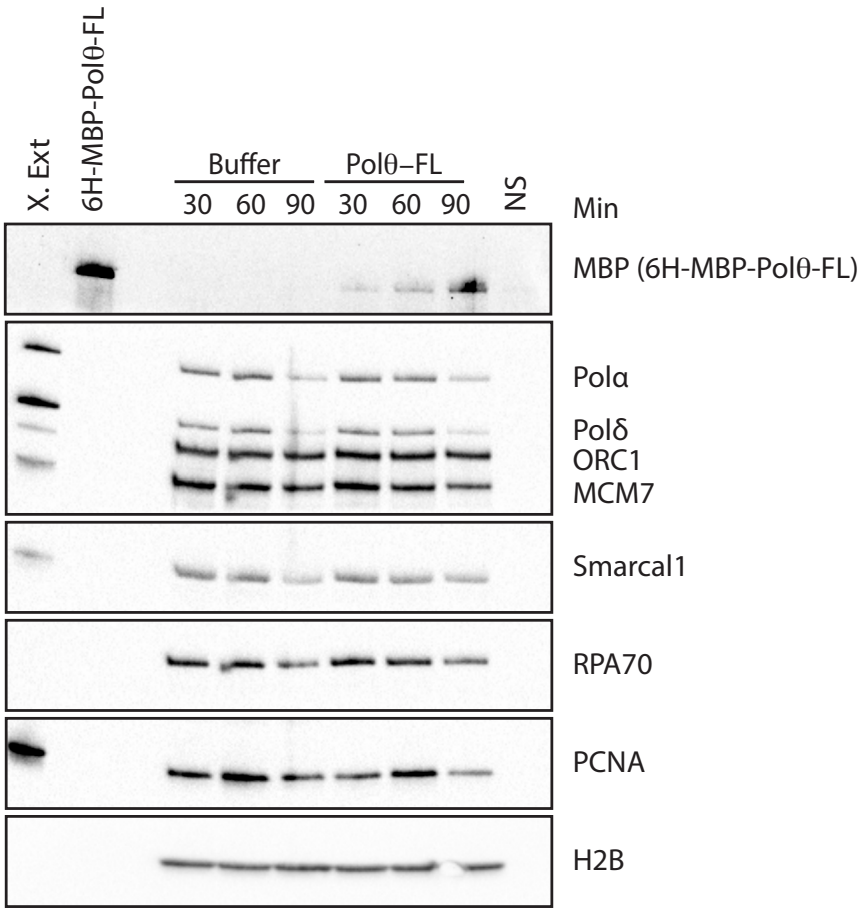


Figure 3.7 6H-MBP-Polθ-FL associates with replicating chromatin.

Chromatin binding time course with or without recombinant 6H-MBP-Polθ-FL (6.5ng/µl). 20ng recombinant 6H-MBP-Polθ-FL was loaded as a positive control in lane 2. Western blotting was carried out with the chromatin fraction from 30 µl of extract incubated with 4000 nuclei/µl for the indicated times. To detect the recombinant protein anti-MBP antibody was used. Western blot analysis shows loading dynamics of different DNA replication and repair factors. NS: no sperm nuclei.

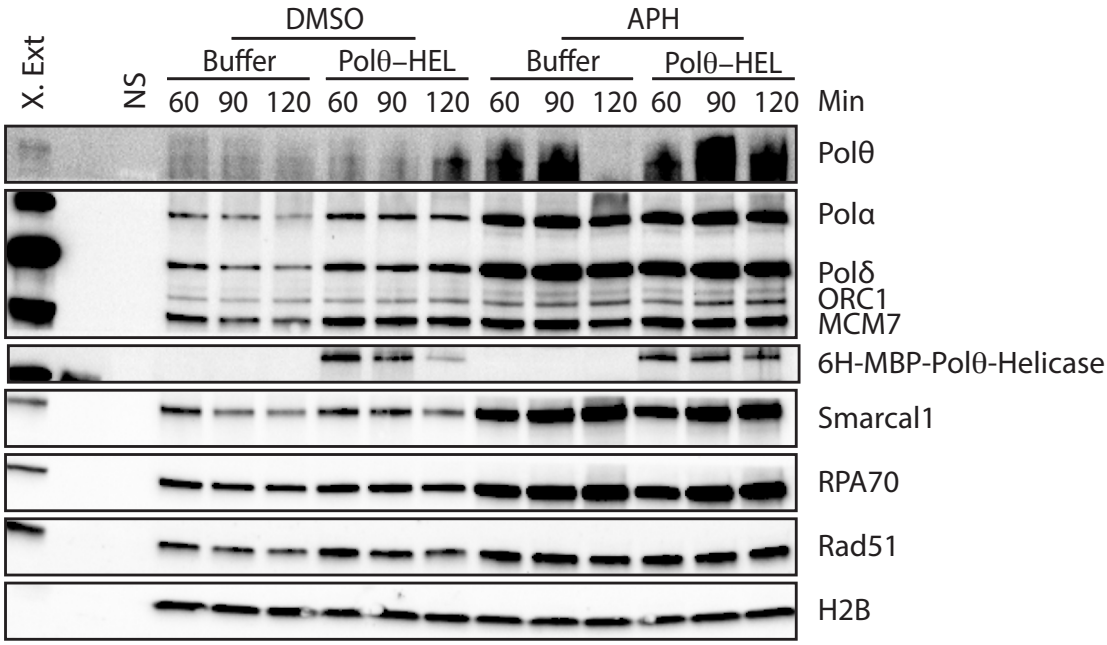


Figure 3.8 6H-MBP-Polθ-Helicase associates with replicating chromatin.

Chromatin binding time course with or without recombinant 6H-MBP-Polθ-Helicase(10.5ng/µl)., and 1.5mM aphidicolin. APH was added at time 45 into the egg extract. Western blotting was carried out with the chromatin fraction from 30 µl of extract incubated with 4000 nuclei/µl for the indicated times. To detect the recombinant protein anti-MBP antibody was used. Western blot analysis shows loading dynamics of different DNA replication and repair factors. NS: no sperm nuclei.

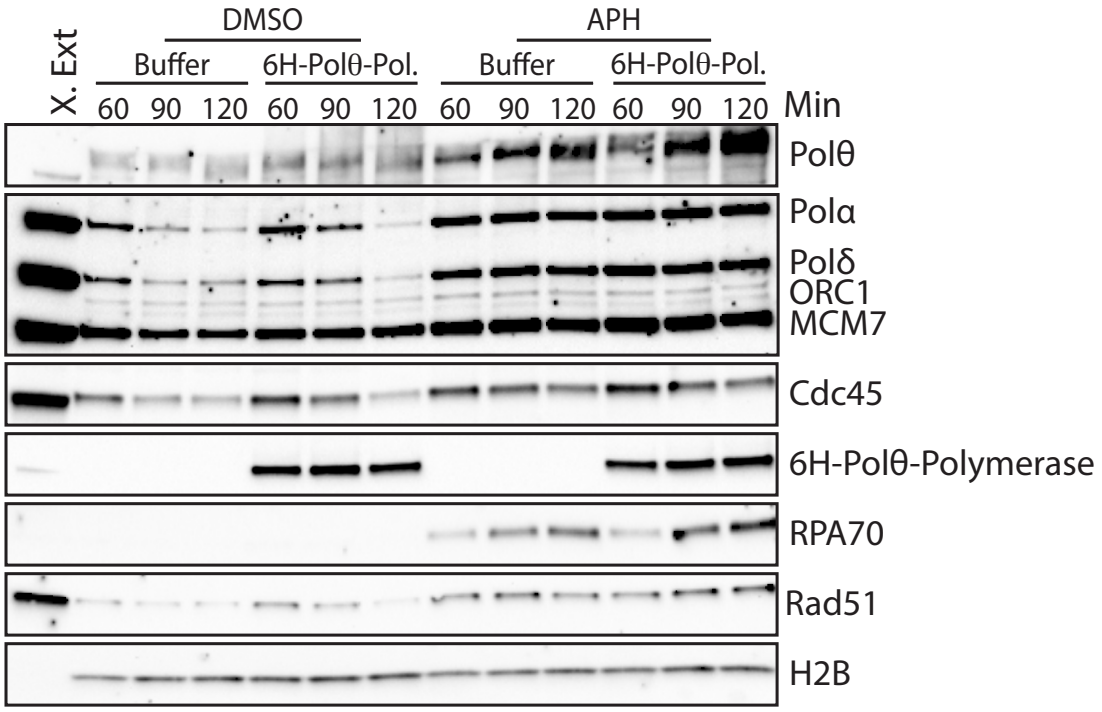


Figure 3.9 6H-Polθ-Polymerase associates with replicating chromatin.

Chromatin binding time course with or without recombinant 6H-Polθ-Polymerase (12.5ng/µl)., and 1.5mM aphidicolin. APH was added at time 45 into the egg extract 45 min after sperm nuclei addition. Western blotting was carried out with the chromatin fraction from 30 µl of extract incubated with 4000 nuclei/µl for the indicated times. To detect the recombinant protein anti-Polθ antibody was used. Western blot analysis shows loading dynamics of different DNA replication and repair factors.

3.6 Polθ is recruited at DNA double strand breaks

Polθ is a crucial player in microhomology mediated end-joining (MMEJ) repair pathway, an alternate route in homologous-recombination (HR) deficient cells that appear to defend against DNA double strand breaks (DSBs)[51]. Biochemical studies imply that HR and MMEJ share the same substrate: resected DSBs with short single-stranded DNA (ssDNA) 3' overhangs bound by RPA[57]. To reconfirm the role of Polθ at DSBs, we performed a chromatin binding assay using restriction enzymes EcoRI and EcoRV to mimic DSBs by creating blunt and sticky ends respectively on the *Xenopus laevis* sperm DNA. These treatments lead to an increase in Polθ loading on the chromatin as compared to the mock treatment (Fig.3.10), suggesting that in egg extracts Polθ responds to DNA double strand breaks as previously described in the literature using mammalian cells. We also observed that Polθ stays on the chromatin up to later time points unlike Polα and Polδ which disassociate from the chromatin after one complete cycle of DNA replication at time 60/90 mins based on the quality of the egg extract. The enrichment of Polθ on the chromatin is higher upon treatment with EcoRV as compared to EcoRI which could imply that Polθ has a stronger affinity for blunt ends as opposed to resected 3' overhangs, or it could be a technical limitation which is specific for these enzymes. The experiment needs to be repeated with another set of restriction enzymes which cut in blunt and sticky end fashion.

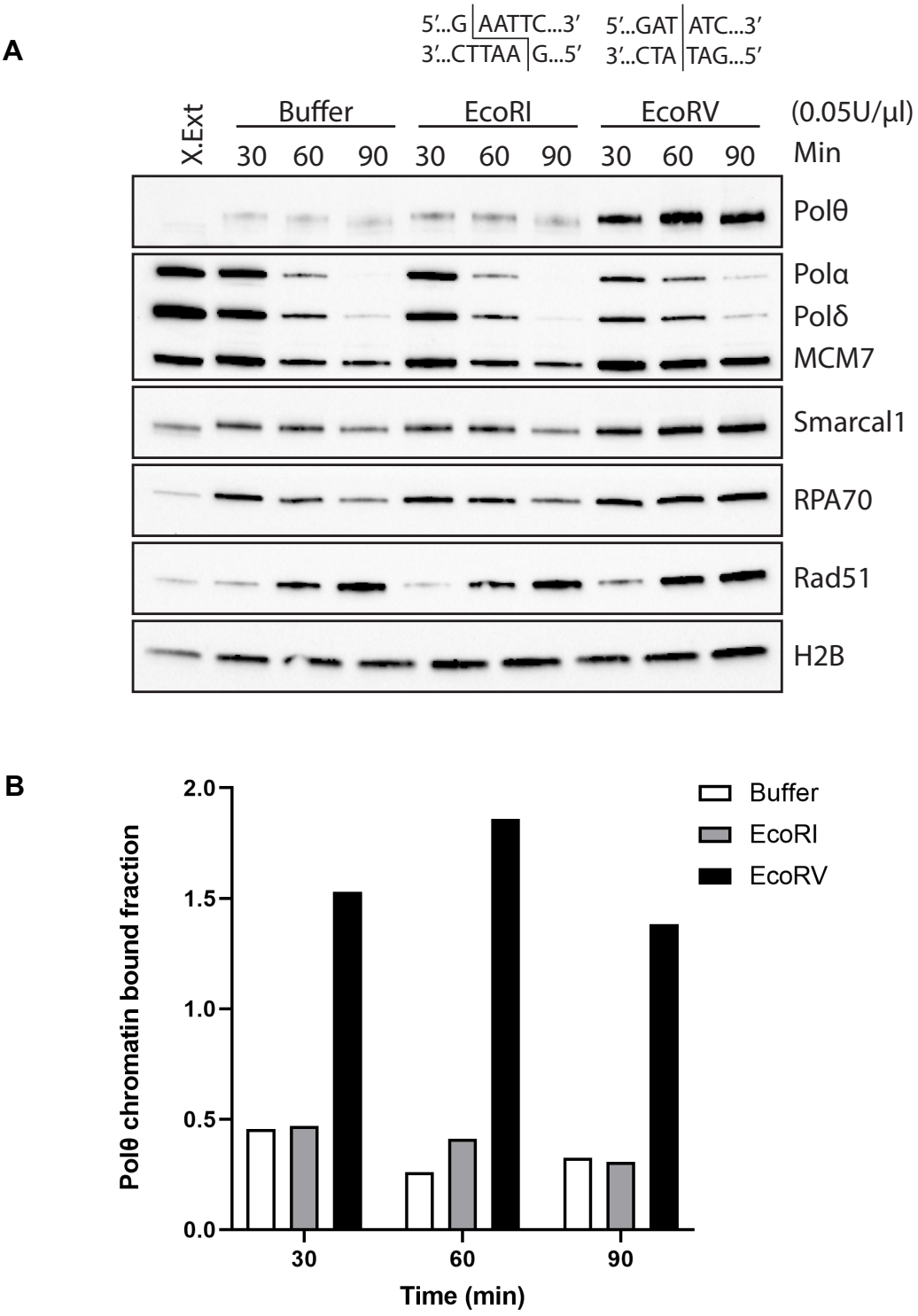


Figure 3.10 Polθ is enriched on the chromatin upon DNA double strand breaks. (**A**) Chromatin binding time course with or without EcoRI/EcoRV (0.05U/µl), added at time 0 into the egg extract. Western blotting was carried out with the chromatin fraction from 30 µl of extract incubated with 4000 nuclei/µl for the indicated times. (**B**) Histograms represents the quantificationof chromatin bound Polθ (normalized with H2B) carried out through Fiji software.

3.7 Endogenous Polθ is enriched at stalled replication forks induced by aphidicolin

Polθ is synthetically lethal with BRCA mutations[51]. The explanation in the literature points out at Polθ functions in a backup DNA repair pathway (MMEJ) in the absence of HR. However, our lab has previously demonstrated that homologous recombination DNA repair factors such as Rad51, BRCA2 and the MRN complex are also required to ensure complete and faithful replication[84, 92]. BRCA1/2 defects are also associated with replicative defects (gaps and collapsed forks), contributing to the essential role of BRCA1/2 in replication in addition to their role in HR[115, 116]. Given the widespread overexpression of Polθ in cancer[68, 69], we asked whether Polθ could also suppress BRCA1/2 replicative defects in addition to promoting DSB repair. Therefore, to understand the involvement of Polθ in replication, we performed a chromatin binding experiment in normal and replication stress conditions. Replication stress was induced by high dose of aphidicolin (APH). APH is a potent inhibitor of DNA replication and strongly binds to the binary complex of Polα-DNA, thereby stalling replication fork [114, 117, 118]. Chromatin binding time course experiment showed an enrichment in the loading of endogenous Polθ on the chromatin upon APH (1.5mM) treatment as compared to the control, suggesting a role of Polθ at the stalled or collapsed replication forks (Fig.3.11).

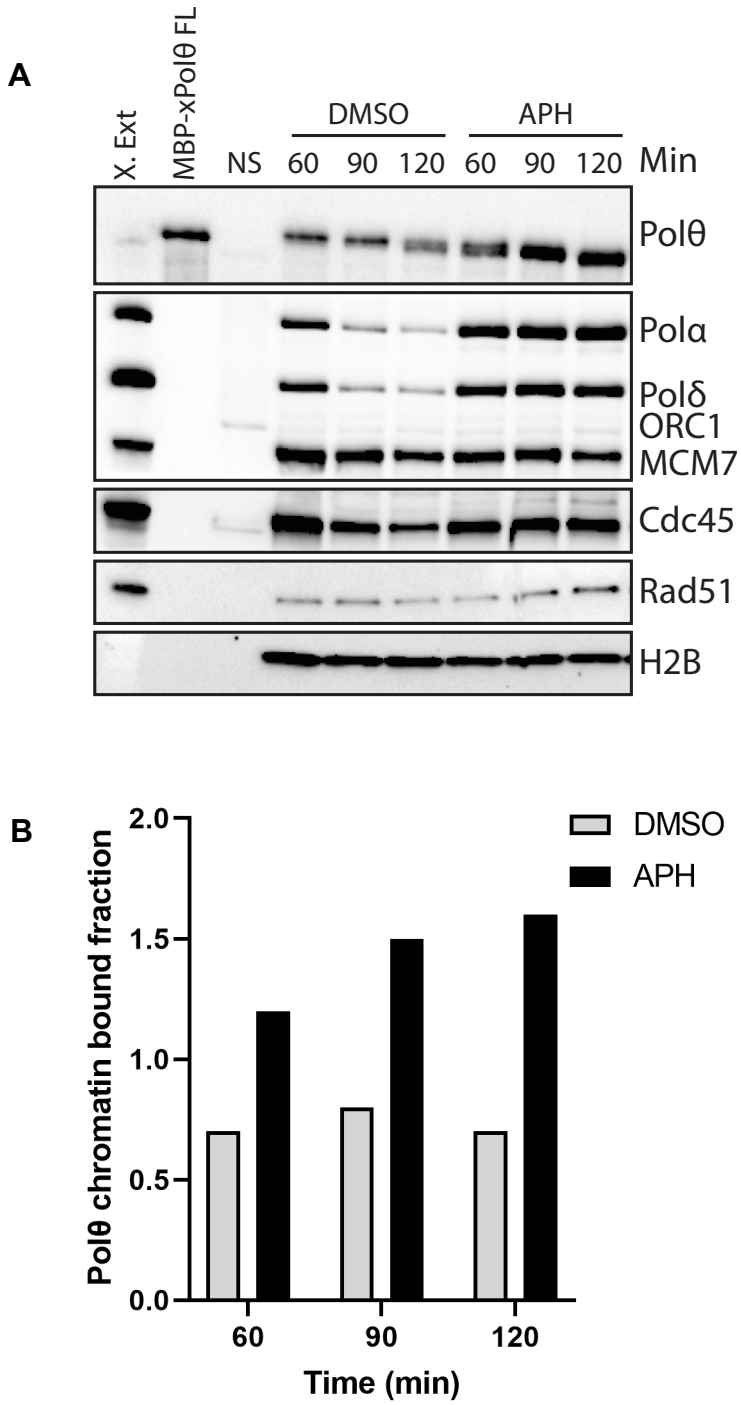


Figure 3.11 Polθ is enriched on the chromatin upon replication stress induced by Aphidicolin. (**A**) Chromatin binding time course with or without Aphidicolin (1.5mM), added at time 45 into the egg extract. Western blotting was carried out with the chromatin fraction from 30 µl of extract incubated with 4000 nuclei/µl for the indicated times. NS: no sperm nuclei. (**B**) Histogram represents the quantification of chromatin bound Polθ with respect to H2B, carried out through Fiji software.

3.8 Aphidicolin does not inhibit DNA synthesis by Polθ

Chromatin binding time course experiment showed an increase in the loading of Polθ on the chromatin upon treatment with high dose (1.5mM) of APH, implying a role of Polθ at the stalled replication forks (Fig.3.11). Then we asked whether Polθ is catalytically active or not at the stalled replication forks, if its enrichment of Polθ at the stalled replication forks could be due to a mere chelation at the forks, like it happens for Family B polymerases such as Polα and Polδ or it is catalytically active at the stalled replication forks. To test this hypothesis, we performed a fluorescence-based oligo extension assay using the purified recombinant proteins 6H-MBP-Polθ-FL and 6H-Polθ-Polymerase in APH titration background. The complete protocol is summarized in section2.6. 6H-MBP-Polθ-FL extended the GC rich primer even in the presence of 3mM APH. Then we examined if the Polθ polymerase domain alone can also catalyse DNA synthesis on these primers in APH background. Indeed, 6H-Polθ-Polymerase was independently sufficient to elongate the annealed primers even in the presence of APH. In fact, recombinant polymerase domain seems to catalyse the reaction faster when compared to the full length. To summarize, Polθ is able to polymerize even in the presence of high dose of Aphidicolin up to 3mM (Fig.3.12). Hence the data confirm that Polθ is actively engaged at stalled forks in replication stress response, likely polymerizing at the ssDNA gaps generated at the fork.

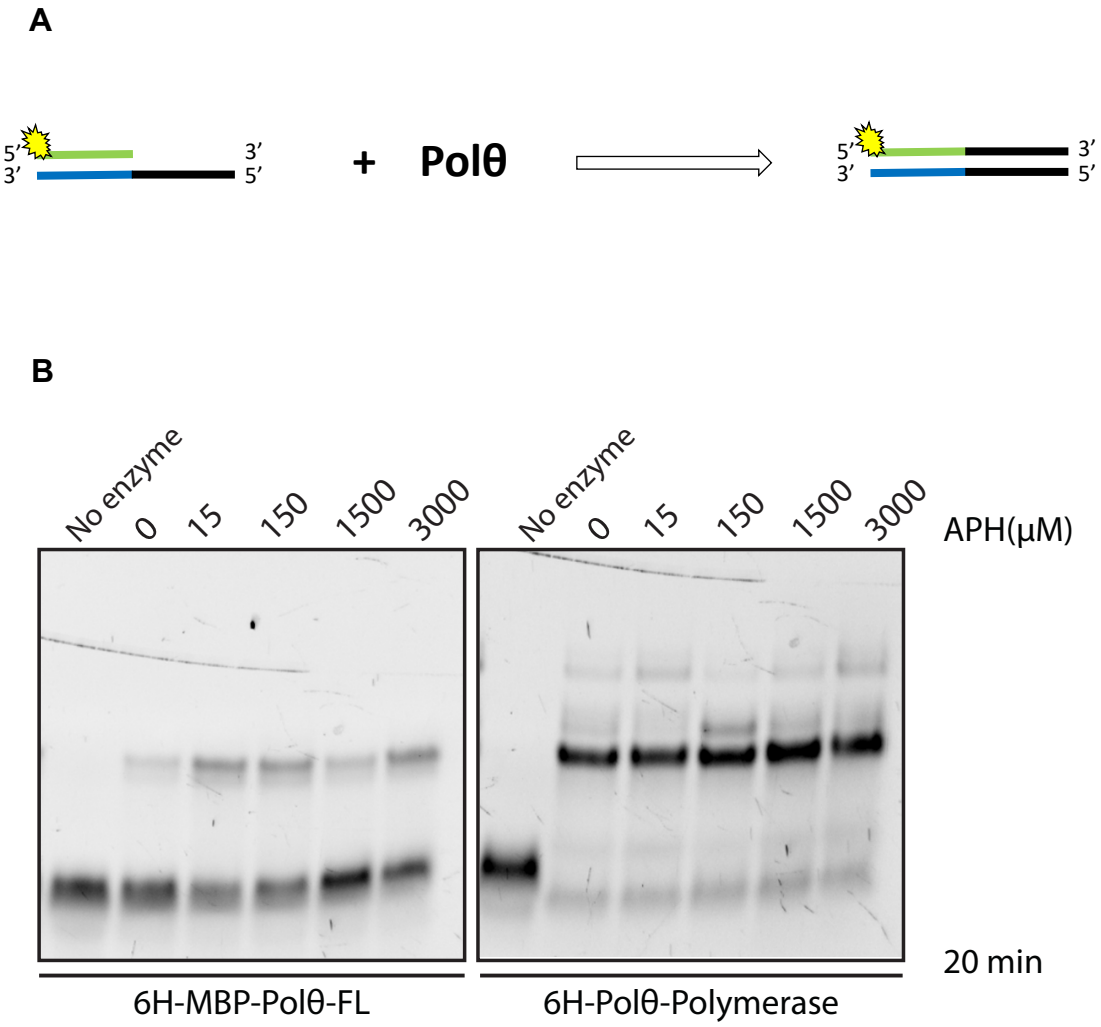


Figure 3.12 Aphidicolin does not inhibit DNA synthesis by Polθ. (**A**) Schematic representation of the the *in vitro* fluorescent-oligo extension assay. (**B**) Denaturing TBE-urea-polyacrylamide gel electrophoresis showing elongated primer products in the presence of APH. The gel was imaged using 520/30 filter on ChemidocXRS+ imager after UV excitation.

3.9 Polθ is located at the replication fork

To ask whether the enrichment of Polθ on the chromatin upon APH treatment is a specific enrichment at the replication fork, we performed iPOND experiments. iPOND, as the name reveals is a technique for isolation of proteins on nascent DNA[119, 120]. iPOND allows us to study proteins on the newly synthesized DNA with a spatial and temporal resolution[104, 121, 122]. The schematic representation of the iPOND experiment performed using *Xenopus laevis* egg extract is shown in Fig.3.13A.

In initial stages of *Xenopus* embryos, replication origins are activated in clustered and origins within each cluster are stochastically fired every 5-15kb to complete fast and faithful duplication of the genome[123, 124]. In somatic mammalian cells, origins are fired every 5—150kb, hence slowing down the overall duplication of the genome[10, 125]. However, the replication fork progression rate is ~1.2kb/min in the Xenopus system as opposed to ~2kb/min in mammalian cells[104, 124, 125]. Considering the similar fork progression rate in Xenopus embryonic and somatic mammalian cells, we adapted the iPOND protocol used in mammalian cells for the Xenopus system[104].

Extracts were supplemented with 1.5mM APH 40 min after nuclei addition and pulse-labeled for 10 min with biotin-dUTP as indicated. At 0', 15', 30' and 45' min chromatin was fractionated, sonicated and nascent chromatin was pull-down with streptavidin beads. The eluted samples were immunoblotted for Polθ, Polα, Polδ, Polε, ATR, H2B as shown in Fig.3.13B. P-Chk1 activation upon APH treatment is also shown in total extract. Interestingly, we observed a reproducible binding of Polθ at the replication fork, together with other components of replisome, upon replication stress conditions. These results emphasize the role of Polθ in replication stress, beside its role in DSB repair.

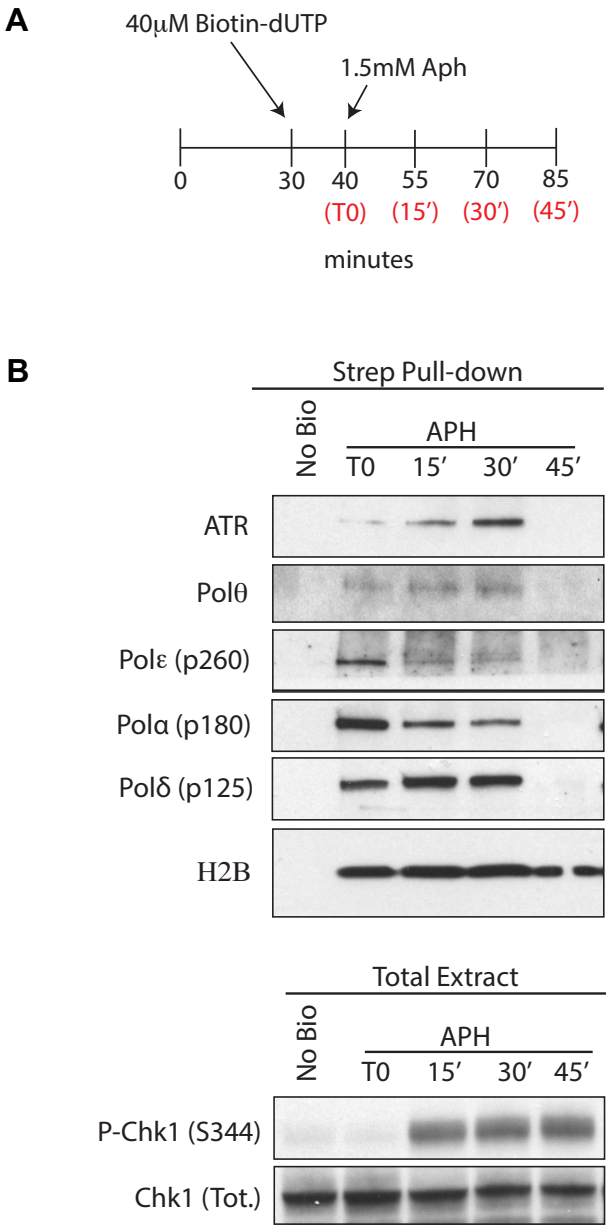


Figure 3.13 Polθ is at the replication fork in replication stress conditions

(**A**) Schematic representation of the iPOND assay. (**B**) iPOND showing proteins bound to chromatin containing nascent DNA following biotin pull-down with streptavidin beads. Extracts were supplemented with 1.5mM APH 40 min after nuclei addition and pulse-labeled for 10 min with biotin-dUTP as indicated. At 0', 15', 30' and 45' min chromatin was fractionated and nascent chromatin was pull-down with streptavidin beads. The eluted samples were analyzed by WB. P-Chk1 activation upon APH treatment is also shown in total extract.

3.10 Polθ depletion does not affect affect bulk DNA synthesis in *Xenopus laevis* extract

To verify the role of Polθ in DNA replication, we assessed the replication efficiency and loading of replicative polymerases on the chromatin of *Xenopus laevis* extracts depleted of Polθ or by chemically inhibiting Polθ in comparison with mock-depleted extracts. As shown in Fig.3.14, Polθ was immunodepleted >95% from the *X. laevis* egg extract (for detailed protocol see section 2.11). These extracts were subjected to a chromatin binding assay in which the level of chromatin bound proteins were assessed by WB. Mock and Polθ depleted extracts were also treated with 1.5mM APH to induce replication stress. The chromatin binding results show that Polθ depletion does not overall affect the loading of replicative polymerases on the chromatin. Polθ is enriched on the chromatin upon APH as shown previously in Fig. 3.11. A slight increase in the levels of chromatin bound RPA70 and Rad51 was also observed in Polθ depleted extracts. A slight increase in the levels of RPA70 was also observed upon chemical inhibition of Polθ (5μM) as shown in Fig. 3.14B suggesting the presence of ssDNA upon Polθ inhibition. Polθ inhibited extracts were also subjected to DNA replication assay to measure the DNA replication efficiency based on the incorporation of α^{32}P – dCTP. The assessment from the non-denaturing agarose gel electrophoresis suggests that Polθ inhibition does not overall affect the rate of DNA replication in *Xenopus laevis* egg extracts (Fig.3.15A). These experiments confirmed that Polθ is not required for bulk DNA replication.

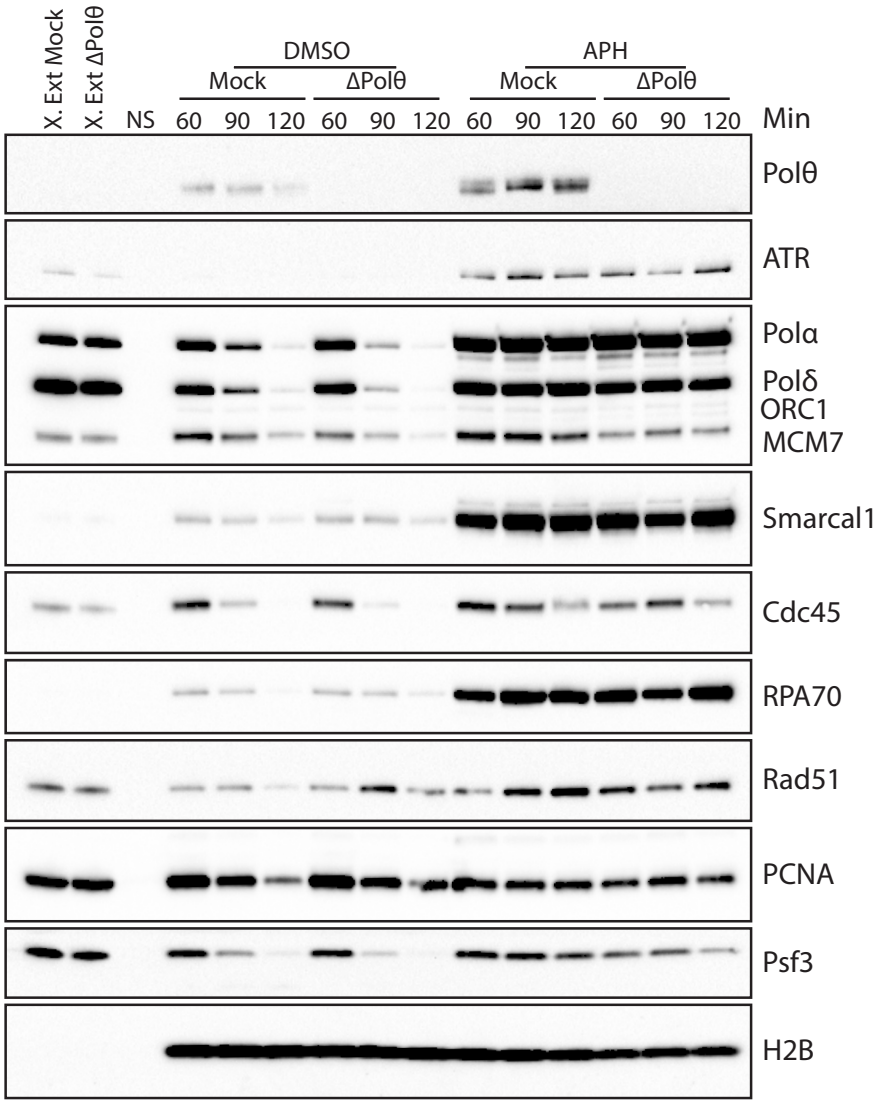


Figure 3.14 Polθ depletion does not affect loading of replicative polymerases in *Xenopus laevis* egg extract. Chromatin bound fractions from mock depleted extract and Polθ depleted extract (using 0.4μg/μl Anti-Polθ Rabbit 29047) in the presence or absence of 1.5mM APH, are loaded in 4-15% SDS-PAGE gel and immunoblotted to test the level of Polθ depletion and for other replication factors. NS: no sperm nuclei.

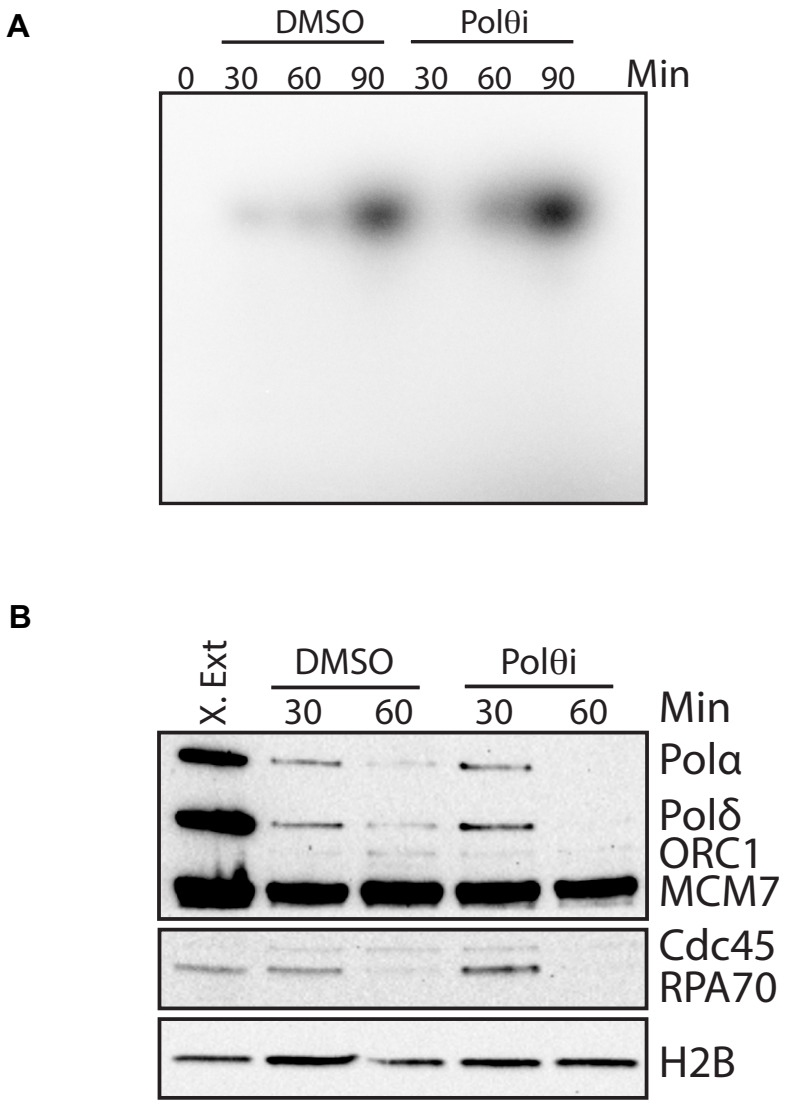


Figure 3.15 Chemical inhibition of Polθ does not affect bulk DNA synthesis in *Xenopus laevis* egg extract. (**A**) Interphase extract was supplemented with sperm nuclei (4000 nuclei/μl) and 0.1μl of α^{32}P – dCTP (3000Ci/mmol) in the presence or absence of Polθi (5μM) for indicated time points and DNA synthesis was monitored by Neutral Agarose Gel Electrophoresis. (**B**) Chromatin bound fractions from mock and Polθi extract, are loaded in 4-20% SDS-PAGE gel and immunoblotted to test the level chromatin bound replication factors.

3.11 Aphidicolin induced fork-stalling leads to large ssDNA accumulation at the fork and replication fork reversal in Xenopus laevis egg extracts

Aphidicolin is a highly specific inhibitor of DNA Polymerase α and strongly binds to the binary complex of Polα-DNA, thereby stalling replication forks[114, 126]. Since Family B polymerases are structurally similar[127, 128], Aphidicolin also weakly inhibits DNA Polymerase δ and ε without affecting the exonuclease activity of Pol ε[129, 130]. Considering the specificity of aphidicolin towards Family B polymerases, we set out to know the length of ssDNA behind and at the fork when the replicative polymerases are inhibited by Aphidicolin. Knowing the length of ssDNA at the replication fork in the presence of Aphidicolin will allow us to understand the role of Polθ at the ssDNA gaps because as we have shown in Fig. 3.12, Polθ is not inhibited by aphidicolin. Hence, we titrated the concentration of APH and monitored the level of ssDNA proteins on the chromatin and Polα accumulation. 1.5mM APH at time 60 mins into the replication reaction was optimized to induce fork stalling but without disturbing the nuclear assembly in *Xenopus laevis* egg extracts. The samples were collected 60 mins after APH treatment to be processed for electron microscopy sample preparation (for detailed protocol see section2.12). A representative normal fork and stalled fork with ssDNA gap on the daughter strand is shown in Fig.3.16 and Fig.3.17 respectively. Aphidicolin mediated fork stalling induced ssDNA gaps of an average length of 0.25-0.75kb, and nearly 15-20% of the forks showed extensive ssDNA gaps of 1.5-2.5kb (Fig. 3.18). Fork stalling by 1.5mM APH treatment also led to the formation of about 15% reverse fork formation (Fig.3.22).

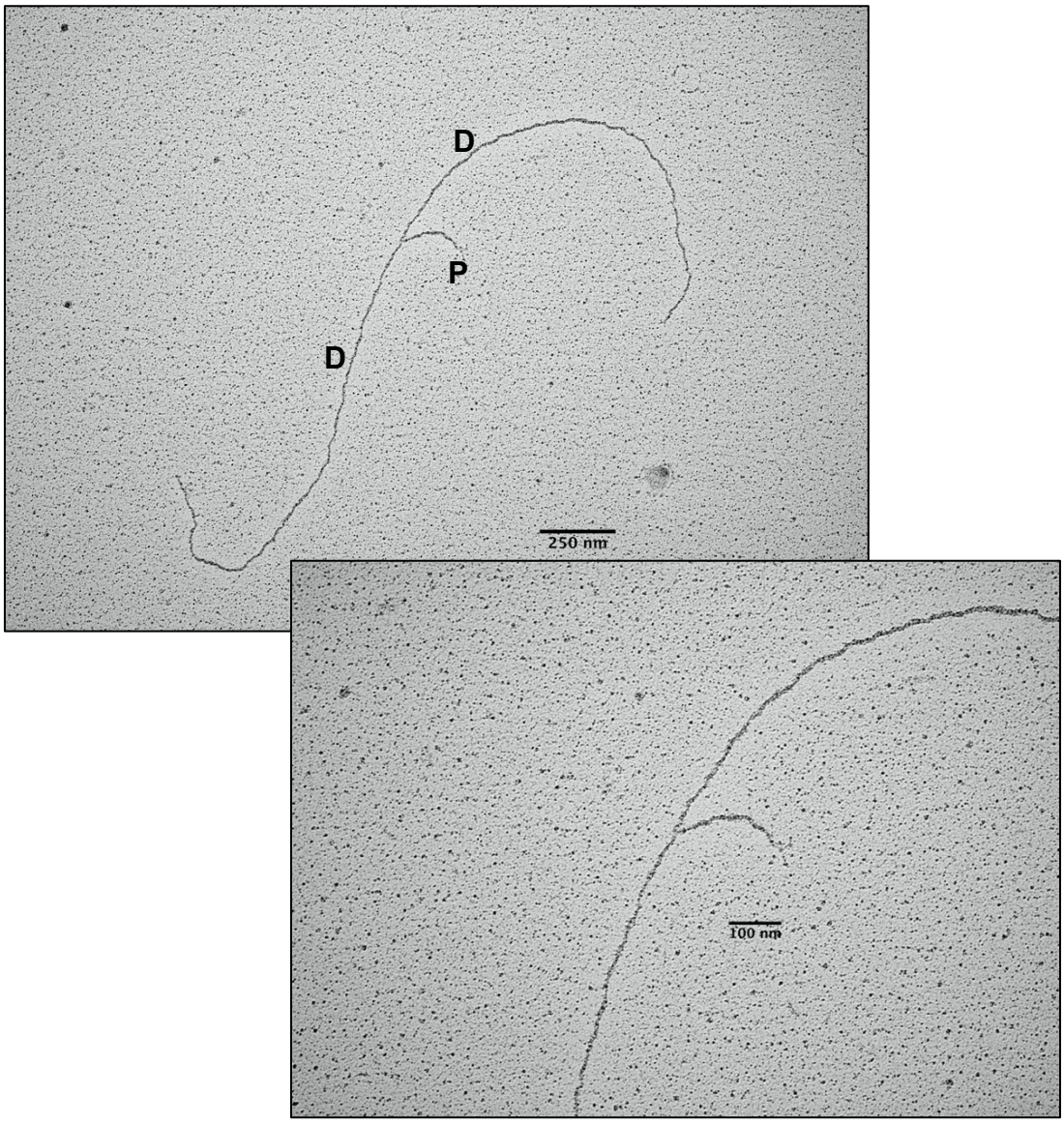


Figure 3.16 Electron Microscopic visualisation of a normal replication fork. Representative replication intermediate isolated from *Xenopus* egg extracts from control extracts. Letter P indicate the parental strand, D indicates the daughter strands.

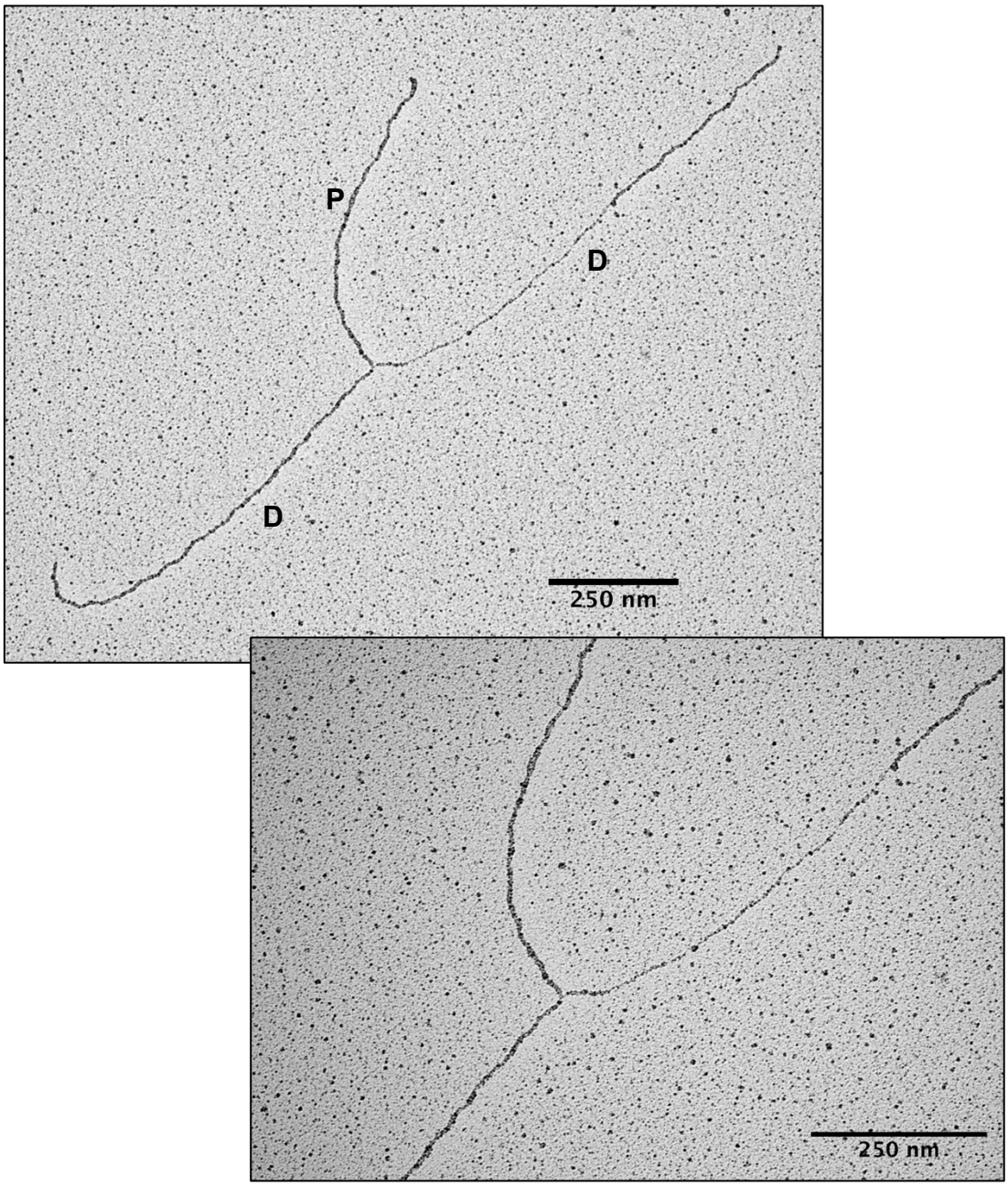


Figure 3.17 Electron Microscopic visualisation of ssDNA accumulation at the replication fork in APH treated extracts. Representative replication intermediate isolated from *Xenopus* egg extracts treated with 1.5mM APH at time 60min into the reaction. Letter P indicate the parental strand, D indicates the daughter strands.

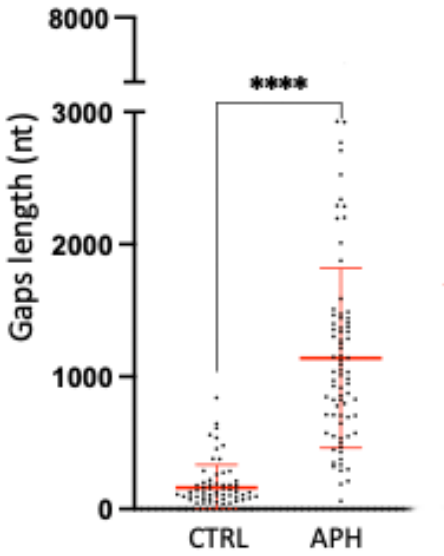


Figure 3.18 Aphidicolin induced fork-stalling leads to large ssDNA accumulation at the fork. Scattered plot distribution showing ssDNA gap length in nnucleotides (nt), obtained from Electron Microscopic data of >300 replication forks, in control and 1.5mM APH treated egg extracts.

3.12 ssDNA at replication forks induced by aphidicolin is suppressed by overload of Polθ full length

Since we observed an enrichment of endogenous Polθ on the chromatin upon replication stress conditions induced by aphidicolin treatment (Fig.3.11). We also proved that the polymerase activity of Polθ is not blocked by aphidicolin (Fig.3.12). Furthermore, we also observed an enrichment of ssDNA binding proteins such as RPA70 and Rad51 on the chromatin upon Polθ depletion, suggesting the possibilty of larger ssDNA gaps in the absence of Polθ. So we envisaged although Polθ does not play a role in bulk DNA replication, could it be involved in preventing large ssDNA at the replication fork under replication stress conditions. To test our hypothesis, we performed an electron microscopy (EM) experiment by overloading 6H-MBP-Polθ–FL (6.5ng/μl) to the Xenopus extract in 1.5mM aphidicolin background to induce replication stress. We then analyzed EM DNA replication intermediates isolated from the above mentioned treatment. Electron micrographs showing ssDNA accumulation at the replication fork upon 6H-MBP-Polθ–FL overload is shown in Fig.3.19. Aphidicolin mediated fork stalling induced ssDNA gaps of an average length of 0.25-0.75kb, whereas upon Polθ overload the average length of ssDNA gaps reduced to <0.25kb (Fig.3.20). The experiment was repeated three times for statistical significance. Overall the data clearly suggests that Polθ overload is required to prevent ssDNA gaps accumulation at the forks upon replication stress.

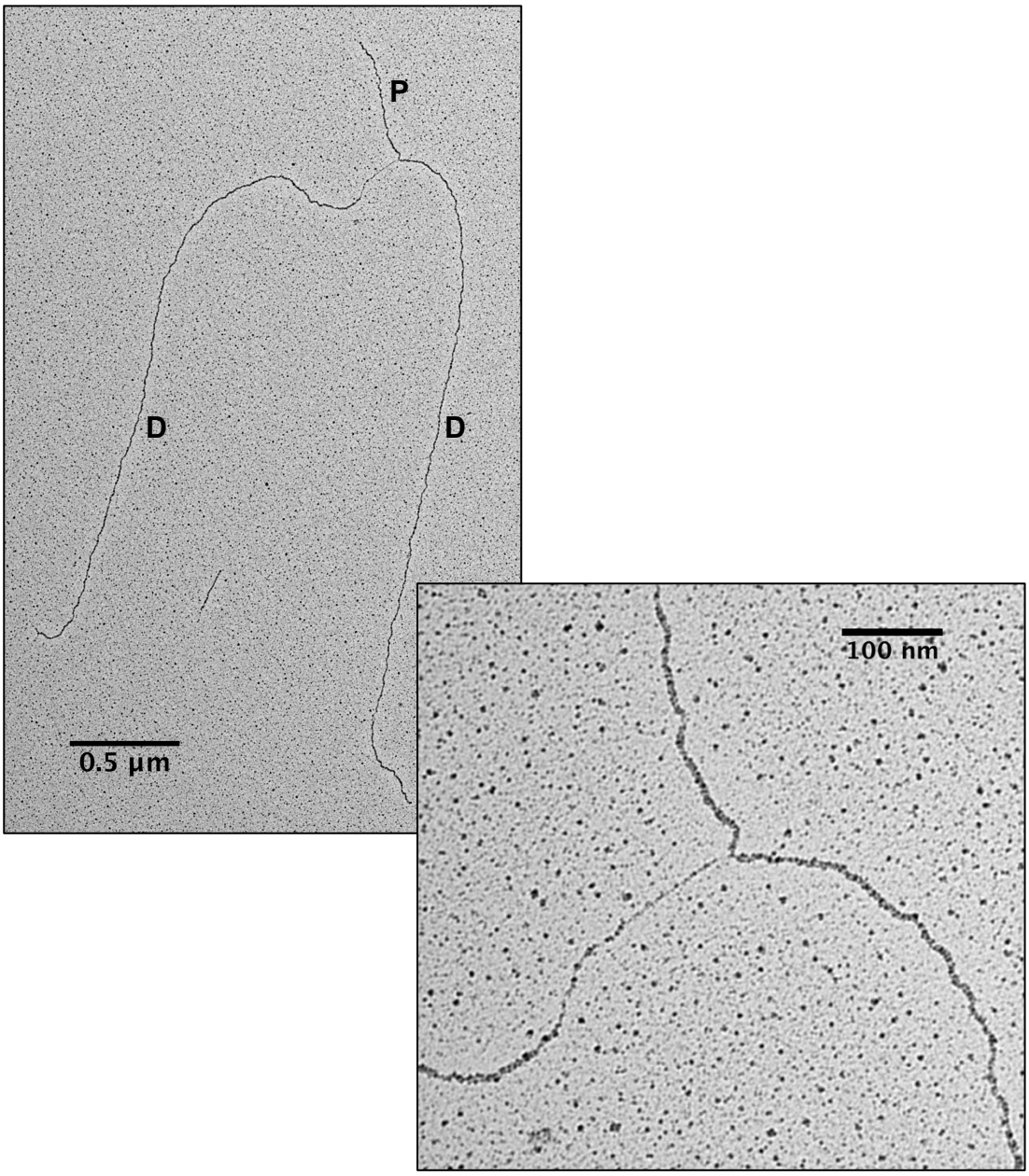
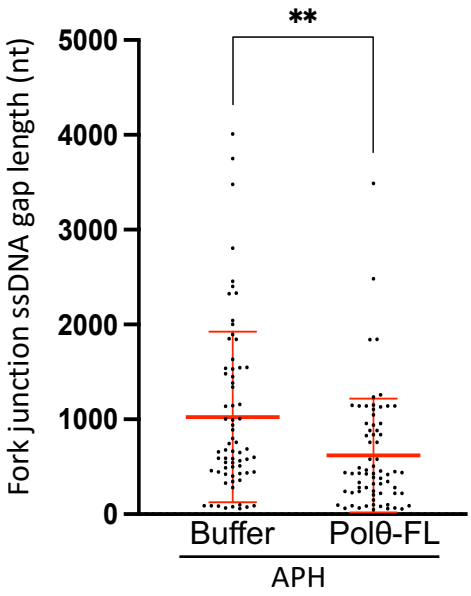


Figure 3.19 Electron Microscopic visualisation of ssDNA accumulation at the replication fork in Polθ overload. A Representative electron micrograph showing a replication intermediate isolated from *Xenopus* egg extracts supplemented with 6H-MBP-Polθ–FL (6.5ng/µl) at time 0 min and 1.5mM APH at time 60min into the reaction. Letter P indicate the parental strand, D indicates the daughter strands.

A



B

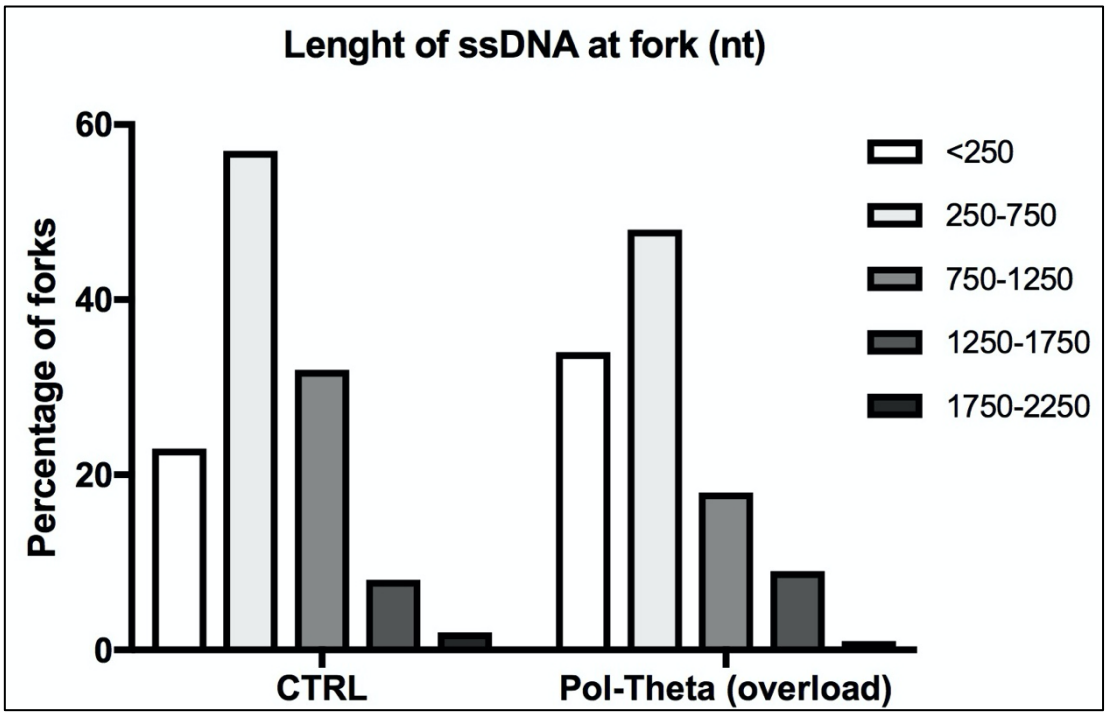


Figure 3.20 ssDNA at replication forks induced by APH is suppressed by overload of Polθ full length. (**A**) Scattered plot distribution showing ssDNA gap length in nucleotides (nt), obtained from Electron Microscopic data of >300 replication forks, in control and 1.5mM APH treated egg extracts. (**B**) Histograms showing the distribution of length of ssDNA at fork as measured in nucleotides in control and Polθ-FL overload conditions upon APH treatment.

3.13 Reverse forks induced by aphidicolin are partially suppresses by overload Polθ of full length

Replication stress is defined as the slowing down or stalling of the DNA replication fork and it may happen due to many circumstances such as depletion of the nucleotide pool, double stand breaks, lesions on the template DNA, conflicts between replication and transcription machineries etc. [5, 77, 131]. One of the ways cells cope with replication stress is by replication fork reversal, as a method to promote DNA damage tolerance and repair during DNA replication [78, 132]. Reverse forks are four-way junction structures formed by the coordinated annealing of the newly synthesized daughter strands[79, 133]. And since our in vitro experiments showed a decrease in reverse fork number upon overexpression of Polθ (data not shown), we hypothesized overexpression of Polθ would prevent fork regression in replication stress conditions mimicked by the *Xenopus* egg extract treated with APH. To test our hypothesis, we did an electron microscopy experiment mimicking Polθ overexpression by supplementing purified recombinant 6H-MBP-Polθ-FL in the replicating Xenopus extract. 45 minutes later, replicating extract was supplemented with 1.5mM APH to induce replication stress. An hour later samples were collected and processed to be observed under electron microscope. We observed a slight reduction in reverse fork (RF) number upon Polθ overload (Fig.3.22) The experiment was repeated three times for the statistical significance of the results.

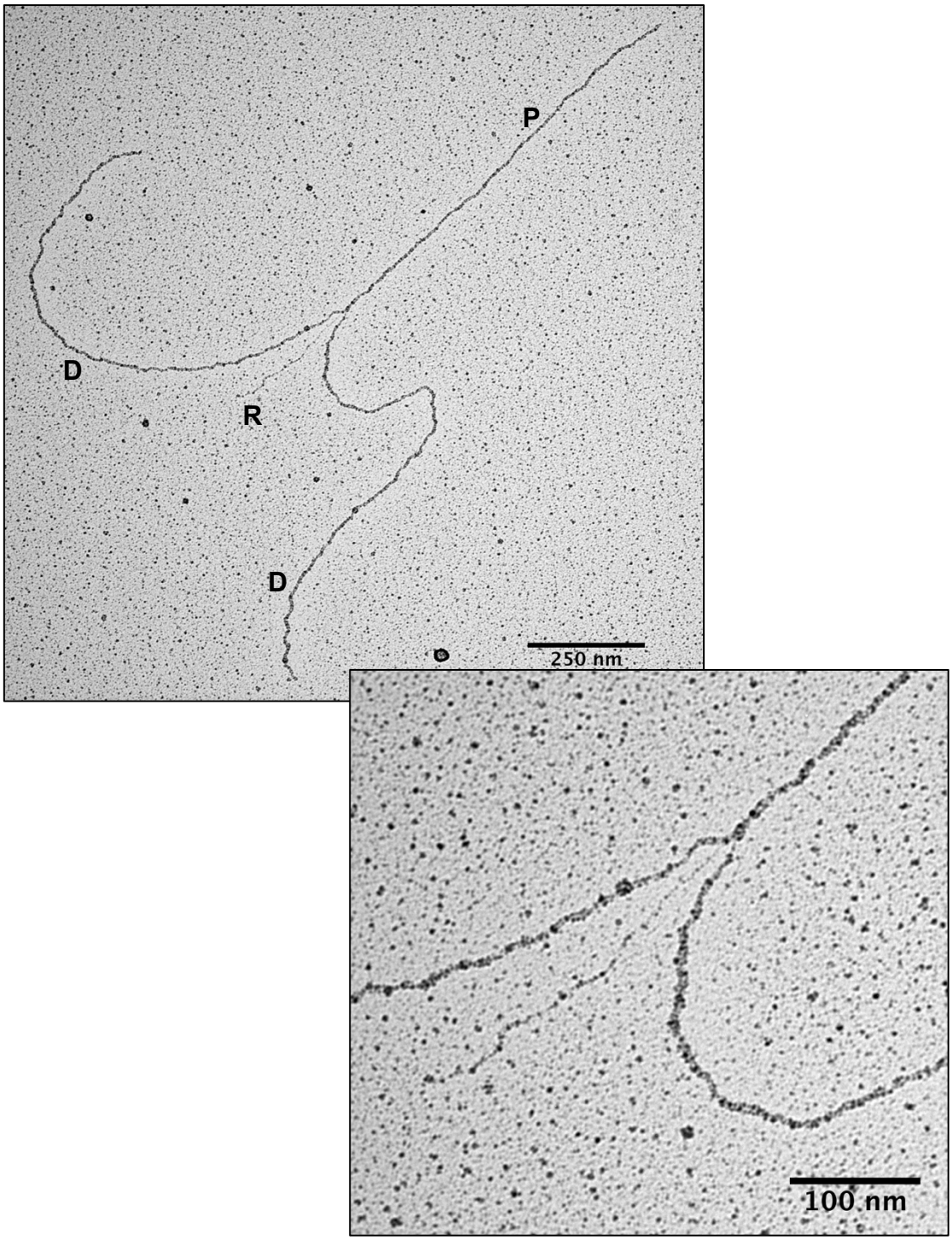


Figure 3.21 Overexpression of Polθ counteracts reverse fork formation.

A representative electron micrograph showing four branched structure after overload of Polθ in replication stress conditions caused by high dose of APH treatment. Letter P indicate the parental strand, D indicates the daughter strands, and R indicates the reversed fork.

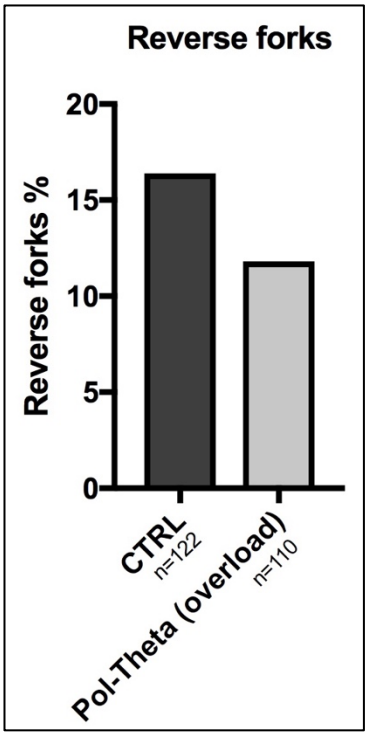


Figure 3.22 Reverse forks induced by aphidicolin are partially suppresses by Polθ overload. Graph showing reverse fork quantification in control and Polθ-FL overload conditions upon APH (1.5mM) treatment.

Chapter 4. Discussion

4.1 Polθ fills in ssDNA gaps at the replication forks

Somatic mutation theory of cancer implies all cancers are caused by somatic mutations in normal cells. Moreover, cancer genome sequencing projects have identified different 'signature' mutations that are recurrent in a specific cancer type [134, 135]. Tumor cells carrying inactive BRCA1/2 mutations display hypermutagenicity, which is evenly distributed across all genome[136-139]. This pattern of somatic mutations cannot be explained by simple role of BRCA1/2 in repairing DSBs. We hypothesize this hypermutagenicity in BRCA1/2 defiecient tumors arises from the accumulation of ssDNA gaps, which are then filled by low fidelity polymerases including Polθ. This polymerase in particular, would be extremely well suited to perform this task as it localizes right at replication forks, where ssDNA gaps arise in the absence of BRCA1/2[115, 140]. Previous finding from our lab have elucidated the role of BRCA2/Rad51 in normal replication process by biochemical analyses of replicated chromatin in *Xenopus* extract and visualization of replication intermediates by electron microscopy. Although bulk DNA replication is unaffected in either Rad51 or BRCA2 depleted extracts, two types of ssDNA gaps are clearly observed in more than 50% of the replication forks. These ssDNA gaps are seen either behind the replication fork and at the replication fork junction [84, 92]. These findings have also been confirmed in mammalian cells defective for BRCA1 and BRCA2 [141, 142]. Gaps behind the replication forks accumulate mainly because of Mre11-dependent degradation of the nascent DNA in the absence of Rad51 [92]. Gaps at fork junctions instead do not depend upon Mre11 activity. As it has been previously shown that Rad51 interacts with polymerase α thereby controlling lagging strand DNA synthesis,

it is possible, however in the absence of BRCA1/2 and/or RAD51, ssDNA accumulates due to incomplete lagging strand DNA synthesis. Alternatively, ssDNA on lagging strands could accumulate due to increased fork speed driven by unrestrained leading strand progression. In this case the lagging strand would struggle to keep up with the speed of the leading strand and would therefore accumulate ssDNA gaps. In either condition a backup polymerase which is able to fill in the gaps could prevent the deleterious accumulation of ssDNA at forks, which ultimately become the substrate of nucleases resulting in fork cleavage.

In order to investigate the involvement of Polθ at the ssDNA at the forks, we analyzed ssDNA gaps in 6H-MBP-Polθ-FL overexpression conditions, obtained by the addition of the recombinant FL protein, in APH treated conditions. APH induces the formation of ssDNA gaps at fork junctions, mimicking what happens when BRCA1/2 and RAD51 are not bound to DNA. Strikingly, we observed a reduction in ssDNA gaps at fork junction upon 6H-MBP-Polθ-FL overexpression. So, it is possible that Polθ fills in ssDNA gaps to allow the lagging strand to be completely replicated in the absence of functional BRCA1/2. In physiological conditions the gaps on the lagging strand could be due to the presence of DNA lesions such as abasic sites, which can be formed on ssDNA by the SMUG1 enzyme[143, 144]. Abasic sites accumulating on the ssDNA would be exposed to the attack of DNA processing enzymes such as APE1, which cleaves DNA on ssDNA containing abasic sites, leading to the formation of DSBs. Consistent with this, suppression of SMUG1 has been shown to alleviate the lethality associated with replication stress in BRCA1/2 defective cells[145]. Polθ would be useful in this context also due to its ability to bypass and replicated across abasic sites[146, 147].

4.2 Polθ prevents reverse fork formation

Replication stress is defined as the slowing down or stalling of replication fork progression and DNA synthesis that results in an accumulation of persistent ssDNA at the fork. Given our results it is likely that Polθ fills the gaps at fork junctions by extending the 3' end of stalled okazaki fragments on the lagging strands or 3' of the leading strand (See Figure 4.1). Polθ could work on normal template like in the case of DNA synthesis stalled by APH or by bypassing and synthesizing across an abasic site, which could be the cause of an unprovoked stalling event. The filling-in of the gaps, which are more frequent in BRCA1/2 cells, might be important to prevent further processing of stalled forks. In particular, Polθ could prevent the formation of persistent ssDNA at forks. Persistent ssDNA at the fork is remodeled into reverse forks by the coordinated annealing of the newly synthesized daughter strands into a four-way junction structure like a Holliday junction. Although it is still unclear whether replication fork reversal is a protective or a pathological condition for the cell. In a positive outlook, when DNA replication is challenged, transient replication fork reversal could be a way to prevent ssDNA exposure and subsequent cleavage by DNA exonucleases including SLX4, Mus81, Mre11 and others, thereby maintaining genome stability. Since Polθ fills in the gaps, persistent ssDNA is shortened thereby sequestering the substrate for chromatin remodelers to form a reverse fork (See Fig.4.1). Preliminary results show that the reverse forks in the absence of Polθ are cleaved by DNA endonucleases. Therefore, it is likely that overexpression of Polθ helps cells to cope with replication stress by preventing the formation and subsequent processing of reversed forks since these cells mostly lack BRCA1/2 which are the main players in resolving reverse fork structures.

4.3 Conclusion and future perspectives

To conclude, in this study, using Xenopus egg extract-based system and electron microscopy, we uncovered the role of Polθ in coping with DNA replication stress. We first generated and validated the specificity of Anti-Polθ antibody. We characterized the low-fidelity polymerase activity of Polθ even in the presence of Polα inhibitor - Aphidicolin. We showed an enrichment of Polθ on the chromatin in replication stress conditions generated by treatment with APH. We also demonstrated both that Polθ is at the replication fork, using iPOND assay. We furthermore demonstrated by electron microscopy that Polθ fills in ssDNA gaps at the replication fork and counteracts fork reversal upon replication stressed induced by APH. So, our working hypothesis (see Fig.4.1) is that the synthetic lethal relationship between *POLQ* and *BRCA*, is not limited to DSB repair but it also prevents DSB by ssDNA gap fill-in and thereby suppressing reverse fork formation. In breast and ovarian cancers Polθ overexpression provides an advantage to tumor cells in coping replication stress resulting from chemotherapeutic drugs. This basic knowledge can be used to understand how to selectively target compensatory functions of Polθ to neutralize BRCA defective tumor growth.

In the future, we plan to study, using electron microscopy, the differential roles of Polθ at leading or lagging strands since ssDNA gaps are mostly observed on one side of the replication fork. We are also interested in understanding if the reduction in reverse fork number upon Polθ overexpression could be attributed to Polθ helicase-like domain. In addition, we plan to study the structural consequences of Rad51 and Polθ absence upon replication stress.

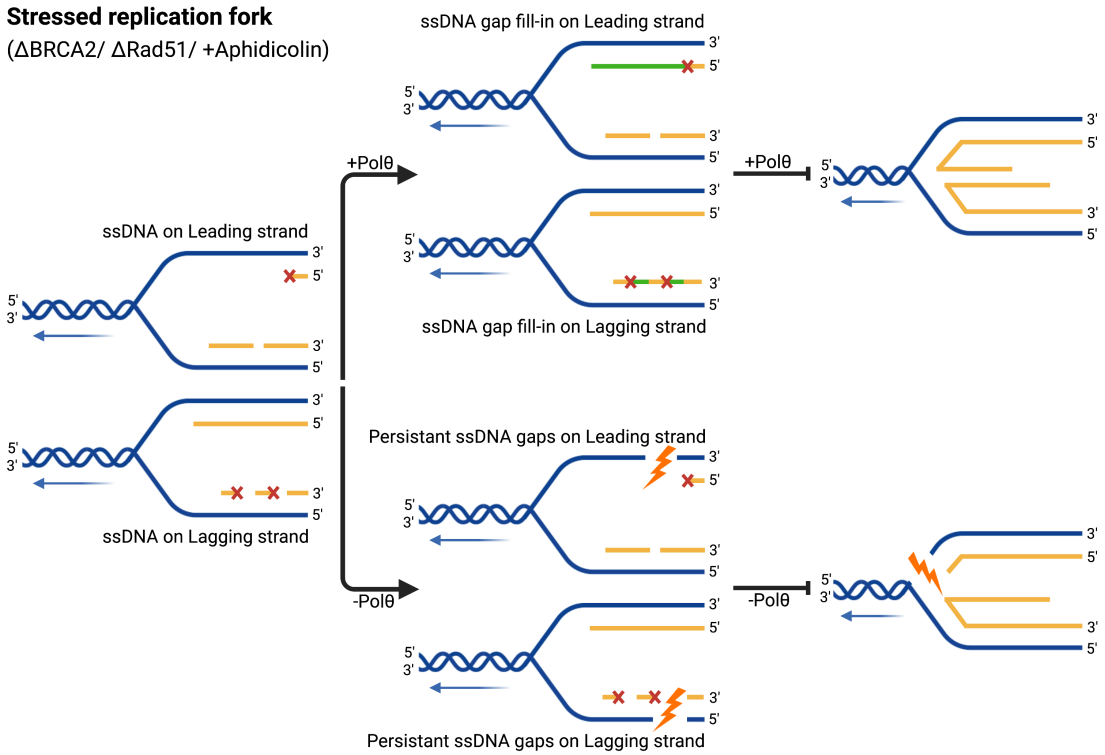


Figure4.1 Model for Polθ function at stressed replication forks. Replication stress leads to accumulation of ssDNA gaps at the fork either on the leading strand or the lagging strand. Overload of recombinant Polθ leads to ssDNA gap fill-in at the replication fork, and a reduction in reverse fork formation. However, in the absence of both Polθ and BRCA1/2, these persistent ssDNA gaps and reverse fork could be targeted by endonucleases and Holiday junction resolvases leading to genome stability. Hence, we hypothesize Polθ does not only repair DNA DSBs but also prevents their formation by ssDNA gap fill-in.

References

1. Meselson, M. and F.W. Stahl, *The Replication of DNA in Escherichia Coli.* Proc Natl Acad Sci U S A, 1958. **44**(7): p. 671-82.

2. Watson, J.D. and F.H. Crick, *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid.* Nature, 1953. **171**(4356): p. 737-8.

3. Bell, S.P. and A. Dutta, *DNA replication in eukaryotic cells.* Annu Rev Biochem, 2002. **71**: p. 333-74.

4. Dewar, J.M. and J.C. Walter, *Mechanisms of DNA replication termination.* Nat Rev Mol Cell Biol, 2017. **18**(8): p. 507-516.

5. Zeman, M.K. and K.A. Cimprich, *Causes and consequences of replication stress.* Nat Cell Biol, 2014. **16**(1): p. 2-9.

6. Preston, B.D., T.M. Albertson, and A.J. Herr, *DNA replication fidelity and cancer.* Semin Cancer Biol, 2010. **20**(5): p. 281-93.

7. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.

8. Blow, J.J. and A. Dutta, *Preventing re-replication of chromosomal DNA.* Nat Rev Mol Cell Biol, 2005. **6**(6): p. 476-86.

9. Rao, P.N. and R.T. Johnson, *Mammalian cell fusion: studies on the regulation of DNA synthesis and mitosis.* Nature, 1970. **225**(5228): p. 159-64.

10. Fragkos, M., et al., *DNA replication origin activation in space and time.* Nat Rev Mol Cell Biol, 2015. **16**(6): p. 360-74.

11. Nishitani, H. and Z. Lygerou, *DNA replication licensing.* Front Biosci, 2004. **9**: p. 2115-32.

12. Stinchcomb, D.T., K. Struhl, and R.W. Davis, *Isolation and characterisation of a yeast chromosomal replicator.* Nature, 1979. **282**(5734): p. 39-43.

13. Theis, J.F. and C.S. Newlon, *The ARS309 chromosomal replicator of Saccharomyces cerevisiae depends on an exceptional ARS consensus sequence.* Proc Natl Acad Sci U S A, 1997. **94**(20): p. 10786-91.

14. Leonard, A.C. and M. Mechali, *DNA replication origins.* Cold Spring Harb Perspect Biol, 2013. **5**(10): p. a010116.

15. Vashee, S., et al., *Sequence-independent DNA binding and replication initiation by the human origin recognition complex.* Genes Dev, 2003. **17**(15): p. 1894-908.

16. Masuda, T., S. Mimura, and H. Takisawa, *CDK- and Cdc45-dependent priming of the MCM complex on chromatin during S-phase in Xenopus egg extracts: possible activation of MCM helicase by association with Cdc45.* Genes Cells, 2003. **8**(2): p. 145-61.

17. Jones, M.L., et al., *Structure of a human replisome shows the organisation and interactions of a DNA replication machine.* EMBO J, 2021. **40**(23): p. e108819.

18. Yao, N.Y. and M. O'Donnell, *SnapShot: The replisome.* Cell, 2010. **141**(6): p. 1088, 1088 e1.

19. Baker, T.A. and S.P. Bell, *Polymerases and the replisome: machines within machines.* Cell, 1998. **92**(3): p. 295-305.

20. O'Donnell, M. and H. Li, *The Eukaryotic Replisome Goes Under the Microscope.* Curr Biol, 2016. **26**(6): p. R247-56.

21. Coster, G. and J.F.X. Diffley, *Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading.* Science, 2017. **357**(6348): p. 314-318.

22. Burhans, W.C., et al., *Identification of an origin of bidirectional DNA replication in mammalian chromosomes.* Cell, 1990. **62**(5): p. 955-65.

23. Montanari, M., et al., *Role of geminin: from normal control of DNA replication to cancer formation and progression?* Cell Death Differ, 2006. **13**(7): p. 1052-6.

24. Li, A. and J.J. Blow, *Cdt1 downregulation by proteolysis and geminin inhibition prevents DNA re-replication in Xenopus.* EMBO J, 2005. **24**(2): p. 395-404.

25. Saxena, S. and A. Dutta, *Geminin-Cdt1 balance is critical for genetic stability.* Mutat Res, 2005. **569**(1-2): p. 111-21.

26. Lutzmann, M., D. Maiorano, and M. Mechali, *A Cdt1-geminin complex licenses chromatin for DNA replication and prevents rereplication during S phase in Xenopus.* EMBO J, 2006. **25**(24): p. 5764-74.

27. Lujan, S.A., J.S. Williams, and T.A. Kunkel, *DNA Polymerases Divide the Labor of Genome Replication.* Trends Cell Biol, 2016. **26**(9): p. 640-654.

28. Chilkova, O., et al., *The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA.* Nucleic Acids Res, 2007. **35**(19): p. 6588-97.

29. Burnham, D.R., et al., *The mechanism of DNA unwinding by the eukaryotic replicative helicase.* Nat Commun, 2019. **10**(1): p. 2159.

30. Zheng, L. and B. Shen, *Okazaki fragment maturation: nucleases take centre stage.* J Mol Cell Biol, 2011. **3**(1): p. 23-30.

31. Perera, R.L., et al., *Mechanism for priming DNA synthesis by yeast DNA polymerase alpha.* Elife, 2013. **2**: p. e00482.

32. Pellegrini, L., *The Pol alpha-primase complex.* Subcell Biochem, 2012. **62**: p. 157-69.

33. Wrighton, K.H., *DNA replication: Keeping up with the leader.* Nat Rev Mol Cell Biol, 2010. **11**(1): p. 4.

34. Fairman, M.P., *DNA polymerase delta/PCNA: actions and interactions.* J Cell Sci, 1990. **95 (Pt 1)**: p. 1-4.

35. Balakrishnan, L. and R.A. Bambara, *Okazaki fragment metabolism.* Cold Spring Harb Perspect Biol, 2013. **5**(2).

36. Stodola, J.L. and P.M. Burgers, *Mechanism of Lagging-Strand DNA Replication in Eukaryotes.* Adv Exp Med Biol, 2017. **1042**: p. 117-133.

37. Okazaki, R., et al., *Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains.* Proc Natl Acad Sci U S A, 1968. **59**(2): p. 598-605.

38. Sun, H., et al., *Error-prone, stress-induced 3′ flap-based Okazaki fragment maturation supports cell survival.* Science, 2021. **374**(6572): p. 1252-1258.

39. Bae, S.H., et al., *RPA governs endonuclease switching during processing of Okazaki fragments in eukaryotes.* Nature, 2001. **412**(6845): p. 456-61.

40. Bae, S.H. and Y.S. Seo, *Characterization of the enzymatic properties of the yeast dna2 Helicase/endonuclease suggests a new model for Okazaki fragment processing.* J Biol Chem, 2000. **275**(48): p. 38022-31.

41. Leman, A.R. and E. Noguchi, *The replication fork: understanding the eukaryotic replication machinery and the challenges to genome duplication.* Genes (Basel), 2013. **4**(1): p. 1-32.

42. Patel, P.H. and L.A. Loeb, *Getting a grip on how DNA polymerases function.* Nat Struct Biol, 2001. **8**(8): p. 656-9.

43. Doublie, S. and K.E. Zahn, *Structural insights into eukaryotic DNA replication.* Front Microbiol, 2014. **5**: p. 444.

44. Hubscher, U., G. Maga, and S. Spadari, *Eukaryotic DNA polymerases.* Annu Rev Biochem, 2002. **71**: p. 133-63.

45. Burgers, P.M., *Polymerase dynamics at the eukaryotic DNA replication fork.* J Biol Chem, 2009. **284**(7): p. 4041-5.

46. Steitz, T.A., *DNA polymerases: structural diversity and common mechanisms.* J Biol Chem, 1999. **274**(25): p. 17395-8.

47. Russell, H.J., et al., *The 3'-5' proofreading exonuclease of archaeal family-B DNA polymerase hinders the copying of template strand deaminated bases.* Nucleic Acids Res, 2009. **37**(22): p. 7603-11.

48. Bebenek, A. and I. Ziuzia-Graczyk, *Fidelity of DNA replication-a matter of proofreading.* Curr Genet, 2018. **64**(5): p. 985-996.

49. Sharief, F.S., et al., *Cloning and chromosomal mapping of the human DNA polymerase theta (POLQ), the eighth human DNA polymerase.* Genomics, 1999. **59**(1): p. 90-6.

50. Wood, R.D. and S. Doublie, *DNA polymerase theta (POLQ), double-strand break repair, and cancer.* DNA Repair (Amst), 2016. **44**: p. 22-32.

51. Ceccaldi, R., et al., *Homologous-recombination-deficient tumours are dependent on Poltheta-mediated repair.* Nature, 2015. **518**(7538): p. 258-62.

52. Seki, M., F. Marini, and R.D. Wood, *POLQ (Pol theta), a DNA polymerase and DNA-dependent ATPase in human cells.* Nucleic Acids Res, 2003. **31**(21): p. 6117-26.

53. Maga, G., et al., *DNA polymerase theta purified from human cells is a high-fidelity enzyme.* J Mol Biol, 2002. **319**(2): p. 359-69.

54. Brambati, A., R.M. Barry, and A. Sfeir, *DNA polymerase theta (Poltheta) - an error-prone polymerase necessary for genome stability.* Curr Opin Genet Dev, 2020. **60**: p. 119-126.

55. Patterson-Fortin, J. and A.D. D'Andrea, *Exploiting the Microhomology-Mediated End-Joining Pathway in Cancer Therapy.* Cancer Res, 2020. **80**(21): p. 4593-4600.

56. Black, S.J., et al., *DNA Polymerase theta: A Unique Multifunctional End-Joining Machine.* Genes (Basel), 2016. **7**(9).

57. Mateos-Gomez, P.A., et al., *The helicase domain of Poltheta counteracts RPA to promote alt-NHEJ.* Nat Struct Mol Biol, 2017. **24**(12): p. 1116-1123.

58. Gaillard, H., T. Garcia-Muse, and A. Aguilera, *Replication stress and cancer.* Nat Rev Cancer, 2015. **15**(5): p. 276-89.

59. Ramsden, D.A., J. Carvajal-Garcia, and G.P. Gupta, *Mechanism, cellular functions and cancer roles of polymerase-theta-mediated DNA end joining.* Nat Rev Mol Cell Biol, 2021.

60. Higgins, G.S. and S.J. Boulton, *Beyond PARP-POLtheta as an anticancer target.* Science, 2018. **359**(6381): p. 1217-1218.

61. Mateos-Gomez, P.A., et al., *Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination.* Nature, 2015. **518**(7538): p. 254-7.

62. Mengwasser, K.E., et al., *Genetic Screens Reveal FEN1 and APEX2 as BRCA2 Synthetic Lethal Targets.* Mol Cell, 2019. **73**(5): p. 885-899 e6.

63. Helleday, T., *The underlying mechanism for the PARP and BRCA synthetic lethality: clearing up the misunderstandings.* Mol Oncol, 2011. **5**(4): p. 387-93.

64. Farmer, H., et al., *Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy.* Nature, 2005. **434**(7035): p. 917-21.

65. Lord, C.J. and A. Ashworth, *PARP inhibitors: Synthetic lethality in the clinic.* Science, 2017. **355**(6330): p. 1152-1158.

66. Bryant, H.E., et al., *Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase.* Nature, 2005. **434**(7035): p. 913-7.

67. Lord, C.J. and A. Ashworth, *Mechanisms of resistance to therapies targeting BRCA-mutant cancers.* Nat Med, 2013. **19**(11): p. 1381-8.

68. Higgins, G.S., et al., *Overexpression of POLQ confers a poor prognosis in early breast cancer patients.* Oncotarget, 2010. **1**(3): p. 175-84.

69. Shinmura, K., et al., *POLQ Overexpression Is Associated with an Increased Somatic Mutation Load and PLK4 Overexpression in Lung Adenocarcinoma.* Cancers (Basel), 2019. **11**(5).

70. Schrempf, A., J. Slyskova, and J.I. Loizou, *Targeting the DNA Repair Enzyme Polymerase theta in Cancer Therapy.* Trends Cancer, 2021. **7**(2): p. 98-111.

71. Lemee, F., et al., *DNA polymerase theta up-regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability.* Proc Natl Acad Sci U S A, 2010. **107**(30): p. 13390-5.

72. Li, J., et al., *Depletion of DNA Polymerase Theta Inhibits Tumor Growth and Promotes Genome Instability through the cGAS-STING-ISG Pathway in Esophageal Squamous Cell Carcinoma.* Cancers (Basel), 2021. **13**(13).

73. Pan, Q., et al., *Knockdown of POLQ interferes the development and progression of hepatocellular carcinoma through regulating cell proliferation, apoptosis and migration.* Cancer Cell Int, 2021. **21**(1): p. 482.

74. Woodward, A.M., et al., *Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress.* J Cell Biol, 2006. **173**(5): p. 673-83.

75. Ge, X.Q., D.A. Jackson, and J.J. Blow, *Dormant origins licensed by excess Mcm2-7 are required for human cells to survive replicative stress.* Genes Dev, 2007. **21**(24): p. 3331-41.

76. McIntosh, D. and J.J. Blow, *Dormant origins, the licensing checkpoint, and the response to replicative stresses.* Cold Spring Harb Perspect Biol, 2012. **4**(10).

77. Byun, T.S., et al., *Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint.* Genes Dev, 2005. **19**(9): p. 1040-52.

78. Neelsen, K.J. and M. Lopes, *Replication fork reversal in eukaryotes: from dead end to dynamic response.* Nat Rev Mol Cell Biol, 2015. **16**(4): p. 207-20.

79. Quinet, A., D. Lemacon, and A. Vindigni, *Replication Fork Reversal: Players and Guardians.* Mol Cell, 2017. **68**(5): p. 830-833.

80. Atkinson, J. and P. McGlynn, *Replication fork reversal and the maintenance of genome stability.* Nucleic Acids Res, 2009. **37**(11): p. 3475-92.

81. Neelsen, K.J., et al., *Visualization and interpretation of eukaryotic DNA replication intermediates in vivo by electron microscopy.* Methods Mol Biol, 2014. **1094**: p. 177-208.

82. Moldovan, G.L., B. Pfander, and S. Jentsch, *PCNA, the maestro of the replication fork.* Cell, 2007. **129**(4): p. 665-79.

83. Mailand, N., I. Gibbs-Seymour, and S. Bekker-Jensen, *Regulation of PCNA-protein interactions for genome stability.* Nat Rev Mol Cell Biol, 2013. **14**(5): p. 269-82.

84. Kolinjivadi, A.M., et al., *Smarcal1-Mediated Fork Reversal Triggers Mre11-Dependent Degradation of Nascent DNA in the Absence of Brca2 and Stable Rad51 Nucleofilaments.* Mol Cell, 2017. **67**(5): p. 867-881 e7.

85. Betous, R., et al., *SMARCAL1 catalyzes fork regression and Holliday junction migration to maintain genome stability during DNA replication.* Genes Dev, 2012. **26**(2): p. 151-62.

86. Betous, R., et al., *Substrate-selective repair and restart of replication forks by DNA translocases.* Cell Rep, 2013. **3**(6): p. 1958-69.

87. Hishiki, A., et al., *Structure of a Novel DNA-binding Domain of Helicase-like Transcription Factor (HLTF) and Its Functional Implication in DNA Damage Tolerance.* J Biol Chem, 2015. **290**(21): p. 13215-23.

88. Kile, A.C., et al., *HLTF's Ancient HIRAN Domain Binds 3' DNA Ends to Drive Replication Fork Reversal.* Mol Cell, 2015. **58**(6): p. 1090-100.

89. Vujanovic, M., et al., *Replication Fork Slowing and Reversal upon DNA Damage Require PCNA Polyubiquitination and ZRANB3 DNA Translocase Activity.* Mol Cell, 2017. **67**(5): p. 882-890 e5.

90. Zellweger, R., et al., *Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells.* J Cell Biol, 2015. **208**(5): p. 563-79.

91. Schlacher, K., et al., *Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11.* Cell, 2011. **145**(4): p. 529-42.

92. Hashimoto, Y., et al., *Rad51 protects nascent DNA from Mre11-dependent degradation and promotes continuous DNA synthesis.* Nat Struct Mol Biol, 2010. **17**(11): p. 1305-11.

93. Lemacon, D., et al., *MRE11 and EXO1 nucleases degrade reversed forks and elicit MUS81-dependent fork rescue in BRCA2-deficient cells.* Nat Commun, 2017. **8**(1): p. 860.

94. Mijic, S., et al., *Replication fork reversal triggers fork degradation in BRCA2-defective cells.* Nat Commun, 2017. **8**(1): p. 859.

95. Ray Chaudhuri, A., et al., *Replication fork stability confers chemoresistance in BRCA-deficient cells.* Nature, 2016. **535**(7612): p. 382-7.

96. Sakofsky, C.J. and A. Malkova, *Break induced replication in eukaryotes: mechanisms, functions, and consequences.* Crit Rev Biochem Mol Biol, 2017. **52**(4): p. 395-413.

97. Hashimoto, Y., F. Puddu, and V. Costanzo, *RAD51- and MRE11-dependent reassembly of uncoupled CMG helicase complex at collapsed replication forks.* Nat Struct Mol Biol, 2011. **19**(1): p. 17-24.

98. Sannino, V., et al., *Xenopus laevis as Model System to Study DNA Damage Response and Replication Fork Stability.* Methods Enzymol, 2017. **591**: p. 211-232.

99. Costanzo, V., K. Robertson, and J. Gautier, *Xenopus cell-free extracts to study the DNA damage response.* Methods Mol Biol, 2004. **280**: p. 213-27.

100. Newport, J. and M. Dasso, *On the coupling between DNA replication and mitosis.* J Cell Sci Suppl, 1989. **12**: p. 149-60.

101. Lopes, M., *Electron microscopy methods for studying in vivo DNA replication intermediates.* Methods Mol Biol, 2009. **521**: p. 605-31.

102. Gillespie, P.J., A. Gambus, and J.J. Blow, *Preparation and use of Xenopus egg extracts to study DNA replication and chromatin associated proteins.* Methods, 2012. **57**(2): p. 203-13.

103. Aze, A., et al., *Centromeric DNA replication reconstitution reveals DNA loops and ATR checkpoint suppression.* Nat Cell Biol, 2016. **18**(6): p. 684-91.

104. Sirbu, B.M., F.B. Couch, and D. Cortez, *Monitoring the spatiotemporal dynamics of proteins at replication forks and in assembled chromatin using isolation of proteins on nascent DNA.* Nat Protoc, 2012. **7**(3): p. 594-605.

105. Costanzo, V., et al., *Mre11 protein complex prevents double-strand break accumulation during chromosomal DNA replication.* Mol Cell, 2001. **8**(1): p. 137-47.

106. Seki, M., et al., *High-efficiency bypass of DNA damage by human DNA polymerase Q.* EMBO J, 2004. **23**(22): p. 4484-94.

107. Trenz, K., A. Errico, and V. Costanzo, *Plx1 is required for chromosomal DNA replication under stressful conditions.* EMBO J, 2008. **27**(6): p. 876-85.

108. Clift, D., et al., *A Method for the Acute and Rapid Degradation of Endogenous Proteins.* Cell, 2017. **171**(7): p. 1692-1706 e18.

109. Gan, H.H., et al., *Analysis of protein sequence/structure similarity relationships.* Biophys J, 2002. **83**(5): p. 2781-91.

110. Pearson, W.R., *An introduction to sequence similarity ("homology") searching.* Curr Protoc Bioinformatics, 2013. **Chapter 3**: p. Unit3 1.

111. Larsen, N.B., et al., *Replication-Coupled DNA-Protein Crosslink Repair by SPRTN and the Proteasome in Xenopus Egg Extracts.* Mol Cell, 2019. **73**(3): p. 574-588 e7.

112. Beagan, K. and M. McVey, *Linking DNA polymerase theta structure and function in health and disease.* Cell Mol Life Sci, 2016. **73**(3): p. 603-15.

113. Costanzo, V. and J. Gautier, *Xenopus cell-free extracts to study DNA damage checkpoints.* Methods Mol Biol, 2004. **241**: p. 255-67.

114. Sheaff, R., D. Ilsley, and R. Kuchta, *Mechanism of DNA polymerase alpha inhibition by aphidicolin.* Biochemistry, 1991. **30**(35): p. 8590-7.

115. Panzarino, N.J., et al., *Replication Gaps Underlie BRCA Deficiency and Therapy Response.* Cancer Res, 2021. **81**(5): p. 1388-1397.

116. Cong, K., et al., *Replication gaps are a key determinant of PARP inhibitor synthetic lethality with BRCA deficiency.* Mol Cell, 2021. **81**(15): p. 3128-3144 e7.

117. Ikegami, S., et al., *Aphidicolin prevents mitotic cell division by interfering with the activity of DNA polymerase-alpha.* Nature, 1978. **275**(5679): p. 458-60.

118. Carme, F., M. Rosa, and E. Josep, *Chromosome aberrations induced by aphidicolin.* Mutat Res, 1999. **430**(1): p. 47-53.

119. Sirbu, B.M., et al., *Analysis of protein dynamics at active, stalled, and collapsed replication forks.* Genes Dev, 2011. **25**(12): p. 1320-7.

120. Kliszczak, A.E., et al., *DNA mediated chromatin pull-down for the study of chromatin replication.* Sci Rep, 2011. **1**: p. 95.

121. Alabert, C., et al., *Nascent chromatin capture proteomics determines chromatin dynamics during DNA replication and identifies unknown fork components.* Nat Cell Biol, 2014. **16**(3): p. 281-93.

122. Lopez-Contreras, A.J., et al., *A proteomic characterization of factors enriched at nascent DNA molecules.* Cell Rep, 2013. **3**(4): p. 1105-16.

123. Blow, J.J., et al., *Replication origins in Xenopus egg extract Are 5-15 kilobases apart and are activated in clusters that fire at different times.* J Cell Biol, 2001. **152**(1): p. 15-25.

124. Hyrien, O., C. Maric, and M. Mechali, *Transition in specification of embryonic metazoan DNA replication origins.* Science, 1995. **270**(5238): p. 994-7.

125. Herrick, J. and A. Bensimon, *Global regulation of genome duplication in eukaryotes: an overview from the epifluorescence microscope.* Chromosoma, 2008. **117**(3): p. 243-60.

126. Baranovskiy, A.G., et al., *Structural basis for inhibition of DNA replication by aphidicolin.* Nucleic Acids Res, 2014. **42**(22): p. 14013-21.

127. Zhu, W. and J. Ito, *Family A and family B DNA polymerases are structurally related: evolutionary implications.* Nucleic Acids Res, 1994. **22**(24): p. 5177-83.

128. Malik, R., et al., *Structure and mechanism of B-family DNA polymerase zeta specialized for translesion DNA synthesis.* Nat Struct Mol Biol, 2020. **27**(10): p. 913-924.

129. Cheng, C.H. and R.D. Kuchta, *DNA polymerase epsilon: aphidicolin inhibition and the relationship between polymerase and exonuclease activity.* Biochemistry, 1993. **32**(33): p. 8568-74.

130. Dillon, L.W., A.A. Burrow, and Y.H. Wang, *DNA instability at chromosomal fragile sites in cancer.* Curr Genomics, 2010. **11**(5): p. 326-37.

131. Rickman, K. and A. Smogorzewska, *Advances in understanding DNA processing and protection at stalled replication forks.* J Cell Biol, 2019. **218**(4): p. 1096-1107.

132. De Septenville, A.L., et al., *Replication fork reversal after replication-transcription collision.* PLoS Genet, 2012. **8**(4): p. e1002622.

133. Qiu, S., et al., *Replication Fork Reversal and Protection.* Front Cell Dev Biol, 2021. **9**: p. 670392.

134. Stok, C., et al., *Shaping the BRCAness mutational landscape by alternative double-strand break repair, replication stress and mitotic aberrancies.* Nucleic Acids Res, 2021. **49**(8): p. 4239-4257.

135. Alexandrov, L.B., et al., *Signatures of mutational processes in human cancer.* Nature, 2013. **500**(7463): p. 415-21.

136. Janavicius, R., *Founder BRCA1/2 mutations in the Europe: implications for hereditary breast-ovarian cancer prevention and control.* EPMA J, 2010. **1**(3): p. 397-412.

137. Peshkin, B.N., M.L. Alabek, and C. Isaacs, *BRCA1/2 mutations and triple negative breast cancers.* Breast Dis, 2010. **32**(1-2): p. 25-33.

138. Roy, R., J. Chun, and S.N. Powell, *BRCA1 and BRCA2: different roles in a common pathway of genome protection.* Nat Rev Cancer, 2011. **12**(1): p. 68-78.

139. Gorodetska, I., I. Kozeretska, and A. Dubrovska, *BRCA Genes: The Role in Genome Stability, Cancer Stemness and Therapy Resistance.* J Cancer, 2019. **10**(9): p. 2109-2127.

140. Nagaraju, G. and R. Scully, *Minding the gap: the underground functions of BRCA1 and BRCA2 at stalled replication forks.* DNA Repair (Amst), 2007. **6**(7): p. 1018-31.

141. Taglialatela, A., et al., *Restoration of Replication Fork Stability in BRCA1- and BRCA2-Deficient Cells by Inactivation of SNF2-Family Fork Remodelers.* Mol Cell, 2017. **68**(2): p. 414-430 e8.

142. Feng, W. and M. Jasin, *Homologous Recombination and Replication Fork Protection: BRCA2 and More!* Cold Spring Harb Symp Quant Biol, 2017. **82**: p. 329-338.

143. Di Noia, J.M., C. Rada, and M.S. Neuberger, *SMUG1 is able to excise uracil from immunoglobulin genes: insight into mutation versus repair.* EMBO J, 2006. **25**(3): p. 585-95.

144. Liu, Z.J., et al., *Sequencing abasic sites in DNA at single-nucleotide resolution.* Nat Chem, 2019. **11**(7): p. 629-637.

145. Taglialatela, A., et al., *REV1-Polzeta maintains the viability of homologous recombination-deficient cancer cells through mutagenic repair of PRIMPOL-dependent ssDNA gaps.* Mol Cell, 2021. **81**(19): p. 4008-4025 e7.

146. Hogg, M., et al., *Lesion bypass activity of DNA polymerase theta (POLQ) is an intrinsic property of the pol domain and depends on unique sequence inserts.* J Mol Biol, 2011. **405**(3): p. 642-52.

147. Yousefzadeh, M.J. and R.D. Wood, *DNA polymerase POLQ and cellular defense against DNA damage.* DNA Repair (Amst), 2013. **12**(1): p. 1-9.